

团 体 标 准

T/ISC XXX—XXXX

Claw 类企业级智能体能力要求

Capability requirements of claw enterprise-level agents

（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
Claw 企业级智能体能力要求	1
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
3.1 智能体 ai agent	1
3.2 企业级智能体 enterprise-level agent	1
3.3 Claw 类智能体 claw-type agent	1
3.4 大模型 large-scale model	2
3.5 强化学习 reinforcement learning	2
3.6 自然语言理解 natural language understanding	2
3.7 提示词 prompt	2
3.8 技能 Skill	2
4 符号和缩略语	2
5 能力概述	3
6 模型接管能力	3
7 感知交互能力	3
7.1 知识处理	3
7.2 意图理解	3
7.3 上下文管理	4
8 规划执行能力	4
8.1 任务规划	4
8.2 任务执行	4
8.3 执行操作	4
8.4 学习能力	5
9 技能 Skills 平台与工具链能力	5
9.1 技能注册与生态	5
9.2 技能审核机制	6
9.3 技能管理运营	6
9.4 技能元数据规范	6
9.5 二次开发与扩展	6
9.6 基础技能库	6
10 可信安全能力	7
10.1 权限管理	7
10.2 数据安全	7
10.3 执行透明	7
10.4 行为风险可控	7

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、中国第一汽车集团有限公司、天翼云科技有限公司、中国移动通讯集团浙江有限公司、中移信息技术有限公司、中原银行股份有限公司、中国联合网络通信有限公司软件研究院、浪潮通用软件有限公司。

本文件主要起草人：刘思颖、崔晨星、罗欧、徐恩庆、陈韵、陈天澈、武沛多、魏旭、郑少斌、王爱娇、张巍、李文伟、史露强、蒋通通、王淼鑫、吉文越、郭艺娟、古英杰、申佳、王力强、叶志欣、顾永杰、白阳、叶剑、蒋传辉。

Claw 企业级智能体能力要求

1 范围

本文件规定了Claw类企业级智能体能力要求，包括模型接管能力、感知交互能力、规划执行能力、技能Skills平台与工具链能力、可信安全能力方面。

本标准适用于指导企业级用户规划实施并运维Claw类企业级智能体使用，以及指导服务提供方设计、研发此类产品，也可作为第三方测评机构评估使用。

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。

- [1] GB/T 41867-2022 信息技术 人工智能 术语
- [2] GB/T 5271.28-2023 信息技术词汇 第28部分：人工智能
- [3] GB/T 45288.1—2025 人工智能 大模型 第1部分：通用要求
- [4] GB/Z 185.1—2026 人工智能 智能体互联 第1部分：总体架构

3 术语和定义

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是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。

3.1

智能体 ai agent

感知环境，并主动采取行动以实现特定目标的实体。

注：本文件中的智能体特指人工智能智能体，一般为软件系统。

[来源：GB/Z 185.1—2026，3.1]

3.2

企业级智能体 enterprise-level agent

基于大语言模型或基础模型构建，部署于企业业务环境中，具备自主感知、推理规划、多步骤任务执行及跨系统工具调用能力的软件实体。

3.3

Claw类智能体 claw-type agent

具备直接操作计算机系统、应用程序或外部设备能力的智能体，可通过自然语言指令自主完成文件处理、代码编写与执行、浏览器操作、终端命令运行、鼠标键盘控制等具体任务。

3.4

大模型 large-scale 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:参数量规模达到行业公认的大型模型标准,通常指参数量达到亿级或以上的深度学习模型。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]

3.5

强化学习 reinforcement learning

一种通过与环境交互,学习最佳行动序列,使回报最大化的机器学习方法。

[来源:GB/T 41867—2022, 3.2.25]

3.6

自然语言理解 natural language understanding

通过对功能单元从已传入的功能单元中的自然语言形式的文本或语音中的提取信息,并产生对给定文本或语音及其表示的描述。

[来源:GB/T 5271.28-2023, 28.01.18]

3.7

提示词 prompt

提示语

使用大模型进行微调或下游任务处理时,插入到输入样本中的指令或信息对象。

[来源:GB/T 45288.1—2025, 3.5]

3.8

技能 Skill

是指封装了特定功能或服务的可插拔模块,具备标准化的输入输出接口,可被智能体动态发现、加载与调用,用于扩展智能体在信息检索、内容创作、数据分析、系统操作等领域的能力。

4 符号和缩略语

下列符号和缩略语适用于本文件。

A/B测试: A/B测试 (A/B Testing)

API: 应用程序接口 (Application Programming Interface)

ID: 标识符 (Identifier)

IVR: 交互式语音应答 (Interactive Voice Response)

JSON: JavaScript对象表示法 (JavaScript Object Notation)

OA: 办公自动化 (Office Automation)

OCR: 光学字符识别 (Optical Character Recognition)

RBAC: 基于角色的访问控制 (Role-Based Access Control)

SCA: 软件成分分析 (Software Composition Analysis)

SQL: 结构化查询语言 (Structured Query Language)

SSH: 安全外壳协议 (Secure Shell)

URL: 统一资源定位符 (Uniform Resource Locator)

5 能力概述

Claw类企业级智能体能力应包括模型接管能力、感知交互能力、规划执行能力、技能Skills平台与工具链能力、可信安全能力等几方面，其能力要求框架如图1所示。

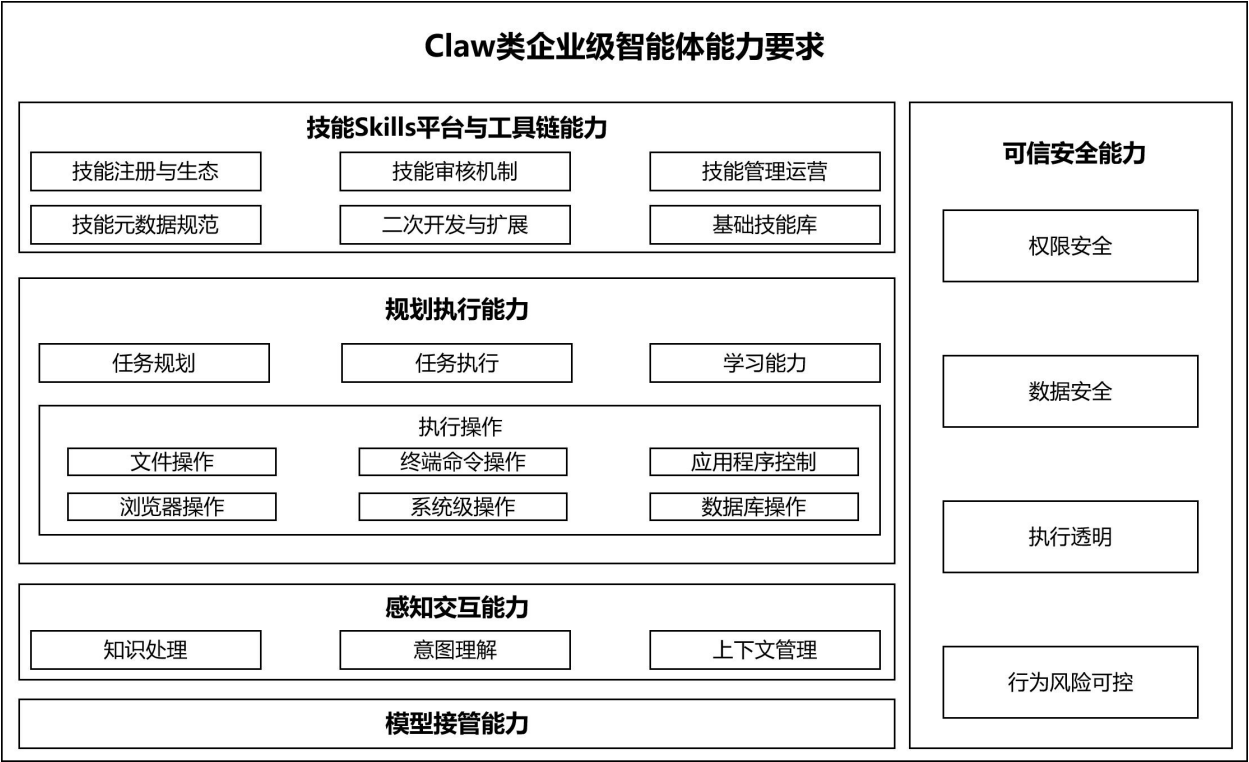


图 1 Claw 类企业级智能体能力要求框架

6 模型接管能力

- a) 应支持主流大模型的接入，如可接入Claude、GPT、DeepSeek、Qwen等主流AI模型及本地开源模型，提供匹配的模型注册、配置与调用接口，允许用户可根据需求自主选择、切换与组合使用模型；
- b) 宜支持模型版本管理与灰度发布能力，支持多版本模型并存、A/B测试及按流量比例灰度切换；
- c) 宜具备模型路由与负载均衡能力，支持根据任务类型、模型性能、成本及响应时效等策略自动分发请求至最优模型实例；

7 感知交互能力

7.1 知识处理

- a) 应支持多源异构知识数据的统一接入，如结构化数据、非结构化文档、半结构化数据及实时流数据；
- b) 应支持知识数据的预处理与向量化能力，如内容分块、向量化、索引构建及元数据管理等；

7.2 意图理解

- a) 应支持从用户指令中准确识别任务目标、操作对象及约束条件，具备单意图与多意图联合识别能力；
- b) 应支持槽位填充与必填信息补全，对缺失关键参数的任务指令，能够主动引导用户补充或提供默认值建议；
- c) 应支持区分事实性查询和操作性指令及对话式交互，针对不同类型请求采取相应的处理策略与响应模式；
- d) 宜支持复杂条件意图识别，能够解析包含否定、条件、比较及时间约束等复杂语义的用户指令；

7.3 上下文管理

- a) 应支持对历史交互记录的存储和检索功能，能够根据用户需求快速定位并调用过往对话中的关键信息；
- b) 应支持按短期记忆，如对话历史、任务状态等；
- c) 宜支持跨对话的长期记忆，如用户偏好、常用操作路线等；
- d) 宜支持记忆生命周期管理，如设置记忆有效期、自动归档及清理过期对话痕迹，避免信息冗余、隐私泄露等；
- e) 宜支持上下文窗口压缩与摘要机制，在上下文长度超出模型限制时，自动进行关键信息提取与历史摘要；

8 规划执行能力

8.1 任务规划

- a) 应具备任务理解能力，支持对用户输入任务目标的意图及关键信息准确识别；
- b) 应具备任务分解能力，支持将复杂任务拆解为多个目标明确、步骤清晰的子任务；
- c) 应具备路径规划能力，支持基于可配置参数将复杂任务生成最优原子操作序列；
- d) 宜具备动态调整能力，支持依据实时数据与业务变化自适应调整已规划的任务路径；
- e) 宜具备任务优先级评估能力，支持按紧急度、重要性等因素对任务进行优先级排序；

8.2 任务执行

- a) 应具备任务执行能力，支持依据规划结果自主完成数据处理、文件操作、系统指令及网络请求等操作；
- b) 应具备人机协作能力，支持任务执行过程中人工干预，如补充任务数据、修改任务内容或打断任务进程等；
- c) 宜具备跨系统协同能力，支持通过动态适配异构API接口与常规协议，实现与其他系统或智能体的任务级联与数据互通；
- d) 宜具备状态监控能力，支持可视化实时展示任务进度、任务状态、资源占用及异常告警情况等；

8.3 执行操作

8.3.1 文件操作

- a) 应支持目录的浏览与切换支持能够获取文件列表、文件的创建、读取、写入、删除、移动、复制操作；
- b) 应支持文本文件、JSON、Markdown等常见格式的读取与写入；
- c) 宜支持文件内容检索与比对，包括关键词搜索、正则匹配、文件差异比较等；

8.3.2 终端命令操作

- a) 应支持在终端环境中执行系统命令；
- b) 宜支持通过SSH协议连接远程机器并执行命令；
- c) 宜支持命令执行结果的捕获与结构化解析；
- d) 宜支持命令执行的安全管控，命令白名单、危险操作拦截、执行超时控制等；

8.3.3 应用程序控制

- a) 应支持通过命令行或协议唤起方式启动应用程序；
- b) 应支持关闭指定的应用程序；支持窗口管理操作，包括移动、缩放、切换等；

8.3.4 浏览器操作

- a) 应支持打开指定URL的网页；
- b) 宜支持页面导航操作，包括前进、后退、刷新；
- c) 宜支持页面元素的识别与交互，包括点击按钮、填写表单等；

8.3.5 系统级操作

- a) 应支持系统剪贴板的读取与写入；
- b) 宜支持系统通知的发送；
- c) 宜支持系统定时任务的创建与管理，定时脚本执行、周期性任务调度等；
- d) 宜支持系统服务管理操作，服务的启动、停止、状态查询及配置修改；

8.3.6 数据库操作

- a) 应支持主流数据库类型的连接与访问；
- b) 应支持SQL语句的执行与结果解析；
- c) 宜支持数据库元数据查询；

8.4 学习能力

- a) 宜支持能够基于积累的用户交互数据、任务执行数据等，通过机器学习算法（如监督学习、无监督学习、强化学习等）进行模型训练和优化，不断提升自然语言理解、任务执行等能力；
- b) 宜支持用户反馈机制，能够根据用户对任务执行结果的评价（如满意、不满意、改进建议等），对自身行为和策略进行调整和优化；

9 技能 Skills 平台与工具链能力

9.1 技能注册与生态

- a) 应支持接入官方技能商店及社区开发者开发的技能模块，一键安装、启用，无需复杂配置，可快速扩展功能范围；
- b) 应支持建立技能注册中心，对所有已注册技能进行统一标识、索引和检索管理，支持按功能分类、输入输出类型、审核状态等维度筛选；
- c) 应支持对内部技能仓库设置访问权限控制，如基于角色（RBAC）的读取、安装、管理等分级权限，限定仓库可见范围至特定组织、部门或用户组；
- d) 应支持对第三方Skills技能进行严格的安全审核与签名认证，防止恶意代码注入；

9.2 技能审核机制

- a) 应支持建立技能审核流程，技能模块需经过人工审核或自动化审核（含静态分析与动态行为检测）方可发布启用；
- b) 应具备多维度的审核机制，如涵盖业务合理性、边界安全性、输出质量等不同维度；
- c) 应支持对通过审核的技能进行数字签名认证，运行时校验签名完整性，拒绝加载未经签名或签名失效的技能包；
- d) 宜支持建立技能安全漏洞应急响应机制，对已发布技能发现高危漏洞时，能够快速下架、推送修复版本并通知受影响租户；
- e) 宜支持审核结果结构化记录，如记录审核人、审核时间、审核意见、通过或驳回状态等，并提供申诉或重新提交审核流程；

9.3 技能管理运营

- a) 应提供技能运营管理后台，支持对技能进行上线、下线、禁用、版本更新等全生命周期操作，并记录所有变更的操作日志；
- b) 宜支持技能使用数据统计，包括调用次数、调用成功率、平均响应时延、错误分布等核心指标，为运营决策提供数据依据；
- c) 宜支持技能评价与反馈机制，允许用户对已使用技能提交评分与意见，运营人员可基于反馈进行质量优化和分类治理；
- d) 宜提供重复技能合并策略，支持基于语义相似度的重复技能检测，识别同一能力的不同表达方式，并对重复技能进行合并；

9.4 技能元数据规范

- a) 应具备唯一标识符，每个技能应具有全局唯一的标识符（ID），用于技能的精确引用、版本追踪和依赖管理；
- b) 应具备名称与描述信息，技能应包含人类可读的名称和功能描述，描述应清晰说明技能的用途、适用场景和核心能力，便于用户理解和选择；
- c) 应支持输入参数规范，技能应通过结构化Schema定义其输入参数的名称、类型、必填性、取值范围及默认值，支持JSON Schema等标准化格式，确保参数传递的准确性和一致性；
- d) 应支持输出格式定义，技能应明确定义其输出内容的结构和格式，包括返回数据类型、字段含义、可选的可视化呈现方式等；
- e) 宜支持版本管理，技能应包含版本号信息，支持语义化版本控制（如Major.Minor.Patch），便于技能的迭代更新和兼容性管理；
- f) 宜支持提供者信息，技能应包含开发者或提供方信息，包括名称、联系方式、官方文档链接等，便于用户获取技术支持和反馈问题；
- g) 宜支持依赖关系声明，技能应声明其依赖的外部服务、第三方库或其他技能，便于部署时的环境检查和依赖安装；

9.5 二次开发与扩展

- a) 应支持用户自定义开发Skills技能模块，提供标准化的开发接口和文档，方便用户根据自身需求开发专属功能，开发的技能模块可正常集成到智能助手；
- b) 宜支持提供Skills技能的模拟调试环境，便于开发者在安全环境下验证功能；

9.6 基础技能库

- a) 宜具备信息检索类技能库，如搜索引擎、学术文献检索，具备实时联网抓取能力；
- b) 宜具备内容创作类技能库，如智能PPT生成、文生图/绘本、播客/视频脚本生成等能力；
- c) 宜具备深度分析类技能库，如多步骤研究分析（Agent模式），能够输出深度报告；
- d) 宜具备系统运维类技能库，如日志分析、系统监控、故障排查辅助等；
- e) 宜具备数据处理类技能库；如数据清洗、格式转换、数据可视化等；

10 可信安全能力

10.1 权限管理

- a) 应支持基于角色的访问控制（RBAC），通过定义不同的用户角色并分配相应的权限级别；
- b) 应支持根据业务需求的变化实时调整用户的权限设置，保证权限管理的灵活性和适应性；
- c) 宜支持对智能体工具调用权限的细粒度管控，按技能、数据源及操作类型进行授权与审计；

10.2 数据安全

- a) 应采用加密算法对存储和传输中的数据进行加密处理，保护数据免受未经授权的访问或泄露；
- b) 应支持数据分类分级管理，对敏感数据（如财务数据、客户数据、商业机密）进行自动识别、标记、脱敏及访问控制；
- c) 应支持定期执行数据备份操作，并制定完善的数据恢复计划；

10.3 执行透明

- a) 应支持安全审计追踪能力，如所有智能体调用记录、配置变更、权限调整等关键操作；
- b) 应支持记录完整的安全操作日志，如命令输入、文件访问、网络请求等，且审计日志不可篡改，便于事后追溯；
- c) 宜支持模型决策可解释性，对关键业务场景的智能体决策提供推理过程、依据来源及置信度说明；

10.4 行为风险可控

- a) 应支持内容安全能力，如多级内容审核机制、权威知识验证技术、全维度伦理审查体系等；
- b) 应支持提示词安全防御能力，具备对提示词注入攻击（Prompt Injection）、越狱攻击（Jailbreak）及提示词泄露的检测与防护机制；
- c) 宜支持风险预判能力，在执行指令前应进行风险评估，对可能产生破坏性后果的操作发出告警，如数据删除、权限变更、资金转账等）发出告警或进行人工确认或多级审批；