



为什么要投资人工智能伦理与治理？

五个真实世界的起源故事

与圣母院—IBM技术伦理实验室合作



IBM如何提供帮助

客户可以利用ibm深厚的行业、职能和技术专业知识；企业级技术解决方案；以及基于科学的研发创新，来发掘人工智能、分析和数据的潜力。

更多信息：

ibm.com/services/人工智能

来自 IBM 软件的 AI 解决方案 ibm.com/Watson

来自 IBM 研究院的 AI 创新®
research.ibm.com/人工智能

巴黎圣母院-IBM技术伦理实验室
techethicslab.nd.edu/



关键点

衡量人工智能道德价值的组织可能会领先一步。我们的整体人工智能道德框架考虑了三种类型的ROI。

■ 拥抱人工智能伦理至关重要。

这不仅仅关乎损失规避。75%的高管认为人工智能伦理是重要的竞争优势来源。¹ 超过85%的受访消费者、公民和员工重视人工智能伦理。²

■ 长期来看，积极主动的人工智能伦理战略能为整个组织创造价值。

大部分公司（54%）期望人工智能伦理在战略上非常重要³ 随着高管提到涉及20个不同的业务职能。⁴

■ 投资人工智能伦理具有创造可量化价值的潜力。

衡量人工智能伦理价值的组织可能领先一步。我们的整体人工智能伦理框架考虑了三种类型的ROI：经济影响（有形的）、声誉影响（无形的）和能力（真实期权ROI）。

生成式人工智能正在改变行业，但其令人目眩的崛起也引发了重大的伦理问题。在潜在利益与伦理和监管影响之间取得平衡至关重要。

但是这不简单。在IBM商业价值研究所 (IBM IBV) 的研究中，80%的企业领导者认为AI可解释性、伦理、偏见或信任是生成式AI采用的主要障碍。⁵ 还有一半人说他们的组织缺乏治理和结构来管理生成式人工智能的伦理挑战。⁶

面对这种不确定性和风险，许多首席执行官都在暂停行动。超过一半 (56%) 的人将生成式人工智能的主要投资推迟到他们明确人工智能标准和法规之前。⁷ 并且 72% 的高管表示他们的组织实际上会因为伦理问题而放弃生成式人工智能的益处。⁸

然而，有一条前进的道路——如果高管们拓宽他们的视野，将人工智能伦理视为一个机会。更好的是：*持续的研究表明，投资人工智能伦理有潜力创造可量化的效益。*

为了解锁这种潜力，组织需要在评估人工智能伦理投资的投资回报率时采用新的视角。在本报告的第一部分，我们确定了三种适用于人工智能伦理的投资回报类型——换句话说，一个整体的人工智能伦理框架。在第二部分和第三部分，我们探索了两种不同但有价值的方法来论证当前的人工智能伦理投资。（我们计划在2025年进行更多研究，以更深入地探讨量化问题，从而在此基础上开展工作。）最后，我们提供了一份行动指南，用于在组织内部使整体人工智能伦理框架落地。

我们还包含了来自人工智能伦理一线的五位高管的故事，作为 IBM IBV、Notre Dame—IBM 技术伦理实验室、IBM 人工智能伦理委员会和 IBM 隐私与负责任技术办公室之间持续合作项目的一部分。一些访谈是与牛津经济学合作进行的。

探索全面的AI伦理框架⁹

人工智能伦理与治理投资可广泛涵盖企业范围，从人工智能伦理委员会到基于设计的伦理方法，从综合治理方案到涵盖人工智能伦理与治理的培训计划，以及其他诸多努力。¹⁰（参见第13页的“人工智能伦理：前线故事”。另参阅我们的IBM IBV研究报告）[企业AI治理指南](https://ibm.co/ai-governance) 在 ibm.co/ai-governance。）

那么，组织如何开始衡量此类举措的影响呢？我们开发了一个整体的AI伦理框架来满足这一需求，并通过与30多家组织进行一系列广泛的对话来验证它。这种方法可以帮助组织理解他们在AI伦理和治理方面的投资价值。

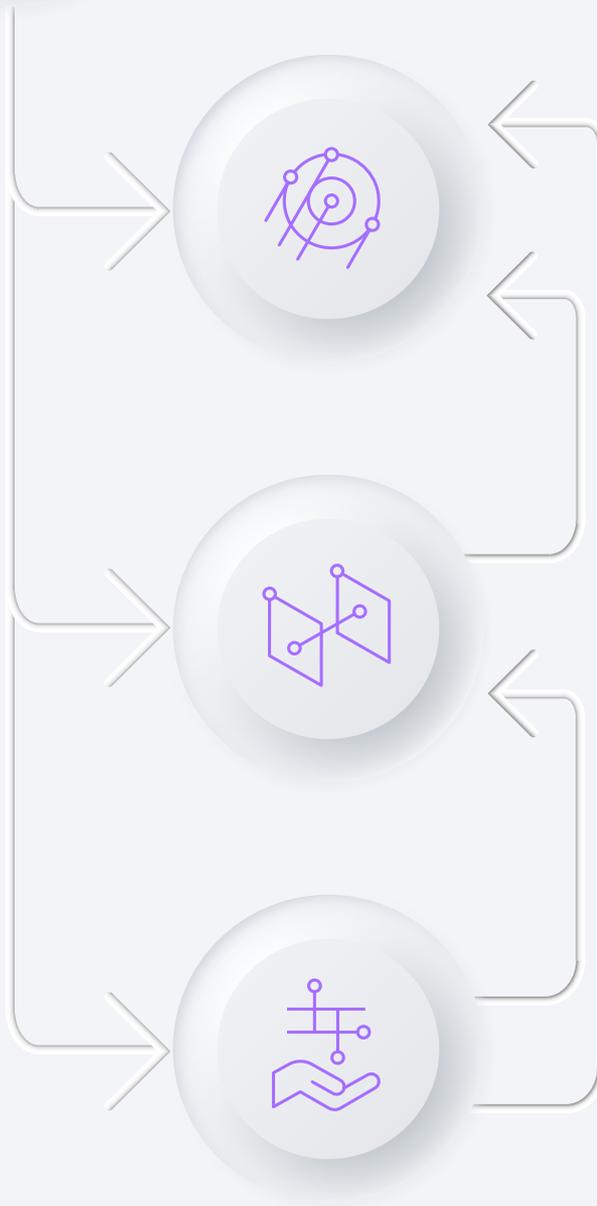
传统上，投资是通过仅以财务术语计算ROI来证明其合理性的。AI伦理投资则更具挑战性来评估，它既提供有形和无形的益处，也有助于建立长期能力。

“我们的工作不仅要为组织的使命做出贡献，还要为组织的利润率做出贡献，”SAS数据伦理实践部的副总裁雷吉·汤森德指出。“否则，它看起来就像是慈善事业，而慈善事业不会得到长期的资助。”

我们开发了一个全面的AI伦理框架，通过与超过30个组织进行的一系列广泛对话来验证它。



一个整体的人工智能伦理框架确定了组织在人工智能伦理投资方面应考虑三种ROI类型。



经济影响 (有形ROI)

指的是人工智能伦理投资带来的直接财务收益，例如成本节约、收入增加或资本成本降低。例如，一个组织通过投资人工智能风险管理可能避免监管罚款。

声誉影响 (无形ROI)

可能涉及重要的但难以量化的因素，例如支持积极回报或对股东、政府、员工和客户对组织声誉产生影响的组织品牌和文化。例如包括改善环境、社会和治理 (ESG) 得分；增加员工留任率；以及积极媒体报道。

能力 (实物期权投资回报率)

指建立能力所带来的长期效益，这些能力最初为人工智能伦理而设立，可以在整个组织中传播更广泛的价值。例如，技术基础设施或特定的伦理平台可能允许组织以能够进一步降低成本和创新的方式进行现代化。

来源：“人工智能伦理的投资回报：整体框架”。第57届年度HICSS系统科学会议论文集。2024年1月。

上述的整体人工智能伦理框架描述了三种理解人工智能伦理投资对利益相关者影响的途径：通过经济回报的直接途径，以及通过能力和声誉的间接途径。该框架涵盖了组织在人工智能伦理方面进行投资时存在的关联、利益相关者以及潜在回报。¹¹

从宏观角度看，这种方法在实践中可能如何运作？考虑投资建立人工智能伦理委员会的基础设施和人员。这项投资有助于避免监管罚款（有形影响）；增强客户信任、合作伙伴认可和商业机会（无形影响）；并有助于推动管理系统工具的开发，从而改进自动化文档和数据处理（能力）。整体人工智能伦理框架展示了人工智能伦理如何在组织中的实践和成果方面相互交织。

整体人工智能伦理框架说明了人工智能伦理如何在整个组织中贯穿始终，无论在实践方面还是在结果方面。

“损失厌恶”的价值

什么是人工智能伦理？富达投资负责数据政策的高级副总裁简洁地表示：“这是以负责任的方式使用人工智能技术，以便在与我们的客户、潜在客户和其他客户沟通时能够区分对错。”

在最近的IBM IBV研究中，72%的高管表示，如果他们认为潜在的好处可能带来伦理代价，他们会从生成式人工智能计划中退一步。这些相同的组织在收入增长方面更有可能表现优异——这种关联性很难被忽视。¹²

然而，高贵的AI意图往往被更多地谈论，而不是付诸行动。虽然我们研究中超过一半的组织公开支持AI伦理原则，但不到四分之一的人已经将其操作化。¹³不到20%的人强烈同意他们的组织在人工智能伦理方面的行动和实践符合（或超过）其声称的原则和价值观。¹⁴

“想去做是好事，但你需要真的去做，”一家全球金融服务公司的负责人工智能治理的高级领导说。“但去做，你需要资源，这需要资金。更重要的是，你需要高级管理层的决心。”

那么，投资人工智能伦理的商业合理性是什么？这通常始于一种损失规避方法：避免与监管合规相关的成本，或保留可能因客户将业务转向优先考虑人工智能伦理的企业而流失的收入。

高贵的AI意图往往被谈论多于付诸行动

。

这些动机反映了短期策略这一事实，并不会削弱其重要性¹⁵。损失厌恶会立即产生效果。正如一家全球金融服务公司负责人工智能治理的资深领导者所说，“商业案例的全部内容都是关于降低声誉风险。”

损失规避的例子包括：¹⁶

监管依据

避免监管罚款。

避免法律费用。

实施必要的合规技术机制。

启用业务以满足合规要求。

客户/合作伙伴/竞争对手的论证

消除利益相关者的担忧。

避免对企业模式构成威胁。

满足特定的客户要求或需求。

保护品牌声誉。

与竞争对手保持同步。

来自人工智能监管机构的一击

人工智能法规是行动的催化剂。欧盟人工智能法案是主要实体发布的首部全面人工智能法规。德意志电信的一位策略经理表示：“欧盟人工智能法案可能会改变全球人工智能伦理的面貌。例如，如果一家美国公司与我们合作，他们也需要遵守欧盟人工智能法案。”

欧盟的努力仅仅是开始。人工智能伙伴关系、人工智能全球伙伴关系、世界经济论坛、联合国以及经济合作与发展组织（OECD）等组织都发布了关于负责任的人工智能方法的准则和指南。¹⁷在人工智能治理中心对11个国家超过13,000人的调查中，91%的人同意人工智能需要得到仔细管理。¹⁸

鉴于对法规、监管和负责任的人工智能方法的重视，关注损失规避不仅合乎情理，而且是必要的。

利用人工智能伦理创造价值

投资人工智能伦理的好处不仅在于成本避免或损害控制。它们还有助于构建有用的能力和切实的创新，从而能够支持一个组织的长期战略。¹⁹ 这种价值创造可能比损失规避更为间接，并且需要扩展对ROI的看法。它也不会一蹴而就，并且可能需要时间才能看到可衡量的成果。

但那些对人工智能伦理有深入理解的机构可以利用这些投资：²⁰

- 允许长期计划以负责任的方式扩展人工智能。
- 打造独特且有价值组织能力，能够实现差异化。
- 提高员工效率或生产率。
- 与价值观保持一致，成为行业领导者。
- 抓住市场机遇。
- 保护弱势个人和社区。
- 提升客户满意度。
- 展现可信度和成熟度。
- 支持环境、社会及管治 (ESG) 工作。
- 长期风险管理能力提升。
- 创新以获得竞争优势。

随着人工智能技术的成熟，组织不仅可以将其融入其运营中，还可以将这项技术用于新的创新。一家领先的医疗保健和消费品零售商的高级主管解释道：“基于我们从人工智能角度为创造和丰富客户体验所采取的措施，我们已经看到了在品牌采用率、销售额增长、客户保留率和客户增长率方面的回报。”

融合两者的最佳之处

采用包含损失规避和价值创造的全面方法的组织将更有效率、更有效、更成功——同时也更符合道德。



来源：“基于人工智能伦理和治理投资的投资回报率：从损失厌恶到价值创造。”加利福尼亚管理评论。2024年7月29日。

富达投资集团的高级副总裁观察到：“公司没有意识到的是，前期投资实际上能带来显著的ROI，不仅从伦理角度来看，而且从实施任何用例的总成本角度来看。因为如果你不打好这个基础，你会在所有人逐个实施一个支柱时花更多的钱，并且无法从中获得任何复用效益。”

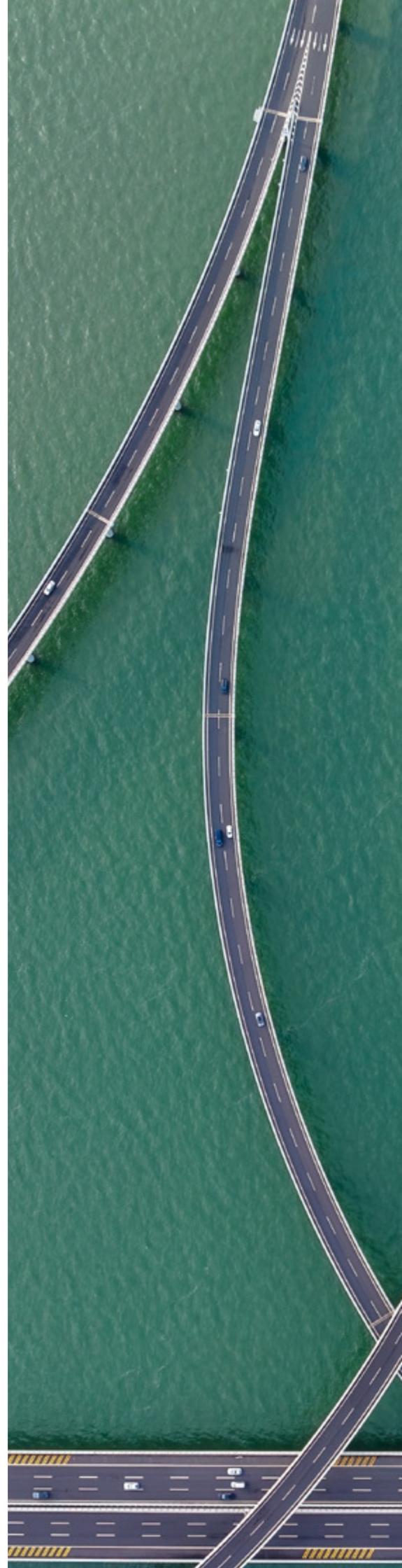
当然，这一演进的一个初步步骤是实际开发与组织战略一致并支持其战略的人工智能用例。德意志电信的战略经理指出，“你可以为客户创造人工智能解决方案，或者你可以为你的内部基础设施创造人工智能解决方案。”

起跑时，采取“防御性”的损失规避姿态来避免我们描述的陷阱，比如监管罚款、诉讼费用和声誉风险，这很本能也很合理。

但肥沃的土壤可以在转向价值创造的枢纽中找到。组织需要创建具有多功能性的技术、基础设施和平台，以支持人工智能伦理倡议，并推动更广泛的企业创新。

获取这些策略的支持和预算可能很棘手。为了说服持怀疑态度的人并克服障碍，组织应明确指出潜在的价值，包括经济回报的指标。这可以通过一个过程来实现，即组织在制定潜在投资计划时，识别相关的损失规避和价值创造的正当理由，然后评估这些潜在投资。²¹ 本质上，使用整体的AI伦理框架。

肥沃的土地可以在价值创造的转向中找到。组织需要创建具有多功能性的技术、基础设施和平台，以支持人工智能伦理计划，并推动更广泛的企业创新。



操作指南

如何收获人工智能伦理投资的回报

投资人工智能伦理不仅是正确的事，也可以是一个明智的商业决策。通过使用整体人工智能伦理框架，组织可以在人工智能伦理方面分配资源做出明智的选择，从而帮助提高人工智能程序的整体可信度和潜力。

根据 IBM IBV 研究，75% 的高管将道德视为重要的竞争优势来源。²² 经济学人信息社的一项研究表明，这些结果得到了印证，它指出可以通过产品质量、人才招聘和留住以及新的收入来源来获得竞争优势。²³

这些研究强调了采取主动方法对待人工智能伦理的关键性。组织必须考虑人工智能的治理与以往技术治理有何不同，将其贯穿于其文化、生态系统和客户互动的每一个角落。

一位领先的健康和消费品零售商的高级总监说：“您基于人类的思考来教育AI引擎，因为他们从伦理角度来看是更好的判断者。”

循着这条思路，SAS的Reggie Townsend观察到：“我们拥有一群来自不同背景和生活经历的人。我们努力工作，但我们更要用心工作。我不雇佣那些对我们做的事情没有热情的人。我们团队中有充满激情的人，我们将这份激情带到工作中。这从根本上非常重要。”

这里是我们优化人工智能伦理投资的五步指南

1

聘请您最精明的AI伦理专家，向企业高管们解释损失规避和价值观生成方法在AI伦理方面的差异。帮助高管设想利用AI伦理技术、平台和基础设施进行更广泛使用的潜力。

2

识别适用于当前人工智能使用案例的人工智能伦理和治理的具体价值生成理由。包括负责任地改进对客户的回答以及提高员工生产力和工作满意度的能力。

3

思考人工智能用例的预期利益相关者影响，并识别潜在指标。这些包括：

- 直接经济回报（例如，扩大客户群的价值）
- 无形声誉收益（例如，客户评论带来的媒体价值）
- 来自实物期权的能力和知识回报（例如，提高客户响应质量从而带来更多首次联系解决）。

4

制定一个能够实现价值创造合理化的AI伦理实施策略。根据行动3中的分析，全面识别潜在回报。这样做可以帮助优化您在人工智能伦理和治理方面的潜在回报，同时使利益相关者、生态系统和社会受益。

5

将价值创造转化为竞争优势。专注于价值创造能够在监管合规是常态的环境中提供竞争优势。

关于整体人工智能伦理框架的更多信息及行动，请参考《人工智能伦理与治理投资的回报率：从损失规避到价值创造》，《加州管理评论》，网址为<https://cmr.berkeley.edu/2024/07/on-the-roi-of-ai-ethics-and-governance-investments-from-loss-aversion-to-value-generation/>，以及《人工智能伦理投资的回报：整体框架》，网址为<https://arxiv.org/abs/2309.13057>。



人工智能伦理

图灵奖获得者

德国电信

以内部治理和教育为准备，
面向欧盟人工智能法案

德国电信的数据计划与通过人工智能应用实现数据变现以及监控欧盟人工智能法案密切相关。公司的一名战略经理领导一个团队，该团队几乎参与了公司中所有的人工智能对话，因此能够对公司的人工智能伦理方法提供整体概述。

德国电信组建了一支由高级管理人员组成的核心团队，负责评估当前和未来的人工智能计划——实际上是一个有组织的治理小组。该小组最重要的目的是帮助确保公司内部和外部的数据隐私和安全程序合规——包括欧盟人工智能法案。

“人工智能的核心是数据。它是任何人工智能产品的基本要素，”该经理表示，并补充说他认为数据也是人工智能伦理方法的关键组成部分。在将任何数据纳入其产品之前，该组织会考虑谁接触这些数据以及如何保护客户数据。除了其客户之外，德国电信还必须在可持续性和能源实践方面保护某些数据类别。

关键的是，德意志电信大力投资于教育和员工了解人工智能及其伦理使用，通常以内部研讨会的形式，包括与欧盟人工智能法案相关的培训。

“培训同事绝对是投资回报，因为它缩短了市场能力时间，我们也能研发出更多创新产品，”他说。而且，随着其持续改进的努力，德国电信体验到更多的创新和客户信任增强。

富达投资

通过重新定位用例实施获得投资回报率

负责的AI计划嵌入到富达投资每个AI用例的阶段中，从稳健的数据管理实践开始，并融入由公司AI卓越中心驱动的动态审查流程。这家金融服务公司在这些计划上投入了大量资金，使AI伦理成为其基础支柱之一——而不是一个合规性复选框练习。

Fidelity的每条业务线都有一支专门用于AI用例开发和供应商管理的团队。这项工作在外部顾问的专业知识指导下进行，并由公司合规和风险官员积极监控，他们接受了专门的AI培训。

卓越人工智能中心参与该流程的每个步骤，从供应商选择到模型评估。它位于富达的数据职能部门，并包括公司每个业务部门代表的参与，其职责范围从风险合规和审计到法律甚至信息安全。该流程还使富达能够自信地回答客户对其人工智能使用和治理方面日益增长的信息需求。

对负责任的人工智能倡议的阻力是不可避免的，因为它们可能会延迟项目或限制用例。“你必须解释，控制之所以如此重要，并非仅仅是一些随机的合规政策，而是如果我们搞错了，这会对公司产生后果，”一位负责公司数据政策的副总裁表示。富达通过将这些倡议定位为人工智能项目成功的关键要素，并简化整体治理流程，已经能够最大限度地减少反对意见。

“你必须解释清楚，控制之所以如此重要，并不仅仅是一份随机的合规政策，而是如果我们搞错了，公司会面临后果。”

Fidelity Investments 负责数据政策的副总裁。

SAS

确保一个为我们所有人打造的AI驱动未来

雷吉·汤森德是SAS公司数据伦理实践部的副总裁，他领导着一个负责协调全球组织内负责任创新原则、运营工作流程和治理结构的团队。一切始于疑问和调查。

受脆弱人群风险和人工智能日益复杂性的启发，Townsend和 closest 同事开始在 SAS 深入研究负责任的人工智能和数据伦理。他们得到了 SAS 领导层的授权，将公司长期以来对负责任创新的承诺正式化。因此，SAS 创建了数据伦理实践 (DEP)。

秉持“设计即道德”的哲学，DEP指导公司帮助员工和客户部署以数据驱动、促进人类福祉、赋权与公平的系统。这种方法迫使个人回答三个基本问题：

- 为了什么目的？
- 为了什么？
- 它可能对谁失效？

该团队帮助构建值得信赖的人工智能能力和工作流程，以帮助客户和开发者追求其负责任的人工智能目标。DEP提供的人工智能治理咨询服务正在帮助客户负责任地实施人工智能。DEP还为员工在产品开发、市场营销等方面提供关键建议。

当 Townsend 的角色和团队被创建时，希望他们的工作能够加强产品的可信度、流程和人员。反过来，这将提升该品牌作为值得信赖的 AI 领导者的声誉。

利润当然很重要。但是根据Townsend的说法，他的团队的指导原则是，无论SAS软件出现在哪里，它都不会造成伤害。“有时候，”Townsend观察道，“你不得不采取行动，因为采取行动是正确的。”

有时，你必须采取行动，因为这是正确的行动。

雷吉·汤森德兼副总裁，数据伦理部门，SAS

全球金融服务公司

用降低声誉风险来证明正向回报

对于一位负责全球金融服务公司人工智能治理的高级领导而言，人工智能发展与伦理始于教育。他提倡举办讨论伦理原则和价值观的工作坊——赋予领导层权衡取舍的能力。“我们需要以令领导层感兴趣的方式谈论人工智能，而不仅仅是流程和程序，”他观察到。

在讨论如何衡量人工智能伦理投资回报时，高级领导提出了“令人毛骨悚然的红线”隐喻。通常，组织发现自己正处于一种境地：他们正在做一件完全合法且利润丰厚的事情，但仍然不确定其行为是否符合伦理——一种感觉仿佛越过了“令人毛骨悚然的红线”。

在这种情况下，他说组织必须通过当前和未来世代的眼光来审视这一活动，并结合所有全面的政策考虑。只要这些考虑得到充分涵盖，组织就应该感到放心，即没有跨越“诡异界线”。

他也指出，“声誉风险是证明正向回报的关键因素。我们的目标是应用数据和人工智能伦理原则，以降低声誉风险。”

例如，他的团队对贷款定价进行了伦理公平性审查，该定价涉及信用评分算法。在进行此次审查时，团队分析了模型的全部165个特征，询问是否存在任何潜在因果机制，使得特定的数据特征可能与个人偿还贷款的能力相关。

最终，由于不存在因果关系，移除了三个数据特征，从而避免了在使用这项AI技术时缺乏公平性。

我们需要以能引起领导层兴趣的方式谈论人工智能，而不仅仅是流程和程序。

一家全球性金融服务公司负责人工智能治理的高级领导

一家领先的保健品和消费品零售商

以全面的AI伦理与治理策略驱动成功

该组织的资深总监发起了一项人工智能计划，通过供应商和内部产品来提供解决方案。与他的最近一次谈话涵盖了三个主要运营领域。

一个严格的治理流程。零售商的人工智能治理小组是一个集中的机构，有助于确保所有人工智能计划都完成其所需的批准步骤。为此，它召开会议，项目团队在这些会议上展示他们如何使合规措施与小组的控制计划保持一致。如果获得批准，项目将进入下一步。主任指出，作为一个大型企业，与大量合作伙伴、供应商、客户和其他生态系统打交道，在构建其人工智能能力时必须极其谨慎。

人工智能伦理引擎。无论零售商投资于基于saas、供应商或开源的产品，他们都会确保所有伦理参数在部署之前得到满足。其内部审计流程被称为“人工智能伦理引擎”。在与供应商合作时，该组织首先进行背景调查，查看其行业、客户、声誉和能力状况。这个过程可能持续两到四个月。一旦零售商选择了供应商，它就会进行试点。如果成功并满足伦理措施，合作就会继续。

利益相关者成功。该组织已大量投资人工智能能力，以提升客户互动体验，并推动市场战略和客户增长。总监指出：“人工智能本身或人类本身单独都不能成功，但如果你把两者结合起来，结果就是成功和准确的。”

在这个特定的零售商那里，用于客户服务的AI能力，例如，不会取代客户服务员工。相反，该组织投资于为这些员工提供额外的技能，从而实现员工留任。这种方法可以为顾客、员工和公司的经济效益创造利益。

人工智能本身或人类本身都不能成功，但如果你把这两者结合起来，结果就是成功和准确的。

。

一家领先的保健品和消费品零售商的高级总监

作者

尼古拉斯·伯兰特

学术项目高级副院长，信息技术、分析与运营教授，圣母大学，门多萨商学院 [linkedin.com/in/berente/](https://www.linkedin.com/in/berente/) nberente@nd.edu

马丽安娜·加纳皮尼

副教授，哲学联盟学院 [linkedin.com/in/marianna-b-ganapini-769624116/](https://www.linkedin.com/in/marianna-b-ganapini-769624116/) marianna@logicanow.com

马利亚莱娜·贝维拉卡瓦

Notre Dame大学莫纳什商学院分析学博士研究生，[linkedin.com/in/marialena-bevilacqua-6848b9132/](https://www.linkedin.com/in/marialena-bevilacqua-6848b9132/) mbevilac@nd.edu

布莱恩·戈林

合作伙伴，人工智能研究主管IBM商业价值研究所 [linkedin.com/in/brian-c-goehring-9b5a453/](https://www.linkedin.com/in/brian-c-goehring-9b5a453/) goehring@us.ibm.com

黑瑟·多米诺

全球领导者，负责人工智能倡议，IBM副总监，圣母大学—IBM技术伦理实验室 [linkedin.com/in/heatherdomin/](https://www.linkedin.com/in/heatherdomin/) hesill@us.ibm.com

弗朗切斯卡·罗西

IBM Fellow 和 AI 道德全球领袖 IBM Research [linkedin.com/in/francesca-rossi-34b8b95/](https://www.linkedin.com/in/francesca-rossi-34b8b95/) Francesca.Rossi2@ibm.com

贡献者

Sara Aboulhosn, Angela Finley, Rachna Handa, Jungmin Lee, Stephanie Meier, 和 Lucy Sieger

关于研究洞察

研究洞察是基于事实针对企业高管在关键公共部门和私营部门问题上的战略性见解。它们基于我们自身原始研究分析的调查结果。如需了解更多信息，请联系IBM商业价值研究所，邮箱地址为 iibv@us.ibm.com。

IBM商业价值研究院

二十年来，IBM商业价值研究院一直是IBM的思想领导智库。激励我们的，是产出基于研究、融合科技的战略洞见，以帮助领导者做出更明智的商业决策。

从商业、技术和社会的交汇点这一独特位置出发，我们每年对数千名高管、消费者和专家进行调查、访谈并互动，将他们的观点整合为可信、鼓舞人心且可行的洞察。

为了保持联系和获取信息，请在 ibm.com/ibv 网站上订阅IBV的电子邮件通讯。您也可以在 <https://ibm.com/ibv-linkedin> 上找到我们在领英上的页面。

在不断变化的世界中，正确的伙伴

在IBM，我们与客户合作，将商业洞察、先进研究和科技结合起来，让他们在快速变化的环境中获得独特的优势。

相关报告

企业人工智能治理指南

IBM商业价值研究院。2024年10月。 ibm.com/ai-governance

CEO的生成式人工智能指南：负责任的AI与伦理

IBM商业价值研究所。2023年10月。 ibm.com/ceo-generative-ai-responsible-ai-ethics

行动中的AI伦理

IBM商业价值研究院。2022年4月。 ibm.com/ai-ethics-action

笔记和来源

1 高林, 布莱恩, 弗朗西斯卡·罗西, 和贝丝·拉德登。
AI伦理实践：企业可信AI发展指南 IBM商业价值研究所。2022年4月。 <https://ibm.co/ai-ethics-action>

2 同上。

3 同前。

4 同前。

5 2023 商业价值生成式人工智能 市场调查报告。369位全球高管。2023年4月/5月。未公开信息。

2023年IBM商业价值研究院下一代设计调查。212位全球CxO。2023年7月。未公开信息。

7 *2023年CEO研究。人工智能时代CEO的决策：有意而行动* IBM商业价值研究所。2023年6月。 <https://ibm.co/c-suite-study-ceo>

8 *CEO的生成式人工智能指南：客户和员工体验。* IBM商业价值研究院。2023年8月。 <https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/ceo-generative-ai/employee-customer-experience>

9 贝维拉克瓦, 玛丽亚莱娜, 尼古拉斯·贝尔恩特, 赫瑟·多明, 布莱恩·戈林, 弗朗西斯卡·罗西。“人工智能伦理投资回报：整体框架。”《第57届年度HICSS系统科学会议论文集》。2024年1月。 <https://arxiv.org/abs/2309.13057>

10 “建立人工智能信任。” ibm案例研究。2023年。 <https://www.ibm.com/case-studies/ibm-cpo-pims> ; “ibm首席隐私官如何基于公司的可信人工智能框架扩展自动化以应对人工智能监管要求。” ibm。2023年。 <https://www.ibm.com/downloads/cas/7PWAWRQN> ; “ibm承诺三年内培训200万人学习人工智能, 重点关注代表性不足的群体。” ibm新闻中心。2023年9月18日。 <https://newsroom.ibm.com/2023-09-18-IBM-Commits-to-Train-2-Million-in-Artificial-Intelligence-in-Three-Years-with-a-Focus-on-Underrepresented-Communities>

11 鲍维拉夸, 玛丽亚莱娜, 尼古拉斯·贝尔恩特, 海瑟·多明, 布莱恩·戈林, 和弗朗西斯卡·罗西。“人工智能道德投资的回报：一个整体框架。”第57届年度HICSS系统科学会议论文集。2024年1月。 <https://arxiv.org/abs/2309.13057>。“经合组织人工智能原则概述。”经合组织。访问于2024年11月15日。 <https://oecd.ai/en/ai-principles>

12 *CEO的生成式人工智能指南：客户和员工体验。* IBM商业价值研究院。2023年8月。 <https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/ceo-generative-ai/employee-customer-experience>

13 高林, 布莱恩, 弗朗西斯卡·罗西和贝丝·拉德登。
AI伦理实践：企业可信AI发展指南 IBM商业价值研究所。2022年4月。已发布和未发布数据。<https://ibm.co/ai-ethics-action>

14 同前。

15 多明, 海瑟, 弗朗切斯卡·罗西, 布莱恩·戈林, 玛莉安娜·加纳皮尼, 尼古拉斯·贝尔恩特, 和玛丽亚莱娜·贝维拉克。 “论人工智能伦理和治理投资的回报率：从损失规避到价值创造。” *加利福尼亚管理评论* 2024年7月29日。 <https://cmr.berkeley.edu/2024/07/on-the-roi-of-ai-ethics-and-governance-investments-from-loss-aversion-to-value-generation/>

16 详见原文。

17 贝维拉夸, 玛丽亚莱娜, 尼古拉斯·贝尔恩特, 赫瑟·多明, 布莱恩·戈林, 和弗朗西斯卡·罗西。“人工智能伦理的投资回报：整体框架。”第57届年度HICSS系统科学会议论文集。2024年1月。 <https://arxiv.org/abs/2309.13057>。经济合作与发展组织。(2023)。经济合作与发展组织人工智能原则概述。经济合作与发展组织人工智能(AI)原则 - OECD.AI。 <https://oecd.ai/en/ai-principles>

18 德雷克斯勒, 诺艾米, 大卫·麦卡菲里, 劳伦·卡恩, 凯特·梅斯, 马库斯·安德尔斯球格, 艾伦·达福, 迈克尔·C·霍罗威茨和宝宝张。“初步调查结果：美国和欧洲公众越来越普遍并日益同意人工智能需要得到谨慎管理。” *人工智能治理中心* 2023年4月17日。 <https://www.governance.ai/post/increasing-consensus-ai-requires-careful-management>

19 高林，布莱恩，弗朗西斯卡·罗西和贝丝·鲁登。
AI伦理实践：企业可信AI发展指南 IBM商业价值研究所。2022年4月。<https://ibm.co/ai-ethics-action>

20 多明，海瑟，弗朗西斯卡·罗西，布莱恩·戈林，玛莉安娜·甘帕尼尼，尼古拉斯·贝尔恩特和玛丽亚莲娜·贝维拉卡瓦。“关于人工智能伦理和治理投资的回报率：从损失规避到价值创造。” *加利福尼亚管理评论*。
2024年7月29日。<https://cmr.berkeley.edu/2024/07/关于人工智能伦理和治理的投资回报率：从风险规避到价值创造的投入/>

21 同上。

22 格林，布赖恩，弗朗西斯卡·罗西，贝丝·鲁登。
AI伦理实践：企业走向值得信赖的AI指南。IBM商业价值研究所。2022年4月。<https://ibm.co/ai-ethics-action>

23 “保持领先——负责任人工智能的商业案例。”经济人智库。2020年10月8日。<https://www.eiu.com/n/ staying-ahead-of-the-curve-the-business-case-for-responsible-ai/>

© 版权所有 IBM 公司 2024

IBM公司新奥尔丘路阿蒙克，
纽约州10504

美国制造 | 2024年12月

IBM、IBM标志、ibm.com是国际商业机器公司 (International Business Machines Corp.) 的商标，在世界许多司法管辖区注册。其他产品和服务名称可能是IBM或其他公司的商标。有关IBM商标的当前列表可在网站上找到，网址为：ibm.com/legal/copytrade.shtml中的“版权和商标信息”。

本文件自首次发布之日起即为最新版本，IBM 可随时更改。并非所有产品和服务都在 IBM 运营的每个国家/地区提供。

本文件中的信息按“原样”提供，不提供任何保证，无论是明示的还是暗示的，包括但不限于对适用性、特定用途适用性以及任何非侵权保证或条件的保证。IBM产品的保证根据其提供协议的条款和条件进行。

本报告仅供一般性指导之用。它并非旨在替代详细研究或专业判断的行使。IBM对任何依赖本出版物而遭受的任何损失概不负责。

本报告中使用的数据可能源自第三方来源，IBM不对其独立进行核实、验证或审计。使用此类数据的结果按“原样”提供，IBM不作任何明示或暗示的陈述或保证。

