

白皮书

生命科学MDM中的 自主AI

数据治理新纪元

索瓦尼亚·布卡帕坦姆·蒂鲁马拉，高级总监，产品与战略，全球主数据管理，IQVIA
FRANCESCA D'ANGELO，信息管理业务总监，IQVIA

目录

执行摘要	1
代理AI在MDM中的变革作用	1
比较代理式人工智能与传统 MDM 方法 2	
人工智能驱动MDM解决方案的关键评估标准	3
数据质量管理	3
数据剖析与编目	3
治理与合规	4
可伸缩性与性能	4
人工智能准备与整合	4
持续的数据治理和自我优化的治理	5
智能规则管理	5
通过自我优化的治理实现持续精炼	5
由人类专业知识指导的自动化	6
一个永续优化的 MDM 新时代	6
参考文献	7
关于作者	8

执行摘要

生命科学组织正处于一场变革的转折点：将具有能动性的人工智能（AI）集成到主数据管理（MDM）中，正在重新定义客户和产品数据的处理方式。与传统需要持续监督的AI系统不同，具有能动性AI可以自主运行：设定目标、做出决策并对其采取行动，只需最少量的人类干预。对于那些处理复杂数据集、面临严格监管和高风险挑战的制药和生物技术公司而言，这项创新承诺在数据质量、运营速度和整体流程敏捷性方面带来无与伦比的改进。

医药/医疗保健行业早期采用——截至2024年，约23%的组织已实施了AI代理——说明了这一转变的势头。¹

行业分析师预测，到2028年，近三分之一的企业软件将具备代理式AI功能，而从2024年的几乎为零。²

这份白皮书概述了自主AI在现代MDM项目中的作用，将其与传统的数据管理工具和常规的AI自动化方法进行对比，并提供了选择或升级MDM解决方案的评估标准。现实世界的案例叙述展示了自主AI如何促进持续的数据治理、智能规则管理和自我优化的数据管理。最后，我们推荐一个分阶段的“爬行-行走-奔跑”路线图，详细说明了生命科学团队在组织和技术方面的考量，使他们能够在确保合规性、可扩展性和持久商业价值的同时，利用这些突破。

代理AI在MDM中的变革作用

自主人工智能正在通过部署能够协作执行多步骤流程、运用高级推理（通常由大型语言模型提供支持）以解决复杂数据挑战的自主代理，革新主数据管理（MDM），使其成为动态数字劳动力，为MDM带来智能、自主和实时决策能力。这种创新范式使得数据清理、匹配、丰富和治理检查等常规任务能够实现全天候运行且人工干预极少，从根本上将MDM从手动、被动式流程转变为主动、自动驾驶系统。对于承担着庞大、多样且高度监管数据的生物科技公司——从临床试验结果到患者记录和产品信息——自主人工智能通过快速整合数据并简化监管报告编制等流程，提供了一种变革性的解决方案，最终让人类专家专注于战略计划，同时确保数据得到高效和准确的管理。



对比自主型人工智能与传统 MDM 方法

传统的MDM长期以来依赖于固定规则、手动数据管理和孤立的自动化脚本。MDM中早期的AI或机器学习计划主要是人领导的；它们通过自动化常规任务或标记异常来协助，但人操作员仍然保持总体控制。相比之下，代理型AI从根本上“AI领导”，部署一套自主代理，它们协作执行复杂的多步骤工作流，并在预定义的治理参数内进行决策，尽管人类输入对于监督和例外情况仍然是必要的。这种演变可以总结如下：



传统 MDM (人工/基于规则的)

- 依赖于预设规则和显著的人工干预来确保数据质量。
- 修正往往具有反应性——只在问题发生后才进行处理——并且这些系统除非规则被手动更新。
- 因此，传统系统既可以耗时，又难以扩展。



传统人工智能辅助MDM

- 结合机器学习和自动化来支持人类决策。
- 例如，机器学习模型可以检测重复项、标记异常值或建议合并和标准化，但它们仅在狭义定义的任务范围内运行。
- 仍需要人工批准或分析这些建议，因为系统不管理完整的工作流程或重新设计流程。



代理 AI驱动MDM

- 采用自主代理那些能整体管理数据的解读上下文和适应意外状况无需预编程规则。
- 一个能动系统可以，对于例如，识别新产品变体和智能集成将其纳入现有层次结构无需人工干预。
- 这些代理准确摄取数据，自主识别和修复不一致性，从过去的纠错中学习，甚至能在错误发生前预先阻止错误——从而在整个MDM生命周期中协调决策。

从依赖人类的过程转向由人工智能领导的范式具有变革性。将人工智能仅仅视为执行现有任务更快方式的组织，有错过充分实现其潜在机会的风险。自主式主数据管理通过使数据管理员从耗时且重复的任务中解放出来，重新定义了数据管理，让他们能够将注意力转移到更高价值的业务活动上。结果，公司可以实现前所未有的数据质量、一致性和速度水平。例如，一个原生的自主式人工智能主数据管理平台证明，传统方法需要大团队审核和纠正数据——这是一个缓慢且容易出错的流程——而自主式人工智能则持续且大规模地执行这些任务，有效地将主数据转变为自我改进、高价值的资产。

人工智能驱动MDM解决方案的关键评估标准

在使用代理式人工智能评估客户和产品MDM平台时，制药和生物技术组织应评估以下关键标准，以确保MDM计划在有效利用人工智能的同时保持信任、合规性和绩效。



数据剖析与编目

它意味着什么：

一个由人工智能驱动的MDM解决方案必须擅长数据剖析——理解数据内容、质量及其关系——并构建一个强大的数据目录。代理式人工智能可以通过揭示元数据、关系和异常情况来自动化新数据源的分析，无需人工干预。

它为何重要：

这项能力在生命科学领域尤其重要，因为数据来源于不同的系统（例如，研究实验室、临床试验、CRM、ERP）。强大的特征分析和目录支持增强了数据治理，并使组织为有效的AI集成做好了准备。

要寻找什么：

- 具备人工智能驱动数据发现代理等特性至关重要，这些代理扫描数据库以识别数据资产并评估其可信度。
- 一些平台提供数据发现代理，能够迅速识别用于分析的合规、相关数据，有效充当自主数据目录管理员。
- mdm解决方案应持续分析数据以进行质量或结构变化，并相应地更新元数据/目录，确保新数据集能被管理员和下游用户快速理解。

示例：

一个由人工智能驱动的数据发现代理，它扫描数据库以识别合规且相关的用于分析的数据，有效地充当自主数据目录管理员。



数据质量管理

它意味着什么：

平台应提供持续的数据质量监控和实时错误补救，而不是依赖定期清理。

它为何重要：

高质量、可信的数据对人工智能性能至关重要——代理的好坏取决于它们处理的数据。

要寻找什么：

- 评估该工具是否使用人工智能来识别并自动纠正数据问题（例如，重复项、不一致性、缺失字段）或至少建议更正。
- 领先解决方案结合机器学习与基于规则的方法，在初始阶段（例如创建时捕获重复的客户记录）防止错误，并批量清理遗留数据，带来可量化的改进，如提升数据质量评分或降低错误率。

示例：

某供应商的新代理系统有一个数据质量代理，该代理持续监控和纠正MDM系统和存储库中的问题。



治理与合规

它意味着什么：

尽管人工智能驱动的决定越来越多，但强大的数据治理仍然至关重要，并且必须扩展到人工智能代理本身。

它为何重要：

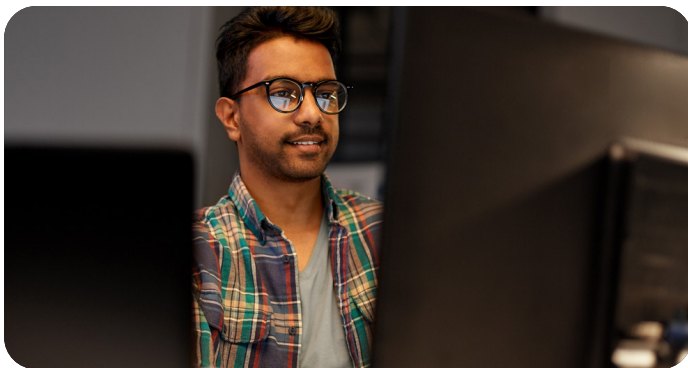
“数据信任悖论”——确保人工智能决策正确、合乎道德且合规——通过提供可解释的人工智能和完整的审计追踪的稳健治理框架得到解决。

要寻找什么：

- 评估代理式 MDM 解决方案，确保它们支持 AI 驱动的行动的政策执行、透明度和可审计性。
- 工作流必须通过执行访问控制并保护数据隐私来支持合规性（GDP R、HIPAA、FDA规则等），即使流程自动化。
- 寻找内置的治理功能，包括基于角色的控制、对高影响变更的审批工作流以及数据管理员用于监督 AI 活动的仪表板——确保治理是一个基础要素，而不是附加功能。

示例：

如果一个AI代理合并客户记录，平台应该记录关于是谁或什么进行了更改（代理与人类），理由（如应用的规则或学习到的模式），并提供审核或撤销的选项。



可伸缩性与性能

它意味着什么：

考虑到生命科学环境涉及大型复杂数据集——全球客户记录、广泛的产品目录以及数百万的交易数据点——MDM解决方案必须能够扩展以实时管理大量和多样的数据域，而不会降低性能。

它为何重要：

人工智能代理擅长并行处理，允许多个任务或数据集同时处理，从而加快了吞吐量。

要寻找什么：

- 验证底层平台是企业级的，能够支持关键任务负载，提供低延迟响应，并确保高可用性。
- 理想解决方案利用云基础设施和优化算法，以便添加更多数据或代理不会影响性能。
- 检查基准测试或案例研究，了解系统可以处理多少记录或域名，是否支持实时数据流，以及它是否可以在水平方向（更多处理能力/代理）和垂直方向（管理复杂数据模型和深层次结构）上进行扩展。



人工智能准备与整合

它意味着什么：

一个关键——尽管有时会被忽视——的标准是MDM平台和您的数据实践的人工智能就绪能力。人工智能就绪能力意味着该平台必须拥有将数据传输给人工智能模型和代理的必要集成能力（API、连接器）、提供上下文丰富的元数据架构，甚至可能还内置机器学习模型或访问高级基础模型（例如针对特定任务的大语言模型集成）。

它为何重要：

一项针对1203名数据管理领导者的调查，在2024年7月发现，那些未能认识到AI就绪型数据需求与传统数据管理之间巨大差异的组织，将危及他们AI工作的成功。³

要寻找什么：

- 评估 MDM 解决方案是否考虑了人工智能。一些现代工具提供无代码的人工智能界面，允许非技术用户配置代理或以自然语言查询数据。
- 同时也要考虑平台与您现有的AI或分析堆栈的集成程度。它能否连接数据科学平台或为产品分类或客户细分等功能接受预训练模型？
- 在组织层面，人工智能准备程度也取决于员工和流程成熟度，包括对数据管理员和IT团队进行培训，以便他们能够解释、调整并将人工智能代理的输出结果整合到业务工作流程中。
- 一个适合人工智能的环境可能包括用于安全测试的沙盒，或模拟模式来在全面部署前评估代理的影响。
- 最终，这一标准的核心在于您的MDM系统是否已准备好将关键功能委托给AI，同时保持强大的技术和管理基础，确保强大的代理MDM平台与成熟的数据文化相结合，能够自信地应对这些挑战。

持续数据治理与自我优化治理

MDM中代理式AI最强大的优势之一是其能够实现持续数据治理的能力。

在传统环境中，数据管理员会定期审核记录，或在出现问题时进行干预，但自主式人工智能系统则充当着全天候、不知疲倦的守护者，实时监控、清理和组织数据。

以比人类团队快得多、精确得多和规模大得多的速度、准确性和规模运行——而且不需要休息——这些代理可以实时捕捉和解决数据质量问题，将数据管理转变为嵌入式的、持续的系统功能，而不是定期的手动任务。

智能规则管理

这项自主管理的关键方面是智能规则管理。传统的MDM依赖静态业务规则（例如标准化地址或识别重复记录），而代理式AI系统则可以随着时间的推移学习和适应这些规则。AI代理不仅可以执行数据策略，还可以对其进行完善——检测新模式并相应地调整操作。代理式MDM“从过去的错误中学习，以持续提高准确性。”

例如，如果一个AI代理错误地合并记录，并且人类纠正了错误，一个设计良好的系统将结合该反馈以避免未来的错误，甚至根据持续的数据分析提出新的规则或匹配逻辑。

此外，这些代理并非应用一刀切的规则，而是能进行上下文感知处理——通过推断如何根据已知模式映射或标准化新字段来适应新的数据源和格式，所有这一切都不需要额外的编程。



通过自我优化治理的持续改进

自我优化治理是指系统通过反馈和分析，能够持续完善其治理流程——例如数据质量检查和合规性验证——的能力。AI代理的持续运行会生成大量元数据，这些元数据可以被用来精细调整治理策略。

例如，如果某个代理反复将某个特定数据字段标记为有问题，那么这种见解可以促使更新数据录入流程以防止未来出现的问题。

此外，先进的MDM解决方案现在采用预测分析来预见并解决数据问题，以避免其对下游流程产生的影响。在实践中，一个AI代理可能会检测到数据模式逐渐的变化（例如，关键字段空值激增），并自动向数据所有者发出警报或触发清理流程。到2026年，MDM解决方案将主动

使用人工智能和预测分析来预见和解决数据问题，通过减少60%的人工干预来确保数据完整性。⁴

由人类专业知识指导的自动化

至关重要，持续的管理和自我优化的治理并不会消除人类监督的需求——它们重新定义了它。数据管理员和治理团队从亲自清理数据转变为作为人工智能的监督者和导师，设定目标，定义高级规则，并在需要时审查人工智能驱动的决策。**人类监督对于验证关键变更和做出道德或战略判断仍然至关重要，即使代理式MDM自动化了大部分流程。**

尽管自动化功能强大，但在代理式主数据管理（agentic MDM）中，人工监督仍然至关重要。这种共生伙伴关系利用人工智能处理规模和复杂性的能力，同时依赖人类判断做出最终决策。当运作得当，这种由专家人工指导驱动、持续不断的人工智能管理，将生成始终可靠、随时可供任何业务或人工智能应用使用的主数据。

一个始终优化的MDM新纪元

通过结合实时数据质量管理、自适应规则学习和主动治理，代理式人工智能将master data management（mdm）转变为一个动态的、由人工智能主导的功能——它能够在复杂的领域中毫不费力地进行扩展，同时保持合规性和性能。随着生命科学组织面临不断增长的数据挑战和监管需求，采用这种变革性的方法不仅能够提高运营效率，还能够带来战略优势。

艾昆纬随时准备帮助您评估、实施并成功运用下一代智能数据治理。

为了探索自主AI如何提升您的MDM计划并创造持久的企业价值，请联系我们 iqvia.com/mdm

参考文献

1.

Forrester. (2024). *AI代理的状态，2024年。*

[AI代理的状态，2024 | 福瑞斯特](#)

2.

库兰, R., 约瑟夫, L., 霍普金斯, B., & 莱克莱尔, C. (2025, 3月11日)

代理式人工智能是下一个竞争前沿。

福雷斯特 [代理式AI是下一个竞争前沿](#)

3.

Gartner®. (2025年2月20日) Gartner. 缺乏人工智能 (AI) 准备数据使人工智能 (AI) 项目成为风险。[新闻稿]. Gartner. [缺乏AI准备数据使AI项目面临风险](#)

4.

Gartner®. (2023年8月7日).主数据管理解决方案市场指南.Gartner.

[主数据管理解决方案](#)

[市场指南](#)

GARTNER是Gartner，Inc.和/或其在美国和国际上的附属公司注册商标和服务标志，并许可可在此使用。版权所有。

关于作者



索瓦尼亚·布卡帕坦姆·蒂鲁马拉

高级总监，产品与战略，全球主数据管理，IQVIA

索瓦尼亚是一位经验丰富的产品管理高级总监，拥有超过17年在主数据管理（MDM）、数据治理以及生命科学、医疗保健、消费者健康和医疗科技领域的全球经验。目前，她领导IQVIA全球信息管理业务的战略与创新，推动开发符合不断变化的市场需求和客户期望的前沿解决方案。

她成功地领导了大型MDM和数据迁移项目，为顶尖制药公司提供解决方案，并在80多个国家进行了部署。她的专长涵盖数据建模、产品战略、跨职能协作和敏捷项目交付。作为一位公认的数字化变革大使，她推动了企业范围内的转型倡议和流程优化工作。

索瓦尼亚也处于将自主AI集成到MDM和数据治理的前沿，利用智能代理自动化数据质量、治理和合规监控——提高运营效率和决策。

她毕业于贝塔斯曼信息技术学院（BITS Pilani）电气与电子工程学士学位，获得哈佛商学院在线的“准备就绪证书”，并且是认证Scrum Master。



FRANCESCA D'ANGELO

信息管理业务总监，IQVIA

弗朗西丝卡在生命科学领域拥有20多年的经验

该行业，拥有15年以上的信息管理专业经验。她擅长科学、战略和协作的交叉领域，利用她的专业知识提供创新解决方案，推动科学进步并改善患者结果。

作为IQVIA全球信息管理业务的负责人，Francesca担任领域专家，为客户、内部团队和合作伙伴提供战略指导和建议支持。她的工作推动产品采纳和长期发展。

最近，她承担了将先进的AI模型整合到信息管理产品组合中的挑战，旨在提高生命科学组织中商业和预商业运营的效率。

联系我们
iqvia.com

