

F R O S T & S U L L I V A N

沙利文



头豹
LeadLeo

2023年中国数据管理解决方案 市场报告

湖仓一体、数智融合、云原生、数据治理

2023年6月

头豹研究院
弗若斯特沙利文咨询（中国）

研究框架

◆ 数据管理解决方案技术动态跟踪	5
• 技术动态围绕的两大趋势主题	
• AI 带来数据管理技术革新	
• 云原生化进程加深	
• 方法论助力完善数据管理解决方案	
• 构建安全的数据流通与共享能力	
◆ 中国数据管理解决方案市场分析	18
• 市场落地的机遇与挑战	
• 供给端市场落地布局分析	
• 供应商行业领域覆盖图谱	
◆ 中国数据管理解决方案行业竞争分析	27
• 中国数据管理解决方案行业评价维度	
• 中国数据管理解决方案行业综合评分表现	
• 中国数据管理解决方案行业领导者	
◆ 名词解释	36
◆ 方法论	37
◆ 法律声明	38

图表目录

• 技术动态围绕的两大趋势主题	-----	5
• 人与 AI、数据与 AI 的关系	-----	8
• AI 为数据管理解决方案带来的智能化功能	-----	9
• AI 擅长与不擅长的能力	-----	10
• 传统云托管与云原生下的数据管理解决方案	-----	12
• 云原生数据管理解决方案的价值	-----	12
• 供需端云原生与 Serverless 方面动态	-----	13
• Serverful 单体架构、Serverful 微服务架构、Serverless 架构各自的优势	-----	14
• 三大数据管理方法论简介	-----	15
• DataOps 与 Data Fabric 可支撑 Data Mesh 落地	-----	16
• 数据安全法与个人信息保护法相关内容	-----	17
• 数据对经济、社会、企业价值提升	-----	19
• DCMM能力等级参评单位占比及参评单位数量变化	-----	20
• 数据管理相关政策整理	-----	21
• 用户对 Serverless 的顾虑与解决方法	-----	22
• 不同的行业需求因业务与基础能力不同而有所差异	-----	23
• 云计算产品降价对供应商生态发展和技术的影响	-----	24
• 实现行业场景化能力增强的供应商代表	-----	25
• 数据管理解决方案行业覆盖图谱	-----	26
• 中国数据管理解决方案竞争力评分维度	-----	28
• 中国数据管理解决方案综合竞争表现	-----	30
• 中国数据管理解决方案领导者	-----	31

1.1 技术动态围绕的两大趋势主题

关键发现

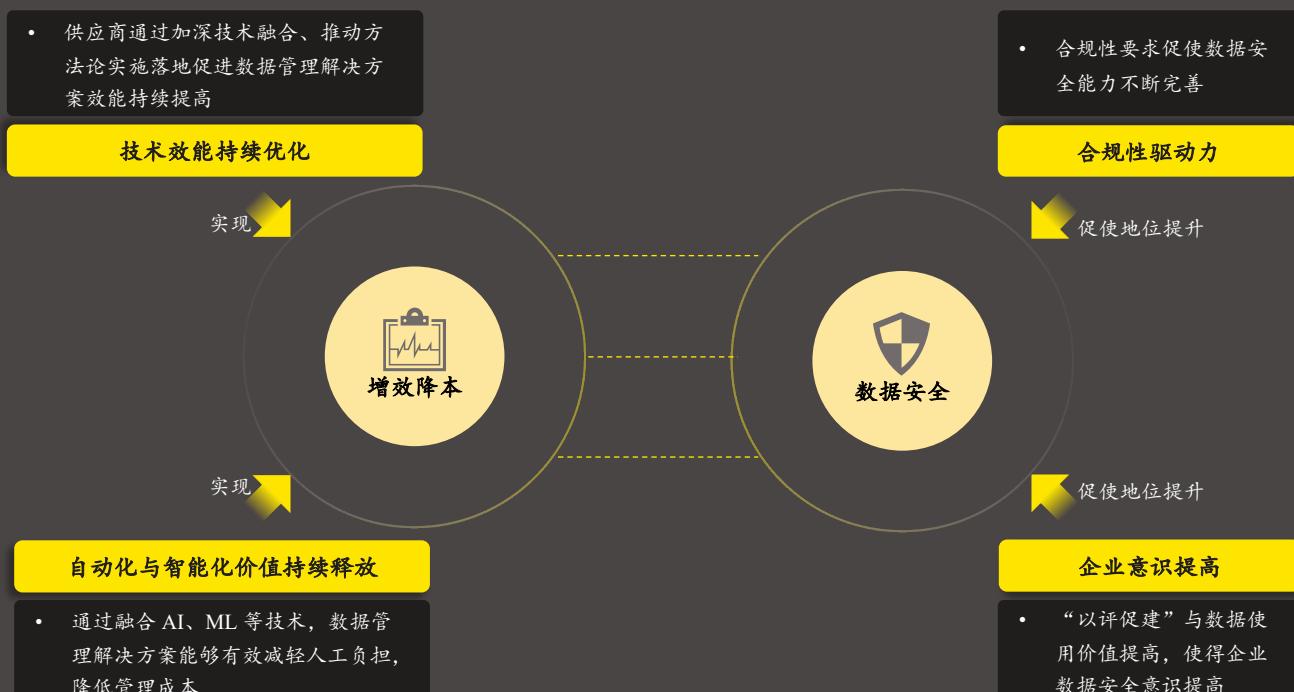
供应商的技术动态围绕增效降本和数据安全两大主题进行，以加强数据管理解决方案商业落地效果，加强数据管理解决方案与企业数据管理能力可持续发展建设的匹配度，促进数据价值流动与释放

□ 增效降本和数据安全两大主题

随着近年来数据管理技术的不断优化，加上国家政策不断加强对数据领域发展的关注，中国企业加强了数据管理技术落地实践，数据管理水平正在逐步提高。不过，企业面临着技术落地后仍无法充分将数据价值释放的问题，这主要源于企业在数字化转型过程中，需要带着现有IT架构、组织架构进行技术迭代，一味地追求新技术而没有考虑企业固有属性，会导致无法将数据价值与公司业务发展连通，如技术复杂化加重数据孤岛、技术复杂化扩大数据团队与业务团队的沟通成本等。

在上述背景下，我们关注到供应商的技术动态围绕增效降本和数据安全两大主题进行：通过技术融合与方法论实施，增强解决方案性能效果和面向场景的落地效果，并降低成本；通过丰富数据安全功能，助力数据流通。这些布局将加强数据管理解决方案与企业数据管理能力可持续发展建设的匹配度，提高方案落地效果。

技术动态围绕的两大趋势主题



01

增效降本：自动化与智能化价值持续释放

□ AI 技术提高，扩大其在数据管理领域发挥的作用

早期自动化与智能化的实现主要基于预设的规则与模型，能够一定程度上减轻数据管理团队的压力，但可实现的功能与影响的范围还很有限。随着 AI 技术提高，可实现的智能化功能得到扩展，如预测性功能、自适应功能、自然语言理解相关功能等，这些功能不仅能够让企业进一步减轻团队压力，更好地进行数据洞察，降低数据管理运维成本，还能够降低数据使用门槛，扩大数据使用对象范围，从而助力企业同时解决技术落地与组织协作的问题，让数据变得能用、好用、易用，使数据能够真正地赋能企业业务发展。

02

增效降本：技术效能持续优化

□ 加强云原生发展，将方法论融入解决方案，持续优化技术效能

数据湖、湖仓一体、HTAP 等技术能够解决企业业务发展所带来的问题，因此也在近几年得到推行与落地。这些技术的潜在价值显著，但却尚未能完全释放，这当中的阻力主要源于新技术的落地需要与现有 IT 架构与组织架构构建新的体系，常见问题包括新技术与现有 IT 架构的协同需要一定学习成本，技术落地后运维压力不减反增；技术落地提高了体系复杂度，但数据使用的流程未作更新或难以更新，导致技术落地难以实现。

针对这些问题，供应商给出云原生与结合方法论的解法，为技术价值的释放提供动力。比如说，云原生能够使解决方案更轻便、性能天花板更高，既减轻运维压力，也能提高性能；方法论从全局的角度为构建新的体系提供指引，从而促进数据管理效率以及组织协同效率的优化。基于这些选择，企业将能逐步扫除阻力，充分利用技术的效能。

03

数据安全：合规性驱动与企业意识提高使数据安全能力地位提高

□ 数据安全能力影响着数据价值能否得到充分挖掘与释放

数据安全的核心问题是如何让数据合规且安全地在内部与外部流动起来，而流动的程度则决定着数据价值是否能够被真正地挖掘与释放。在企业数据管理意识逐步提高与合规性驱动下，数据安全的重要性得到重视，因此投入意愿也有所提高。供应商通过提供数据安全管理、数据共享的方法与工具，帮助企业进行数据安全能力建设，促进数据能够合规、安全、顺畅地流动，以得到充分利用与价值挖掘。

1.2 数智协同推动技术革新

关键发现

AI与数据天然具有正向协同关系，因此AI与数据管理融合形成技术革新具有必然性。通过使用AI，数据管理解决方案不仅能够优化技术落地效果，还能促进团队间协作以优化组织协同能力，以充分挖掘与释放数据价值。

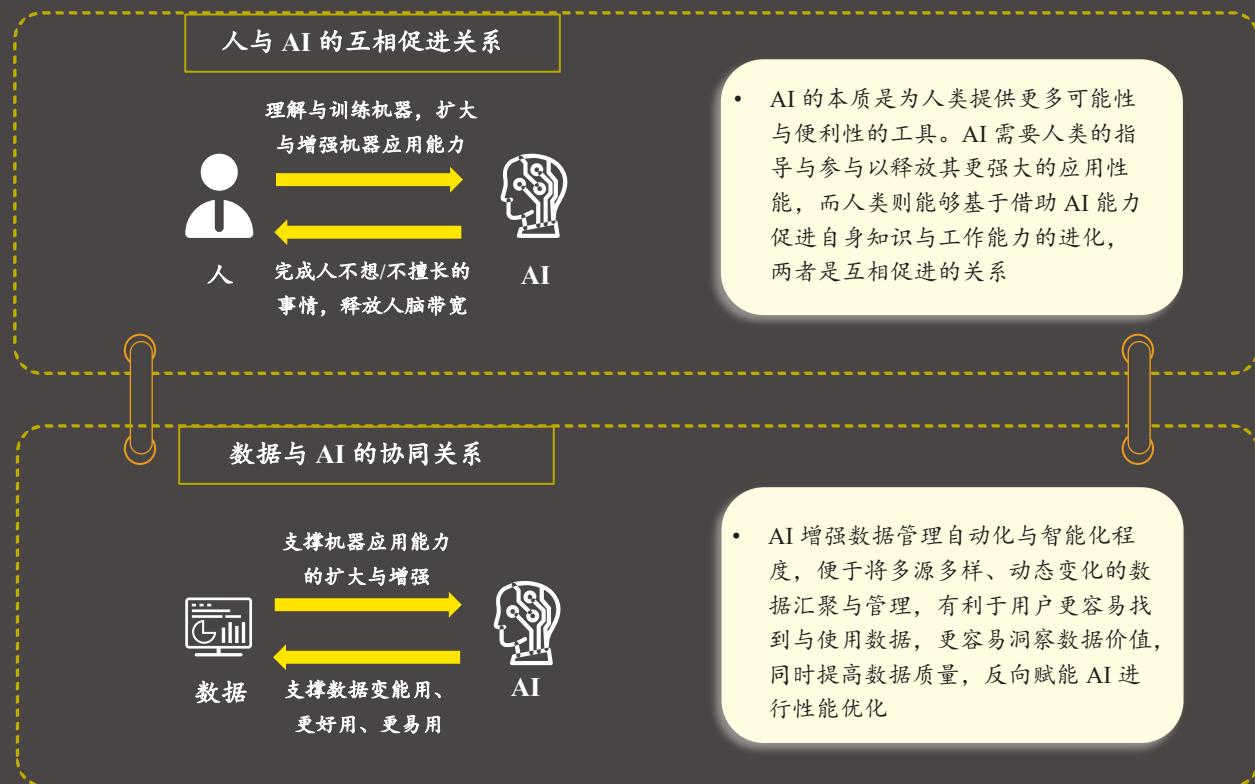
□ AI与数据管理天然具有互相赋能的关系

在数据复杂性、数据量、业务快速变化的环境之下，企业需要更有效率的数据管理解决方案，AI为数据管理带来技术革新成为必然，同时两者并非单向赋能关系，而是正向协同关系，更是进一步加强了两者融合的必然性。

AI赋予数据各参与者更强大的工具与决策支持，减轻了复杂工作、枯燥工作给数据团队带来的脑力带宽消耗，帮助团队将精力更多地花费在业务交付，而非工具的管理，实现数据管理的效果与效率优化。与此同时，数据管理的优化促使数据质量提高、多源多样的数据得到有效管理，使AI性能与训练效率提高得到支撑，因此AI与数据管理天然具有互相赋能的关系。

随着AI大模型发展促使场景化应用能力增强，数智融合形成的新技术功能也已开始实现落地。沙利文认为，目前“数智融合”趋势已被推向了关注度的顶峰，而在这一趋势中的下一阶段将是从“数智协同（AI for DMS）”向“数智统一治理（AI in DMS）”发展。

人与AI、数据与AI的关系



来源：沙利文



www.leadleo.com
400-072-5588

FROST & SULLIVAN

沙利文

□ AI 不仅能优化技术，也能促进组织协同

AI 融入数据管理解决方案形成的功能主要能够在加强数据管理技术落地效果与组织协同两大方面带来价值：

- **加强数据管理技术落地效果：**AI 为数据管理功能带来扩展与延伸，形成过去人工干预和手动操作无法实现和难以实现的功能。比如说，基于 AI 对 SQL 访问情况与习惯进行分析，可对数据库和配套产品的性能问题进行排查，并提供优化建议。在无 AI 的环境中，这些性能问题可能难以被发现，或性能问题非显著，企业没有立项解决的动力，导致问题一直存在与堆积，直到性能问题变得显著的时候才进行紧急解决。此外，AI 技术还能通过数据驱动不断学习与优化，进而适应不断变化的数据管理需求与挑战，提高数据管理技术在实际应用的能力。
- **组织协同：**AI 技术的导入可降低数据使用的门槛，促进数据平民化，如业务人员可使用自然语言转 SQL 的功能进行数据获取，减少了沟通成本与流程时间，提高数据分析的效率。过去数据获取的流程可能需要逐个需求提交申请，存在数据团队不理解需求或多个需求排队，导致效率低下。同时，AI 技术也能加强数据工程师与数据科学家的合作，如利用基于 AI 的数据流监测，可及时、准确的发现问题，减少双方的摩擦，高效进行数据开发。

AI 为数据管理解决方案带来的智能化功能



□ AI 对数据管理从业者的冲击

区别于利用ML和AI的增强数据管理技术，GPT-4等大语言模型展示出强大能力后，对于数据分析师、数据科学家、DBA、数据开发工程师等数据管理从业者而言，将会有什么冲击？

在数据管理领域，建议对大语言模型的关注点放在“接替工作”而非“替代人类”，即接管人类不想做/不擅长的事情，释放人的脑力带宽去专注到更重要的业务交付上。从职业所需能力维度来看，AI 经过训练与动态优化后，能够基本上完成基本能力的工作，但在缺乏人类动态指引与监督难以进行需要深度处理能力的工作，因为这受到多因素的影响，如对错标准不一、沟通协作与同理心、多维度知识经验与技能的运用等。

AI 擅长基本能力去完成工作，但不擅长自主进行处理能力

OpenAI 研究报告中职业所涉及的 O*NET 基础能力

处理能力

批判性思维——基于逻辑与推理，判断得出问题的多种解决方案、结论或方式的优劣势

主动学习——洞察新信息在现在与未来的问题解决与决策的影响

学习策略——经过训练与指导，选择和使用合适的方法和步骤进行新事物的学习与传授

反省能力——监视与评价自己、他人和组织的行为，以做出对行为的优化或修正

基本能力

阅读理解——理解与工作相关文件中的句子和段落意思

积极倾听——对别人所说的话给予充分的关注，花时间去理解注意要点，适当提问，不在不合适的时候打断别人

写作能力——根据听众的需要进行有效的书面沟通

说话能力——与他人交谈以有效地传达信息

数学能力——用数学来解决问题

科学能力——运用科学的规则和方法来解决问题

OpenAI 研究指出的人类难以被AI替代的能力



OpenAI 研究指出人类最容易被AI替代的能力



跨职能技能

编程能力——为各种目的编写计算机程序

现阶段的数据管理流程与架构将于 AI 的融入迎来升级，人类需要做的是理解与学习 AI，利用 AI 赋能决策，这要求：

1. 企业需加强培训提高企业员工的 AI 素养，重点加强批判性思维与科学解决问题的能力；
2. 数据团队与 AI 团队应学会如何进行协作，优化工作流程和组织架构，以获得有效的数智协同成果；
3. 数据各参与方应学会利用 AI，不断优化自身数据管理、运用的经验与知识，在AI进化的同时，实现自身的完善与进化；
4. 企业要进一步推进高质量的数据治理转型工程，打通“信息孤岛”、开发利用“黑暗数据”、转化“数据资产”价值。“无治理、不分析”：需获得海量高质量的数据才能孵化出可信的AI。

01

数智融合带来的关键趋势：数据管理领域的职位将进一步细化

□ 人类仍为数智融合时代的关键角色，需适应不断变化的数据管理格局

参考软件工程领域，“软件工程师”的职位开始发展出如DevOps工程师或SRE网站可靠性工程师等更专业化细分的角色。随着数据工程领域也逐渐技术成熟化和场景复杂化后，预计会出现更多根据业务需求或目标定义的职位：

- 数据质量工程师 (Data reliability engineers) 将负责保证数据的质量
- 数据产品经理 (Data product managers) 将推动数据的应用和盈利
- DataOps 工程师 (DataOps engineers) 将专注于数据的治理和提高效率
- 数据架构师 (Data architects) 将专注于解决数据孤岛问题和长期投资

而既有的数据分析师、数据科学家、DBA、数据开发工程师等数据管理岗位，或许将面临一些来自AI自助式流程替代的冲击而对工作内容和岗位责任有所改动。但我们认为原有的这些职位并不会消失，虽然ChatGPT能力受到了大量关注，但这个过程还是极其依赖于专业数据从业者进行审查和调整。而面对大型语言模型的低可解释性、不确定的道德与合规、不确信的数据隐私安全、低准确性的质量控制等隐患，人类在数据管理里的角色仍然至关重要。尽管基础设施可能会有所变化，自动化技术可能会改变我们花费时间和注意力的方式，但在可见的未来中，专业的数据管理从业人员仍然将在数据管理的过程中发挥主导作用。

02

数智融合带来的关键趋势：数智统一治理 AI in DMS

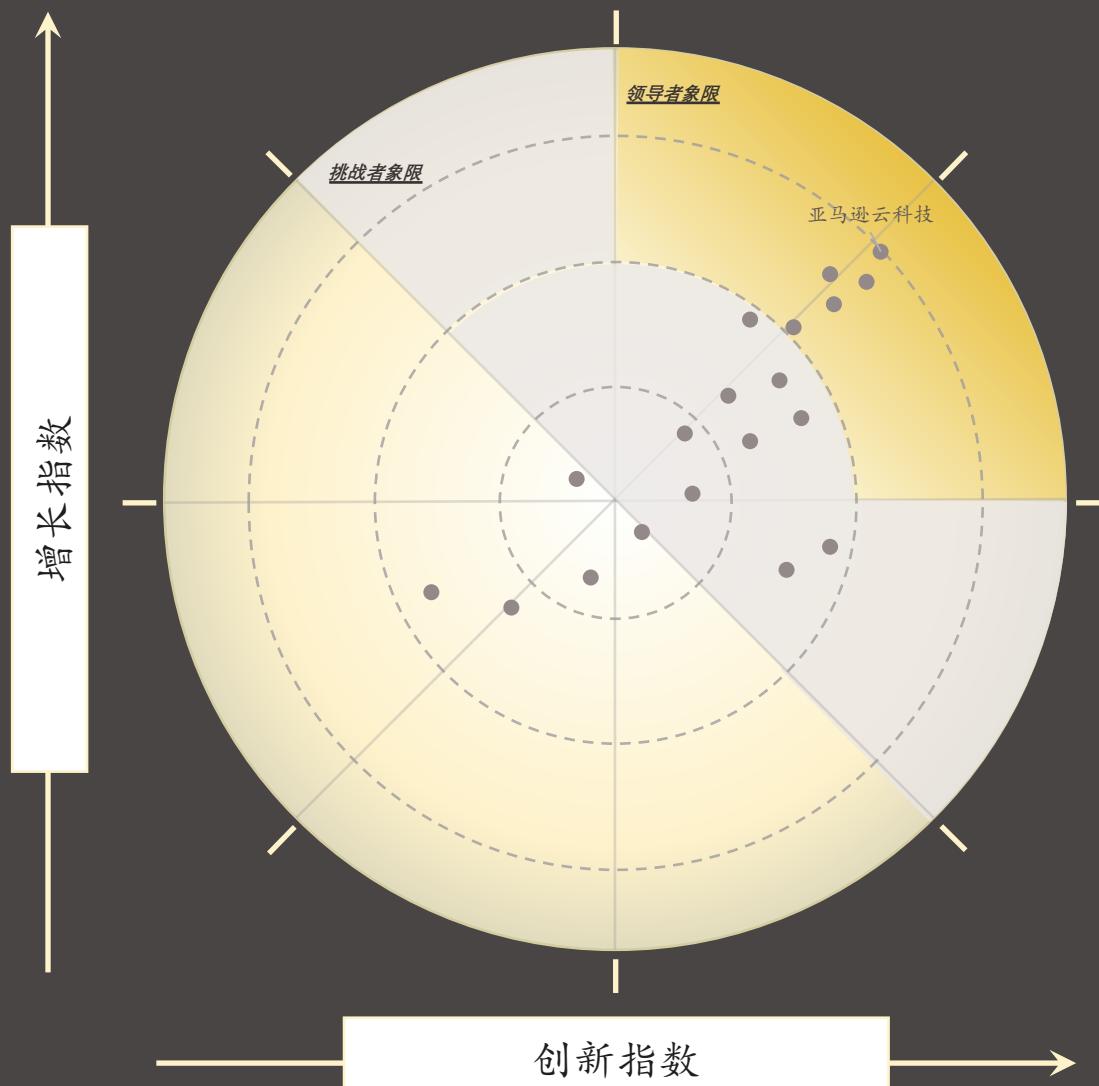
□ AI in DMS 将能显著优化数据管理解决方案使用体验

AI in DMS (Artificial Intelligence in Data Management Solutions) 不仅仅是在数据管理系统中使用AI技术，而且还包括将AI平台作为数据管理系统的内置功能，并将AI平台从一个辅助工具变成一个核心中台的定位。这种深度耦合使得数据管理系统能够更好地处理和理解大量的数据，从而提高数据分析的效率和准确性。

在这个数智统一平台的环境中，可以覆盖从数据基座到AI开发再到各类数据分析应用的数据价值挖掘全生命周期。这种在同一平台上统一生产的方式，不仅可以提高数据处理的效率，还可以确保数据的一致性和准确性。

AI in DMS 将成为一种新的数据管理模式，它通过将AI技术深度集成到数据管理系统中，实现数据和智能的深度融合，从而提高数据管理系统的整体效率和准确性。而由于数智双系统的一体化，流数据的数智应用、实时数据可视化、业务需求驱动数据与AI开发等，都将成为新的使用体验。

■ 2022年中国数据管理解决方案市场综合竞争表现——Frost Radar (弗若斯特雷达)TM



注：圆环按由内向外递增的逻辑对应由低至高的综合评分，竞争力由“创新指数”以及“增长指数”综合得出。结论仅适用于该阶段数据管理解决方案市场发展情况。

□ 纵坐标代表“增长指数”：

- 衡量竞争主体在数据管理解决方案增长维度的竞争力，位置越靠上方，数据管理解决方案的性能与兼容性、安全保障、服务支持、产业链生态、商业落地成熟度等市场增长能力及水平越强

□ 横坐标代表“创新指数”：

- 衡量竞争主体在数据管理解决方案创新维度的竞争力，位置越靠右侧，数据管理解决方案的数据仓库与数据湖技术改进、云原生创新性、湖仓一体创新性、方法论创新性等能力及水平越强

领导者：亚马逊云科技

关键发现

亚马逊云科技发力云原生数据战略，持续提高数据管理解决方案技术先进性，并通过加强产品间联动的性能、结合自动化和机器学习提供的智能服务，优化解决方案落地用户端的使用体验与效果，是中国数据管理解决方案的领导者

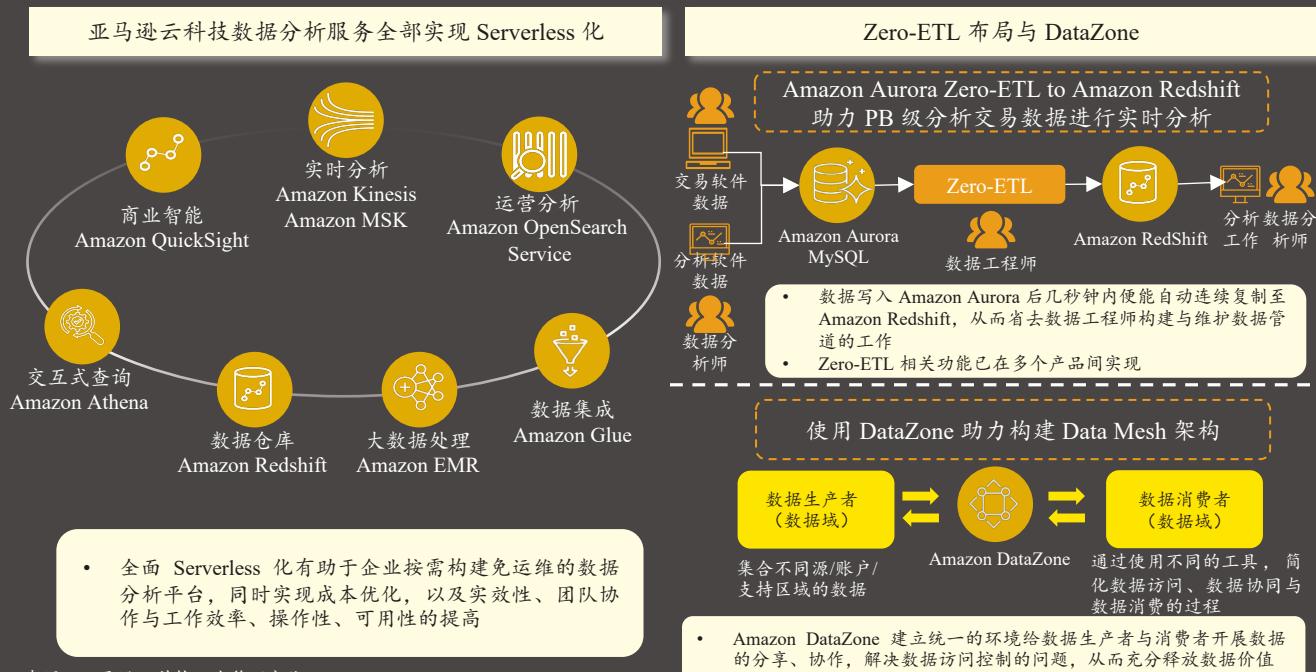
完善云原生产品矩阵，持续优化用户体验与方案落地效果

企业对于数据的分析与管理会随着时间而动态变化，这伴随着数据技术引入困难、数据管理成本提高等方面的难题。亚马逊云科技建立了完善的云原生产品矩阵帮助企业应对这些难题，其中，数据分析服务于今年正式实现全面 Serverless 化，有助于通过免运维服务的联动，让企业实现增效降本，同时能够更迅速地适应企业业务规模、逻辑、需求的变化。与此同时，亚马逊云科技产品智能化程度也对优化用户体验起到至关重要的作用，如 Amazon Glue FindMatches 通过机器学习能力可自动识别数据集中的重复或匹配记录，实现关联客户记录、产品目录去重、欺诈侦测的功能。

实现跨组织、高效的数据一体化融合

针对如何提高企业数据分析效率的问题，亚马逊云科技给出“Zero-ETL”的解法，通过减少不同服务间手动迁移或转换数据的工作，帮助企业减少或免去构建与维护 ETL 数据管道的时间。此外，跨组织进行数据共享的效率容易受到数据找不到、数据访问控制过严或过松等方面的影响，Amazon DataZone 通过建立统一的环境，帮助企业进行数据分类、发现、管理、共享和分析，能够显著降低组织间协作的阻力，提高效率。

亚马逊云科技产品持续升级，致力于提升用户体验



名词解释

- ◆ **数据仓库：**是为企业所有级别的决策制定过程，提供所有类型数据支持的战略集合。
- ◆ **数据湖：**是一个集中式存储库，允许您以任意规模存储所有结构化和非结构化数据。
- ◆ **湖仓一体：**是一种全新的开放式架构，能够兼顾拥有类似数据结构的数据湖所具备的灵活性和可扩展性优势，以及数据仓库的数据管理功能。
- ◆ **微服务：**是一种架构风格和设计模式，提倡将应用分割成一系列细粒度的服务，每个服务专注于单一业务功能，运行于独立部署的进程中，服务之间边界清晰是一种架构风格和设计模式，提倡将应用分割成一系列细粒度的服务，每个服务专注于单一业务功能，运行于独立部署的进程中，服务之间边界清晰。
- ◆ **单体架构：**是一种传统的架构风格和设计模式，所有的业务模块在单点运行、开发、部署。
- ◆ **Serverless（无服务器部署）：**无服务器部署通过FaaS+BaaS提供服务，允许用户在不构建运维一个复杂的基础设施的情况下进行开发，运行和管理应用程序。
- ◆ **容器：**是一种轻量级、可移植、自包含的软件打包技术，使应用程序可以在几乎任何地方以相同的方式运行。开发人员在自己笔记本上创建并测试好的容器，无须任何修改就能够生产系统的虚拟机、物理服务器或公有云主机上运行。
- ◆ **DataOps：** DataOps 是面向数据流水线，以实现从数据采集、清洗加工、到分析的全数据链路最大程度自动化运行的一种理念，而并非一项技术或工具体系。如今，单纯的业务应用已不存在，而数据驱动型应用开始影响企业的经验状况，企业需要一套理念来快速响应甚至是引领业务需求，DataOps 获得发展机会。从组成部分看， DataOps 包含方法论、技术工具和组织文化三大板块，三者的相互配合能达到提高数据管理效率的目的。
- ◆ **Data Fabric：** Data Fabric 为一种新兴的数据管理设计理念，可实现跨异构数据源的增强数据集成和共享，通过对现有的、可发现和可推断的元数据资产进行持续分析，来支持数据系统跨平台的设计、部署和使用，从而实现灵活的数据交付。
- ◆ **Data Mesh：** Data Mesh 是使用分布式体系结构框架的数据管理方法，旨在解决传统数据管理的瓶颈和挑战。它提倡将数据从单一的中心化管理模式转变为分散的、自治的数据单元，通过 Data Mesh 来实现数据的分散、协作和自治。

方法论

- ◆ 头豹研究院布局中国市场，深入研究19大行业，持续跟踪532个垂直行业的市场变化，已沉淀超过100万行业研究价值数据元素，完成超过1万个独立的研究咨询项目。
- ◆ 头豹研究院依托中国活跃的经济环境，研究内容覆盖整个行业发展周期，伴随着行业内企业的创立、发展，扩张，到企业上市及上市后的成熟期，头豹各行业研究员积极探索和评估行业中多变的产业模式，企业的商业模式和运营模式，以专业视野解读行业的沿革。
- ◆ 头豹研究院融合传统与新型的研究方法论，采用自主研发算法，结合行业交叉大数据，通过多元化调研方法，挖掘定量数据背后根因，剖析定性内容背后的逻辑，客观真实地阐述行业现状，前瞻性地预测行业未来发展趋势，在研究院的每一份研究报告中，完整地呈现行业的过去，现在和未来。
- ◆ 头豹研究院密切关注行业发展最新动向，报告内容及数据会随着行业发展、技术革新、竞争格局变化、政策法规颁布、市场调研深入，保持不断更新与优化。
- ◆ 头豹研究院秉承匠心研究，砥砺前行的宗旨，以战略发展的视角分析行业，从执行落地的层面阐述观点，为每一位读者提供有深度有价值的研究报告。

法律声明

- ◆ 本报告著作权归头豹所有，未经书面许可，任何机构或个人不得以任何形式翻版、复刻、发表或引用。若征得头豹同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“头豹研究院”，且不得对本报告进行任何有悖原意的引用、删节或修改。
- ◆ 本报告分析师具有专业研究能力，保证报告数据均来自合法合规渠道，观点产出及数据分析基于分析师对行业的客观理解，本报告不受任何第三方授意或影响。
- ◆ 本报告所涉及的观点或信息仅供参考，不构成任何证券或基金投资建议。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告或证券研究报告。在法律许可的情况下，头豹可能会为报告中提及的企业提供或争取提供投融资或咨询等相关服务。
- ◆ 本报告的部分信息来源于公开资料，头豹对该等信息的准确性、完整性或可靠性不做任何保证。本报告所载的资料、意见及推测仅反映头豹于发布本报告当日的判断，过往报告中的描述不应作为日后的表现依据。在不同时期，头豹可发出与本报告所载资料、意见及推测不一致的报告或文章。头豹均不保证本报告所含信息保持在最新状态。同时，头豹对本报告所含信息可在不发出通知的情形下做出修改，读者应当自行关注相应的更新或修改。任何机构或个人应对其利用本报告的数据、分析、研究、部分或者全部内容所进行的一切活动负责并承担该等活动所导致的任何损失或伤害。