ICS 35. xxx
CCS Lxx

团　　体　　标　　准

T/ISC XXX—XXXX

企业级大模型和应用全栈服务能力

Enterprise-Level Large Models and Full-Stack Application Service Capabilities

在提交反馈意见时，请将您知道的相关专利与支持性文件一并附上。

（征求意见稿）

2025-12-24

XXXX－XX－XX 发布　　XXXX－XX－XX 实施

中 国 互 联 网 协 会　　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：中国信息通信研究院、

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：

企业级大模型和应用全栈服务能力

1. 范围

本文件规定了企业人工智能应用成熟度评价模型。

本文件适用于指导第三方测评机构对企业建设大模型和应用全栈服务能力成效进行阶段性评估。

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 5271.1-2000 信息技术 词汇 第1部分：基本术语

GB/T 5271.28-2001 信息技术 词汇 第28部分：人工智能 基本概念与专家系统

3. 术语和定义

3.1 模型训练 model training

利用训练数据，基于机器学习算法，确定或改进机器学习模型参数的过程。[来源：GB/T 41867-2022，3.1.2]

3.2 微调 fine-tuning

为提升人工智能模型的预测精确度，一种先以大型广泛领域数据集训练，再以小型领域数据集继续训练的附加训练技术。[来源：GB/T 41867-2022，3.1.2]

3.3 训练数据 training data

用于训练机器学习模型的输入数据样本子集[GB/T 41867-2022]

3.4 模型评估 model evaluation

通过既定的各类人工智能任务评估指标，对训练生成的模型进行质量评判。

3.5 结构化数据 structured data

一种数据表示形式，按此种形式，由数据元素汇集而成的每个记录的结构都是一致的并且可以使用关系模型予以有效描述。[来源：GB/T 35295-2017，定义 2.2.13]

3.6 非结构化数据 unstructured data

相对于结构化数据（即行数据，存储在数据库里，可以用二维表结构来逻辑表达实现的数据）而言，不方便用数据库二维逻辑表来表现的数据即称为非结构化数据。[来源：GB/T 35295-2017，定义 2.1.25]

4. 概述

表 1 大模型和应用全栈服务能力总体架构

能力域	能力项
基础设施	计算
	存储
	网络
数据	数据采集
	数据存储
	数据治理
	数据管理
	数据评估
	数据安全
模型	模型训练
	模型推理
	模型评估
	模型管理
	模型安全
应用	应用开发
	应用部署
	应用监控
	应用管理
	应用安全

5. 基础设施建设能力

基础设施维度规范了企业建设的基础设施时应该具备的计算、存储、网络能力。

5.1 计算

——应支持通用计算资源，例如X86、ARM等架构的CPU通算资源；
——应支持GPU、NPU等智算资源；
——应支持提供基于容器的智能计算集群；
——应支持硬件资源的虚拟化；
——应支持计算资源弹性扩缩容。

5.2 存储

——应支持挂载高效读写硬盘，例如SSD、HDD等多种类型硬盘；
——应支持存储资源管理及弹性扩缩容。

5.3 网络

——应支持多类型网络接口；
——宜应支持高性能网络交换机等硬件设备；
——宜支持高性能RDMA网络；
——应具备提供网络资源管理能力，包括：

- 应支持创建、删除、编辑、查询虚拟私有网络；
- 应支持私有网络IP地址管理功能，如IP地址的自动和手动分配；
- 应支持创建、删除、编辑、查询虚拟子网；
- 宜支持查看私有网络拓扑信息；
- 宜支持虚拟网关和路由。

6. 数据建设能力

数据维度规范了企业应该具备的数据采存治管评等能力。

6.1 数据采集

——应支持多种数据源的导入，包括内置数据和外部知识数据（通过API或其他方式接入外部知识源，如搜索引擎、开放数据库等）；
——应支持多模态数据的导入，包括文字、文档、图片、音频、视频、结构化数据等知识数据模态。

6.2 数据存储

——应支持多种数据类型的存储，包括结构化数据（如表格、关系型数据等）、半结构化数据（如JSON、XML、API接口数据等）和非结构化数据（包括文档、图片、音频和视频等）的存储；
——应支持多种存储方式：

- 应支持使用传统数据库系统存储结构化数据；
- 应支持使用分布式存储系统存储非结构化数据；
- 应支持使用图数据库存储知识图谱数据；
- 应支持使用向量数据库存储向量数据。

6.3 数据处理

6.3.1 数据清洗

——应支持去重功能，消除数据集内部及不同数据集之间的重复数据，确保数据的唯一性和一致性。数据去重是指通过特定的算法和技术手段，识别并移除数据中完全相同或高度相似的记录，避免数据冗余，提高数据的准确性和处理效率；
——应提供过滤功能，识别并移除敏感、不当或不符合要求的数据，确保数据的合规性和安全性。数据过滤是指依据预设的规则和标准，对数据进行筛选和审查，剔除其中的非法、有害、低质量或与目标不相关的内容，从而保障数据的合法性和可用性；
——应支持转换功能，为模型训练提供标准化的数据格式。数据转换是指通过对数据进行一系列的处理操作，如格式调整、编码转换、数据归一化等，将原始数据转换为适合模型训练和分析的统一格式，来提升数据质量、增强数据的一致性，同时通过敏感信息替换等方式保护数据隐私。

- 文本转换：支持格式转换、编码转换、多语种互译、文本结构化（如 QA 对提取、实体识别）、敏感词替换、文本摘要生成等；
- 图像转换：支持背景处理、人脸模糊、分辨率增强、尺寸调整、格式转换（如 RGB 标准化）等；

- 视频转换：支持格式转换、视频分割、关键帧提取、分辨率调整、标签生成（基于帧内容分析）、动态物体追踪等；
- 音频转换：支持高斯加噪、语音增强、格式转换（如 WAV 转 MP3）、语音文本转录、采样率调整等；
- 多模态转换：支持跨模态数据提取（如图片 OCR 文字识别、公式结构解析、视频中语音与画面同步标注）。

6.3.2 数据标注

6.3.2.1 文本标注

——应支持对文本进行自动化标注。

- 应支持对象自动提取功能，自动标注文本中的人名、机构名、地点等实体信息；
- 应支持标签自动生成功能，对从文本中提取出的内容分配语义标签（如词性、情感、主题等）以及实体间的关联关系；

——应支持对文本数据进行人工标注。

6.3.2.2 图片标注

——应支持对图片进行自动化标注；

- 应支持对象自动提取功能，从图片中定位并提取出具体的实体对象；
- 应支持标签自动生成功能，对从图片种提取出的对象或内容分配标签；
- 应支持图生文功能，基于图像内容自动生成结构化描述文本（如物体识别、场景描述、属性标注），并确保生成文本与图像语义的一致性。

——应支持对图片进行人工标注。

6.3.2.3 视频标注

——应支持对视频进行自动化标注；

- 应支持关键帧自动化提取，基于内容变化率自动标注视频关键帧时间戳；
- 应支持行为分析标注，识别并标注特定业务行为（如客户接待、设备操作）；
- 应支持多模态关联标注，同步标注视频画面与对应音频文本内容；

——应支持对视频进行人工标注。

6.3.2.4 音频标注

——应支持对音频进行自动化标注；

- 应实现语音转写，将音频自动转写为带时间戳的文本；
- 宜支持声纹特征标注，区分并标注不同说话人的声纹特征；
- 应具备情感分析标注，根据语调自动标注积极/消极/中立情绪；

——应支持对音频进行人工标注。

6.3.2.5 标注审核

——应支持标注结果的质量审核机制包括抽样复核、校验规则校对等方式，确保标注数据的准确性和一致性。

6.3.3 数据配平

——应支持数据比例均衡功能，支持各类别样本的比例自动或手动调整，确保数据集中各类别数据的分布合理，避免模型训练中的偏差问题；

——应提供数据采样功能，对过多或过少的数据类别进行合理调整，优化数据集结构；

——宜支持数据分布分析功能，帮助识别并纠正数据不平衡问题，提出配平建议或操作支持。

6.3.4 数据增强

数据增强是在学习模型时通过对原始数据进行变换、扰动或组合，生成额外的训练样本，来扩充训练数据集，提高模型的泛化能力。

6.3.4.1 文本数据增强

——应支持对文本进行同义词替换处理；

——应支持对文本进行随机插入、随机删除、句式结构调整、语序变换等处理。

6.3.4.2 图片数据增强

——应支持对图片进行平移、翻转、旋转、缩放处理；

——应支持对图片进行裁剪处理；

——应支持对图片进行色彩（亮度、饱和度、对比度等）变换处理；

——宜支持对图片进行噪声注入。

6.3.4.3 音频数据增强

——应支持对音频进行重采样处理；

——应支持对音频进行变声（如调整音量、音高、速度）处理；

——应支持对音频进行音频切割、音频拼接处理；

——应支持对音频进行加噪/降噪处理；

——宜支持对音频的时域/频域进行扰动处理。

6.3.4.4 视频数据增强

——应支持对视频进行帧率、分辨率调整；

——应支持对视频进行裁剪；

——应支持对视频的色彩、色调进行调整。

6.3.5 数据混合拆分

——应支持多数据集混合功能，将来自不同来源、格式、类型、归属不同数据集的知识数据统一处理，形成完整的知识体系；

——应提供数据融合功能（例如结构映射、格式转换、语义对齐），确保不同数据集之间的数据格式和结构兼容性；

——应支持数据集拆分功能，将大数据集拆分成多个子集。

6.3.6 文档切分

文档切分是在生成知识库时将大块文档切分成小块，用于增加模型推理时从数据库召回的准确性。

——应支持对多种格式的文档进行切分，包括但不限于word、excel、pdf等；

——应支持多种分段配置方式，包括但不限于自定义配置分段策略(固定大小、基于结构、上下文组合)、分段最大长度设置等；

——应支持多种分段策略，包括但不限于固定长度分段、自适应分段。

6.4 数据管理

6.4.1 数据目录

——应支持数据可视化管理能力，提供数据目录功能，方便用户管理和使用；
——应支持数据分类分级管理能力，提供数据分类分组、数据分级等功能。

6.4.2 质量管理

——应支持数据质量评估功能，针对文本、图片、音频、视频等不同模态的数据，制定相应的质量评估指标（如清晰度、完整性、一致性等），确保数据的整体质量；
——应支持数据质量修复功能，通过自动化或人工干预方式，对低质量数据进行修复或替换，提升整体数据质量；
——宜提供数据质量报告功能，生成详细的质量分析报告，帮助用户全面了解数据质量状况并优化数据处理流程。

6.4.3 标签管理

——应支持标签用途管理功能，明确标签的具体用途（如用于模型蒸馏、多轮对话训练等），确保标签与任务目标的高度匹配；
——宜支持标签来源管理功能，记录标签的生成方式（如自动化标注、人工标注）及来源（如内部标注团队、第三方服务），确保标签的可追溯性和可靠性；
——宜提供多模态标签关联功能，将文本、图片、音频、视频等不同模态的标签进行关联，形成统一的标签体系，提升数据的综合应用能力。

6.5 数据安全

——应建立多维度数据分类分级机制，明确敏感数据标识规则与访问边界，实现训练数据、推理数据及用户隐私数据的差异化管控；
——应采用动态脱敏技术对训练数据集进行特征值模糊化处理，并在数据存储与传输过程中实施端到端加密策略，确保原始数据不可还原；
——应通过硬件级隔离或虚拟化技术实现推理环境与核心业务系统的物理 / 逻辑分离，防范模型输出泄露敏感信息；
——应建立数据操作全链路日志追踪机制，对数据导出、模型参数修改等高危行为实施双人复核与实时告警；
——应部署 AI 驱动的异常流量监测系统，实时识别非授权数据访问、模型逆向工程等攻击行为。

7. 模型建设能力

模型维度规范了企业在构建模型时，应该具备的训练、推理、评估、管理等维度的能力。

7.1 模型训练

7.1.1 模型训练

——应支持训练任务部署在NPU/GPU多种硬件资源上；

——应支持对训练任务进行创建、查询、管理、停止、删除等生命周期管理；
——应支持提供多种主流训练框架，例如，Tensorflow、PyTorch等；
——应支持利用数据集进行模型训练；
——应支持模型训练的版本管理，可对指定版本进行重新训练；
——应支持单机多卡模式并行训练；
——宜支持多机多卡模式分布式训练；
——宜支持实时输出模型训练信息，包含训练时长、迭代次数等；
——宜支持模型训练模板配置，包含超参设置等；
——宜支持模型训练过程的可视化，包含收敛、损失函数等参数值。

7.1.2 模型优化

7.1.2.1 模型微调

——应支持创建和管理精调任务，提供全参微调、Lora、Qlora等精调方式。

7.1.2.2 模型量化

——宜支持对模型进行剪裁、轻量化、蒸馏等操作，以降低模型计算复杂度的能力。

7.1.3 训练任务加速

具备提供适用于智能计算的网络、存储等优化和对模型训练加速的工具或组件，具体要求如下：
——应支持提供面向大量IO吞吐场景的高性能存储服务；
——应支持提供高性能网络、通信加速工具；
——应支持提供针对AI训练加速工具。

7.1.4 断点续训

——应支持模型在训练出现故障导致任务中断后，快速恢复。

7.2 模型评估

——应支持模型评估测试数据集管理，包含测试数据集的创建、维护和查看等功能；
——应支持自定义评估指标库，涵盖分类任务（准确率、F1值）、检测任务（mAP、IoU）、生成任务（BLEU、ROUGE）等场景化指标；
——应支持基于测试集对模型进行评估；
——应支持查看评估报告；
——应支持筛选及导出评估分析结果，便于用户进行进一步的分析和处理。

7.3 模型推理

7.3.1 模型推理

——应支持多种深度学习模型服务化框架，以支持不同模型的部署和运行；
——应支持模型推理服务的弹性资源调度，包括手动、自动扩容缩容，以适应不同负载下的服务需求。

7.3.2 模型部署

——应支持多种部署方式，例如云端和分布式部署，来适应不同硬件环境，满足多样化的部署需求；
——应支持多种模型精度部署，如 BF16、INT8 等，以适应不同计算能力和资源限制。

7.3.3 提示词管理

——应支持预置提示词；
——宜具备提示词模板制定和模板选择功能。

7.4 模型管理

7.4.1 模型服务管理

——应支持模型服务的添加、删除等编辑功能；
——应支持生成API Key、API接口。

7.4.2 模型监控

——应支持对模型所占用的资源使用进行监控，包括计算、存储、网络；
——应支持查看调用总量、调用失败、调用tokens数，输入、输出tokens；
——应支持按分钟、按小时、按日查看各监控项曲线走势图；
——宜支持查看错误次数、错误占比、错误码等调用失败数据。

7.5 模型安全

——应支持模型训练安全，包括但不限于提供多种手段防范外部恶意攻击、面对恶意攻击（如数据投毒、后门攻击）的应具备相应的防御性能力；
——应支持在大模型推理时的实时监控与预警的能力，以及对大模型进行恢复与应急处理。

8. 应用建设能力

应用维度规范了企业在开发AI应用时，应该具备的开发、部署、监控、管理等维度的能力。

8.1 应用开发

——应支持集成开发环境，实现代码级应用开发；
——应支持基于低无代码的可视化应用开发，以实现复杂业务逻辑的集成；
——应支持将开发环境中的容器打包成镜像，用户可自定义镜像名称和版本。

8.2 应用部署

——应支持以API、SDK的方式部署应用，提供灵活的集成选项；
——应支持应用部署策略，包括但不限于：蓝绿部署、滚动升级等，且支持自动回滚，以确保部署过程的稳定性和可靠性；
——应支持以容器镜像的方式部署应用，以提高部署的便捷性和一致性。

8.3 应用管理

——应支持对应用API调用接口进行管理；
——应支持应用的全生命周期管理，包括（批量）部署、删除、版本升级、更新补丁等。

8.4 应用监控

——应支持对平台部署应用的效果进行监控，包括准确率、召回率等指标；

——应支持对平台服务的性能进行监控，包括平台吞吐率、响应时延、并发能力、资源占用率、运行稳定度等。

8.5 应用安全

——应支持对 API 调用行为监测系统，对异常高频访问、非授权数据提取等行为实施实时阻断与溯源；

——应支持涉黄、涉爆、涉政、涉恐等违规敏感内容的审核；

——应实现基于语义理解的上下文关联审核，对隐晦表述等高风险内容进行深度挖掘与分级处置。