

## 白皮书

# 使用数据虚拟化和人工智能实现数据 驱动型制造

作者：Mike Ferguson，  
来自 Intelligent Business Strategies  
2024 年 3 月

研究赞助商：

denodo 

## 目录

引言：致力于成为数据驱动型企业 .....	3
隐藏复杂性：对数据编织的需求 .....	5
数据驱动型企业的逻辑数据管理 .....	5
制造业中的主要业务问题 .....	7
制造业中的问题 .....	7
制造业的首要业务优先事项 .....	8
解决制造业问题所需的洞察 .....	10
制造业中所需的机器学习模型和人工智能 .....	11
所需的数据源数量不断增加 .....	13
在制造业中使用数据虚拟化 .....	15
智能制造 .....	15
智能制造洞察 .....	16
数字主线 .....	16
数字孪生 .....	17
工业物联网 .....	17
合规与风险 .....	18
供应链管理 .....	19
建立供应链控制塔，以实现可视性 .....	20
通过供应链指挥中心实现主动优化 .....	20
结论 .....	21

## 引言：致力于成为数据驱动型企业

制造业公司正面临着更高的客户期望、材料短缺、通货膨胀和能源成本上升等问题

买家提出了个性化要求，并希望了解材料的原产地，以及制造过程中的可持续性情况

人们对数据和人工智能带来的益处寄予厚望，发展智能制造已成为当务之急

数据和人工智能正在成为推动智能制造的核心力量

为了缩短实现价值的时间，需要利用业务就绪型数据建立高质量、安全、合规的数据基础

数据复杂性是阻碍成功的主要障碍

如今，大多数制造业公司的高管们都面临着疫情后经济时代的重大挑战。这包括应对客户和潜在客户的更高期望，以及地缘政治局势紧张、材料短缺、通货膨胀和能源成本上升等问题，所有这些问题都会影响其供应链和制造成本。

面对客户和潜在客户，为了赢得业务并维持客户忠诚度，制造商面临着向其提供更多服务的压力。例如，对于 B2C（企业对消费者）制造商而言，允许客户在购买前配置产品已成为常态，并且符合客户的期待。然而，客户如今还希望从制造商那里获取更多信息，以决定是否购买产品。例如，他们希望更清楚地了解材料来源，以及其打算购买的产品在制造过程中是否符合可持续性原则。个性化、公平贸易供应链实践，以及赢得“环保买家”的信任如今已变得至关重要。

此外，智能制造已成为议程中的优先事项，而这需要在机器人技术、物联网、云计算和人工智能等领域进行大量数字化转型投资。在此背景下，高管们如今已认识到，数据和人工智能对于帮助其实现更加智能化的运营至关重要。

然而，他们对数据、商业智能、机器学习和人工智能 (AI) 所能带来的益处抱有极高期望。考虑到围绕人工智能的炒作，这种期望达到了非常高的水平。在他们看来，数据和人工智能可以提供个性化服务、削减成本、提高效率、优化制造流程和供应链、自动执行任务、改善质量控制、降低风险，并帮助他们制定更及时、更准确的决策，从而抓住机会。他们还将数据和人工智能视作颠覆竞争市场的手段，并希望借此在新市场中发掘新产品和新商业模式的机遇。简而言之，数据和人工智能现在已被视为未来智能制造的核心。

对于制造业公司而言，若要从数据和人工智能中获得最大的投资回报率，就必须建立高质量、安全、合规的数据基础，以便在此基础上建立竞争洞察力，并充分利用人工智能。但如果认为这是一件水到渠成的事，那将是严重的疏忽。数据需要满足高质量、合规、“业务就绪”等条件，并且能够在整个企业（甚至外部）共享，以帮助缩短实现价值的时间，并达成制造业高管设定的愿景和战略性业务目标。

然而，成功的道路上存在着障碍。其中一项主要挑战就是数据源越来越多，导致数据复杂度增加。如今，数据可能存在于本地、多种云环境、软件即服务 (SaaS) 应用程序提供商数据中心以及边缘位置的数据存储中。如图 1 所示，现在，许多公司的数据都持久化存储在了真正的混合多云分布式数据环境中。这些数据存储在云端和本地的关系数据库、NoSQL 数据库、云存储、文件系统和内容管理系统中。数据也可能来自边缘设备的流式传输，或通过多种不同的机制（如 API、文件传输、数据市场和数据洁净室）从多个外部数据提供商和供应商处摄取。因此，我们面临着两难的困境：一方面，高管期望数据和人工智能提供

全球多家工厂的物联网数据量不断增长，同时需要遵守不同地区的数据治理法规，这些都增加了数据的复杂性

通往智能制造的道路，带来更高的敏捷性和速度；但另一方面，实现这一目标所需的数据变得越来越难以查找、治理和整合。此外，对于在全球范围内拥有多家制造工厂的全球制造商而言，这更具挑战性（尤其是在边缘位置），因为其物联网部署在世界各地多家工厂的多条生产线上。在这种情况下，除了数据源的数据复杂性外，还涉及数据量、数据规模以及必须遵守不同地区数据治理法规的问题。为了继续运营，这些都是必须应对的问题。

许多公司现已拥有混合多云分布式数据环境，但仍希望快速利用数据和人工智能来提高竞争力

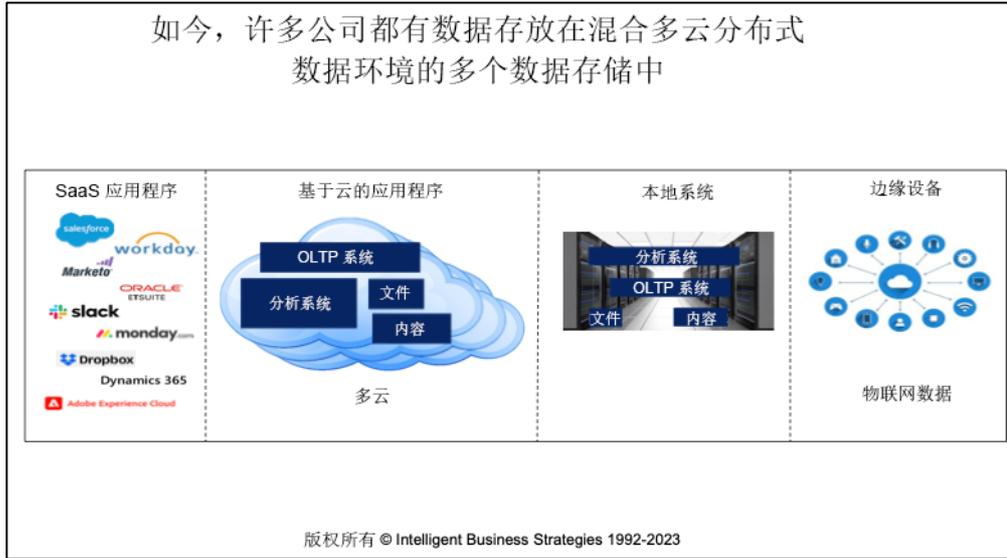


图 1

本白皮书着眼于此，并探讨数据虚拟化如何通过提供一个平台来帮助解决这个问题，从而让制造业公司能够缩短实现价值的时间、满足高管的期望，并实现成为数据驱动型智能制造组织的目标。

## 隐藏复杂性：对数据编织的需求

鉴于许多制造业公司都面临着数据复杂性的问题，需要部署一些关键的技术组件来加速构建数据基础，使公司能够建立洞察力、进行预测、生成建议并实施人工智能自动化，以帮助打造数据驱动型企业。这就产生了对数据编织软件的需求。

数据编织软件可在整个分布式环境中连接数据，使公司能够查找、规划和治理数据

数据编织是一种数据管理平台软件，能够连接混合多云分布式数据环境中的各种数据存储和数据源。它提供全面的功能，可支持自动化的数据和关系发现、数据分类、治理、清理、转换、整合和数据共享。它可以被多名 IT 专业人员和公民数据工程师使用，用来开发经过编排的管道，以生产业务就绪型数据产品，并以合规的方式在整个企业内安全共享。

它通常包括数据目录、对协作开发的支持和数据市场

数据编织软件支持与各种数据源的安全连接、数据目录、协作开发、DataOps、性能和可扩展性、可延展性、统一数据治理和数据市场。

它的作用是隐藏复杂性，简化数据访问，并加速数据工程管道的开发，以生产可复用的数据产品，这些数据产品可用于创建机器学习模型、人工智能机器人和数据仓库，为商业智能分析和报告提供支持。相关内容如图 2 所示。

它隐藏了分布式数据环境的复杂性

多个数据生产者团队可以使用通用的数据编织，构建整合数据的管道，以生产可复用的数据产品

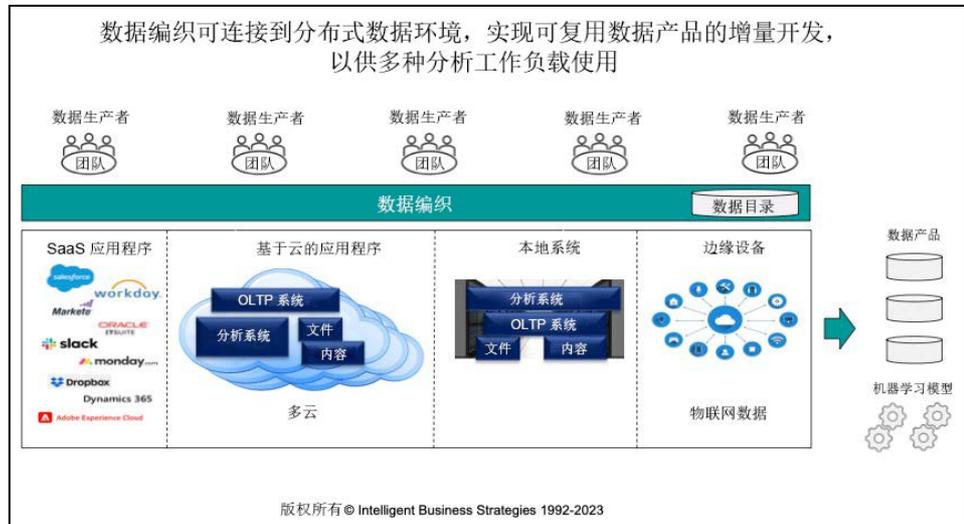


图 2

## 数据驱动型企业的逻辑数据管理

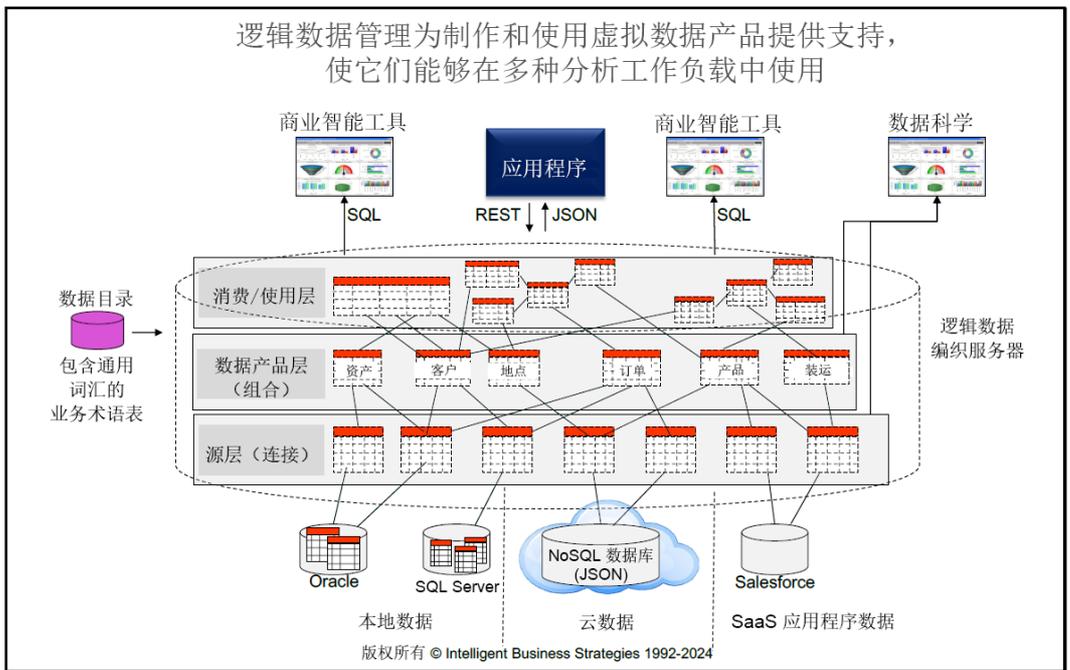
逻辑数据管理软件可用作数据编织平台，基于多个底层数据存储中的数据，创建整合的虚拟数据产品。

使用业务术语表中定义的通用业务数据名称，可以定义数据实体的虚拟视图，从而创建数据产品。平台的数据目录可提供业务术语表功能，以及创建各个数据产品所需的源数据元数据。然后，多个数据源中已识别的数据可以映射到用于描述数据产品的各个虚拟视图的业务术语和模式上。这种方法允许您逐步创建一层语义链接的虚拟数据产品，这些产品可以：

逻辑数据管理支持针对不同的分析应用场景，创建和使用虚拟数据产品

- 快速组装成针对不同分析工作负载（例如数据仓库、数据集市、数据科学）进行优化的不同虚拟模式。
- 通过查询为数据科学中的机器学习模型开发提供特征数据。
- 用作创建其他虚拟数据产品的输入。
- 通过不同的接口（例如 SQL、REST、ODATA、GraphQL）被制造生产线上的其他应用程序使用。

相关内容如图 3 所示。



定义虚拟视图层有助于快速整合来自多个来源的数据，并以易于理解的业务术语形式提供给使用者

图 3

数据产品可以快速组合，以创建其他虚拟数据结构

请注意，这种快速“组装”语义链接的数据产品以创建其他虚拟模式的能力，标志着生产力的重大变化，这种变化提高了敏捷性，缩短了价值实现的时间。它为启动多个分析项目提供了业务就绪型数据，并实现了基于组件的模式开发。此外，它还支持创建通用语义层，以便为所有工具和应用程序提供具有相同通用业务数据名称的数据。

如果需要，也可以持久化存储数据

## 制造业中的主要业务问题

制造业公司可以从使用逻辑数据管理软件中获益

要了解数据虚拟化和人工智能在制造业中的价值，首先必须了解制造业公司当前面临的问题。然后，我们可以了解业务中的优先事项，以及帮助解决这些问题所需的洞察力。在此基础上，我们可以了解所需的数据源，以及如何利用上述逻辑数据管理来整合这些数据源中的数据，从而轻松建立所需的洞察力。

### 制造业中的问题

制造业公司必须应对生产成本上升、劳动力短缺、材料短缺、供应链的复杂性和中断，以及改善质量控制的需求

目前，制造业存在一些问题。其中包括：

- 生产成本上升（原材料、能源、劳动力）
- 劳动力短缺问题
- 材料短缺、供应链的复杂性和中断问题
- 产品生命周期和制造过程中的质量控制
- 数字化转型，例如对物联网、人工智能和自动化技术的采用
- 网络安全

数据和人工智能可用于帮助降低成本、提高效率、减少浪费并提高自动化程度

和其他所有行业一样，近年来，无论是材料成本、能源成本还是劳动力成本，都在不断攀升。制造商通常会将这些成本以价格上涨的形式转嫁给客户，但他们也可以采取主动措施，利用数据和人工智能来提高效率、减少浪费并提升自动化程度，以此来降低成本。自动化也有助于缓解劳动力短缺问题。除此之外，还可以采取更多措施，最大限度地降低因设备故障、产品错误、装运错误等原因造成的意外运营成本。

数据和人工智能还可用于帮助监控和优化供应链与生产流程

供应链的中断和复杂性问题也持续困扰着制造业公司。提升供应商的多元化程度将有所帮助，但在帮助制造商了解需求、供应商表现和供应链运营情况方面，数据和人工智能将同样变得至关重要，其将使制造商能够进行预测、采取行动，并避免中断。数据和人工智能可用于监控和优化供应链，降低原材料成本，并改善材料规划。根据销售数据对所需材料做出预测，有助于优化材料库存、缩短交货时间，以及避免特定制造工厂为克服材料短缺问题而进行材料转移。类似地，数据和人工智能也可用于监控能源成本、生产吞吐量，并最大限度地减少停机时间。

通过建立数字主线，为产品的设计、制造和分销所涉及的所有系统提供一致的产品和材料数据，从而提高产品质量

产品质量也是生产中的一个问题。造成此类问题的一个原因是，需要将材料清单和产品数据提供给产品端到端生产流程中的所有相关系统。然而，通常做法是，随着生产流程的推进，手动将这些数据重复录入到不同系统中。这可能会导致数据录入错误，进而造成产品质量问题。帮助解决此问题的一种方法是，建立从产品设计到产品制造、分销和使用的数字主线，从而在整个流程中按需共享和提供整合的数据。

## 制造业的首要业务优先事项

制造业公司正在拥抱智能制造，提供产品定制、接单生产、物联网实时数据，和人工智能自动化服务

从前端部门到制造，再到后端部门和供应链，整个流程都需要整合

为了克服这些问题，制造业公司正在进行重大转型，以拥抱并采用智能制造实践。为此，需要端到端地整合整个业务流程，有能力为客户提供定制产品的服务，以满足其个性化需求，并建立接单生产的能力。此外，如果要想实现智能制造的愿景，就需要更多数据。因此，理所当然地，许多制造业公司都在投资物联网和人工智能，以便捕获数据，这将使他们能够监控和优化制造流程，提高产品质量，并在供应链中建立弹性。但事情并没有这么简单。制造业公司需要连接前端流程、制造流程、后端流程以及供应链，从而端到端地整合整个流程。公司还需要在整个运营过程中提供整合的数据、商业智能、机器学习和人工智能自动化，才能真正拥抱智能制造。相关内容如图 1 所示。

要真正拥抱智能制造，就需要在整个运营过程中整合数据，以实现实时可视性，以及商业智能、机器学习和人工智能

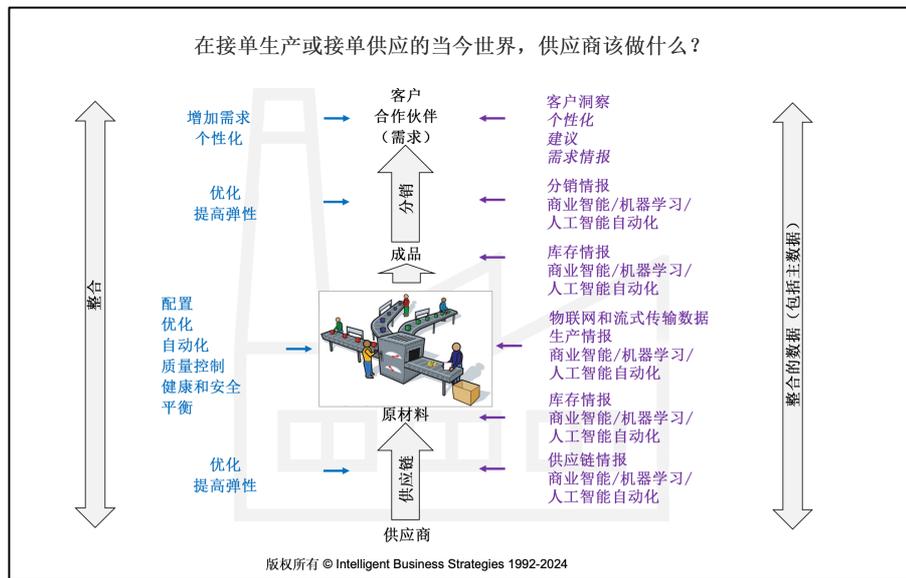


图 4

鉴于这种要求，制造业公司目前正优先重视以下领域：

优先领域	理由
产品定制	<ul style="list-style-type: none"> <li>“动态客户档案”</li> <li>更具竞争力</li> <li>接单生产</li> <li>先购买，再生产</li> <li>更优秀的客户服务</li> </ul>
降低成本	<ul style="list-style-type: none"> <li>提高盈利能力和竞争优势</li> </ul>
提高端到端效率	<ul style="list-style-type: none"> <li>以更少的人员数量实现更高的吞吐量和可扩展性</li> </ul>
智能工厂	<ul style="list-style-type: none"> <li>更准确的需求预测</li> <li>提高自动化和设备性能</li> <li>人工智能检测的预防性维护</li> <li>人工智能辅助的质量控制，将产品缺陷率降至最低</li> </ul>

改善客户体验、建设智能工厂、提高效率、加强质量控制、打造弹性供应链、确保合规性、应对网络安全问题以及实现可持续性，这些都是制造业中的优先事项

	<ul style="list-style-type: none"> <li>● 数字孪生</li> <li>● 劳动生产率跟踪</li> <li>● 在从设计到生产、销售、使用和处置/回收的整个过程中，通过数字主线跟踪产品的性能和使用数据</li> </ul>
改善产品质量	<ul style="list-style-type: none"> <li>● 无需重新制造产品，降低运营成本</li> <li>● 减少报废和浪费</li> <li>● 提高客户满意度</li> <li>● 增强对品牌的信心</li> </ul>
减少浪费	<ul style="list-style-type: none"> <li>● 降低成本，减少能耗</li> </ul>
供应链弹性和优化	<ul style="list-style-type: none"> <li>● 加强材料规划</li> <li>● 改善库存管理</li> <li>● 防止供应链和生产中断</li> <li>● 避免因延误而造成客户不满</li> </ul>
定制/个性化	<ul style="list-style-type: none"> <li>● 提供更个性化的服务，让客户能够根据自己的具体喜好定制产品</li> <li>● 接单生产模式可以改善现金流</li> </ul>
监管合规性	<ul style="list-style-type: none"> <li>● 健康和安全</li> <li>● 产品溯源</li> <li>● 客户信心</li> </ul>
可持续性	<ul style="list-style-type: none"> <li>● 利用可再生能源</li> <li>● 优化资源消耗</li> <li>● 回收和合规性</li> <li>● 吸引环保买家</li> </ul>

智能工厂需要自动化、物联网、云计算和人工智能来降低成本并提高效率

需要关于供应商和供应链的高质量、整合化、低延迟的数据和情报，以便更快地将产品推向市场

还需要实时监控事件，以进行检测和预测，给出建议并及时采取行动，以避免中断、防止浪费、最大限度地提高产量、优化库存并改进质量控制

考虑到这些优先事项，智能制造的概念就是建立智能工厂，通过使用自动化、物联网 (IoT)、云计算和人工智能 (AI) 技术，降低成本，并改善效率和可持续性。区块链等其他技术也有助于提供整个供应链的可追溯性和实时可视性。

在供应链方面，制造商希望以一致、快速的方式引入供应商，并更好地对供应商进行管理，以便更快地将产品推向市场。后者需要高质量、丰富、经过整合的供应商数据和详细的供应商情报。此外，整合的数据还可以提高供应链的透明度。制造商还希望使用人工智能和低延迟数据来提高供应链的敏捷性、可持续性和弹性。

这一切都需要能够访问和共享整个供应链的数据。这还需要对事件进行实时监控，以检测和预测问题，并在这些事件和问题对业务造成重大影响之前及时采取行动，从而避免中断、防止浪费、完成工单，以能够按时交货。共享经过整合的数据还有助于对产品材料和产品来源进行认证，从而提供产品溯源和合规信息。这还能让制造商优化库存，并向客户展示其对产品质量的控制。此外，在发生产品召回的情况下，还能够通过供应链中的详细数据，以更快的速度和更低的成本解决争议。

## 解决制造业问题所需的洞察

为了建立更丰富、更全面的洞察，从而为更好的决策提供支持，帮助解决问题，经过整合的数据必不可少

为解决这些问题，并满足业务中的首要需求，制造业公司需要整合和利用数据，以生成更丰富、更全面的洞察，从而为上述各个高优先级领域的决策提供支持。这意味着需要整合一系列不同内部和外部数据源中的数据，以建立与下表所示领域相关的洞察。

情报类型	所需的洞察
市场情报	<ul style="list-style-type: none"> <li>• 单一客户视图</li> <li>• 客户细分</li> <li>• 个性化</li> <li>• 营销自动化</li> </ul>
需求规划和预测	<ul style="list-style-type: none"> <li>• 销售预测</li> </ul>
生产效率情报	<ul style="list-style-type: none"> <li>• 生产计划</li> <li>• 设备性能水平</li> <li>• 生产速率</li> <li>• 工单完成情况</li> <li>• 瓶颈</li> <li>• 制造事件（例如设备故障）</li> </ul>
生产成本情报	<ul style="list-style-type: none"> <li>• 制造过程中的材料成本、能源消耗水平和成本、劳动力成本</li> </ul>
生产质量情报	<ul style="list-style-type: none"> <li>• 废品率、缺陷率、模式、趋势、根本原因分析</li> </ul>
库存水平和预测性库存管理	<ul style="list-style-type: none"> <li>• 原材料</li> <li>• 成品</li> </ul>
供应商情报	<ul style="list-style-type: none"> <li>• 供应商位置</li> <li>• 供应商关系</li> <li>• 提供的材料</li> <li>• 供应商交货表现</li> <li>• 供应商交货时间</li> <li>• 供应商风险与合规性</li> <li>• 供应商可持续性</li> <li>• 供应商总支出（旨在更好地针对折扣、价格和支付条款进行谈判）</li> </ul>
供应链情报	<ul style="list-style-type: none"> <li>• 各工厂的材料库存</li> <li>• 交货短缺</li> <li>• 内部事件（例如，新订单、订单变更、订单取消）</li> <li>• 外部事件（例如，罢工、天气、拥堵）</li> </ul>
可持续性	<ul style="list-style-type: none"> <li>• ESG 报告</li> </ul>

对市场、需求预测、生产效率、生产成本、产品质量、库存水平、供应商、供应链和可持续性的洞察必不可少

## 制造业中所需的机器学习模型和人工智能

机器学习和人工智能自动化也是必需的  
两者都需要高质量且及时的整合数据

除了从商业智能 (BI) 中获得洞察外，制造商还需要利用机器学习和人工智能来帮助其满足业务的首要需求。

和商业智能一样，机器学习和人工智能也需要经过整合的数据，用来构建必要的预测性和规范性模型，以检测和预测事件、识别问题和机会，并提出建议和/或采取自动化措施，以帮助优化供应链和制造运营、降低风险、改善库存水平等。

制造业的许多领域都需要利用机器学习 (ML) 和人工智能。所需的模型类型示例如下。

模型类型	业务用途
预测性和预防性维护模型	避免因设备故障造成的意外停机
预测性库存模型	避免短缺，并确保按时向客户提供高质量产品
质量控制图像识别模型	控制产品质量，检测生产过程中的质量问题，确保采取纠正措施并避免延误
规范性设备优化模型	确保设备性能达到最佳运行状态，避免延误并最大限度地提高产量
用于降低健康和安全风险的风险模型	预测和检测风险问题，并采取行动，以确保工作场所的健康和安全
用于动态规划、预测和调度的规范性模型	不断优化端到端流程，以满足需求，并使业务能够得到扩展
用于优化供应链的外部事件预警模型	监控和预测外部事件，并采取行动，以避免供应链和制造流程发生中断
网络安全检测模型	检测和预防网络安全事件造成的意外停机和业务中断

要预防设备故障、预测库存、改进质量控制、最大限度地提高产量、改善健康与安全、优化供应链和实施网络安全措施，机器学习技术必不可少

预测性和预防性维护模型可基于资产数据和物联网数据（如生产过程中使用的机器的温度、压力和振动读数）进行训练。

质量控制图像识别模型可从部署在生产线特定阶段的摄像头中获取实时数据，并充当废品顾问，向机器操作员推荐设备参数，以减少浪费。

除了所示的传统机器学习示例之外，在制造业中也可以使用生成式人工智能。尽管仍处于早期阶段，但生成式人工智能已可用于：

- 在产品的设计过程中生成设计备选方案
- 生成代码，对机器进行编程，从而帮助节省引入自动化技术所需的时间和精力

生成式人工智能可在设计、设备编程、预防性维护和材料处理工作中提高效率

- 生成文本或图像，为预防性维护工程师提供逐步指导，以缩短维护时间，确保正确的维护规程得到遵守，并充分利用设备。这可能包括所需的任何替换部件的清单
- 将设备操作员的自然语言提示转化为一系列操作，以便在机器人等设备上执行材料搬运等任务
- 更快地生成产品信息和营销内容，帮助更迅速地将产品推向市场

传统的机器学习和生成式人工智能对于实现智能工厂和智能制造至关重要。所有这些模型的开发和后续再训练都需要采集、整理、转换和整合多个数据源中的数据，这意味着数据工程对于成功至关重要。这一点不仅适用于初始模型开发，也适用于模型的再训练，可确保随着时间的推移，模型仍可保持准确性。

## 所需的数据源数量不断增加

目前，为了提供分析所需的数据，制造业公司需要的数据源数量正在迅速增加

商业智能、机器学习和人工智能在制造业中的应用，从根本上说，依赖于捕获和整合供应链及端到端制造运营流程中多个数据源的数据。

需要访问数据，才能提供：

- 供应商、材料、产品、资产和客户的全方位视图
- 用于数据仓库和数据集市的数据，使业务分析师能够生成前面讨论过的各类商业智能
- 数据科学中用于开发模型和构建生成式人工智能机器人的数据
- 实时事件流处理和流分析，用于监控供应链运作、生产线运营和客户事件（如更改订单、取消订单），以及产品售出后的性能和使用情况

为了提供数据以支持这些分析，所需的数据源包括：

- 客户关系管理系统数据
- ERP 系统数据
  - 例如，生产计划、材料需求计划、车间活动、工单、生产日程安排和库存数据
- 采购/供应链管理系统数据
- 可编程逻辑控制器 (PLC)
- 制造执行系统 (MES)
- 监控控制和数据采集 (SCADA) 系统
- 质量控制系统
- 视觉系统
- 工业物联网数据（从生产线设备流式传输），通常采用 MQTT 格式
- 能源监控系统
- 资产管理系统
- 产品生命周期管理系统
- 库存管理系统
- 配送系统（用于运输）
- 天气数据
- 外部供应商数据
- 主数据（客户、产品、供应商、资产等）

为分析工作负载提供支持的数据来自于各种内部和外部来源，以及分布在混合多云分布式数据环境中的设备

还要注意的，在许多情况下，越来越多正在使用的交易应用程序都是 SaaS 应用程序。因此，所需的数据可能分散在多个 SaaS 应用程序、本地数据存储、多个不同云中的不同数据存储以及来自物联网设备和点击流数据的大量实时流数据中。这与图 1 所示的分布式数据环境一样。

需要对数据进行整合，以进行分析，并且还需要提供有关客户、产品、供应商、材料、工厂和资产的全方位视图

需要一个包含通用业务数据名称的业务术语表，以共享数据并在整个供应链中获得可视性

还需要整合实时物联网数据与资产、产品、材料及供应商的数据，以持续监控供应链、制造和分销过程

这些系统的数据需要得到整合，不仅是用于分析，还是为了创建客户、产品、供应商、材料、工厂和资产主数据的全方位视图，以推动实现跨系统的一致性。其他需要的关键数据实体包括客户订单、工单、原材料和成品库存数据以及装运数据。然而，这不仅仅是数据整合的问题。为了建立对数据的信心，还需要一个包含通用业务数据名称（通用业务词汇表）、数据质量和数据沿袭信息的业务术语表。

包含通用业务数据名称的业务术语表对于在整个供应链中获得可视性至关重要，因为数据由多个组织提供，相同的数据有许多不同的物理数据名称。通用业务数据名称有助于确保每个人都能清楚了解所共享数据的含义。如果制造商要建立数字主线，以便在端到端设计和制造流程中跨系统共享数据，情况也是如此。

此外，在物联网数据方面，可能需要实时访问这些数据，并将其与其他数据（例如主数据）整合，以实现实时生产和供应链运营的监控。图 5 显示了实时物联网数据处理与传统数据仓库数据整合之间的对比。在智能制造中，两者都是必需的。

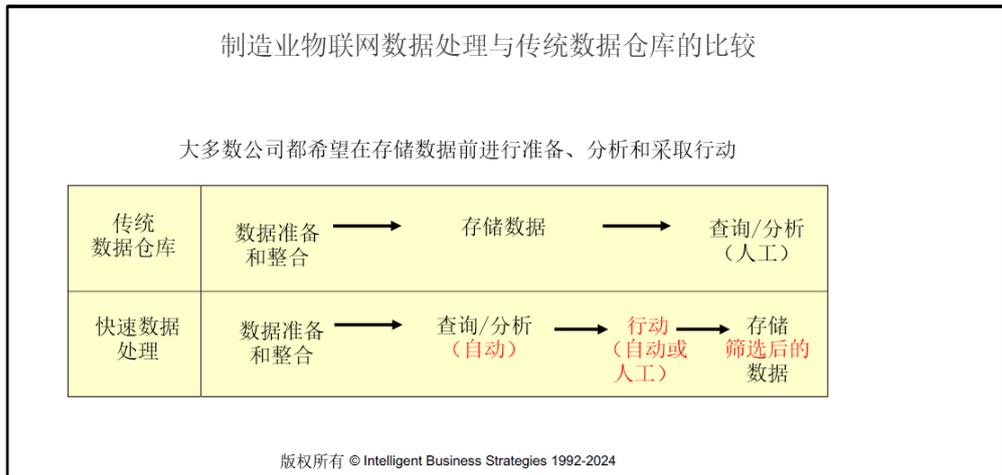


图 5

## 在制造业中使用数据虚拟化

要克服的主要挑战是需要具备快速设计和整合数据的能力

考虑到所需的洞察力，以及制造业公司需要从越来越多的数据源中获取数据，建立快速规划和整合数据的能力成为了必须解决的一项最重大挑战。

在每一个领域，无论是客户全方位实时档案、供应链、制造流程、库存管理、分销，还是风险管理和合规性，都需要按需对数据进行整合。

具有数据虚拟化功能的数据编织软件可实现按需整合数据

数据编织软件可以连接到广泛的所需数据源，以整合特定类型分析所需的数据。在这里，逻辑数据管理的数据虚拟化功能发挥着关键作用。这些功能有助于整合所需的数据，为上述所有制造业应用场景提供洞察。有三个关键示例。

### 智能制造

智能工厂的目标是防止意外的设备故障，实现任务自动化，降低成本，提高产品质量，避免错误以及优化制造流程，从而提高效率和降低成本。建立智能工厂需要商业智能、机器学习和人工智能技术。

这意味着需要捕获、整理和整合数据，从而实现：

- 在制造流程执行和完成工单时，所有与制造流程相关的系统都能按需获取一致且经过整合的数据（例如通过数字主线方式）
- 建立历史数据，并使其可用于分析，以产生与整个端到端制造流程相关的情报（如图 4 所示）。这包括创建关于生产效率、生产成本和生产质量的情报
- 使数据可用于构建机器学习模型和人工智能机器人
- 使数据可用于部署机器学习模型和人工智能机器人
- 实时监控现场运营，包括监控设备性能、效率、成本和质量控制
- 使数据可用于构建数字孪生，以供模拟和分析使用

需要对数据进行整合，以便在整个制造流程中按需提供一致的数据，并为多种分析工作负载提供支持

图 6 展示了如何利用逻辑数据管理的数据虚拟化功能整合不同数据源的数据，以帮助实现智能制造。这可以使用图 3 中的分层方法来完成，该方法包括创建数据产品，或以逻辑数据仓库的形式整合多个分析系统，也可以将两者相结合。

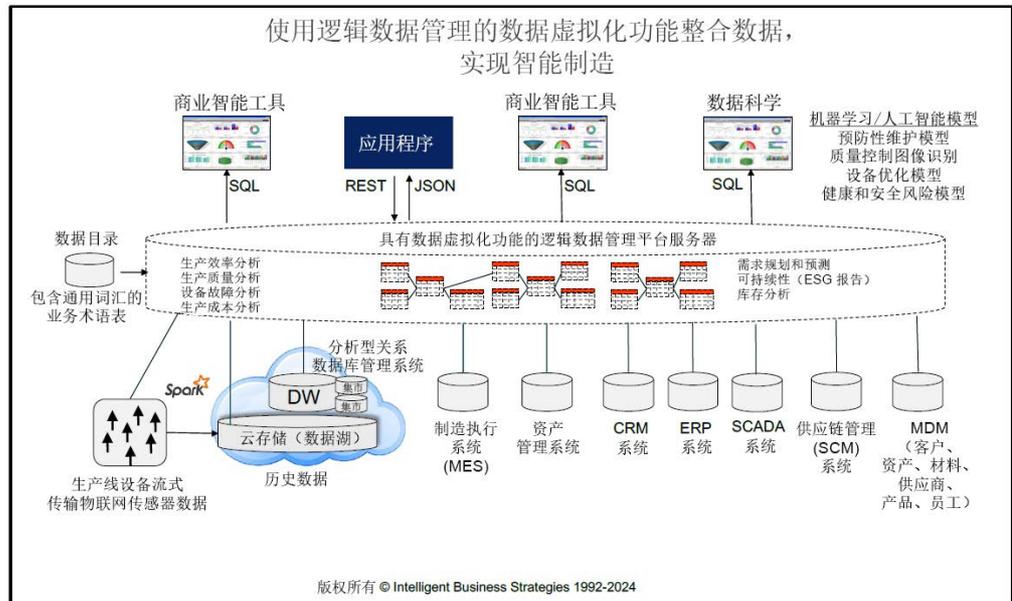


图 6

可以按需访问有关销售、工单、材料交付、库存、机器及其在制造流程中的表现的数据，以及有关成本、效率和产量的历史数据

在这里，您可以了解如何访问以下方面的数据：

- 客户需求及订单
- 要制造的产品、工单、材料交付情况、库存和使用情况、运输情况，这些信息都来自 ERP 系统
- 制造资产，例如设备、零件、生产线等
- 产品制造方式，例如来自 MES 系统的运营调度、生产活动和资源分配数据
- 工业设备性能（流式传输物联网数据）
- SCADA 系统中的受控机械日志
- 生产绩效和生产质量的历史数据（数据仓库/数据湖屋）
- 客户、产品、供应商、材料、资产等

主数据有助于提供上下文

可能已经存在现有的分析系统，例如数据仓库

## 智能制造洞察

来自运营和分析数据源的数据可以被整合，并创建虚拟视图，以提供全面和近乎实时的视图，从而支持各种不同的分析（如分析关于营销、生产效率、生产质量、库存和运输的端到端情报）。

多个虚拟视图可为端到端商业智能、机器学习和人工智能提供数据

## 数字主线

另一项关键要求是提高数据质量，以防止在生产中产生意外错误、延误和成本。这包括从产品设计到生产的整个流程中，在多个系统中获取一致的产品数据。这个过程被称为数字主线。在生产过程中，需要将这些数据的不同整合视图实时提供给生产流程中不同阶段的系统和人员进行查询。

数据可以被整合，并在数字主线中共享，以防止在生产中产生意外错误、延误和成本

需要在数字线程中共享的数据可能包括：

- 产品数据（包括材料清单）

使用通用业务词汇表，  
可在数字主线中提供一  
致的整合数据虚拟视图

- 使用的材料和零件/组件的来源
- 工程变更、产品变体、替代材料等
- 工单、机器操作员指令数据以及与工单或生产批次相关的物联网数据
- 流程阶段测试和检查数据，如合格/不合格结果、纠正措施等

需要注意的是，数字主线要想取得成功，就必须创建通用的业务词汇表，因为许多不同的数据源会对相同的数据使用不同的物理数据名称。如果没有通用的业务数据名称，几乎不可能为所有需要数据的系统提供一致的数据。此外，通过整合的虚拟视图提供对这些数据的访问能力，还可以实时监控需求、工单、生产计划和库存水平，以提供生产流程的端到端运营报告。

这也有利于建立实时可  
视性和生成运营报告

## 数字孪生

随着捕获的数据越来越多，这些数据可以与物联网数据整合，并用于构建数字孪生，以帮助在生产流程中分析和优化各种复杂设备的性能，并优化整个制造流程。可以创建不同类型的数字孪生，其益处显著，包括：

还可以提供数据来创建  
不同类型的数字孪生，  
以帮助及早发现设计缺  
陷、减少浪费、改善产  
量和可持续性，并缩短  
产品上市时间

- 缩短上市时间
- 及早发现产品设计缺陷
- 减少废料浪费
- 提高制造产出率
- 改善可持续性

## 工业物联网

工业物联网传感器和设备的使用是所有智能工厂的核心所在，它们可嵌入机器内部并部署到整个工厂。这些传感器和设备可持续发射湿度、温度、压力、振动等实时数据，以持续监控设备状态、设备性能、环境条件（如能源使用）、健康和安全风险等。原始物联网数据是必不可少的，可确保设备以最大产能在最佳状态运行，以避免延误并最大限度地提高吞吐量，同时确保将成本降至最低，并使员工健康和​​安全不受损害。产品质量也可以得到监控。

可以访问物联网数据，  
并将其与资产数据相结  
合，以提供制造运营的  
实时可视性，供自动分  
析和情境感知使用

大部分此类工作都需要在数据源地附近进行，仅仅是因为可生成的原始物联网数据量巨大。然而，物联网数据通常需要在运行时与资产和材料数据整合，以提供上下文，从而实现实时预测和规范性分析。此外，使用通用业务数据名称提供原始物联网信息数据，以及使用不会降低物联网读数精度的数据类型也很重要。

数据虚拟化可以连接到 **Kafka** 等物联网流式传输数据源以及静态数据，以使用通用业务数据名称提供实时制造数据的整合虚拟视图。这样就可以创建情境感知仪表板和机器学习模型，以用于自动分析，从而：

这可用于预测设备故障、瓶颈、健康和安全风险以及质量问题，以便及时采取行动，从而避免出现问题，保护员工，并确保一切得到优化

- 预测设备故障并安排预防性维护，以避免出现意外停机
- 检测瓶颈并提出调整建议，以优化设备性能
- 检测生产过程中的质量问题，确保采取纠正措施，避免报废和延误
- 预测和检测风险，并采取行动，确保工作场所的健康和安全

## 合规与风险

在全球经济中，许多制造商不仅在许多不同的国家/地区销售产品，还可能在许多不同的国家/地区拥有制造这些产品的工厂。因此，全球制造商必须遵守越来越多特定于国家/地区的法规和法律，这涉及广泛的义务。这包括与以下方面相关的法规和法律：

制造商必须遵守越来越多的法规和法律，这涉及广泛的义务

- 产品安全
- 健康和安​​全
- 环境影响
- 数据处理和保护
- IT 安全和安保
- 公平贸易与竞争
- 出口管制
- 反腐败法
- 雇佣法

药品、医疗器械、太阳能电池板等制造领域的法规尤为严格。这其中包括环境安全方面的法规。例如，在欧盟，欧洲职业安全与健康署多年来一直在实施《欧洲机械指令》，该指令适用于以下产品的制造：机械、可互换设备、安全组件、起重配件、链条、绳索和织带、可拆卸机械传动装置，以及打算安装到其他机械或其他机械组装在一起的半成品机械（如驱动系统）。

跨多个法律辖区的监管负担沉重，需要来自许多不同数据源的数据

不言而喻，应对这种跨多个法律辖区的监管负担是一项重大挑战，因为它需要整合许多不同数据源的数据（这些数据源在很多情况下位于不同的地理区域），以便进行监管报告和审计。可追溯性和透明度也至关重要。

数据虚拟化可提供合规报告和风险监控所需数据的整合虚拟视图

为解决这一问题，可以使用数据虚拟化创建实时数据的整合虚拟视图，以提供合规报告和风险监控所需的数据。可以创建不同的虚拟视图，以便在不同的法律辖区进行报告，无需复制数据。这就向业务用户隐藏了数据整合的复杂性，使其能够快速、及时地提供合规报告。

## 供应链管理

制造商希望在整个供应链中获得实时可视性，并具备抵御中断的能力

多元化采购非常重要

创建通用业务数据名称对于实现数据共享至关重要

数据虚拟化可用于创建整合的数据虚拟视图，以帮助提供有关供应商的情报以及对供应链事件和运营的实时可视性

供应链的实时虚拟视图可按需提供整合数据，以用于情境感知仪表盘和使用人工智能的自动化监控

这有助于避免中断，并使供应链保持在最优状态

在供应链方面，制造商面临的挑战是如何实时了解整个供应链的情况，并在其中建立弹性。为此，制造商需要通过以下方式打造适应未来的供应链：

- 投资区块链等先进技术，提高透明度和可追溯性
- 实现采购多样化，并投资本地生产，为每家制造工厂设立多个提供相同材料的供应商
- 创建通用词汇表，在其中收录经过商定的业务友好型数据名称，以便在整个供应链中理解和共享数据
- 连接多个内部和外部数据源，为整个供应链中的实时数据创建通用业务词汇表虚拟视图（无需复制或合并数据），以实现供应链的实时可视性
- 创建整合的虚拟数据产品，以便构建一套全面的：
  - 最新供应商情报，包括供应商的位置和与制造工厂的距离、所提供的材料、备选供应商、供应商交货表现、供应商总支出、供应商交货时间、供应商风险和合规性等信息
  - 实时供应链情报，包括各工厂的材料库存、交货和交货短缺情况、客户订单和订单变更、交货车辆位置、外部事件等的实时视图
- 访问实时虚拟视图，以创建始终最新的运营情境感知报告和仪表盘，让每家制造工厂都能查看其供应链情况
- 通过构建并使用机器学习模型，实时自动监控供应链，从而：
  - 自动检测和预测中断
  - 将预测与供应商情报及实时供应链情报结合使用，触发自动警报、建议和行动，以防止中断并保持供应链处于最优状态
  - 自动采取行动，避免外部事件干扰供应链和制造流程

图 7 展示了如何利用逻辑数据管理的数据虚拟化功能，将供应链中不同数据源的数据与内部数据仓库整合在一起。这有助于建立供应链的实时可视性，并提供在整个供应链中进行监控和及时决策所需的数据，从而将供应链保持在最优状态。这包括提供实时情境感知报告。

供应链的实时可视性有助于及早发现或预测需求变化和问题，并采取行动，避免中断

还可以监控供应商的表现和支出

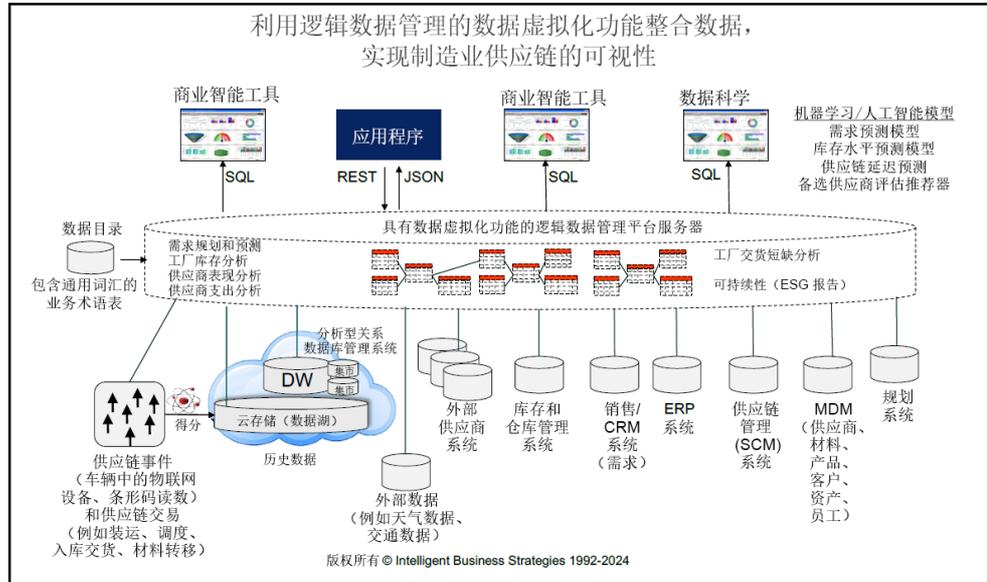


图 7

### 建立供应链控制塔，以实现可视性

可以建立供应链控制塔，提供不同的个性化供应链视图，实现端到端情境感知

通过整合来自内部和外部数据源的数据，还可以建立供应链控制塔，其中包含一系列相互关联、面向供应链中不同相关方的个性化仪表盘。这些相关方包括您的供应商、您的组织（即制造商）、您的分销商以及客户。这些不同的个性化视图可以显示整个供应链的实时指标和事件，以实现端到端的情境感知。这也可以与您自己的制造工厂数字孪生、客户网络以及实时 GPS 数据相结合，以显示不同制造工厂和客户位置的状态。

### 通过供应链指挥中心实现主动优化

供应链指挥中心不仅包括控制塔，还包括事件监控、警报、建议、任务、自动化、任务分配和协调等功能，以确保供应链处于最优状态

使用逻辑数据管理软件还可以创建供应链指挥中心。它超越了供应链控制塔的功能范围，提供了数据整合、供应链可视性、工作协调、建议、操作员任务的自动化分配，以及任务自动化等功能，可实现弹性和主动的端到端供应链优化。

例如，通过规范性机器学习模型访问整合的数据，可以在整个供应链中提供实时警报和建议，自动分配任务并触发自动化，以帮助所有相关方保持供应链的优化。其结果是实现整合的感知和响应能力。

还可创建虚拟数据视图，为客户的全方位营销、销售和服务、网络安全等提供支持

除了这些示例之外，通过使用逻辑数据管理来发现和整合数据，还可以在制造业中进行许多其他类型的分析。这包括利用客户全方位数据平台中的数据进行面向市场营销、销售和客户的分析，以及网络安全分析。

## 结论

*如果制造业公司希望成为由数据和人工智能驱动的企业，就必须能够及时整合多个数据源中的数据*

*逻辑数据管理隐藏了分布式数据环境的复杂性，提高了敏捷性，并且能够加快数据集成的速度*

*多个数据生产者团队可以创建可重复使用的虚拟数据视图，从而缩短实现价值的时间*

如果制造业公司想要成为由数据和人工智能驱动的企业，其未来发展在很大程度上取决于能否及时地从越来越多的数据源中整合数据。现在，要构建高质量、安全、合规和丰富的数据基础，以帮助产生新的洞察、建议和预测，这一要求已变得至关重要。这一基础将使他们能够建立所需的洞察，提供对供应链、端到端制造流程、数字主线、数字孪生和实时生产线运营监控的可视性。

主要的挑战在于克服不断增加的数据源和数据存储所造成的复杂性，这些数据源和数据存储分散在真正的混合多云分布式数据环境中。通过部署逻辑数据管理软件，既能够连接这些数据源，又可以构建虚拟数据产品，同时允许源数据保留在原处，这将有助于提高敏捷性，克服数据工程中的瓶颈，并及时逐步提供所需的洞察，从而改善业务成果。

逻辑数据管理提供了这些功能，因此是制造业公司在为智能制造和数据驱动型企业选择数据管理技术时应该考虑的技术之一。

## 关于 Intelligent Business Strategies

Intelligent Business Strategies 是一家独立的研究、教育和咨询公司，其目标是帮助企业了解和利用商业智能、机器学习、高级分析、数据管理、大数据和企业业务集成等领域的新进展。这些技术共同帮助企业转变成为 *智能企业*。

### 作者



Mike Ferguson 是 Intelligent Business Strategies Limited 的管理总监。作为独立的 IT 行业分析师和顾问，他专攻 BI/分析和数据管理。凭借 40 多年的 IT 经验，Mike 为数十家企业提供了有关商业智能/分析、数据策略、技术选择、企业架构和数据管理方面的咨询服务。Mike 也是欧洲最大的数据和分析会议 Big Data LDN 的会议主席，以及 EDM Council CDMC 执行顾问委员会的成员。他撰写过大量文章，也在世界各地的各类活动中发表过演讲。此前，他是关系模型的发明者 Codd and Date Europe Limited 的总裁和联合创始人，还担任过 Teradata 公司的 Teradata DBMS 首席架构师以及 Database Associates 的欧洲管理总监。他教授一系列热门大师课程，包括：数据战略、数据架构现代化、分布式数据环境治理方法、实施数据网络的实用指南、数据目录和嵌入式分析、智能应用程序和人工智能自动化。



电话: (+44)1625 520700

网址: [www.intelligentbusiness.biz](http://www.intelligentbusiness.biz)

电子邮件: [info@intelligentbusiness.biz](mailto:info@intelligentbusiness.biz)

使用数据虚拟化和人工智能实现数据驱动型制造  
版权所有 © 2024, Intelligent Business Strategies  
保留所有权利

本白皮书使用的所有屏幕截图和图表均来自 Denodo，  
版权和知识产权仍归 Denodo 所有



Denodo 公众号



数据编织咨询群