



中华人民共和国国家标准化指导性技术文件

GB/Z 226—2026

人工智能 模型即服务(MaaS)技术架构

Artificial intelligence—Model as a service(MaaS) technical framework

2026-08-27 发布

国家市场监督管理总局 发布
国家标准化管理委员会

目 次

前言 III

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 缩略语 1

5 技术架构 1

 5.1 概述 1

 5.2 用户层 2

 5.3 访问层 2

 5.3.1 概述 2

 5.3.2 访问控制 2

 5.3.3 访问管理 3

 5.4 模型层 3

 5.4.1 概述 3

 5.4.2 数据管理 3

 5.4.3 模型开发与交付 3

 5.4.4 模型应用 4

 5.5 资源管理层 4

 5.6 综合层 4

 5.6.1 概述 4

 5.6.2 运营运维支撑 5

 5.6.3 安全 5

 5.6.4 集成 5

前 言

本文件为规范类指导性技术文件。

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。

本文件由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本文件起草单位：中国电子技术标准化研究院、浪潮通信信息系统有限公司、中国移动通信集团有限公司、中国南方电网有限责任公司超高压输电公司、阿里云计算有限公司、深圳思谋信息科技有限公司、上海文镔信息科技有限公司、浪潮电子信息产业股份有限公司、中移雄安信息通信科技有限公司、浙江大华技术股份有限公司、浪潮软件科技有限公司、百度云智(北京)科技有限公司、西门子(中国)有限公司、宁波宁帆信息科技有限公司、中国科学院信息工程研究所、上海商汤智能科技有限公司、中移(苏州)软件技术有限公司、国家能源投资集团有限责任公司信息技术分公司、华为云计算技术有限公司、中电信数智科技有限公司、山东省人工智能研究院、北京神舟航天软件技术股份有限公司、广电运通集团股份有限公司、深信服科技股份有限公司。

本文件主要起草人：徐洋、杨雨泽、杨语澈、崔洪志、鲍薇、曹汐、彭凌、张传刚、彭骏涛、秦日臻、梁秉豪、李学武、肖红梅、李睿宇、陈鹏光、仲凯韬、邸贺亮、杨佳丽、孔维生、张峰、刘增志、郑子木、王珂琛、生若谷、章杰标、刘井强、宣善明、王静、刘紫君、吴庚、李旭东、周良杰、张天霖、曹晓琦、陈利明、单珂、密伟、杨波、尚啸、高天雷、潘风文、高宇、徐金涛。

人工智能 模型即服务(MaaS)技术架构

1 范围

本文件给出了技术实现视角的模型即服务(MaaS)各类过程的详细描述。

本文件适用于对运用 MaaS 进行模型开发、使用、管理提供指导。

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/Z 221—2026 人工智能 模型即服务(MaaS)参考架构

3 术语和定义

GB/Z 221—2026 界定的以及下列术语和定义适用于本文件。

3.1

模型即服务 model as a service

提供模型开发、使用和管理能力及所需资源且可规模化复制的服务模式。

4 缩略语

下列缩略语适用于本文件。

AI:人工智能(Artificial Intelligence)

API:应用程序编程接口(Application Programming Interface)

MaaS:模型即服务(Model as a Service)

SDK:软件开发工具包(Software Development Kit)

5 技术架构

5.1 概述

图 1 是 GB/Z 221—2026 给出的 MaaS 的各类过程的总体结构。该总体结构表述了运用 MaaS 进行模型开发、使用、管理时所需的各类过程。这些过程按用户层、访问层、模型层、资源管理层,以及综合层予以归纳。

用户对基于 MaaS 开发的产品的使用,首先是通过用户层与访问层的交互接口进行操作,收到用户层的指令和数据后,访问层进行初步处理后,将相关数据和指令转发至其他相关层,执行相应的过程,直至完成该次任务,将执行结果反馈给用户。执行任务中可能需要执行不同过程,例如可能涉及资源管理类过程,在完成相关资源管理类过程后,将信息通过访问层传递给用户层,并通过用户层的交互界面最终反馈给用户。