

# IBM Cloud Pak for Data

智能自动化您的数据和 AI 策略,无论何处都可以在合适的时间将正确的数据连接到正确的人员。

## 目录

- IBM Cloud Pak for Data: 任何数据,任何云,任何地方
- 新一代的 IBM Cloud Pak for Data
- IBM Cloud Pak for Data 的部署模型
- IBM Cloud Pak for Data 的顶级用例
- 后续步骤

## 引言

当今环境多变难料,各个企业必须更加智慧,响应速度更快,智能运营并与世界交流,弹性响应市场变化,灵活优化成本,进行创新。在数据的推动下,AI 正在让龙头企业实现转型,传递价值。[最近的一项研究](#)发现,数据驱动型企业在收入和盈利能力方面更胜一筹的可能性要高出 178%。

然而,要想在业务中成功扩展 AI,必须克服数据复杂性。如今,企业正在努力管理和维护跨越公有、私有和本地云的海量数据。70% 的全球受访者表示,他们的公司正在从 20 多个不同的数据源中拉取数据,为其 AI、BI 和分析系统提供信息。此外,三分之一的人认为数据复杂性和数据孤岛是应用 AI 最大障碍。数据的寿命,即数据最相关并最有价值的时间,正在急剧缩短,这一事实让混合数据图景更加复杂。<sup>1</sup>

解决方案?灵活、有弹性的云原生平台支持客户通过可信的数据和 AI 来预测和自动处理结果。

# IBM Cloud Pak for Data: 任何数据,任何云,任何地方

IBM Cloud Pak® for Data 是一个完全集成的数据和 AI 平台,让企业能够通过释放生产力和降低复杂性来加速 AI 驱动的转型。收集、组织和分析数据;然后在协作的平台体验中将 AI 注入整个业务。IBM Cloud Pak for Data 设计为云原生,建立于 Red Hat® OpenShift® 容器平台之上,并充分利用其潜在资源和基础架构优化管理的优势。该解决方案可以部署在任何云上,并充分支持 AWS、Azure、Google CloudPlatform、IBM Cloud® 和私有云部署等多云环境。它的关键集成功能跨越了从数据管理和数据运营到业务分析和 AI 的整个分析生命周期。主要优点包括:

- **单一的统一平台**  
根据您的需求,将数据管理、数据治理、数据科学和 AI 功能整合到一个直观的集成平台中。
- **内置治理**  
使用自动化的端到端治理,帮助在整个组织内实施策略和规则,并快速响应不断变化的法规。
- **可扩展且可定制**  
从不断增长的专利、第三方和开源服务目录中灵活部署数据和 AI 服务,构建最适合您需求的平台。
- **预建的 AI 和行业应用**  
借助 IT 运营、客户服务、风险和合规以及金融运营的行业解决方案,快速创新。
- **专为混合云设计**  
由于其云原生设计和 RedHat OpenShift 基础,几乎可以在任何环境中部署该平台,无论是在本地还是在云上。

## 新一代的

## IBM Cloud Pak for Data

最新版本的 IBM Cloud Pak for Data 凭借新的 AI 赋能能力进一步为整个平台注入智能自动化,这些新 AI 赋能能力是平台内新的智能Data Fabric(数据经纬)的核心部分。

这种智能Data Fabric(数据经纬)使用 AI 来自动化复杂的数据管理任务,并跨多个环境中普遍发现、集成、编目、保护和治理数据。

利用智能Data Fabric(数据经纬)实现工作流程自动化

### AutoAI

自动化数据准备、模型开发和特征工程,在几分钟内找到并部署性能最佳模型。简化 AI 生命周期管理,更快地建立模型,加速部署,开发 AI,以具备更加广泛的技能。

若要了解更多 AutoAI 的相关信息, [请点击此处](#)。

### AutoSQL

自动访问、更新和统一分布在分布式存储和云中的数据,无需数据移动或复制。AutoSQL 是一种高性能的通用查询引擎。它简化了数据图景,支持跨不同数据源(包括数据仓库、数据湖和流动数据)中使用相同的查询与平台现有数据,全部都无需手动更改。

若要了解更多 AutoSQL 的相关信息, [请点击此处](#)。

### AutoCatalog

为用户相关性实现整个混合数据和云图景中发现、编目和充实数据自动化。向更多人提供业务即用数据。

### AutoPrivacy

在混合云环境中,自动执行跨混合数据生态系统的通用数据和使用策略,在实现数据使用的同时降低风险。

若要了解更多 IBM Cloud Pak for Data 可以提供的统一隐私框架的相关信息, [请点击此处](#)。

## IBM Cloud Pak for Data 的部署模型

自三年多前发布 IBM Cloud Pak for Data 以来,IBM 不断推进新功能和额外的部署和消费模型,包括:

- IBM Cloud Pak for Data:一个可以在任何云上运行的客户管理软件平台。
- IBM Cloud Pak for Data System:一个预配置的超融合系统,将存储、计算、网络和软件相结合,并将私有云的部署时间减少到几个小时。
- IBM Cloud Pak for Data as a Service:一种用于IBM Cloud Pak for Data 服务启动套件的“即付即用”订阅模式,在 IBM Cloud 基础结构上充分管理。

# IBM Cloud Pak for Data 的顶级用例

## 数据存储融合

**简化数据图景，支持以更快的速度从数据中提取价值，具有成本效益。**

数据驱动数字转型；企业需要以足以影响业务成果的速度从其数据中快速提取准确和可操作的洞察。数据驱动的洞察能够激发企业范围内的创新，发现新产品或市场的机遇，让销售人员能够进行更有意义的讨论，并确定可以改进的内部流程。

数据存储融合包含将数据从传统到现代数据经纬的迁移，以促进和加快数字转变；它涉及优化数据资产，以更好地利用不同的数据类型、结构、容量和速度。

IBM 数据存储融合让企业能够跨越所有数据景观收集多维数据，提供支持关键任务应用程序的业务即用数据，为高级分析提供动力，并实现实时运营高效，同时减少基础架构。

## DataOps

**快速为 AI 交付业务即用数据。**

随着业务和市场需求的变化，企业需要确保能够快速响应，支持远程工作、计划削减成本和其他必要事项，同时保持其竞争优势。越来越多的企业发现，数据是他们最大的资产。为了获取企业数据的价值，企业必须正视当前混合和多云数据源的现实，并设计现代信息架构，以适应现实中大多数情况中各种各样的数据源和格式。

DataOps 数据集成和治理能力让企业能够从其所有数据中获取价值，这就要求企业随时随地访问和信任结构化和非结构化数据。允许验证规则作为摄取过程的一部分自动运行，有助于减少孤岛，确保近实时地使用有意义和有价值的信息。

革命性的云原生 IBM Cloud Pak for Data 平台统一了跨越整个数据和数据分析生命周期的市场领先服务，并作为企业数据经纬，帮助整合数据，用于分析和 AI。它包括以前由 IBM Information Server 平台提供的功能，这些功能现在可作为 IBM Cloud Pak for Data 上的 DataStage® 和 IBM Watson® Knowledge Catalog 云即用服务使用。

## 数据隐私和安全

**为混合云构建一个普遍的隐私框架。**

大多数企业都面临着两项往往相互矛盾的当务之急：让数据更易于为分析和 AI 所用，同时确保安全和合规。由于最近数据类型和数据源激增，以及针对信息使用方式和禁止方式的响应规则和法规，管理这些看似矛盾的事项的优先级愈发复杂。更加聪明的企业正在寻求新的战略和技术，以帮助将其不同的数据安全、隐私和治理实践转变为更加全面的方法，了解和监管整个企业的敏感数据。

IBM 正在简化这一过程，帮助客户充分了解和敏感数据在其企业中的利用方式。在一个可以在任何云中部署的协作平台中，集成包含数据安全、混合数据管理、治理、风险和合规在内的领先解决方案，IBM 提供跨客户业务的 PII 实时视图。除了访问敏感数据的人员、原因、对成果的影响这一普遍视图之外，客户还获得了加快降低所有数据和 AI 风险所必需的普遍政策和法规执行工具。

利用智能自动化加速敏感数据的收集、编目和屏蔽，构建可信、合规的数据基础，让您的团队和模型可以轻松访问这些数据。将每位数据用户转变为内置 AI 的风险和合规专家，了解最新法规、自然语言和环境，增强团队的技术和理解。

## ModelOps

**实现 AI 生命周期的自动化，提高 AI 投资回报。**

到 2023 年，70% 的 AI 工作负载将使用应用程序容器或使用无服务器编程模型构建，这就需要 DevOps 文化。<sup>2</sup>

模型操作 (ModelOps) 是在应用程序中实施模型的一种原则性方法。ModelOps 同步应用程序和模型管道之间的节奏。通过 ModelOps，可以使用边缘混合云的数据、模型和资源来优化数据和 AI 投资。ModelOps 涵盖了跨云优化模型和应用程序使用的端到端生命周期，针对机器学习和深度学习模型、优化模型和其他运营模型，与持续集成和持续交付 (CICD) 相结合。

为构建 ModelOps 实践，需要平台来简化和自动化组织数据的 AI 生命周期；构建、运行和管理模型；并优化决策。IBM Cloud Pak for Data 是构建 ModelOps 实践和加速 AI 投资跨任何云的价值理想平台。



凭借值得信赖的、合规的 AI 实现创新。

AI 正开始发挥其潜力。然而，随着应用率的上升，当企业 试图治理、管理和保护数据以及根据该数据训练的模型时，数据生命周期中出现了新的差距。“AI 治理”一词是指制定策略的组织方法，以有效设计、部署和监测 AI 赋能的模型和算法，重点关注公平、问责、透明度、安全和隐私，以帮助确保结果公平。

几十年来，IBM一直处于 AI 的最前沿，而 IBM Cloud Pak for Data 平台现在可以应对数据和 AI 生命周期的每个阶段。它带有内置的治理和质量工具，如 IBM Watson Knowledge Catalog，以及使用 Watson Studio 的模型自动化和专门构建的 AI 模型风险管理工具，包括 Watson OpenScale™。AI 的基础是数据。凭借建立信任和确保一致合规的统一平台，AI可以真正用于实现创新。

## 后续步骤

采取后续步骤，了解 IBM Cloud Pak for Data 的更多相关信息。

[阅读Data Fabric\(数据经纬\)白皮书](#)

[立即注册，获取免费试用产品](#)

[访问平台页面](#)

© Copyright IBM Corporation 2021

国际商业机器中国有限公司  
北京市朝阳区北四环中路27号  
盘古大观写字楼25层  
邮编: 100101

美国出品  
2021 年 5 月

IBM、IBM 徽标、ibm.com、IBM Cloud Pak、IBM Cloud、DataStage、IBM Watson 和 OpenScale 是 International Business Machines Corporation 在美国和/或其他国家或地区的商标或注册商标。其他产品 和服务名称可能是 IBM 或其他公司的商标。IBM 商标的最新列表可以在ibm.com/trademark上查看。

Red Hat 和 OpenShift 是 Red Hat 公司及其子公司在美国和其他国家的注册商标。

本文档为自最初公布日期起的最新版本，IBM 可随时对其进行修改。IBM 并不一定在开展业务的所有国家或地区提供所有 产品和服务。

这里讨论的性能数据是在特定操作条件下得出的。实际结果可能会有差异。用户负责评估并验证与 IBM 产品和程序配合使用的任何其他 产品或程序的运行。本文档中的信息“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，不包含任何有关适销、适用于某种特定用途的保证 以及有关非侵权的任何保证或条件。IBM 的产品 根据其提供时所依据协议的条款和条件获得保证。

客户负责确保对适用的法律和法规的合规性。IBM 不提供任何法律咨询，也不声明或保证其服务或产品将确保客户遵循任何法律 或法规。

良好安全实践声明：IT 系统安全性涉及通过 防御、检测和响应来自企业内部和外部的不正当 访问来保护系统和信息。不正当的访问可能导致 信息被篡改、破坏或盗用，或者导致您的系统遭到误用而攻击别人。任何 IT 系统或产品都不应被认为是完全安全的，而且没有任何单一产品、服务或安全措施在防止不正当的使用或访问方面是完全有效的。IBM 系统、产品和服务旨在成为 合法、全面的安全方法的一部分，它必定涉及额外的操作程序，并且可能需要其他系统、产品或服务配合才能获得最好的效果。IBM 不保证任何系统、产品 或服务免受任何一方的恶意或非法行为为干扰，或帮助您企业免受任意一方恶意或非法行为的攻击。

关于 IBM 未来方向和意向的声明仅表示目标和目的，可能随时更改 或撤销，恕不另行通知。

- 1 IBM Global AI Adoption Index 2021 Executive Summary, 2021.
- 2 Gartner Artificial Intelligence, Data Science & Machine Learning, 2020.