



从“治理数据”到“数据价值”

——医院数据价值释放探索

陈金雄

2026年HIT设计者/工程师讲坛课件
版权归演讲者所有

目录



01

理清概念

02

治理策略

03

价值实现

04

前瞻展望

2026年HIT专家网工程师讲坛课件
版权归演讲者所有



01

理清概念

2026年HIT专家网工程师讲坛课件
版权归演讲者所有

要不要做数据治理？



蔡华法-AI转型 作者

已声明原创 >

广东 22 天前

先做好数据再搞 AI？大错特错！未来 AI 强大时不需要数据治理，AI 能去除噪音还原真实。企业常犯的策略错误：想把数据大海全部烧干，这样治理三年都做不好。正确思路：从用到治到建再到用的循环——AI 应用需要哪杯水就先用，不是清理所有数据。数据应从 AI 系统获取而非旧系统，质量更高代价更小。最终 AI 建设的过程，本身就可以补足数字化，Let it be，欢迎交流讨论！

#AI #人工智能 #AI转型 #蔡华法

认识数据治理

围绕业务场景和技术条件

联机事务应用：

如果没有患者主索引、药品主数据管理，临床业务怎么开展？

联系分析应用：

是否一定要做前结构化处理或前治理？差数据会影响应用吗？

病历的前结构化与病历的后结构化，在AI时代那种方式更好？

AI与数据治理的关系

Ai对数据治理肯定会带来巨大的价值。数据不是最怕脏，而是怕不真实，不全面。



宁义: ∞先做好数据再搞AI? 大错特错! 未来AI强大时不需要数据治理, AI能去除噪音还原真实...



不真实、缺胳膊少腿的数据，怎么用的？



即使用AI来数据治理，也是有成本的，成本应该还不低。如果能在源头尽可能规范数据，对业务本身还是非常有价值的。



昨天 13:05

你撤回了一条消息

很多数据是要实时驱动业务的，比如检查、检验结果、实时采集的生命体征等，不能等AI治理了再去用。过马路要看红绿灯和斑马线，如果红绿灯坏了，是敢过还是不敢过呢？离开具体业务场景和具备的技术条件，谈数据怎么治理，以及要不要治理，没办法谈到一起。



你撤回了一条消息

包括成本效益和应用的时效性等



为什么数据叫好不叫座？

- 应用目标不明确
- 基础数据不健全
- 数据治理不到位
- 医院参与不充分
- 应用效果不明显

医院端：数据没有发挥应有的价值

产业端：没有诞生有影响力的企业

数据是一种产品吗？

80%

LIS / PACS 专科系统

60%

HIS 核心业务系统

40%

数据层面

数据产品没有"一劳永逸"的解法

把数据能力寄托于"开箱即用、包治百病"的产品，并不现实

数据是一种服务

数据的本质是"服务"，不是可以简单采购的"产品"。

从数据采集、治理到应用的全流程，更多是"个性化服务"。

医院掌握主动权

- 理解业务场景
- 主导治理方向

企业提供赋能助力

- 规划咨询
- 解决方案
- 运营服务



02

治理策略

2026年HIT专家网工程师讲坛课件
版权归演讲者所有

数据赋能的核心：数据-业务-场景 相融合

赋能业务（业务场景、业务痛点、业务价值）

为应用，业务场景，好食客

数据业务化（汇聚、治理、数据资产、组件化、API）

后治理，数据处理，好厨师

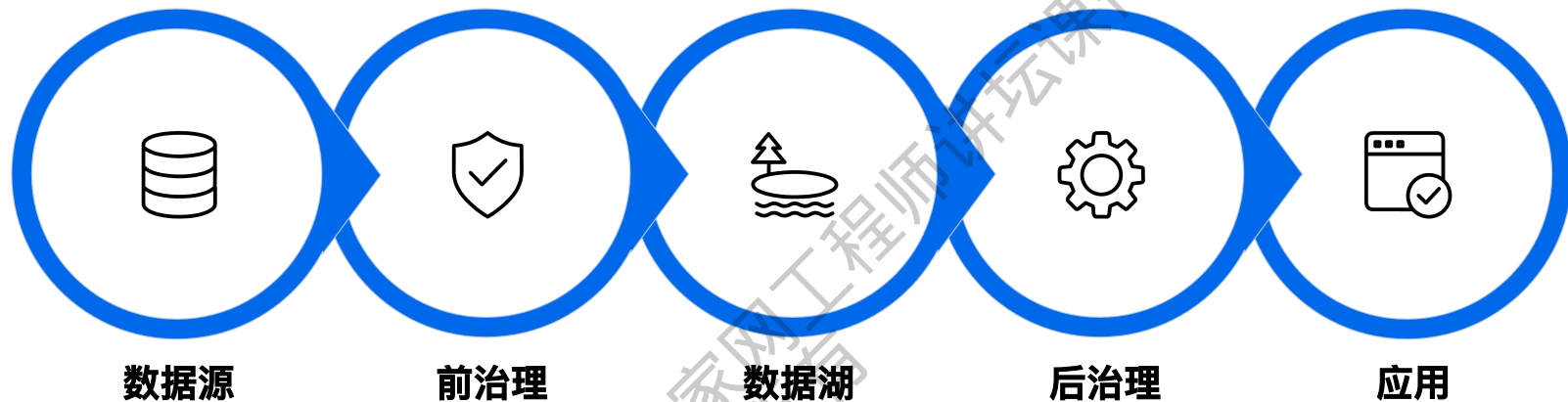
业务数据化（主索引、主数据、数据元素、数据链）

前治理，数据基础，好食材

普遍误区：非常看重后治理，指望找个好厨师做出满汉全席既不重视食材好坏，也不关注食客到底喜欢吃什么。

正确做法：医院自己主导，做好前治理，明确业务场景、痛点、需求和目标
后治理其实成本很高，不一定全部治理，根据需要来定制化治理寻找合作伙伴，一起合作来做这些事情，而不是包治百病的产品

数据治理全景图



- 主索引、主数据管理：以前治理为主；后治理承担必要的标准桥接与补齐。
- 前治理：尽可能全量，建设主数据、修复系统、规范业务、夯实数据基础。
- 数据湖：原始数据全量沉淀，先不用，但不丢。
- 后治理：围绕具体业务需求和业务场景建模、ETL，不追求全量治理。
- 应用：临床、管理、服务、科研、AI。

前治理全量，数据湖兜底，后治理按需，三层协同释放数据价值。

数据前治理

目的：改进系统与管理的源头工程，实时支撑联机事务应用。

- 以数据为镜，发现并修复生产系统缺陷。
- 规范业务流程，改进医院管理。
- 让数据"出生即标准"：准确、一致、完整、有效。
- 尽可能在源头消除主数据、标准、数据链等基础问题。

01

系统修复

修复生产系统缺陷

02

数据提质

确保准确、一致与完整

03

工作提效

优化流程提升效率

04

业务优化

聚焦业务价值与管理

前治理做好，后治理才能把主要精力放在业务价值，而不是反复处理基础性数据问题。

"三个1"治理体系

一套治理标准

让数据质量可度量、可管理。

一个质控平台

让治理过程自动化、常态化，
7×24 小时在线值守。

一套运行机制

依靠响应处理、根源分析和
持续改进形成闭环。

标准、平台、机制协同联动，形成数据管理闭环。

一套治理标准——建立指标体系

DAMA治理体系 × 医疗行业理解

横向五维质量 + 纵向主数据实体 = 可执行质检规则

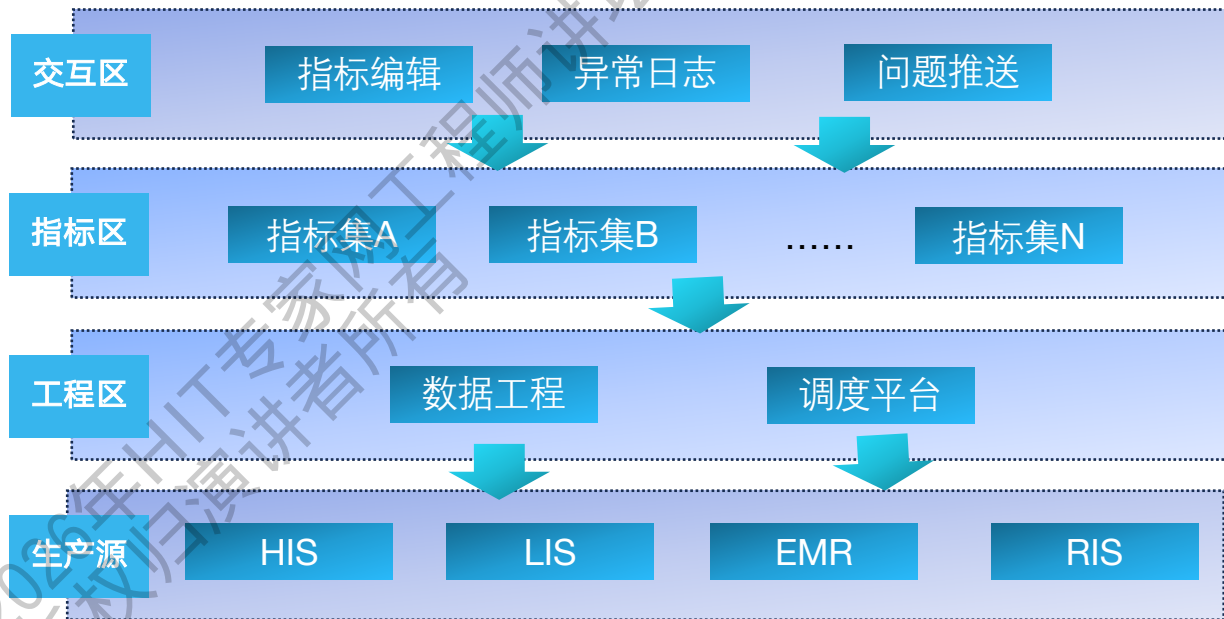
业务流	完整性	准确性	一致性	唯一性	有效性	典型异常案例
患者流	关键字段齐全	身份信息相符	跨系统一致	一人一档	编码格式合规	收费记录关联不到医嘱
医嘱流	全周期完整	执行内容相符	系统流转一致	医嘱号唯一	频次途径合规	医技类收费记录关联不到设备
供应链	采购消耗可追溯	库存实物相符	出入库一致	一物一码	编码有效合规	账单无法找到对应收费明细项目
费用流	明细完整可追溯	金额计算准确	账财医保一致	流水号唯一	规则编码合规	退费明细无法找到对应正向收费明细
主数据	标准目录完整	标准映射准确	版本保持一致	编码全局唯一	标准状态有效	
主索引	业务记录可关联	匹配算法准确	跨院区映射一致	主索引全院唯一	生成规则合规	
数据展示	指标模块齐全	前后台口径一致	多端展示一致	无重复渲染	格式规范	
费用流 × 完整性: 关联关系缺失						质检规则的价值 发现异常 → 追溯根因 → 改进系统与流程

标准不是表格里的标签：横向五维质量与纵向主数据实体的交叉点，才能形成可执行的质检规则。

一个建设质控平台

质控平台框架示意图

- 脚本管理
- 定时查询
- 异常侦测
- 问题推送



一套运行机制

自动侦测

定时探查并识别异常数据。

问题推送

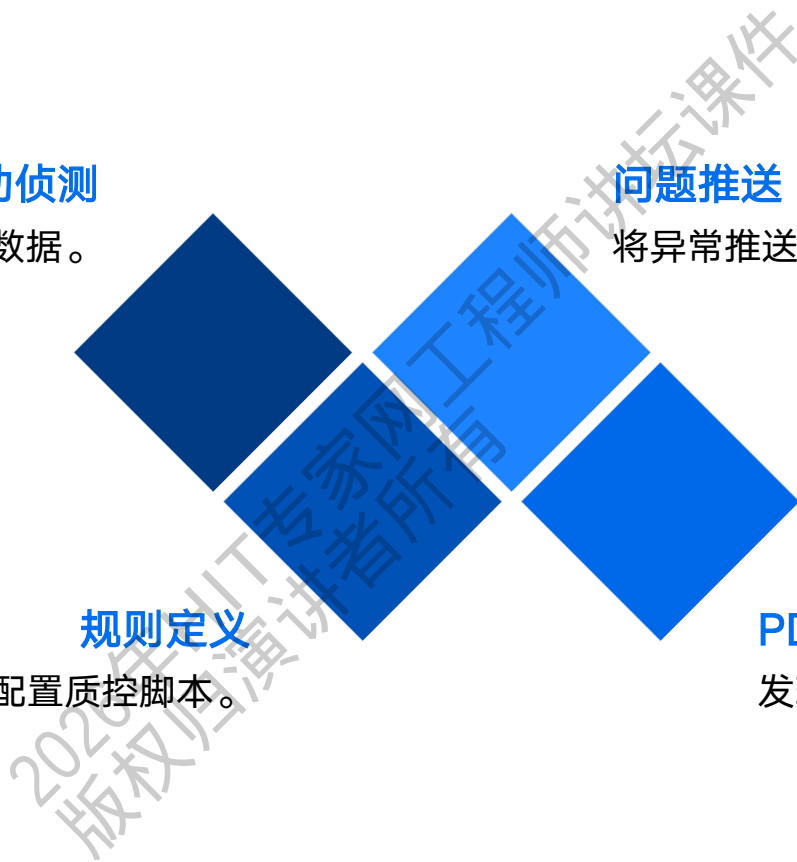
将异常推送责任人处理回溯。

规则定义

多方协作配置质控脚本。

PDCA闭环

发现、分析、改进并验证效果。



前治理成效

35%

耗时缩短

报表开发与数据校对耗时同比缩短 35%。

57%

异常减少

数据异常发生率较上季度减少 57%。

22

缺陷发现

发现系统缺陷与接口问题 22 处，已修复 19 处。

20%

上报减少

临床上报系统问题数同比减少 20%。

200

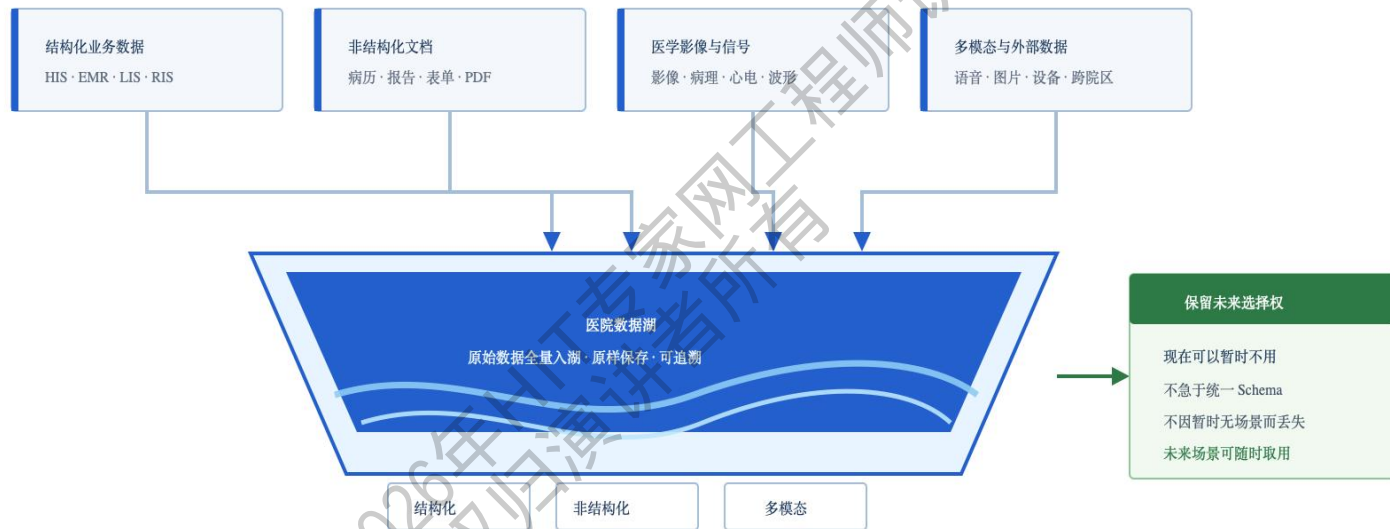
质控指标

200 个在用质控指标，每年 60 余次发现异常，覆盖 2 个生产系统。

某医院推动前治理后，2024年下半年到2025年上半年数据，2026年以来基本没有报错。

数据湖：先不用，但不丢

原始数据全量沉淀，保留未来医院业务创新的选择权



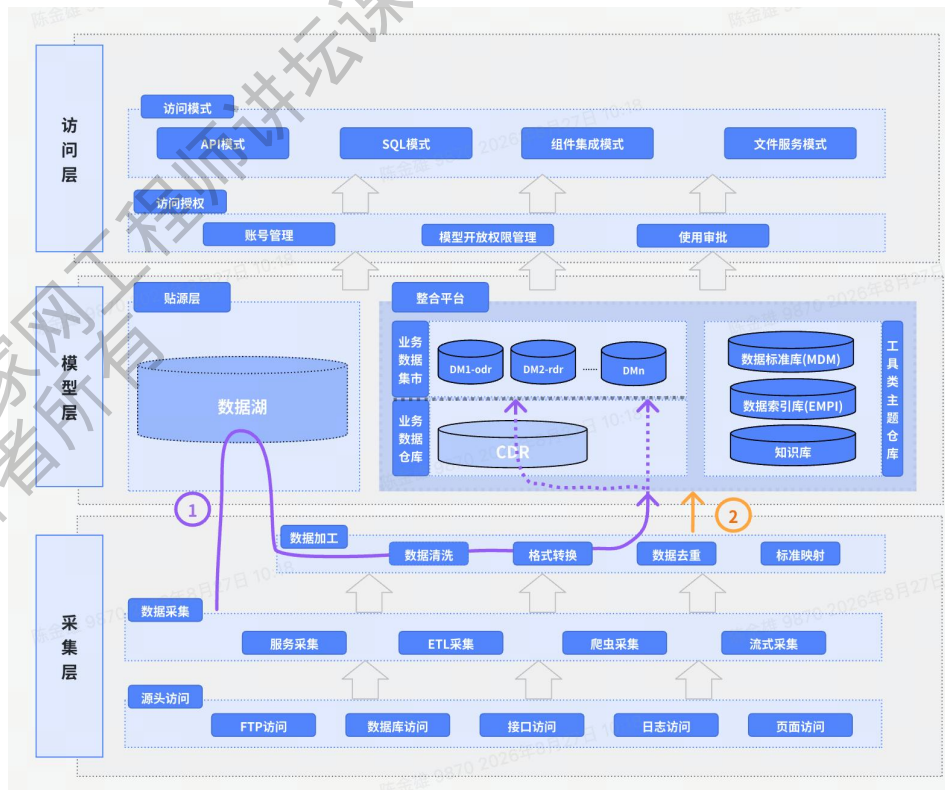
先沉淀原始资产，再在明确业务场景下按需建模和加工

数据后治理：场景建模，按需治理

- 哪个场景需要，就为哪个场景建模。
- 抽取 → 转换 → 加载，产出临床、科研、运营主题数据集。
- 没必要全量治理；
- 后治理的目标是业务价值，不是重复处理基础问题。

注明：

后治理既可以通过数据湖，也可以从源头直接采集。



复杂肿瘤患者智能导诊：基于历史病历的专家匹配

➤ 业务问题

复发难治患者往往携带多家医院的历史病历、影像、病理、检验和治疗记录。如何快速了解患者既往诊疗经历，辅助判断是否适合本院进一步诊疗，并推荐匹配的专科专家？

➤ 案例价值

- 减少患者重复描述和资料反复整理；
- 帮助医院快速识别适合进一步评估的患者；
- 提高患者与专科专家的匹配效率；
- 为专家接诊和多学科会诊提供结构化依据。

➤ 技术挑战

- 患者资料治理：OCR/AI 提取病情阶段、既往治疗、疗效与风险特征
- 院内多院区临床数据治理：统一患者、疾病、诊疗、专家等主数据和口径
- 双侧标准桥接：患者病情特征与专科能力、专家专长、诊疗经验映射
- 患者—专科—专家匹配推荐
- 为接诊准备和多学科会诊提供参考

通过数据后治理，满足数据应用的质量规范要求

挑战

1.病历语种翻译统一

2.病历要素拆解建模

3.数据标准转化统一

4.数据质量提升

处理

AI辅助处理/人工审核调整

按专病要素建立知识库

设定核心标准/完成项目映射

多源数据印证修补

成效

提升可读性/降低信息提炼难度

自由文本转为结构化信息，便于检索

知识内容口径统一，优化数据利用速度

修复历史数据缺陷，提升数据价值



03

价值实现

2026年HIT专家网工程师讲坛课件
版权归演讲者所有

价值一：复杂资料变成病情摘要（临床驾驶舱）

当前最稳定的业务能力，是让医生更快看懂复杂患者

资料接入

PDF / 图片 / Word / 多语言
OCR、医学翻译、原文留痕

病例画像

基本信息、诊断、风险标签
关键指标统一呈现

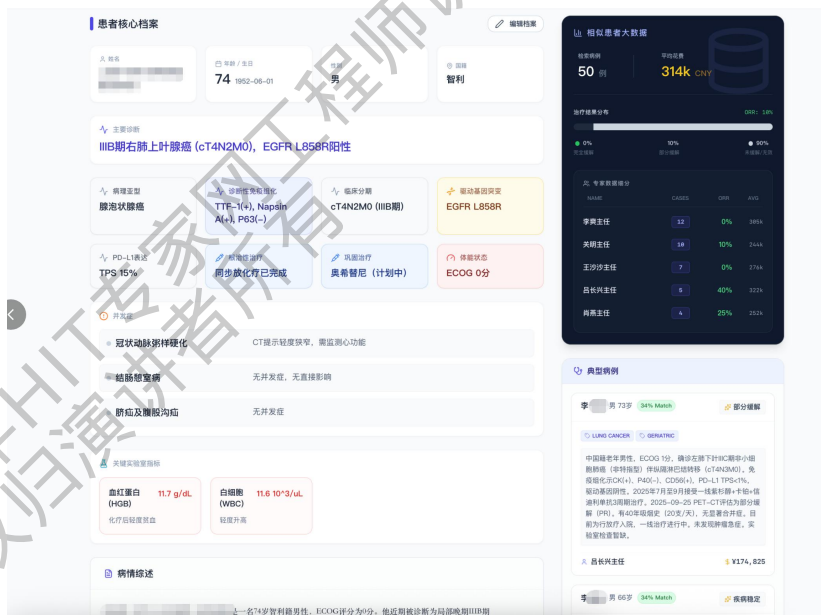
证据综述

多份资料归并成摘要
来源引用，支持人工修订

时间轴与缺项

诊断 → 治疗 → 评估 → 复发
提示还需要补充的资料

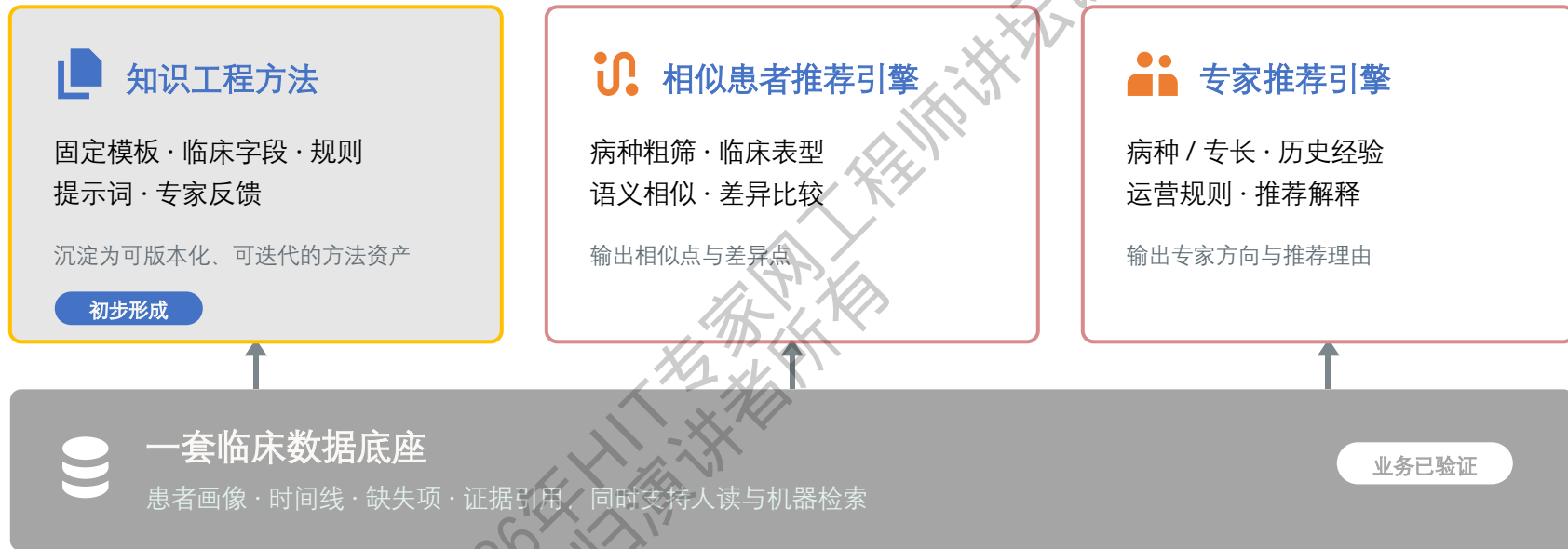
真实运行界面·病历整理工作台



客服确认后

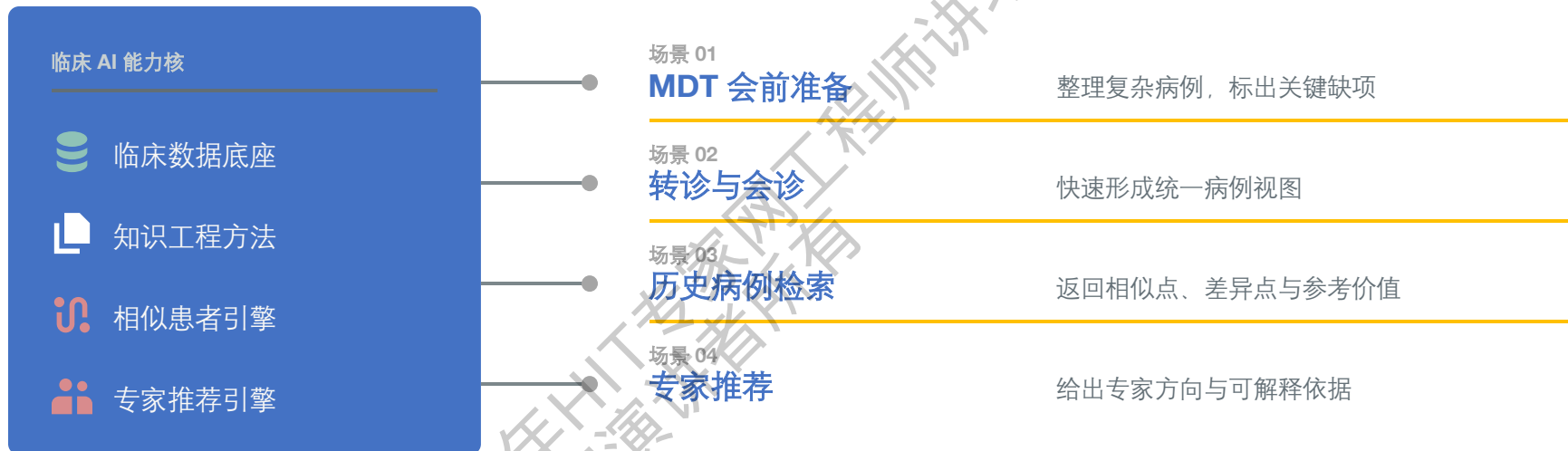
复制摘要 / 导出 DOCX

价值二：四类可复用能力



价值三：能力延展，从“客服”变成“临床场景”

同一套底座、方法与引擎，可以嵌入现有业务，不复制整套系统。



辅助信息理解与经验检索，不替代医生诊断和治疗决策。

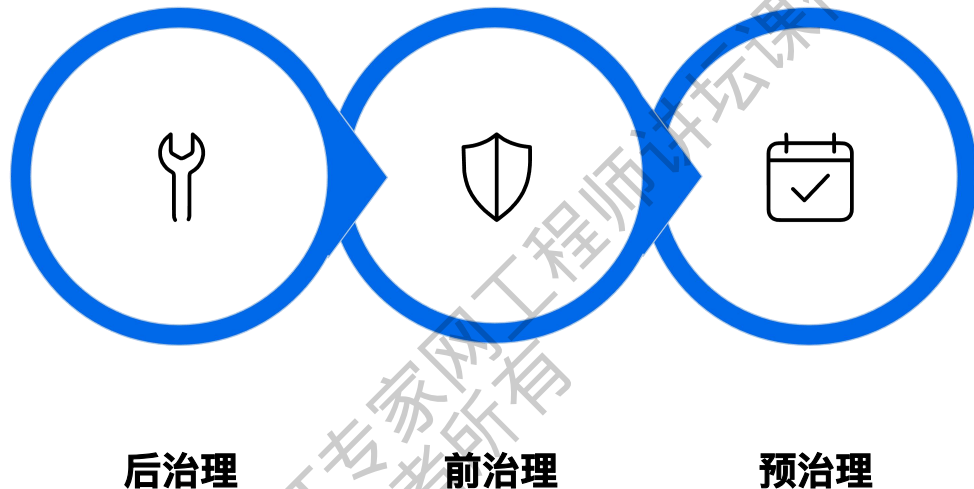


04

前瞻展望

2026年HIT专家网工程师讲坛课件
版权归演讲者所有

预治理：AI 时代，把数据质量前移到软件测试阶段



预治理发生在软件测试阶段，在数据真正进入生产系统之前验证数据规范，它有两个目的：

1. AI 编程时代，代码越来越不可见，无法像过去一样逐行阅读和验证；数据验证因此成为新的质量闸门。
2. 在前治理之前解决数据质量问题，把不规范数据扼杀在萌芽状态；治理越早，成本越低，效果越好。

对工程师的启示：除功能验证外，还要把数据规范、数据链和数据结果纳入测试标准。

越早治理，成本越低，效果越好。

结语

从治理数据到数据价值

认知

数据是服务，医院主导。

治理

预治理 → 前治理 → 后治理。

价值

场景牵引，多模态资产。

感谢倾听！

