TC609

全国数据标准化技术委员会技术文件

TC609-4-2026-XX

城市全域数字化转型　城市智能体建设及应用指南

Citywide digital transformation-Guidelines for the construction and application of city intelligent agent

（征求意见稿）

2026-XX-XX 发布　　2026-XX-XX 实施

全国数据标准化技术委员会　发布

目　　次

前　　言 II
1 范围 1
2 规范性引用文件 1
3 缩略语 1
4 术语和定义 1
5 建设原则 2
6 总体架构 2
6.1 概述 2
6.2 智能感知和连接层 3
6.3 数据层 4
6.4 能力层 5
6.5 应用层 9
7 城市智能体实施路径 11
7.1 场景识别 11
7.2 打造高质量数据集 15
7.3 构建 AI 能力 15
7.4 搭建智能体 17
7.5 智能体运营 19
参　考　文　献 21

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。

本文件由全国数据标准化技术委员会（SAC/TC609）提出并归口。

本文件起草单位：重庆信息通信研究院、北京思特奇信息技术股份有限公司、中国电子技术标准化研究院、数族科技（南京）股份有限公司、安阳数智科技有限公司、浙江省质量科学研究院、联想（北京）有限公司、数融智联（徐州）信息科技有限公司、广州智能科技发展有限公司、联通数据智能有限公司、国家数据发展研究院、北京新瑞锦世科技有限公司、大连深方信息科技有限公司、中移雄安信息通信科技有限公司、杭州高新区(滨江)区块链与数据安全研究院、华为技术有限公司、建筑材料工业信息中心、南京理工大学、深圳开鸿数字产业发展有限公司、广州维视达数字科技有限公司、山东未来集团有限公司、中国人民大学、南威软件股份有限公司、浪潮智慧城市科技有限公司、武汉云计算科技有限公司、扬州市大数据管理中心、国家发展和改革委员会创新驱动发展中心(数字经济研究发展中心)、深圳市有方科技股份有限公司、北京建筑大学、山西远大纵横科技有限公司、中国质量认证中心有限公司、嘉兴市文诺财经大数据技术研究院、中电信数政科技有限公司、联通（江苏）产业互联网有限公司。

城市全域数字化转型 城市智能体建设及应用指南

1 范围

本文件给出了城市智能体的建设原则、总体架构及实施路径。

本文件适用于城市智能体的设计、开发活动，可为城市智能体建设及应用提供指导。

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 信息安全技术 个人信息安全规范

GB/T 45288.1-2025 人工智能 大模型 第1部分：通用要求

TC609-5-2025-01 高质量数据集 建设指南

TC609-5-2025-02 高质量数据集 格式要求

TC609-5-2025-03 高质量数据集 分类指南

TC609-5-2025-04 高质量数据集 质量评测规范

TC609-XX-202X-XX 城市全域数字化转型 城市智能中枢 城市人工智能公共服务平台总体要求

TC609-4-202X-XX 城市全域数字化转型 城市感知体系建设导则

3 术语和定义

下列术语和定义适用于本文件。

3.1

高质量数据集 high-quality dataset

经过采集、加工等数据处理，可直接用于开发和训练人工智能模型，能有效提升模型性能及泛化能力的数据集合。

[来源：TC609-5-2025-01，3.2，有修改]

3.2

城市智能体 city intelligent agent

能够感知城市运行状态并利用工具采取行动以实现特定目标的代理。

注1：智能体以大模型为智能底座，具备自主感知、理解、规划、决策、记忆、行动和使用工具的能力。

注2：智能体能够自动化执行复杂任务，具有自主性、交互性、反应性和适应性等基本特征。

3.3

大模型 large-scale model

大规模深度学习模型 large-scale deep learning model

基于大量数据训练得到，具有复杂计算架构，能处理复杂任务，且具备一定泛化性的深度学习模型。

注：大模型的参数量由其功能和模态决定，一般不低于1亿。大模型训练使用的数据总量受参数量的影响，达到收敛的大模型的参数量的对数与其训练数据总量的对数成正比。

[来源：GB/T 45288.1—2025，3.1]

4 缩略语

下列缩略语适用于本文件。

5G：第五代移动通信技术（5th Generation Mobile Communication Technology）

AI:人工智能(Artificial Intelligence)

API：应用编程接口（Application Programming Interface）

GIS:地理信息系统(Geographic Information System)

IPv6：互联网协议第6版（Internet Protocol version6）

QoS：服务质量（Quality of Service）

RAG：检索增强生成（Retrieval Augmented Generation）

SLA：服务水平协议（Service Level Agreement）

5 建设原则

5.1 以人为本，需求导向

以市民需求为核心，避免技术堆砌，通过智能技术提升就业、医疗、教育等公共服务精准度。

5.2 数据驱动，智能协同

以数据整合为基础，建立统一数据中心，打破数据共享壁垒，发挥智能体智能决策优势，通过大数据、人工智能等技术实现城市“可感知、会计算、能思考” 。

5.3 安全可靠，可持续发展

构建“安全型”城市智能体，确保数据安全与系统稳定性，同时符合低碳环保要求，形成节约型建设模式。

5.4 开放合作，生态共建

政府引导与市场运作结合，鼓励社会资本参与，同时推动产学研合作，探索构建“协同型”智慧平台，强化跨领域资源整合。

6 总体架构

6.1 概述

城市智能体总体架构由感知和连接、数据层、能力层、应用层等要素组成，见图1。

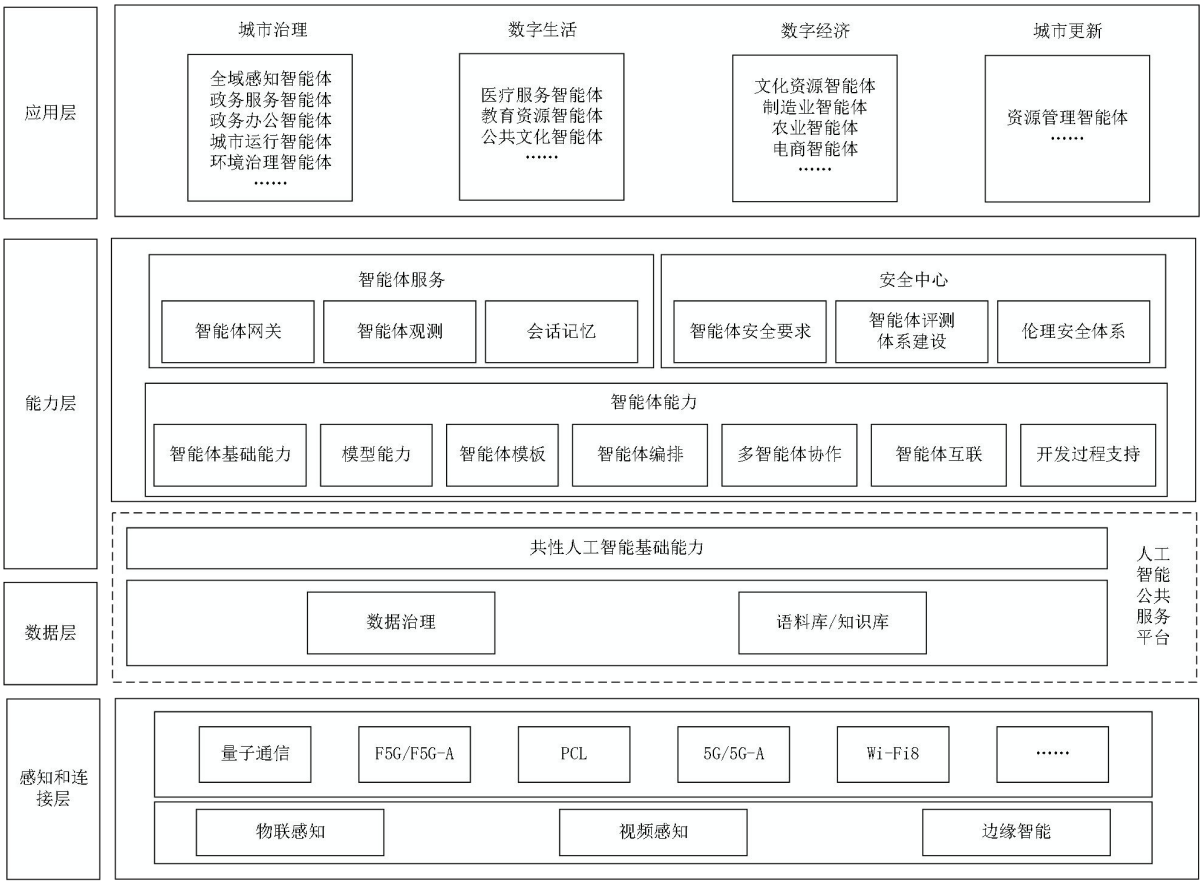

图 1　城市智能体总体架构

6.2　智能感知和连接层

6.2.1　智能感知

智能感知考虑如下因素：

a)　多维采集能力：具备视频、音频、物联传感、射频识别及空间定位等多模态数据的采集能力，支持对城市部件、人员、车辆、事件及环境参数的精准识别与监测；

b)　边端处理能力：感知终端宜具备边缘计算能力，支持在源端进行数据清洗、特征提取、隐私脱敏及实时分析，实现“感算一体”；

c)　控制交互能力：感知终端具备可配置的控制接口，支持响应上层指令，实现对末端设备的远程控制、自动调节及状态校准；

d)　感知终端符合 TC609-4-202X-XX 中 6.1 章节的要求。

6.2.2　智能连接

智能连接考虑如下因素：

a)　泛在网络接入：具备异构网络融合接入能力，符合 TC609-4-202X-XX 中 6.2 章节的要求，支持包括但不限于光纤、以太网、公/专网的 4G/5G/6G/NB-IoT、高低轨卫星通信、无线自组网、新一代 Wi-Fi/星闪等多种传输协议，以及以及 IPv6、TCP/IP、UDP、HTTP/HTTPS、MQTT/MQTTS、CoAP、RTSP 等应用层协议；

b) 标准化数据接入：支持主流的数据格式，具备统一的设备标识与注册机制，确保不同厂商、不同类型感知设备的互联互通；
c) 数据流通安全可信：支持基于数据分级分类要求或信息系统安全等级要求，提供对应等级的数据可信流通服务，确保数据在网络中流通的安全可信；
d) 高并发与高性能：满足城市级高密度连接需求，并为关键业务提供微秒级时延与确定性带宽保障（SLA）；
e) 内生安全：建立设备入网的实时鉴权与信任机制，具备防范非法终端接入、数据加密传输及异常流量清洗的能力，确保链路安全可信。

6.3 数据层

6.3.1 概述

城市智能体的数据治理及语料库和知识库可基于大数据平台与人工智能公共服务平台提供的服务，若未建公共服务平台，按照6.3.2、6.3.3的要求。

6.3.2 数据治理

对数据进行治理，以确保数据的完整性、准确性、一致性、多样性、可靠性、安全性与合规性，数据能够被准确地用于模型的训练和部署，同时保护数据的隐私和安全。城市智能体数据治理工作要求如下：

a) 标准与模型体系：在全域范围内统一技术、业务与安全元数据模型，确立人、事、物、地点、时间与组织等主数据标准及其参照规范。以跨部门统一本体与术语为基础，构建可被机器理解的语义框架，为知识抽取、规则表达与推理提供稳定支撑；
b) 数据资产盘点与目录治理：围绕全域数据资源，建立统一的资产台账与主题目录，覆盖政务、城市治理与公共服务等板块。通过敏感度、共享级别与开放级别的分类分级，明确数据的权属边界与管理责任，确保后续共享与使用有据可依、可查可审；
c) 采集与接入：面向政务业务系统、物联网感知终端等开展有序接入，通过明确字段定义、质量阈值、更新频率与服务级别，与事件流、批处理结合，满足不同场景的实时性与批量处理需求；
d) 清洗、标注与结构化：对文本开展去噪、去重与版式解析，对图像与视频进行质量筛选与标签规范化；采用弱监督、半监督与主动学习相结合的标注机制，可引入人工复核保障准确性。围绕实体、关系、事件与时空要素开展结构化抽取，并以图数据库或 RDF 等形态沉淀为可复用的知识资产。重点针对现有异构数据，采用 ETL 工具和数据融合技术解决格式与语义差异，将现有非结构化历史数据转化为结构化数据；同步处理新增异构数据，实现新旧数据格式统一、语义互通；
e) 数据脱敏：对涉及个人隐私和敏感信息的数据，采用加密、掩码、匿名化等技术进行脱敏处理，脱敏后数据无法关联至具体个人；对政务敏感数据采用分级脱敏策略，根据访问权限控制数据可见范围，脱敏处理符合 GB/T 35273 要求；
f) 质量评估与持续改进：构建覆盖准确率、完整率、时效性、一致性与可用性的质量指标体系，结合语料与知识评测集形成自动化评估闭环，定期识别覆盖不足、偏差与漂移问题，推动数据与知识的持续修复与优化；
g) 生命周期与版本治理：建立创建、变更、冻结、归档与销毁的全生命周期管理机制，确保数据资源在不同阶段有明确的处置要求。通过版本追溯与差异对比，支撑模型再训练与知识回滚，降低演进过程中的风险。

6.3.3 语料库/知识库

6.3.3.1 概述

语料库/知识库是城市智能体的知识支撑模块，为场景识别、决策支持等核心功能提供数据与知识服务。语料库作为原始数据储备中枢，汇聚政务服务、城市治理等多领域多模态数据，为场景识别提供基础训练数据与实时输入；知识库则作为结构化知识加工中心，通过知识抽取、关联与推理，形成可直接支撑决策的规则库、案例库与模型参数，赋能智能问答、自动审批、风险预警等核心应用。两者协同构成“数据-知识-应用”的完整闭环,为城市智能体提供基础支撑。

6.3.3.2 组成

语料库/知识库在城市智能体中，主要分为以下四个类别：

a) 政务领域基础语料库：城市智能体中汇聚政务服务相关原始数据的基础资源池，该语料库为智能问答、自动审批等应用提供基础训练数据，助力提升政务服务效率，为决策提供支撑服务；
b) 城市治理基础语料库：面向城市运行管理的动态数据汇聚中心，聚焦城市交通、环境、安防等领域的实时与历史数据，支撑交通拥堵预警、环境质量监测、公共安全事件识别等核心场景；
c) 部门知识库：政府各职能部门内部的专业化知识管理系统，聚焦部门业务规则、审批流程、历史决策等结构化知识的沉淀与应用，提升跨部门协作效率。功能包括知识结构化录入、版本控制、跨部门知识协同，以及基于业务场景的知识智能推送；
d) 公共知识库：面向社会公众开放的标准化知识服务平台，汇聚法律法规、公共服务信息、城市基础数据等普惠性知识资源。通过城市智能体问答功能解答市民办事疑问，同时促进公共知识的普及与传播，提升城市服务的可及性与透明度。

6.3.3.3 技术要求

语料库、知识库满足以下技术要求，包括但不限于：

a) 互联互通要求：系统支持主流数据传输协议与接口标准，提供标准化 API 接口供外部系统调用，确保不同知识库与语料库之间以及与城市智能体其他模块之间的兼容对接。满足不同场景下的数据时效性需求，同步过程中具备冲突检测与自动解决能力，保障数据一致性；
b) 数据质量要求：符合 TC609-5-2025-01、TC609-5-2025-02 中的基本规定，从完整性、准确性、一致性、时效性四个维度建立通用保障机制。系统内置数据质量监控工具，对异常数据进行自动告警与人工干预提示；
c) 安全合规要求：知识与语料管理的基本底线，构建多层次安全防护体系。访问控制支持权限配置，确保知识访问的可追溯性；数据传输与存储需采用国密算法加密，敏感信息需进行脱敏处理，脱敏后数据不得关联个人身份或敏感业务信息；系统满足数据安全相关法律法规要求，具备完善的操作审计日志，支持安全事件的追溯与分析。

6.4 能力层

6.4.1 共性人工智能基础能力

城市智能体的共性人工智能基础能力可基于城市人工智能公共服务平台提供的标准化、普惠化人工智能服务，对模型、算力、数据集等统一规划、建设、管理，详见TC609-4-202X-XX。

6.4.2 智能体能力

智能体能力提供智能体各种特性的开发支持，包括：

a) 智能体基础能力。主要包含大模型感知、认知、执行和学习的各类能力：
   1) 支持智能体目标设定、安全约束设定等；
   2) 支持提示词模板、知识检索 RAG，工具注册和工具访问、各类认知推理框架；
   3) 感知层面支持各类多模态输入的相互转化能力，比如语音转文字等。

b) 模型能力。针对城市智能体，构建分层分级、协同融合的模型服务架构，支撑城市治理各领域的智能化应用需求：
   1) 城市智能体模型体系。城市智能体模型体系宜包括以下能力：
      ——通用大模型。语言大模型：支持中文、英文等多语言处理；多模态大模型：支持文本、图像、音频、视频等多模态数据的统一理解和生成，具备跨模态推理能力；代码生成模型：专门用于自动化脚本生成、配置文件生成等技术任务的专用模型；
      ——垂直模型。垂直模型包括：城市治理大模型；公共服务大模型；产业经济大模型；生态环保大模型；
      ——场景模型。针对特定的场景，可以提供相应的场景模型。场景模型适应范围更小，基于特定的场景数据和目标进行微调。比如：交通调度、安防监控、环境监测等领域的预训练模型。
   2) 城市智能体模型能力要求。
      ——基础能力要求。多语言支持：支持中文、英文等主要语言的理解和生成；多模态处理：具备文本、图像、音频、视频的综合处理能力；长文本理解：具备处理长篇文档和复杂文本的能力，支持上下文关联理解和跨段落信息整合；实时推理：提供快速响应的推理服务，满足实时业务场景的性能要求，支持高并发访问；
      ——质量指标要求。准确性要求：在核心业务任务中保持高准确率，确保输出结果的可靠性和正确性，满足实际应用场景的质量标准；召回能力：具备良好的信息检索和内容发现能力，能够全面获取相关信息，避免重要信息遗漏；一致性要求：对同类问题和相似场景保持稳定一致的处理结果，确保系统输出的可预测性和可靠性；安全性保障：建立完善的内容安全检测机制，有效识别和过滤有害、违规内容，确保输出内容的安全合规；时效性保证：建立及时的知识更新机制，确保信息的时效性和准确性，快速响应数据变化和业务需求；鲁棒性要求：在面对异常输入、边界情况或系统压力时，保持稳定的性能表现和服务质量；
      ——模型推理。模型推理支持在线实时推理与离线批量推理，具备多种主流模型格式的统一部署与服务化能力。通过资源动态调度与请求队列管理，保障高并发、大数据量场景下的低延迟与高可用。

c) 智能体模板：
   1) 对于各类城市的行业场景，提供各类成熟的行业智能体和领域知识库模板。基于模板可以快速开发相应的行业场景智能体；
   2) 支持各种类型智能体的模板，包含工作流类，自主调度类，智能问答类等；
   d) 智能体编排：支持可视化拖拽式或者基于代码的工作流设计，允许开发者通过流程方式定义智能体的思考路径、任务拆解逻辑和决策分支，实现复杂城市业务逻辑的自动化处理。

e) 多智能体协作：
   1) 支持各类多智能体协作模式，包含主席智能体和执行智能体的经典模式；
   2) 支持基于 A2A 等标准协议框架。

f) 智能体互联：

1) 支持智能体之间使用标准通信协议（如 MCP、A2A、ANP 及扩展其他协议）进行通信和交互；
2) 基于智能体资产管理，支持城市智能体的自动发现。自动发现的智能体必须是经过智能体评测中心评测通过的安全智能体。

g) 开发过程支持：
1) 支持各类相关资产的注册和使用，包含提示词、工具等；
2) 支持智能体的调试、评估、发布等。

6.4.3 智能体服务

智能体服务要点如下。

a) 智能体网关。接入网关在保证处理所有流入流量的前提下，确保安全、可靠和高可用：
1) 统一的 API 入口，负责请求路由、组合和协议转换；
2) 统一处理身份认证、权限鉴定、访问限流和请求日志；
3) 提供智能体运行地址注册，用于智能体路由。

b) 智能体观测。可观测性是监控和理解智能体生态系统端到端行为的过程，包括 AI 智能体可能与大语言模型和外部工具进行的任何交互。主要包含以下特性：
1) 实时汇聚智能体的各种运行数据，包含智能体访问的知识、各类工具、大模型，提示词、推理决策信息，还包含智能体相互的调用日志、跨域互联互通日志等；
2) 制定智能体服务的核心指标，包含性能，可用性，准确性等。对于日志进行智能分析，及时发现大模型的处理性能，效率和可靠性等问题。同时提供日志给安全中心，进行智能体实时安全监控和防护；
3) 通过可视化方式对于智能体运行展现。

c) 会话记忆。对于会话的记忆是提升个性化智能的重要手段，主要特性包含：
1) 支持会话的短期、中期和短期记忆。不同的记忆可以采用不同的存储系统；
2) 支持会话的摘要存储，提升会话的存储效率和使用效率；
3) 支持会话敏感信息的脱敏、加密等措施，提升存储安全。

d) 智能体资产管理。用于城市范围内不同级别的智能体资产管控，包含城市级、行业级、企业级等。
1) 资产管理支持智能体的注册和发布。注册的智能体应是经过评测后的安全智能体；
2) 资产管理的智能体可以被查询和订阅。

6.4.4 安全中心

6.4.4.1 智能体评测体系建设

智能体评测体系建设要点如下。

a) 评测规则：依据《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等法律法规，制定覆盖城市智能体全生命周期的安全评测规则；

b) 评测范围。
1) 技术安全评测：包括系统漏洞、数据加密、访问控制、应急响应等；
2) 数据安全评测：包括数据采集、传输、存储、使用、销毁全流程安全；
3) 功能质量评测：包括智能决策准确性、容错能力、抗干扰能力等；
4) 合规性评测：包括法律法规符合性、行业标准符合性等；

5) 供应商管理对第三方组件、服务和外包环节进行全生命周期管理。对供应商进行安全资质评估、定期复评和风险告警，建立合同中的安全条款与违约责任。

c) 评测机制。

1) 建立定期评测制度，每季度进行常规评测，重大版本更新后进行专项评测；

2) 引入第三方评测机构，确保评测结果的客观性和公正性；

3) 建立评测结果反馈机制，形成问题整改闭环管理；

4) 根据要求完成 AI 备案并向网信办开放测试接口，确保所有向公众提供服务的城市智能体符合网信办相关要求。

d) 评测平台：搭建智能化评测平台，实现自动化扫描、渗透测试、风险评估等功能。建立内容安全检测能力，能够根据网信办发布的测试数据集进行内容安全合规的自查自纠，确保持续安全合规性。

6.4.4.2 伦理安全体系建设

伦理安全体系建设要点如下。

a) 伦理安全框架制定。

1) 依据《新一代人工智能伦理规范》等，制定城市智能体伦理安全框架；

2) 确立公平性、透明性、可解释性、问责性、隐私保护五大伦理原则。

b) 伦理安全管理机制。

1) 建立伦理安全审查委员会，负责智能体伦理安全评估和决策；

2) 建立伦理安全风险评估制度，在智能体设计、开发、部署各阶段进行伦理评估；

3) 建立伦理安全投诉处理机制，接受公众对智能体伦理问题的投诉和建议。

c) 重点领域伦理安全保障。

1) 公共服务领域：确保智能体提供的公共服务公平公正，不歧视任何群体；

2) 社会治理领域：保障智能体决策过程透明可解释，避免暗箱操作；

3) 个人信息处理：严格保护个人隐私，避免过度收集和滥用个人信息；

4) 算法决策领域：定期评估算法公平性，防止算法歧视和偏见；

5) 价值观对齐：算法设计、模型训练和优化、提供服务过程中，采取训练数据筛选、价值观对齐、输出校验等方式，有效规避产生民族、信仰、国别、地域、性别等歧视的风险。

d) 伦理安全教育宣传。

1) 对智能体开发、运营人员进行伦理安全培训；

2) 向公众宣传智能体伦理安全理念，提高公众伦理安全意识；

3) 发布伦理安全白皮书，公开智能体伦理安全管理情况。

6.4.4.3 智能体安全要求

智能体安全要求如下。

a) 智能体通用安全要求：

1) 网络安全要求至少满足 GB/T 22239-2019 中第二级安全技术要求；

2) 训练数据安全符合 GB/T 45654-2025 中第 4 章的要求；

3) 模型安全符合 GB/T 45654-2025 中第 5 章的要求。

b) 智能体感知安全要求：

1) 感知数据涉及个人信息保护符合 GB/T 35273 要求，进行脱敏或去标识；

2) 在各种预期和非预期条件下（如传感器噪声、人为干扰），感知系统都能持续提供完整和准确的信息。

c) 智能协同安全要求：
   1) 智能体之间通信应采用加密协议，如 HTTPS；
   2) 智能体协同需要验证对方智能体身份可信；
   3) 智能体工具调用应保证调用接口的安全。
d) 智能体行动安全要求：
   1) 智能体向用户明确告知任务执行所申请的操作系统权限、访问用户数据的目的、处理方式等内容，并获得用户许可；
   2) 智能体执行涉及有重要影响重要任务，引入人为因素参加决策。
e) 智能体其他要求：建立 AI 内容安全围栏，确保输入、输出内容符合社会伦理、国家及地方法律法规的要求。

6.5 应用层

6.5.1 城市治理

包括但不限于：

a) 全域感知智能体：对摄像头视频流、烟感报警器状态数据、智能井盖液位与位移传感器信号等异构物联网设备进行实时解析与特征提取；通过计算机视觉算法（如目标检测、行为分析）、时序异常检测模型及多模态数据关联引擎，自动识别高空抛物轨迹、消防通道车辆违停占压、燃气浓度突变泄漏等风险事件；基于预设风险等级规则自动生成结构化处置工单，实时推送至网格员移动终端，同时触发社区门禁系统联动锁闭、广播系统分区播报预警信息，形成“感知-识别-预警-处置”的闭环响应链；依托平台汇聚社区建筑信息模型（BIM）、人口热力、车辆轨迹等静态与动态数据，构建高精度社区数字镜像，通过强化学习算法模拟停车位潮汐使用、垃圾清运路径规划、应急通道占用等场景，动态优化资源调度策略，最终实现从“被动响应”到“主动预防”、从“经验驱动”到“数据驱动”的治理范式跃迁；
b) 政务服务智能体：建立统一资源目录，实现企业基本信息、信用状况、社保缴纳等数据的实时归集、脱敏与可信共享；对企业提交的营业执照、法定代表人身份证明、经营场所证明等申请材料进行自动解析、要素提取与交叉验证，通过规则引擎内置的审批规则与风险预警模型，完成材料真实性核验、经营范围合规性审查、地址库比对等检查，真正达成“数据多跑路、群众少跑腿”的治理目标；
c) 政务办公智能体：集成多通道语音识别引擎与声纹身份认证技术，实时采集会议音频流并转写为结构化文本，自动提取议题要点、发言人观点及决策结论，生成带任务清单的会议纪要；依托政策文件、法规条例与历史公文的知识图谱，当用户输入公文主题关键词时，智能体通过语义检索与篇章生成模型自动匹配政策依据、引用相关法条、套用法定格式，输出符合公文格式标准的初稿；自动拆解决议事项为可执行任务，智能匹配责任部门与责任人，通过政务微信、短信等多渠道推送督办提醒，实时抓取任务系统回填数据生成执行进度看板，推动政务办公从“人找事”向“事找人”、从“经验驱动”向“算法驱动”的智能化跃迁；
d) 城市运行智能体：实时汇聚政务服务业务数据、交通领域视频监控与微波雷达数据、环境部门空气质量监测站点数据以及卫健系统疫情流调数据等，构建覆盖城市运行全要素的“数据湖泊”；基于城市信息模型（CIM）搭建高精度数字孪生平台，将多元数据映射为三维可视化模型，通过深度学习算法对交通拥堵模式、空气污染扩散路径、疫情传播链进行推演与预测，生成包含警力部署方案、信号灯配时优化策略、公交应急接驳线路、污染源管控清单、封控区划定建议等在内的智能处置预案；通过跨部门联动指挥系统将预案拆解为可执行任务，基

于政务微信、指挥专网同步推送至交管、环保、卫健、城管等责任部门，同时模拟不同处置策略效果，动态优化资源调度，实现从“人脑决策”到“算法驱动”的范式转变；

e) 环境治理智能体：集成高分辨率卫星遥感影像、无人机倾斜摄影数据及地面微站传感器网络，对 PM2.5、COD、土壤重金属等环境指标进行采样与时空配准，构建覆盖全域的“生态环境数字孪生体”，实现大气污染传输路径、水体污染扩散范围、土壤污染风险区划精确定位；部署基于深度学习的多光谱图像解析算法与视频行为识别模型，通过垃圾分类投放点的摄像头实时抓拍，一旦检测到混投行为，立即通过边缘计算节点触发定制化语音提醒；运用强化学习算法对垃圾分类时空分布热力图进行挖掘分析，动态规划最优收运路径，同时结合气象预测与污染溯源模型，提前预警重污染天气过程，推送至生态环境、城管、交通等部门协同执行。

6.5.2 数字生活

包括但不限于：

a) 医疗服务智能体：实时对接 120 急救系统的车辆定位与出车状态、医院 HIS 系统的急诊床位占用与专科救治能力等，同时融合居民电子健康档案、可穿戴设备监测数据，构建覆盖“急救-诊疗-用药”全链条的动态医疗数据湖；针对急救场景，综合考虑救护车实时位置、路况拥堵指数、医院接诊饱和度、患者生命体征等参数，动态生成“最优急救路径+最近可用医院+预开放绿色通道”的立体调度方案；针对慢病管理，构建跨机构的疾病风险预测模型，在数据不出域的隐私保护前提下，对居民健康档案进行联合训练，并自动生成个性化干预建议（如用药调整、生活方式指导），通过加密推送至对应家庭医生移动端；系统持续追踪急救患者院内救治结果与慢病患者依从性数据，反哺模型迭代优化，推动医疗服务从“被动应对”向“主动健康”、从“经验决策”向“算法驱动”的智能化转型；

b) 教育资源智能体：通过实时汇聚学校教师的专业资质、授课时段与教学评价数据，课程的科目、难度、学分与教室设备兼容性信息，以及学生的选课偏好、学业画像与升学目标数据，构建覆盖“师-生-课-场”全要素的教育资源数字孪生体；针对课程编排难题，将教师时间窗、教室容量、设备需求、学生志愿匹配度等约束条件转化为目标函数，生成满足大部分学生首选志愿、教师课时负荷均衡度合理的最优课表；在课后服务场景，基于学生兴趣标签（艺术、科技、体育等类）与社团资源池（师资、场地、器材）构建双边匹配模型，为每位学生生成个性化课后服务清单；系统追踪课程实施效果、学生成绩提升与课后服务满意度数据，动态调整资源配置策略，推动教育资源从“校际割据”向“区域融通”、从“经验配置”向“算法驱动”的现代化治理转型；

c) 公共文化智能体：运用计算机视觉算法对博物馆客流进行实时热力图渲染与聚类分析，识别拥堵节点与冷门展区，通过强化学习模型动态优化展品布局与参观路径，在图书馆场景中，自动识别热门书籍并触发智能补货预警，同时融合室内定位与路径规划技术为读者提供 AR 导航服务；针对大型文化活动，智能体依托数字孪生平台提前导入场地 BIM 模型与人群行为参数，仿真模拟人流的疏散过程，自动生成包含最优撤离路线、警力部署节点、医疗救援点位的分级应急预案，当实时监测到人流密度超过安全阈值时，自动触发广播引导与闸机限流联动机制，推动公共文化服务从“经验式管理”向“算法式治理”、从“被动响应”向“主动预防”的智慧化转型。

6.5.3 数字经济

包括但不限于：

a) 文化资源智能体：运用激光三维扫描、倾斜摄影测量与光谱成像技术对文物表面微痕、古建筑木结构节点进行全要素数字化采集，生成云数据并自动拓扑重构为可交互式数字模型；将古籍的繁体竖排文字自动解析为分镜脚本，智能匹配历史服饰、建筑场景素材库，批量生成动态漫画与数字人讲解视频；支持用户上传手机拍摄的文物局部照片、语音描述或输入文本，系统在数字文化资源中定位关联藏品，推动文化遗产从“静态保护”转向“活态传承”、从“小众研究”走向“大众沉浸”的智能化演进。；

b) 制造业智能体：实时接入生产线 PLC、SCADA 等系统以及部署在机床、机械臂、传送带上的振动、温度、电流传感器，同步汇聚 ERP 订单数据、MES 工艺参数与 WMS 物料状态，形成生产全要素数字孪生底座；针对设备运维，构建设备健康度画像，预测轴承磨损、电机过热等潜在故障，自动触发维保工单至工程师终端并同步调整生产排程；针对排程优化，将订单交期、设备产能、工装夹具可用性、物料齐套率等约束条件转化为多目标优化函数，动态生成可执行的生产计划；在质量管控环节，部署工业相机与边缘 AI 计算节点，对产品表面瑕疵、尺寸偏差进行在线检测，推动制造模式从“事后抢修”向“事前预防”、从“刚性排产”向“柔性调度”、从“抽样质检”向“全检追溯”的智能化跃迁；

c) 农业智能体：依托部署于田间的物联网传感器网络实时监测土壤墒情（温湿度、EC 值、氮磷钾含量）、气象微站逐分钟更新的风速、光照、降雨量数据，以及无人机多光谱成像与卫星遥感获取的作物 NDVI 植被指数、病虫害侵染特征，构建覆盖万亩级农田的数字化种植底座；融合深度学习算法（如气象预测、病虫害识别）与作物生长机理模型，将多源异构数据输入精准种植决策引擎，动态生成包含灌溉量精确到立方米、施肥配方细化到克级配比、施药窗口精确到小时级的个性化农事方案，并通过知识图谱关联历年产量数据与市场价格波动规律；智能体实时抓取电商平台、批发市场、期货交易所的农产品价格数据，结合物流成本、仓储容量与农户产能，运用强化学习算法预测最优定价区间与上市窗口，自动匹配盒马鲜生、美团优选等 B 端渠道或社区团购 C 端资源，实现“订单农业”模式，推动农业生产从“经验依赖”转向“数据驱动”、从“盲目种植”转向“按需生产”的智能化转型；

d) 电商智能体：实时采集趋势、社媒舆情、海关进出口数据及竞品动态，构建需求预测模型，挖掘品类潜力，结合企业产能、利润率与物流成本，自动生成“目标市场-潜力产品-定价策略”三维推荐方案，推动交易从“经验选品”向“算法驱单”、从“人工跟单”向“无人值守”的智能化升级。

6.5.4 城市更新

包括但不限于资源管理智能体：融合高分辨率卫星遥感影像、无人机倾斜摄影、地面物联网传感器等多源数据，基于GIS引擎与BIM技术构建自然资源三维数字孪生模型，实时监测土地利用变化、森林蓄积量动态、河流水质与水量等指标，通过深度学习算法自动识别违规占用耕地、非法采伐、水体污染等异常事件；整合区域内电力、燃气管道压力、热力管网温度等异构能源数据，构建多能互补的区域能源互联网调度平台，运用混合整数规划与强化学习算法对源-网-荷-储进行一体化优化，实现风光储协同出力、冷热电联供动态调节与需求侧响应柔性控制，推动资源管理从“被动巡查”转向“主动预警”、能源调度从“经验平衡”转向“算法最优”的智慧化跃升。

7 城市智能体实施路径

7.1 场景识别

7.1.1 场景分析

7.1.1.1 场景识别原则

场景识别原则如下。

a) 全面性原则：

1）覆盖城市治理全生命周期，包括规划、建设、管理、运营等各个阶段；

2）涵盖政府、企业、市民等全部利益相关方的需求；

3）统筹考虑城市治理、数字生活、数字经济、城市更新多维度发展目标。

b) 系统性原则：

1）基于城市智能体总体架构，确保场景设计与技术架构的一致性；

2）注重场景间的关联性和协同性，避免孤立建设；

3）遵循从简单到复杂、从局部到整体的渐进式发展路径。

c) 实用性原则：

1）以解决实际问题为导向，聚焦城市治理的痛点难点；

2）充分考虑技术可行性和经济合理性；

3）确保场景应用的可操作性和可持续性。

d) 标准化原则：

1）遵循国家相关技术标准和规范要求；

2）建立统一的场景描述规范和评价体系；

3）确保不同城市间场景的可比较性和可推广性。

7.1.1.2 场景分类体系

场景分类体系如下。

a) 一级分类（按应用领域）：

1) A 类：城市治理场景

——A1 政务服务：网上办事、一网通办、跨域通办等；

——A2 应急管理：风险预警、应急指挥、资源调度等；

——A3 民意速办：网格管理、服务热线、一表通等。

2) B 类：数字生活场景

——B1 民生服务：教育、医疗、养老、就业等公共服务；

——B2 消费服务：智慧商圈、电子商务、消费维权等；

——B3 社会保障：社会救助、社会保险、社会福利等。

3) C 类：数字经济场景

——C1 产业服务：企业服务、产业链协同、创新孵化等；

——C2 智能监测：经济监测、政策服务、企业服务等。

4) D 类：城市更新场景

——D1 城市基础设施：城市巡检、车路协同等；

——D2 智慧社区：智慧物业、惠民服务、高楼消防等。

b) 二级分类（按技术复杂度）：

——基础级：数据查询、信息展示、简单交互；

——应用级：业务办理、流程优化、智能推荐；

——智能级：预测分析、决策支持、自动处置；

——协同级：跨部门联动、多方协同、一体化服务。

c) 三级分类（按服务对象）：

——市民端（C 端）：面向个人用户的服务场景；
——企业端（B 端）：面向法人单位的服务场景；
——政府端（G 端）：面向政府部门的管理场景。

7.1.1.3 场景识别方法论

场景识别方法论如下。

a) 需求调研方法：
   1) 问卷调查法：设计标准化问卷，收集各类用户需求；
   2) 深度访谈法：与关键利益相关方进行一对一访谈；
   3) 焦点小组法：组织专题讨论，深入挖掘潜在需求；
   4) 观察研究法：实地观察用户行为和业务流程。
b) 问题分析方法：
   1) 鱼骨图分析法：系统分析问题的根本原因；
   2) 帕累托分析法：识别影响最大的关键问题；
   3) SWOT 分析法：评估内外部环境和发展机遇；
   4) 价值链分析法：梳理业务价值创造过程。
c) 场景设计方法：
   1) 用户过程图：描绘用户完整的服务体验过程；
   2) 业务流程图：梳理标准化的业务处理流程；
   3) 系统架构图：设计支撑场景的技术实现架构；
   4) 数据流图：明确场景中的数据流转关系。

7.1.1.4 场景优先级评估

场景优先级评估要点如下。

a) 评估维度：
   1) 紧迫性：问题的严重程度和解决的迫切性；
   2) 影响面：涉及的用户规模和社会影响范围；
   3) 可行性：技术实现难度和资源投入要求；
   4) 效益性：预期产生的经济和社会效益；
   5) 示范性：对其他场景的引领和带动作用。
b) 评估方法：
   1) 采用层次分析法（AHP）确定各维度权重；
   2) 建立评分标准，对每个场景进行量化评估；
   3) 综合计算优先级得分，形成场景建设序列。

7.1.2 数据需求分析

7.1.2.1 数据需求分析方法

数据需求分析方法如下。

a) 场景驱动分析：
   1) 输入分析：识别场景运行所需的输入数据；
   2) 处理分析：确定数据处理和转换的要求；
   3) 输出分析：明确场景产出的数据成果；

4) 反馈分析：分析数据使用效果的反馈机制。

b) 业务流程分析：

1) 流程梳理：详细描述业务办理的完整流程；

2) 节点分析：识别各流程节点的数据需求；

3) 接口分析：明确系统间的数据交换接口；

4) 异常分析：考虑异常情况下的数据处理需求。

c) 技术能力分析：

1) 算法需求：分析 AI 算法对训练数据的要求；

2) 模型需求：确定模型训练和推理的数据规模；

3) 性能需求：评估数据处理的性能指标要求；

4) 扩展需求：考虑未来技术发展的数据需求。

7.1.2.2 数据质量要求

数据质量要求如下。

a) 准确性要求：

1) 建立数据质量评估指标体系；

2) 设定不同类型数据的准确率阈值；

3) 建立数据校验和纠错机制。

b) 完整性要求：

1) 定义数据完整性的评判标准；

2) 建立数据缺失检测和补全机制；

3) 确保关键数据字段的完整性。

c) 一致性要求：

1) 统一数据标准和编码规范；

2) 建立数据一致性检查机制；

3) 确保跨系统数据的一致性。

d) 时效性要求：

1) 根据业务需求确定数据更新频率；

2) 建立数据时效性监控机制；

3) 确保关键数据的及时更新。

7.1.2.3 数据安全与隐私保护

数据安全与隐私保护要点如下。

a) 数据分级分类：

1) 建立数据分级分类标准；

2) 实施差异化的安全保护措施；

3) 建立数据使用权限管控机制。

b) 隐私保护技术：

1) 采用数据脱敏技术保护个人隐私；

2) 运用差分隐私技术进行数据分析；

3) 实施数据最小化使用原则。

c) 合规性要求：

1) 严格遵守《数据安全法》等法律法规；

2) 建立数据处理合规审查机制；
3) 确保跨境数据传输的合规性。

7.2 打造高质量数据集

高质量数据集建设按照生命周期有序开展，包括数据需求、数据规划、数据采集、数据预处理、数据标注、模型验证等阶段。各阶段主要按以上顺序逐步开展，同时会对其他阶段进行反馈，或者会在其他阶段反馈下进行迭代优化。高质量数据集建设方法可参照TC609-5-2025-01。

7.3 构建 AI 能力

7.3.1 概述

城市智能体的AI能力构建是实现智能化城市治理的核心环节，遵循分层构建、能力复用、持续优化的原则，建立涵盖基础算力、模型训练、知识构建、能力封装的完整AI能力体系。

7.3.2 基础 AI 能力建设

7.3.2.1 算力基础设施

算力基础设施建设要点如下。

a) 计算资源配置
   1) GPU/TPU 集群：配置 AI 计算能力；
   2) CPU 集群：支持大规模并发推理和数据处理；
   3) 存储系统：实现高速存储，支持模型和数据的快速读写；
   4) 网络架构：建设高带宽、低延迟的内部网络，支持分布式训练。
b) 云边协同架构
   1) 云端训练中心：承担大模型训练和知识库构建任务；
   2) 边缘推理节点：部署在各业务场景，实时 AI 推理服务；
   3) 协同调度机制：实现云边资源的统一调度和负载均衡。

7.3.2.2 AI 开发环境

AI开发环境建设要点如下。

a) 模型开发环境
   1) 支持主流深度学习框架；
   2) 提供可视化模型设计和调试工具；
   3) 集成自动化机器学习（AutoML 等）能力。
b) 数据处理工具链
   1) 数据标注平台：支持文本、图像、视频等多模态数据标注；
   2) 数据增强工具：提供数据合成、变换、扩充等功能；
   3) 特征工程平台：支持特征提取、选择和组合。

7.3.3 大模型训练与优化

7.3.3.1 基础大模型构建

基础大模型构建要点如下。

a) 预训练模型选择

1) 通用语言模型；
2) 多模态模型：支持文本、图像、语音的统一理解和生成；
3) 代码生成模型：支持城市治理相关的自动化脚本生成。

b) 模型训练策略
1) 分布式训练：采用数据并行和模型并行相结合的训练方式；
2) 混合精度训练：提升训练效率；
3) 梯度累积和检查点：确保大规模模型训练的稳定性。

7.3.3.2 领域模型微调

领域模型微调要点如下。

a) 城市治理领域适配
1) 收集城市管理、公共服务、应急响应等领域的专业语料；
2) 构建领域特定的训练数据集。

b) 多任务学习框架
1) 文本理解：支持政策解读、舆情分析、投诉处理等任务；
2) 决策支持：提供资源调度、风险评估、预测分析等能力；
3) 交互对话：实现自然语言的人机交互和多轮对话。

7.3.3.3 模型评估与优化

模型评估与优化要点如下。

a) 评估指标体系
1) 准确性指标：精确率、召回率、F1 分数等；
2) 效率指标：推理延迟、吞吐量、资源利用率；
3) 安全性指标：对抗攻击防护、隐私保护能力。

b) 持续优化机制
1) 在线学习：基于实际应用反馈持续优化模型；
2) A/B 测试：对比不同模型版本的性能表现；
3) 模型蒸馏：将大模型知识迁移到轻量化模型。

7.3.4 知识库与语料库构建

7.3.4.1 城市知识图谱

城市知识图谱要点如下。

a) 知识体系设计
1) 实体识别：城市设施、组织机构、人员、事件等；
2) 关系定义：空间关系、时间关系、因果关系、从属关系；
3) 属性描述：静态属性和动态属性的完整刻画。

b) 知识获取与更新
1) 结构化数据抽取：从数据库、API 等获取标准化知识；
2) 非结构化数据挖掘：从文档、网页、社交媒体提取知识；
3) 专家知识录入：录入领域专业知识。

7.3.4.2 多模态语料库

多模态语料库建设要点如下。

a) 文本语料
   1) 政策法规文档：政策文件和法规条文；
   2) 业务流程描述：覆盖城市治理各领域的标准作业程序；
   3) 历史案例库：积累典型事件处理案例。
b) 多媒体语料
   1) 图像数据：城市监控、卫星遥感、无人机航拍等；
   2) 音频数据：热线电话、会议录音、现场采集等；
   3) 视频数据：监控录像、执法记录、培训材料等。

7.3.4.3 知识质量保障

知识质量保障要点如下。

a) 质量控制机制
   1) 数据清洗：去除重复、错误、过时的知识条目；
   2) 一致性检查：确保知识间的逻辑一致性和完整性；
   3) 专家审核：建立专家评审机制，确保知识的准确性。
b) 版本管理与追溯
   1) 知识版本控制：记录知识的创建、修改、删除历史；
   2) 来源追溯：标记知识的来源渠道和可信度等级；
   3) 更新机制：建立定期更新和实时更新相结合的机制。

7.4 搭建智能体

7.4.1 AI 能力封装与服务化

7.4.1.1 能力组件化

能力组件化要点如下。

a) 原子能力
   1) 感知能力：文本理解、图像识别、语音识别等；
   2) 认知能力：推理分析、知识问答、决策支持等；
   3) 生成能力：文本生成、图像生成、报告生成等。
b) 能力接口标准化
   1) API 接口规范：定义统一的调用接口和参数格式；
   2) 数据格式标准：规范输入输出数据的结构和编码；
   3) 服务质量保障：定义响应时间、可用性等 SLA 指标。

7.4.1.2 能力编排与组合

能力编排与组合要点如下。

a) 智能体架构选型：根据场景复杂度，确定采用多智能体架构或单智能体架构。
b) 智能体能力开发：开发智能体关键能力，包括意图感知、任务规划、任务执行、持续优化的技术实现。
c) 功能验证
   1) 通过模拟测试，验证其意图理解、任务规划、调度执行的正确性、对目标的达成度以及与其他智能体的协同效率。

2) 建立人工干预机制，允许开发与领域专家对智能体的不当决策进行纠正，并以此作为优化其模型的重要反馈数据。

7.4.1.3 能力运营与监控

能力运营与监控要点如下。

a) 性能监控
   1) 实时监控：监控 AI 服务的调用量、响应时间、错误率；
   2) 资源监控：跟踪计算资源、存储资源的使用情况；
   3) 业务监控：监控 AI 能力在具体业务场景中的效果。
b) 运维管理
   1) 自动扩缩容：根据负载自动调整服务实例数量；
   2) 故障恢复：提供服务降级、熔断、重试等容错机制；
   3) 版本管理：支持 AI 模型和服务的灰度发布和回滚。
c) 持续学习与优化：基于智能体在运行过程中产生的日志以及接收到的人工反馈，持续优化智能体。

7.4.2 AI 能力安全与合规

7.4.2.1 模型安全

模型安全要点如下：

a) 对抗攻击防护：
   1) 输入验证：对输入数据进行安全检查和过滤；
   2) 对抗训练：在训练过程中加入对抗样本提升鲁棒性；
   3) 异常检测：识别和拦截恶意输入和异常请求。
b) 隐私保护：
   1) 差分隐私：在模型训练中加入噪声保护个人隐私；
   2) 联邦学习：支持多方数据协作而不泄露原始数据；
   3) 数据脱敏：对敏感数据进行匿名化和脱敏处理。

7.4.2.2 合规管理

合规管理要点如下：

a) 算法审计：
   1) 公平性评估：检测模型是否存在歧视性偏见；
   2) 可解释性：提供模型决策的解释和依据；
   3) 透明度：公开模型的基本信息和使用范围。
b) 数据合规：
   1) 数据分类分级：按照敏感程度对数据进行分类管理；
   2) 访问控制：实施基于角色的数据访问权限控制；
   3) 审计日志：记录数据访问和使用的完整日志。

7.4.3 AI 能力评估与认证

7.4.3.1 能力成熟度评估

能力成熟度评估要点如下：

a) 评估维度：
   1) 技术成熟度：算法先进性、模型性能、系统稳定性；
   2) 应用成熟度：场景覆盖度、用户满意度、业务价值；
   3) 管理成熟度：治理体系、运营能力、安全保障。
b) 评估方法：
   1) 定量评估：基于客观指标的量化评估；
   2) 定性评估：专家评议和用户反馈；
   3) 第三方评估：引入独立机构进行客观评估。

7.4.3.2 认证与标准化

认证与标准化要点如下：

a) 能力认证：
   1) 功能认证：验证 AI 能力是否满足功能要求；
   2) 性能认证：验证 AI 能力是否达到性能指标；
   3) 安全认证：验证 AI 能力是否符合安全标准。
b) 标准化建设：
   1) 技术标准：制定 AI 能力的技术规范和接口标准；
   2) 管理标准：建立 AI 能力的管理制度和流程规范；
   3) 服务标准：定义 AI 能力的服务质量和交付标准。

7.5 智能体运营

7.5.1 安全运营

定期开展城市智能体安全测评。

7.5.2 资产运营

资产运营要点如下：

a) 通过安全评测的智能体，注册接入城市智能体资产管理。
b) 使用智能体资产管理的注册智能体，需要通过订阅或者订购流程。

7.5.3 业务价值运营

业务价值运营要点如下：

a) 基于城市全域数字化转型的总体目标和指标体系，计算智能体的业务价值，以及对于城市全域数字化转型的贡献度。
b) 基于业务价值，优化城市智能体资产。

7.5.4 运行维护

运营服务体系构成如下：

a) 运营管理模块：
   1) 智能体运行监控：实时监测智能体系统运行状态、性能指标、资源使用情况；
   2) 故障管理：建立故障预警、诊断、修复机制，实现故障快速处置；
   3) 性能优化：基于运行数据进行性能分析，优化系统配置和资源分配；
   4) 版本管理：负责智能体版本更新、升级、回滚等管理工作。

b) 服务保障模块：

1) 服务级别管理：制定服务级别协议（SLA），保障服务可用性、响应时间等指标；

2) 客户服务：建立多渠道服务支持体系，提供技术咨询、问题解答等服务；

3) 培训服务：为用户和运维人员提供操作培训、技术培训等。

参 考 文 献

[1] 关于深化智慧城市发展 推进城市全域数字化转型的指导意见（发改数据〔2024〕660号）
[2] 深化智慧城市发展 推进全域数字化转型行动计划（发改数据〔2025〕1306号）
[3] 国家数据基础设施建设指引（发改数据〔2024〕1853号）
[4] 生成式人工智能服务管理暂行办法
[5] 重庆市人工智能赋能超大城市现代化治理三年行动计划（2025—2027年）
[6] 中华人民共和国科学技术部 新一代人工智能伦理规范（2021年9月25日）