



# 大开“智”界：人工智能监管日趋严格

2022年人工智能将面临日趋严格的监管审查，影响波及多个不同行业

**监**管虽然常常滞后于技术创新，目前看来却正在赶超人工智能应用的发展，包括机器学习、深度学习和神经网络。德勤全球预测，2022年将出现更加广泛的讨论，以探索如何更加系统地开展人工智能监管，多项监管提案将被提出——尽管这些提案很可能将在2023年或之后才会变成实质性的执行法规实施生效。一些辖区可能甚至试图完全禁止人工智能的整个分支领域，如公共场所的人脸识别、社会评分和闩下技术等。

## 监管的重要性毋庸置疑，但将如何开展？

正常情况下，预测是力求准确和量化的，但对于监管变化而言通常难以实现。尽管如此，我们仍有充分的理由相信，未来人工智能监管将变得更加普遍、更加严格。截至2021年，欧盟提出了详细的提案，<sup>1</sup>美国联邦贸易委员会亦发布了政策性文件，<sup>2</sup>均旨在加大对人工智能的监管力度。同时，中国也正围绕科技公司推出多项法律法规，其中有部分涵盖人工智能监管。<sup>3</sup>

为何是现在，而非更早？我们认为有多重原因：

- 相比仅仅五年之前，2022年人工智能的性能将会大大增强。得益于速度远远更快的专用处理器、更好的软件以及更大的数据集，人工智能可从事比以往更多的工作，也更具经济效益。<sup>4</sup>因此，人工智能正迅速广泛普及——这反过来亦会导致更严的监管审查。
- 部分监管机构担心人工智能会产生公平、偏见、歧视、多元化和隐私方面的影响。例如，机器学习是当前人工智能背后的基本工具，已经受到监管机构及其他有权机构对其潜在社会偏见影响的密切审查。<sup>5</sup>
- 人工智能监管是地缘政治层面一项有力的竞争性工具。如果一个国家或地区能够设定人工智能监管的全球性标准，则在这一国家或地区运营的企业将获得有利的竞争优势，而外部企业则将处于劣势。

一些监管机构曾直言不讳地指出人工智能可带来的危害。例如，在2021年8月的一份白皮书中，美国联邦贸易委员会委员Rebecca Kelly Slaughter写道：“越来越多的证据显示，算法决策可在就业、信用、医疗和住房等多个高风险经济领域造成具有偏见、歧视性和不公平的结果。”<sup>6</sup>她接着表示，虽然联邦贸易委员会拥有一些现成的工具可用于更好地监管人工智能，“推出新的法规可有助于更有效地解决人工智能和算法决策造成的危害。”<sup>7</sup>

寻找开展人工智能监管的有效方法将极具挑战性。一个根本的问题在于，许多人工智能计算并不具有合理解释——算法做出了决定，但我们并不清楚其做出具体决定的原因。相比上个世纪常用于支持决策制定的可合理解释且可审查的技术，这种不透明性

使得人工智能监管更加难以有效开展。监管法规的意图是防止人工智能驱动型决策产生负面的不利结果，如偏见和不公平，但由于负责做出这些决策的人工智能系统很难被理解和审查，因而无法准确预测负面不利结果何时将会出现——直到民众或机构受到影响。

另一个潜在的问题是训练数据的质量。欧盟人工智能监管法规草案明确表示“训练、验证和测试数据集应当具有相关性、代表性，不存在任何错误且完整全面。”但是对于机器学习所需的数据规模，这一标准——尤其是“不存在任何错误且完整全面”——设置了一个极高的门槛，许多企业和用例可能无法满足要求。<sup>8</sup>

随着人工智能广泛应用，所有人都有理由关心人工智能将如何受到监管，因为相关法律法规将影响人工智能应用可带来多大程度的利益和危害。以下主要利益相关方尤其应当关注：

**人工智能工具使用者。**监管机构很可能将会严打存在危害各阶层民众的算法偏见或其他问题的用例。多项调查显示，人工智能编码形式的偏见存在性别、种族、性向、财富或收入等多方面歧视。这种偏见常常会进一步使弱势群体处于更加不利的地位。这是因为人工智能实际上并非100%人工，而是需要经过数据集进行训练的，这些数据集可反映人类自身的偏见。由此产生的结果是，经过这些数据集训练的人工智能并未消除人类偏见，反而常常将之进一步放大。

一个有关数据集驱动型偏见的知名案例是，一家公司想要雇用更多的女性员工，却发现其所使用的人工智能工具在不停地招聘男性。无论该公司如何努力尝试均无法消除这一偏见，原因就在于训练数据，因此这家公司便完全停止使用这种人工智能工具。<sup>9</sup>

人工智能监管法规将在不同程度上影响不同行业和职能对人工智能工具的使用。例如，在人力资源领域的人工智能应用——具体而言是招聘或绩效管理——将受到深刻的影响。<sup>10</sup>人工智能驱动型决策在招聘、雇佣、晋升、纪律、解约及薪酬方面出现问题的情况已经出现了不少案例。

监管机构亦可能会特别关注审核用户原创内容的互联网平台，其中许多平台严重依赖于人工智能开展审核工作。若不采用人工智能，每天实时审核成百上千万份内容基本是不可能实现的，至少是在成本上是难以承受的。然而，2020年一项调查称算法审核系统“依然缺乏透明、无法解释且所知甚少”，同时“可进一步加剧而非缓解许多与平台制定的内容政策有关的现存问题。”<sup>11</sup>

**金融服务行业也将面临巨大影响，因为该行业在各领域——从信用评级、贷款和抵押，到保险和理财——均采用人工智能协助开展工作。**

从行业角度看，公共行业——健康、教育、政府福利、分区划定、公共安全、刑事司法系统等等——可能将受到深刻的影响。例如，针对法律实施和刑事司法的公共场所人脸识别已经被广泛使用，但这正是欧盟监管法规寻求禁止（除部分例外情况）的技术之一。<sup>12</sup>对于私营医疗和教育行业而言，监管亦将成为一大问题，会对评分等级、奖助学金、学生贷款和纪律措施等事务带来影响。金融服务行业也将面临巨大影响，因为该行业在各领域——从信用评级、贷款和抵押，到保险和理财——均采用人工智能协助开展工作。

物流、采矿、制造、农业等行业受到的影响可能较小。这些行业的人工智能算法当然也会产生问题，但均往往集中于准确性和错误方面，而非造成偏见。然而，尽管这些问题对人类产生直接危害的几率很小，但亦可能会对环境造成影响。

**人工智能工具供应商。**不少科技公司以销售人工智能工具或解决方案的单一业务为生。部分这些工具或解决方案包括有可能受到严苛监管或被禁止的人工智能技术子集；某些甚至仅仅涵盖这些子集。还有更多科技公司提供包含人工智能组件或功能的综合性解决方案，有可能会受到监管的影响。超大规模业者尤其需要紧密关注监管动向。他们均拥有“人工智能即服务”业务，可能在不同程度上受到影响；监管法规可能禁止他们在一些地域销售某些服务，或者相关企业甚至可能要对客户使用其人工智能服务负责。

**同时属于供应商的人工智能使用者。**许多科技互联网平台和应用均大量采用其直接销售或用于执行自身业务模式（或两者）的相同人工智能技术。这些技术中，常见的包括人脸识别、舆情探测和行业预测，均是可能会引发争议的人工智能功能。

**监管机构及社会。**法规制定机构本身也面临多重挑战，需要有效平衡快速变化的技术发展与一系列利

益相关方顾虑之间的关系。他们将需要清晰地阐述全球和国家政策目标，从而制定符合自身实际的相关法律、法规和行为规范。一个灵活敏捷、以改进为导向的监管策略，将比缺乏弹性、以规则为导向的法律更加有效。最后，尽管监管机构和整体社会目标是相互联系的，但彼此又互为分开、截然不同，有时未必能保持一致。

## 小结

未来两年，我们可能会见证不同情景的发展。

首先，利益相关方受通过实施的监管法规影响，可能会在部分辖区关闭由人工智能驱动的功能，或完全停止在部分辖区的运营——或继续正常运营，受到处罚，然后支付相关罚款作为业务运营的成本。

第二，欧盟、美国和中国等大型重要市场有可能会通过相互冲突的人工智能监管法规，企业无法——遵循。

第三，某一套人工智能监管法规有可能会成为占据主导的“黄金标准”，正如欧盟针对隐私保护的《通用数据保护法案》一样，有助于简化跨境合规的复杂度。

第四，人工智能供应商及平台甚至有可能抱团组建联合体，引领针对人工智能工具如何使用及如何提升透明度和可审查性的探讨——采用一定程度的自我监督，以期弱化监管机构认为需要从上至下开展监管的认知。

即便实际出现的是最后一个情景，监管机构也不太可能完全坐视不管。几乎可以断定，未来短期内将有更多监管法规实施生效。尽管无法准确预知将会出现何种法规，但很大可能将会对人工智能应用带来巨大的影响。

## 尾注

1. European Commission, "Proposal for a regulation of the European Parliament and of the council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts," EUR-Lex, April 21, 2021.
2. Elisa Jillson, "Aiming for truth, fairness, and equity in your company's use of AI," Federal Trade Commission, April 19, 2021.
3. Arjun Kharpal and Evelyn Cheng, "The latest target of China's tech regulation blitz: algorithms," CNBC, September 9, 2021.
4. 举个例子，采用专用人工智能芯片可带来数以千计的提升。Saif Khan and Alexander Mann, *AI chips: What they are and why they matter*, Center for Security and Emerging Technology (CSET), April 2020.
5. James Manyika, Jake Silberg, and Brittany Presten, "What do we do about the biases in AI?," *Harvard Business Review*, October 25, 2019.
6. Rebecca Kelly Slaughter, "Algorithms and economic justice: A taxonomy of harms and a path forward for the Federal Trade Commission," *Yale Journal of Law & Technology*, August 2021.
7. Ibid.
8. European Commission, Artificial Intelligence Act.
9. Jeffrey Dastin, "Amazon scraps secret AI recruiting tool that showed bias against women," Reuters, October 10, 2018.
10. Tom Simonite, "New York City proposes regulating algorithms used in hiring," *Wired*, January 8, 2021.
11. Robert Gorwa, Reuben Binns, and Christian Katzenbach, "Algorithmic content moderation: Technical and political challenges in the automation of platform governance," *Big Data & Society* 7, no. 1 (2020).
12. 欧盟委员会，《人工智能法》。

## 关于作者

**Duncan Stewart | Canada | [dunstewart@deloitte.ca](mailto:dunstewart@deloitte.ca)**

Duncan Stewart is the director of research for the Technology, Media & Telecommunications (TMT) industry for Deloitte Canada. He presents regularly at conferences and to companies on marketing, technology, consumer trends, and the longer-term TMT outlook.

**Paul Lee | United Kingdom | [paullee@deloitte.co.uk](mailto:paullee@deloitte.co.uk)**

Paul Lee is a UK partner and the global head of research for the technology, Media & telecommunications (TMT) industry at Deloitte. In addition to running the TMT research team globally, Lee manages the industry research team for Deloitte UK.

**Ariane Bucaille | France | [abucaille@deloitte.fr](mailto:abucaille@deloitte.fr)**

Ariane Bucaille is Deloitte's global Technology, Media & Telecommunications (TMT) industry leader and also leads the TMT practice and the TMT Audit practice in France. She has more than 20 years of experience and is a chartered and certified public accountant.

**Gillian Crossan | United States | [gicrossan@deloitte.com](mailto:gicrossan@deloitte.com)**

Gillian Crossan is a principal in Risk & Financial Advisory, Deloitte & Touche LLP, and leads the global technology industry sector. She has been with Deloitte for more than 25 years and has worked across sectors including energy, health care, consumer products, and technology.

## 致谢

The authors would like to thank the following individuals for their contributions to this chapter:

**Beena Ammanath, Ralf Esser, Lukas Kruger, Susie Samet, and Nick Seeber.**

## 关于德勤科技、传媒和电信行业中心

德勤科技、传媒和电信行业 (TMT) 中心专注于研究并发表洞察, 以帮助企业领导者清晰了解其业务选择。在新技术和新趋势背景下, 本中心的研究将协助企业高管简化复杂的业务问题, 并提出明智策略, 提升企业长久竞争优势并赢得商业胜利。本中心将作为值得信赖的顾问, 帮助高管更好地识别风险, 获悉商业回报, 赢取关键机遇, 从而在快速变化的TMT环境中解决棘手挑战。

### 联系我们

了解有关科技、传媒和电信行业中心的更多信息并获取最新研究和洞察报告, 请访问 [www.deloitte.com/us/tmtcenter](http://www.deloitte.com/us/tmtcenter)。

### 订阅

如您想接收TMT行业电子邮件, 请访问<https://my.deloitte.com/subscriptions.html>, 选择您感兴趣的领域进行订阅。

### 关注我们

敬请关注 [@DeloitteTMT](https://twitter.com/DeloitteTMT)。

德勤科技、传媒和电信行业汇聚了全球最顶级的行业专家, 组成全球最大的专业团队之一, 协助各类形态和规模的企业在数字化时代蓬勃发展, 成就辉煌。德勤科技、传媒和电信行业专家致力于为企业丰富的定制化服务, 帮助他们顺应变革趋势, 抢占行业先机, 所服务的客户遍布全球, 覆盖全价值链。敬请联系作者或访问[www.deloitte.com](http://www.deloitte.com), 了解更多信息。

# Deloitte.

## Insights

敬请登陆 [www.deloitte.com/insights](http://www.deloitte.com/insights) 订阅德勤洞察最新资讯。



敬请关注 @DeloitteInsight

### 参与人员

**编辑:** Junko Kaji, Preetha Devan, Prodyut Ranjan Borah, Rupesh Bhat, Arpan Kumar Saha, Ribhu Ranjan, Emma Downey, Nairita Gangopadhyay, Blythe Hurley, and Aparna Prusty

**创意:** Jaime Austin, Sylvia Yoon Chang, Govindh Raj, Sanaa Saifi, and Rishwa Amarnath

**推广:** Maria Martin Cirujano

**封面设计:** Jaime Austin

### 关于德勤

Deloitte (“德勤”) 泛指一家或多家德勤有限公司, 及其全球成员所网络和它们的关联机构。德勤有限公司 (又称“德勤全球”) 及其每一家成员所和它们的关联机构均为具有独立法律地位的法律实体。德勤有限公司并不向客户提供服务。请参阅 [www.deloitte.com/about](http://www.deloitte.com/about) 了解更多信息。

### 关于本刊物

本通讯中所含内容乃一般性信息, 任何德勤有限公司、其全球成员所网络或它们的关联机构 (统称为“德勤组织”) 并不因此构成提供任何专业建议或服务。在作出任何可能影响您的财务或业务的决策或采取任何相关行动前, 您应咨询合格的专业顾问。

我们并未对本通讯所含信息的准确性或完整性作出任何 (明示或暗示) 陈述、保证或承诺。任何德勤有限公司、其成员所、关联机构、员工或代理方均不对任何方因使用本通讯而直接或间接导致的任何损失或损害承担责任。德勤有限公司及其每一家成员所和它们的关联机构均为具有独立法律地位的法律实体。

CQ-035SC-21

© 2021。欲了解更多信息, 请联系德勤全球。