

T/WEA

团 体 标 准

T/WEA XXXX—2026

医院大数据与人工智能一体化建设规范

Specification for integrated construction of big data and AI in
hospitals

(征求意见稿)

2026 - XX - XX 发布

2026 - XX - XX 实施

中国智慧工程研究会 发布

目 次

前言 III

引言 V

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 总体原则与建设框架 4

5 数据治理体系与数据资源建设要求 7

6 数据平台与服务共享能力建设要求 11

7 人工智能平台与模型全生命周期管理要求 15

8 场景应用建设与评估验收要求 19

9 安全与合规保障要求 24

10 运行维护与持续改进 25

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国智慧工程研究会提出并归口。

本文件起草单位：

本文件主要起草人：

全国团体标准信息平台

引 言

随着医疗健康领域数字化转型持续深化，医院信息化建设正从“业务系统上云上网、数据汇聚共享”逐步演进为以数据要素为核心驱动的治理与应用体系建设。医院大数据与人工智能（以下简称“AI”）的融合应用，正在成为提升医疗质量安全、优化诊疗效率、强化运营精细化管理、支撑科研创新与区域协同的重要技术路径。与传统单点系统建设不同，大数据与AI一体化建设强调以统一数据底座与治理体系为前提，通过标准化的数据采集与汇聚、规范化的数据管理与服务、可控可审计的算法与模型全生命周期管理，以及面向临床、管理与科研的场景化应用闭环，实现“数据—知识—决策—行动—反馈”的持续优化。

当前医院在推进大数据与AI建设过程中，普遍存在以下问题：一是数据来源多、系统异构程度高，数据标准不统一、主数据不一致、口径不一致导致数据可用性不足，难以支撑跨系统、跨院区、跨机构的可靠分析与建模；二是数据治理体系不健全，数据资产目录、质量规则、元数据管理、血缘追溯、权限分级与共享审批等机制缺失或落地不足，造成“数据可汇聚但不可用、可用但不可控”的风险；三是AI应用“重模型、轻治理”，模型训练数据不可追溯、性能评估不完整、上线后漂移监测与回退机制缺乏，临床可解释性不足，难以形成可信可控的闭环应用；四是安全与合规压力显著，个人信息与健康信息的处理活动涉及范围广、链路长，若缺乏全域的身份认证、访问控制、审计留痕、脱敏与共享治理，易引发合规与安全风险；五是建设路径与组织机制不清晰，业务部门、信息部门、科研部门与供应商之间协同不足，导致重复建设、能力割裂与投入产出不匹配。

为指导医院规范开展大数据与AI一体化建设，本文件围绕顶层架构、数据治理、平台能力、算法与模型治理、安全合规、运行维护与持续改进等关键要素，提出统一的建设要求与实施要点，旨在为医院规划建设、项目实施、能力验收与运行运营提供可遵循的规范依据。本文件强调以业务需求为牵引、以数据治理为基础、以AI全生命周期管理为核心、以安全合规为底线，通过标准化、体系化方法提升医院数据与智能能力的可持续运营水平。

本文件所称“医院大数据与人工智能一体化建设”，是指在统一的技术与治理框架下，面向医疗服务、医技药事、护理与质控、运营管理、科研与教学、区域协同与监管报送等业务领域，构建数据汇聚与治理平台、数据服务与分析平台、AI开发训练与部署平台以及场景应用体系，并建立覆盖组织制度、流程规范、技术标准、安全合规与评估改进的整体体系。本文件不以具体产品形态、厂商实现方式为限定，重点规定能力要求、治理要求与可验证的关键控制点，支持不同规模医院按分期路线推进建设。

使用本文件时，医院宜结合自身发展定位、学科特色、信息化基础与数据资源状况，开展现状评估与能力差距分析，明确建设目标与阶段成果，并在实施过程中建立跨部门协同与持续改进机制，确保建设成果可落地、可度量、可运营。

医院大数据与人工智能一体化建设规范

1 范围

本文件规定了医院大数据与人工智能一体化建设的总体原则、建设目标与能力框架，明确了数据治理体系、数据资源与数据标准、数据汇聚与存储处理、数据服务与共享交换、分析建模与可视化、人工智能平台能力、算法与模型全生命周期管理、场景应用建设与评估验收、运行维护与持续改进，以及安全与合规保障等方面的要求与实施要点。

本文件适用于各级各类医院（含综合医院、专科医院、中医医院及其他具备临床诊疗服务能力的医疗机构）开展院内大数据与人工智能平台的一体化规划、建设、改造与运行管理，也适用于医院集团、医联体牵头单位对成员单位开展一体化能力建设指导与协同治理的相关工作。对新建医院，本文件可用于顶层设计与分期实施；对存量医院，本文件可用于能力现状评估、差距分析、平台整合与能力升级

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 1.1—2020 标准化工作导则 第1部分：标准化文件的结构和起草规则

GB/T 22239—2019 信息安全技术 网络安全等级保护基本要求

GB/T 35273—2020 信息安全技术 个人信息安全规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

医院大数据 hospital big data

医院在医疗服务、医技药事、护理质控、运营管理、科研教学与保障运行等活动中产生或汇聚的多源异构数据集合，具有规模大、类型多、更新快、价值密度相对低但可通过治理与分析挖掘价值等特征。

3.2

一体化建设 integrated construction

在统一的组织治理、技术架构、标准体系与安全合规框架下，对数据汇聚与治理平台、数据服务与分析平台、人工智能开发训练与部署平台以及场景应用体系进行协同规划、建设与运营，实现能力复用、口径一致、链路可追溯与持续迭代。

3.3

数据治理 data governance

为保障数据可用、可信、可控与可持续而建立的组织制度、流程规范与技术机制的集合，通常包括数据标准、主数据、元数据、数据质量、数据安全、数据共享、数据资产管理与数据生命周期管理等内容。

3.4

数据标准 data standard

为实现数据语义一致与交换共享而制定的规范集合，包括数据元定义、编码体系、数据字典、指标口径、接口字段定义与消息规范等。

3.5

数据元 data element

用于描述信息交换或数据管理的最小数据单元，通常由名称、定义、数据类型、表示格式、取值范围、编码规则等属性构成。

3.6

元数据 metadata

用于描述数据的数据，包括数据的定义、结构、来源、口径、质量规则、权限属性、更新频率、责任人、血缘关系等信息。

3.7

主数据 master data

在多个业务系统间共享使用、相对稳定且需统一管理的关键基础数据集合，通常包括患者、人员、机构科室、诊疗项目、药品、耗材、检验检查项目、设备资产等对象的基础信息及其编码映射关系。

3.8

数据质量 data quality

数据在准确性、完整性、一致性、及时性、唯一性、有效性等方面满足既定业务需求与规则约束的程度。

3.9

数据资产目录 data asset catalog

对数据资源及其属性（包括主题域、数据集、口径、责任人、共享等级、服务接口、质量状态等）进行统一登记、检索与管理的目录体系。

3.10

数据服务 data service

将数据以可复用、可治理、可监测的方式对外提供的能力形态，包括数据查询服务、数据接口服务、指标服务、标签服务与数据订阅推送服务等。

3.11

数据湖 data lake

面向多源异构数据的集中存储与管理方式，可保留数据原始形态并支持按需加工与多样化分析计算，强调弹性扩展与多类型数据承载能力。

3.12

数据仓库 data warehouse

面向主题域分析的集成数据存储与管理方式，通常通过建模、清洗、汇总与口径固化形成面向分析的结构化数据集合，强调口径一致与可复核性。

3.13

机器学习 machine learning

通过数据学习规律并用于预测、分类、聚类、推荐等任务的算法方法集合，模型性能通常通过训练数据与评价指标进行度量。

3.14

深度学习 deep learning

以多层神经网络为代表的机器学习方法，常用于医学影像识别、文本理解、时序预测等场景，具有特征自动学习能力强、对数据与算力依赖较高等特点。

3.15

特征工程 feature engineering

对原始数据进行清洗、转换、编码、衍生与选择，以形成可用于模型训练与推理的特征集合的过程。

3.16

特征库 feature store

对特征定义、计算逻辑、版本、血缘、复用与在线离线一致性进行统一管理的能力体系，为模型训练与推理提供一致的特征服务。

3.17

模型 model

通过算法训练得到的参数化或规则化表达，用于对输入数据产生预测、分类、评分或建议输出，可包括统计模型、机器学习模型与规则模型等。

3.18

训练 training

使用训练数据对模型参数进行学习与拟合的过程，通常需要明确数据来源、样本定义、标签口径、训练策略与超参数配置。

3.19

验证 validation

在训练过程中或训练后，使用验证数据对模型进行调参与性能评估的过程，用于选择模型与防止过拟合。

3.20

测试 testing

在独立于训练与验证的数据集上对模型性能进行最终评估的过程，用于衡量模型在未见数据上的泛化能力。

3.21

部署 deployment

将模型以服务、组件或嵌入方式上线到目标运行环境（如院内业务系统、影像工作站、数据平台等）并支撑实际业务调用的过程。

3.22

推理 inference

已部署模型在运行时对输入数据进行计算并输出结果的过程，通常对时延、稳定性、可用性与审计留痕有要求。

3.23

模型全生命周期管理 model lifecycle management

对模型从需求提出、数据准备、训练验证、测试评估、上线部署、运行监测、更新迭代到退役下线的全过程实施规范化管理的活动集合。

3.24

可解释性 explainability

模型输出能够以可理解方式说明其关键影响因素、证据与推理路径的特性，用于支持临床人员判断、降低误用风险并满足合规与审计需要。

3.25

机器学习运维 machine learning operations, MLOps

指在机器学习模型的全生命周期内，综合应用软件工程、数据工程与运维工程的技术与管理体系，实现从数据采集、模型开发、验证、部署到监控和迭代更新的持续集成与持续交付过程。

4 总体原则与建设框架

4.1 总体原则

4.1.1 业务牵引与价值导向

建设应以临床诊疗、护理与质控、医技药事、运营管理、科研教学与区域协同等真实需求为牵引，明确优先级与阶段性目标，避免“为平台而平台”。应用场景应形成可验收的闭环条目，包含输入数据、输出结果、触发条件、责任角色、处置路径与效果指标。

4.1.2 数据治理优先与标准统一

应先建立数据标准体系、主数据与元数据管理、数据质量规则与血缘追溯机制，再推进数据资产服务化与AI建模应用。数据口径、编码映射与指标定义应统一管理，避免多口径并存导致分析结论不可复核、模型训练不可重复。

4.1.3 统一架构与能力复用

应采用平台化、服务化架构构建数据与AI能力，减少点对点接口与烟囱式系统重复建设。数据汇聚、治理加工、数据服务、特征服务、模型服务、监测审计等共性能力应统一沉淀与复用。

4.1.4 全生命周期管理与可信可控

应对数据与模型实施全生命周期管理。数据侧应覆盖采集、加工、发布、共享、归档与销毁；模型侧应覆盖需求、数据准备、训练验证、测试评估、上线部署、运行监测、更新迭代与退役下线。关键过程应留痕可追溯，确保可审计、可回退。

4.1.5 安全合规内生性

应将网络安全、数据安全、个人信息保护、访问控制、审计留痕、脱敏与共享审批等要求融入架构设计与能力实现，并在运行阶段持续验证有效性，避免“建成后再补合规”。

4.1.6 分期实施与持续运营

应根据医院信息化基础、数据资源与组织能力，分期推进建设。宜遵循“先数据底座与治理、后服务化共享、再AI平台与应用”的路径，并建立运营组织、指标体系与持续改进机制，实现由“项目交付”向“能力运营”转型。

4.2 建设目标与总体架构

4.2.1 建设目标

医院大数据与人工智能一体化建设目标应至少包括：

- a) 形成统一的数据汇聚与治理底座，实现多源异构数据的标准化接入、统一治理与质量可控；
- b) 形成统一的数据服务体系，实现数据资产目录化、服务化与可监测调用；
- c) 形成统一的AI开发与运行平台，实现数据—特征—模型—服务的工程化闭环与全生命周期管理；
- d) 形成可复制的场景应用体系，实现临床与管理关键场景的规模化落地与效果可评估；

e) 形成安全合规与审计追溯体系，实现“可授权、可控制、可审计、可追责”。

4.2.2 总体架构分层

医院一体化建设总体架构宜按“数据源层—接入汇聚层—存储计算层—治理加工层—服务共享层—AI平台层—应用层—运营安全层”的逻辑分层组织，各层职责边界应清晰，并通过统一标准、统一身份与统一审计实现贯通。该分层结构应与图1相符合。



图1 医院大数据与人工智能一体化建设总体架构示意图

4.3 能力框架与建设内容

医院大数据与人工智能一体化建设能力框架宜覆盖数据侧、AI侧与运营侧的关键能力。能力框架应便于形成建设清单、验收要点与运营指标，并支撑分期推进。一体化建设能力框架与关键能力要素见表1。

表1为示例结构，医院可结合自身实际扩展子能力或细化要素，但能力域应覆盖数据治理、服务共享、AI工程化与安全合规等关键内容。

表1 一体化建设能力框架与关键能力要素

能力域	子能力	关键能力要素示例
数据接入与汇聚	多源接入	标准接口/消息接入；批量/实时采集；接入校验；增量同步
存储与计算	湖仓一体/分层存储	数据湖承载多类型；仓库固化口径；实时/离线计算；资源弹性
数据治理	标准与主数据	数据标准、编码映射；主数据统一；元数据登记；血缘追溯
数据质量	质量规则与监测	缺失/重复/逻辑校验；质量评分；问题工单闭环；整改验证
数据服务与共享	目录与服务化	数据资产目录；数据API/指标服务；订阅分发；共享审批与留痕
AI平台与工程化	训练部署与运维	数据准备/标注；特征库；训练验证；模型库；推理服务；漂移监测
场景应用	业务闭环落地	临床/质控/运营/科研场景；触发与处置；效果评估；回退机制
运营与安全合规	贯穿治理	统一身份与权限；审计留痕；分级分类与脱敏；安全事件响应；合规审计

4.4 分期建设路线与依赖关系

- 4.4.1 第一阶段：数据底座与治理体系建立。以数据接入汇聚、数据标准与主数据、元数据与血缘、基础质量规则与审计能力为重点，确保数据“可汇聚、可识别、可追溯”。
- 4.4.2 第二阶段：数据服务化与分析能力增强。以数据资产目录、指标口径固化、数据API/指标服务、订阅分发与共享审批为重点，推动数据由“可用”向“可复用、可规模化供给”转变。
- 4.4.3 第三阶段：AI平台工程化与重点场景落地。以特征工程与特征库、训练验证流水线、模型库与部署推理、监测回退与可解释性为重点，优先落地高价值、可闭环的场景应用。
- 4.4.4 第四阶段：协同扩展与持续运营优化。面向医联体、区域协同、科研平台与监管报送等扩展应用，加强跨机构数据共享治理与联合建模能力，并通过运行监测与评估机制持续优化。
- 4.4.5 分期路线的依赖关系应清晰。未建立统一口径、血缘追溯与安全审计能力前，不宜直接开展高影响的临床AI上线应用，以降低误用风险与合规风险。

5 数据治理体系与数据资源建设要求

5.1 治理组织、职责与工作机制

医院应建立院级数据治理组织体系，形成“统一领导、分级负责、业务主导、技术支撑、安全合规贯穿”的治理机制。数据治理组织应至少包括决策层、管理层与执行层，并与医疗质量管理、信息安全管理、科研管理等体系协同。

数据治理职责宜至少覆盖以下内容：

- a) 数据治理决策与制度制定：明确治理目标、范围、制度体系与考核机制；
- b) 数据标准与口径管理：统一数据元、编码体系、指标口径与接口字段语义；
- c) 主数据管理：统一关键对象编码、映射、分发与变更控制；
- d) 元数据与血缘管理：登记数据来源、加工规则、口径版本与影响分析；

- e) 数据质量管理：制定质量规则、监测指标与整改闭环；
- f) 数据资产管理：建立数据资产目录、共享等级与数据服务目录；
- g) 数据共享与对外提供管理：建立审批、授权、脱敏与留痕机制；
- h) 安全与合规管理：数据分级分类、访问控制、审计追溯与风险处置。

为保证治理可落地，医院应建立常态化工作机制，至少包括：数据治理例会与专题会、数据标准评审与发布流程、主数据变更流程、数据质量问题工单闭环机制、数据资产上架与下架机制、共享审批与对外提供登记机制、年度评估与持续改进机制。

如需以清单化方式固化治理职责与交付物，可按示例1组织治理要素。医院可结合实际设置岗位与组织形态，但应确保“标准、主数据、元数据、质量、资产、共享、安全”均有明确责任主体。

示例1：数据治理职责分工与交付物清单

治理要素	牵头责任	主要参与	主要交付物
数据标准与指标口径	数据治理办公室/信息部门	临床、医技、药事、护理、运营	数据元与编码规范；指标口径手册；接口字段字典
主数据管理	信息部门/数据管理岗	各业务条线	主数据对象清单；编码映射表；变更记录与版本
元数据与血缘	数据平台团队	各系统管理员	元数据登记；血缘链路；数据集说明与影响分析
数据质量	数据治理办公室	业务部门、平台团队	质量规则库；监测报表；问题工单与整改验证
数据资产目录与服务	数据资产管理员	数据使用方、平台团队	数据资产目录；数据服务目录；服务SLA与调用监测

5.2 数据标准体系建设要求

医院应建立统一的数据标准体系，作为数据汇聚、治理加工、数据服务与模型训练的共同依据。数据标准体系至少应包含数据元标准、编码体系、数据字典、指标口径、接口字段定义与消息规范等内容，并形成可维护、可版本化的标准管理机制。

数据标准体系建设应满足以下要求：

- a) 统一语义：对同一业务概念应统一定义，避免同名异义或异名同义；
- b) 统一编码：对关键对象与关键字段应明确编码规则与取值范围，并建立跨系统映射关系；
- c) 统一口径：对核心指标应明确计算口径、数据来源、统计周期、过滤条件与版本信息，保证可复核；
- d) 版本管理：标准变更应具备评审、发布、生效、兼容与回溯机制，确保历史口径可追溯；

e) 与系统联动：数据标准应与ETL/ELT加工规则、数据服务定义、特征定义与模型输入输出字段保持一致性绑定，避免“标准在文档、实现各自为政”。

医院在制定数据标准时，应优先覆盖对AI与运营分析影响最大的领域，包括但不限于：患者主索引与就诊链路、诊断与手术操作、检验检查项目与结果表达、用药与处置、护理评估、费用结算、质控事件、资源排班等。

5.3 主数据管理要求

医院应建立主数据管理机制，保证跨系统的关键对象一致性，为数据汇聚、指标口径统一与AI训练样本一致性提供基础。主数据对象宜至少包括：患者主索引、人员与资质、机构科室、诊疗项目、药品、耗材、检验项目、检查项目、设备资产等。

主数据管理应至少满足以下要求：

- a) 统一主编码：为每类主数据对象建立唯一标识与主编码体系，明确编码规则与生命周期；
- b) 映射与对齐：对存量系统编码差异应建立映射关系，并配置持续同步与一致性校验；
- c) 权责与变更控制：明确主数据维护责任与审批流程，变更应留痕并具备回滚能力；
- d) 分发与消费机制：主数据应通过统一服务或标准分发机制提供给业务系统、数据平台与AI平台使用，避免各自维护造成漂移；
- e) 质量与一致性监测：应对主数据重复、缺失、冲突、失效等进行监测与整改闭环。

涉及患者主索引的，医院应确保同一患者在不同系统、不同院区或不同就诊渠道下可进行准确匹配与合并，并对合并、拆分、更正等高风险操作配置更严格的审批与审计策略。

5.4 元数据、数据资产目录与血缘追溯要求

医院应建设覆盖数据全链路的元数据管理与血缘追溯能力，保证数据“可发现、可理解、可追溯、可评估影响”。元数据管理对象至少应包括数据集、表、字段、指标、接口服务、标签、特征与模型输入输出字段等。

元数据与血缘管理应满足以下要求：

- a) 统一登记：数据资产上架前应完成元数据登记，至少包含定义、来源系统、更新频率、责任人、权限属性、质量状态与使用约束；
- b) 血缘贯通：应记录从源系统到数据湖、数据仓库、数据服务、特征库、模型训练集与推理输入的加工链路与规则版本；
- c) 影响分析：标准、口径、加工规则或接口变更时，应支持影响范围分析并支撑变更评审；
- d) 可审计：对对外提供数据、用于科研的数据集、用于模型训练的数据集，应能追溯来源、处理过程与授权依据；
- e) 与服务联动：数据资产目录应与数据服务目录联动，明确服务接口、SLA、调用监测与下线机制。

数据资产目录的分类宜按主题域组织，并明确共享等级与使用边界，支撑跨部门复用与审批管理。数据资产目录与服务目录应作为数据服务化与AI工程化的重要入口，其建设要求应与图1中“服务共享层”的定位相一致。

5.5 数据质量管理要求

医院应建立数据质量管理体系，形成“规则—监测—工单—整改—验证—固化”的闭环机制，保障数据用于分析决策与模型训练时的可信度与可重复性。

数据质量管理应至少包括：

- a) 质量维度与指标：定义准确性、完整性、一致性、及时性、唯一性、有效性等维度，并明确监测指标与阈值；
- b) 质量规则库：建立字段级、记录级、主题域级质量规则，包括缺失、重复、范围、逻辑一致性、编码匹配与跨表一致性校验；
- c) 质量监测与告警：对关键主题域与关键指标实施持续监测，异常应触发告警并进入整改闭环；
- d) 问题闭环与责任：质量问题应形成工单，明确责任部门、整改时限与验证方式；
- e) 质量分级与适用性声明：用于临床决策支持或高影响模型训练的数据集应满足更高质量要求，并形成数据集质量说明与适用性声明；
- f) 质量固化：对频发问题应追溯根因并在源头采集、业务流程或数据标准层面固化改进，避免长期依赖事后修补。

数据质量规则、监测指标与整改闭环应与表1能力框架中的“数据质量”子能力一致衔接，并为AI训练验证提供可量化的数据可靠性依据。

5.6 数据共享、对外提供与使用边界管理要求

医院应建立数据共享与对外提供管理机制，确保数据在共享交换、科研使用、第三方合作与区域协同等场景下满足最小必要、目的限定、授权可控与审计可追溯要求。

数据共享与对外提供应至少满足以下要求：

- a) 分级分类与共享等级：对数据集与字段设置共享等级与访问边界，明确可共享范围、共享方式与审批要求；
- b) 审批与授权：数据共享与对外提供应按流程审批，明确申请目的、使用范围、期限、接收方责任与禁止事项；
- c) 脱敏与最小化：按数据敏感程度配置脱敏策略与最小必要字段集，避免过度提供；
- d) 留痕与审计：对共享与对外提供应记录数据集版本、字段清单、脱敏方式、授权依据、提供时间、接收方与使用期限，并支持追溯；
- e) 结果回流与复核：对联合建模或外部分析结果，应明确回流方式、复核机制与使用边界，避免未经验证直接进入临床流程；

f) 违规处置：建立违规使用监测与处置机制，明确责任追究与整改要求。

共享与对外提供管理应与图1“运营安全层”的统一身份、访问控制与审计能力贯通实现，保证链路可控。

5.7 数据生命周期管理要求

医院应建立数据生命周期管理机制，覆盖数据从产生采集到加工发布、共享使用、归档保存与安全销毁的全过程管理，以支撑合规要求、成本优化与风险控制。

数据生命周期管理应至少包括：

a) 分类分层存储策略：明确原始数据、明细数据、汇总数据、标签特征与训练样本等不同层级数据的存储策略与保留周期；

b) 数据版本与可重复性：对关键数据集、训练样本集与特征定义应实施版本管理，保证分析结果与模型训练可复现；

c) 归档与保全：对需要长期保存的数据应明确归档规则与检索机制，保证可追溯；

d) 安全销毁与权限回收：对超过保留期限或授权到期的数据，应按规定销毁或不可逆处理，并回收访问权限；

e) 成本与容量治理：结合数据冷热分层、压缩与归档策略优化存储成本，并确保性能满足关键业务需求。

6 数据平台与服务共享能力建设要求

6.1 总体要求

医院数据平台建设应满足以下总体要求：

a) 平台化与服务化：数据接入、加工、治理、服务发布与共享审批应平台化实现，减少点对点脚本与人工操作依赖；

b) 分层分域：数据应按来源层、贴源层、明细层、汇总层、指标层（或标签特征层）等分层组织，主题域划分应与业务领域一致；

c) 批流一体（按需）：应同时支持批量与实时数据处理能力，实时能力的建设范围应与业务需求相匹配；

d) 可观测与可审计：应具备链路监测、数据质量监测、接口调用监测、作业运行监测与审计追溯能力；

e) 弹性与高可用：应支持容量扩展与性能弹性，关键链路应具备高可用与容灾能力；

f) 安全合规内生：统一身份认证、访问控制、分级分类、脱敏与共享审批应贯穿平台全链路。

6.2 数据接入与汇聚能力要求

6.2.1 数据源接入范围与接入清单

医院应建立数据源接入清单，至少覆盖HIS、EMR、LIS、PACS、手术麻醉、输血、护理、药事、病理、运营管理、设备与物联等系统，并明确每类数据源的接入方式、更新频率、增量规则、字段映射、质量校验与责任人。

接入清单应与第5章数据标准、主数据与元数据登记要求一致，接入前应完成数据源字段语义梳理与标准映射准备。

6.2.2 接入方式与数据传输机制

数据接入应支持以下方式并按需组合使用：

- a) 标准接口与服务调用：适用于业务系统可提供标准化接口的场景；
- b) 消息与事件驱动：适用于状态变更、实时通知与高频事件数据汇聚；
- c) 批量抽取：适用于历史数据回溯、低频更新与数据量较大的表级同步；
- d) 文件或对象存储导入：适用于影像、文档、结构化报表等离线交换数据；
- e) 数据库日志或变更捕获：适用于增量数据同步与对一致性要求较高的场景。

接入链路应配置传输加密、完整性校验、失败重试、断点续传与幂等处理机制，避免重复写入或数据丢失。对关键主题域数据，应配置接入侧数据校验规则与异常告警策略。

6.2.3 贴源层建设要求

医院宜建设贴源层（或ODS层）用于承接源系统数据，贴源层应尽可能保持源数据结构与语义，并记录来源系统、采集时间、增量标识、版本信息与关键审计字段，为后续治理加工与血缘追溯提供依据。

贴源层应至少满足：

- a) 保留原始字段与必要的业务键；
- b) 记录采集批次与采集时间戳；
- c) 支持按主数据映射进行关键对象统一标识（如患者主索引映射）；
- d) 对重大变更（字段新增、含义变化、枚举扩展）具备版本兼容与变更留痕能力。

6.3 存储与计算能力要求

6.3.1 存储体系与分层策略

医院应根据数据类型与使用场景建立分层存储体系，至少应覆盖结构化数据、半结构化数据与非结构化数据（如文档、报告、影像相关元数据等）。存储体系可采用数据湖、数据仓库或湖仓结合方式实现，但应满足以下要求：

- a) 主题域可扩展：支持按主题域扩展数据集与分区策略；
- b) 历史可追溯：支持历史快照或版本记录，满足口径回溯与模型可复现；
- c) 冷热分层：支持冷热分层与归档，平衡成本与性能；

- d) 一致性保障：关键数据应具备一致性校验机制，防止跨层数据不一致。

6.3.2 计算体系与资源治理

数据平台应支持离线计算与按需实时计算能力，并具备统一的资源调度与隔离策略。对关键计算任务应配置优先级、资源配额、超时与重试策略，避免资源争用影响关键业务链路。

计算体系应满足：

- a) 支持批处理任务的可编排与可追溯；
- b) 支持实时处理的低时延与稳定运行；
- c) 支持交互式分析与探索（用于科研分析与数据开发）；
- d) 支持作业级监控、失败告警与自动恢复策略。

6.3.3 备份与容灾

医院应对平台关键组件与关键数据集配置备份与容灾策略，至少包括：

- a) 数据备份：贴源层、核心主题域明细层、关键指标层、训练样本集等；
- b) 配置与元数据备份：数据标准、主数据映射、元数据血缘、权限策略、作业编排配置等；
- c) 容灾演练：定期演练恢复流程，验证恢复时间目标与恢复点目标满足业务要求；
- d) 备份审计：备份与恢复操作应留痕可审计。

6.4 数据加工建模与指标口径固化要求

6.4.1 数据分层加工与主题建模

医院应建立数据分层加工与主题建模机制，将贴源数据加工为可复用的主题域明细数据与汇总数据，并固化关键指标口径。加工规则应与数据标准、主数据与元数据管理一致，确保口径可复核、血缘可追溯。

主题建模应至少满足：

- a) 明确主题域范围与边界；
- b) 明确核心业务实体与关联关系（如患者—就诊—医嘱—检验检查—用药—费用等链路）；
- c) 明确时间维度与事件维度（发生时间、记录时间、入库时间等）；
- d) 明确口径版本与变更兼容策略。

6.4.2 统一指标体系与指标服务

医院应对运营分析、质控监测与AI训练常用指标建立统一口径体系，明确指标定义、计算逻辑、数据来源、过滤条件、统计周期与版本信息，并通过指标服务方式对外提供，避免不同报表、不同模型使用不同口径。

指标口径固化后，变更应纳入变更控制与影响分析流程，并通过元数据与血缘记录支撑回溯。

6.4.3 加工质量校验与异常处理

数据加工链路应配置质量校验与异常处理机制，至少包括：

- a) 输入校验：源数据缺失、主键冲突、编码不匹配等；
- b) 过程校验：行数对账、关键字段分布漂移、重复率异常等；
- c) 输出校验：指标值合理性、跨层一致性、口径一致性等；
- d) 异常处置：异常告警、工单流转、补数重跑与结果复核。

加工链路的质量校验规则应与第5章数据质量规则库一致衔接。

6.5 数据服务与共享交换能力要求

6.5.1 数据资产目录与服务目录联动

医院应将可共享、可复用的数据集与指标纳入数据资产目录，并将可调用的数据能力纳入数据服务目录。目录应体现数据集定义、口径、版本、质量状态、共享等级、责任人、服务接口与SLA等信息。

数据服务发布前应完成元数据登记、质量评估与权限策略配置，其发布流程应与第5章共享审批与对外提供管理要求一致。

6.5.2 数据服务形态与接口规范

数据服务宜至少包括：

- a) 数据查询服务：面向明细数据或主题数据的查询；
- b) 指标服务：面向统一口径指标的查询与订阅；
- c) 标签服务：面向人群标签、画像标签等的获取；
- d) 数据订阅与推送：面向事件或周期性数据分发；
- e) 数据交换服务：面向跨系统或跨机构的数据交换。

数据服务应明确接口字段定义、返回码、分页策略、限流策略、幂等策略与版本管理机制，并支持调用监测与审计追溯。

6.5.3 共享审批、授权与审计

数据服务调用应在统一身份认证与访问控制体系下进行授权管理。对敏感数据、跨部门共享、对外提供等场景，应在审批通过后开通相应权限，并记录授权依据、期限与使用边界。数据服务调用应记录调用方、调用时间、查询条件范围、返回数据规模与结果状态等审计信息，满足追溯要求。

6.6 标签体系、特征工程与特征服务要求

医院在开展AI建模与精细化运营分析时，应建立可治理的标签体系与特征服务能力，以实现可复用、可一致、可追溯的特征供给。

标签与特征体系应满足：

- a) 标签与特征定义标准化：明确定义、来源、计算逻辑、更新频率与适用范围；
- b) 标签分级与权限控制：对敏感标签（如疾病、用药、风险评分等）实施更严格的访问控制与审计；
- c) 在线离线一致性：用于训练与推理的特征计算逻辑应一致，避免训练推理偏差；
- d) 版本与血缘：特征定义与计算链路应可追溯，支持回溯与复现；
- e) 质量监测：对特征缺失率、分布漂移、异常值比例等实施监测，并形成整改闭环。

6.7 数据订阅分发与消息通知要求

医院应按需建设数据订阅分发能力，为业务系统联动、运营看板更新与AI推理触发等提供支撑。订阅分发应满足：

- a) 支持按主题、按事件、按周期订阅；
- b) 支持订阅范围控制与最小必要字段集；
- c) 支持失败重试、幂等处理与死信队列（或等效机制）；
- d) 支持订阅调用监测、延迟监测与异常告警；
- e) 支持订阅关系与权限的审计追溯。

订阅分发与消息通知链路的监测与处置应纳入运行维护闭环管理。

6.8 平台可观测性与服务等级要求

医院应建立平台可观测性体系，覆盖数据接入、加工链路、服务调用与资源运行。可观测性应至少包括：

- a) 数据接入监测：采集延迟、成功率、增量一致性、异常条数；
- b) 作业运行监测：作业成功率、耗时、资源占用、失败原因；
- c) 数据质量监测：关键规则触发情况与趋势；
- d) 服务调用监测：QPS、成功率、时延、错误码与限流触发；
- e) 审计监测：异常访问、批量导出、高频调用等风险行为。

医院应对关键数据服务设定服务等级目标（SLO）或服务等级协议（SLA）要求，并建立告警阈值与处置流程。服务等级管理应与第8章运行维护与持续改进机制相衔接。

6.9 互联互通与外部对接要求

医院如需对接区域平台、医联体平台、科研协作平台或监管报送平台，应在统一标准、共享审批、脱敏策略与审计留痕前提下开展对接，并明确对接范围、数据项清单、交换频率、数据质量要求、异常处理与责任边界。对外对接接口应纳入接口版本管理与调用监测体系，避免对外变更影响院内关键链路。

7 人工智能平台与模型全生命周期管理要求

7.1 总体要求

医院建设人工智能平台与开展模型全生命周期管理应满足以下要求：

- a) 工程化与平台化：模型开发、训练、部署与运维应平台化管理，减少个人化脚本与不可控流程，确保可重复、可交付；
- b) 数据可追溯：模型训练与推理所用数据、标签口径、特征定义与加工链路应可追溯，满足血缘回溯与合规审计；
- c) 评估可复核：模型性能评估应具备独立测试、指标口径明确、样本定义清晰、结果可复核；
- d) 风险可控：对临床高影响应用应配置更严格的上线门槛、灰度策略、回退机制与人工兜底；
- e) 运行可监测：模型上线后应持续监测性能、输入分布、漂移与异常，并形成处置闭环；
- f) 权限与审计：对数据集、标注任务、训练作业、模型发布与推理调用等关键活动应实施最小权限与审计留痕。

模型全生命周期管理的阶段划分、关键控制点与交付物应与表2相符合。

表2 模型全生命周期阶段与关键控制点及交付物

阶段	关键控制点示例	主要交付物示例
需求与场景定义	业务目标、适用范围、输入输出、风险等级、验收指标	场景需求说明；风险分级；验收指标定义
数据准备与标注	数据来源合法性、样本定义与口径、标签一致性、数据集版本	训练/验证/测试数据集；标签规范；数据集说明书
特征工程与特征库	特征定义标准化、在线离线一致性、特征版本与血缘	特征定义清单；特征计算逻辑；特征版本记录
训练与验证	训练配置可复现、超参记录、过拟合控制、验证策略	训练配置与日志；验证结果；模型候选版本
测试与评估	独立测试集、指标体系、分层评估、公平性与鲁棒性	测试报告；评估指标明细；适用性声明
登记与发布	模型库登记、版本管理、审批流程、回退点	模型登记信息；发布审批记录；回退方案
部署与推理	服务SLA、接口规范、灰度策略、权限与审计	推理服务接口说明；灰度方案；调用审计策略
运行监测与漂移	性能监测、漂移检测、告警阈值、处置闭环	监测看板；漂移报告；告警与工单记录
更新迭代与退役	再训练触发条件、版本切换、退役策略与归档	迭代记录；版本切换验证；退役与归档记录

7.2 数据准备与标注管理要求

7.2.1 数据来源与合规性

用于模型训练、验证、测试与推理的数据应来源合法、目的明确、权限合规。医院应明确数据使用目的、使用范围、使用期限与接触人员范围，并通过统一审批授权机制开通访问权限。对涉及个人信息与健康信息的数据处理活动，应落实最小必要、访问控制、脱敏与留痕要求。

7.2.2 样本定义、标签口径与数据集版本

医院应明确样本定义与纳入排除标准，形成可复核的数据集构建规则。标签定义应口径一致、可追溯，标注规范应包括标签含义、边界规则、冲突处理、质检方法与一致性评价方式。训练集、验证集与测试集应进行版本化管理，并记录来源、抽样方式、时间范围、关键分层变量分布与质量说明。

7.2.3 标注流程、质检与一致性控制

人工标注或半自动标注应建立流程化管理，至少包括任务分配、标注执行、质检抽查、争议仲裁与版本发布。对多标注者场景，应配置一致性评价指标与阈值，低一致性标签应触发复核与规则修订。标注数据的变更、回滚与历史版本应可追溯。

7.2.4 数据脱敏与隔离

对用于科研与模型训练的敏感数据应按策略脱敏或去标识化，并在数据平台与AI平台层面实施环境隔离与访问隔离。训练环境与生产环境之间的数据流转应具备审批、记录与审计机制，防止未经授权的数据复制与外泄。

7.3 特征工程、特征库与一致性要求

医院应建立特征工程规范与特征库，实现特征定义、计算、复用与治理。特征管理应满足以下要求

- a) 特征定义标准化：特征应明确名称、定义、来源、计算逻辑、统计窗口、更新频率、适用范围与权限属性；
- b) 在线离线一致性：用于训练与推理的特征应采用一致的计算逻辑与口径，避免训练推理偏差；
- c) 版本与血缘：特征应版本化管理，记录依赖数据集、加工链路与变更历史，支持回溯与复现；
- d) 质量监测：对特征缺失率、分布漂移、异常值比例等进行监测，异常应进入整改闭环；
- e) 权限与审计：对敏感特征与高风险标签应实施更严格的访问控制与审计。

7.4 训练、验证与测试评估要求

7.4.1 训练可复现与过程留痕

模型训练应记录训练数据集版本、特征版本、代码版本、超参数配置、随机种子（如适用）、训练环境与算力资源信息，确保训练过程可复现。训练日志、关键中间产物与模型权重文件应可追溯并纳入版本管理。

7.4.2 验证策略与过拟合控制

医院应根据模型类型与样本特征选择合适的验证策略，并配置过拟合控制策略。验证结果应形成记录，包含指标、分层表现与误差分析。

7.4.3 独立测试与分层评估

模型上线前应在独立测试集上完成评估。评估应不仅包含总体指标，还应按关键分层维度（如年龄段、科室、病种、设备类型、采集方式、时间段等）进行分层评估，识别性能差异与潜在风险。对临床高影响场景，宜增加回顾性验证与前瞻性试运行验证，并形成适用性声明与限制条件说明。

7.4.4 指标体系与阈值门槛

模型评估指标应与场景目标一致，并明确指标口径与计算方法。医院应为不同风险等级场景设置上线门槛与否决项，并在发布审批中执行。

7.5 模型登记、版本管理与发布审批要求

医院应建设模型库（模型登记与管理能力），对模型元信息进行统一管理。模型登记信息宜至少包括：模型名称、用途与场景、输入输出字段定义、训练数据集与特征版本、训练配置摘要、评估报告、适用范围与限制、风险等级、上线记录、责任人、版本号与回退点。

模型发布应执行审批流程。对临床高影响模型，审批宜包含临床业务负责人、质量管理与安全合规负责人参与，并明确上线范围、灰度策略、回退条件与人工兜底方案。模型版本切换应记录切换时间、影响范围、验证结论与复核结果。

7.6 部署推理、服务治理与系统集成要求

7.6.1 部署形态与接口规范

模型部署可采用服务化推理、嵌入式部署或混合方式实现。无论采用何种部署形态，均应明确接口规范、输入输出字段口径、错误码、超时策略与重试策略，并将推理服务纳入统一身份认证、访问控制与审计体系。

7.6.2 灰度发布与回退机制

模型上线宜采用灰度发布策略，逐步扩大范围并持续验证效果。应明确回退触发条件（如性能下降、漂移严重、误报漏报显著、临床反馈不可接受等）与回退流程，确保在风险出现时能够快速恢复至上一稳定版本或人工流程。

7.6.3 人工确认与临床闭环

对临床决策支持类输出，医院应明确模型输出在业务流程中的定位（提示、建议或强提醒等），并配置必要的人工确认、解释呈现与处置闭环机制，避免模型输出被误认为最终诊断或自动决策。模型输出、人工采纳与否、处置结果应留痕可追溯，为持续评估与再训练提供依据。

7.7 运行监测、漂移管理与再训练要求

模型上线后医院应建立持续监测机制，至少包括：

- a) 性能监测：对关键指标进行持续评估，必要时采用抽样复核或金标准复查；

- b) 输入分布与漂移检测：监测关键特征分布变化、缺失率变化与采集方式变化；
- c) 运行稳定性监测：监测推理时延、错误率、资源占用与调用峰值；
- d) 业务效果监测：监测模型对闭环完成率、处置时效、质量安全指标的影响；
- e) 告警与处置闭环：设置告警阈值，触发后进入工单处置与复盘改进。

再训练或模型更新应设置触发条件，包括但不限于：模型性能持续下降、漂移达到阈值、业务规则或诊疗路径发生重大变化、数据标准或口径变更、设备或采集流程变化、出现重大安全风险或投诉事件等。再训练过程应遵循表2对应阶段的控制点要求，并保持版本可追溯与可回退。

7.8 可解释性、可靠性与安全控制要求

对临床与管理高影响模型，医院应配置可解释性与可靠性控制措施，至少包括：

- a) 解释呈现：输出应提供必要的解释信息或证据要点（如关键影响因素、阈值触发原因、相关历史信息摘要等）；
- b) 适用性边界：明确模型适用人群、适用科室、适用设备、适用时间范围与限制条件；
- c) 异常输入处理：对缺失、异常或超范围输入应采取保护性策略（拒绝输出、提示复核或降级到规则/人工流程）；
- d) 误用防控：界面与流程应避免将模型输出呈现为最终诊断结论，必要时增加提示与确认；
- e) 安全审计：模型输出与采纳路径应可审计，支持事后追溯与责任界定。

7.9 MLOps 流水线与审计合规要求

医院应建立MLOps工程化能力，实现数据—特征—训练—评估—发布—部署—监测的流水线管理，并与数据治理与安全合规体系贯通。MLOps能力宜包括：

- a) 代码、数据、特征与模型版本统一管理；
- b) 自动化训练与评估流水线（按需）；
- c) 模型发布审批与自动化部署（含灰度与回退）；
- d) 运行监测、漂移检测与告警；
- e) 全链路审计留痕与报告生成；
- f) 权限分离与最小权限控制（开发、审核、发布、运维职责分离）。

对涉及对外合作、联合建模或第三方算法引入的，应明确责任边界、数据处理边界、模型知识产权与合规要求，并将第三方接入与模型调用纳入统一审计与风险管理。

8 场景应用建设与评估验收要求

8.1 总体要求

医院开展大数据与人工智能场景应用建设应满足以下要求：

a) 场景闭环与责任明确：场景应定义触发条件、责任角色、处置路径与闭环判定条件，确保“输出可用、流程可承接、责任可追溯”；

b) 数据与口径可复核：场景所用数据集、指标口径、标签特征与模型输入输出字段应可追溯并保持一致；

c) 风险分级与管控：应对场景风险等级进行分级管理，高影响场景应配置更严格的上线门槛、验证方式与回退策略；

d) 人机协同与可解释：对临床决策支持类场景，应提供必要解释信息与证据要点，明确模型输出的业务定位，避免替代临床判断；

e) 持续监测与迭代：上线后应持续监测效果与风险，触发后进入优化闭环，必要时按表2要求实施模型再训练或策略调整；

f) 合规先行：涉及个人信息与健康信息处理的场景，应落实最小必要、目的限定、授权控制、脱敏与审计留痕。

8.2 场景遴选、分级与建设清单

8.2.1 场景遴选原则

医院应基于业务价值与可落地性遴选场景，优先选择数据基础较好、闭环可定义、效果可度量、风险可控的场景，避免仅以“技术可做”作为立项依据。场景遴选宜综合考虑以下因素：

- a) 医疗质量安全改善潜力；
- b) 业务高频程度与效率提升空间；
- c) 数据可得性与质量可控性；
- d) 业务流程承接能力与组织协同成熟度；
- e) 合规可行性与隐私风险；
- f) 实施复杂度与依赖条件（接口、终端、人员培训等）。

8.2.2 场景分级管理

医院应对场景进行风险分级管理。分级依据宜包括：对患者安全影响程度、对诊疗决策影响程度、误报漏报后果、可解释性要求、替代人工程度、合规敏感性等。

对风险较高的临床场景（如辅助诊断、危重风险预警、用药禁忌强提醒、影像检出与分级等），应采用更严格的验证策略、灰度发布与回退机制，并设置人工确认与兜底流程。

对运营管理与资源调度类场景，可在确保数据口径一致与权限合规前提下加快迭代，但仍应保证指标可复核与结论可追溯。

8.2.3 场景建设清单与路线图

医院应形成场景建设清单与分期路线图，至少包括：场景名称、业务目标、适用范围、数据依赖、模型或规则类型、集成系统、风险等级、验收指标、上线策略与责任人。清单应纳入统一治理与变更控制，并与图1所示总体架构中的“应用层—AI平台层—服务共享层”贯通一致。

8.3 场景应用建设流程要求

8.3.1 需求定义与流程嵌入

场景需求定义应明确业务流程嵌入点与处置闭环，至少包括 触发条件、输入数据、输出内容、呈现方式、责任角色、处置动作、时限与升级路径、闭环判定条件、审计留痕点、异常分支与降级策略。

场景需求定义完成后，应进行流程一致性校验，确保院内制度、岗位职责与系统流程可承接，避免出现“模型输出无法被执行”“结果无处落地”的情况。

8.3.2 数据准备与可用性评估

场景建设前应对数据可用性进行评估，至少包括 数据来源清单、字段口径与标准映射、关键对象主数据一致性、缺失与异常情况、历史覆盖范围、标注可行性（如适用）、合规审批与脱敏策略。

数据准备、标注与数据集版本管理应符合第7章要求，模型全生命周期阶段控制点与交付物应与表2相符合。

8.3.3 方案设计与实现

场景实现可采用规则、统计分析、机器学习或深度学习等方式，医院应根据风险等级与可解释性要求选择合适技术路线。方案设计应明确：

- a) 采用规则或模型的边界与分工；
- b) 特征来源与计算逻辑；
- c) 输出阈值、分级策略与触发频率控制；
- d) 与业务系统的集成方式与接口约束；
- e) 运行监测指标与告警策略；
- f) 回退方案与人工兜底流程。

对需要跨系统协同的场景，应优先通过数据服务、指标服务与模型服务实现复用，避免重复实现与口径漂移。

8.3.4 联调测试与试运行

场景上线前应进行联调测试与试运行验证。测试应覆盖典型流程与异常分支，包括接口失败、数据缺失、权限不足、超时、重复触发、撤销与回退等情况。

对高风险临床场景，宜在限定范围内开展试运行，形成试运行报告，内容至少包括 模型性能、误报漏报分析、临床采纳情况、处置闭环效果、用户反馈与改进建议。

8.4 系统集成、人机协同与结果呈现要求

8.4.1 集成方式与调用治理

场景应用与业务系统集成应明确调用方式（同步/异步）、接口字段口径、错误码与超时策略、重试与幂等机制，并纳入统一身份认证、访问控制与审计体系。

模型推理服务、指标服务与数据服务的调用应纳入服务治理，具备限流、熔断、降级与告警能力，保证关键业务链路稳定。

8.4.2 结果呈现与可解释

临床决策支持类输出应以“提示/建议/预警”等形式呈现，并提供必要解释信息与证据要点，明确适用范围与限制条件。

场景输出不应以界面文案、流程设置或默认选项误导使用者将模型输出视为最终诊断结论。对强提醒或拦截类提示，应配置确认机制与理由记录，并纳入审计留痕。

8.4.3 处置闭环与反馈采集

场景应配置处置闭环机制，至少记录 触发时间、触发依据摘要、接收对象、处置动作、处置时间、关闭原因、最终结果与复核信息。

场景应采集人工反馈（采纳/未采纳、原因、结果）以支撑效果评估与持续迭代。反馈数据应纳入数据治理与质量管理体系，保证可用于再训练或规则优化。

8.5 效果评估指标与验收要求

8.5.1 评估原则

场景效果评估应同时关注“技术性能、业务效果、风险与合规”三个维度，确保不仅模型指标达标，而且能够产生可度量的业务收益并维持风险可控。评估指标应口径明确、数据来源可追溯、计算方法可复核。

8.5.2 技术性能评估

技术性能评估应根据场景类型选择合适指标，例如分类、回归、检出、分割、排序推荐或异常检测等任务指标。评估应包含总体指标与关键分层维度表现，并记录评估数据集版本、样本定义与阈值选择依据。

对临床高风险场景，除离线评估外宜开展试运行阶段的在线效果评估，结合人工复核形成更贴近真实环境的性能判断。

8.5.3 业务效果评估

业务效果评估应围绕闭环效率、质量安全与资源利用等方面设置指标，例如：处置时效提升、闭环完成率提升、漏报相关不良事件减少、重复检查减少、床位周转改善、平均住院日变化、成本控制效果等。

业务效果指标应与管理目标一致，并在上线前定义基线期与对比期，避免仅凭主观感受判断效果。

8.5.4 风险与合规评估

医院应对场景进行风险与合规评估，至少包括：误报漏报风险、误用风险、数据安全与隐私风险、权限边界与审计完整性、对外合作边界等。对高风险场景，应明确否决项与回退触发条件。

8.5.5 验收要求与资料归档

场景验收应以“可用、可控、可追溯、可回退”为基本要求。验收应至少包含：

- a) 场景需求与流程说明、适用范围与限制条件；
- b) 数据集与特征版本说明、口径与血缘追溯信息；
- c) 模型评估报告与试运行报告；
- d) 集成接口说明、权限策略与审计策略；
- e) 上线与灰度方案、回退方案与兜底流程；
- f) 运行监测指标与告警策略；
- g) 用户培训材料与操作指引。

上述资料应纳入档案化管理，并与表2所列模型生命周期交付物形成映射关系，保证后续复核与再训练可追溯。

8.6 风险控制、灰度发布与回退机制

医院应针对不同风险等级场景配置差异化的上线与控制策略。高风险场景应至少具备：

- a) 灰度发布：限定科室、限定人群、限定时段或并行运行方式逐步扩大；
- b) 人工确认与兜底：明确人工复核路径与责任人；
- c) 回退机制：设置性能下降、漂移超阈值、临床反馈不可接受、重大投诉或不良事件等回退触发条件，并明确回退流程；
- d) 事件复盘：对重大异常应复盘并形成整改措施，必要时触发模型再训练或规则重构，过程应符合表2要求

8.7 上线推广与运营管理

场景上线后医院应建立运营管理机制，至少包括：

- a) 培训与宣贯：按角色开展培训，强调使用边界、处置流程与合规要求；
- b) 运行监测：监测调用量、响应时延、错误率、触发量、处置闭环率、采纳率等指标；
- c) 反馈与迭代：建立用户反馈渠道与迭代节奏，形成“反馈—分析—优化—验证—固化”的闭环；

- d) 价值复盘：定期复盘业务效果与投入产出，决定扩面、优化或退役；
- e) 变更控制：上线后的阈值调整、规则更新、模型版本切换应纳入变更控制与审计留痕。

9 安全与合规保障要求

9.1 总体要求

医院应将网络安全、数据安全与个人信息保护要求嵌入平台架构与业务流程，落实统一身份认证、访问控制、审计追溯、数据分级分类、脱敏与共享审批等机制，确保数据与模型“可授权、可控制、可审计、可追责”。

9.2 数据分级分类与最小必要

医院应对数据集、字段、标签与特征实施分级分类管理，明确共享等级与使用边界。数据处理与共享应遵循最小必要原则，按目的限定配置可访问范围、字段范围与使用期限，并保留授权依据与审计记录。

9.3 身份认证、权限控制与职责分离

医院应建立统一身份认证与授权体系，按“人员—角色—职责—权限—数据范围”实施访问控制。对高风险操作（批量导出、跨域访问、训练数据集复制、模型发布与回退、规则阈值调整等）应实施更严格认证策略、职责分离与强化审计。

9.4 审计留痕与安全监测

医院应对数据访问、数据服务调用、共享审批、模型训练与发布、推理调用、配置变更等关键活动实施全链路审计。应建立异常访问与高风险行为监测机制，对异常告警形成处置闭环并定期复盘。

9.5 脱敏、去标识化与对外提供控制

医院在科研使用、联合建模、对外提供或第三方合作场景下，应按分级分类要求配置脱敏或去标识化策略，明确接收方责任、数据使用范围与期限，并对数据交付、结果回流与销毁回收进行登记与审计。

9.6 第三方接入与算法引入管理

医院引入第三方平台、算法或开展联合建模时，应明确数据处理边界、接口调用边界与责任划分，并对第三方接入实施网络边界控制、认证授权、审计留痕与应急处置机制。第三方算法上线前应完成安全评估与风险评估，必要时设置灰度与回退。

9.7 安全事件响应与合规审计

医院应建立安全事件分级响应机制，覆盖发现、研判、处置、恢复、复盘与改进。应定期开展合规审计与专项检查，对发现的问题形成整改闭环并验证有效性。

10 运行维护与持续改进

10.1 运维组织与职责

医院应建立覆盖数据平台、AI平台、应用系统与安全合规的运维组织，明确应用运维、数据运维、接口运维、模型运维与安全运维职责边界，并建立值守、响应与升级机制。对关键业务链路和高影响场景，应配置联合保障与快速响应通道。

10.2 运行监测与可观测性

医院应建立端到端可观测性体系，至少覆盖数据接入延迟与成功率、作业运行状态、数据质量指标、数据服务调用量与时延、推理服务稳定性与错误率、模型性能与漂移、审计与异常访问等。关键阈值与告警策略应形成处置闭环，并定期复盘优化。

10.3 故障、变更与配置管理

医院应建立故障分级响应与根因分析机制，形成“发现—处置—恢复—验证—复盘”闭环。数据标准、主数据映射、加工规则、接口字段、特征定义、模型版本与阈值策略等关键配置变更应纳入变更控制，实施影响评估、测试验证、灰度发布与回退策略，并保留审计留痕。

10.4 模型运行治理与迭代

模型上线后应持续监测性能与漂移，设置再训练或更新触发条件。模型版本切换应执行审批与验证，必要时保留并行运行与人工兜底机制。对低采纳率、高误报或高投诉的模型/规则，应开展专项评估并优化或退役。

10.5 数据质量与资产运营

医院应开展数据质量常态化治理，建立质量规则、监测与工单整改闭环。数据资产目录与数据服务目录应持续维护，明确责任人、共享等级、SLA与下线机制，确保数据供给可控、可复用、可追溯。

10.6 安全运营与合规复核

医院应开展账号权限复核、异常访问监测、第三方接入复核与安全漏洞管理。应定期开展合规审计与专项检查，对发现问题形成整改闭环并验证有效性。

10.7 定期评估与持续改进

医院应定期开展平台能力与应用效果评估，综合技术指标、业务指标与风险合规指标形成评估报告，输出改进清单、责任分工与验证方式。有效改进措施应固化为标准、规则与配置基线，并推广复制。