

团 体 标 准

T/SA XXXX—XXXX

融媒体智能交互与数据要素流通技术体系 架构

Integrated media intelligent interaction and data element circulation technology
system architecture

（征求意见稿）

XXXX—XX—XX 发布

XXXX—XX—XX 实施

福建省标准化服务行业协会 发 布

目 次

前言 II

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 总体架构 2

 4.2 总体架构组成 3

 4.3 各层级核心功能 3

 4.4 多智能体核心能力域 5

 4.5 架构运行流程 5

5 智能交互层 5

 5.1 总体技术要求 6

 5.2 多智能体协同技术要求 6

 5.3 多模态交互技术要求 7

 5.4 智能推荐技术要求 8

 5.5 虚拟交互技术要求 8

6 数据资源层 9

 6.1 数据分类与范围 9

 6.2 数据湖技术要求 9

 6.3 数据采集技术要求 10

 6.4 数据处理技术要求 10

 6.5 数据流通技术要求 11

7 技术支撑层 11

 7.1 人工智能引擎技术要求 11

 7.2 云计算平台技术要求 12

 7.3 大数据处理平台技术要求 12

 7.4 区块链技术要求 13

 7.5 通信网络技术要求 13

8 安全与合规 14

 8.1 网络安全要求 14

 8.2 数据安全要求 14

 8.3 隐私保护要求 14

 8.4 合规要求 15

 8.5 安全测评要求 15

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××××提出。

本文件由福建省标准化服务行业协会归口。

本文件起草单位：

本文件主要起草人：

融媒体智能交互与数据要素流通技术体系架构

1 范围

本文件规定了融媒体智能交互与数据要素流通技术体系的术语和定义、体系架构总体框架、智能交互层技术要求、数据资源层技术要求、技术支撑层要求、管理保障体系及安全与合规要求等核心技术内容，明确了融媒体场景下智能交互实现与数据要素安全高效流通的全流程技术标准。

本文件适用于融媒体领域内涉及智能交互与数据要素流通的各类活动及相关主体，具体包括媒体内容生产机构、技术研发企业、平台运营方、数据服务提供商等参与融媒体业务的组织和单位。涵盖从融媒体智能交互中台的搭建、多智能体协同运作，到数据湖的建设、多源异构数据的采集与集成，以及数据要素在不同系统和业务环节间的流通与应用等场景；无论是传统媒体的数字化转型，还是新兴融媒体业务的开展，只要涉及智能交互技术和数据要素流通过程，均在本规范的适用范畴之内。

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 26816 信息资源核心元数据
GB/T 35273 信息安全技术 个人信息安全规范
GB/T 37988 信息安全技术 数据安全能力成熟度模型
GB/T 43707 科学数据溯源元数据
GY/T 354 广播电视台融合媒体云平台 总体架构
GY/T 355 广播电视台融合媒体云平台 接口协议
YD/T 3740 互联网新技术新业务安全评估 互联网资源协作服务
T/AI 126.3 数据湖 第3部分：数据流通管理规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

融媒体 converged media

整合报纸、广播、电视、互联网等多种媒体形态，融合内容生产、传播、交互、服务等环节，依托数字化技术实现跨平台、跨终端、全场景内容服务的媒体形态。

3.2

智能交互 intelligent interaction

基于人工智能、自然语言处理、计算机视觉等技术，实现用户与融媒体平台之间多模态（文字、语音、图像、手势）、个性化、实时化的互动交流，包括内容推荐、智能问答、语音控制、虚拟交互等服务。

3.3

数据要素流通 data element circulation

融媒体数据要素在合规前提下，在平台内部、平台之间及跨行业场景中，实现采集、清洗、加工、共享、交易、应用的全流程流通过程，核心是实现数据价值最大化，符合国家“数据要素”政策框架要求。

3.4

技术体系架构 technical system architecture

支撑融媒体智能交互与数据要素流通的技术框架，包括智能交互层、数据资源层、技术支撑层等功能层级，以及管理保障体系、安全与合规体系，各层级协同实现融媒体业务全流程技术支撑，与GY/T 354“总体架构”表达保持一致。

3.5

多智能体协同 multi-agent collaboration

融媒体智能交互中台内，多个具备独立功能的智能体\，通过统一的编排机制实现信息交互、任务分配、决策协同的运作模式，固化协同、记忆、规划、执行四大核心能力域。

3.6

智能体编排 agent orchestration

通过标准化规则和算法，对多智能体的任务流程、交互逻辑、资源分配进行动态调度与管理，确保智能体间高效协同响应业务需求的过程。

3.7

数据湖 data lake

用于存储、管理融媒体多源异构数据（结构化、半结构化、非结构化）的集中式数据存储与处理平台，具备数据汇聚、治理、分析等功能，适配融媒体业务场景特性，符合 GB/T 43707 要求。

3.8

数据脱敏 data desensitization

通过屏蔽、替换、加密等技术手段，对敏感数据\进行处理，确保数据在流通中不泄露个人隐私或商业秘密的技术过程，符合GB/T 37988要求。

3.9

智能推荐 intelligent recommendation

基于用户画像、内容标签、行为数据，通过算法模型自动为用户推荐个性化融媒体内容的技术功能。

3.10

接口协议 interface protocol

定义融媒体系统间数据传输、交互的标准规则，包括数据格式、传输方式、权限校验等内容，遵循GY/T 355接口协议风格，是实现系统互操作的核心依据。

3.11

虚拟交互场景 virtual interaction scenario

基于虚拟现实（VR）、增强现实（AR）、数字人等技术构建的虚拟融媒体交互环境，如虚拟演播厅、虚拟发布会、虚拟客服等。

3.12

数据要素血缘 data element lineage

记录数据要素从采集、加工、流转到应用的全流程链路信息，包括数据来源、处理过程、关联关系等，支持数据追溯与审计，符合T/AI 126.3相关要求。

3.13

智能代理 intelligent agent

具备自主感知、决策、执行能力的软件实体，是多智能体协同体系的基本构成单元，具备协同、记忆、规划、执行核心能力。

4 总体架构

4.1.1 核心原则

4.1.1.1 政策适配性

深度契合国家“数据二十条”等“数据要素”政策框架，确保技术体系架构符合《数据安全法》《个人信息保护法》等法律法规要求，实现政策落地与技术实践协同。

4.1.1.2 技术先进性

采用人工智能、大数据、5G、区块链等前沿技术，参考GY/T 354、YD/T 3740 等行业标准，确保架构具备前瞻性与可扩展性；借鉴多智能体互联互通团标趋势，强化智能体协同能力。

4.1.1.3 数据驱动性

以数据要素为核心，贯穿智能交互、内容生产、传播运营全流程，通过规范的数据治理挖掘数据价值，符合数据要素流通的政策导向。

4.1.1.4 安全合规性

将安全与合规要求嵌入各层级，构建覆盖数据全生命周期的安全防护体系，保障数据要素安全流通，对齐“数据入湖安全技术要求”类团标。

4.1.1.5 开放兼容性

支持与现有融媒体系统、第三方平台的对接，遵循GY/T 355接口协议风格与T/AI 126.3数据管理要求，消除技术壁垒；兼容蓝光存储、对象存储等多形态存储方式。

4.1.1.6 用户中心性

围绕用户需求设计智能交互功能，优化交互流程与体验，实现个性化、场景化服务供给。

4.1.1.7 标准衔接性

与GB/T 43707、GB/T 26816深度衔接，扩展媒体场景元数据模型；统一数据采集、映射、质量与血缘要求，符合T/AI 126系列标准。

4.2 总体架构组成

4.2.1 三大功能层级

4.2.1.1 智能交互层

作为用户与融媒体平台的直接交互入口，核心功能包括多模态交互实现、智能推荐服务、虚拟交互场景搭建、多智能体协同编排，是提升用户体验的关键层级；重点固化多智能体协同、记忆、规划、执行能力域。

4.2.1.2 数据资源层

负责融媒体数据要素的全生命周期管理，核心功能包括数据湖建设、数据采集、数据存储、数据处理、数据流通、数据应用，是技术体系的数据核心；严格遵循GB/T 43707、T/AI 126.3等标准要求。

4.2.1.3 技术支撑层

为上层层级提供基础技术支撑，核心包括人工智能引擎、云计算平台、大数据处理平台、区块链技术、通信网络，是保障上层功能实现的基础；支撑多智能体协同与数据要素流通的底层技术需求。

4.2.2 两大保障体系

4.2.2.1 管理保障体系

涵盖技术管理、运维管理、人员培训等内容，确保技术体系规范运行、高效运维，支撑产业协同与创新生态构建。

4.2.2.2 安全与合规体系

覆盖网络安全、数据安全、隐私保护、合规审计等要求，为智能交互与数据要素流通提供全流程安全保障；覆盖鉴权、脱敏、留痕、备案等安全要求。

4.3 各层级核心功能

4.3.1 智能交互层

4.3.1.1 多模态交互模块

支持文字、语音、图像、手势等交互模态的解析与响应，实现多模态信息的转换与协同。

4.3.1.2 多智能体协同模块

明确多智能体的功能边界、交互规则与编排机制，固化协同、记忆、规划、执行四大能力域，实现推荐、问答、控制等智能体的高效协作。

4.3.1.3 智能推荐模块

基于多维度数据实现个性化内容推荐，支持冷启动、负反馈处理与多场景适配。

4.3.1.4 虚拟交互模块

构建 VR/AR 场景与数字人交互功能，提供沉浸式交互体验。

4.3.1.5 交互数据采集模块

采集用户交互行为、偏好等数据，为数据资源层提供输入；遵循数据采集标准化要求。

4.3.2 数据资源层

4.3.2.1 数据湖平台

构建融媒体专属数据湖，规范元数据模型（衔接GB/T 43707与GB/T 26816）与治理规则，兼容多形态存储。

4.3.2.2 数据采集模块

通过多方式采集多源异构数据，确保数据来源合法、格式规范；统一采集标准与映射规则，符合T/AI 126.3 要求。

4.3.2.3 数据处理模块

实现数据清洗、脱敏、标注、建模等治理操作，形成标准化数据要素；保障数据质量与血缘可追溯。

4.3.2.4 数据流通模块

通过标准化接口实现数据安全共享与交易，支持平台内与跨平台流通；遵循GY/T 355接口协议风格。

4.3.2.5 数据应用模块

为智能交互、内容生产、运营决策提供数据服务。

4.3.3 技术支撑层

4.3.3.1 人工智能引擎

提供NLP、计算机视觉、推荐算法等AI能力，支撑智能交互与数据处理；为多智能体协同提供算法支撑。

4.3.3.2 云计算平台

提供弹性计算、存储资源，支持系统按需扩展与高效部署；兼容多形态存储资源管理。

4.3.3.3 大数据处理平台

提供批处理、流处理等数据计算能力，支撑海量数据高效处理；满足数据湖数据治理与分析需求。

4.3.3.4 区块链组件

实现数据溯源、版权保护与交易存证，保障数据流通可信；支撑数据要素流通的溯源与确权。

4.3.3.5 通信网络适配

支持5G、4G、Wi-Fi 等网络环境，优化交互响应与数据传输效率；符合YD/T 3740要求。

4.3.4 保障体系

4.3.4.1 管理保障体系

包括技术标准管理、系统运维管理、人员能力培养，确保标准落地与体系稳定运行；建立标准更新与衔接机制。

4.3.4.2 安全与合规体系

包括网络边界防护、数据分级保护、隐私合规管控、操作审计追溯，防范安全风险；覆盖鉴权、脱敏、留痕、备案全流程。

4.4 多智能体核心能力域

多智能体协同体系固化四大核心能力域，确保智能交互高效响应，具体要求见表1。

表1 多智能体核心能力域要求表

能力域	核心功能	技术要求	应用场景
协同能力	智能体间任务分配、信息交互、结果整合	支持基于事件流的实时通信；遵循统一接口协议；协同响应时间≤200ms	多模态交互协同、跨场景服务联动
记忆能力	用户交互历史存储、偏好挖掘、场景记忆	支持长期数据存储与快速检索；记忆数据保留符合合规要求；用户画像更新延迟≤10s	个性化推荐、连续对话、场景化服务适配
规划能力	复杂任务拆解、流程优化、冲突协调	支持动态任务规划与调整；冲突解决准确率≥90%；任务完成率≥95%	多步骤交互执行、跨模块业务协同
执行能力	指令解析、操作执行、结果反馈	指令执行响应时间≤300ms；执行准确率≥98%；结果反馈及时率 100%	语音控制、智能操作、服务自动触发

4.5 架构运行流程

4.5.1 数据采集阶段

技术支撑层提供采集工具，数据资源层通过“自动采集+人工采集+用户贡献”方式，收集交互、内容、业务数据，执行合规校验与格式标准化；遵循T/AI 126.3数据采集要求。

4.5.2 数据治理阶段

采集数据接入数据湖，完成元数据注册（衔接GB/T 43707与GB/T 26816）、清洗、脱敏、建模，形成标准化数据要素，满足“准确率≥99.5%、完整性≥99%”要求；建立数据血缘追溯机制。

4.5.3 智能交互阶段

智能交互层调用AI引擎，基于标准化数据要素提供多模态交互服务，多智能体协同模块通过编排机制实现任务协同，四大能力域协同运作，确保交互实时性。

4.5.4 数据流通阶段

在安全合规保障下，通过标准化接口（遵循GY/T 355风格）实现数据要素流通，平台内流通延迟≤1s，跨平台流通执行溯源与脱敏；符合数据要素流通政策与安全标准。

4.5.5 管理保障阶段

两大保障体系全程介入，监控运行状态、管控安全风险，确保各环节规范高效；定期开展标准衔接与体系优化。

5 智能交互层

5.1 总体技术要求

5.1.1 性能要求

5.1.1.1 响应时间

文字交互≤300ms,语音交互≤1s,图像交互≤2s,VR/AR 交互≤300ms,多智能体协同响应≤200ms。

5.1.1.2 准确率

语音识别清晰环境≥95%、嘈杂环境≥90%，自然语言理解≥92%，推荐点击率≥15%，多智能体协同任务完成准确率≥90%，指令执行准确率≥98%。

5.1.1.3 并发能力

支持同时在线用户≥10 万，峰值并发请求≥5万次/秒，系统可用性≥99.9%。

5.1.1.4 兼容性

支持PC端、移动端（iOS/Android）、智能终端，适配不同系统与屏幕尺寸；支持与现有融媒体系统无缝对接。

5.1.2 体验要求

5.1.2.1 个性化

基于用户画像适配交互模态、内容风格与响应方式，支持自定义偏好设置；利用多智能体记忆能力域实现长期偏好跟踪。

5.1.2.2 连续性

支持跨终端会话同步，用户切换设备后可延续交互流程，无需重复操作；依赖多智能体协同与记忆能力保障。

5.1.2.3 容错性

对用户输入错误（语音模糊、文字错别字）具备修正能力，降低操作门槛；通过多智能体规划能力优化交互容错机制。

5.1.2.4 可定制

支持针对政务、新闻、文化等场景定制交互流程与功能模块；通过智能体编排机制实现场景化适配。

5.2 多智能体协同技术要求

5.2.1 智能体划分与功能边界

核心智能体类型及功能边界明确，避免功能重叠，具体见表2。

表2 核心智能体类型及功能边界表

智能体类型	核心功能	能力域支撑	交互对象
推荐智能体	个性化内容推荐、兴趣挖掘、冷启动处理	协同、记忆、规划、执行	用户、内容管理系统、数据资源层
问答智能体	智能问答、知识查询、多轮对话	协同、记忆、规划、执行	用户、知识库系统、交互数据采集模块
控制智能体	语音控制、设备联动、操作执行	协同、规划、执行	用户、终端设备、业务系统
虚拟交互智能体	数字人交互、VR/AR 场景控制	协同、执行	用户、虚拟场景引擎
编排智能体	多智能体任务调度、冲突协调、流程优化	协同、规划、执行	其他所有智能体、业务调度系统

5.2.2 协同机制

5.2.2.1 信息交互机制

采用标准化事件流格式实现智能体间信息传递，支持同步/异步通信；事件流包含消息标识、发送方、接收方、内容体、时间戳等核心字段。

5.2.2.2 任务分配规则

基于任务类型、智能体负载、响应优先级动态分配任务；优先级分为P0（紧急）、P1（高）、P2（中）、P3（低）四级。

5.2.2.3 冲突协调策略

当多个智能体竞争同一资源或任务时，采用“优先级优先 + 负载均衡”策略协调；冲突解决延迟≤100ms。

5.2.2.4 能力域固化

严格落实协同、记忆、规划、执行四大能力域要求，其中记忆能力需支持用户交互历史、偏好数据的长期存储（符合数据留存合规要求）与快速调用。

5.2.3 接口协议

遵循GY/T 355“接口协议”风格，定义统一的API与事件流规范，支持JSON/XML数据格式；API设计符合RESTful规范。核心接口类型及要求见表3。

表3 多智能体协同核心接口表

接口类型	核心功能	请求参数	响应要求
智能体注册接口	智能体接入协同体系、能力声明	智能体 ID、类型、能力描述、接口地址	注册结果、分配标识、通信密钥
任务分发接口	编排智能体向执行智能体分配任务	任务 ID、类型、参数、优先级	任务接收确认、执行进度回调地址
信息交互接口	智能体间数据传递、状态同步	消息 ID、发送方 ID、接收方 ID、内容	消息接收确认、处理结果（同步 / 异步）
结果整合接口	汇总多智能体执行结果、返回给用户	任务 ID、各智能体执行结果	整合后结果、置信度、建议操作

5.3 多模态交互技术要求

5.3.1 文字交互

支持中文、英文等多语言交互，支持emoji表情识别与回应，多轮对话能力≥10轮；利用问答智能体协同能力实现对话连贯性。具备关键词提取、意图识别、语义相似度匹配功能，准确响应用户查询、咨询、指令类需求；意图识别准确率≥92%。

5.3.2 语音交互

5.3.2.1 语音识别

支持普通话及5种以上方言（粤语、四川话、上海话等），具备静音检测与端点检测功能；清晰环境识别准确率≥95%，嘈杂环境≥90%。

5.3.2.2 语音合成

支持语速（0.8~1.2倍速）、语调、音色（≥5种）调整，合成语音自然度≥4.5分（5分制）；支持情感语音合成（喜悦、中性、严肃）。

5.3.2.3 语音控制

支持“播放新闻”“暂停视频”等20种以上常用指令，识别准确率 $\geq 93\%$ ；通过控制智能体实现指令执行与反馈。

5.3.3 图像交互

5.3.3.1 图像识别

支持物体、场景、文字、二维码识别，准确率 $\geq 94\%$ ；人脸识别仅用于授权场景，准确率 $\geq 99\%$ ，符合GB/T 35273个人信息保护要求。

5.3.3.2 图像搜索

支持以图搜图功能，响应时间 $\leq 3s$ ，相似结果准确率 $\geq 90\%$ （前10条）；结合推荐智能体实现关联内容推荐。

5.3.3.3 手势交互

支持点击、滑动、暂停等10种以上常用手势，识别距离 $\leq 3m$ ，误识别率 $\leq 5\%$ ；通过控制智能体实现手势指令映射与执行。

5.4 智能推荐技术要求

5.4.1 算法要求

5.4.1.1 支持协同过滤、内容推荐、Transformer深度学习推荐等多算法融合，可按场景自动切换；推荐智能体具备算法选择与优化的规划能力。

5.4.1.2 基于用户行为、兴趣标签、场景信息（时间、地点、设备）实现多维度个性化推荐；利用记忆能力域存储用户长期偏好。

5.4.1.3 推荐多样性基尼系数 ≥ 0.7 ，用户行为发生后推荐列表更新时间 $\leq 10s$ ；支持推荐结果解释功能，提升用户信任度。

5.4.2 功能要求

5.4.2.1 冷启动支持

新用户推荐点击率 $\geq 8\%$ ，支持兴趣标签引导功能；推荐智能体通过协同能力调用内容标签数据。

5.4.2.2 负反馈处理

用户标记“不感兴趣”后，同类内容推荐占比 $\leq 5\%$ （30天内）；记忆能力域记录用户负反馈历史。

5.4.2.3 多场景适配

支持首页、详情页、搜索结果、离线推送等多场景推荐，适配准确率 $\geq 90\%$ ；通过规划能力域实现场景化推荐策略调整。

5.5 虚拟交互技术要求

5.5.1 数字人交互

5.5.1.1 形象要求

支持写实型、卡通型风格，面部表情 ≥ 10 种，肢体动作 ≥ 20 种，形象逼真度 ≥ 4.2 分；支持自定义形象配置。

5.5.1.2 交互能力

对话响应时间 $\leq 1.5s$ ，表情同步延迟 $\leq 200ms$ ，多轮对话准确率 $\geq 90\%$ ；虚拟交互智能体协同问答智能体实现对话能力。

5.5.1.3 应用场景

支持虚拟主播、客服、向导，虚拟主播直播稳定性 $\geq 99.9\%$ ；支持多平台同步直播。

5.5.2 VR/AR 交互

5.5.2.1 场景构建

VR场景分辨率≥4K，帧率≥60fps，沉浸感评分≥4.3分；AR 场景识别准确率≥95%；支持场景自定义编辑。

5.5.2.2 交互功能

支持用户自由移动、互动操作，响应时间≤300ms，支持多人协同（≤5人）；通过虚拟交互智能体实现多人协同调度。

6 数据资源层

6.1 数据分类与范围

6.1.1 数据分类

融媒体数据要素按类型分为三类，具体内容见表4，分类规则符合 T/AI 126.3要求。

表4 融媒体数据要素分类及内容表

数据类型	核心内容	示例数据	流通优先级	敏感级别
用户数据	基本信息、行为数据、偏好数据	脱敏手机号、点击记录、兴趣标签	中	敏感
内容数据	元数据、内容本体、标签数据	新闻标题、视频文件、关键词标签	高	一般
业务数据	传播数据、运营数据、交易数据	曝光量、用户活跃度、付费记录	低-中	一般-敏感

6.1.2 数据范围

- 6.1.2.1 支持结构化（数据库表）、半结构化（JSON/XML）、非结构化（音视频/图片/文本）数据，音视频支持 MP4、MP3、AVI 等主流格式；文本支持 TXT、DOC 等格式。
- 6.1.2.2 支持 PB 级存储与处理，单平台日均数据增量≥10TB 时稳定运行；兼容海量小文件与大文件存储。
- 6.1.2.3 包括平台采集、用户上传、第三方接入、合作伙伴共享，来源需可追溯；采集过程符合 GB/T 35273 个人信息保护要求。

6.2 数据湖技术要求

6.2.1 元数据模型

与GB/T 43707、GB/T 26816深度衔接，扩展媒体场景元数据字段，包括数据来源、格式、敏感级别、内容标签、血缘信息、流通记录等。元数据注册完整率≥99%，更新频率≤24小时，支持语义检索与血缘追踪；元数据存储符合T/AI 126.3要求。元数据分类及核心字段见表5。

表5 融媒体元数据分类及核心字段表

元数据类型	核心字段	数据来源	更新频率
核心元数据	数据 ID、名称、类型、格式、大小、创建时间	数据采集模块	实时
溯源元数据	采集来源、采集工具、处理流程、操作人员	数据处理模块、流通模块	实时更新
媒体特色元数据	内容标签、传播渠道、交互量、版权信息	内容管理系统、交互数据采集模块	每小时更新
流通元数据	共享对象、流通方式、脱敏状态、使用目的	数据流通模块	实时

6.2.2 存储适配

6.2.2.1 兼容蓝光存储、对象存储（S3/OSS）、关系型数据库、分布式文件系统等多形态存储；按数据热度自动分级存储，热数据（3个月内）SSD存储、冷数据HDD/蓝光存储自动迁移。

6.2.2.2 存储系统年故障率 $\leq 0.01\%$ ，数据多副本（ ≥ 3 份）存储，单副本损坏可快速恢复（恢复时间 ≤ 1 小时）；支持数据存储加密（AES-256），符合安全要求。

6.2.2.3 存储架构支持横向扩展，扩展过程不影响业务运行；支持存储资源弹性伸缩，响应时间 ≤ 5 分钟。

6.3 数据采集技术要求

6.3.1 采集方式

6.3.1.1 自动采集

通过爬虫、API接口、传感器采集，频率可配置（实时/定时），单接口采集速率 ≥ 1000 条/秒；API接口遵循GY/T 355风格。

6.3.1.2 人工采集

支持编辑手动录入，具备格式校验与完整性校验功能；支持批量导入与审核流程。

6.3.1.3 用户贡献

用户上传内容需经智能+人工审核，审核通过率 $\geq 95\%$ ；审核标准符合内容安全规范。

6.3.1.4 采集标准化

统一采集字段映射规则，支持数据格式自动转换；采集工具具备身份认证与权限控制功能。

6.3.2 技术要求

6.3.2.1 采集精度

结构化数据准确率 $\geq 99.8\%$ ，非结构化数据（图文提取）准确率 $\geq 95\%$ ；支持采集数据质量实时校验。

6.3.2.2 容错能力

网络中断、格式错误时支持断点续采，重试次数 ≤ 5 次；采集失败数据自动记录，支持手动补采。

6.3.2.3 合规性

采集第三方数据需获得授权，签订数据使用协议；用户信息采集符合GB/T 35273要求，明确告知用户采集目的与用途。

6.3.2.4 血缘记录

采集阶段记录数据初始血缘信息，包括来源地址、采集时间、采集设备、操作人员等，支持全流程追溯。

6.4 数据处理技术要求

6.4.1 数据治理

6.4.1.1 清洗

支持去重、缺失值填充、异常值剔除，1000万条数据清洗时间 ≤ 1 小时，清洗准确率 $\geq 99.5\%$ ；清洗规则可配置、可扩展。

6.4.1.2 脱敏

支持静态与动态脱敏，敏感数据识别准确率 $\geq 98\%$ ，脱敏后保留业务可用性；脱敏算法符合GB/T 37988要求，支持多种脱敏策略（屏蔽、替换、加密）。

6.4.1.3 标注与建模

智能标注准确率 $\geq 90\%$ ，支持人工修正；用户画像、内容标签模型迭代周期 ≤ 7 天；模型训练数据符合合规要求。

6.4.1.4 数据质量控制

建立数据质量评估指标体系，包括准确率、完整性、一致性、时效性，质量达标率 $\geq 99\%$ ；支持数据质量异常报警与自动修复，修复成功率 $\geq 95\%$ 。

6.4.2 数据血缘管理

6.4.2.1 血缘记录

完整记录数据从采集、清洗、加工、流通到应用的全流程血缘信息，包括数据流转路径、处理规则、关联数据等。

6.4.2.2 血缘查询

支持按数据ID、名称、时间范围查询血缘信息，查询响应时间 $\leq 3s$ ；支持可视化展示血缘链路。

6.4.2.3 血缘应用

血缘信息用于数据追溯、问题定位、合规审计，保留期限 ≥ 3 年；符合T/AI 126.3数据血缘要求。

6.5 数据流通技术要求

6.5.1 流通模式

6.5.1.1 内部流通

平台内业务系统间流通延迟 $\leq 1s$ ，支持实时数据同步；采用标准化接口与事件流实现数据传输。

6.5.1.2 跨平台流通

通过标准化接口实现，接口兼容性 $\geq 95\%$ ，支持数据格式转换；遵循GY/T 355接口协议风格。

6.5.1.3 数据交易

交易记录上链存证（基于区块链技术），交易数据需脱敏处理，支持收益分配智能合约；交易流程符合“数据要素”政策要求。

6.5.2 流通管控

6.5.2.1 权限控制

基于RBAC模型实现字段级权限管控，不同角色仅能访问授权数据；支持权限动态调整与审计。

6.5.2.2 流向追溯

记录数据来源、流通路径、使用目的、操作人，追溯记录保留 ≥ 3 年；支持按数据ID、时间、对象查询流通记录。

6.5.2.3 熔断机制

数据泄露、接口故障时自动熔断，停止流通并报警；熔断触发后支持手动恢复，恢复前需安全校验。

6.5.2.4 合规备案

跨平台流通数据需进行合规备案，备案信息包括数据类型、流通范围、脱敏方式、安全措施等。

7 技术支撑层

7.1 人工智能引擎技术要求

7.1.1 核心功能

7.1.1.1 NLP 引擎：

支持分词、语义理解、文本生成、情感分析，处理速度 ≥ 1000 句/秒；支撑问答智能体、推荐智能体的自然语言处理需求。

7.1.1.2 CV 引擎

支持图像识别、目标检测、OCR、视频分析，1080P 图像处理速度 ≥ 100 张/秒；支撑图像交互、虚拟交互智能体的视觉处理需求。

7.1.1.3 推荐引擎

支持多算法融合，响应时间 $\leq 100\text{ms}$ ，推荐准确率 $\geq 85\%$ ；为推荐智能体提供算法支撑。

7.1.1.4 多智能体引擎

提供智能体注册、编排、协同的底层技术支撑，支持 ≥ 100 个智能体同时运行；固化协同、记忆、规划、执行能力域的技术实现。

7.1.2 性能要求

7.1.2.1 并发处理

支持同时处理 ≥ 1 万条请求，峰值并发 ≥ 5 万条/秒；支持请求优先级调度。

7.1.2.2 模型迭代

支持快速部署、灰度发布，模型更新不影响业务运行；支持模型版本管理与回滚。

7.1.2.3 兼容性

支持TensorFlow、PyTorch等主流框架，模型格式支持ONNX；支持与第三方AI服务接口对接。

7.2 云计算平台技术要求

7.2.1 计算资源

计算资源要求如下：

- a) 支持容器化（Docker/K8s）部署，容器启动时间 $\leq 30\text{s}$ ，调度延迟 $\leq 500\text{ms}$ ；支持弹性扩缩容，响应时间 ≤ 5 分钟；
- b) 计算资源利用率 $\geq 70\%$ ；支持CPU、GPU混合调度，满足AI模型训练与推理需求；
- c) 支持多租户隔离，租户间资源互不干扰；提供资源使用监控与统计功能。

7.2.2 存储资源

存储资源要求如下：

- a) 支持对象存储、块存储、文件存储，对象存储容量可无限扩展，块存储IOPS ≥ 10 万；兼容蓝光存储等多形态存储；
- b) 存储加密：静态数据AES-256加密，传输数据TLS 1.3加密，密钥每3个月更换一次；支持密钥安全管理与备份；
- c) 存储备份：支持自动备份与手动备份，备份策略可配置；备份数据保留期限可设置，支持异地备份。

7.3 大数据处理平台技术要求

7.3.1 处理功能

7.3.1.1 批处理

基于Hadoop/Spark，吞吐量 $\geq 100\text{GB/小时}$ ，处理延迟 ≤ 2 小时；支持大规模数据清洗与建模。

7.3.1.2 流处理

基于Flink，延迟 $\leq 500\text{ms}$ ，吞吐量 ≥ 10 万条/秒；支持实时数据处理与分析。

7.3.1.3 数据治理

支持数据血缘分析、质量监控、元数据管理，血缘分析准确率 $\geq 98\%$ ；与数据湖平台深度协同。

7.3.1.4 数据查询

支持SQL、NoSQL查询，复杂查询响应时间 $\leq 10\text{s}$ ；支持多维度数据分析与可视化。

7.3.2 性能要求

性能要求如下：

- a) 扩展性：集群节点可横向扩展，最大规模 ≥ 100 节点，扩展后性能线性提升；支持动态扩缩容；
- b) 稳定性：年故障率 $\leq 0.1\%$ ，故障自动转移，恢复时间 ≤ 10 分钟；支持集群健康监控与预警；
- c) 兼容性：支持多种数据格式与存储系统接入；符合 T/AI 126.3 数据处理要求。

7.4 区块链技术要求

7.4.1 核心功能

7.4.1.1 账本功能

支持数据流通、权限变更、交易记录上链，账本不可篡改、可追溯；支撑数据要素流通的溯源与确权。

7.4.1.2 智能合约

支持数据交易自动执行、收益分配、权限管控，合约执行准确率 $\geq 99.9\%$ ；合约代码需经过安全审计。

7.4.1.3 身份认证

支持分布式身份认证（DID），跨平台身份互通，认证响应 $\leq 500\text{ms}$ ；为数据流通提供可信身份支撑。

7.4.1.4 存证功能

支持内容版权、数据交易、操作日志存证，存证时间 $\leq 10\text{s}$ ；存证数据可用于合规审计。

7.4.2 性能要求

交易吞吐量 ≥ 1000 笔/秒，区块确认时间 ≤ 3 秒，支持并行交易；满足数据流通高频交易需求。支持SM2/SM3加密算法，拜占庭容错率 $\geq 1/3$ ；具备抗篡改、抗攻击能力。支持与现有融媒体中心、数据流通系统对接；支持联盟链部署模式。

7.5 通信网络技术要求

7.5.1 网络适配

支持5G、4G、Wi-Fi 6，5G环境下下行速率 $\geq 1\text{Gbps}$ ，端到端延迟 $\leq 10\text{ms}$ ；符合YD/T 3740要求。支持网络切换无感知，切换延迟 $\leq 100\text{ms}$ ，弱网环境下智能调整服务策略。支持边缘计算节点部署，降低交互响应延迟；边缘节点与云端数据同步延迟 $\leq 1\text{s}$ 。

7.5.2 网络安全

支持SDN网络架构，VLAN隔离、VPN接入，网络攻击识别准确率 $\geq 99\%$ ；防范DDoS、SQL注入等常见攻击。

实时监测网络带宽、延迟、丢包率，异常时自动报警并调整传输策略；支持网络流量管控与优先级调度。

支持数据传输加密与完整性校验；网络设备具备身份认证与权限控制功能。

8 安全与合规

8.1 网络安全要求

8.1.1 边界安全

部署防火墙、IDS/IPS，拦截SQL注入、XSS、DDoS等攻击，拦截率 $\geq 99\%$ ；定期更新防护规则库。网络边界采用“最小权限”原则，仅开放必要端口，每月进行端口扫描；禁止高危端口对外开放。远程访问需通过VPN双因素认证，接入日志保留 ≥ 1 年；建立远程访问审计机制。

8.1.2 内部安全

内部网络按业务分区隔离，分区间通过防火墙控制访问；不同业务系统部署在独立网段。部署网络行为审计系统，审计日志保留 ≥ 6 个月，异常检测准确率 $\geq 95\%$ ；对敏感操作进行重点审计。终端设备统一管理，安装安全软件，定期更新病毒库；禁止未经授权设备接入内部网络。

8.2 数据安全要求

8.2.1 分级保护

数据按敏感程度分为三级，敏感数据需加密存储+动态脱敏+严格授权；分级结果定期审核。重要数据采用加密传输与存储，访问需双人授权；敏感数据处理需全程审计，禁止未经授权导出。数据分级保护措施与数据湖管理深度融合，确保不同级别数据采用差异化保护策略。

8.2.2 全生命周期安全

8.2.2.1 采集安全

采集工具安全认证，传输加密（TLS 1.3），数据校验完整性；采集权限严格管控。

8.2.2.2 存储安全

静态数据AES-256加密，传输TLS 1.3 加密，密钥定期更换；存储介质安全管理，报废介质物理销毁。

8.2.2.3 使用安全

敏感数据使用需审批，全程审计，禁止未经授权导出；数据使用环境安全加固。

8.2.2.4 销毁安全

物理销毁或多次覆写，销毁记录保留 ≥ 1 年；数据销毁需经过审批流程。

8.2.2.5 安全追溯

全生命周期操作日志保留 ≥ 3 年，支持按数据ID、操作人、时间范围追溯。

8.3 隐私保护要求

8.3.1 个人信息保护

遵循“合法、正当、必要”原则，仅采集业务相关信息，明确告知用户目的与用途；获得用户明示同意，支持授权撤回。个人信息处理符合GB/T 35273，实现匿名化、去标识化；个人信息存储期限不超过必要时间，用户注销账号后30日内删除或匿名化。禁止向第三方非法提供个人信息，确需提供的需获得用户单独同意并进行脱敏处理；建立个人信息处理合规审计机制。

8.3.2 隐私增强技术

采用联邦学习、安全多方计算等技术，在不泄露原始数据前提下实现数据共享分析；联邦学习模型训练准确率与集中式偏差 $\leq 5\%$ 。安全多方计算结果准确率 $\geq 99\%$ ；采用差分隐私技术，在数据发布与共享中保护个人隐私。虚拟交互场景中，对用户生物特征数据（人脸、语音）进行加密存储与脱敏使用；生物特征数据仅用于授权场景。

8.4 合规要求

8.4.1 法律法规合规

符合《数据安全法》《个人信息保护法》《网络数据安全条例》等要求；满足“数据二十条”政策中数据要素流通相关要求。满足网络安全等级保护二级及以上要求，核心平台需达三级；按等级保护要求开展测评与整改。跨境数据流通需通过安全评估，遵守数据出境相关规定；建立跨境数据流通合规审查机制。

8.4.2 合规审计

自动化审计系统实时记录数据操作、系统访问行为，审计日志保留 ≥ 2 年；审计内容包括数据采集、处理、流通、销毁全流程。每季度自查，每年第三方评估，整改项完成期限：一般项 ≤ 30 天，严重项 ≤ 15 天；自查与评估报告留存归档。针对多智能体协同过程中的数据交互、权限变更进行专项审计；确保智能体操作合规可追溯。

8.5 安全测评要求

8.5.1 定期测评

每年1次全面测评，系统重大升级后专项测评；测评内容包括技术安全、管理安全、合规性。

8.5.2 测评要求

由具备资质机构执行，测评内容包括网络安全、数据安全、隐私保护、接口安全等；测评机构需具备相关行业测评资质。

8.5.3 整改验证

测评发现的问题需限期整改，整改后验证通过方可上线；重大安全问题整改后需进行专项测评。
