ICS 35. xxx
CCS Lxx

团　　体　　标　　准

T/ISC XXX—XXXX

人工智能数据智能体 总体要求

Artificial intelligence data agents — General requirements

（征求意见稿）

XXXX－XX－XX 发布　　XXXX－XX－XX 实施

中 国 互 联 网 协 会　　发布

目　　次

前　　言 III
引　　言 IV
1　　范围 1
2　　规范性引用文件 1
3　　术语和定义 1
3.1 1
3.2 1
4　　符号和缩略语 1
5　　总体原则 2
5.1　　有效原则 2
5.2　　可靠原则 2
5.3　　可信原则 3
5.4　　安全原则 3
5.5　　可控原则 3
5.6　　协同原则 3
6　　数据任务有效性与结果质量要求 3
6.1　　数据任务理解要求 3
6.2　　任务规划与资源选择要求 4
6.3　　数据活动执行要求 4
6.4　　结果校验与质量要求 4
6.5　　结果交付与任务反馈要求 5
7　　性能与可靠性要求 5
7.1　　任务执行性能要求 5
7.2　　运行稳定性要求 6
7.3　　异常识别与容错处理要求 6
7.4　　任务恢复与执行一致性要求 6
7.5　　资源管理与运行监测要求 7
8　　结果可信与过程可追溯要求 8
8.1　　数据来源与处理依据要求 8
8.2　　结果依据与结论核查要求 8
8.3　　不确定性与能力边界说明要求 9
8.4　　执行过程追溯要求 9
8.5　　记录与审计支持要求 9
9　　安全与合规要求 10
9.1　　数据安全与个人信息保护要求 10
9.2　　身份认证与访问控制要求 10
9.3　　模型、工具及执行环境安全要求 11

9.4 输入输出安全与攻击防护要求 11
9.5 安全监测与事件处置要求 11
9.6 合规要求 12
10 自主行为与人工干预要求 12
10.1 任务授权与自主执行边界要求 12
10.2 操作风险分级与事前确认要求 13
10.3 执行过程监督要求 13
10.4 人工接管与任务中止要求 13
10.5 状态恢复与操作回退要求 14
11 互操作与协同性要求 14
11.1 数据平台与工具互操作要求 14
11.2 接口调用与信息交换要求 14
11.3 数据语义一致性要求 15
11.4 多智能体任务协同要求 15
11.5 人机协作与任务交接要求 16
12 测试与验证 ..... **错误！未定义书签。**
12.1 测试方案与验证条件 ..... **错误！未定义书签。**
12.2 数据任务有效性与结果质量验证 ..... **错误！未定义书签。**
12.3 性能与可靠性验证 ..... **错误！未定义书签。**
12.4 结果可信与过程可追溯验证 ..... **错误！未定义书签。**
12.5 安全与合规验证 ..... **错误！未定义书签。**
12.6 自主行为与人工干预验证 ..... **错误！未定义书签。**
12.7 互操作与协同性验证 ..... **错误！未定义书签。**
12.8 测试记录与结果判定 ..... **错误！未定义书签。**
参考文献 17

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。

本文件是T/CIITA x《……》第i部分。T/CIITA x已经发布了以下部分：

——第1部分：……；

——……；

——第n部分：……。

本文件替代T/CIITA x.i《……》，与T/CIITA x.i相比，除结构调整和编辑性改动外，主要技术变化如下：

a)

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：

本文件主要起草人：

本文件及其所代替文件的历次版本发布情况为：

——

引　言

数据已成为数字经济发展的重要基础性资源，人工智能技术与数据处理、数据治理及数据服务的深度融合，正在推动数据处理模式向智能化、自动化和协同化方向发展。随着人工智能技术从模型能力提升向智能体应用拓展，数据智能体作为人工智能技术应用于数据领域的重要形态，能够围绕数据任务开展任务理解、任务规划、数据处理、工具调用及结果校验等活动，为提升数据处理效率、促进数据资源开发利用和推动数据业务智能化转型提供新的技术支撑。

当前，数据智能体在设计开发、系统集成、数据处理、任务执行、安全保障及运行管理等方面仍需要进一步明确通用要求。建立统一的数据智能体总体要求，有助于规范数据智能体的技术实现与应用过程，提升不同产品、解决方案及应用系统之间的兼容性、互操作性和运行可靠性，促进数据智能体规范化建设与应用。

本标准规定了数据智能体在总体设计、数据处理与结果、集成与互操作、性能与可靠性、安全与合规、运行管理、测试与验证及人员管理等方面的要求，适用于企业、政府机构、科研院所等各类组织开展数据智能体的设计、开发、部署、运行及管理，可为数据智能体产品研发、解决方案建设和应用实施提供参考依据。

人工智能 数据智能体 总体要求

1 范围

本标准规定了数据智能体的总体要求。

本标准适用于企业、政府机构、科研院所等各类组织开展数据智能体的设计、开发、部署、运行及管理，也可作为数据智能体产品、解决方案及应用系统的研发、建设、选型和应用的参考依据。

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 45288.1-2025 人工智能 大模型 第1部分：通用要求

GB/Z 242-2026 人工智能 智能体技术要求

GB/T 35273—2020 信息安全技术 个人信息安全规范

GB/T 43697—2024 数据安全技术 数据分类分级规则

3 术语和定义

下列术语和定义适用于本文件。

3.1 元数据 metadata

描述数据的结构、类型、含义、来源、关联关系及其他相关属性的数据。

3.2 数据血缘 data lineage

描述数据在来源、处理、转换及结果形成过程中，输入数据、处理步骤与输出数据之间关联关系的信息。

4 符号和缩略语

下列符号和缩略语适用于本文件。

API 应用程序接口（Application Programming Interface）

5 总体框架

5.1 框架组成

数据智能体总体要求由总体原则、数据任务执行要求、执行保障与全过程控制要求构成，其关系见图1。该框架描述逻辑能力及要求之间的关系，不限定具体产品的部署形态、软件分层或技术实现。

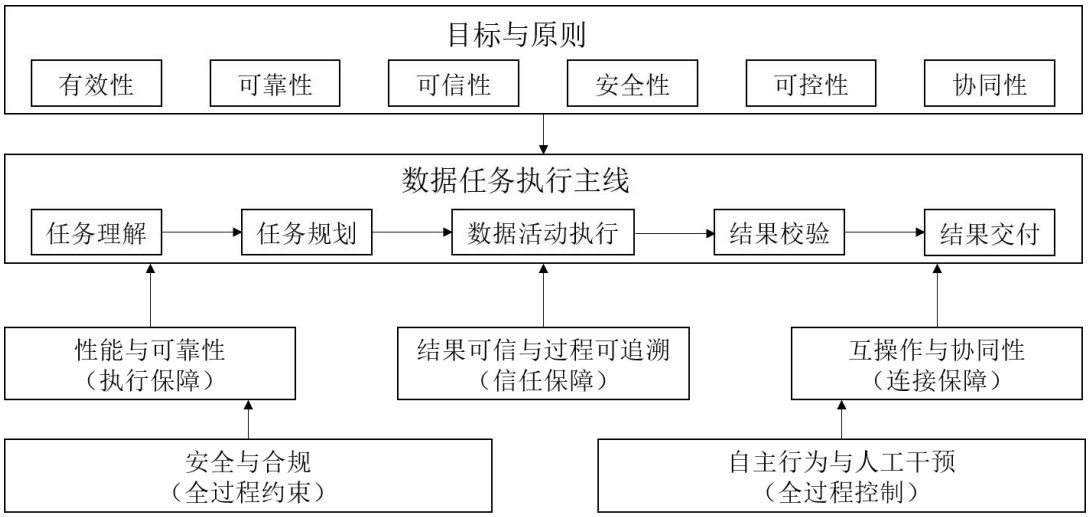


图1 数据智能体总体要求框架

5.2 要求之间的关系

总体原则明确数据智能体的基本目标。数据任务有效性与结果质量要求围绕任务理解、规划、执行、校验及交付构成任务主线；性能与可靠性要求保障任务连续、稳定执行；结果可信与过程可追溯要求支撑结果依据核查和过程审计；互操作与协同性要求支撑跨平台、工具、智能体及人员的信息交换和任务衔接。安全与合规要求、自主行为与人工干预要求贯穿任务全过程。

5.3 适用范围与能力声明

提供方应明确声明数据智能体支持的数据任务类型、输入输出、数据源和工具、运行条件、外部依赖及自主执行范围。通用要求适用于所声明能力的设计和实现；针对特定任务或功能的要求，应在具备相应能力时落实。

6 总体原则

6.1 有效性原则

数据智能体应准确理解数据任务的目标、作用对象、预期结果及约束条件，根据任务需要组织数据资源、模型和工具开展相应活动，形成符合任务要求的结果。对于无法完成或仅部分完成的任务，应如实反馈任务完成情况。

6.2 可靠性原则

数据智能体应在预期运行条件下稳定执行数据任务，并能够识别和处理数据异常、工具调用失败、服务不可用等情况。任务发生中断、重试或恢复时，应保持任务状态和数据处理结果的一致性，避免产生非预期的数据变更或重复操作。

6.3 可信性原则

数据智能体形成的结果应具有可核查的依据，关键数据来源、处理过程和执行行为应可追溯。对于数据质量不足、依据不充分、结论存在不确定性或任务超出自身能力范围的情况，应作出说明，不应将推断结果表述为已核实的事实。

6.4 安全性原则

数据智能体应在数据访问、模型调用、工具调用、任务执行和结果交付等过程中保护数据及相关资源，防范数据泄露、越权操作和恶意指令等风险。涉及数据安全、个人信息保护及其他合规事项时，应遵守适用的法律法规和管理要求。

6.5 可控性原则

数据智能体应在明确的任务范围和授权边界内开展自主或半自主活动，不应自行扩大可访问资源或可执行操作的范围。对于可能产生重大影响的操作，应支持事前确认；在执行过程中，应支持必要的人工监督、接管、中止和后续处置。

6.6 协同性原则

数据智能体应能够按照任务需要与数据平台、模型服务、外部工具、其他智能体及人员进行协作，保持交互信息的语义一致和任务状态的有效传递。在任务分工、执行衔接和结果交接过程中，应明确各参与方的职责边界及必要的交接信息。

7 数据任务有效性与结果质量要求

数据智能体在数据处理中的典型执行动线见图2。任务步骤应根据所承担的数据活动确定；当任务信息不足、结果校验未通过或执行条件变化时，应进行澄清、调整计划、重新执行或转交人工处理。图中动线不要求所有任务均执行全部步骤，也不要求采用固定的内部编排方式。

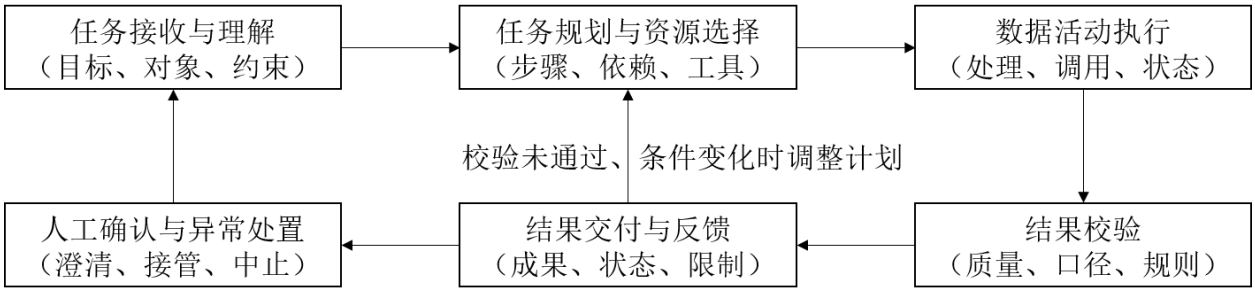


图2 数据智能体数据任务执行动线示意图

7.1 数据任务理解要求

数据智能体应能够理解数据任务的目标、作用对象、预期结果及相关约束，识别完成任务所需的数据和处理要求，内容包括但不限于：

a）宜支持接收自然语言、结构化指令或API等方式提交的数据任务，识别任务目标、数据对象、处理范围、预期输出及完成条件。

b）宜支持结合数据目录、元数据、业务术语及历史任务信息，理解任务涉及的数据字段、指标口径、时间范围和业务含义。

c）宜识别任务描述中可能影响执行结果的歧义、信息缺失或要求冲突，并通过请求澄清或明确执行假设等方式进行处理。

d）宜识别任务是否处于自身声明的能力范围内；对于无法完成的任务，应反馈原因，不应将其作为已完成任务处理。

7.2 任务规划与资源选择要求

数据智能体宜根据任务目标和执行条件，确定任务步骤及其依赖关系，选择适当的数据资源、模型和工具，形成可执行的任务方案，内容包括但不限于：

a）宜支持对涉及多个数据活动的任务进行分解，明确各步骤的处理目标、输入输出及执行顺序。

b）宜支持根据任务需要识别和选择数据源、数据集、元数据、处理规则、模型服务及外部工具，并判断其与任务要求的适配性。

c）宜支持结合数据可用性、工具能力、处理时效及资源条件，确定数据任务的执行方式；当必要的资源或条件不满足时，宜调整任务方案或反馈任务受阻情况。

d）宜支持根据任务执行过程中的数据情况、中间结果或工具返回信息，调整后续执行步骤，确保调整后的方案仍符合任务目标和约束条件。

7.3 数据活动执行要求

数据智能体应按照任务目标和执行方案，组织并执行相应的数据活动，识别执行过程中的数据问题和步骤状态，内容包括但不限于：

a）宜支持根据所承担的任务类型，开展数据采集、标注、治理、发现、传输、交付、加工、分析、运维及安全合规等相关数据活动。

b）宜支持协调数据资源、处理工具和模型服务执行任务，并根据前序步骤的处理结果确定后续步骤的输入和执行条件。

c）宜支持在数据接入和处理过程中，识别缺失、重复、异常、格式不一致及语义冲突等问题，并根据任务要求采取相应的处理措施。

d）宜识别任务步骤的执行状态；当关键步骤未完成、执行结果不符合后续处理条件，或数据问题可能影响任务目标时，应调整执行方式、停止相关步骤或反馈问题。

e）对涉及数据变更或结果交付的任务，宜根据实际执行情况区分已执行、未执行及执行失败的操作，不应将未执行的操作反馈为已完成。

7.4 结果校验与质量要求

数据智能体应依据任务目标、处理规则和预期输出，对数据活动结果进行校验，识别影响结果使用的质量问题，内容包括但不限于：

a）宜支持根据任务类型和结果形式，对结果的准确性、完整性、一致性、有效性及格式符合性等进行适用性检查。

b）对查询、计算、转换、关联、聚合等处理结果，宜支持校验关键处理条件、指标口径和计算结果，识别因参数设置、数据匹配或工具返回异常造成的结果偏差。

c）对数据标注、治理、发现等任务的结果，宜支持依据相应的标注规范、治理规则或发现条件，检查结果是否符合任务要求。

d）宜对校验发现的缺失、重复、冲突、格式错误或明显不合理的结果进行标识，并根据问题性质采取修正、重新处理或提交复核等措施。

e）对未达到任务约定质量要求的结果，不应将其作为符合要求的最终结果直接交付；无法修正时，宜说明存在的问题及其影响。

7.5 结果交付与任务反馈要求

数据智能体应按照任务要求交付处理结果，并反馈任务完成情况、结果适用范围及需要关注的问题，内容包括但不限于：

a）宜支持根据任务要求，以结构化数据、文件、统计结果、分析报告或接口返回结果等形式交付成果，并对交付内容、格式、粒度和时间范围进行适配。

b）宜明确反馈任务的完成状态，区分已完成、部分完成和未完成等情况；对于部分完成或未完成的任务，应说明已完成内容、未完成内容及原因。

c）宜对交付结果中已发现且可能影响使用的数据质量问题、处理限制或未满足的任务条件予以说明。

d）宜支持接收用户或相关系统对任务结果的反馈，并根据反馈对相关处理步骤或交付结果进行必要的调整，反馈调整后的结果及任务状态。

8 性能与可靠性要求

8.1 任务执行性能要求

数据智能体宜具备与任务复杂程度、数据规模和交付时效相适应的执行性能，并能够对任务执行过程中的关键性能指标进行监测，内容包括但不限于：

a）宜根据数据任务类型、输入数据规模、处理步骤及交付要求，设定任务响应时间、执行时长和结果交付时长等性能指标。

b）宜支持对任务接收、任务规划、数据处理、模型及工具调用等环节的耗时进行监测，识别影响任务执行效率的关键步骤。

c）宜支持在预期负载下执行多个数据任务，并根据任务优先级、步骤依赖关系和资源需求进行调度，避免任务间资源竞争造成执行超时或任务积压。

d）当数据规模、任务复杂程度或外部服务响应情况导致任务无法在预期时间内完成时，宜反馈任务进度及预计影响，不应将尚未完成的任务反馈为已完成。

8.2 运行稳定性要求

数据智能体应在声明的运行环境和预期负载条件下保持稳定运行，并能够识别影响任务连续执行的运行状态变化，内容包括但不限于：

a）宜支持对数据智能体服务、模型服务、数据源连接及外部工具等关键依赖的可用状态进行检查，识别依赖服务不可用或响应异常的情况。

b）宜支持在长时间运行、批量任务执行及多任务并发等预期场景下维持任务处理能力，避免因任务积压、状态异常或资源耗尽导致服务持续不可用。

c）宜支持对任务排队、执行中、等待外部响应、已完成及执行失败等状态进行管理，避免任务状态丢失或长期停留在无法推进的状态。

d）当运行环境或关键依赖发生变化，可能影响正在执行的任务时，宜识别受影响的任务及步骤，并反馈相应的运行状态。

8.3 异常识别与容错处理要求

数据智能体应识别数据任务执行过程中的异常，并根据异常类型、任务状态和操作影响采取适当的处理措施，内容包括但不限于：

a）宜支持识别数据源不可用、数据格式异常、模型或工具调用失败、通信超时、资源不足及结果返回异常等情况，并定位受影响的任务步骤。

b）宜支持根据异常性质采取重试、调整执行顺序、切换可用资源、降级处理、暂停或终止任务等措施，避免异常在后续步骤中持续传播。

c）对可重试的操作，宜设定重试条件、次数或间隔；对可能产生数据写入、传输、发布等外部影响的操作，应识别重试可能导致的重复执行风险。

d）当异常无法自动处理，或继续执行可能造成结果错误、数据状态不一致时，宜停止受影响的步骤，并反馈异常原因、已完成内容及后续处理需要。

e）宜记录异常发生环节、异常类型、采取的处理措施及处理结果，以支持问题定位和后续改进。

8.4 任务恢复与执行一致性要求

数据智能体宜具备任务中断后的恢复能力，并在重试、恢复或重新调度过程中保持任务状态和数据处理结果的一致性，内容包括但不限于：

a）宜支持保存任务标识、执行计划、已完成步骤、关键中间结果及必要的执行状态，为任务恢复提供依据。

b）任务因服务故障、接口中断或资源不足等原因中断后，宜支持从适当的步骤继续执行；无法继续执行时，应明确记录中断原因和已完成操作。

c）在任务恢复、重试或重新调度前，宜识别已执行的数据写入、修改、传输或交付操作，避免重复执行造成数据重复、结果重复或状态冲突。

d）对存在步骤依赖关系的任务，宜确保后续步骤使用的输入与前序步骤的实际执行结果一致，不应将失败、未完成或未经确认的中间结果作为正常结果继续使用。

e）当任务仅部分完成时，宜保持任务状态与实际执行情况一致，并说明已完成和未完成的步骤。

8.5 资源管理与运行监测要求

数据智能体宜根据数据任务的执行需求管理计算、存储、网络及模型和工具服务等资源，并对影响任务可靠运行的状态进行监测，内容包括但不限于：

a）宜支持识别不同数据任务对计算、存储、网络带宽、模型调用及外部工具的资源需求，并根据任务优先级和执行情况分配资源。

b）宜支持监测任务数量、执行时长、资源使用情况、接口响应状态及关键组件健康状态，识别资源不足、任务积压和性能下降等情况。

c）当资源使用达到预设限制或可能影响任务正常执行时，宜采取任务排队、并发限制、资源调整或告警等措施。

d）宜支持在任务完成、失败或终止后释放不再使用的临时资源，避免资源长期占用影响后续任务执行。

e）宜记录与任务性能和运行可靠性相关的监测信息，为性能评估、故障定位和资源配置调整提供依据。

8.6 性能与可靠性考评参数

提供方应根据声明的任务类型、数据规模、任务复杂程度和运行条件，确定适用的性能与可靠性考评参数。考评参数至少应覆盖时延、任务处理能力、执行成功情况、运行可用性和故障恢复情况；确不适用的参数应说明理由。参数定义及统计口径见表1。本文件不规定统一数值阈值，具体目标值应由提供方与部署方依据应用需求确定。

表1 性能与可靠性考评参数

参数类别	考评参数	定义及统计口径
时延	任务响应时间	从接收有效任务请求至返回任务受理状态的时间；不以受理响应代替结果交付。
时延	任务执行时长	从任务开始执行至形成完成、部分完成或失败等终态的时间；应区分成功和失败任务。
时延	端到端交付时长	从接收有效任务请求至完成约定结果交付的时间；排队、外部服务和人工等待时间应分别记录。
处理能力	任务吞吐量	规定观察周期内完成且满足约定质量要求的任务数与观察时长之比；不同任务类别分别统计。
处理能力	并发任务数	在满足声明的时延和质量目标条件下，同时执行的任务数量；应注明并发方式和负载组成。
执行结果	任务完成率	满足预设完成条件且通过结果校验的任务数占有效测试任务总数的比例；部分完成和失败不计为完成。
执行结果	结果质量达标率	按预先确定的样本或结果单元，满足约定质量规则的数量占被检查总数的比例；应明确检查单元及抽样方式。

参数类别	考评参数	定义及统计口径
稳定运行	服务可用率	观察周期内可提供约定服务的时长占约定服务总时长的比例；计划维护等排除项应事先约定。
故障恢复	任务恢复成功率	预设可恢复故障场景中，恢复后满足任务目标且未产生非预期重复操作的任务数占恢复测试任务总数的比例。
故障恢复	恢复时间	从故障被检出至恢复约定任务处理能力的时间；应注明故障类型、恢复点及恢复后状态。
执行一致性	非预期重复操作发生率	重试、恢复或重新调度中发生非预期重复写入、传输或交付的任务数占相关测试任务总数的比例。
资源效率	单位任务资源消耗	规定条件下完成一项任务所消耗的计算、存储、模型调用量或费用等；应明确计量单位和外部依赖。

9 结果可信与过程可追溯要求

9.1 数据来源与处理依据要求

数据智能体宜支持识别和管理任务所使用的数据来源及处理依据，建立结果与相关数据、规则之间的关联，内容包括但不限于：

a）宜支持记录任务所使用的数据源、数据集或其他数据资源的来源标识、获取时间、数据范围及版本等信息。

b）宜支持识别任务涉及的元数据、业务定义、指标口径、质量规则和处理规则，并将其与相应处理步骤关联。

c）宜支持识别不同数据来源在时间范围、字段含义、统计口径等方面的差异，并对可能影响结果的差异进行核查或说明。

d）宜支持记录数据清洗、转换、关联、聚合等活动所使用的输入数据、关键处理条件及输入输出之间的关联关系。

e）对数据来源或处理依据无法确认的情况，宜提示相关限制，避免将未经核实的数据或规则作为确定依据。

9.2 结果依据与结论核查要求

数据智能体宜支持将处理结果和分析结论与其形成依据关联，并提供必要的核查信息，内容包括但不限于：

a）宜支持将查询结果、统计结果和分析结论关联至相应的数据范围、查询条件、计算方法及关键中间结果。

b）宜支持将数据标注、分类、质量检查和治理结果关联至所使用的规则、规则版本及执行结果。

c）宜支持区分直接来源于数据的事实与基于数据形成的推断、解释或建议，避免将推断内容表述为已核实事实。

d）宜支持根据任务需要，通过重新查询、重新计算、规则复核或人工复核等方式核查重要结果，并记录核查中发现的差异。

e）宜支持识别缺少充分依据的结论并作出说明，避免生成无法对应数据来源、工具执行结果或计算过程的依据描述。

9.3 不确定性与能力边界说明要求

数据智能体宜支持识别影响结果判断和使用的不确定因素，并说明结果的适用条件及自身能力限制，内容包括但不限于：

a）对数据缺失、数据过时、质量不足或来源冲突等情况，宜说明相关问题及其对结果可能产生的影响。

b）对预测、估计、关联分析等具有不确定性的结果，宜明确其结果性质，避免将预测或估计表述为确定事实，或将相关性表述为因果关系。

c）当业务定义、指标口径或处理条件存在多种解释时，宜说明任务执行所采用的解释或假设及其适用限制。

d）当任务超出数据智能体声明的能力范围，或现有数据、工具及方法不足以支持可靠结论时，宜反馈能力或条件限制。

e）宜根据任务风险和结果用途，对需要进一步核查或人工判断的结果作出提示。

9.4 执行过程追溯要求

数据智能体宜支持对影响结果形成的关键执行过程进行追溯，以便核查任务步骤、数据处理活动及结果形成路径，内容包括但不限于：

a）宜支持关联任务请求、执行计划、实际执行步骤和最终结果，识别计划与实际执行情况之间的差异。

b）宜支持记录关键的数据访问、处理规则应用、模型调用和工具调用活动，以及相应的执行状态和输入输出摘要。

c）对任务重规划、处理规则调整、工具切换或人工干预等情况，宜记录调整内容及其影响的任务步骤。

d）对涉及数据修改、结果发布或交付的任务，宜支持区分计划执行、实际执行和执行失败的操作，并关联相应的处理结果。

e）宜支持依据追溯信息查询结果形成路径，定位与结果异常有关的数据、处理步骤或工具调用。

9.5 记录与审计支持要求

数据智能体宜对任务执行和结果形成过程保留必要记录，并提供记录查询和核查能力，内容包括但不限于：

a）宜支持记录任务标识、执行时间、关键操作、执行状态、异常情况、结果交付及人工干预等信息，并将相关记录与对应任务关联。

b）宜支持按照任务标识、数据对象、工具调用或结果版本等条件检索记录，为结果复核和问题定位提供依据。

c）宜采取适当措施保护记录的完整性，防止关键记录被未经授权修改或删除。

d）宜根据过程核查需要确定记录内容，避免在记录中不必要地保存原始敏感数据、个人信息及敏感调用参数。

e）宜结合任务类型、结果影响程度和适用的管理要求，确定记录范围和保存期限，并支持提供审计所需的信息。

10 安全与合规要求

10.1数据安全与个人信息保护要求

数据智能体宜根据任务所涉及的数据类型、敏感程度和使用目的，对数据及个人信息实施相应保护，内容包括但不限于：

a）数据智能体应支持接收和使用依据GB/T 43697—2024确定的数据分类分级结果，并据此配置数据访问、处理和交付限制；未获得必要分类分级信息时，应提示相关人员确认。

b）宜支持在数据接入、处理、存储、传输和交付过程中，采取加密、脱敏、访问限制等与数据敏感程度相适应的保护措施。

c）宜遵循最小必要原则使用任务数据，控制传入模型上下文、工具参数及外部服务的数据范围，减少无关数据和敏感信息的暴露。

d）涉及个人信息处理时，部署方应依据GB/T 35273—2020中适用的个人信息处理要求确定控制规则；数据智能体应支持执行个人信息访问、使用、传输、输出及保存期限等控制规则。

e）宜支持对任务执宜支持对任务执行过程中产生的临时数据、副本和缓存进行管理，并在其使用目的结束后按相关要求进行清理或处置。

10.2身份认证与访问控制要求

数据智能体宜支持对任务发起者、数据智能体本身及其调用的相关服务进行身份识别，并依据任务授权控制对数据和工具的访问，内容包括但不限于：

a）宜支持对用户、数据智能体、外部系统及工具等访问主体进行身份识别，并结合身份认证结果提供相应的访问能力。

b）宜支持依据任务需求、主体身份和授权范围，控制数据智能体可访问的数据资源、可调用的模型与工具以及可执行的操作类型。

c）宜支持区分数据读取、修改、删除、对外传输等操作权限，并对不同数据对象和操作类型实施相应的访问控制。

d）宜支持识别授权失效、权限变更或越权访问等情况，并限制受影响的数据访问及工具调用。

e）宜支持在跨系统调用和多智能体协作过程中传递必要的身份与授权信息，避免因任务转交或工具调用扩大原有权限范围。

10.3模型、工具及执行环境安全要求

数据智能体宜对任务执行所依赖的模型服务、外部工具和执行环境进行安全管理，降低不可信组件或高权限能力带来的风险，内容包括但不限于：

a）宜支持管理可调用模型、工具和连接器的来源、版本、用途及权限范围，并识别未经配置或不符合使用要求的调用对象。

b）宜支持对模型及工具调用的输入参数、执行权限和返回状态进行校验，减少错误参数、异常响应或不可信返回内容对后续任务的影响。

c）对代码执行、数据库写入、文件操作和网络访问等可能改变数据或系统状态的能力，宜采取隔离、权限限制或执行范围约束等措施。

d）宜支持识别模型、工具及连接器的版本变更或配置变更可能带来的安全影响，并在投入使用前开展必要的检查。

e）宜支持保护任务执行中使用的访问凭据、接口密钥和其他敏感配置信息，避免其通过提示内容、工具返回结果或任务输出泄露。

10.4输入输出安全与攻击防护要求

数据智能体宜对任务输入、外部数据及输出结果进行安全检查，区分可信任务指令与不可信内容，防范恶意内容干扰任务执行，内容包括但不限于：

a）宜支持对用户输入、数据源内容、文档、网页及工具返回结果中的恶意指令、异常参数等风险内容进行识别和处理。

b）宜支持区分经授权的任务指令与作为处理对象的数据内容，避免将数据中包含的提示注入内容作为任务指令执行。

c）宜支持识别可能诱导数据智能体泄露敏感信息、绕过访问控制或调用非授权工具的输入内容，并采取限制、拦截或请求复核等措施。

d）宜支持对模型生成内容、数据处理结果和对外输出进行检查，识别其中可能包含的个人信息、敏感数据或超出任务授权范围的内容。

e）对可能影响后续自动执行的工具返回结果或外部内容，宜支持在使用前进行必要的格式、来源和内容校验，降低错误或恶意内容在任务链中传播的风险。

10.5安全监测与事件处置要求

数据智能体宜具备安全风险监测和事件处置能力，及时识别任务执行中的异常访问、异常调用和敏感数据异常流动，内容包括但不限于：

a）宜支持监测数据访问、模型与工具调用、跨系统传输及结果输出中的异常行为，识别越权访问、异常调用频次和敏感数据异常流动等风险。

b）宜支持依据风险类型和影响程度采取告警、限制调用、暂停受影响任务或移交人员处置等措施，防止风险进一步扩大。

c）宜支持记录安全事件发生的时间、关联任务、涉及资源、异常行为及处置结果，为事件调查和影响分析提供依据。

d）对可能涉及数据泄露、凭据暴露或工具滥用的事件，宜支持采取隔离相关资源、撤销或更新凭据、限制相关权限等处置措施。

e）宜支持根据安全事件的处置结果，调整相关安全配置、调用策略或防护规则，并对调整效果进行验证。

10.6合规要求

数据智能体宜结合应用场景、数据类型和任务活动，识别并落实适用的法律法规、标准及组织管理要求，内容包括但不限于：

a）宜支持识别数据采集、使用、加工、传输、提供和公开等活动涉及的合规要求，并将相关限制纳入任务执行条件。

b）宜支持根据个人信息保护、数据安全、网络安全及人工智能应用等方面的适用要求，确定数据处理和结果交付的相关控制措施。

c）涉及跨组织、跨区域或通过外部服务处理数据的任务，宜支持识别数据提供范围、使用目的及传输条件等方面的限制。

d）宜支持对任务执行所依据的授权、处理规则及合规控制措施进行必要记录，为合规检查和问题核查提供依据。

e）当任务要求与适用的合规限制存在冲突时，宜提示相关风险，并支持调整任务方案或移交相关人员处理。

11 自主行为与人工干预要求

11.1 任务授权与自主执行边界要求

数据智能体宜根据任务目标和授权条件，确定自主执行的数据活动、资源范围及操作边界，内容包括但不限于：

a）宜支持识别任务授权的数据对象、处理范围、可调用工具、可执行操作及结果交付对象，并将其作为任务规划和执行的约束条件。

b）宜支持区分可自主执行、经人工确认后执行及不予执行的操作，并根据任务类型和应用场景配置相应的执行方式。

c）当任务规划、工具调用或任务转交可能改变原有执行范围时，宜支持检查其与任务授权的一致性，避免通过拆分步骤或调用外部工具扩大自主执行范围。

d）当用户提出的任务要求超出已有授权，或执行过程中需要新增数据对象、操作类型及交付范围时，宜支持暂停相关步骤并请求补充授权或调整任务要求。

e）宜支持对任务授权条件及自主执行策略进行配置和更新，并识别授权变更对正在执行任务的影响。

11.2 操作风险分级与事前确认要求

数据智能体宜依据操作对象、影响范围和结果可逆性等因素，对数据活动进行风险分级，并对高风险操作设置事前确认机制，内容包括但不限于：

a）宜支持识别数据删除、批量修改、权限变更、对外发布、重要数据传输等可能产生较大影响的操作，并确定相应的确认条件。

b）宜支持在高风险操作执行前，向具备相应职责的人员呈现操作对象、数据范围、拟采取的操作、预期结果及可能产生的影响。

c）宜支持对操作对象、影响范围和处理结果进行预览或模拟，辅助人员判断是否继续执行。

d）宜支持根据确认结果执行、调整或取消相关操作；在未取得所需确认的情况下，宜保持相关步骤处于等待状态。

e）当执行前的实际条件与已确认的操作范围或预期影响发生明显变化时，宜支持重新提示并请求确认。

11.3 执行过程监督要求

数据智能体宜提供任务执行过程中的必要状态和关键操作信息，支持相关人员了解任务进展并及时发现执行偏差，内容包括但不限于：

a）宜支持展示任务目标、执行计划、当前步骤、已完成步骤、待执行步骤及任务状态等信息。

b）宜支持对数据修改、工具调用、结果发布等关键操作提供执行前后状态及处理结果，便于人员检查其是否符合任务目标。

c）宜支持识别任务执行过程中的计划偏离、关键条件变化、结果校验未通过及其他需要人员关注的情况，并进行提示。

d）宜支持相关人员在任务执行过程中查看必要的中间结果、异常信息和待确认事项，为调整任务提供依据。

e）宜支持根据任务风险等级和自主执行程度，配置不同的监督节点、提示方式和人工复核条件。

11.4 人工接管与任务中止要求

数据智能体宜支持在任务受阻、执行偏离或需要人工判断时，由相关人员介入、接管或中止任务，内容包括但不限于：

a）宜支持相关人员对运行中的任务发出暂停、恢复、调整或终止指令，并反馈指令执行结果。

b）当任务目标不明确、关键执行条件不满足、结果存在重大偏差或操作风险超出预设范围时，宜支持暂停后续自主执行并请求人工处理。

c）宜支持在人工接管时提供任务当前状态、已执行操作、相关中间结果、待处理步骤及已识别风险等信息，减少交接过程中的信息缺失。

d）宜支持人员在接管后调整任务目标、处理规则或执行步骤，并在继续执行前识别调整内容对已有结果和后续操作的影响。

e）任务被中止后，宜支持停止尚未开始的相关操作，并反馈已完成、正在执行及未执行的步骤，以及需要进一步处置的事项。

11.5 状态恢复与操作回退要求

数据智能体宜针对人工干预或任务中止后的状态处置，提供必要的恢复、撤销或回退支持，降低自主操作对数据和业务造成的非预期影响，内容包括但不限于：

a）对可能改变数据或系统状态的任务，宜支持在执行前识别可回退条件，并根据操作性质采取预览、备份、版本保存或补偿等措施。

b）在任务暂停、接管或中止后，宜支持识别已生效的数据修改、外部调用及结果交付等操作，为人员判断后续处置方式提供依据。

c）对具备回退条件的操作，宜支持按照确认的范围执行撤销、恢复或补偿，并检查回退后的数据及任务状态。

d）对无法回退或只能部分回退的操作，宜在执行前提示其影响，并在任务中止或处置时说明已产生的结果及剩余风险。

e）宜支持记录人工确认、接管、中止、恢复及回退等过程中的关键操作和处理结果，便于后续核查。

12 互操作与协同性要求

12.1 数据平台与工具互操作要求

数据智能体宜根据数据任务需要，与数据源、数据平台及相关工具进行集成，并利用其提供的数据资源和服务能力，内容包括但不限于：

a）宜支持通过API、数据库连接器、消息接口或文件交换等方式，与数据库、数据仓库、大数据平台、数据服务平台及文件存储系统等进行交互。

b）宜支持识别可用数据平台和工具提供的功能、适用数据类型、输入输出要求及调用条件，并根据任务需要选择相应能力。

c）宜支持利用数据目录、元数据管理、数据质量、数据血缘等平台能力，获取完成任务所需的数据定义、处理规则及资源信息。

d）宜支持识别平台或工具的版本、配置及能力变化，并评估相关变化对现有数据任务的影响。

e）宜支持对不同平台和工具的调用结果进行适配，使其能够作为后续任务步骤的输入。

12.2 接口调用与信息交换要求

数据智能体宜依据交互对象的接口规范，完成任务请求、执行状态和处理结果等信息的传递，内容包括但不限于：

a）宜支持管理接口地址、协议类型、调用方式、参数定义、数据格式和版本等信息，为任务规划和工具调用提供依据。

b）宜支持按照接口定义组装请求参数、发起调用并解析返回结果，识别接口返回的执行状态、错误信息及结果数据。

c）宜支持在不同系统的数据格式、字段名称和参数结构之间进行必要的映射或转换，并检查转换后信息的完整性。

d）宜支持交换任务标识、任务状态、处理结果及必要的上下文信息，以便交互双方关联同一任务及其执行过程。

e）当接口变更、返回内容不符合约定或交互信息不完整时，宜支持识别受影响的任务步骤，并反馈接口适配或信息交换问题。

12.3 数据语义一致性要求

数据智能体宜识别不同数据源、平台及协作主体之间的数据语义差异，保持任务执行和结果交换中的数据含义一致，内容包括但不限于：

a）宜支持识别跨系统使用的业务术语、字段含义、数据类型、计量单位、时间范围及指标口径等信息。

b）宜支持根据元数据、业务规则或映射关系，对不同系统中含义相同但名称或表达方式不同的数据进行语义匹配。

c）对名称相同但含义、统计范围或计算方法不同的数据，宜支持识别差异，避免直接合并、比较或替代使用。

d）宜支持在数据转换和结果交换过程中保留影响数据理解的必要说明，包括指标定义、统计周期、单位及适用范围等。

e）当语义冲突无法依据现有信息消除，且可能影响任务结果时，宜提示冲突内容并请求相关人员或系统确认。

12.4 多智能体任务协同要求

数据智能体宜支持与其他智能体协同完成涉及多个数据活动或专业能力的任务，明确任务分工、执行依赖及结果交接条件，内容包括但不限于：

a）宜支持根据任务目标和各智能体的能力范围，划分子任务，并明确子任务的输入、预期输出和完成条件。

b）宜支持向其他智能体传递完成子任务所需的任务要求、数据范围、处理规则及必要的上下文信息。

c）宜支持获取协同任务的接收、执行、完成或失败等状态，并根据任务依赖关系协调后续步骤。

d）宜支持对其他智能体返回的结果进行格式、内容和完成状态检查，识别结果缺失、任务重复或结果冲突等情况。

e）当协同任务发生中断或参与方的执行结果不一致时，宜支持反馈受影响的任务范围，并协调重新执行、调整分工或请求人工处理。

12.5 人机协作与任务交接要求

数据智能体宜支持与任务发起者、数据管理人员、业务人员及其他相关人员协作，在任务提出、执行、复核和交接过程中保持信息完整，内容包括但不限于：

a）宜支持人员提供任务目标、数据范围、处理规则、结果要求及修正意见，并将相关信息用于后续任务执行。

b）当任务涉及业务口径确认、数据质量判断或结果解释等需要人员参与的事项时，宜支持提出明确的问题，并提供作出判断所需的相关信息。

c）宜支持向人员交接任务进展、已完成步骤、中间结果、待处理问题及已识别的结果限制，避免因交接信息不足影响后续处理。

d）宜支持接收人员的核查意见和处理反馈，并关联至相应任务步骤或结果，以便继续执行或修正结果。

e）宜支持明确智能体与人员在任务执行、结果复核和成果使用等环节的协作关系，避免任务交接后出现处理遗漏或重复。

参 考 文 献

[1] GB/T45907—2025 人工智能 服务能力成熟度评估

[2] GB/T36073—2025 数据管理能力成熟度评估模型

[3] GB/Z 242-2026 人工智能 智能体技术要求

[4] GB/T43580—2023 区块链和分布式记账技术 存证通用服务指南