ICS 点击此处添加 ICS 号
CCS 点击此处添加 CCS 号

T/ISC

团 体 标 准

T/ XXXX—XXXX

互联网的关键领域 AI 训练数据可信质量要求

Trustworthy quality requirements for AI training data in key fields of the internet

（征求意见稿）

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX－XX－XX 发布 XXXX－XX－XX 实施

发 布

目 次

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。

本文件由中国互联网协会提出并归口。

本文件起草单位：珠海市人民医院、中国联通国际有限公司、南京南瑞信息通信科技有限公司、中国联通（香港）创新研究院、南京航空航天大学、中国信息通信研究院、武汉卓讯互动信息科技有限公司、北京邮电大学、北京航空航天大学、深圳鹿野人工智能有限公司、中科数测科技有限公司、华兴中科标准技术（北京）有限公司。

本文件主要起草人：王涵、肖雨果、于向荣、刘书博、朱世顺、魏兴慎、曹永健、张恺、张吉、陈文弢、静静、马若龙、邵彦华、刘亚卓、宋宗维、刘欣然、周鸣一、黎立、孙海波、董坤、董婧一、成瑾、李华、任国静、丁月。

互联网的关键领域AI训练数据可信质量要求

1　范围

本文件规定了互联网的关键领域AI训练数据可信质量的基本要求，数据采集、预处理、标注、合成数据、存储与传输、使用与处置全生命周期各环节的可信质量要求，以及数据可信质量评估方法和管理保障要求。

本文件适用于互联网关键领域AI系统的训练数据提供者、处理者及使用者开展数据可信质量管理相关活动，也可用于第三方评估机构进行可信数据质量评估。

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 35273—2020　信息安全技术　个人信息安全规范

GB/T 40685　信息技术服务　数据资产　管理要求

GB/T 45652　网络安全技术　生成式人工智能预训练和优化训练数据安全规范

GB/T 45674—2025　网络安全技术　生成式人工智能数据标注安全规范

3　术语和定义

下列术语和定义适用于本文件。

3.1

关键领域　critical domain

涉及国家安全、国民经济命脉、重要民生、重大公共利益等领域，如金融、能源、医疗、交通等。

3.2

AI训练数据　AI training data

用于训练和优化人工智能模型的数据集，包括文本、图像、音频、视频等模态数据。

注：参考GB/T 45652中预训练数据和优化训练数据的定义。

3.3

数据可信质量　data trustworthiness quality

数据在真实性、准确性、完整性、一致性和安全性等方面的综合属性，满足特定人工智能应用场景的功能性、安全性及合规性要求。

3.4

数据标注　data annotation

通过人工操作或使用自动化技术机制，将特定信息（如标签、类别或属性）添加到数据样本的过程。

[来源：GB/T 45674—2025,3.3，有修改]

3.5

个人信息　personal information

以电子或者其他方式记录的能够单独或者与其他信息结合识别特定自然人身份或者反映特定自然人活动情况的各种信息。

[来源：GB/T 35273—2020,3.1]

4　概述

互联网的关键领域AI训练数据可信质量管理覆盖数据采集、预处理、标注、存储、传输、使用和处置全生命周期各环节，以组织战略为指引，以业务为主线、以价值为导向，通过管理对象、管理过程和管理保障的协同联动，实现数据资产保值增值。管理框架如图1所示。

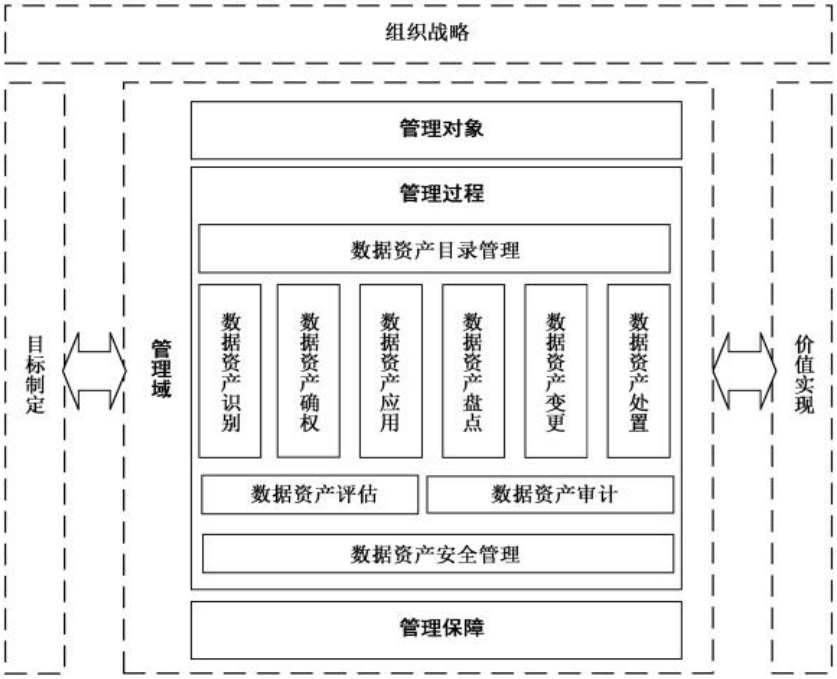

图 1　数据质量管理框架

注：图1参考GB/T 40685中数据资产管理框架，体现了组织战略、目标制定、管理域和价值实现的逻辑关系。

5　数据可信质量基本要求

5.1　基本原则

开展AI训练数据可信质量管理活动，应遵循以下原则：

a) 合法合规性：数据处理活动应符合国家法律法规和行业监管要求，并符合 GB/T 35273—2020 中关于个人信息保护的规定；
b) 目的明确：数据使用目的应明确、清晰、具体，不得超范围使用，参考 GB/T 35273—2020 中目的明确原则；
c) 质量可控：应建立数据质量评价机制，明确质量指标和阈值，参考 GB/T 40685 中数据质量评估要求；
d) 安全防护：应采取技术和管理双重措施，保障数据的保密性、完整性和可用性，符合 GB/T 45652 中的通用安全要求。

5.2　数据分类分级

应按照GB/T 40685中数据资产信息要素的要求，对AI训练数据进行分类分级，并采取差异化保护措施。涉及个人敏感信息的，应符合GB/T 35273—2020中个人敏感信息的处理要求。

6　数据采集阶段可信质量要求

6.1　采集任务规划

6.1.1　数据采集任务规划应作为数据采集工作的前置环节，由算法需求方、业务专家及数据管理人员共同参与，将模型训练需求转化为明确的数据采集任务，并形成正式的数据采集方案。

6.1.2　采集任务规划阶段应明确数据的类型、规模及分布要求，制定数据完整性、清晰度、准确性等质量标准。采集方案应对数据来源、采集方式、采集流程及责任分工进行说明，同时明确采集环境和设

备条件。
6.1.3　数据采集方案应在实施前完成评审。评审人员应包含业务专家、数据管理人员及安全管理人员。

6.2　数据来源合规性

6.2.1　数据采集活动应确保数据来源合法，并符合国家有关网络安全、数据安全及个人信息保护相关法律法规要求。对于来自第三方机构或平台的数据，应对数据提供方资质进行审核，并确认数据获取方式合法合规。
6.2.2　涉及个人信息的数据采集，应明确数据采集目的，并依法取得数据主体授权或满足法律规定的数据处理条件。数据采集应遵循数据最小化原则，仅采集实现既定业务目标所必需的数据。
6.2.3　应建立数据来源登记机制，对数据来源机构、数据采集时间、采集方式及授权情况进行记录，以确保数据来源可追溯。

6.3　数据真实性验证

6.3.1　应建立数据真实性验证机制，确保采集数据真实反映客观对象或业务场景。
6.3.2　对于人工采集的数据，应通过抽样审核或复核机制对采集结果进行检查；对于设备采集或自动采集的数据，应确保采集设备处于正常校准状态，并对采集数据进行格式校验、异常检测及重复检测。
6.3.3　对于发现存在明显异常或疑似伪造的数据，应及时进行标记、核查或剔除。必要时，可通过多源数据比对或交叉验证等方式提高数据真实性。

6.4　数据采集过程控制

6.4.1　数据采集活动应按照既定采集规范开展，并建立过程监控机制。
6.4.2　采集活动应记录采集时间、采集地点、采集设备及采集人员信息。对于自动化采集系统，应建立运行监控机制，对系统运行状态和数据采集情况进行监测。
6.4.3　在大规模数据采集任务开始前，宜开展小规模试采集，以验证采集流程和采集设备是否满足质量要求。试采集结果应进行评估，并根据评估结果对采集方案进行必要调整。

6.5　跨境数据采集合规性

6.5.1　涉及跨境数据采集、传输或使用的，应符合国家有关数据出境安全管理的法律法规及监管要求。
6.5.2　在开展跨境数据采集或使用前，应对数据出境的必要性、数据类型及潜在风险进行评估，并依法履行相应的安全评估、备案或其他合规程序。
6.5.3　对涉及个人信息或重要数据的跨境数据处理活动，应采取必要的安全保护措施，包括数据脱敏、访问控制及加密处理等，以降低数据泄露风险。
6.5.4　应对跨境数据流动进行记录，记录内容宜包括数据来源、出境方式、接收方信息及数据使用范围等，以支持审计与监管要求。
6.5.5　对跨境数据的使用应限定在合法授权范围内，不得超出原定用途进行扩展使用。

6.6　数据安全与隐私保护

6.6.1　数据采集过程中应采取必要的安全措施，防止数据泄露或非法使用。采集系统应具备访问控制和身份认证机制，并采用安全通信协议进行数据传输。
6.6.2　涉及敏感数据的采集活动，应采取加密存储、数据脱敏或匿名化处理等技术措施，以降低数据安全风险。
6.6.3　涉及个人信息的数据采集，应严格遵守国家个人信息保护相关法律法规，并建立相应的数据保护和管理机制。

6.7　数据采集可追溯性

6.7.1　应建立完整的记录机制，对数据来源、采集时间、采集方式及采集人员等信息进行记录。
6.7.2　数据采集系统应保存数据采集日志，以支持数据审计与问题追溯。
6.7.3　宜建立数据血缘（Data Lineage）管理机制，使训练数据能够追溯至其原始采集来源和采集过程。

6.8　数据采集质量记录

6.8.1　数据采集活动应建立数据采集质量记录机制，对数据采集过程中的关键质量信息进行记录，以支持数据质量评估、审计和问题追溯。
6.8.2　数据采集质量记录宜包括数据来源信息、采集时间、采集地点、采集设备或系统信息、采集人员或采集系统标识，以及采集过程中的质量检查结果等内容。数据采集质量记录宜以日志或元数据形式保存，并与训练数据建立关联关系。
6.8.3　数据采集质量记录应按照相关数据管理和网络安全要求进行安全存储和管理，并在数据生命周期内保持可查询和可追溯。

6.9　数据版权与知识产权合规性

6.9.1　数据采集和使用活动应符合国家有关著作权、知识产权及数据合规管理相关法律法规要求，不得侵犯他人合法权益。
6.9.2　对来源于互联网公开数据的训练数据，应评估其使用是否符合合理使用原则或相关法律规定，并避免将未经授权的数据直接用于商业用途。
6.9.3　对包含代码的数据集，应遵循相关开源许可证要求，明确数据的使用范围、修改方式及再分发条件，不得违反开源协议约定。
6.9.4　对图像、音频、视频等多媒体数据，应确认其版权归属及授权情况。对于未明确授权的数据，不应直接用于模型训练或数据集构建。
6.9.5　在使用第三方数据或公开数据构建训练数据集时，应保留数据来源信息及授权依据，以支持后续审计与合规验证。
6.9.6　对基于已有数据生成的合成数据，应评估其是否存在对原始数据的直接复现或版权侵权风险，并采取必要措施降低相关风险。

7　数据预处理阶段可信质量要求

7.1　数据清洗与过滤

7.1.1　应对原始数据进行清洗，以识别和处理缺失数据、重复数据及明显错误数据。
7.1.2　数据清洗过程中应根据业务需求和模型训练需求，对无效数据、异常数据或与任务目标无关的数据进行过滤或标记。
7.1.3　数据清洗规则应形成规范文档，并在数据处理过程中保持一致执行。

7.2　数据标准化处理

7.2.1　应对不同来源的数据进行标准化处理，确保数据格式、编码方式及单位表示的一致性。
7.2.2　数据标准化处理应统一数据结构和字段定义，并建立数据字典或元数据说明，明确数据字段含义及取值范围。
7.2.3　在多源数据融合场景下，应对不同数据源之间的数据进行规范化转换。

7.3　安全预处理措施

7.3.1　应根据数据敏感程度采取必要的安全处理措施，以防止数据泄露或非法使用。
7.3.2　对涉及个人信息或敏感数据的数据集，应进行脱敏处理，包括匿名化、去标识化或数据扰动等方式。
7.3.3　不应生成可重新识别个人身份的信息组合，并应对处理后的数据进行安全风险评估。

7.4　数据质量校验

7.4.1　数据预处理完成后，应对数据质量进行校验，以确保数据满足模型训练和业务应用需求。
7.4.2　数据质量校验宜包括数据完整性检查、数据一致性检查及数据格式合规性检查。
7.4.3　对发现存在明显质量问题的数据，应进行修正、重新处理或剔除。

7.5　数据处理可追溯性

7.5.1 数据预处理活动应建立数据处理记录机制，对数据处理过程中的关键操作进行记录。
7.5.2 数据处理记录宜包括数据来源、处理时间、处理规则及处理人员或处理系统信息。
7.5.3 数据管理系统宜建立数据血缘（Data Lineage）管理机制，使数据能够追溯到原始数据及其处理过程。

7.6 数据偏差控制

7.6.1 在数据预处理过程中，应识别可能存在的数据偏差，以降低模型训练中的系统性偏差风险。
7.6.2 应分析数据在不同类别、地域、人群或时间维度上的分布情况，避免样本分布严重失衡。
7.6.3 对存在明显偏差的数据集，可通过样本补充、数据重采样或加权处理等方式进行调整。

7.7 数据预处理质量记录

7.7.1 数据预处理活动应建立质量记录机制，对数据处理过程中的关键质量信息进行记录。
7.7.2 数据预处理质量记录宜包括数据处理时间、处理规则、处理结果及质量检查情况等内容。
7.7.3 数据预处理质量记录应按照相关数据管理和网络安全要求进行安全存储，并在数据生命周期内保持可查询和可追溯。

8 数据标注阶段可信质量要求

8.1 任务规划

8.1.1 标注需求分析与规范制定

8.1.1.1 任务定义

应与算法需求方、业务专家协作，将模型训练需求转化为明确、无歧义的数据标注任务。任务定义应至少包括：

a) 标注对象：明确数据的具体形态（如图像、文本、音频、视频等）；
b) 标注类型：明确任务类型，包括但不限于分类、框选、分割、关键点、文本生成、对话排序、内容安全性评估等；
c) 标签体系：定义完整的标签集合及层级关系；
d) 标签定义：对每个标签进行精确定义，确保语义唯一。

8.1.1.2 标注规范撰写

应制定详尽的《数据标注规范》文档，作为标注活动的执行基准。该文档应包含以下内容：

a) 正例与反例：为每个标签提供清晰、多样的图文或文本示例；
b) 边界与极端情况：明确定义模糊、疑难、遮挡、截断、小目标、光照变化、歧义语句等特殊情况的处理规则；
c) 格式与精度要求：对标注结果的格式、坐标精度、命名规则、编码格式等作出具体规定；
d) 版本记录：包含规范的版本号、修订日期、修订内容及审批人。

8.1.1.3 偏见与公平性评估

在任务规划阶段，应分析数据分布是否存在地域、性别、种族、年龄等维度的偏差。对于存在显著偏差的数据集，应制定相应的采样加权策略或补充采集计划，以确保数据的公平性与代表性。

8.1.1.4 规范评审与更新

《数据标注规范》应由算法、业务及数据质量管理等多方人员共同评审确认。在标注过程中，应建立规范的动态更新机制，所有对规则的修订都应有版本记录，并及时同步给全体相关人员。

8.1.2 人员与工具选择

8.1.2.1 人员能力要求与培训

应符合下列规定：

a) 人员筛选：应根据任务复杂度和专业性，筛选具备相应背景知识和技能的标注人员。对于关键领域（如医疗影像、金融票据、法律文本），宜优先选择具备行业知识的专业人士。
b) 岗前培训：应对所有标注人员进行岗前培训，内容包括但不限于标注规范、工具使用、安全保密协议；
c) 考核上岗：培训后应进行考核，合格后方可上岗，并保存培训与考核记录。

8.1.2.2 平台与工具要求

应符合下列要求：
a) 功能要求：应选择功能完善、稳定可靠的数据标注平台或工具。平台应具备用户权限管理、任务分配、进度跟踪、质量检查及数据版本控制等功能；
b) 安全要求：平台应符合国家网络安全与数据安全要求，特别是涉及个人信息或重要数据的处理，应优先选择支持私有化部署或境内可信云服务的平台。工具应具备操作留痕、防截屏、防复制等安全功能；
c) 效率要求：工具应支持高效、精确的标注操作，并能集成自动化质检规则（如检查标注框尺寸是否为零、标签是否在预定义集合内等）。

8.1.3 质量与安全策略

8.1.3.1 质量度量衡定义

应在任务开始前，明确定义衡量标注质量的核心指标，并设定可接受的质量阈值。核心指标应包括但不限于：
a) 准确率（Accuracy）、精确率（Precision）、召回率（Recall）；
b) 交并比（IoU，适用于检测/分割任务）；
c) 标注者间一致性（Inter-Annotator Agreement,IAA）；
d) 不良率（Defect Rate，适用于生产管理）。

8.1.3.2 质检策略设计

应设计多层次的质检流程，例如：标注员自检、交叉评审（共识机制）、质检员抽检/全检、专家终审等。

8.1.3.3 数据安全与合规

应符合下列规定：
a) 法律法规遵守：应严格遵守《中华人民共和国网络安全法》、《中华人民共和国数据安全法》及《个人信息保护法》等法律法规；
b) 数据分类分级：在规划阶段必须完成数据分类分级，并对涉及个人信息的数据进行脱敏处理（如人脸模糊化、证件号码遮蔽），除非业务场景有明确的豁免许可和授权；
c) 访问控制：应制定严格的数据访问控制策略，确保标注人员只能在授权环境下接触到完成脱敏处理的、任务所需的最少数据。严禁数据被随意下载、拷贝或外传。

8.2 任务实施

8.2.1 任务执行与过程监控

8.2.1.1 试标注与校准

在大规模标注开始前，应组织核心标注人员进行小范围的试标注。符合下列要求：
a) 试标注样本量宜不少于总任务量的 1%或不少于 50 条；
b) 试标注结果应经算法专家与质检负责人评审，标注人员间一致性（IAA）应达到预定阈值（如 0.8 以上）；
c) 试标注后应复盘并迭代《数据标注规范》，未通过校准的人员不得参与正式标注。

8.2.1.2 金标准植入

应在正式任务包中混入已知正确答案的“金标准”数据（Golden Set）。通过监测标注人员对金标准数据的标注准确率，实时监控人员状态。若连续金标准测试不合格，应暂停该人员任务并重新培训。

8.2.1.3 过程跟踪

管理人员应通过标注平台实时监控任务进度、人员效率和初步质量表现，及时发现并介入处理异常情况。

8.2.1.4 动态答疑与知识库构建

应建立高效的沟通渠道（如即时通讯群组、答疑文档），供标注人员随时提问。宜建立每日晨会或夕会机制，对所有典型问题的解答都应被记录、整理，并补充到《数据标注规范》或形成FAQ知识库，确保全体人员理解一致。

8.2.2 安全与保密执行

8.2.2.1 环境安全

标注活动应在受控的网络环境中进行。对于涉密或高度敏感的数据，应采用物理隔离或虚拟桌面（VDI）等技术方案，确保数据不离开安全域。远程标注场景下，应对终端环境进行安全检测。

8.2.2.2 行为审计

标注平台应记录所有用户的关键操作日志，包括登录、数据访问、标注、修改、导出等，日志保存期限应符合法律法规要求，以备安全审计和问题追溯。

8.2.2.3 保密协议履行

应确保所有接触数据的人员均已签署保密协议，并定期进行安全意识宣贯，强调违规操作的严重后果。

8.2.2.4 应急预案

应制定数据安全与质量事故应急预案。一旦发生数据泄露、大规模标注错误或平台故障，应立即启动应急响应，包括阻断传播、追溯源头、上报管理层及通知受影响方。

8.3 任务评审

8.3.1 评审机制

8.3.1.1 共识校验

宜采用多人独立标注同一批数据的方式，通过计算标注者间一致性，对标注结果进行一致性评估。对于不一致的标注结果，应由质检人员或领域专家进行复核或仲裁。标注一致性指标宜根据任务类型选择适当的评价方法，包括但不限于Kappa系数、一致率或其他统计指标。

标注一致性应达到预设质量阈值。对于一般任务，IAA指标宜不低于0.75；对于关键领域或高风险应用场景，IAA指标宜不低于0.80。当标注一致性低于预设阈值时，应分析原因并采取改进措施，包括优化标注规范、加强人员培训或调整标注策略。

8.3.1.2 抽样检查

对于大规模数据集，质检员应按预定比例（如5%～20%）对标注完成的数据进行随机抽样检查。宜采用动态抽样策略，若初期错误率高，应自动提高抽检比例。抽检不合格的批次应被整体驳回，要求标注员返工。

8.3.1.3 全量检查

对于准确性要求极高或样本量较小的关键任务，应进行100%的全量评审。

8.3.1.4 自动化校验

应利用脚本或平台功能，对数据进行逻辑性、格式化、一致性检查，自动发现明显错误。校验规则应包括但不限于：

a) 标注框宽高不为零；
b) 标签 ID 在预定义字典内；
c) 多边形标注点不自交；
d) 关键点位数量符合定义。

8.3.2 质量反馈与迭代

8.3.2.1 错误归因与分析

质检发现的每个错误都应被记录在案，并分析其产生原因。错误类型应标准化分类，如：理解类错误、操作类错误、规范类错误、工具类错误。

8.3.2.2 闭环反馈

应建立“质检-反馈-修改-复核”的闭环流程。质检结果应及时、具体地反馈给对应的标注员，并指导其完成修正。

8.3.2.3 持续改进

定期对质检发现的共性问题进行总结，用于优化《数据标注规范》、改进培训材料或调整标注策略。

8.4 交付验收

8.4.1 交付物清单

交付的不仅是标注结果文件，而应是一个完整的交付包，至少包括：

a) 标注数据：按照约定格式产出的最终版标注文件；
b) 《数据标注规范》：最终版本的规范文档；
c) 质量报告：包含本次任务的各项质量指标（准确率、IAA 等）的量化结果、质检过程概述、发现的主要问题及结论；
d) 数据描述文件：描述数据集的元信息，如数据来源、采集时间、数据总量、各类别样本分布统计等；
e) 合规文档：包括数据授权许可协议、知识产权声明或脱敏后的操作日志（如需）。

8.4.2 验收标准与流程

8.4.2.1 验收标准

验收方应依据任务规划阶段确定的质量阈值和交付物清单进行验收。

8.4.2.2 验收抽检

验收方有权对交付的数据集进行独立抽检，以验证质量报告的真实性和准确性。若验收抽检错误率超过约定阈值，验收方有权拒收整批数据并启动熔断机制。

8.4.2.3 交付方式

数据的传输应采用加密通道（如SFTP/HTTPS）或离线加密介质，确保传输过程的安全性。

8.4.2.4 验收确认

验收合格后，双方应签署正式的验收确认单。对于不合格项，应明确拒绝原因并启动返工或修正流程。

8.5 后期维护

8.5.1 版本控制与可追溯性

8.5.1.1 版本管理

已交付的数据集应被视为一个特定版本进行归档。任何后续的修正或增补都应作为新的版本发布，并附带详细的变更日志（Change Log）。

8.5.1.2 数据血缘与可追溯性

应建立数据血缘（Data Lineage）管理机制，确保每一个标注结果都能追溯到其原始数据、标注人员、评审人员、标注规范版本以及产生的具体时间。这一追溯链条是实现责任界定和问题排查的基础。

8.5.2 错误修正与数据迭代

8.5.2.1 错误反馈渠道

应为数据使用者（如算法工程师）建立一个便捷的渠道，用于反馈在模型训练或测试中发现的标注错误。

8.5.2.2 迭代更新流程

对使用者反馈的错误，应有专人进行核实、修正，并整合到下一版本的数据集中。此过程应遵循前述的评审和版本管理要求。

8.5.2.3 概念漂移应对

随着业务发展，现实世界中的概念可能发生变化。应定期审视存量数据的标签体系和标注规范是否依然适用，并根据需要启动对历史数据的重新标注或增量标注。

8.5.3 数据处置

8.5.3.1 存储与归档

标注完成的数据及其相关文档应按照数据安全管理规定进行加密存储和备份归档，并明确数据保留期限。

8.5.3.2 安全销毁

超出保留期限或根据法律法规、合同约定需要销毁的数据，应采用不可恢复的方式进行彻底销毁（如多次覆写、物理粉碎等）。销毁完成后，应出具由安全负责人签署的销毁报告，并保留销毁记录。

9 合成数据质量要求

9.1 基本要求

9.1.1 合成数据是指通过人工智能模型或算法生成的数据，包括但不限于生成式模型产生的文本、图像、音频和视频等数据。对于采用合成数据作为训练数据的数据集，应对其质量进行专门管理。

9.1.2 合成数据应具备与真实数据一致的语义合理性和结构一致性，不应包含明显错误、逻辑冲突或不符合业务场景的内容。

9.1.3 合成数据应经过真实性验证或质量评估，必要时应通过人工审核或与真实数据对比的方式进行校验，以确保其能够支持模型训练目标。

9.2 偏差与公平性控制

9.2.1 在生成和使用合成数据时，应识别和评估数据偏差风险，包括但不限于类别分布偏差、语义偏差及模型放大偏差等问题。

9.2.2 对发现存在明显偏差的合成数据，应通过调整生成策略、引入真实数据或重新采样等方式进行修正。

9.2.3 在关键领域应用中，应对合成数据的公平性进行评估，避免因数据偏差导致模型产生歧视性或不合理决策。

9.3 版权与合规性要求

9.3.1 合成数据的生成过程应符合相关法律法规要求，不得侵犯他人知识产权或使用未经授权的数据

进行训练生成。
9.3.2 对来源于真实数据的生成模型，应评估其输出内容是否存在对原始数据的直接复现或泄露风险。
9.3.3 在对外提供或共享合成数据时，应明确数据来源、生成方式及使用限制。

9.4 数据标识与可追溯性

9.4.1 合成数据应进行标识或标注，与真实数据进行区分。
9.4.2 数据管理系统应记录合成数据的生成模型、生成时间及生成参数等信息。
9.4.3 应建立合成数据的可追溯机制，使其能够追溯至生成模型及生成过程。

9.5 使用控制

9.5.1 在模型训练中使用合成数据时，应控制其使用比例，避免对模型训练产生不合理影响。
9.5.2 合成数据宜与真实数据结合使用，并通过实验验证其对模型性能的影响。
9.5.3 对于关键领域应用，应对使用合成数据训练的模型进行额外评估，以验证其可靠性和安全性。

10 数据存储与传输可信质量要求

10.1 数据存储安全

10.1.1 数据存储系统应具备安全防护能力，能够防止未经授权的访问、篡改或删除，并符合相关网络安全与数据安全管理要求。
10.1.2 对涉及敏感信息或关键领域数据的训练数据，应采取加密存储、访问控制和权限管理等安全措施，以降低数据泄漏风险。
10.1.3 数据存储系统宜建立备份与恢复机制，定期进行数据备份，并具备在系统故障或数据损坏情况下的数据恢复能力。

10.2 数据访问控制

10.2.1 数据存储系统应建立访问控制机制，对不同用户设置相应的访问权限，并对数据访问行为进行管理。
10.2.2 数据访问应遵循最小权限原则，仅允许授权人员访问与其工作任务相关的数据。
10.2.3 数据访问系统应记录访问日志，包括访问时间、访问人员及访问操作，以支持安全审计和问题追溯。

10.3 数据传输安全

10.3.1 数据在传输过程中应采用安全通信协议或加密技术，以防止数据在传输过程中被窃取或篡改。
10.3.2 对涉及敏感信息或关键领域数据的传输，应采取身份认证和传输完整性校验机制，以确保数据来源可信和数据内容完整。
10.3.3 数据传输过程应建立传输记录或日志，以支持传输过程的监控和追溯。

10.4 数据完整性保护

10.4.1 数据存储与传输过程中应采取措施确保数据完整性，防止数据在存储或传输过程中被篡改或损坏。
10.4.2 宜采用校验码、哈希值或数字签名等技术，对数据进行完整性验证。
10.4.3 当检测到数据完整性异常时，应及时进行核查并采取修复或重新传输等措施。

10.5 数据可追溯性

10.5.1 数据存储与传输活动应建立记录机制，对数据存储位置、数据版本及传输过程进行记录。
10.5.2 数据管理系统宜建立数据血缘（Data Lineage）管理机制，使数据能够追溯至其生成来源及处理过程。
10.5.3 在发生数据安全事件或质量问题时，应能够通过相关记录追溯数据存储和传输过程。

10.6 数据存储与传输质量记录

10.6.1 数据存储与传输活动应建立质量记录机制，对数据存储和传输过程中的关键质量信息进行记录。
10.6.2 数据质量记录宜包括存储位置、数据版本、备份情况、传输时间及传输方式等信息。
10.6.3 数据存储与传输质量记录应按照相关数据管理和网络安全要求进行安全存储，并在数据生命周期内保持可查询和可追溯。

11 数据使用与处置可信质量要求

11.1 数据使用管理

11.1.1 数据使用活动应符合数据采集和数据处理的既定目的，并遵循合法、正当和必要原则。未经授权，不得将数据用于超出原定用途的场景。
11.1.2 数据使用过程中应建立数据访问与使用管理机制，对数据使用人员、使用范围及使用行为进行管理。
11.1.3 对涉及敏感信息或关键领域数据的使用活动，应采取必要的安全控制措施，包括权限控制、数据脱敏及使用环境限制等。

11.2 数据共享与对外提供

11.2.1 数据共享或对外提供活动应符合国家有关数据安全管理规定，并经过相关管理部门或数据所有方批准。
11.2.2 对外提供数据时，应明确数据使用范围、数据保护责任及数据安全要求，并通过协议或合同形式进行约定。
11.2.3 涉及个人信息的数据共享或对外提供，应依法采取匿名化或去标识化处理措施，以防止个人信息泄露。

11.3 数据处置管理

11.3.1 当数据达到保存期限或不再满足业务需求时，应按照数据管理制度进行规范处置。
11.3.2 数据处置应采取安全措施，以防止数据在处置过程中被恢复或非法获取。
11.3.3 数据处置过程应符合相关法律法规和组织数据管理制度的要求。

11.4 数据销毁

11.4.1 对需要销毁的数据，应采用不可恢复的方式进行处理，例如数据覆盖、物理销毁或安全删除等。
11.4.2 对涉及敏感信息或重要数据的销毁活动，应在受控环境中进行，并由相关管理人员监督执行。
11.4.3 数据销毁完成后，应记录销毁时间、销毁方式及责任人员等信息。

11.5 数据使用与处置可追溯性

11.5.1 数据使用和处置活动应建立记录机制，对数据访问、使用及处置过程进行记录。
11.5.2 数据管理系统应保存相关操作日志，以支持安全审计和问题追溯。
11.5.3 在发生数据安全事件或质量问题时，应能够通过相关记录追溯数据使用和处置过程。

11.6 数据使用与处置质量记录

11.6.1 数据使用与处置活动应建立质量记录机制，对数据使用和处置过程中的关键质量信息进行记录。
11.6.2 数据质量记录宜包括数据使用人员、使用时间、使用目的及处置方式等信息。
11.6.3 数据使用与处置质量记录应按照相关数据管理和网络安全要求进行安全存储，并在数据生命周期内保持可查询和可追溯。

12 数据可信质量评估方法

12.1 评估原则

12.1.1 数据可信质量评估应遵循客观性原则，评估过程应基于真实数据和明确的评估指标，避免主观判断对评估结果产生影响。

12.1.2 数据可信质量评估应遵循可重复性原则，评估方法和评估流程应形成规范文档，以确保不同评估人员能够获得一致或相近的评估结果。
12.1.3 数据可信质量评估应遵循全面性原则，应综合考虑数据在真实性、准确性、完整性、一致性和安全性等方面的表现。

12.2 评估指标

12.2.1 对标注数据集的评估，可采用标注一致性指标，例如标注者间一致性或专家复核一致性等指标。
12.2.2 在涉及敏感数据或关键领域数据的场景中，还应评估数据安全性和合规性，以确保数据处理活动符合相关法律法规要求。

12.3 评估方法

数据可信质量评估可采用自动化评估与人工评估相结合的方法，以提高评估效率和评估准确性。
12.3.1 自动化评估可通过数据校验规则、数据统计分析或脚本检测等方式，对数据格式、数据分布及数据完整性进行检查。
12.3.2 人工评估可通过专家审核、抽样检查或交叉评审等方式，对数据内容质量和语义准确性进行评价。

12.4 评估流程

12.4.1 数据可信质量评估应建立规范化评估流程，包括评估准备、指标计算、结果分析及评估报告形成等环节。
12.4.2 在评估过程中，应根据数据规模和数据类型确定合理的抽样比例，以保证评估结果具有代表性。
12.4.3 评估完成后，应形成数据质量评估报告，对评估方法、评估指标及评估结果进行说明。

12.5 抽样方法

12.5.1 数据可信质量评估应根据数据规模、数据类型、类别分布和应用场景选择适当的抽样方法。抽样方法应有助于保证评估样本的代表性和评估结果的可信性。
12.5.2 对分布较均匀、样本结构相对单一的数据集，可采用随机抽样方法。随机抽样应确保样本被抽取的机会基本一致，避免因主观选择导致评估偏差。
12.5.3 对类别分布不均衡、来源多样或包含多个业务场景的数据集，宜采用分层抽样方法。分层抽样应根据类别、来源、时间、区域、风险等级或其他关键属性进行分层，并分别抽取样本。
12.5.4 对于少数类样本、边界样本、异常样本或高风险样本占比较低的数据集，宜采用补充抽样或重点抽样方法，以提高评估对关键风险的识别能力。
12.5.5 当训练数据集规模较大时，抽样数量应与数据规模、评估目标和风险等级相匹配。对于关键领域高风险场景，宜适当提高抽样覆盖程度。
12.5.6 抽样过程应形成记录，记录内容宜包括抽样方法、抽样依据、样本规模、样本分布及执行时间等信息。

12.6 第三方评估

12.6.1 数据可信质量评估可由数据提供方或使用方开展，也可委托第三方评估机构进行独立评估。第三方评估应保持独立性和客观性，不得存在利益冲突。
12.6.2 第三方评估机构应具备相应的技术能力和管理能力，并满足以下基本条件：应具备数据质量评估、数据安全或人工智能相关领域的专业技术能力；应具备完善的质量管理制度和内部控制机制；应具备符合国家相关规定的数据安全保障能力。
12.6.3 第三方评估机构宜通过国家或行业相关资质认证，或具备等效能力证明，包括但不限于信息安全、数据管理或软件测试等相关领域的认证资质。
12.6.4 第三方评估活动应按照统一的评估规范和流程开展，并形成完整的评估记录和评估报告。评估报告应包括评估依据、评估方法、评估过程及评估结果等内容。
12.6.5 应建立第三方评估机构的监督机制，对评估活动的规范性、独立性和结果可靠性进行监督。对于存在违规行为或评估质量不符合要求的机构，应采取暂停合作、整改或取消评估资格等措施。
12.6.6 第三方评估结果可作为数据质量认证、数据治理改进及模型训练数据选择的重要依据。

12.7 评估记录与持续改进

12.7.1 数据可信质量评估活动应建立评估记录机制，对评估过程和评估结果进行记录。
12.7.2 评估记录宜包括评估时间、评估人员、评估方法及评估结果等信息。
12.7.3 数据管理组织应根据评估结果持续改进数据管理流程，以提升数据可信质量。

12.8 多主体责任划分

12.8.1 在AI训练数据管理过程中，涉及数据提供方、数据处理方及模型训练方等多主体时，应明确各主体在数据质量管理中的职责分工。
12.8.2 数据提供方应对数据来源的合法性、数据真实性及数据完整性负责，并提供必要的数据来源说明和授权依据。
12.8.3 数据处理方应对数据预处理、标注及数据质量控制过程负责，确保数据处理过程符合相关规范要求，并对处理过程中引入的数据质量问题承担相应责任。
12.8.4 模型训练方应对训练数据的选择、使用及适用性评估负责，应对因数据使用不当或数据质量不满足要求而导致的模型性能问题承担相应责任。
12.8.5 在多主体协作场景下，各方应通过协议或合同形式明确数据质量责任、数据安全责任及风险承担机制。
12.8.6 当数据质量问题导致AI系统异常或风险事件时，应根据数据生命周期各阶段的责任划分进行追溯，并由相关责任主体承担相应责任。

13 管理保障要求

13.1 组织管理

13.1.1 组织应建立数据质量管理组织体系，明确数据管理责任部门及相关岗位职责。
13.1.2 组织应指定数据管理负责人或数据质量负责人，统筹协调数据采集、处理、存储、使用及处置等环节的管理工作。
13.1.3 组织应建立跨部门协作机制，促进业务部门、技术部门及数据管理部门之间的信息共享与协同管理。

13.2 制度建设

13.2.1 组织应建立数据管理制度，包括数据采集管理、数据处理管理、数据安全管理及数据质量管理等制度。
13.2.2 数据管理制度应明确数据生命周期各阶段的管理要求，并规定数据管理流程和责任分工。
13.2.3 数据管理制度应根据法律法规变化和业务发展情况进行定期评估和更新。

13.3 技术保障

13.3.1 组织应建立支持数据质量管理的信息技术平台，对数据采集、处理、存储及使用过程进行管理。
13.3.2 数据管理系统应具备数据质量监测、数据访问控制及日志记录等功能。
13.3.3 组织应采取必要的技术措施保障数据安全，包括身份认证、访问控制、数据加密及安全审计等。

13.4 人员管理与培训

13.4.1 组织应对参与数据管理和数据处理活动的人员进行管理，并明确其岗位职责。
13.4.2 组织应定期开展数据安全、数据质量及相关法律法规培训，提高相关人员的数据管理能力。
13.4.3 对涉及敏感数据处理的岗位人员，应加强安全管理，并签署保密协议。

13.5 监督与审计

13.5.1 组织应建立数据质量监督机制，对数据管理活动进行定期检查。
13.5.2 组织应开展数据质量审计，对数据质量管理制度执行情况进行评估。
13.5.3 对发现的数据质量问题或安全风险，应及时采取整改措施，并进行持续改进。

13.6 管理记录与持续改进

13.6.1 数据管理活动应建立记录机制，对数据管理过程中的关键活动进行记录。
13.6.2 管理记录宜包括数据质量检查记录、审计记录及问题整改记录等内容。
13.6.3 组织应根据数据质量评估和监督审计结果，持续优化数据管理流程，以提升数据可信质量水平。