



推动人工智能发展：驾驭风险，攻克难关

作者：Beena Ammanath、David R. Novak、Siri Anderson、Abha Kulkarni

德勤综合研究中心、德勤人工智能研究院联合发布

随着人工智能逐渐深入各行各业，由此产生的风险也引发出各种顾虑，导致人工智能应用放缓。所幸的是，这些顾虑能得到有效缓解和应对。

概览

人工智能时代已经全面来临。¹自2017年开始，德勤每年发布《企业人工智能应用现状分析》报告，以估量人工智能技术在全球范围及各行各业的快速发展。在2020年7月发布的最新一期报告中，大多数受访者表示大幅增加了人工智能投资，超过四分之三的受访者认为人工智能将在未来三年为自己所在的企业带来重大变革。此外，人工智能投资也为企业创造出更多可衡量的价值：提升流程效率、改善决策、提高员工生产力、改进产品和服务。²这些价值提升了企业对人工智能的认知——四分之三的受访者表示，人工智能具有重要战略意义，较前一次调研结果上升10%。

但另一方面，对人工智能风险的顾虑与这种上升趋势形成反差：56%的受访企业表示，出于对新兴风险的顾虑，他们已经或计划放缓人工智能的应用。鉴于人工智能的价值和战略意义已得到普遍认可，这一比率造成了强烈反差。为了更好地理解这些问题，德勤将受访企业的顾虑归总为三类：对人工智能决策流程的**信心**、人工智能和数据使用的**伦理规范**，以及**市场不确定性**。³

本文将针对上述问题进行详细阐述，并提供专业见解协助企业有效管理相关风险。由此，企业可对内加强持续投资和创新的信心；对外则进一步获得利益相关方的信任，确保人工智能战略得到有效执行。

出于对新兴风险的顾虑，**56%的受访企业表示已经或计划放缓人工智能的应用。**

人工智能发展面临的挑战

近年来，人工智能风险日益引起商业领袖和普通民众的关注。有关人工智能是否会削减或创造工作岗位的争论从未停止；而对人工智能偏见或失败报道也屡见不鲜。⁴2020年7月，在公众的一片反对与批评声中，成立仅五天的人工智能初创公司 Genderify 被迫关闭。该公司称其人工智能可以通过分析用户的名字、用户名和电子邮件地址来识别其性别，并宣称这是一种增强客户数据的能力。据报道，该系统充斥着对性别的误解，并基于错误的假设做出决定。例如，在包含“科学家”的记录中，识别为女性的概率仅为4.3%。但事实上，如今女性在科学领域的占比远高于此。⁵而包含“教授”的记录中，识别为男性的概率高达98.4%，但男性占美国终身教授的比例仅为74%（该比例本身就反映出一种不平等）。⁶

不难想象，诸如此类的事件以及事件曝光之后引发的公众抵制，将导致许多企业高管更为谨慎，重新考虑如何在企业内部、为客户和员工以及在生态系统中推行人工智能。为了更准确地了解企业的最大顾虑源自何处，我们请受访者对顾虑事项进行评分，评分标准从“极少”到“极大顾虑”。列出的风险归总为三类：*信心*、*伦理规范*和*市场不确定性*。

信心是指企业普遍认为一项人工智能工具本身可靠，包括信任人工智能提供的洞察和决策，信任系统的安全性，以及能够理解、证明或解释人工智能的决策过程。

伦理规范是指从社会而非企业的角度审视使用人工智能的益处和正确性，包括数据隐私、公平、偏见，以及可能造成的失业问题。**市场不确定性**则涵盖那些不在企业直接控制范围内的因素，比如监管环境的变化以及公众或员工的意见。这些因素不一定与人工智能的本质相关，但会影响人工智能的实施和应用。

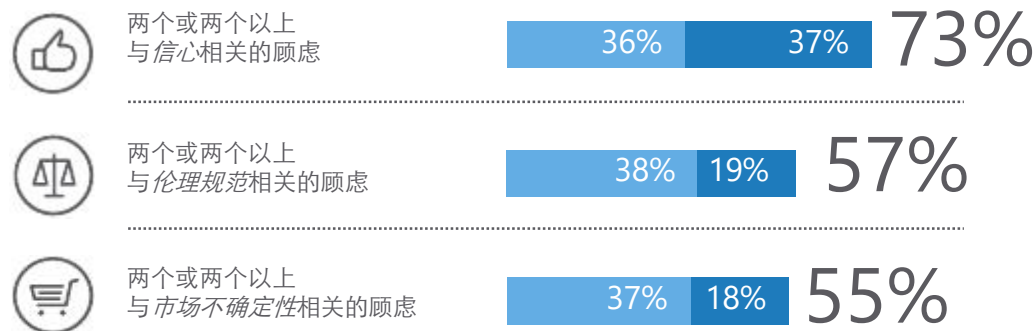
图1反映了在每个类别中，将两个或两个以上的顾虑事项评为“重大”或“极大顾虑”的受访者比例。

图 1

对人工智能风险的重大与极大顾虑

■ 重大顾虑 ■ 极大顾虑

对人工智能风险的顾虑



注：本图反映在每个类别中将两个或两个以上顾虑事项评为“重大”或“极大顾虑”的受访者比例。

来源：德勤分析

信心

对人工智能系统的信心是最大的顾虑所在，73%的受访企业高管在信心这一项中选择了两个或两个以上重大或极大顾虑。而在计划放缓人工智能应用的受访者中，有85%选择了两个或两个以上与信心相关的重大或极大顾虑。

可以想见，信心不足造成的顾虑会引起不安。这些顾虑与公司产品和服务的质量密切相关，因此会大大影响一个企业的声誉以及公众对其的看法，进而影响企业推出的每一个产品。例如，医疗机构可以利用人工智能的巨大潜力来识别癌症类型并确定最佳治疗方案。⁷然而，人工智能决策

引擎有可能推荐不安全的治疗方案或错误识别癌症类型。这会给患者造成严重后果，使企业面临风险并带来投资损失。一旦失败，则可能让企业付出比投资损失更为昂贵的代价——打击员工持续创新的信心，甚至破坏患者和客户的信任。

在上述案例中，该公司据称建立了保障措施，以确保患者安全，从而降低破坏患者信任的可能。另外，相关错误还被用来识别系统故障，改进未来的算法。⁸如果从一开始就充分规避风险，这将转化为一股积极的力量，推动人工智能发展成熟。

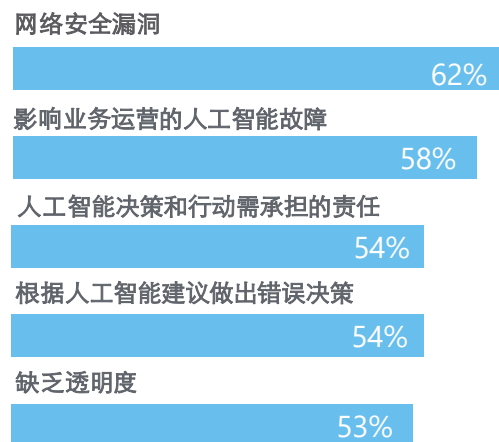
图 2

对人工智能信心的顾虑

■ 重大或极大顾虑



总体而言，您的企业对以下人工智能潜在风险的顾虑程度如何？



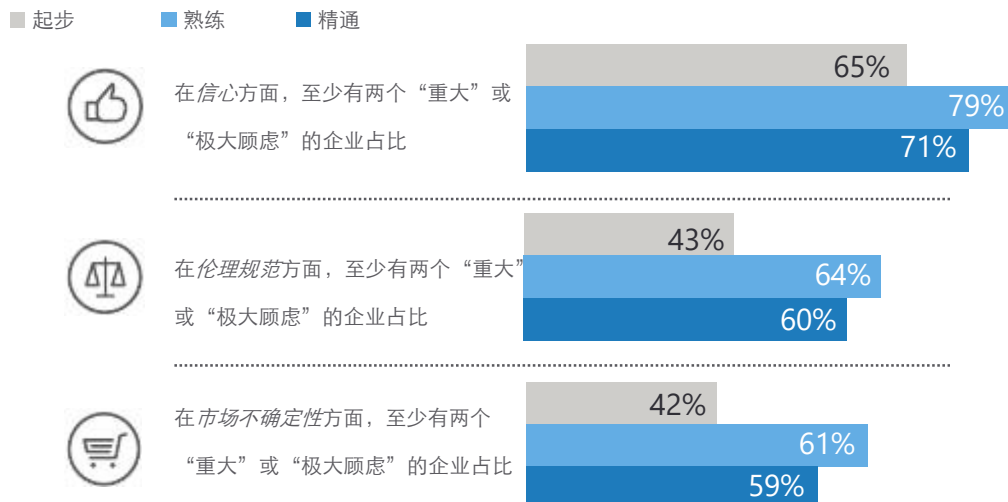
来源：德勤分析

人工智能成熟度与企业管控预期风险的能力

整体来看，随着受访企业人工智能能力的成熟度上升，对**信心**、**伦理规范**和**市场不确定性**的顾虑呈现为铃型曲线。

图 3

成熟度如何改变顾虑程度



来源：德勤分析

低成熟度和高成熟度的企业展现出较低程度的顾虑，而中等成熟度的企业在三个类别中均表现出最高程度的顾虑。原因可能在于低成熟度的企业还没有全面看到风险，许多项目仍处于概念验证或试点阶段。当企业达到中等成熟度时，这些挑战逐渐明晰，但企业尚未具备应对这些挑战所需的能力。最后，当企业达到高度成熟时，也掌握了更多降低这些风险的技能，因此顾虑程度往往会相应降低。

事实上，有证据表明，随着企业的人工智能能力日益成熟，企业可通过开发流程、采取行动或提升技能来降低风险并实现更积极的成果，其顾虑程度也随之下降。

（参见侧栏“人工智能成熟度与企业管控预期风险的能力”）。

与信心相关的另一个重大挑战是，如今的人工智能系统往往存在可解释性和决策能力之间的权衡。一些机器学习模型异常复杂，即便是训练有素的数据科学家也难以准确理解算法是如何做出决策的。神经网络的广泛运用让这个挑战变得更为突出，进一步涉及企业解释决策、减少错误和满足监管需求的能力。

为了解决这个问题，计算机科学领域正在发展可解释的人工智能，旨在开发能够更好地进行自我解释的人工智能模型。

目前，可解释的人工智能可以提供人工智能程序做出决策的相关信息，包括人工智能程序的优缺点、制定决策所采用的具体标准，并针对各类决策的可信度提出建议。美国商务部最近提出了评估人工智能可解释性的四个原则：1. 人工智能系统需要为每项输出提供附带的证据或理由；2. 人工智能系统必须为个人用户提供有意义的、可理解的解释；3. 人工智能系统提供的解释必须正确反映系统输出的过程；4. 系统只应在其设定条件下运行，不应在没有充足信心的情况下向用户提供决策。⁹掌握了人工智能的可解释性，企业也能更好地在符合伦理规范的前提下运用人工智能。

伦理规范

57%的受访者称在伦理规范方面有两个或

两个以上的重大或极大顾虑。而在计划放缓人工智能应用的受访者中，有73%声称在伦理规范方面至少有两个重大或极大顾虑。

人工智能在伦理规范方面存在特有的问题，尤其就公平和偏见而言，即人工智能赖以建立的数据本身可能是不完整、有偏见或不平等的。出于这个原因，即使公司本意为善，也很难根除数据本身自带的偏见。

在构建人工智能、分析系统输出以及应用系统决策的过程中，相关人员也可能受到无意识偏见的影响，因此需要采取慎重措施，尽可能地减少偏见。在最近的一个医疗案例中，一个基于医疗支出的风险预测算法最终在其结果中表现出种族偏见，使白人患者比黑人患者更有可能从额外的医疗计划中获益。¹⁰虽然该算法准确地反映了这两个群体的医疗支出，但没有考虑到黑人患者往往身处经济劣势，从而影响他们的医疗支出：价格高昂的急诊室就诊费，参保率更低等。¹¹

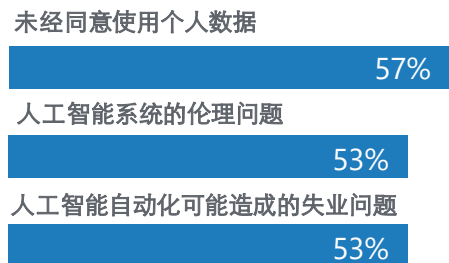
图 4

对人工智能伦理规范的顾虑

■ 重大或极大顾虑



总体而言，您的企业对以下人工智能潜在风险的顾虑程度如何？



来源：德勤分析

这可能会导致完全不同的需求水平呈现相同的支出水平，而这正是该系统在分析数据时所忽略的。在上面的例子中，黑人患者的风险评分不准确，被排除在额外医疗计划之外的比例更高。

重要的是，程序设计者应全面检测基于代理的决策和创建的变量，制定保障措施来评估这种设计是否反映了利益相关者的实际情况。此外，这些例子也表明，有必要纳入多元化观点，并对工程师和决策者开展培训，应对算法设计中可能产生的意外结果。

市场不确定性

最后，55%的受访者表示在市场不确定性方面有两个或两个以上的重大或极大顾虑。这一比例排在第三位，但明显与投资行为有着巨大关联。而在有计划放缓人工智能

应用的受访者中，有71%表示至少有两个与市场不确定性相关的重大顾虑。

对市场不确定性的顾虑涵盖范围非常广，从公众舆论到监管环境等等。57%的受访者表示其重大或极大顾虑与针对人工智能技术不断推出的新规定和法规有关，例如数据隐私、人脸识别和决策透明度等。¹² 52%的受访者表示其重大或极大顾虑在于，客户一旦发现人工智能应用中存在缺陷或侵犯隐私，就会发起强烈抵制。还有一半以上的受访者担心，使用人工智能系统将导致员工产生负面想法。

如何驾驭人工智能应用中的风险？

人工智能带来的巨大变革和潜在风险似乎令人生畏。我们的调研中，每十位受访者

图 5

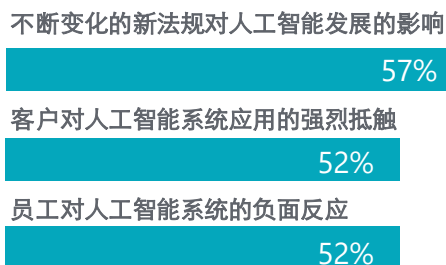
对人工智能市场不确定性的顾虑

■ 重大或极大顾虑



来源：德勤分析

总体而言，您的企业对以下人工智能潜在风险的顾虑程度如何？



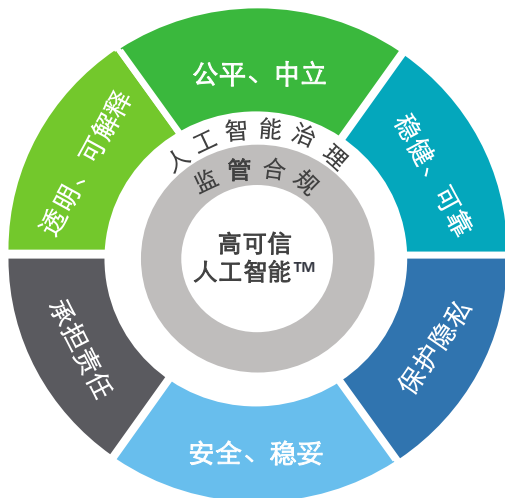
中仅有四位表示做好了充分准备。¹³ 鉴于人工智能技术的快速发展，这种不确定性必将在一段时间内持续存在。然而，企业可以借鉴现有的一些框架，并从先前的技术变革中吸取经验，帮助驾驭当前的风险。

运用可信赖的人工智能，减少对信心和伦理规范的顾虑

在上文分析的三类顾虑中，信心和伦理规范是企业关注的重点，而这两点最终都指向人工智能应用是否可信的问题。为了增强对人工智能的信心，企业应确保其工具和解决方案是透明、可靠、安全的，并制定相应的问责制度。要减少对伦理规范的顾虑，通常需要评估方案的设计是否在意愿上和效果上都实现了公平，以及数据的使用是否遵循了明确的隐私保护标准。企业应该制定保障措施来管控人工智能风险，涵盖图6中列出的六大关键维度。

图 6

高可信人工智能框架



来源：德勤“高可信人工智能服务”

对于信心和伦理规范的顾虑本质上归结为企业对人工智能的运用是否可信。企业实施这一框架的能力在很大程度上取决于负责任的数据管理、严格的治理标准，以及借助全面多样的观点集合来识别和应对有害假设。

德勤建议：注重数据管理和治理

对企业而言，有效处理数据对人工智能的质量和可解释性至关重要，可以提升对人工智能的信心，并增强企业管控伦理规范的能力。即使在今天，40%的受访企业仍表示在一系列数据实践中的成熟度为“低”或是“中等”，近三分之一的高管认为与数据相关的挑战是阻碍其公司人工智能举措的三大问题之一。¹⁴ 在数据治理方面，仅有五分之一的受访企业将例行监测、管理和提高数据质量作为正式数据治理工作的一部分，而只有12%的企业认为他们的数据是最新的，仅9%的企业认为他们的数据是准确的。¹⁵

倘若缺少有效的数据管理和治理流程，企业就很难降低风险。专注于这两大基础要素，企业才能成功实现可信赖的人工智能所需的六个维度，减少对信心和伦理规范的顾虑，为人工智能应用扫清障碍。

德勤建议：纳入多元化观点

公司实施保障措施的能力还取决于企业内部技术方和业务方的多元化思维。数据能力和治理流程为评估各种标准和法规的合规性创建了待查点，但多元化的观点和思维才能让重要见解浮出水面。

多元化观点有助于在产品的设计和审查阶段发现潜在问题，而不是在产品推出后才后知后觉。

驾驭市场不确定性

对*市场不确定性*的顾虑更多来自于外部环境中尚不明确的因素，因此需要采取不同的方法进行应对。伴随新兴技术和重大转型出现的是不断变化的法规和随时波动的舆论。虽然企业不可能完全掌控这些不确定因素，但企业高管依然可以采取的措施，更有效地驾驭风险

德勤建议：重视变革管理和有效沟通

公众和员工的意见并不会影响人工智能的表现。经过深思熟虑的变革管理和有效沟通对于成功驾驭市场不确定性至关重要。未来几年，人工智能可能会改变许多工作岗位，因此帮助员工和客户学习新的工作及互动方式至关重要。Humana就是一个例子，该公司部署了人工智能客服助理来接听让人工客服不堪重负的每月数百万个来电。Humana开发了人工智能助理来处理基本的信息请求，占来电总量的60%。事实上，配备自然语言理解能力的人工智能机器人还能做更多：协助员工在幕后收集信息，然后提供给Humana呼叫中心的工作人员。而这项新技术被定义为人工客服助手，减轻了员工可能感受到来自人工智能的威胁。

该公司专注于利用人工智能来支持而非取代员工，包括与员工的沟通（例如恰当表达同理心）以及与人工智能协作解决问题。这些做法有助于最大化地实现成功。¹⁶

了解员工的看法，同时有计划地推行技术以减轻其颠覆程度，往往是获得认可的关键。企业可考虑与沟通和营销团队合作，确保员工和客户都能充分理解并看到人工智能带来的价值。采取以用户为中心的方法，而不是为了节约成本强行推进，有助于避免抵触和反击。

德勤建议：预测应对监管变化所需的能力

欧洲议会和欧盟于2016年制定《通用数据保护条例》(GDPR)，要求于2018年实施，企业有两年时间来了解和筹备。在GDPR生效前的几个月，对监管合规习以为常的行业，比如金融业，往往准备得较为充分。这类企业已经建立了相应的技能和运营能力，可应对与合规相关的必要变革。目前，有关新出台人工智能政策和可能推出技术法规的讨论日益激烈，企业应积极建立这些能力，从而既能参与法规的制定，又能及时应对新法规的出台。

结论

人工智能蕴含着巨大的变革潜能。本次受访的技术高管中，逾九成认为人工智能将是下一次科技革命的核心所在，并且人工智能将逐渐承担琐碎的工作，让人类有更多自由去追求更富创造性的工作和娱乐。

人工智能对当今世界的潜在价值不容忽视，商业领袖不能因过度谨慎而退缩。通过谨慎地开发流程和提升能力，可以帮助企业减轻对风险和挑战的顾虑。即刻行动，打造竞争优势，攻克发展难关，助力企业赢在未来。



方法论

本次调研于2019年10月至12月期间开展，九个国家的2737名IT和业务线高管参与了调研。所有受访企业都采用了人工智能技术，受访者须满足以下任一条件：

- 负责确定人工智能技术的支出和/或批准人工智能投资
- 负责制定人工智能技术战略
- 负责管理或监督人工智能技术的实施
- 是人工智能技术领域的专家
- 可影响或制定有关人工智能技术的决策

受访者中，47%为IT高管，其余为业务线高管。70%是首席高管，包括：首席执行官、总裁和所有者（35%）；首席信息官和首席技术官（32%）；以及其他首席高管（3%）。

为了更准确地了解企业最大的顾虑源自何处，调研要求受访者对一系列顾虑事项进行评分，评分标准从“极少”到“极大”。我们将列出的风险分为三类：信心、伦理规范和市场不确定性。

图 7

对人工智能顾虑的类别

类别	对人工智能的顾虑
信心	缺乏透明度——不能解释人工智能系统如何做出具体决策 人工智能决策和行动需承担的责任 根据人工智能建议做出错误决策 影响业务运营的人工智能故障 网络安全漏洞
伦理规范	未经同意使用个人数据 人工智能系统的伦理问题 人工智能自动化可能造成的失业问题
市场不确定性	不断变化的新法规对人工智能发展的影响 客户对人工智能系统应用的强烈抵触 员工对人工智能系统的负面反应

尾注

1. Beena Ammanath, David Jarvis, and Susanne Hupfer, *Thriving in the era of pervasive AI: Deloitte's State of AI in the Enterprise, 3rd Edition*, Deloitte Insights, July 14, 2020.
2. Ibid.
3. Ibid.
4. For a deeper look, see Meredith Broussard, *Artificial Unintelligence: How Computers Misunderstand the World* (Cambridge: MIT Press, 2018).
5. Catalyst, "Women in science, technology, engineering, and mathematics (STEM): Quick take," August 4, 2020.
6. Bridget Turner Kelly, "Though more women are on college campuses, climbing the professor ladder remains a challenge," Brookings Institution, March 29, 2020; Synced, "AI-powered 'Genderify' platform shut down after bias-based backlash," July 30, 2020.
7. Adam Conner-Simons and Rachel Gordon, "Using AI to predict breast cancer and personalize care," *MIT News*, May 7, 2019.
8. Jo Cavallo, "Confronting the criticisms facing Watson for oncology," *ASCO Post*, September 10, 2019.
9. National Institute of Standards and Technology, "NIST asks A.I. to explain itself," August 18, 2020.
10. Ziad Obermeyer et al., "Dissecting racial bias in an algorithm used to manage the health of populations," *Science* 366, no. 6464 (2019): pp.447–53.
11. Ibid.; Jamila Taylor, "Racism, inequality, and health care for African Americans," Century Foundation, December 19, 2019.
12. Mark MacCarthy, "AI needs more regulation, not less," Brookings Institution, March 9, 2020.
13. Ammanath et al., *Thriving in the era of pervasive AI*.
14. Karthik Ramachandran, *Data management barriers to AI success*, Deloitte Insights, August 7, 2020.
15. MIT SMR Connections and SAS, "How trust delivers value in data, analytics, & AI," January 15, 2019.
16. Kim S. Nash, "Artificial intelligence helps Humana avoid call center meltdowns," *Wall Street Journal*, October 27, 2016.
17. Ponemon Institute, "The race to GDPR: A study of companies in the United States & Europe," April 2018.
18. Edelman, "2019 Edelman AI survey," March 2019.

致谢

感谢 **David Jarvis, Susanne Hupfer, Brenna Sniderman** 和 **Natasha Buckley** 为本报告谋篇布局。

关于作者

Beena Ammanath | bammanath@deloitte.com

Beena Ammanath现任Deloitte Consulting LLP 人工智能主管总监，是一名在人工智能和数字化转型方面拥有丰富全球经验的资深高管。她是人工智能公司Humans For AI Inc. 的创始人兼首席执行官，并合著了《人工智能变革商业》（AI Transforming Business）一书。她常驻加利福尼亚州Pleasanton。

David R. Novak | danovak@deloitte.com

David R. Novak 博士是德勤综合研究中心的研究经理。他擅长在人与科技、人工智能和可持续发展领域发展领先理念。在加入德勤之前，诺瓦克曾担任传播学教授、战略传播顾问以及数据分析领域的工作。

Siri Anderson | sianderson@deloitte.com

Siri Anderson是德勤综合研究中心的研究员和作家，专注于发展未来工作和数字现实领域的领先理念。Anderson曾领导德勤数字公司的赞助和内容项目。在加入德勤之前，Anderson曾在科技和医疗行业担任编辑和战略传播职务。

Abha Kulkarni | abhakikulkarni@deloitte.com

Abha Kulkarni是德勤印度服务有限公司德勤综合研究中心的高级分析师。她通过分析各行业和地区的劳动力数据来支持“未来工作”。她主要关注应用定量研究技术，实现数据驱动的研究见解。

联系我们

德勤以专业洞察，助您引领变革。如欲了解攻克难关的创新见解，请随时联系我们。

德勤全球行业领导人

Beena Ammanath

德勤人工智能研究院执行总监 | Deloitte & Touche LLP

+1 925 474 7139 | bammanath@deloitte.com

德勤中国领导人

赵文华

德勤人工智能研究院中国区领导人

德勤中国首席数字官

+86 21 6141 1568 | josechu@deloitte.com.cn

关于德勤人工智能研究院


德勤人工智能研究院汇集人工智能领域顶尖智慧，凭借前沿研究和创新，推动“赋能时代”的人机协作，助力组织机构实现人工智能转型。研究院的成立旨在促进人工智能的对话和发展，激励创新，挑战现状。德勤人工智能研究院与由行业思想领袖、学术权威、初创企业、研发团队、企业家、投资者和创新者组成的生态系统协作，结合德勤在人工智能应用领域的深度经验，助力组织机构实现人工智能转型。研究院专注于人工智能多个重点领域，当前研究重心主要集中在伦理道德、创新、全球发展、未来工作和人工智能案例研究等。敬请访问[Deloitte.com](https://www.deloitte.com)。

关于德勤综合研究中心

德勤综合研究中心专注于对跨行业、跨职能的关键业务问题，包括从新兴技术的快速变化到人类行为的一致因素，生成全新的观点。我们以新的方式审视变革性，以各种形式传递新思想，如研究文章、短视频、面对面的研讨会和在线课程。敬请访问[Deloitte.com](https://www.deloitte.com)。

Deloitte. Insights

敬请登陆www.deloitte.com/insights订阅德勤洞察最新资讯。

 敬请关注 @DeloitteInsight

德勤洞察

编辑: Matthew Budman、Hannah Backman, Rupesh Bhat 和 Nairita Gangopadhyay

创意: Victoria Lee 和 Govindh Raj

推广: Maria Martin Cirujano

封面: Daniel Hertzberg

关于德勤洞察

德勤洞察发布原创文章、报告和期刊，为企业、公共领域和非政府机构提供专业洞察。我们的目标是通过调查研究，利用整个德勤专业服务机构的专业经验，以及来自学界和商界作者的合作，就企业高管与政府领导人所关注的广泛议题进行更深入的探讨。

德勤洞察是Deloitte Development LLC旗下出版商。

关于本刊物

本刊物中所含内容仅为一般性信息，任何德勤有限公司、其成员所或它们的关联机构（统称为“德勤组织”）并不因此构成提供任何会计、商业、财务、投资、法律、税务或其他专业建议或服务。本刊物不能代替此等专业建议或服务，读者亦不应依赖本文件中的信息作为可能影响自身财务或业务决策的基础。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。

任何德勤网络内的机构均不对任何方因使用本刊物而导致的任何损失承担责任。

关于德勤

Deloitte（“德勤”）泛指德勤有限公司（一家根据英国法律组成的私人担保有限公司，以下称“德勤有限公司”），以及其一家或多家成员所和它们的关联机构。德勤有限公司与每一个成员所均为具有独立法律地位的法律实体。德勤有限公司（又称“德勤全球”）并不向客户提供服务。在美国，德勤指德勤有限公司、在美国以“德勤”的名义运营的关联机构及其各自的附属公司所属的一家或多家美国成员所。根据公告会计条例及法规，某些服务并不向鉴证客户提供。请参阅www.deloitte.com/about 以了解更多有关德勤有限公司及其成员所的详情。