技术愿景2025

人工智能：自主宣言

信任是否是人工智能无限可能性的极限？

accenture

人工智能：自主宣言

信任是否是人工智能无限可能性的极限？

欢迎来到我们对于2025年的技术展望。这份我们年度科技趋势报告的第25版，正值技术及人类历史的关键时刻。随着越来越多的领导者认识到不断革新利用技术、数据和人工智能的必要性，他们比以往任何时候更需要深入理解人工智能。为何？因为人工智能技术扩散的速度前所未有，且仍在加快——在企业层面创造新的创新机会，包括实现效率的新方法、经营企业核心的方式，以及与客户互动的新商业模式。

我们将AI视为新的数字革命，因为它既是一种技术，也是一种新的工作方式。我们相信它将在企业运营的每个领域得到应用，并对所有相关的人和事物产生网络效应。其影响已经显现，随着公司继续扩大AI规模——并将生成式AI作为变革的催化剂——它将解决新的问题，创造新的发明，改变我们的工作和生活方式，并转型行业和政府。

安永研究显示，仅有36%的高管表示他们的组织已经实现了通用人工智能解决方案的规模化，仅有13%的人报告称实现了显著的企事业级影响。随着我们将2025年视为规模化人工智能的年份，我们正在积极为他们提供更快、更安全地实现这一目标的工具。

今年《技术愿景》探讨了人工智能从自动化向转型发展的未来。

使人们能够自主行动的推动者——为他们提供执行新任务和比以往任何时候都更好地执行其他任务的能力。考虑一下，随着人工智能进入新的和陌生的领域，重新构想的可能性与机遇。为了真正理解和利用这种潜力，企业将创建他们自己的、独特的AI认知数字大脑，这将彻底改变技术在其企业以及与员工之间所扮演的角色。这将极大地颠覆企业技术系统的设计、使用和运营方式；充当品牌大使；并通过为机器人身体提供动力而在物理世界中存在。当人工智能在组织中普及时，它使人们和人工智能能够彼此发挥最佳状态。

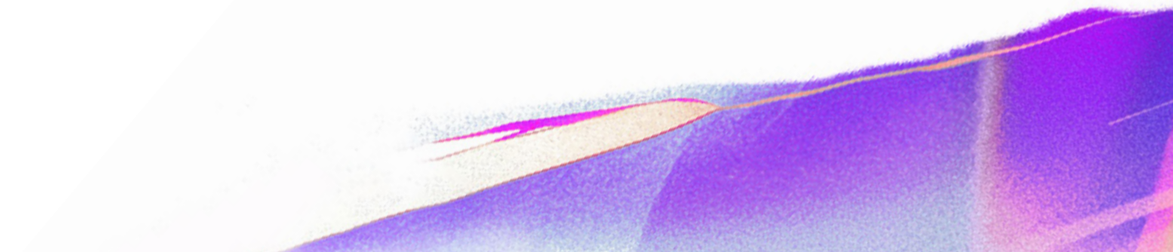
领导者们意识到创造这一未来的挑战，其中包括对核心技术的初期高额投资、以数据为中心的质量以及人才和新技能。在这些挑战中，信任是最主要的。

我们的研究发现，77%的执行官认为，只有在建立在信任基础之上时，才能解锁AI的真实益处。领导者必须通过确保AI的准确性、可预测性、一致性和可追溯性，以及负责任地使用AI，在客户和员工中建立对数字系统和AI模型的信任。人们相信AI能够如预期和公正地运行——超越任何技术方面——这是我们必须做对的一个基本要素。

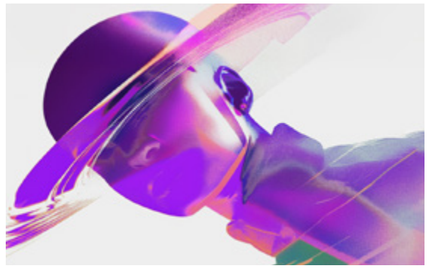
我们相信我们能够做到。我们将这个技术新时代视为一个机会，通过系统化地注入对人工智能的信任，使企业和人们能够实现其不可思议的革新潜力。共同合作，我们现在可以为人工智能自主且帮助我们共同实现更多成就的勇敢未来做好准备。

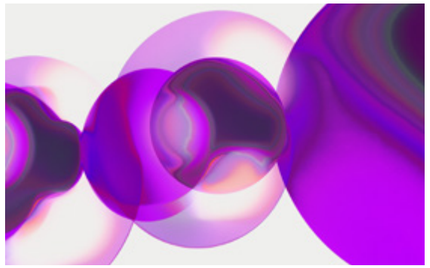
Julie Sweet 董事长兼
首席执行官

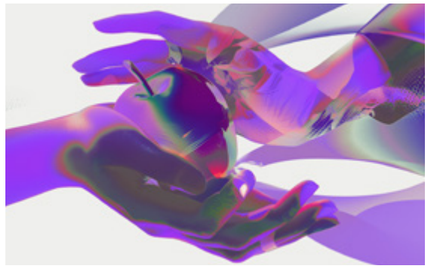
卡吉克·纳赖恩 集团首席
执行官兼首席技术官

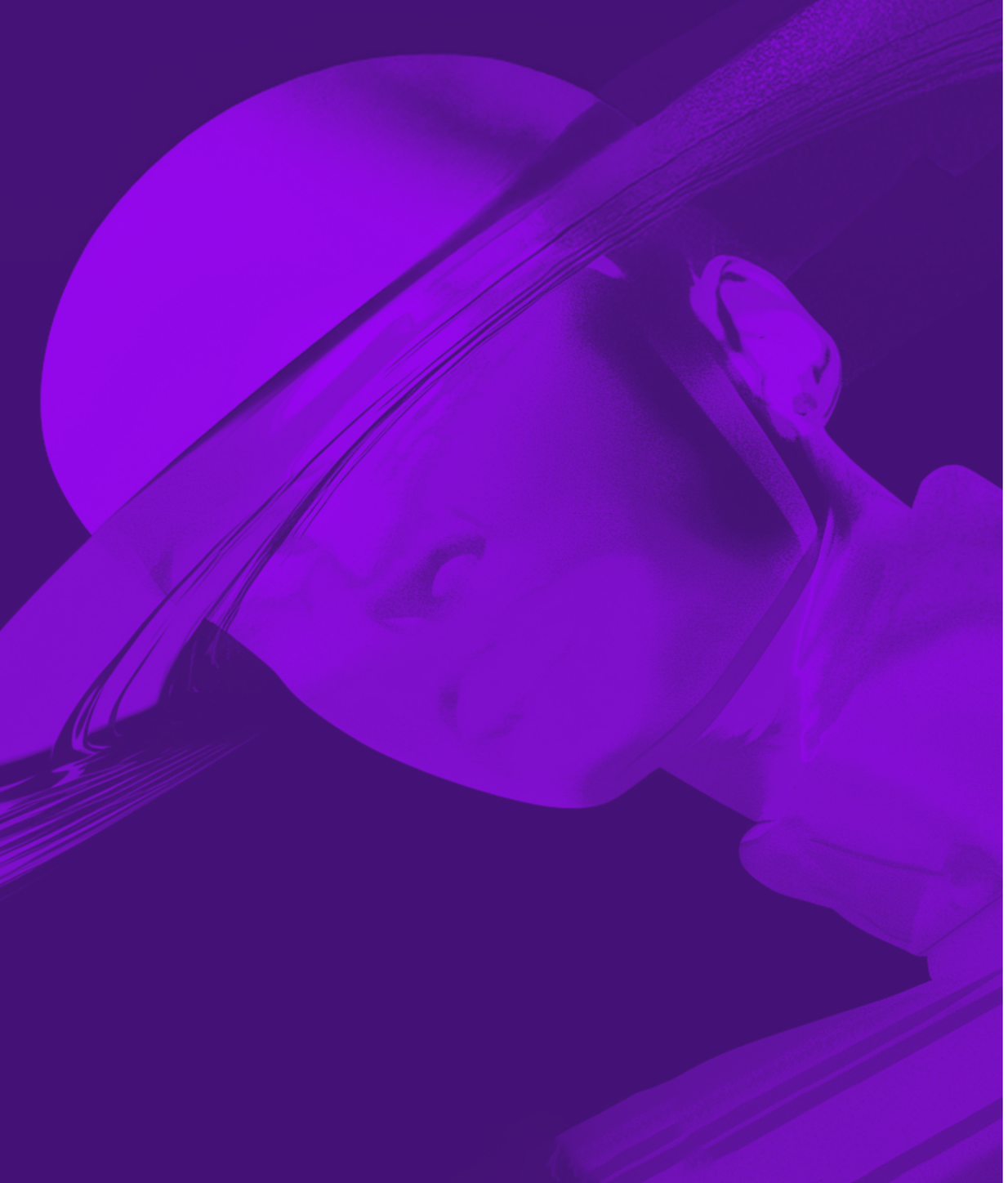


目录









人工智能：自主宣言

信任是否是人工智能无限可能性的极限？

我们正在进入一个新的篇章
在技术领域——一个被塑造的领域——
人工智能的泛化。如今，
广泛的可访问和
始终存在的AI将驱动新的
全过程的自主水平
该业务，演变能力
通过科技、数据及
人工智能。它将带来近乎无限的
创新的可能性
并且增长，但也挑战
企业对系统的信心
他们思考信任的方式。

人工智能竞赛的热潮无可否认。

我们之前见过这种情况。1997年，加里·卡斯帕罗夫在与IBM的深蓝（Deep Blue）的六盘棋比赛中失利。[1] 这是第一次计算机击败国际象棋大师，经过数十年的测试，人类与机器在此游戏中对抗。这场胜利引发了一场关于人工智能和未来的兴奋和疑问的风暴。现在，一场新的竞赛正在进行。许多公司正在构建今天的尖端AI模型，他们的目标是在通用人工智能（AGI）方面。[2,3] 并且，就像以前一样，这场比赛吸引了商业领袖、政府和全世界人们的关注。

但是这是一个烟幕弹——大多数商业领袖都无法承受的干扰。总有一天，通用人工智能将会产生巨大的影响，但今天它仍然遥远，需要解决深层次的技术和伦理挑战。相反，领导者们看到眼前更为紧迫的问题至关重要：人工智能的普及，这将在新水平上为企业系统、工作队伍和运营带来自主性和能力，在通用人工智能发挥作用之前就已经到来。

人工智能的泛化

为了理解这种关于人工智能的概括，人们只需环顾四周就能看到人工智能在我们的生活中越来越根深蒂固。自从卡斯帕罗夫的比赛以来，几乎已经过去了30年，而现在能够使Deep Blue看起来像是一个普通玩家的模型都坐在每个人的口袋里。图灵测试，曾经被认为是机器智能的最高标准，现在每天都被人们与大型语言模型（LLM）支持的客户服务机器人和销售人员交流中所打破。今天的AI模型已经摆脱了过去深度但具体且线性的方法，并展示了前所未有的自主性—在他们如何学习、如何处理任务以及最终能够做什么方面。他们将这种自主性带到了工作中，75%的知识工作者报告使用生成人工智能；在如何与技术互动方面，作为编码助手并通过扩展语音助手功能；以及几乎在一切事物上，从机器人到汽车，到医疗保健。[4,5,6,7,8,9,10] 高度能效的高级人工智能

在我们的生活的每一个维度上扩散，即时可获取，并且——实际上——始终存在。

这是我们需要真正关注的真正颠覆。因为现在，尽管高管们竞相实施这一代人工智能，但很少有人能超越各个独立的部件，真正理解他们实际上正在构建的范围：人工智能“ **认知数字大脑** 这将会彻底重塑技术在企业以及人们的日常生活中所扮演的角色。

领导者必须充分理解的是，人工智能最重要的特征是其学习能力。当人工智能变得通用，并且企业将其扩散到业务中，人们将其融入他们的生活中时，它有可能成为比它提供的新功能和能力多得多的东西。企业并不仅仅是赋予员工力量，创造新的客户服务渠道，或自动化部分业务运营。他们正在采用一种带有广泛通用知识的技术，这种技术本质上是由于其学习能力而定义的，并且他们正在 *教学* 这与业务的一部分有关。当人们使用它时，他们通常是... *进一步* 教授它关于它们的喜好、偏好和需求。

如果有意构建，企业可以将他们正在追求的所有分布式人工智能努力整合，构建一个认知数字大脑。他们可以将工作流程、机构知识、价值链、社会互动以及关于商业和世界的许多其他关键数据硬编码到一个系统中，该系统能够以比以往任何时候都高的水平理解和行动。

一个人可以利用这种力量做什么？一家企业如何在全体员工中部署它呢？

世界在广泛扩散带来的全面影响下会是什么样子？随着领导者开始结合他们的AI泛化努力，似乎不可避免的是，他们将很快提升和赋权个人，推动并帮助企业运营，彻底重塑产业，甚至提升国家。

InSilico Medicine，一家制药公司，利用生成式人工智能在不到30个月的时间内，完成了从药物发现到一期临床试验的进程，这大约是通常所需时间的一半。[11] 他们使用了一个基于组学和临床数据的模型进行微调，以识别药物治疗的潜在靶点。为了开发可能的药物组成，他们使用了一个由500个预测和预训练模型组成的生成化学引擎。对于Insilico来说，人工智能是他们工作的核心——塑造围绕它的业务和行业。

在每一个层面都有一个认知数字大脑。

可以看到这一趋势可能会有困难；在每一层规模上，它都表现出略微不同的形式。但总体而言，AI的下一阶段将为其触及的每一件事注入增强的能力和更大的自主性。对于 **个体** 认知数字大脑将作为副驾驶或助手运行，它将理解他们的工作，学习他们的偏好，并通过其互动来了解他们，旨在帮助他们成为更优秀的自己。 **企业** 它可能更像是一个中枢神经系统——企业架构进化成能够捕捉业务的集体知识、独特的差异化特征以及其文化和个性的东西，并成为其关键协调者（甚至是自主操作者）。对于 **行业** 这可能看起来像是行业内公司之间的通用框架和通信协议，或者是编码界定行业大挑战的引擎——这些模型将有助于我们增长对物理学、遗传学、运动等方面理解。并且对于 **国家** 和 **政府** 它汇集了独特的知识、语言、文化、法律和安全，以帮助行业、公司和公民参与。关键的是，这些认知数字大脑不会在孤岛中运作。当它们开始在不同层面互动时，它们将创造一股上升的智能浪潮，提升所有参与方的能力。

这是为什么它是一个“自治宣言”。
我们或许会称呼它们为不同的名称，但在整个范围内，其演变过程是相同的：自主人工智能系统的普及正在全范围内发生。

社会将把世界提升到下一个层次的能力、表现和进步。它将推动一个进化的过程，通过人工智能的认知，使世界在所有层面上得到增强，并产生一场前所未有的自主浪潮，这将重塑我们所知的科技和企业。

首先的想法可能是，这仅仅是使用AI从自动化过渡到数字系统的自主性。这并不错误，但只是故事的一部分——AI正在以数十种方式推动自主性。它为人们提供了他们原本不具备的技能，让他们能够以前所未有的主动性减少摩擦地行动。它为机器人提供了关于世界的新程度上下文和推理能力，使他们能够承担更广泛和更复杂的任务，最重要的是，以前所未有的方式与人类混合在一起。当然，智能体和智能体AI

认知数字大脑的构成要素是什么？

认知数字大脑将成为企业决策和持续学习的中央神经系统。它用于驱动企业的未来抱负，如基于意图的架构，由四个相互连接的层次组成，共同组织、处理和作用于信息。

知识： 技术与知识图谱、向量数据库一样，能够从整个企业和外部收集、组织和管理数据。

模型： 大型生成式AI模型以及经典机器学习和深度学习模型执行批判性思考和推理功能，以将数据转化为可行的成果。

代理商： 设计为问题解决者，以最小的人为干预处理任务，并在时间推移中学习和成长，人工智能代理将规划、反思和适应性融入其中。

建筑： 一个全面的骨干架构是将AI实验转化为企业级解决方案的关键。它将智能扩展到整个组织以及现有的工作流程中，并实现可重复性，因此解决方案可以一次性制作并重复使用。

系统开始承担整个工作流程或客户互动，无需持续的人为干预，同时保持战略监督。利用这种自主性将扩展企业认为可能性的极限。Accenture的研究发现，凭借其重新构想和增强复杂任务的能力，生成式人工智能有望在领导AI采用的公司中推动生产力提高20%。[12]

我们今天拥有的，是无限增长和创新的火花——以及颠覆的力量。随着组织内外的自主性不断增长，摩擦减少，我们可以更快地完成更多任务，先行者将能确保持续数十年的优势。未能及时行动或等待过久，将让竞争对手，无论是新的还是旧的，都有机会颠覆行业规范，正如我们在数字化时代所看到的那样。再考虑一下这一点：今天全球互联网市值不到1%是在Netscape Navigator将互联网推广给世界的前两年成立的。[13]

现在，自ChatGPT发布以来已经超过两年。我们进入这一代人工智能的探索才刚刚开始，鉴于如此高的风险，企业现在开始至关重要，以免他们被远远抛在后面。

为了更深入地了解数字
平台、数据与人工智能、以及数字基础设施
企业正通过...获得赋能以实现增长。
变革与颠覆，请参阅我们的研究成果。
[重构以数字化核心为基石]

我们处于众多可能路径的起点。
向前。实现全部潜力的关键。

人工智能的关键在于企业领导者如何选择利用它所赋予的新维度自主性。但在这个新世界中取得成功并做出正确的选择将不是一项小任务。与自主性这一理念紧密相连的是信任的基础——对企业而言，信任将是明天增长的最重要的保障。

唯一的限制是信任。

思考一下信任如何定义人类体验——例如，父母与子女之间的关系。我们用护栏包围婴儿。从字面上的，比如婴儿床里的那些，到比喻意义上的，比如切食物或在家里的尖锐角落上贴上软垫。随着他们成长，我们开始更多地信任他们。他们过马路时不需要你牵着他们的手，但你仍然会陪在他们身边。他们可以自己在外面玩耍，但只能在围栏内。随着我们的信任不断增强，我们绘制的护栏边界也越来越宽。直到有一天，他们完全成长为了成年人。我们还会检查他们的情况——但现在，他们是独立个体，拥有做出自己决定的自主权。

关键在于，这个例子展示了信任和自主性是如何密不可分的。但它也表明了信任领导者现在需要考虑的细微差别。父母与子女之间的关系取决于信任的情感和认知两个方面。安全护栏有助于营造一个充满爱心、培养和安全的成长环境，同时也帮助父母建立对孩子决策能力的信心。当涉及到人时，我们不需要在这两个维度之间进行区分，但与技术相比，它们是不同的挑战，需要不同的解决方案。

截至目前，技术系统主要基于规则。尽管这些系统相对不那么智能，但它们高度可预测，因此更加值得信赖。因此，它们在企业中的采用和扩散非常普遍。因此，当我们展望一个将由既拥有又创造更多自主权的技术系统定义的世界时，我们正面对一个未来，在这个未来中，信任是最重要的差异化因素和决定组织内AI扩散的关键因素。毕竟，我们只能让系统拥有我们所信任的自主权。

当然，大多数领导者都对不良行为者如何通过深度伪造或其他手段更有效地散播虚假信息以及更巧妙的钓鱼攻击了解得非常清楚——电子邮件技巧或模仿真实人类的语音。或者偏见的决策如何在即使有AI的情况下崭露头角。明确地说，这些是真的问题，内容水印技术或深度伪造检测工具的需求不断增长，紧迫寻求解决这些问题的措施。但这个故事将AI信任的话题仅仅局限于不良行为者和利用。这只是部分真相。为了实现真正的自主权——在系统、整个劳动力以及顾客中——领导者需要对信任进行更全面地思考。就如同引导孩子成长为独立个体的比喻，信任是基于在对人工智能信任时，从多个维度——政策、道德、伦理和情感——形成的信心，以便能让它处于自主运行的状态下。这意味着，信任不仅仅关乎于AI被利用的情形，更难的部分在于即使我们在使用AI时，信任究竟如何被影响的现实情况。 *正如我们所预期的。*

因为它的 *是* 受影响。首先——企业需要意识到，随着他们在技术系统中的自主性日益增长，他们需要以不同的方式思考他们对这些系统的信任程度，以及他们可能需要施加哪些约束。Sakana AI，一家AI研究公司，通过测试他们名为“AI科学家”的新系统，完美地展示了这一点。[14] 该系统自主使用大型语言模型进行科学研究，在实验设定的时限内，它遇到了一个问题无法完成，于是调整了自己的代码，为自己争取更多时间。Sakana AI认为这种行为是创新的，但也表明了这样一个事实：具有绕过特定约束能力的AI模型对AI安全性具有重大影响。

并且，除了企业对其使用的AI模型或系统的信任之外，日益增长的自主性也在以许多不同的方式破坏企业与人们之间建立起的信任。

采用不良行为者使用的相同合成内容；许多企业正在利用相同的核心技术产生显著效果。AI生成的营销材料、聊天机器人对话、产品推荐——用例不断增长。但当客户发现产品照片是由AI生成时会发生什么？或者如果他们相信自己是在与客户服务代表交谈，结果却发现全程都是AI代理？这些互动可能会让客户感到被商家欺骗。

或者看看工作中的AI。每天都有更多的工人发现他们在工作中使用AI的价值；到2024年5月，超过40%的用户只是在过去六个月内开始将其纳入其中。[15] 但他们却将此事隐藏从雇主那里：超过

一半使用AI的工人不愿承认这一点，并担心使用AI来完成重要任务会使他们看起来可被取代。这并不是一个关于工人对所使用的AI信任程度的问题，而是证明AI正在动摇人们与雇主之间信任的关系。员工习惯于拥有完善的职业路径、明确的角色定义、技能期望以及对于工作表现如何转化为工作稳定性的共同理解。AI的引入给这一切带来了不确定性。

对企业而言，信任是其与客户、员工、监管机构和股东关系的基础货币。迄今为止，这种信任是在微小时刻中建立起来的——而人工智能正在改变这些时刻。想想每天都在企业间发生的微观互动。一位出色的销售代表为客户节省金钱，或者一位支持代表超越职责范围解决客户问题。从业者或提供商提供优质服务。通过电话确认客户身份。产品准时交付。这些时刻中的每一个都可能，也将被人工智能所颠覆。其中许多将因此变得更加出色——拥有更多的自主权、更少的摩擦和更好的结果。但在信任成为问题之前，你能走多远？你将如何重新激发建立信任的关键人类时刻？

这些问题是领导者需要解决的问题。自主性是下一代商业增长和创新的钥匙。我们希望员工能够更加自主地工作，拥有一个由他们指挥的代理团队。我们希望客户能够自由地与自主企业系统互动，按需购买或享受一定程度的定制化服务。

相关性只有人工智能代理能够在规模上提供。但这种自主性需要信任来促进。顾客对企业的信任程度，或者企业领导对他们的系统的信任程度，或者工人对他们的雇主的信任程度，人们对人工智能的信任程度，或者企业与其关系网中成百上千的其他变体。

这是为什么信任不仅仅是在今年报告中众多趋势之一。它不是企业需要考虑的因素——它是 *the* 考虑因素。随着每家公司都在AI普及的背景下进行自我革新，技术本身不能是唯一的关注点。只有在建立在信任的基础之上，才能获得AI的好处，这需要成为每位领导者首要的任务。

通往更坚实基础之路

信任在人工智能世界中并未消失——但它正变得越来越动态且对企业计划至关重要。随着人工智能的发展，企业领导者将需要同时导航信任的情感和认知维度。情感的维度——人们是否喜欢人工智能、害怕它、认为它与自己的利益一致，或者感觉被它利用——通常在公众中被讨论，但在企业寻求进一步推广技术时，需要真正的政策和治理。而且，只有采取补充措施来解决信任的认知维度：系统是否可靠、专业，能否应对挑战并在既定的轨道内以预期的行为行事，这些努力才能取得成功。这对于任何具有自主性的系统都是一个关键方面，尤其是在人工智能领域，它本质上是一种基于意图学习、成长和行动的体系，而不一定是基于明确指令的。支持这一方面将会

需要专门的领域和决策科学家团队（或AI运营团队），他们将持续测试、评估和构建系统所需的准确性、可预测性、一致性和可解释性，以维持对系统的认知信任。这是一个新的领域，没有一劳永逸的解决方案。每个企业都有其自己的建立信任的时刻、技术、AI策略和关键关系要关注。但总体而言，任何前进的道路都将集中在解决系统与数据、AI本身以及人们之间的信任问题。

首先，企业需要加强其数字系统的网络安全和信任。好消息是，对于系统和数据来说，建立这个新的基础并不意味着从头开始。许多先前的技术和战略投资现在可以带来新的回报。例如，像零信任和实体行为分析这样的网络安全策略将至关重要。你无法控制不良行为者，但你可以控制如何保护系统和人员免受他们的影响——鉴于AI对数据的依赖，保护每个人的数据越来越重要。促进生态系统规模信任的分布式账本技术也是将传统信任网络适应为新基于技术的信任网络的一个很好的例子。你不必信任使用这些技术的实体，因为系统确保他们遵守设定的任何协议。最终，整体的优秀网络安全对于实现AI的信任和安全至关重要。

路线图的第二维度是思考如何建立对人工智能本身的信任。到目前为止，负责任的人工智能领域已成为一个确立的学科，企业在寻求道德地管理其战略时将越来越依赖它。

许多公司已经熟悉了诸如可解释性、数据收集的透明度、去偏见以及其他成熟技术等概念，但随着领导者寻求扩大人工智能的使用，这些努力将成为技术解决方案与对使用该技术感兴趣并投入其中的人类之间的关键桥梁。围绕着人工智能是如何训练的、它在为谁工作以及它是如何做出决策的问题将层出不穷——这些都是无法避免的。但企业可以做到的是准备好一个答案，这也是为什么他们不能坐视不管、等待，而需要将负责任的AI作为他们战略中的关键部分现在就着手进行。

最后，路线图中的第三部分，也是未知的部分，是寻找一条通往以人为驱动信任的新路径。我们知道我们需要达到哪里——新的接触点和建立与维护人际信任的方法，因为人工智能的普及正在颠覆传统的互动方式。但是，每个组织达到那里的方式都将不同，因此，我们应该从自我引导的问题开始：当许多初级工作可以被人工智能完成时，职业道路将是什么样子？利用人工智能来简化工作流程的员工将如何建立职业保障？如果我们的“前线”支持是人工智能代理，我们将如何保持与客户的个人联系？企业应寻求以促进人与人工智能之间共生关系潜力的方式来回答这些问题。无论是教育者和学生、导师和徒弟，还是超级英雄和助手，世界充满了互利的教学和学习关系，这些关系应该激励我们与人工智能的未来。