

# 以战略目标为导向， 实现数据现代化



## 前言

“以战略目标为导向，实现数据现代化”是由 Thoughtworks 赞助出品的麻省理工科技评论洞察报告。为了撰写本报告，麻省理工科技评论洞察对各个国家 / 地区各个行业的企业高管开展了一次全球性调查。此外，本报告还参考了对数据战略和数据现代化专家进行的深度访谈。

本报告作者是 Denis McCauley，编辑是 Teresa Elsey，制作人是 Nicola Crepaldi。本报告中的研究内容具有编辑独立性，所表达的观点仅代表麻省理工科技评论洞察的立场。

在此，感谢下列企业高管抽出宝贵时间与我们分享其真知灼见：

**Heath Bland**，埃克森美孚上游数据办公室经理

**James Morgan**，英国皇家财产局首席数据官

**Conrad Pozsgai**，Payback 公司总经理、首席信息官兼首席运营官

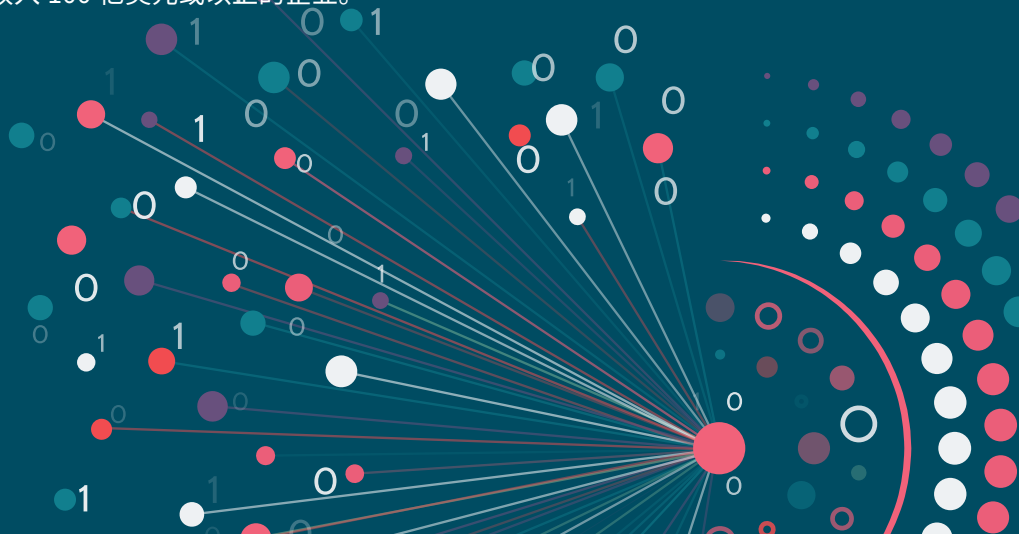
**Danilo Sato**，Thoughtworks 数据与人工智能事业部全球技术主管

**John Spens**，Thoughtworks 北美地区数据与人工智能事业部总经理

### 关于此次调查

麻省理工科技评论洞察于 2024 年 1 月至 2 月开展此次调查，为编制本报告奠定了基础。此次调查样本包括 350 名资深数据和技术主管。

受访者分别任职于美国、英国、德国、新加坡和澳大利亚的企业，样本涵盖了以下八个行业：金融服务、能源、医疗保健和生命科学、制造、公共部门、出版和媒体、零售以及旅游和运输。所有受访者均就职于年收入 5 亿美元或以上的企业，近一半（47%）就职于年收入 100 亿美元或以上的企业。



# 引言

如今，数据日益成为企业成功的关键。一些领先的数字企业能够及时做出明智决策，为客户提供量身定制的体验，并快速发现新机遇，从而脱颖而出。最近，随着人工智能的关注度突飞猛进，对数据的需求也与日俱增。

无论企业希望实现什么目标，如果无法随时获得高质量数据，就不可能取得成功。尽管科技在不断发展，但对于大多数企业而言，获取企业数据并将其转化为可用资产仍是一项巨大的挑战。

虽然提高数据质量和实现有效治理等诸多挑战存在了几十年，但本报告表明，当今的技术领军人物正在探索克服这些挑战的新方法。单靠新技术可能无法解决问题，不过技术型企业可以借助基于云的、现代化的数据和分析解决方案，采用现代化数据工程实践、打造产品导向型团队并实施更全面的数据管理方法。本报告的调查结果揭示了为何数据管理者们必须跳出现有系统的范围，采用各种流程、实践和思维方式来帮助其企业以可重复、可扩展的方式解决长期存在的数据挑战。

Thoughtworks 拥有丰富的数据现代化经验，熟悉各种曾引发软件工程变革的现代工程实践（例如数据即产品，构建演进式架构，采用 CI/CD 和 DataOps 等实践，发展团队和治理结构），深知如何利用这些实践并将其应用于数据标准，从而帮助企业在数据方面做好充分准备，并长期保持这种状态。

有了上述各要素，无论治理要求和数据需求如何变化，企业都始终能快速从数据中获取大量价值。

随着人工智能的发展，越来越多的企业将数据现代化作为核心要务，因此愈加重要的是确保数据战略与更广泛的商业战略紧密结合，同时确保数据管理者们能够清晰阐述各种数据和分析技术可以如何帮助企业业务部门实现其目标。这不仅可以保证数据现代化的变革获得关键的支持，还可以让企业为实现其目标做出更充足的准备，为持续改进、发展和创造价值奠定坚实的基础。

*Kalyanasis Banerjee*

**Thoughtworks 数据与人工智能事业部全球主管**

# 目录

<b>1. 执行摘要</b> .....	5
<b>2. 数据愿景</b> .....	7
数据战略的制定 .....	7
数据战略与业务战略不一致的问题 .....	8
<b>3. 数据现代化的目标</b> .....	10
实现数据现代化所面临的障碍 .....	11
为企业带来的各种好处 .....	13
不成熟的数据能力.....	14
<b>4. 质量管理</b> .....	17
数据质量和时效性改进 .....	18
组织结构和数据团队.....	19
<b>5. 结论</b> .....	21

# 01 执行摘要

**数**据现代化已明确列入很多企业的议程。我们对 350 名资深数据和技术高管进行了调查，有一半以上的受访者表示，他们所在的企业在过去两年中已经实施或正在实施数据现代化项目。此外，有四分之一的企业计划在未来两年内实施此类项目。其他研究也一致表明，企业纷纷加大了对数据资产现代化的投资力度。<sup>1</sup>

无独有偶，随着人们对提升数据能力的关注度激增，他们对人工智能（特别是生成式人工智能）的兴趣也愈发浓厚。实际上，为开发人工智能模型提供支持是本次受访企业致力于实施数据能力现代化改革的首要原因之一。不过，人工智能并非改革的唯一原因，甚至不是主要原因。

本报告旨在阐述企业数据现代化项目的目标，以及相关举措的实施方式。为此，本报告对各行业的资深数据和技术高管展开了调查。研究发现，许多企业在数据现代化方面进行了大量投资，并取得了实质性进展。然而，在许多企业中，数据战略与数据现代化目标似乎还远未达成统一，导致企业的数据和技术团队与其他部门之间出现脱节。数据和技术高管及其团队还可以开展更多工作，了解同事们的数据需求，并就如何满足这些需求积极征求他们的意见。

以下是主要研究结果：

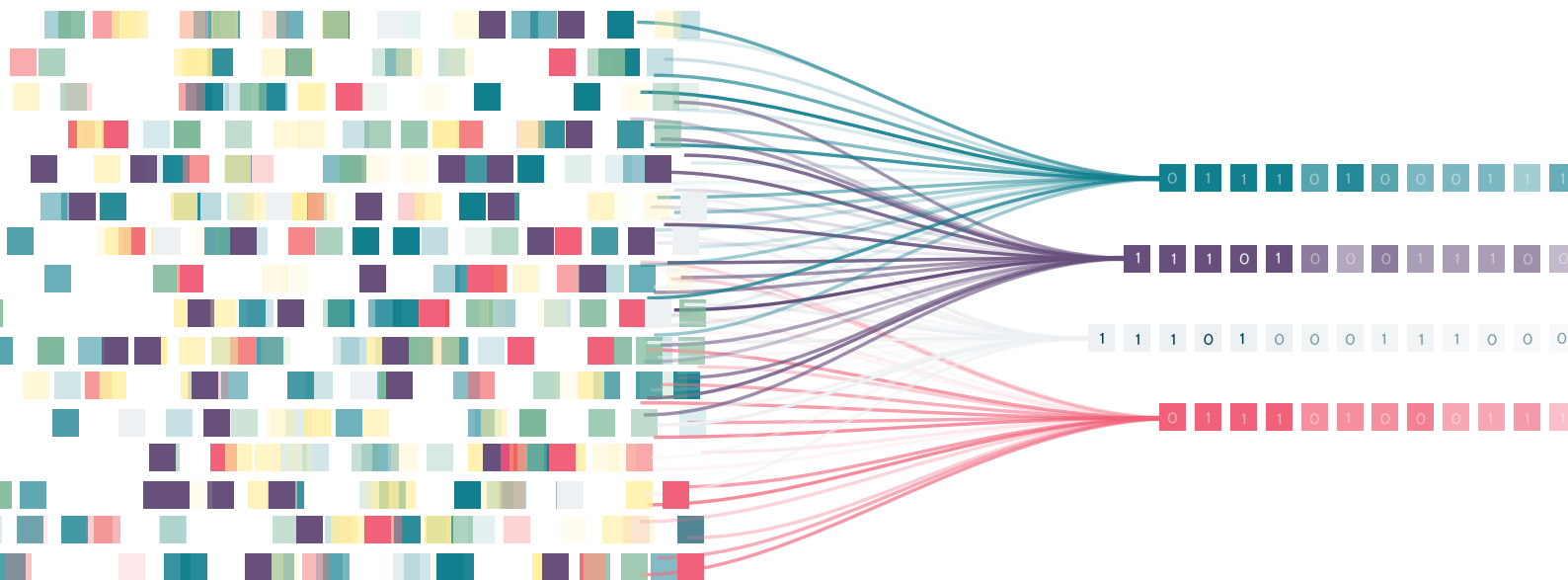
- **人工智能并不是企业推进数据现代化的唯一原因。**实现更明智的决策是数据现代化的首要目标，近一半（46%）的高管将此列为三大驱动因素之一。支持人工智能模型（40%）和脱碳（38%）也是推进数据现代化的主要驱动因素，此外其他驱动因素还包括增强监管合规性（33%）和提高运营效率（32%）。
- **数据战略往往与商业战略脱节。**几乎所有受访企业都认识到了数据战略的重要性。仅有 22% 的企业表示，他们缺乏完善的数据战略。然而，当被问及他们的数据战略是否完全契合关键商业目标时，只有 39% 的企业给出了肯定的回答。数据团队还可以采取更多措施，让其他业务部门和职能部门参与战略讨论：42% 的受访企业表示，他们的数据战略完全由数据或技术团队制定。
- **数据战略为数据现代化铺平道路。**在过去两年中开始进行数据现代化的大多数企业（71%）在此之前早已制定好数据战略，这可能并非巧合。数据现代化目标的实现离不开业务的支持，实施决策需要战略指导，以免增加复杂性或重复工作。

· **数据方面的最大痛点是数据质量和时效性。**高管们指出，数据不达标（41%）和交付不及时（33%）是其数据运营中最需要改进的方面。数据不完整或不准确会导致企业用户质疑数据的可信度。这就解释了为什么在过去两年中，受访企业最普遍采取的数据现代化措施是加强数据治理（45%）。

· **跨职能团队和 DataOps（数据运维）是提高数据质量的关键手段。**许多企业正在落实现代化数据工程实践。近一半（48%）的企业授权跨职能数据团队执行数据质量标准，47%的企业将实施 DataOps 作为重点工作。这些实践与已经成为软件工程领域标准的敏捷方法和产品思维不谋而合，但它们在数据领域才刚刚崭露头角。

· **合规性和安全性问题往往会阻碍数据现代化进程。**合规性和安全性问题是数据现代化面临的主要障碍，各有 44% 的受访企业提到这两大问题。能源、公共部门、运输和金融服务等领域的受访企业尤其经常提到监管合规性问题。高成本是受访企业经常提及的另一个障碍（40%），尤其是在受访的小型企业中。

许多企业正在落实现代化数据工程实践。这些实践与已经成为软件工程领域标准的敏捷方法和产品思维不谋而合，并在数据领域展现出价值。



# 02 数据愿景

在过去十年中，企业纷纷采用人工智能，这也加强了企业对其数据的重视。毕竟，如果没有足够的数据库，人工智能模型几乎无法产生价值，而如果数据质量不高，还可能会造成危害。

最近，随着生成式人工智能模型的兴起和广泛应用，企业进一步认识到数据的巨大潜力，特别是他们可能会在其内部数据、非结构化数据或其他被忽视的数据中发现独特价值。这也揭示了在数据基础尚不成熟的情况下，抓住这些机会向前迈进所面临的诸多困难。

然而，要让管理层相信这些数据就是战略性资产，不应依赖面向企业的人工智能。早在人工智能成为高效商业工具之前，从风险管理到需求预测等许多关键的企业运营工作都要仰仗扎实的数据基础。

## 数据战略的制定

目前，企业普遍认识到了数据的重要性，因此大多数企业都制定了完善的数据战略。在此次调查中，只有约五分之一（22%）的高管表示他们所在的企业没有这样的战略。

然而，这是最近才出现的情况。只有 27% 的受访高管表示，他们所在企业的数据战略已实施了两年以上（见图 1）。

在此次调查中，小型企业（年收入在 5 亿美元到 99 亿美元之间）比大型企业（年收入在 100 亿美元或以上）更有可能缺乏数据战略，而且已实施数据战略两年以上的可能性要小得多。这是不同规模的企业在应对措施方面出现的多种差异之一。这些差异表明，企业规模和资源在数据战略和数据现代化方面具有重要影响。

“数据战略应满足技术和技能需求，但更应与企业的总体业务目标挂钩，否则将无法给企业带来真正的价值。”

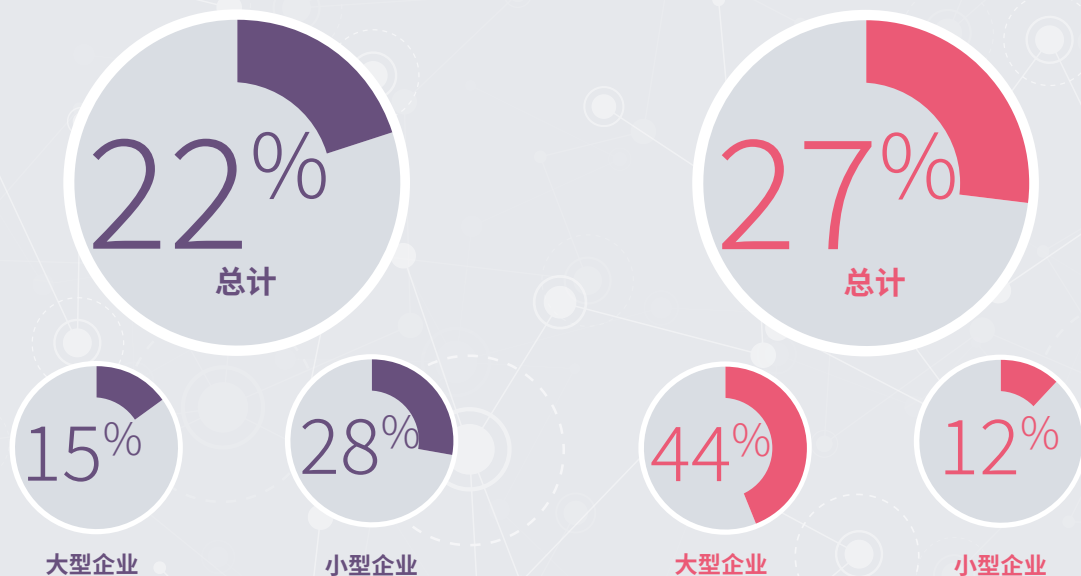
**Danilo Sato**, Thoughtworks 数据与人工智能事业部全球技术主管

图 1：大多数企业都已制定数据战略，但这是最近几年才出现的情况

以下哪项陈述适用于您所在企业的数据战略？

我们尚未制定完善的数据战略

我们已制定成熟的战略，并实施了两年以上时间



来源：麻省理工科技评论洞察调查，2024 年

英国皇家财产局首席数据官 James Morgan 表示：“每个企业都需要数据战略。愿景固然重要，但如果无法确定如何达到目标，那么有愿景也毫无意义。” Morgan 的团队根据数据战略制定了一项实施计划：“我们不可能面面俱到，因为业务重点和需求会发生变化，但必须制定一项战略实施计划。”

Thoughtworks 北美公司数据和人工智能事业部总经理 John Spens 认为，“如果没有利用数据解决企业问题的统一方法，那么数据团队将永远处于被动状态。大型企业的各个业务部门拥有可观的预算，如果缺乏战略，就会导致重复工作和产生不必要的复杂性。”他还指出：“正因为如此，许多公司发现自己在多个云平台上运营，但这并不是根据深思熟虑的战略决定的，而是因为各部门的云方案的供应商不同”

## 数据战略与业务战略不一致的问题

大多数受访企业都制定了数据战略，但相当一部分企业并没有在企业层面落实完全统一的数据战略（见图 2）。在 40% 的受访企业中，各业务部门都有自己的数据战略，但这并不一定意味着缺乏统一性。例如，能源供应商埃克森美孚上游数据办公室经理 Heath Bland 表示，“虽然业务线数据战略可能是在企业、业务能力和业务线 / 资产层面制定的，但都由中央数据办公室进行协调。这种方法可以确保不同的战略相互补充、相辅相成，我们认为这是最大限度发挥数据价值的关键。”

相当一部分（39%）的受访高管表示，他们的数据战略与业务战略目标完全一致。然而，对于大多数尚未实现这种一致性的企业，这可能会带来高昂的代价。Thoughtworks 数据与人工智能事业部全球技术主管 Danilo Sato 表示，“这种不一致会导致企业无法发挥数据战略的价值。数据战略当然应满足技术和技能需



图 2：数据战略尚未达成统一

以下哪项陈述适用于您所在企业的数据战略？

我们的数据战略与业务战略目标完全一致



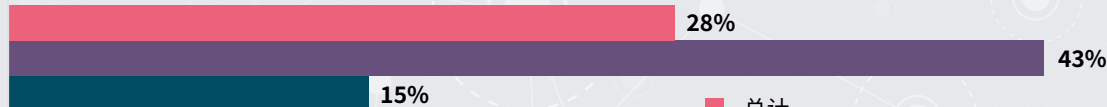
各业务部门有自己的数据战略



数据战略完全由数据或技术团队制定，其他业务部门的参与程度极低



我们的数据战略与人工智能战略密切相关



■ 总计

■ 大型企业 (收入不小于100亿美元)

■ 小型企业 (收入小于100亿美元)

来源：麻省理工科技评论洞察调查，2024 年

求，但更应与企业的总体商业目标挂钩，否则将无法实现企业真正需要的价值。”

造成这种明显缺乏一致性的一个原因是，数据战略往往完全由数据团队或 IT 部门制定，企业其他部门的参与度极低。有 42% 的受访者表示，他们所在的企业正存在这一问题。

首席数据官（在某些情况下是首席信息官）或其团队很有可能负责领导制定企业数据战略，但他们不应孤立地开展此项工作。Spens 说：“作为首席数据官，你可能需要与其他所有高管坐下来探讨，你也是他们的服务提供商，有责任了解他们的需求，并确保你正在制定的战略符合这些需求。”

欧洲客户忠诚度共享计划和营销平台 Payback 的常务董事、首席信息官兼首席运营官 Conrad Pozsgai 指出，这种协调并非一蹴而就。他表示：“如何协调数

据战略并使其与其他业务职能保持一致，这是一项长期存在的挑战，需要进行大量的沟通，不断完善。”

目前致力于发展人工智能能力的企业还会发现，需要密切协调其数据和人工智能战略。超过四分之一（28%）的受访者表示，在他们所在的企业中，数据和人工智能密切相关，不过就这方面而言大型企业（43%）远远高于于小型企业（15%）。

专家们认为，这种整合会变得越来越普遍。Sato 说：“密切协调十分重要。”他举了个例子一个示例：有些团队正在使用基于开源大语言模型的生成式人工智能。他表示：“如果这些模型是使用外部数据训练的，相关团队可能会希望通过整合一些自己的数据加以优化。人人都在设法将自己的专业知识与大语言模型相结合，以实现自己的目的。为此，确实需要将数据和人工智能放在一起考虑。”

# 03 数据现代化的目标

在 各企业着手实现数据能力现代化的过程中，统一且协调的数据战略尤为重要。在此次调查中，大多数企业在过去两年内对其数据资产的多个要素进行了现代化改造（23%），或正在进行现代化改造（31%）。另有 23% 的企业计划在未来两年内进行现代化改造。大多数已进行现代化改造的企业（71%）实施数据战略已超过两年（见图 3）。

数据现代化可能需要进行大量投资。这也就解释了为什么在此次调查中，那些年收入达到或超过 100 亿美元的大型企业比小型企业更有可能已开始进行数据

现代化改造（见图 4）。除了投资之外，还需要具备远见和耐心才能有所回报。Sato 表示，“尽管多年来进行了大量与数据相关的投资，许多企业直到最近才开始报告这些投资的成功率有所提高。”

然而，无论是在财务方面还是在其他关键指标方面，展示早期回报的能力对于让数据现代化持续获得支持至关重要。能够展示早期成果的数据现代化计划往往会有助于全面执行计划。Sato 表示：“虽然对数据现代化的投资通常需要两到五年的时间才见成效，但企业为实现此目标而采取的方法却会产生巨大的影

图 3：数据现代化的现状：所有受访者

以下哪项描述最符合您所在企业目前的数据能力？



来源：麻省理工科技评论洞察调查，2024 年

“我们相信小切片增量式交付方法能尽早展示回报。虽然规模较小，但早期的成功可以增强对数据现代化战略的信心，并赢得内部的支持，有助于推动长期规划的落实。”

**Danilo Sato**, Thoughtworks 数据与人工智能事业部全球技术主管

响。我们不希望前期在基础设施、平台和数据整合工作上做过多的投入，而是会在数据可用后尽快推动实现相关用例，我们相信小步增量式交付方法能尽早展示回报。虽然规模较小，但早期的成功可以增强对现代化战略的信心，并赢得内部的支持，有助于推动长期规划的落实。”

### 实现数据现代化所面临的障碍

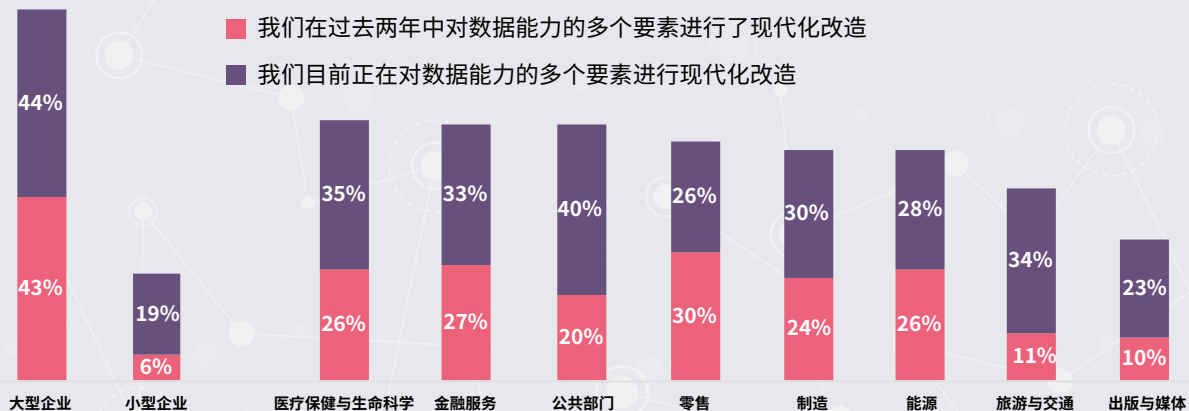
高成本是数据现代化面临的主要障碍之一：有 40% 的受访者是这样认为的。不出意料，这对此次受访的小型企业（46%）来说似乎是一个更沉重的负担。但也有三分之一的大型企业同样表示成本是他们考虑的一个因素（见图 5）。

然而，受访者提及次数最多的数据现代化障碍是监管合规性和安全性问题（两个方面均有 44% 的受访者提及）。在金融服务和医疗保健等监管严格的行业，可能会出于合规性方面的考量而禁止统一存放特定的数据，或对数据存储于公共云中有所限制。对于能源、公共部门、运输和金融服务行业的企业来说，监管是数据现代化进程面临的尤其重大的一个障碍。

安全性问题也会阻碍部分数据现代化工作。在调查中，公共部门的受访者以及金融和零售行业的受访者都特别强调了此类问题，而这些行业都是网络犯罪分子窃取重要个人信息和银行信息的常见目标。

**图 4：数据现代化的现状：按企业规模和行业划分**

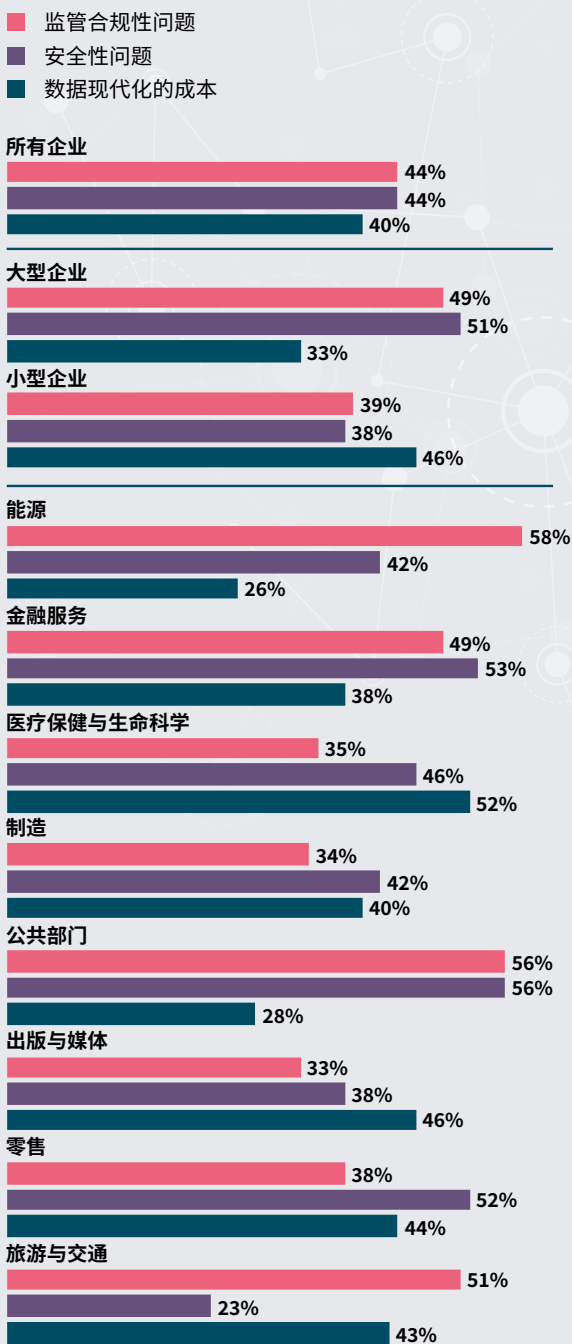
以下哪项描述最符合您所在企业目前的数据能力？



来源：麻省理工科技评论洞察调查，2024 年

图 5：是什么阻碍了数据现代化？

以下哪些项是您所在企业在实现数据能力现代化方面面临的主要障碍？  
(排名前三的回答)



来源：麻省理工科技评论洞察调查，2024 年

## 关于云成本

在讨论数据现代化的成本时，云是一个备受关注的焦点。企业高管们可能会担心，人工智能所需的数据量不断扩大，再加上云随用随付的易扩展性，会导致成本失控。但与云所取代的基础设施相比，云的成本的透明度能让人们对云服务价格一目了然。

随着数据现代化工作的推进，数据的使用率提高，效率低下的问题可能会日益严峻。Thoughtworks 数据与人工智能事业部全球技术主管 Danilo Sato 表示：“无论何种规模的企业，其云服务成本都在攀升。例如，扩大员工对数据的访问权限是数据民主化的关键推动因素，提升了云的使用率，进而迅速暴露出长时间查询或低性能代码的成本。因此，采用成熟的数据设计现代化方法和编写高性能代码，可以带来以前无法实现的成本效率。”

除了良好的数据工程实践外，若要利用云的财务优势，可能还需以更专业的方式研究其成本。FinOps 是“Finance”（财务）和“DevOps”（开发运营）的合成词，指的是使财务专家和 IT 专家协作管理数据、人工智能和云成本的一套新型实践。FinOps 方法有助于企业确信其云成本可监控、可预测并符合企业的总体战略，目标是最大限度地提高企业的云投资回报，而不仅仅是削减成本。

## 为企业带来的各种好处

企业希望通过数据现代化实现什么目标？排名第一的回答是提高整个企业的决策能力。近一半（46%）的受访高管提到了这一点（见图 6）。

这是英国皇家财产局数据现代化工作的核心目标。Morgan 表示：“我们需要做出正确的决策，为了国家的利益优化利用我们所有的资产，无论是伦敦的地块还是房产。这不仅要追求最大经济产出，还要兼顾社会、环境和价值创造效益。”

受访者提到的第二大数据现代化目标是支持人工智能用例的开发（40%）。在这方面，云可能是一个重要的战略选择，它可以提供硬件和基础设施，以“随用随付”的方式进行人工智能模型训练。Pozsgai 说：“支持人工智能模型无疑是推动我们实现数据现代化的一大动力，也是我们上云的部分原因，因为在本地运行模型的成本太高，此外我们还需要云供应商的技术和其他能力。”

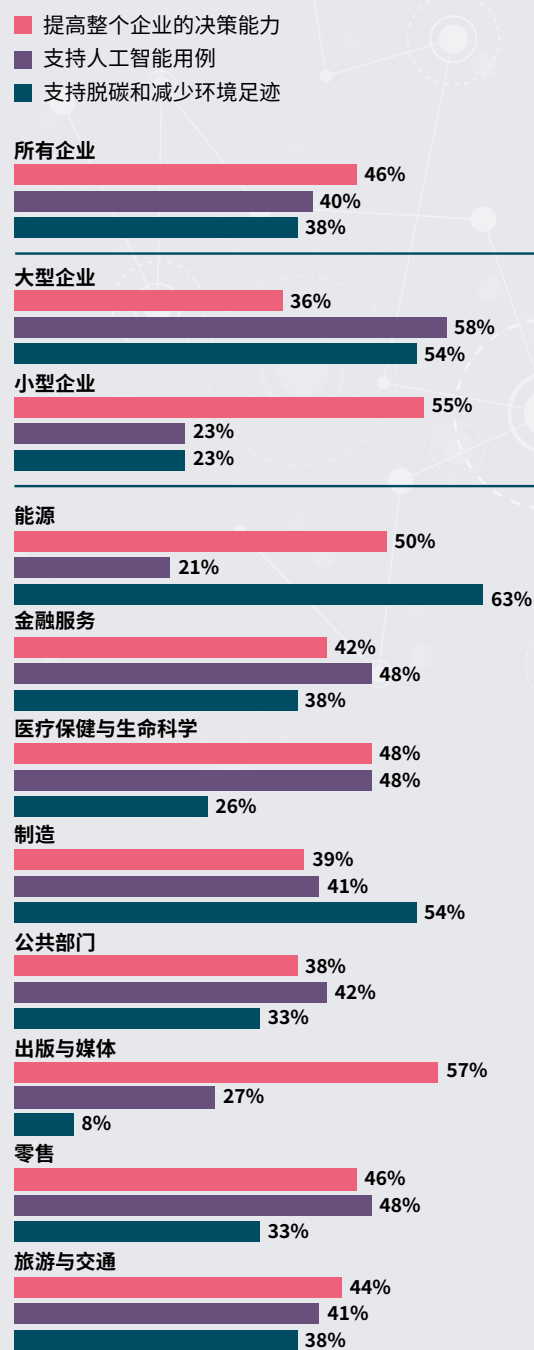
这一点因企业规模不同而有很大差异：受访企业中，大型企业将支持人工智能作为数据现代化目标的可能性是小型企业的两倍半（分别为 58% 和 23%）。

“我们需要做出正确的决策，为了国家的利益优化利用我们所有的资产。这不仅要追求最大经济产出，还要兼顾社会、环境和价值创造效益。”

**James Morgan,**  
英国皇家财产局首席数据官

图 6：数据现代化的主要目标

您所在企业实现数据能力现代化的主要动机是什么？  
(排名前三的回答)



来源：麻省理工科技评论洞察调查，2024 年

## “我们的数据基础设施和流程正在经历巨变,这是向‘产品交付企业’转型的必经之路。”

**Conrad Pozsgai,**

Payback 公司总经理、首席信息官兼首席运营官

排名第三的数据现代化驱动因素 (38%) 是支持脱碳和减少企业的环境足迹。对于那些在过去两年中已进行数据现代化的企业来说,支持人工智能 (70%) 和支持脱碳 (60%) 是他们目前为止的主要动机。其他重要的数据现代化驱动因素包括提高监管合规性 (33%) 和运营效率 (32%)。

### 不成熟的数据能力

随着数据现代化工作的推进,高管们应清晰认识到其现有数据能力的薄弱环节,特别是那些可能危及既定数据现代化目标的实现的环节。此次调查罗列了许多潜在的薄弱环节,其中几项引起了高管们的高度重视,包括数据团队的人才和技能缺口、数据战略(如上所述,这可能是指现有战略不成熟或战略缺失)以及企业数据架构的现状(见图7)。

### Payback 公司: 利用数据加速发展

Payback 公司在业务方面取得了成功,主要原因是他们为德国及其他 10 个国家/地区的零售客户设计了有效的忠诚度计划 and 多渠道营销活动。由 Conrad Pozsgai 带领的公司技术和数据团队责任重大,需以越来越快的速度实现新的计划功能和新的营销活动。Pozsgai 表示:“作为一家技术型企业,我们的主要目标是加快向市场推出新功能。当然,如果我们的速度更快,还可以降低开发成本。”

为了达到理想的速度,Payback 公司对其数据资产进行了大规模的数据现代化改造,包括改用现代化数据工程实践。Pozsgai 表示:“我们的数据基础设施和流程正在经历巨变。这是向‘产品交付企业’转型的必经之路。”

除了采用新的模块化数据架构外,数据现代化还涉及巨大的结构变化。Pozsgai 解释说,数据型企业

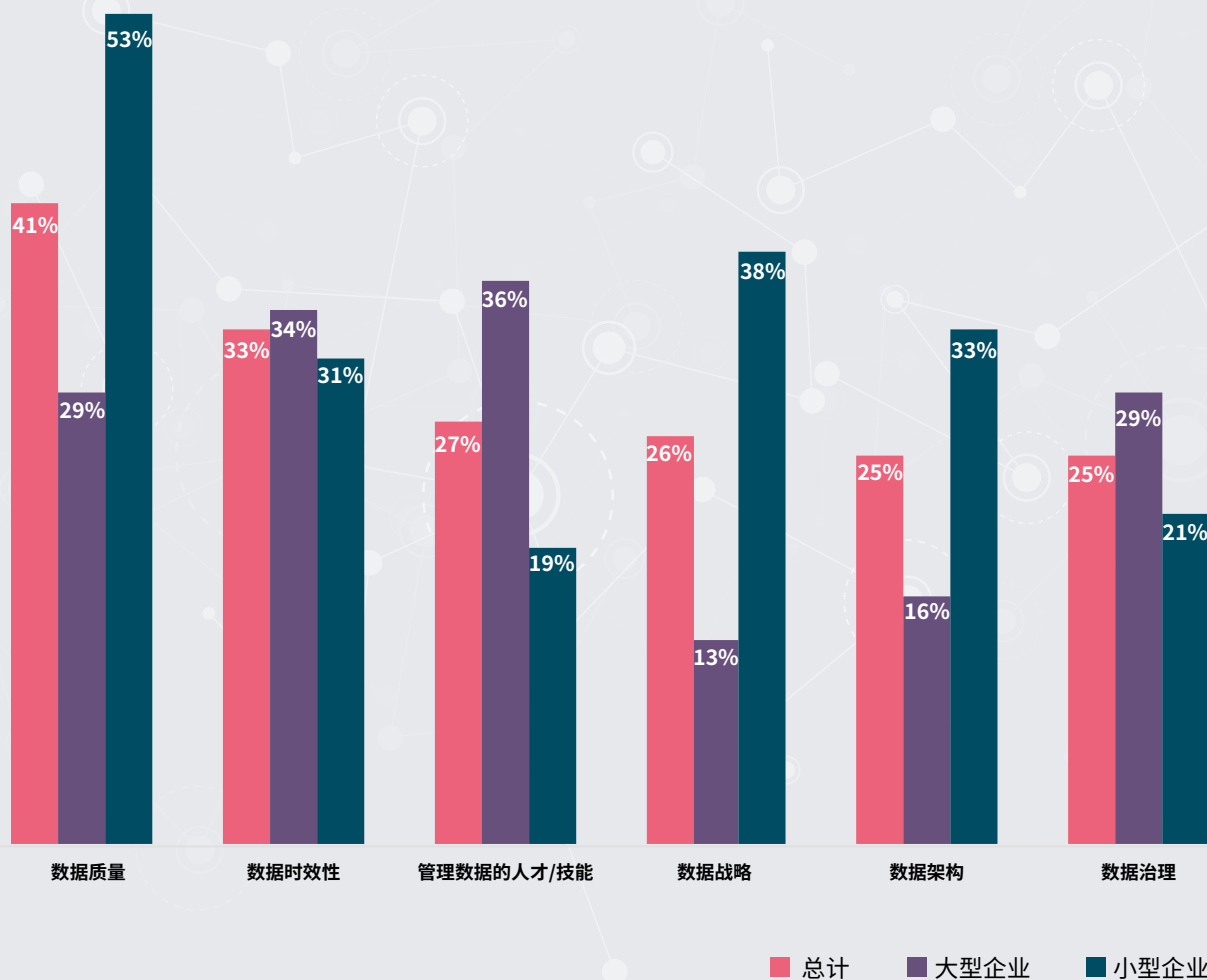
正在从集中式结构过渡到领域驱动型结构:“这就是我们的业务领域,因此我们正逐步将数据团队转移到相同的领域。他们将为某一产品或产品集群开展跨职能工作,这对我们所有的数据人员来说都是一个重大变化。”

通过领域驱动型结构,数据团队获得半独立工作的自主权。Pozsgai 强调这些团队必须能够独立开发数据产品,但这并不能一蹴而就。他说:“目前,我们所有的团队都在全力应对变革,离不开总部的支持。不过,随着时间的推移,这些团队会变得越来越独立。”

数据分析和其他数据工作不能再像过去那样进行。Pozsgai 表示:“事务和通信都具有实时性,越来越多的数据分析工作需即时完成。因此,像其他技术领域一样,我们需将数据纳入整个交付流程中。”

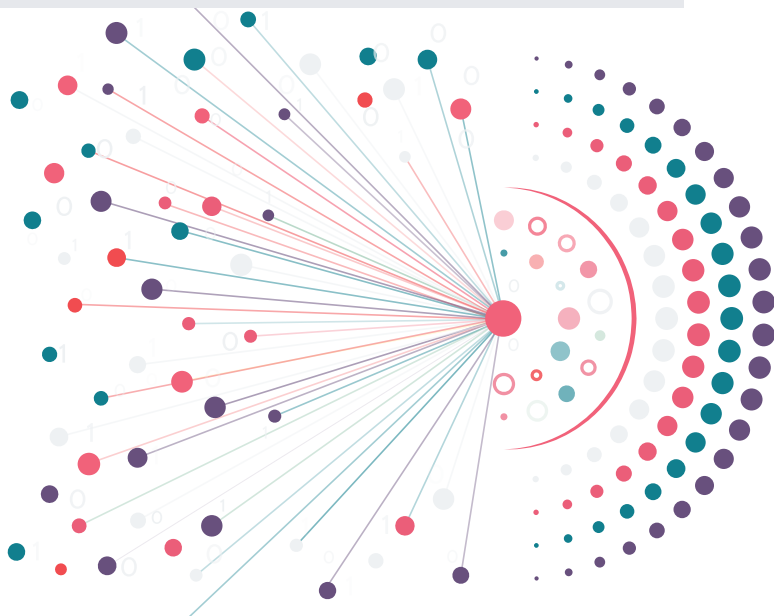
图 7：企业数据能力的薄弱环节

您所在企业数据能力中的哪些要素目前对实现数据现代化目标的支持最不力？（排名前三的回答）



来源：麻省理工科技评论洞察调查，2024 年

然而，受访者最关注的是数据质量（41%）和数据时效性（33%）。有 53% 小型企业受访者认为数据质量是首要薄弱环节。正如下文所述，高管们认为加强数据治理是解决此类数据质量和数据时效性问题的关键。



## 埃克森美孚：数据现代化的效益与痛点

作为全球收入排名第八的公司，埃克森美孚需对其上下游石油、天然气和石化业务及资产中生成的海量数据进行处理和分析。要对如此庞大且复杂的数据资产进行现代化改造，绝非易事。

埃克森美孚上游数据办公室经理 Heath Bland 表示，该公司在实现数据现代化的过程中遇到了若干挑战，其中有两个尤为棘手。其一是供应商解决方案受限。他说：“尽管有云产品，但我们行业的企业可用解决方案的选择范围仍然很窄。由于客源有限而且行业具有独特性，我们的市场无法吸引新的参与者来开发解决方案。”

Bland 指出，另一个棘手挑战是公司数据文化的现状，“这一点，再加上各单位对技术挑战的认识不足，协调工作不到位，而且（数据现代化）缺乏直接、可见的业务价值影响，导致资金不足，资源有限，无法实现数据分析”。

不过，Bland 还表示，埃克森美孚已经开始看到数据现代化带来的切实回报。他列举了五大回报：

- **部署自助服务分析功能**：“我们已经授权用户独立访问和分析数据，实现了人工智能模型训练”。
- **确保企业数据公平合理**：“这意味着要提高我们数据的可查找性、可访问性、互操作性和可重用性”。
- **减少“孤岛化”的影响**：“我们已经能够纵观全局、跨孤岛进行数据融合，从而获得新的洞见。”
- **加强行业数据共享**：“我们采用了 OSDU 数据标准<sup>2</sup>和数据平台，实现行业级数据整合。”
- **实行业数据标准**：“借此我们能够实现业务数字流程的转型和数据现代化，并通过一致性降低数据成本”。

Bland 表示：“我们已经开始看到其中一些好处，但还有些好处需要更多的时间才会表现出来。已取得的进展归功于公司向‘数据即产品’思维方式的转变，以及数据治理的日臻成熟。”

“各单位对技术挑战的认识不足，协调工作不到位，而且（数据现代化）缺乏直接、可见的业务价值影响，导致资金不足，资源受限，无法实现数据分析。”

**Heath Bland,**  
埃克森美孚上游数据办公室经理



# 04 数据质量治理

**M**organ 表示：“保障数据质量是我们最大的痛点。”英国皇家财产局采集的数据类型多样，格式各异，如何确保数据统一存储并准确对其进行估量和利用，是一项长期存在的挑战。

如果因数据缺失、数据重复（数据孤岛会导致这个问题）、手动数据录入错误或其他因素造成数据质量不达标，那么用户就会质疑数据的可信度。Sato 补充表示，时效性也是一个相关因素。他指出：“如果企业用户无法在需要时访问数据，就会降低数据质量。”

Pozsgai 以 Payback 公司的运营情况举例：“我们开展了很多实时营销活动。客户在线下或线上的行为往往会实时触发我们的推送信息。我们必须掌握客户行为的相关数据，以便立即采取行动。”

Bland 表示，提高时效性对石油和天然气行业而言至关重要。他解释说：“实时或近实时的数据访问可提高运营数据和地下数据的价值。随着我们在高级分析和人工智能方面取得进展，数据时效性的重要性将与日俱增，并为业务流程转型创造新机遇。我们已看到在提高网络容量和速度、数字孪生、数据平台、数据采集和流程自动化等方面取得的进步。”

提高数据质量和践行维护确保数据完整性的数据实践，也是良好的数据治理的一大功效。这或许可以解释为什么正在进行数据现代化的企业会优先考虑改善治理。当被问及过去两年为实现数据能力现代化而采取的措施（见图 8），受访者最普遍提到的是升级治理模式（45%）。

**“随着我们在高级分析和人工智能方面取得进展，数据时效性的重要性将与日俱增，并为业务流程转型创造新机遇。”**

**Heath Bland,**  
埃克森美孚上游数据办公室经理

图 8：数据现代化措施的时限

您所在组织何时采取以下措施实现数据能力现代化？

#### 加强数据治理



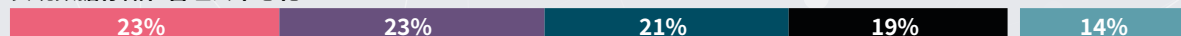
#### 采用全新的数据架构



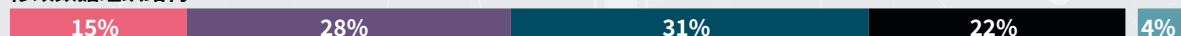
#### 扩大云数据服务的使用范围



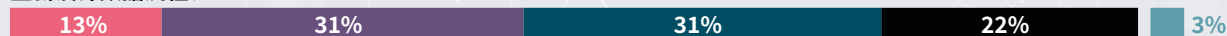
#### 实现数据存储和管理去中心化



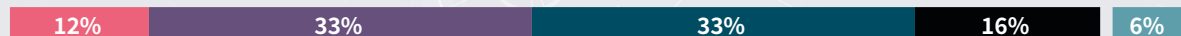
#### 修改数据组织结构



#### 重新设计数据流程、



#### 整合数据库



来源：麻省理工科技评论洞察调查，2024 年

自动化数据治理即利用软件实现治理策略的自动化，是一种新兴技术。Sato 说：“一种现代的实现数据治理的方式是自动化地在平台内支持数据治理。通过自动化，企业可执行或监控其商定的策略（如数据安全、数据质量、访问控制和数据隐私），并将数据治理从流程繁琐的方法转变为赋能方法。”

其次最常见的数据现代化措施是采用全新的数据架构，有 36% 的受访者落实了此项措施。对于理想的数据架构模式，不同的企业有不同的看法。有些企业倾向于采用领域驱动型方法，赋予跨职能数据团队高度自主权。Payback 公司数据现代化计划的一大特点就是向这种模式转变。另一些企业则通过集中式平台和数据库，提高架构统一性。还有一些企业则试图将这两种方式合二为一，采用一种将集中与分散相结合的混合方法。

不过，就治理模式而言，企业在很大程度上达成了共识。为了确保数据质量达到最高水平，整个企业需确立一个总体治理框架。随着 Payback 公司向领域特定结构转型，Pozsgai 认为保持集中数据治理模式显得尤为重要。他说：“当然，领域负责人应为其团队监控数据质量，但不同领域和团队之间存在很多依赖关系，需要确立一个总体治理框架。”

## 数据质量和时效性改进

为了提高数据质量和时效性，受访企业还采取了更具体的措施（见图 9）。为此，近一半（47%）的企业优先考虑实施 DataOps。DataOps 方法的核心原则是自动执行数据质量管理，其中包括测试、监控和异常检测。DataOps 还强调数据团队的跨职能协作。另外，有 48% 的受访者表示，他们的企业都授权了数据团队执行数据质量标准。

图 9：提高数据质量和时效性

在提高各业务部门所用数据的质量和时效性方面，您所在企业的首要任务是什么？

48%

授权跨职能数据质量团队  
执行质量标准

47%

实施DataOps

35%

改善数据验证流程

31%

在整个企业内开展有关数据质量和  
标准的培训

29%

重新制定数据质量标准

23%

实现数据质量管理自动化

来源：麻省理工科技评论洞察调查，2024 年

Spens 认为，监管是提高数据质量的另一个重要因素：“当然，企业必须在相关技术措施上投资，以解决数据质量问题。然而，高层的积极监管同样重要，业务主管们必须非常关注数据质量。”

## 组织结构和数据团队

上述一切都凸显了组织和结构变革对数据现代化的重要性。Sato 表示：“诸多数据现代化实践都关乎重新思考数据团队在企业中的位置及其构成。”受访者非常重视跨职能团队，且此类团队应成为各组织结构必不可少的一部分。他提到：“各项目的推进都离不开跨职能团队。如果将领域专家、数据科学家和工程师以及分析师集中在同一团队，就能减少交接延误，优化流程。这些团队就会更快地实现价值。”

Morgan 指出，在数据团队中，技能等方面的多样性十分重要，对于团队中的数据专家更是如此。他说：“团队中有些人擅长编码，但可能不太了解业务背景，那么就需要擅长沟通的人与他们合作。这些人不仅对技术有所了解，还能清晰讲述业务成果和价值。”

## 什么是 DataOps？

DataOps 之于数据如同 DevOps 之于软件，两者都是基于敏捷方法的理念和原则（而非具体技术），强调自动化和持续改进。DevOps 的主要目标是以更快的速度交付更好的软件。DataOps 的主要目标则是提高数据质量，为业务最终用户提供更好的洞见，同时降低数据管理成本。

DataOps 方法旨在优化数据团队的设计、构建和交付工作，助力加快创建可扩展且灵活可靠的数据产品和架构。采用 DataOps 方法的团队通常会通过自动测试和部署实现持续交付，从而专注于推动反馈和改进的快速循环。

实施 DataOps 还需要进行文化变革，打破数据孤岛，促进各团队和职能部门之间的协作，确保数据为业务带来价值。

受访企业似乎倾向于将结构变革推迟到数据现代化进程的后期。虽然许多企业在过去两年中升级了数据治理并采用了新架构,但只有 15% 的受访者表示他们的企业在这个过程中解决了组织结构问题。大多数受访者 (59%) 表示,这将是他们明年的工作重点,而 22% 的受访者预计至少要一年后才能解决此类变革问题。

数据资产的现代化是一项艰巨的任务,高管们可能会认为过快引入过多变革并非明智之举。改变人员角色和报告结构是最艰难的变革,要克服那些必然会出现的阻力,就可能需要企业的数据和技术高管给予极大的关注。

**“当然,企业必须投资采取相关技术措施,以解决数据质量问题。然而,高层的积极监管同样重要。”**

**John Spens,**  
Thoughtworks 北美地区数据与  
人工智能事业部总经理

## 英国皇家财产局：攻克数据复杂性难题

James Morgan 及其数据专业团队肩负着一项特殊使命：他们需要通过英国皇家财产局的自然资产和建筑资产所产生的数据创造价值，提高效率并优化决策。这些资产种类繁多，包括农田、沿海环境、海底，以及伦敦和其他城市的商业建筑和住宅，不一而足。Morgan 的团队负责管理数据，这些数据涉及海底测绘和建模、农田碳排放、住宅所用材料、商业地产的人流量和营业额以及其他众多类别的数据。Morgan 表示，该企业的数据库系统必须支持关于这些资产的分析、报告、数据科学和决策工作。

为了更高效地管理数据并从中获取更多价值，该企业正在对其整个数据资产进行现代化改造。这项工作的重点之一是采集数据的业务运营系统。Morgan 说：“我们一直在识别目前已采用的哪些人工操作属于权宜之计。我们正致力于确保我们的系统能够获得高效运营所需的一切信息。”他解释说，这些运营系统是他的团队开展相关数据工作的基石，例如分析、报告或数据科学工作。“总而言之，我们关注数据创建和数据变更，需要确保所有架构和系统以及与之相关的业务流程都正确无误，并建立适当的数据管理和治理机制。”

Morgan 补充说：“我们目前就在对数据和分析系统进行现代化改造，包括将其部署到云环境中。我们正在云端建立一个数据生态系统，并在该系统上构建一个数据湖和结构化环境。这是一项艰巨的任务，需将许多传统的孤岛式系统整合到现代云数据架构中。

Morgan 采取一种混合组织模式来管理现代化数据资产。他说：“对于我们这样的企业来说，需构建一个中央环境，在这个环境中维护适当的治理、保护和控制，并且聚集环境管理专家。”不过，这并不意味着所有的数据都需要关联到一个通用模型中。Morgan 还表示：“我们希望在企业内实现数据民主化，将数据交到同事手中，为他们的日常工作和决策提供支持。例如，在许多情况下，我们的海洋环境数据并不需要与我们的伦敦地产资产相关联。将两者相关联并不会带来任何好处，我们为什么要这样做呢？”有时，集中整合数据的成本和时间远远大于收益。

# 05 结论

有些企业正在进行或计划进行数据资产现代化，而本研究报告为这些企业的数据和技术主管提供值得借鉴的丰富经验，其中最重要的经验原则如下：

正确看待人工智能目标。虽然人工智能有望为企业增值，但其他现代化目标也要兼顾。更快、更安全地提供更高质量的数据，这不仅能满足人工智能模型的需求，还能满足许多其他关键业务系统以及其他新兴技术的需求。数据现代化的全面业务案例远不止人工智能。

数据高管必须展示数据现代化的业务价值。全面的数据战略需与业务战略完全协调一致，而孤立地制定数据战略和现代化方法势必会导致精力和资源的浪费。资深数据和技术高管有责任了解如何利用数据帮助业务部门实现其目标，并有针对性地开展相关工作。此外，他们还必须就数据现代化的好处与负责数据投资审批和支持工作的相关业务高进行沟通。

软件工程实践逐渐引入到数据工程领域。如今，跨职能数据团队、DataOps 实践、基于云的解决方案纷纷涌现，而且数据即产品的理念备受关注，这些都呼应了过去二十年来已成为软件工程标准的敏捷实践。在试图加快数据价值交付的过程中，企业发现采用现代化数据工程实践可以解决数据质量和可用性方面的挑战。

数据现代化可能要求针对企业的组织结构和数据架构进行复杂的变革。许多企业倾向于将这些变革推迟到项目的后期阶段以减少项目早期的未知性。然而，现代化举措可通过展示早期价值来获得动力，这意味着高管们可以从敏捷思维中获益，及早行动，做好转型的准备。



## 关于《麻省理工科技评论》洞察

麻省理工科技评论洞察是《麻省理工科技评论》的定制出版部门，后者是全球历史最悠久的科技类杂志，由全球著名科技机构提供支持，针对当今前沿科技和商业挑战开展实时活动和研究。麻省理工科技评论洞察在美国和国外开展定性和定量研究与分析，并出版各种内容，包括文章、报告、信息图表、视频和播客。随着《麻省理工科技评论》全球洞察专家组的不断壮大，我们可以广泛接触到世界各地的高管、创新者和企业家，对他们进行调查和深度采访。

## 关于 Thoughtworks

Thoughtworks 是一家集战略、设计和工程服务于一体的全球技术咨询公司，致力于推动数字创新。我们在 19 个国家设有 48 个办事处，拥有超过 10,500 名员工。在过去 30 年的时间中，我们以技术作为竞争优势，帮助客户解决复杂的业务问题，为客户带来非凡体验。



### 脚注

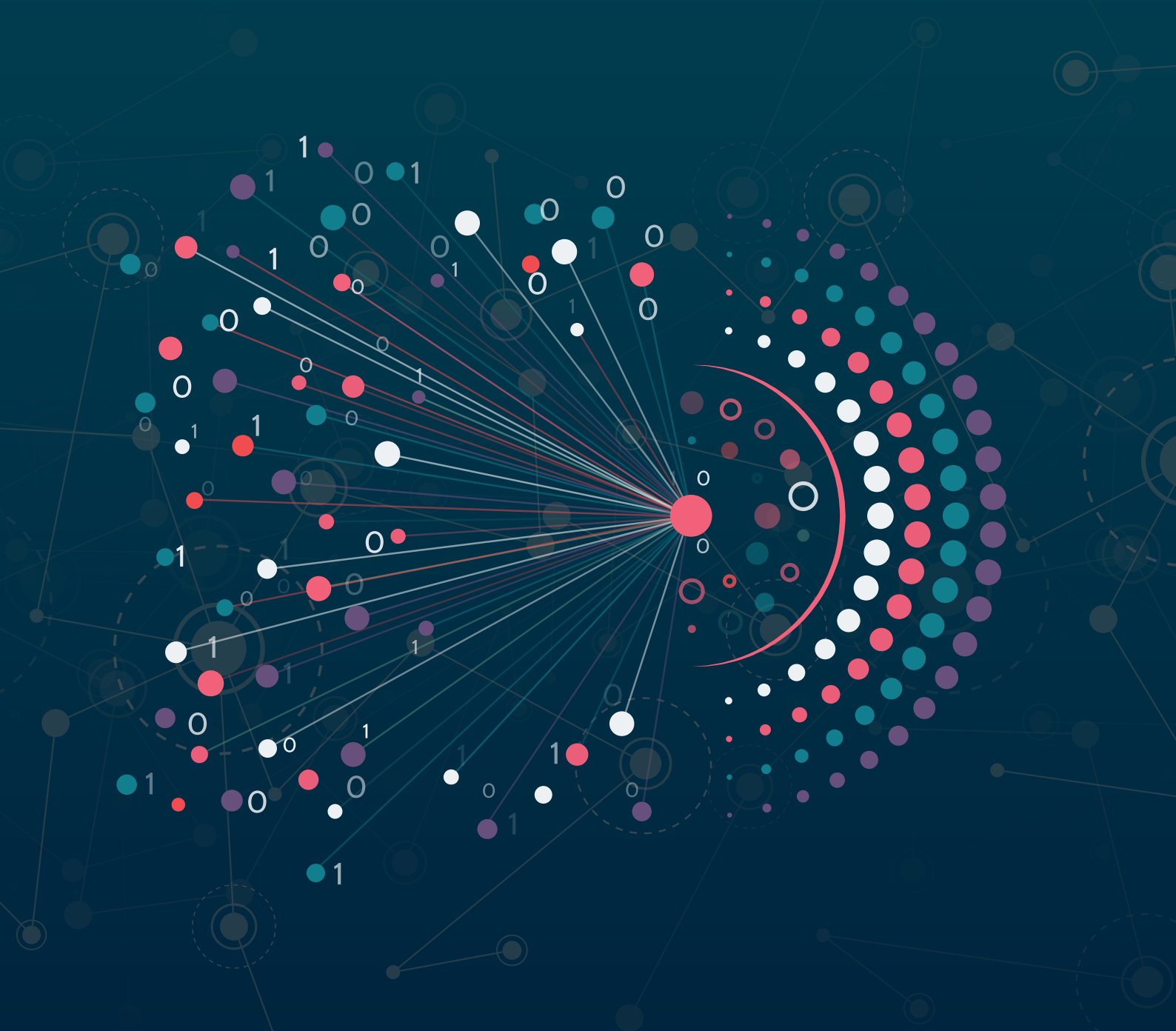
1. 例如，请参见《2024 年人工智能商业预测》，普华永道，<https://www.pwc.com/us/en/tech-effect/ai-analytics/ai-predictions.html>；以及《管理数据成本和复杂性的全新架构》，BCG，2023 年 2 月 7 日，<https://www.bcg.com/publications/2023/new-data-architectures-can-help-manage-data-costs-and-complexity>。
2. OSDU（开源数据宇宙）论坛是一个面向能源行业、基于标准的开源数据平台。

### 插图

插图由 Tim Huxford 提供，素材来自 Shutterstock。

尽管麻省理工科技评论洞察已尽力核实这些信息的准确性，但对于任何人依赖本报告或本报告中的任何信息、意见或结论的行为，本报告概不负责。

© Copyright MIT Technology Review Insights 版权所有，2024 年。保留所有权利。



## 《麻省理工科技评论》洞察

[www.technologyreview.com](http://www.technologyreview.com)

[insights@technologyreview.com](mailto:insights@technologyreview.com)