

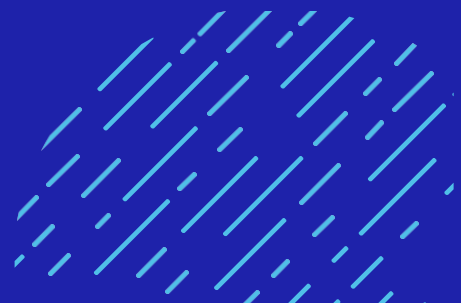


数据就绪：构建规模化医疗AI的坚实基础

HIMSS Market Insights 研究显示，数据就绪才是释放分析与AI潜能的真正关键。

过去几十年间，医疗数字系统的互联互通——即互操作性——已取得长足进步。“但我们逐渐意识到，互操作性的内涵远不止于数据的简单搬运。我们仍有很长的路要走，” InterSystems全球医疗市场战略负责人 Julie Smith表示。

Smith的观察得到了近期一项HIMSS Market Insights的支持。该项由 InterSystems赞助的研究，评估了美国机构的数据集成现状¹。调研深入揭示了为何数据交换不等同于数据就绪、云技术所扮演的角色，以及数据就绪如何为分析与AI项目提供有力支撑。





互操作性不仅关乎信息流转，更关乎信息抵达终端时的可用性。”

JULIE SMITH | InterSystems全球医疗市场战略负责人



数据交换不等于数据就绪

全美范围内，绝大多数医疗行业领导者（86%）一致认为，若缺乏数据准备环节，单纯的互操作性对分析与AI几乎毫无价值（见图1）。三分之二（67%）的受访者表示，当下多数互操作性挑战源于数据就绪不足，而非连接缺失。

70%的领导者主张，数据准备（datapreparation）应被视为互操作性的预期产出，而非一项独立工作。与此同时，关于互操作性的定义，行业内部存在明显分歧：35%倾向于狭义定义，即实现数据在系统间的可靠传输；36%则支持广义定义，即在数据传输基础上，进一步涵盖面向下游应用的清洗、标准化与增强处理。

对互操作性内涵认知的差异，往往会使组织内部关于数据就绪（datareadiness）的讨论复杂化。

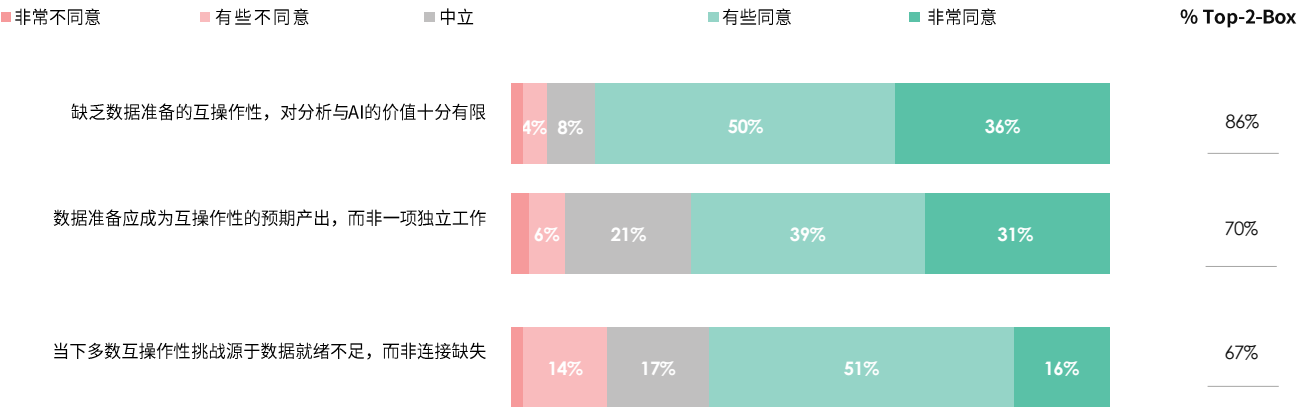
“早期的互操作性主要聚焦于就诊现场的个体患者，” Smith解释道，“但随着数据全面数字化，我们现在有机会将数据用于全人群健康管理。”

分析与AI工具正是应对这一拓展场景的理想选择。然而，Smith指出：“你必须拥有经过清洗、并以可用形式聚合的数据，才能从中提炼洞见、得出正确结论并制定后续行动方案。事实证明，若在数据输入这些工具前未加以筛选和清洗，它们给出的答案往往并不可靠。”

因此，在推进分析与AI应用的过程中，采纳更为全面的互操作性定义至关重要。“互操作性不仅关乎信息流转，更关乎信息抵达终端时的可用性，”她强调。Smith举了一个生动的例子：你用英语写一封信寄到法国。即便信件顺利送达，若收信人不懂英语，数据虽然在技术上完成了交换，却并未产生任何实用价值。

图1. 绝大多数领导者认同，互操作性与数据准备及数据就绪是相辅相成的，能够显著提升其分析与AI项目的产出价值。

请就以下陈述表明您的同意或反对程度。



云端助力，却非数据就绪的“万能药”

近四分之三（73%）的医疗行业领导者表示，即便数据上云，数据准备方面的挑战依然存在（见图2）。三分之二（64%）的受访者认同，云平台能够实现更高效的数据就绪自动化。但仅有逾三分之一（34%）的人认为，上云即可默认提升数据质量。

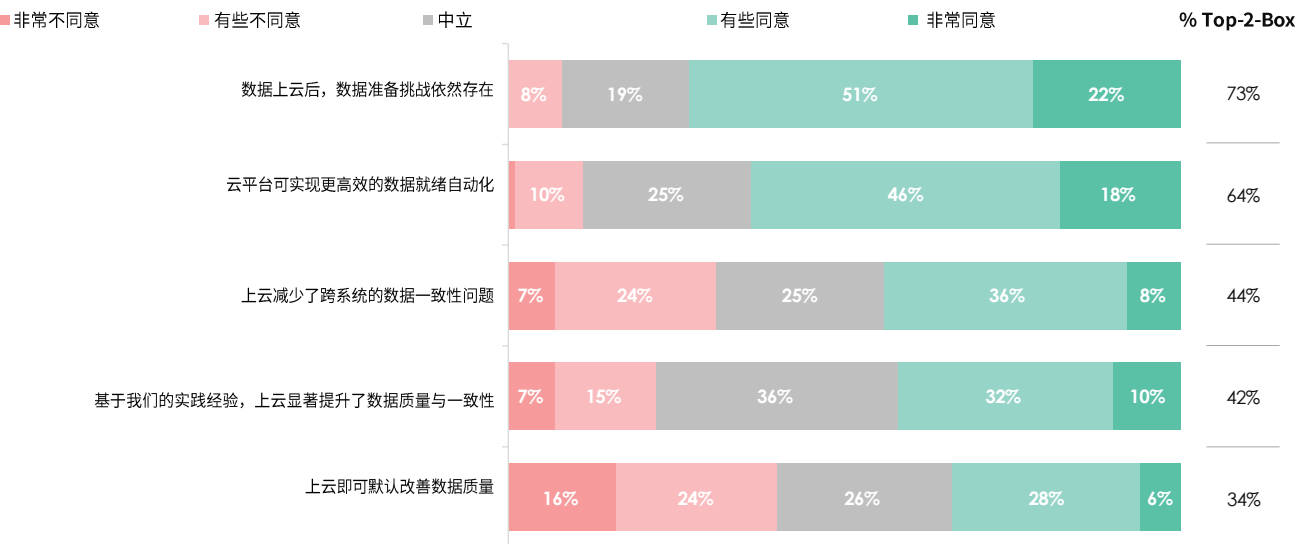
“有一种误解认为，只要上云，数据质量就会自动变好。其实上云除了改变数据存储位置，并不会对数据本身产生任何影响。” InterSystems平台战略与创新总监Jeff Fried指出。他强调，只有当组织采取有针对性的措施——包括启用特定工具与流程以提升数据就绪度——数据质量才会真正改善。

Fried补充道，数据上云也为组织解决数据就绪问题、深挖数据价值创造了新契机，例如：

- 无缝接入更多工具：包括用于提升数据就绪度的各类自动化工具；
- 拓展数据应用场景：支撑分析与AI等高级应用；
- 联动外部专家资源：获得专业支持，助力组织高效管理与利用数据。

图2. 近四分之三的医疗行业领导者认同，数据上云后，数据准备挑战依然存在；但多数人同时也认可，云平台有助于实现更高效的数据就绪自动化。

请就以下陈述表明您的同意或反对程度。



实现数据就绪的策略

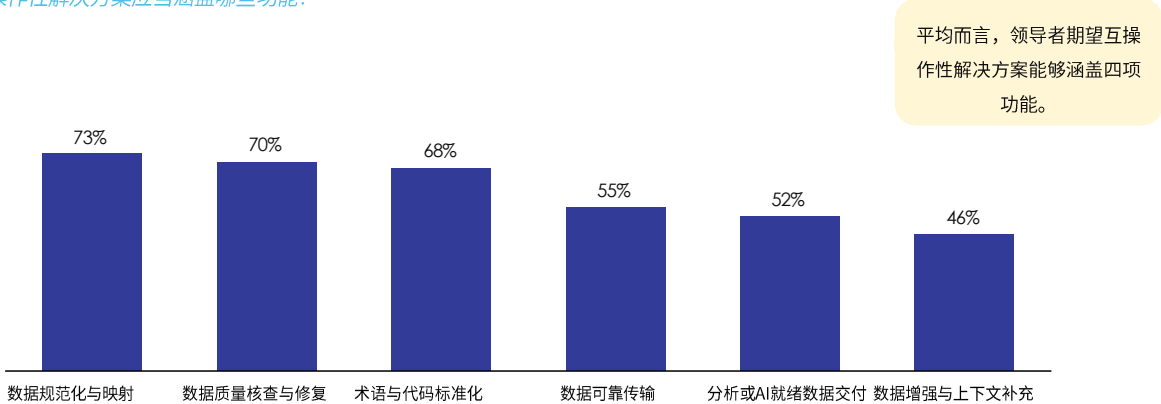
调研显示，领导者当前平均采用两种不同方式为分析与/或AI应用准备数据。57%的受访者依赖自定义脚本或内部自研工具；50%使用第三方ETL（抽取、转换、加载）或数据集成工具；39%依靠IT团队进行人工数据清洗。仅有9%的受访者采用了基本无需人工干预的高度自动化流程。

“遗憾的是，现实情况是，这些方法大多难以规模化扩展，” Smith指出，“我们面对的数据体量极为庞大。如果依赖手动脚本来准备数据，在分析与应用AI所需的规模面前，根本无从谈起。”

在评估包含数据就绪能力的互操作性解决方案时，领导者明确列出了这些方案应具备的具体功能（见图3）。近四分之三（73%）的受访者强调数据规范化与映射；70%看重数据质量核查与修复；68%关注术语与代码标准化；55%提及数据可靠传输；52%要求提供分析或AI就绪数据交付；46%重视数据增强或上下文补充。平均而言，领导者期望互操作性解决方案能够涵盖上述四项功能。

图3. 领导者最常期望互操作性解决方案具备数据规范化与映射、数据质量核查与修复，以及术语与代码标准化功能。

您认为互操作性解决方案应当涵盖哪些功能？



Fried指出，上述诸多能力目前虽已有单点解决方案可供选择——例如，市面上有专门解决术语与代码标准化的独立产品，也有专注于数据增强与上下文补充——即元数据管理——的单独工具。“组织面临的选择是：是采用多个独立产品，还是寻找一个囊括所有这些功能的统一平台？”他表示，“我认为，拥有一个能同时提供这六项能力的统一基础平台更为理想。这样不仅能减少系统组件和潜在的故障连接点，降低学习成本，还能避免数据在多系统间频繁流转带来的风险。”

Fried补充道，在评估具备数据就绪能力的互操作性解决方案时，组织必须确保方案植根于深厚的医疗行业专业知识。“医疗数据特有的广度、深度与精度，使得数据质量核查与规范化、数据增强及元数据管理变得极具挑战性，”他说道。

医疗领域知识不可或缺。举例而言，系统必须能识别出“怀孕”这一诊断不应出现在一名六岁男童的健康档案中。“正因如此，我建议医疗机构应该要求解决方案供应商必须具备医疗领域专业知识。”Fried强调。

实时数据的角色

超过四分之三（77%）的领导者认为，实时或准实时数据流对其组织下一阶段的分析与AI建设至关重要（见图4）。

Smith指出，“实时数据”的内涵与应用需视具体场景而定。例如，急诊室中患者出现心梗，相关数据必须实时捕获并触发警报，方能挽救生命；反之，若研究人员旨在回顾性分析某县的流感发病率，则无需实时获取数据。

“这取决于你要解决什么问题，”她说道。AI所需的数据体量，以及数据传输与分析的成本，都是必须考量的因素。“组织需要优先梳理亟待解决的业务场景，进而明确数据采集的频率与时效要求。”

“



上云除了改变数据存储位置，
并不会对数据本身产生任何影响。”

JEFF FRIED | InterSystems平台战略与创新总监



医疗机构应该要求解决方案供应商必须具备医疗领域专业知识。”

JEFF FRIED

评估组织的数据成熟度

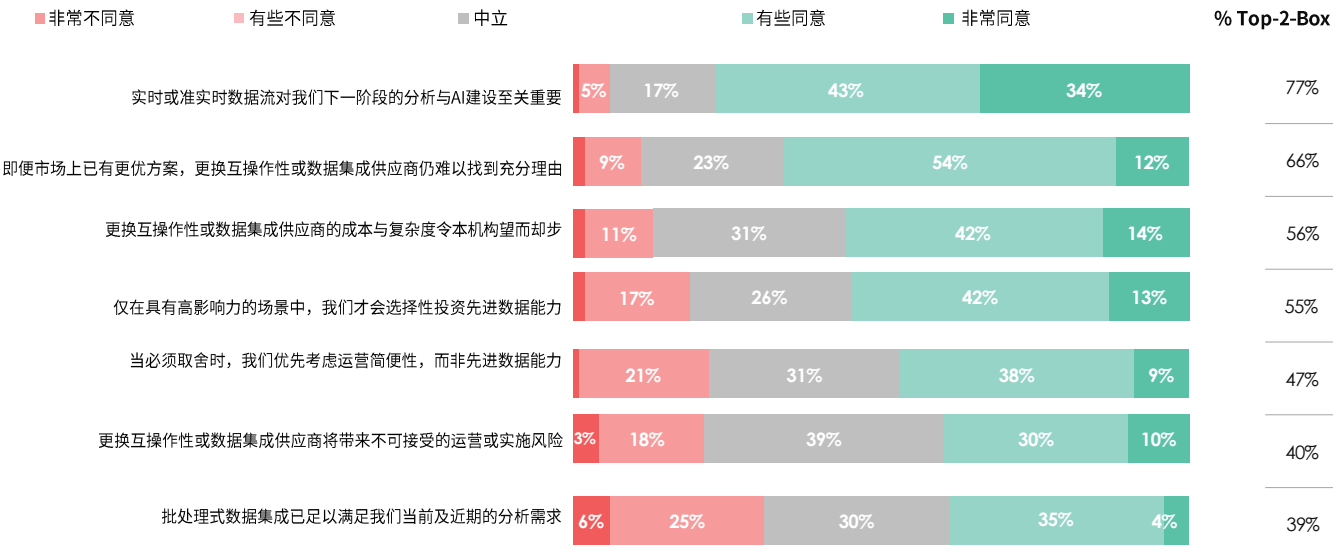
评估组织数据质量成熟度的第一步，是开展数据资产盘点。组织至少需明晰自身拥有哪些数据、数据源自何处，以及存储于何地。

完成数据盘点后，Smith建议，可参照图3所列的数据就绪能力，对标识现有能力部署情况与潜在缺口：

- 数据可靠传输：**原始数据是否已从各类临床数据源（如电子病历EHR、检验系统、药房系统）汇聚至集中环境或数据仓库？是否已部署接口，利用行业标准加速异构患者信息源及格式的接入？
- 术语与代码标准化：**临床代码与术语（如SNOMED CT、LOINC、ICD）是否已对齐至统一标准，确保不同系统“口径一致”？术语与代码标准化是否已实现自动化？
- 数据规范化与映射：**分散的数据结构是否已转换为统一格式（如HL7 FHIR标准），以确保跨系统互操作？各数据元素是否已映射至对应的目标字段？分散数据是否已聚合为单一、纵向的患者全景记录？
- 数据质量核查与修复：**是否已建立流程，用于识别并修复缺失值、重复项或异常值等问题，确保数据在各源头均准确、完整且一致？
- 数据增强与上下文补充：**是否已纳入社会健康决定因素（SDOH）、理赔数据及非临床信息等外部数据，为AI模型增值并提供更优语境？临床笔记等非结构化数据是否已完成索引？
- 分析或AI就绪数据交付：**最终处理后的数据集是否已具备适用于模型训练或实时洞察的优化格式？

图4. 约四分之三的受访者认为，实时或准实时数据流对其组织下一阶段的分析与AI建设至关重要。

请就以下陈述表明您的同意或反对程度。



“关于AI，我们已知的一个事实是：它能给出看似完美的答案，但这些答案往往是错误的。”

JULIE SMITH

Smith强调：“关于AI，我们已知的一个事实是：它能给出看似完美的答案，但这些答案往往是错误的。我们看到的层出不穷的研究都在印证这一点。因此，基于数据产生的结果必须经过客观的衡量与验证，才能满足医疗行业对确定性的严苛要求。”

尽管技术在数据质量与就绪工作中扮演着关键角色，Smith和Fried仍提醒各机构切勿忽视“人”在这一过程中的作用。Fried指出：“说‘数据质量是每个人的责任’很容易——事实也的确如此——但如果人人有责，往往就意味着无人负责。每家机构都应设立专人负责或专班推进数据质量工作，要把它变成某个人的‘本职岗位’。”

如需了解更多，请访问 [InterSystems 官网](#)

参考资料

1. HIMSS Market Insights, 《医疗数据集成调研报告》, 2026年2月 (由InterSystems赞助)。研究对象为美国医疗行业100位中高层领导者 (经理及以上), 且均需参与所在机构数据基础设施、互操作性或分析平台的选择或战略监督工作。InterSystems 未作为研究赞助方公开披露。



关于InterSystems

InterSystems 助力医疗机构在复杂环境中, 依托互操作性、集成与数据就绪能力, 充分释放数据价值。随着医疗机构与支付方正大力推进分析与AI战略, InterSystems 推动数据应用从“简单交换”迈向“可信、标准化、可即用”的新阶段, 为更明智的决策与更优质的诊疗结果提供坚实支撑。