精通人工智能治理

赋能组织以负责任的AI引领

随着全球人工智能的采用激增，组织必须采用治理结构，以指导符合道德、透明和负责任的人工智能使用。本指南提供了实施有效人工智能治理的路线图，其中包括对NTT DATA自身人工智能治理方法的解释，该方法的结构是为了支持组织在人工智能部署的每个阶段。

内容

本指南概述

赋能组织以负责任的AI引领

随着生成式AI（GenAI）等技术的进步，全球对AI的采用正在迅速增长，因此迫切需要采用治理结构，以指导符合道德、透明和负责任的AI使用，并遵守欧洲联盟人工智能法案等新兴法规。

随着人工智能技术成为竞争定位的核心，组织必须处理治理问题，以减轻与数据隐私问题、算法偏见和监管合规性相关的风险，从而保护其声誉和利益相关者的信任。

本指南提供了一种全面的实施有效人工智能治理的方法。

我们探讨人工智能治理的现状及其意义与重要性。我们分析各种治理框架，考察全球最佳实践和标准，包括ISO、经合组织原则和联合国教科文组织指导方针。

鉴于缺乏一个全球公认的AI治理标准，我们倡导一种灵活、模块化的方法，组织可以根据其独特需求进行调整。

我们解释了NTT DATA自身的人工智能治理方法——一种结构化模型，旨在支持组织在人工智能部署的各个阶段，从战略规划、角色定义到风险管理、运营监督。该方法强调负责任的人工智能开发、合规性及风险管理，并由一个专注于人工智能的部门作为支撑，该部门整合了跨学科的专业知识，以推动符合伦理的人工智能计划。

我们提出了一份实用实施方案路线图，从我们的AI成熟度评估开始，逐步过渡到可定制的治理模块。

这种结构化方法使组织能够负责任地采用人工智能，与行业标准、监管框架和组织价值观保持一致，同时培养合规和创新，以实现持续成功。

1. 理解人工智能治理

人工智能治理之旅

虽然人工智能治理的概念已存在十多年了，但近年来它受到了更多的关注。从我们的角度来看，有两个因素显著促成了这种关注度的增加。

第一个因素是 **通用人工智能的广泛涌现** 2023年，及其对人工智能各个方面的影响。这一发展引起了科技公司和大众的关注——到2024年，组织开始有效使用这项新技术，创造了巨大的商业价值。这种兴趣的增长也突显了更广泛范围的人工智能能力。

ntt data 的研究表明，97% 的首席执行官预计仅通用人工智能（genai）将对提高生产力水平产生重大影响——这一观点显示了人们日益增长的信心，相信人工智能有潜力改变业务运营。此外，44% 的高管强烈同意通用人工智能的承诺（以及投资回报率）超过了潜在的安全和法律风险 [1]。

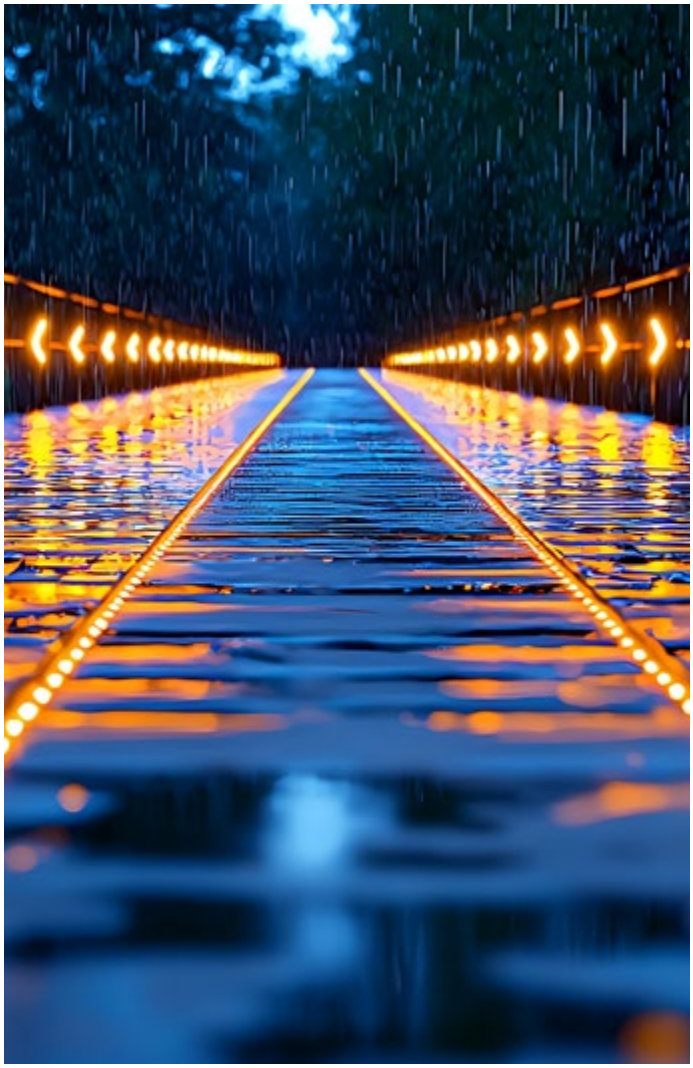
人工智能的采用和发展增长需要实施治理框架、政策及流程，以指导对其使用的有效监督和管理。

第二个因素是监管。在全球范围内，人们已经做出了重大努力来建立监管和保护在使用人工智能时的基本权利的法律框架。在这些举措中，欧盟的人工智能法案（AI法案）作为人工智能监管的里程碑而脱颖而出，它高度重视高风险系统的数据治理政策的制定，并要求组织实施适当的治理框架。

随着数据治理政策的制定，现在是建立人工智能治理政策的理想时机。

共同推进负责任的AI

随着组织面临日益增长的合规成本，特别是在欧盟等地区，全行业合作对于平衡监管需求与创新至关重要。NTT DATA集团呼吁政策制定者、行业领袖和远见卓识者将负责任的AI原则融入治理框架。



通过这样做，企业可以驾驭复杂的法规，降低风险并保持竞争力，同时为所有利益相关者创造长期价值。

尽管人工智能治理目前并非法律要求，可能被视为“锦上添花”，但这种看法将发生转变，人工智能治理将成为一项基本必要。

一个围绕人工智能治理统一认证方法的全球对话正在形成，反映出人工智能政策的关键时刻。全球社会前所未有地见证了政策制定者和国际组织之间这样一种统一而集中的努力，所有人都专注于定义和建立强大的人工智能治理框架。

[1] NTT DATA. 2024. 全球 GenAI 报告：2025 年全球各组织如何掌控其 GenAI 命运。

人工智能治理的定义及其重要性

什么是人工智能治理？

人工智能治理是指一套旨在指导组织内部人工智能系统开发、部署和监督的战略、政策和流程的综合性框架。

它包含两个关键维度：

1. 战略重点：人工智能治理使组织能够利用人工智能的潜力，探索新机遇，增强市场定位并提高竞争力。它支持人工智能在整个运营中的集成和扩展，以与并实现更广泛的战略目标。

2. 负责任的AI开发：AI治理指导并监督AI系统的构建和管理，遵循伦理原则、法律标准和社會價值。它處理透明度、可負責任性、公平性以及風險管理流程的實施等關鍵方面。最終，AI治理促進規範合規以及AI技術的長期可持續性和負責任的使用。

为什么人工智能治理现在比以往任何时候都更加重要

大多数组织不会开发自己的AI系统；相反，它们会依赖从外部供应商那里采购。这一现实强调了在采购、定制和实施过程中建立稳健的治理和控制机制的关键性。

组织必须确保其整合到运营中的人工智能解决方案符合道德标准和监管要求，并与组织价值观保持一致。治理框架在挑选和使用第三方人工智能技术中维护透明度、问责制和公平性方面发挥着关键作用。

未能实施有效的AI治理会让组织面临重大的法律和声誉损害。

人工智能应用中的主要挑战

随着组织越来越多地转向人工智能以提升其运营和推动创新，他们在采用和治理人工智能系统方面遇到了几个挑战。

不受管理的AI使用

采用人工智能能力而不建立任何形式的治理框架通常会导致碎片化工作，存在数据孤岛、流程不完整、监控不足、角色未定义、重复劳动和资源利用效率低下。

随着组织扩展其人工智能能力，他们很快意识到他们当前的方法是不可持续的。

我们看到具有自主性的AI的快速采用速度，突显了不受管束的努力如何能迅速从有益转变为具有重大，甚至可能是负面的后果。

缺乏明确的治理结构，与缺乏监督相关联的风险，包括表现不一致、不符合法规以及道德问题，会更加突出。

在这一阶段，组织通常认识到为人工智能治理奠定坚实基础的需要。

缺乏成熟的治理框架

许多组织在数据治理、道德风险管理、人工智能技术部署中的透明度和问责制等方面面临问题。

一个强有力的治理框架是解决这些复杂性并管理公平性、问责制和监管合规性相关风险的唯一途径。

为了实现全面治理，组织需要提升其方法，超越对单个算法的管理，采用一个同时治理人工智能系统和其实际应用的框架。

进入壁垒

对于在人工智能方面经验有限且通常处于探索人工智能选项早期阶段的组织而言，又会出现另一个挑战： **从哪里开始？**

对于这些企业而言，制定人工智能治理框架对于建立人工智能应用的清晰路线图至关重要。它提供了必要的结构来指导实验，从一开始就嵌入伦理考量，并为人工智能能力的未来发展奠定坚实的基础。

日益增长的ai应用带来了巨大的机遇，但如果缺乏合适的治理结构，组织将面临削弱其寻求实现之根本利益的风险。它们必须在ai旅程的早期就解决这些治理挑战，以确保负责任且有效的实施。

治理在负责任的人工智能发展中的作用

“负责任的AI”是什么意思？

根据国际标准化组织（ISO），“负责任的AI是从伦理和法律角度出发开发及部署AI的一种方法。其目标是安全、可靠且合乎道德地应用AI。负责任地使用AI应提高透明度，同时有助于减少诸如AI偏见等问题。”[2]

在讨论责任和信任时，“负责任的AI”、“符合伦理的AI”和“值得信赖的AI”这些术语经常被混用，但有必要明确这些术语的含义并理解每一个术语的含义，以便恰当使用：

负责任的AI 着重确保人工智能系统以防止对个人和更广泛环境造成伤害的方式开发和部署。责任意味着认识到人工智能对社会的影响，并采取积极措施来保护人权和自由以及社会价值观。

人工智能伦理 指导人工智能决策的原则和道德价值观。由于道德标准可能因组织或地区而异，因此伦理人工智能是一个更具主观性的概念。虽然伦理考量至关重要，但它们高度依赖于文化或制度环境，这可能导致对“什么是对”或“什么是错”的不同解释。鼓励组织制定反映其利益相关者价值观的人工智能伦理政策，并使这些政策透明化。

可信赖人工智能 致力于通过严格的标准和测试在技术本身建立信任，确保系统能按预期运行，并能被信任做出公平和无偏见的决策。

虽然这些术语有重叠，负责任的智能人工学院指出，“负责任的智能”是最全面的，因为它涵盖了伦理原则和技术可信度，但更广泛的承诺在于通过确保智能系统不仅运行正确，而且与保护人民和地球的价值相一致，来保障人权和社会福祉。[3]

[2] ISO. 2023. 构建负责任的AI：如何管理AI伦理辩论。
[3] 人工智能责任研究所。2023。 人工智能 vs 可负责任的人工智能：为什么这很重要？

为什么负责任的AI很重要？

人工智能已深深植根于现代生活的几乎所有方面，并且其增长势头未见减缓。高德纳®指出：“到2028年，人工智能服务市场将达到6090亿美元，以美元计五年复合年增长率为21.4%。增长将受到新的人工智能生成能力以及使用预测分析和决策的传统人工智能技术的双重推动。”[4] GenAI服务市场正经历前所未有的增长，预计将从2023年的47亿美元增长到2028年的2218亿美元。这一增长得益于2023-2028年间GenAI服务的复合年均增长率（CAGR）为115.9%，而相比之下，“经典”AI服务的CAGR为11.3%。

鉴于这种快速扩张，认识到与人工智能相关的潜在益处和风险至关重要。人工智能的不当或不受控制的使用会导致有害的后果。这些后果包括加剧甚至放大偏见；大规模创建虚假或误导性内容，这些内容可用于传播错误信息和操纵公众舆论；侵犯隐私等等。

大多数组织都意识到了算法偏见的问题，但并非所有组织都建立了完善的系统来跟踪偏见和处理隐私风险。NTT DATA的研究显示，只有86%的受访者同意算法偏见仍然普遍存在（96%的首席数据官同意），但只有43%强烈同意他们已经建立了跟踪偏见和隐私风险的系统。[5]

人工智能日益增长的影响力与负面结果的潜在可能性，要求组织实施负责任的AI，而不是简单地采用AI。

最终，这意味着将责任、可靠性和伦理融入设计过程中。例如，一个AI系统可以设计为分析消费者数据以进行广告个性化。而该相同AI系统的负责任设计则应包括通知用户其数据如何及何时被使用、解释AI系统实施背后的原理，并在系统出现故障时提供补救机制等安全措施。

[4] 高德纳公司，预测分析：人工智能服务，全球，C.格雷厄姆等，2024年8月27日。GARTNER是高德纳公司及其附属公司在美国和国际上注册的商标和服务标志，并经其许可在此使用。版权所有。
[5] NTT DATA. 2024 . 全球 GenAI 报告：2025 年全球各组织如何掌控其 GenAI 命运。

通过人工智能治理实现负责任的人工智能

随着人工智能深度嵌入商业和政府运作，组织面临着双重挑战：在促进创新的同时保护公平、隐私和合规性。此时，人工智能治理发挥着关键作用，它弥合了伦理原则和实际应用之间的差距。

平衡创新与问责

负责人工智能发展的核心在于解决偏见、透明度、知识产权和社会影响等挑战，同时促进创新。NTT DATA集团于2019年开创了该组织的AI指南，并自那以后与公司的AI服务业务协同推广该指南。2024年，NTT，Inc.也宣布了建立

NTT AI 章节。

NTT DATA集团通过采用日本商业哲学“三方良好”：对买方好，对供应商好，对社会好来体现这种平衡。“三方良好”这一原则是我们致力于以合乎道德和可持续的方式提供价值的人工智能系统的基石，遵循在所概述的六项基本原则中。 **NTT AI宪章:**

1. 可持续发展

人工智能应对社会、环境和人类做出积极贡献。

2. 人类自主性

尊重个人权利和赋权人类决策是至高无上的。

3. 公平性和开放性

最小化偏差和透明度是建立信任的核心。

4. 安全性

人工智能系统在其整个生命周期中都必须受到威胁保护。

5. 隐私

严格的数据管理协议建立信任并保护个人信息。

6. 沟通与共创

吸引利益相关者可确保创新与社会价值观和人权相一致。

将原则转化为行动：

为使这些原则运作化，NTT DATA集团在四个关键领域实施举措：

1. 以责任为导向的创新

从最初就将道德、安全和可持续性融入人工智能解决方案中

2. 大规模技能提升

对数十万名员工进行人工智能技术和其伦理使用的培训，以确保负责任的实践

3. 多层次治理

建立系统性风险管理框架以应对合规、安全与问责

4. 利益相关者合作

在全球范围内合作，创建一致的AI标准和框架

人工智能治理指导人工智能系统的发展与部署，使其符合社会价值观，通过值得信赖、负责任的人工智能系统降低风险并实现可衡量的商业价值。

负责任的AI作为ROI的驱动力

治理并不仅仅是缓解风险；它也是关于释放价值。

负责任的AI实践可以通过提升客户参与度、开辟新的收入来源和支持可持续增长来驱动ROI。治理框架规定了如何评估AI计划，不仅是为了安全性和合规性，也是为了它们潜在的商业影响。这种双重关注有助于组织提升其在AI价值链中的位置，平衡伦理责任与可衡量的成果。

通过采纳以负责任人工智能为重点的治理结构，组织不仅能够减轻法律和声誉风险，还能在快速发展的AI领域培养信任、问责制和长期成功。

用负责任的AI治理将伦理付诸实践

组织可以通过转变人工智能的开发和实施方式来利用人工智能治理，将人工智能伦理原则付诸实践。

人工智能伦理可实践的五个领域

1. **责任制**
   定义角色和职责，以及人类监督机制，以追究人们对人工智能结果的责任。

2. **公平**
   确保人工智能通过减轻潜在的有害偏见，从设计到部署支持整体健康和包容。

3. **隐私**
   通过工具和流程加强数据治理，防止在人工智能生命周期的数据收集、使用和共享过程中出现潜在的隐私泄露，并保护数据完整性。

4. **透明度**
   确保人工智能利益相关者沟通、人工智能结果可解释性和人工智能系统可审计性。

5. **诚信**
   构建可靠且安全的AI，通过预防潜在滥用和增强保证通过持续评估获得准确结果的验证实践。

2. 有效的AI治理框架和工具

人们已为建立人工智能的指南、政策和规范做出了大量努力。然而，人工智能治理方面没有标准化的框架，也没有任何单一解决方案获得了全球权威机构的认证。这种缺乏标准化提供了灵活性，但也带来了不确定性。组织需要在多种治理选项中寻求方向，却缺乏一条清晰、普遍接受的道路。

我们全球 GenAI 研究的一个有趣的见解突出了这种不确定性的影响：81% 的受访者强烈同意政府 AI 规定不明确会扼杀创新并阻碍对 AI 的投资。 [6] 这“监管混乱的成本”突显了明确方向以释放AI的全部潜力并推动该领域进步的紧迫需要。

在不利方面，缺乏一个广泛认可的标准会削弱信心，使组织不确定于合规性和问责制。这种模棱两可通常导致不一致的实施，组织对最佳实践的解读可能与全球的道德和监管标准不一致。

正面的来看，这种标准化缺失提供了灵活性，使组织能够根据其独特需求和运营环境定制其治理框架。它还允许适应性，因为快速发展的AI技术需要能够与创新发展相匹配的治理策略。此外，组织受益于多样化方法，并享有采用市场解决方案、寻求咨询专长或开发与战略目标一致的定制框架的自由。

缺乏人工智能治理的标准框架给组织带来了挑战和机遇。

> “虽然缺乏认证的AI治理框架带来了挑战，但也为定制化和适应性提供了机遇。组织可以探索各种选项——从咨询服务到内部开发——并利用国际指南来建立负责任和可持续的AI使用治理结构。“

[6] NTT DATA. 2024. 全球 GenAI 报告：2025 年全球各组织如何掌控其 GenAI 命运。

市场上的产品

尽管市场上已有人工智能治理工具，但许多仍处于早期阶段。所提供的解决方案是新的，可能尚未提供有效治理所需的稳健、全面的答案。

通用人工智能因其动态且往往不可预测的输出而具有多层复杂性。针对通用人工智能量身定制的治理工具需要应对诸如管理知识产权问题、确保内容真实性和减轻潜在滥用等独特挑战。此外，仅靠工具无法解决定义流程和责任方面的挑战。

然而，随着市场的成熟，这些人工智能治理工具预计将变得更加可靠和广泛采用。

这将使组织能够利用自动化、可扩展的解决方案来监控和管理其人工智能系统。正如我们之前在数据治理方面所见，我们预计组织将更早地拥有强大的AI治理工具来支持其AI监管。

内部开发

此路径提供最大程度的定制和控制，但需要大量资源以及对人工智能风险和法规的深刻理解。

如果组织在数据治理成熟度方面达到较高水平，并且拥有包括法律、伦理和运营专家以及技术资源的专门AI团队，那么在内部开发专有AI治理框架可能是一个可行的选项。对于这些企业而言，流程要容易得多，因为既定的数据实践为实施有效的AI治理奠定了坚实的基础。

专家合作伙伴

与专家合作伙伴合作通常是大多数组织的最佳选择。

这些合作伙伴带来了丰富的知识、网络和经验。他们可以设计定制的治理方案，这些方案考虑了特定的风险，如偏差放大、意外滥用和版权侵犯，同时符合合规要求和业务目标。

专家合作伙伴也能帮助组织及时了解不断发展的法规和最佳实践，并应对生成式AI、自主型AI以及其他形式AI的复杂性。

探索可用的AI治理框架

人工智能治理并非一个近年来的关注话题。讨论可以追溯到几十年前，当时人们意识到，如果没有强有力的保护措施，人工智能的发展可能会不成比例地扰乱人们的生活，并加剧社会不平等。

作为回应，组织开始制定自己的人工智能应用原则。随后，出现了几个旨在促进服务于社会利益的人工智能发展的举措，例如：

• OpenAI研究院（2015）•人工智能伙伴关系（PAI）（2016）•人工智能伦理与治理倡议（2017）•世界经济论坛第四次工业革命中心（2017）

2019年，重要全球和区域机构发布人工智能原则，揭示了实施中存在的差距，并催化了众多将这些原则转化为实用、可操作的治理机制的努力。

到2020年，政府、倡导团体、国际组织和私营公司已经发布了超过100项伦理指南。[7]

[7] Fjeld，J. 2020. 哈佛大学法学院，伯克曼克莱恩互联网与社会中心。 原则性人工智能：在基于伦理和权利的方法中为人工智能原则建立共识地图

关键框架概述

若干由联合国、经合组织及欧盟等国际组织开发的现有框架和指南，为人工智能治理的最佳实践提供了宝贵的见解，并有助于组织构建自身的框架。

人权焦点

联合国商业与人权指导原则 (UNGPs) [8]

尽管并非专门针对人工智能，但联合国人权保护原则为应对人权风险提供了基础，其中包括人工智能应用所引发的风险。这些原则敦促企业通过进行影响评估、履行尽职调查和建立问责机制来尊重人权。它们为评估人工智能对个人和社区的影响提供了一种现成的手段。

负责任的人工智能

经合组织人工智能原则 [9]

2019年最初被42个国家采用并在2024年5月更新，经合组织人工智能原则强调可信赖人工智能的负责任治理，涵盖人权、透明度、问责制和稳健性。

以可信度为焦点

欧盟委员会人工智能高级专家组的信任人工智能道德指南 [10]

该指南确立了七个关键要求，人工智能系统应满足这些要求才能被认为是值得信赖的。一个具体的评估清单旨在帮助验证每个关键要求的落实情况。和社区。

全球道德标准

联合国教科文组织人工智能伦理建议书 [11]

这份于2021年通过的建议书是全球首个关于人工智能伦理的标准化工具。它强调透明度、问责制和公平性等原则，倡导人权保护，旨在指导各国制定包容性人工智能政策。

[8] 联合国。2011。关于商业和人权指导原则。
[9] 经合组织。2024。经合组织人工智能原则概述。
[10] 欧盟委员会人工智能高级别专家组。2019年。可信赖人工智能的伦理指南。
[11] 联合国教科文组织。2024年。关于人工智能伦理的建议。

实用实现

新加坡的AI治理框架模型 [12]

2019年，新加坡政府个人信息保护委员会（PDPC）发布了其模型人工智能治理框架的第二版。该框架对行业、技术和算法保持中立，将伦理原则转化为人工智能部署的实际应用，从而使组织能够有效实施这些原则。2020年，他们发布了一份补充评估指南以协助实施。此外，还开发了AI Verify工具，以提供一种结构化的方法来评估人工智能系统的可信度，重点关注公平性、透明度和问责制等关键标准。

国际标准

ISO/IEC 42001: 2023 信息技术 — 人工智能 — 管理体系 [13]

ISO/IEC 42001 是一个国际标准，它规定了建立、实施、保持和持续改进人工智能管理体系（AIMS）的要求。它适用于提供或使用基于人工智能的产品或服务的组织，并为人工智能系统的负责任开发和使用提供指导。

监管监督

欧盟人工智能法案 [14]

虽然人工智能法案被归类为条例而非治理框架，但它引入了基于风险的监管方法来监管人工智能应用，为高风险系统制定了具体要求，并旨在为欧盟及其他地区设定标准。在为在欧盟运营的组织定义治理结构时，必须考虑这一监管框架。

[12] 新加坡个人数据保护委员会（PDPC）。2020。新加坡人工智能治理方法。
[13] ISO. 2023. ISO/IEC 42001:2023: 信息技术 — 人工智能 — 管理体系。
[14] 欧盟。2024。法规（欧盟）2024/1689。

一个强健的AI治理框架的关键组成部分

人工智能治理方法可能因行业和公司规模而异。然而，有一些关键组成部分和原则是普遍适用的。一个健全的人工智能治理框架必须：

1. 包含基本原理

2. 包含一个风险管理流程

3.采用人在回路（HITL）方法

通过集成这些组件，组织可以开始应对人工智能的复杂性，确保其系统负责任、合乎道德地开发与实施，并与社会价值观保持一致。

1. 包含基本原理

哈根多夫（2020）的分析揭示，诸如问责制、隐私和公平性等方面的规定大约出现在80%的主要人工智能指南中[15] 这份报告题为“基于原则的人工智能：伦理和权利方法中关于人工智能原则共识的图谱”，确定了在任何治理框架中都应该解决的共同关切领域。[16] 这些包括：

问责制

必须为人工智能系统的发展、部署和成果建立明确的责任线。这确保了实体能够对其行为和决策负责，培养了伦理监护意识。

隐私

保护个人数据是一项基础性要求，需要遵守数据保护法律和道德标准。有效的AI治理框架必须优先考虑隐私保护措施，以维护公众信任和合规性。

公平与非歧视

人工智能系统应该设计为促进公平，减轻偏见，支持所有个人和群体的平等对待。

安全与安全

人工智能系统必须经过设计以确保可靠和安全运行，以最大程度地减少对用户和社会的潜在危害。

透明度和可解释性

深入了解人工智能系统如何决策和运行对于建立信任至关重要。

促进人类价值观

人工智能的发展必须与基本人权和价值观保持一致，以造福人类福祉和社会。

职业道德

在人工智能背景下制定劳动力使用政策，使人工智能专业人士具备履行工作中最高道德标准所必需的技术技能。此类政策培养了责任文化，因为人工智能专家被期望了解新兴法规，为开发透明和负责任的人工智能系统做出贡献，并始终考虑其工作的更广泛社会影响。

2. 风险管理

一个严格的风险管理流程应该是任何人工智能治理框架的一部分，对于识别、评估和缓解与人工智能技术相关的风险至关重要。

nist人工智能风险管理框架，强调公平性、问责制和透明性等原则，是构建此组件的有用资源。该框架提供了一种自愿且结构化的治理方法，使组织能够有效管理人工智能相关风险。[17]

3. 人在回路 (HITL)

人工智能治理的另一个获得普遍共识的组成部分是HITL方法。它强调了在人工智能模型开发和实施的所有阶段都需要人类监督的必要性。

通过引入人工监督，组织可以将伦理问题整合到决策过程中，并主动应对潜在风险。

15 哈根多夫，T. 2020. 人工智能的伦理：对指南的评价。
16 Fjeld，J；Achten，N；Hilligoss，H；Nagy，A和Srikumar，M。2020。 在基于伦理和权利方法的 AI 原则中映射共识。
17 NIST. 2023. 人工智能风险管理框架 .

3. ntt data 的 ai 治理方法

我们的AI治理方法为整个AI生命周期提供全面的监管。它提供了一个结构化模型，组织可以整体或模块地应用，适应不同的成熟度水平，并与行业最佳实践保持一致。这种方法促进了敏捷的持续创新和驱动业务的AI计划。

图1：NTT DATA的AI治理方法

策略：使业务价值与公司战略保持一致

人工智能战略是人工智能治理的基石。它通过优先考虑有影响力的用例、制定清晰的路线图以及培养所有利益相关者积极参与以实现项目成功，使组织能够将人工智能价值与公司目标相结合。

建立人工智能战略基础

组织必须从理解驱动业务的任务和战略目标开始，旨在理解人工智能采用的“现状”场景。

我们帮助客户回答诸如以下问题：

• 什么 **目标** 我们为什么有AI？• 什么 **优先级** 我们有哪些利用人工智能的手段？• 什么与相关 **风险** 人工智能实施方面的？

人工智能发展必须与基本人权和价值观保持一致，以促进人类福祉和社会发展。了解人工智能如何与整体商业战略保持一致，对于培养共同愿景和目标至关重要。

我们利用行业特定的见解和专业知识来发掘机遇，并鼓励人工智能和业务团队进行合作。

统一投资组合管理

人工智能计划必须采用协作方法进行设计，并定期获得业务专家的输入，包括监控其价值创造的能力。这种方法促进了人工智能计划中的一致应用，并有助于构建支持战略目标并实现更大整体价值的解决方案组合。

人工智能用例路线图

识别和评估与战略重点相一致的潜在AI用例，使组织能够根据可行性和影响来优先考虑举措。

在头脑风暴探索不同的用例并定义和优先排序计划后，我们制定了一个定制路线图以开始开发人工智能。该路线图充当战略指南针，提供清晰的方向和关键里程碑，以支持我们的客户实现其目标。

战略目标

人工智能战略基础

为人工智能定义总体愿景、目标和战略方向。

统一投资组合管理

监督和协调人工智能计划，促进业务和技术团队之间的合作，以实现统一的人工智能组合。

人工智能用例路线图

识别和评估与公司目标一致的AI计划，并根据潜在影响和可行性对其进行优先排序。

组织：为AI的成功长期实施量身定制组织模式

创建一个由人工智能驱动的组织包括将人工智能应用于所有业务领域，并协调参与人工智能开发和产业化的团队和角色。这有助于实现良好的治理，并促进人工智能意识的文化。

协作模型

我们帮助组织定义团队结构，以推动人工智能计划，支持业务目标。我们与他们合作，建立高效的多学科团队，包括数据工程师、数据领导者和数据架构师，他们与其他业务领域（如营销、人力资源、物流、采购和销售）有效合作。

根据组织的成熟度、规模和需求，该模型可以是集中的、去中心化的或联邦的，在保持精简的运营方式的同时确保灵活性。

角色和职责

第一步是定义每个团队成员的角色及其所需的胜任力、责任和技能。然后我们建立一种通用的项目管理方法，通过定义项目生命周期内的主要活动来协调不同的角色。

最后，实施了一种报告结构。这包括高层赞助、跨业务线合作的工作团队和委员会，以及潜在风险的升级机制。

人工智能人才与文化

人工智能治理也应包括培养人工智能素养的培训计划和发展项目，提升内部人才和分析能力。

人工智能素养的实践和努力不应仅限于界定人工智能人才和素养需求，还应在整个组织中培养共享的人工智能创新文化。

这些实践支持AI知识的传播，使组织更好地准备利用即将到来的市场趋势和技术进步。它们还有助于通过建立创新和实验的声誉来吸引人才。

战略目标

协作模型

实现一个促进团队协作并提升人工智能人才能力的ai组织模型。

角色和职责

在人工智能生命周期中，根据变更管理策略，定义操作、角色和职责。

人工智能人才与文化

民主化人工智能能力，并为员工做好准备以采用人工智能。

操作：统筹管理支持人工智能生命周期流程、工具和基础设施

通过在整个人工智能计划生命周期中管理运营的标准方法，组织可以快速响应业务机会，并减少人工智能应用程序的上市时间。

人工智能生命周期

为保障人工智能的可扩展实施，我们首先识别并定义人工智能的每个生命周期阶段。我们概述了每个阶段所需的特定行动和精确工具集，并明确了有效的执行角色和职责。我们还提出了潜在的与缓解策略相配套的风险。

这项工作为开发和产业化人工智能解决方案提供了骨干，以加速和扩展所生成的价值。

人工智能治理工具

我们建议组织探索可适应和可扩展的治理工具，以支持持续的模式评估，以便人工智能系统保持合规、透明，并与战略目标和监管标准一致。

组织可以选择开发内部的解决方案以获得完全控制权，尽管成本更高、开发时间更长，也可以利用现有的内部工具，但这可能缺乏某些功能。或者，市场解决方案提供了尖端技术并定期更新，但可能会增加成本并引发数据隐私问题。

战略目标

流程编排

在生产环境中快速翻译和执行人工智能驱动的计划。

人工智能生命周期管理

封装流程、程序和技术工具的预期结构，并明确角色和职责。

人工智能治理工具

提升基础设施并确保技术赋能，以扩展人工智能计划，实现持续敏捷创新。

符合伦理的 AI 部署模块

通过协作，我们共同创建繁荣的人工智能生态系统，这些生态系统不仅满足监管要求，还倡导伦理价值观。在一个快速发展的技术环境中，深入理解伦理人工智能与遵守法律框架一样至关重要。

我们做什么

合规性评估

我们通过评估组织的监管合规性和技术采用情况来评估其人工智能准备情况。我们使用准备情况分数来衡量运营和技术领域的成熟度，突出可行性、风险管理能力和可行的下一步措施，以推进与道德和法律标准的协调。

风险管理

我们通过进行综合评估和设计缓解策略来实施稳健的风险管理框架。这些框架通过旨在在整个人工智能生命周期中最大限度地减少负面影响的政策和协议来应对技术和伦理风险。

人工智能伦理指南

我们的详细指南帮助组织和利益相关者理解人工智能工具、术语和方法，在系统开发和部署中建立包容性和可信度。

AI模型注册中心

我们帮助在整个组织内集中化和记录所有与已部署或正在开发的ai系统相关的信息，创建一个集中的ai模型注册中心。这促进了利益相关者之间的协作，并支持了明智的决策。

政策和程序

为确保系统在道德和法律边界内运行，我们协助建立政策和流程来有效管理人工智能操作。这包括处理安全漏洞、道德问题或运营问题的突发事件管理流程。

AI办公室

我们建议设立一个专门的AI办公室，以监督和协调治理工作。该办公室应整合多学科专业知识，以确保问责制，管理责任，并在减轻声誉、经济和社会风险的同时，遵守道德AI实践。

我们方法的价值

培育负责任的转型

我们的方法超越了合规性，通过嵌入道德素养和培养负责任的AI文化。这确保了系统不仅合规，而且对个人和社会有益，促进信任和包容性。

提升透明度

一个集中的AI注册中心提供了对AI生态系统的全面概述，支持合规性、问责制和战略决策。这种集中式方法促进了更好的治理，并简化了监管报告。

最小化风险

主动的风险评估有助于识别潜在的技术和伦理问题，并在早期减轻风险，从而减少对不可预见挑战的暴露，并保护组织的完整性。

加强治理

政策和程序使组织能够在道德和法律边界内运营人工智能系统，帮助他们预防和快速响应问题，同时实现持续合规和运营稳定。

建立信任

综合指南赋予利益相关方一个共同的知识基础。这增强了合作，建立了公众对人工智能系统的信心，并推动了人工智能在不同环境中的包容性和道德使用。

驱动业务价值

负责任的AI实践不仅确保合规，还通过增强客户信任、支持可持续增长和创造创新机会来释放商业价值，同时最小化声誉风险。

我们与行业标准和最佳实践保持一致

我们的方法是建立在对普适原则和最佳实践的运用之上，借鉴行业领先指南，并适应不同的组织环境。

认识到治理方法因行业、公司规模和具体需求而异，我们的方法论设计得灵活，允许完全或模块化集成。客户可以整体实施我们的方法，或根据需要选择特定的支柱，以便他们只采用最能支持其即时优先事项的组件。这种灵活的结构使组织能够负责任地开发、部署和管理人工智能系统，并随着时间的推移扩展其治理工作。

集中授权

遵循广泛接受的治理标准，我们建议以专门的AI办公室的形式建立集中管理权，该办公室拥有AI监督权并同步ISO/IEC标准等框架。该办公室通过管理安全、合规和风险职责，促进伦理、合规和创新AI。它作为所有AI相关治理活动的枢纽，监督与组织目标和行业法规的一致性。

最小化风险

我们整合了强大的风险管理协议，并开展符合行业最佳实践（例如NIST所概述的实践）的定期合规评估。通过主动的风险评估、事件管理政策和合规性评估，我们帮助组织在满足关键监管要求（包括GDPR和欧盟人工智能法案）的同时，将潜在风险降到最低。我们的方法帮助组织在整个人工智能生命周期中保持安全和合规的环境。

优化AI生命周期

我们管理人工智能生命周期的流程基于行业领先的实践，如aiops，确保标准化的工作流程，以实现快速高效的新一代人工智能实施。这些工作流程不仅支持可靠的运营效率，还减少了人工智能应用的市场推出时间，帮助组织把握人工智能机遇。

负责任的AI开发

我们对负责任人工智能的关注与欧洲委员会的《可信赖人工智能伦理指南》和经合组织的《人工智能原则》等框架相一致。通过伦理素养、集中的人工智能注册机构和全面指南，我们帮助组织以透明和问责的方式开发人工智能解决方案，并以有益于社会的方式行事。这种方法促进了符合法律标准，降低了声誉风险，并建立了公众信任。

人工智能素养

我们认识到培养人工智能素养文化和持续创新的重要性。通过支持我们的客户保持人工智能发展的前沿，我们帮助他们保持竞争力，吸引顶尖人才并创造长期价值。我们对教育、人才发展和健全的员工使用政策的关注使组织能够适应快速变化的人工智能环境，并在其团队中负责任和合乎道德地管理人工智能技术。

我们的方法使组织能够负责任地发挥AI的全部潜力，使每个计划与公认标准保持一致，同时促进增长和适应性。

4. 从蓝图到现实

有效的AI治理实施

有了原则、框架和标准，问题仍然是：我们如何将这一理论转化为可执行的AI治理？

我们的方法确保人工智能治理不仅是理论，而是深度融入组织的运营。我们从评估当前的人工智能成熟度水平开始，然后创建与业务目标相一致的定制化战略和治理结构，支持可持续增长和负责任的人工智能部署。

构建和执行人工智能治理项目：我们的流程

全面且有效的AI治理的3个步骤

01

人工智能成熟度评估

我们开始于 **评估该组织的当前人工智能成熟度** 要了解其优势、差距和治理采用的准备情况。这项评估为改进提供了明确的基准，并能够与长期目标保持一致。

02

定制化AI路线图

根据评估，我们设计了一个 **定制化AI路线图** ．本路线图概述了构建一个支持当前和未来人工智能野心的强大人工智能框架所需的战略举措、里程碑和资源。

03

治理实现

本阶段侧重于 **将人工智能治理融入日常生活实践** 通过建立结构、流程和政策，以促进持续监督、持续改进以及与道德、法律和组织指南保持一致。

1. 人工智能成熟度

任何成功的AI治理项目的基石是对组织当前AI成熟度的全面理解。

我们开发了一项人工智能成熟度评估，通过一系列结构化、深入的与利益相关者的研讨会来识别和定义人工智能治理的关键领域。评估过程包括：

- **三次引导式课程，每次 2 至 3 小时**

  这些会议由跨职能团队举行，包括来自领导层、IT、合规和数据科学部门的相关方。协作方法使我们能够收集关于组织当前能力、目标和与人工智能相关的具体挑战的多样化见解。

- **在4周内进行全面分析和报告**

  本报告涵盖了关键AI治理领域的成熟度分数。每个领域都根据最佳实践和行业标准进行评估，清晰地展现了组织当前的AI能力和改进领域。最终报告是一份综合参考文件，确立了团队作为组织内部AI权威声音的地位。它还通过概述清晰的下一步行动和优先项目，支持预算分配，以提升组织的分析成熟度和AI能力。

我们的AI成熟度评估是通用的

人工智能采用的基础工具

对于没有专门AI结构的组织，该评估是构建AI能力以及确定与其战略目标一致的治理举措优先级的路线图。

提升人工智能成熟度

对于已建立人工智能实验室或中心的组织而言，该评估有助于识别关键治理举措，以帮助提升人工智能成熟度。

2. 定制化人工智能路线图

我们的方法始于定义、实施和管理治理实践，并跨组织明确角色、职责和流程。对于每个核心治理支柱——战略、组织、运营和负责任的AI——我们提供以下内容：

- **核心治理模块**

  每个模块都有清晰的定义，并辅以模板、清单和最佳实践等必要工具，以支持有效的实施。

- **可操作的推荐**

  针对实现短期、中期和长期目标而量身定制的行动，将促进成功推出和人工智能治理实践的持续改进。

3. 治理实施

一旦评估完成并且路线图已交付，我们就进入实施阶段。在此阶段，我们的团队与组织的利益相关者紧密合作，将治理框架嵌入到现有的工作流程和系统中，以符合组织的运营和战略目标。

如何开始

利用我们的服务和参与模式取得成功。

- **一个定制化治理模式**
  我们采用迭代和参与式的方法来开发适合您组织独特需求的模型。我们的方法包括协作研讨会和利益相关者反馈会议。

- **灵活性与可扩展性**
  由于我们的治理方法具有模块化结构，因此它可以随着人工智能的发展和您优先级的改变而进行适应。

- **增强集成**
  我们与您合作，促进跨学科团队之间的协作，并在各职能领域有效嵌入治理实践，促进人工智能治理在您组织的顺利整合。

5. 结论：人工智能治理是一项战略要务

随着组织加速采用人工智能——特别是生成式人工智能和自主人工智能的变革性能力——建立完善的AI治理框架已成为一项战略需求。这不再仅仅是一项主动措施，而是负责任和合乎道德地使用人工智能的基本要求。

生成式人工智能，例如，为创造力、效率和问题解决带来了前所未有的机遇，但也因此在知识产权、数据完整性和偏见缓解等领域带来了独特挑战。

自主AI在任务分配、管理和决策方面提供了新的效率，但不受控制的代理可能会引发未知的安全和流程问题。

通过整合能够平衡战略优先级、监管合规性和道德考量的治理结构，组织可以在有效管理相关风险的同时，发挥AI的全部潜力。

ntt data的ai治理方法提供了一个可扩展和模块化的解决方案，该方案针对生成式ai、自主式ai以及更广泛的ai技术的复杂性进行了定制。

这种方法使组织能够在人工智能成熟度演进的同时，应对优先事项并扩展其治理工作。它支持人工智能计划与战略目标保持一致，并在将伦理原则融入人工智能生命周期的每个阶段的同时，帮助建立明确的角色和责任。

通过在整个组织中嵌入负责任的AI实践，组织可以培养一种以信任、问责制和道德决策为根基的创新文化。

采取全面的人工智能治理方法使您能够在一个由人工智能驱动的世界中负责任地领导。优先考虑透明、公平和可靠的人工智能实践不仅确保了符合当前法规，而且为在一个日益受到人工智能技术塑造的世界中实现可持续增长、竞争力和长期成功铺平了道路。

关于NTT DATA

ntt data是一家超过300亿美元的商务和技术服务领导者，服务了《财富》全球100强企业的75%。我们致力于通过负责任的创新加速客户成功并对社会产生积极影响。我们是世界上领先的AI和数字基础设施提供商之一，在企业级AI、云、安全、连接性、数据中心和应用服务方面拥有无与伦比的能力。我们的咨询和行业解决方案帮助组织和社区自信且可持续地迈向数字未来。作为全球顶级雇主，我们在70多个国家拥有专家。我们还为客户提供接触强大创新中心以及成熟和初创合作伙伴生态系统的机会。ntt data是ntt集团的一部分，该集团每年在研发上投资超过30亿美元。

缩写列表

目标	人工智能管理系统
AIOps	人工智能运维
CAGR	复合年均增长率
GDPR	通用数据保护条例
HITL	人介入
MLOps	机器学习运维
NIST	美国国家标准与技术研究院
PDPC	个人信息保护委员会
联合国教科文组织	联合国教科文组织


NTT DATA