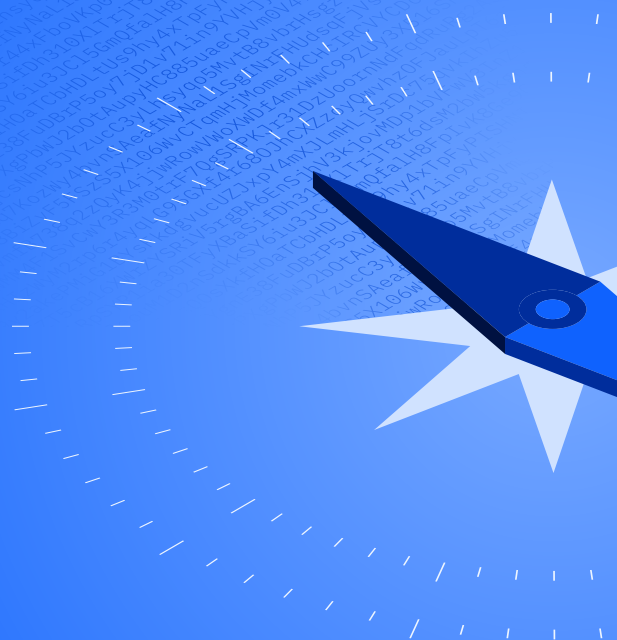


# 伦理工作责无旁贷

不同于以往的任何技术，生成式 AI 正在迅速颠覆商业和社会形态，迫使企业领导者刻不容缓地重新思考其假设、计划和战略。

为了帮助 CEO 们掌握快速变化的形势，IBM 商业价值研究院 (IBM IBV) 发布了一系列有针对性、基于研究数据的生成式 AI 指南，涵盖数据网络安全、技术投资策略和客户体验等主题。

这是本指南的第八部分，重点关注“负责任 AI 与伦理”。



## 人类价值观是负责任 AI 的核心

随着企业纷纷开始探索生成式 AI 在处理各种令人难以置信的新任务方面的能力，CEO 必须充当“掌舵者”，带领企业为生成式 AI 设定明确的 **合理使用方针**。

每个用例都有各自的伦理困境和合规性问题：企业应当如何保护敏感数据？企业应当如何以尊重版权的方式使用 AI？AI 的输出是否存在偏见、歧视，或者根本就是错误的？

尽管解答这些问题需要整个团队的努力，但 CEO 必须为组织设定方向性的伦理方针。CEO 制定的这一路线将定义企业如何在尖端创新和传统的诚信与信任原则之间找到平衡。

CEO 必须实施全面确保透明度和问责制的政策与流程，明确技术的使用方式和场景，以及数据集或底层基础模型的来源。这是一项持续性工作，因为组织需要持续监控和评估其 AI 产品组合，以确保始终符合不断变化的政策和流程。

领导者必须培育一种注重 AI 伦理的文化，旨在优化 AI 的积极影响，同时以优先保障人类能动性和福祉以及环境可持续性的方式，降低为所有利益相关者带来的风险和不利影响。这将是一项仅靠技术无法应对的社会技术挑战。若要取得大规模成功，就必须对组织文化、工作流程和框架进行持续投资。

公共舆论将评判企业的行为是否符合道德规范，以及是否符合消费者价值观。这样一来，就可以主观地衡量公平性和适当性。但是，合规性无法如此衡量。

### IBM 商业价值研究院甄别出了每位领导者都需要了解的三个要点：

1. CEO 不能在 AI 伦理问题上推卸责任。



2. 客户将评判您所做的每一个决定。不要损害他们的信任。



3. 一些企业因监管模糊性而陷入困境。



### 现在，每位领导者都需要采取以下三项行动：

1. 让伦理团队在决策中占有一席之地——而不是仅给予口头上的支持。



2. 通过满足客户期望来赢得信任。



3. 为所有 AI 和数据投资做好伦理和监管准备。



## 1. 战略 + 生成式 AI

需要了解的事项 →

### CEO 不能在 AI 伦理问题上推卸责任。

生成式 AI 就像是“狂野西部的淘金热”。对财富的追逐已经超过了规则和法规，早期的探矿者有机会一夜暴富。

但是，这是以什么为代价呢？如果组织太急于求成，而未考虑复杂的 AI 伦理和数据完整性问题，就可能会因短期利益而损害长期声誉。

高管们了解其中的风险：58% 受访高管认为采用生成式 AI 存在重大伦理风险，如果没有新的治理结构或者至少更加成熟的治理结构，就无法管理这种风险。然而，许多高管都难以将原则付诸实践。尽管 79% 的受访高管表示 AI 伦理对其企业级 AI 方法很重要，但只有不到 25% 的受访高管实施了 AI 伦理的共同原则。

正是因此，CEO 必须掌控全局并为其他人开辟道路。直接向 CEO 寻求 AI 伦理指导的高管人数要比向董事会、总法律顾问、隐私官或风险与合规官寻

求 AI 伦理指导的高管人数多出大约 2 倍。80% 的受访高管表示，企业领导者（而不是技术领导者）应当对人工智能伦理负主要责任。

除了决策以外，CEO 还必须负责向其他领导者普及关于新兴伦理问题的知识。通过将关于可信 AI 的对话提升到其他高级管理层和董事会的层面，CEO 可以确保这些关键利益相关者不会被边缘化。采取积极主动、包容性的方法有助于确保每个人都了解风险以及管理这些风险的明确行动计划。这样一来，组织可以加快行动速度，同时保持领导层协同一致。

尽管 **79%** 的受访高管表示 AI 伦理对其企业级 AI 方法很重要，

但只有不到 **25%** 的受访高管实施了 AI 伦理的共同原则。



## 1. 战略 + 生成式 AI

需要采取的行动 →

让伦理团队在决策中占有一席之地——而不是仅给予口头上的支持。

采取行动，将意图付诸实践。为伦理团队、政策和监测流程提供支持。适时向董事会和外部报告进展情况。

**即使超出了自己的舒适区，也应做到掌控一切。**考虑任命一位首席 AI 伦理官或其他领导者来负责整个企业的 AI 实践，并明确当前高管角色的责任。确保各个业务部门和职能部门的高管遵循共同的 AI 伦理目标。确保在决策中纳入合适的人员，包括风险与信息安全主管。

**创建有效的人机偕行关系。**设定基调并在自动化与增强之间取得适当的平衡。建议创建并采用 AI 设计指南，并将有关算法责任的具体部分纳入企业的商业道德准则。积极制定培训、AI 和数据素养以及变革管理议程。以尊重和关怀的方式对待受影响的员工。

**建立“伦理互操作性”。**发现并吸引注重 AI 的关键技术合作伙伴、学者、初创公司和其他业务合作伙伴，增强您的创新生态系统。确认价值观是企业形象和文化的一部分，并确保所有合作伙伴的价值观相一致。

## 2. 信任 + 生成式 AI

需要了解的事项 →

客户将评判您所做的每一个决定。  
不要损害他们的信任。

建立一个蓝筹品牌需要数十年的时间，而摧毁它只需要几天的时间。在数据泄露和不信任的时代，消费者、员工和合作伙伴对不以诚信行事的企业毫不宽容。

超过一半 (57%) 的消费者表示对企业使用其个人或商业信息的方式感到不安，37% 的消费者曾为了保护隐私而选择更换了品牌。消费者认为，零售、保险和公用事业等许多传统行业的公司在负责任地使用技术方面排名最低。

此外，合作伙伴、投资者和董事会也在密切关注公司，尽管他们似乎更倾向于支持负责任的 AI 发展。受访 CEO 们表示，他们在加速采用生成式 AI 方面感受到来自董事会的更大压力，这一压力是减缓采用生成式 AI 的 6 倍以上。

员工也渴望为与自己价值观相符的企业工作。69% 的受访者表示，他们更愿意接受那些他们认为具有社会责任感的组织的工作机会，45% 的受访者表示他们更愿意接受较低的工资在这样的组织中工作。

总的来说，这些观点展示了为什么拥有更强大数据实践的企业能够创造更多的价值。根据我们的 2023 年首席数据官 (CDO) 研究，此类领先企业中大约 80% 的 CDO 表示，其组织在数据伦理、组织透明度与问责制以及网络安全领域表现出色。

# 45%

的员工表示更愿意接受较低的工资在具有社会责任感的组织中工作。



## 2. 信任 + 生成式 AI

需要采取的行动 →

### 通过满足客户期望来赢得信任。

确保 AI 伦理和数据完整性成为整个组织自上而下的优先事项。建立自下而上的协作信任文化，增强专注度。让伦理成为每个人的责任，并让治理成为一项集体共同目标。

**领先并超越客户的伦理期望。** 认识到客户每天在生活的方方面面都会面临伦理问题。通过明确定义您的伦理价值观来建立信任。广泛、透明地传达您的伦理价值观。一次又一次，反复传达您的伦理价值观。

**始终以人为本。** 对员工进行再培训，确保在工作中合理运用 AI，避免不当运用 AI。针对员工和合作伙伴开展 AI 伦理和偏见识别培训，强调可信 AI 的重要性。明确在哪些情况下应当向专家寻求帮助。让您的团队成为组织内外的伦理管理者，与客户建立信任关系。

**让每个人都负起责任。** 带头担负起个人责任，并期望高管和其他员工也承担个人责任。要求业务和 AI 领导者签名，用他们的个人声誉作为担保 — 当然，首先从您自己做起。将技术伦理作为寻源采购伦理标准的关键部分。将这些承诺公之于众。

### 3. 合规性 + 生成式 AI

需要了解的事项 →

## 一些企业因监管模糊性而陷入困境。

欧盟的 AI 法案即将出台。中国正在制定健全的法规和指导方针。全球的企业领导者都感受到了为此做好准备的压力，但他们并不确定自己要做好哪些准备。

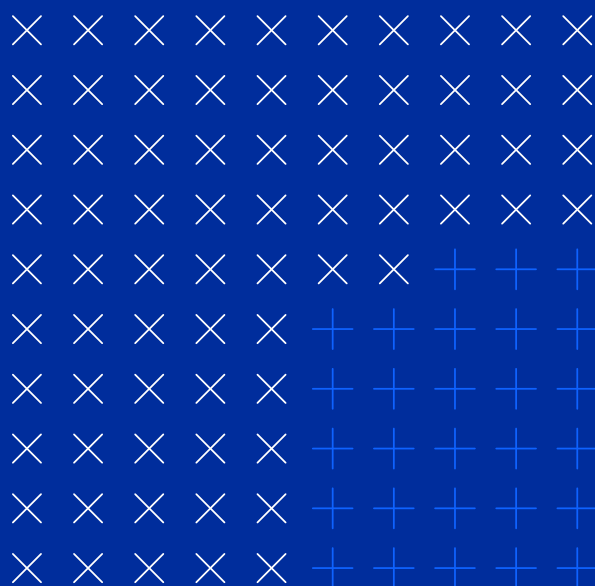
在全球范围内，只有不到 60% 的受访高管认为其组织已经为 AI 监管做好了准备，69% 的受访高管预计会因采用生成式 AI 而受到监管罚款。面对这种不确定性和风险，CEO 们正在踩刹车。超过一半 (56%) 的受访 CEO 推迟了重大投资，等待他们对 AI 标准和法规建立清晰的认识。总体而言，72% 的组织将因伦理顾虑而放弃生成式 AI 带来的收益。

尽管有一些企业仍然深陷在监管泥潭中，但积极解决伦理问题的企业可以满怀信心地继续前进。无论法规如何演变，良好的数据和 AI 治理都是必不可少的，而且从一开始就实施负责任且可信的 AI 将有助于企业在时机成熟时建立良好的合规性。此外，具有强大伦理和治理能力的企业将有机会从竞争中脱颖而出 — 具体来说，四分之三的受访高管将伦理视为差异化竞争优势的来源。

将伦理置于优先位置可以帮助 CEO 在面对法规不确定性时果断行动，并在不损害价值观的情况下享受生成式 AI 的早期收益。或许正是因此，AI 伦理投资呈持续增长之势 — 其占 AI 支出的比例从 2018 年的 3% 增加到 2025 年的近 9%。

# 72%

的高管表示，其组织将因伦理顾虑而放弃生成式 AI 带来的收益。



### 3. 合规性 + 生成式 AI

需要采取的行动 →

## 为所有 AI 和数据投资做好伦理和监管准备。

新兴法规的基本轮廓已经足够清晰，可以为 CEO 的决策提供指引。一旦细节敲定，您就可以重新修正和校准方向。无论最终结果如何，都应当专注于建立可信 AI。良好的治理至关重要。

**沟通、沟通、再沟通。**倡导建立有意义的监管。确保应用场景易于解释，AI 生成的工件清晰可识别，AI 训练保持透明且接受持续批判。

**记录一切，确保可追溯性。**建立归档文化，持续记录组织中使用 AI 的所有实例和相关治理，有效管理风险。通过清单来记录使用 AI 的每个实例，确保 AI 生成的资产可以追溯到基础模型、数据集、提示或其他输入。同时将这些源信息植入到数字资产管理和其他系统中。

**作为掌舵者，随时调整航向。**做好准备，随时根据监管风向的转变和新出台的法规做出调整。



IBM 商业价值研究院

CEO 生成式 AI 行动指南

# 负责任 AI 与伦理

本页分析所依据的统计数据来自 IBM 商业价值研究院联合与牛津经济研究院和 MomentiveAI/SurveyMonkey 合作开展的专项调研、IBM 商业价值研究院已发布的多项研究，以及来自思科的一项外部参考资料。2023 年 8 月至 9 月，针对 200 名美国高管开展了第一项关于生成式 AI 和供应链的调研。2023 年 5 月，针对 300 名美国高管开展第二项调研，主题是生成式 AI 与人才的交叉点。2021 年，针对 10 个国家/地区的 16,349 名消费者开展了第三项调研，了解他们对社会责任和可持续发展的看法。IBM 商业价值研究院已发布的研究包括：《2023 年 CEO 研究：AI 时代的 CEO 决策力》、《2023 年 CDO 研究：化数据为价值和《AI 伦理道德方略》（2022 年）。

## IBM 商业价值研究院

IBM 商业价值研究院 (IBM IBV) 创立二十年来，凭借 IBM 在商业、技术和社会交叉领域的独特地位，我们每年都会针对成千上万高管、消费者和专家展开调研、访谈和互动，将他们的观点综合成可信赖的、振奋人心和切实可行的洞察。

需要 IBV 最新研究成果，请在 [ibm.com/ibv](https://ibm.com/ibv) 上注册以接收 IBV 的电子邮件通讯。您可以在 Twitter 上关注 @IBMIBV，或通过 <https://ibm.co/ibv-linkedln> 在 LinkedIn 上联系我们。

访问 IBM 商业价值研究院中国官网，免费下载研究报告：<https://www.ibm.com/ibv/cn>

扫码关注 IBM 商业价值研究院



官网



微博



微信公众号



微信小程序



© Copyright IBM Corporation 2023

国际商业机器 (中国) 有限公司  
北京市朝阳区金和东路 20 号院 3 号楼  
正大中心南塔 12 层  
邮编: 100020

美国出品 | 2023 年 11 月

IBM、IBM 徽标、ibm.com 和 Watson 是 International Business Machines Corporation 在世界各地司法辖区的注册商标。其他产品和服务名称可能是 IBM 或其他公司的商标。以下 Web 站点上的“Copyright and trademark information”部分中包含了 IBM 商标的最新列表：[ibm.com/legal/copytrade.shtml](https://ibm.com/legal/copytrade.shtml)。

本档为自最初公布日期起的最新版本，IBM 可能随时对其进行更改。IBM 并不一定在开展业务的所有国家或地区提供所有产品或服务。

本文档内的信息“按现状”提供，不附有任何种类的（无论是明示的还是默示的）保证，包括不附有关于适销性、适用于某种特定用途的任何保证以及非侵权的任何保证或条件。IBM 产品根据其提供时所依据的协议条款和条件获得保证。

本报告的目的仅为提供通用指南。它并不旨在代替详尽的研究或专业判断依据。由于使用本出版物对任何企业或个人所造成的损失，IBM 概不负责。

本报告中使用的数据可能源自第三方，IBM 并未对其进行独立核实、验证或审查。此类数据的使用结果均为“按现状”提供，IBM 不作出任何明示或默示的声明或保证。

6JAM3VB7-ZHCN-00