

团 体 标 准

T/CCUA 061—2025
T/BSCEA 006—2025

系统与软件工程 软件开发造价评估平台 能力要求

Systems and software engineering -Software development cost evaluation platform -
Capability requirements

2025 - 10 - 30 发布

2025 - 12 - 01 实施

中 国 计 算 机 用 户 协 会
北 京 软 件 造 价 评 估 技 术 创 新 联 盟 发布

目 次

前言 II

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 缩略语 2

5 评估平台能力框架 2

6 评估平台能力等级 2

7 L1 级能力要求 3

 7.1 核心功能能力域 3

 7.2 管理与支撑功能能力域 4

 7.3 非功能能力域 4

8 L2 级能力要求 4

 8.1 核心功能能力域 4

 8.2 管理与支撑功能能力域 5

 8.3 非功能能力域 6

9 L3 级能力要求 6

 9.1 核心功能能力域 6

 9.2 管理与支撑功能能力域 7

 9.3 非功能能力域 8

10 L4 级能力要求 8

 10.1 核心功能能力域 8

 10.2 管理与支撑功能能力域 9

 10.3 非功能能力域 10

附录 A （资料性） 评估平台能力等级各能力项主要差异 11

附录 B （资料性） 典型应用场景 14

参考文献 16

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。

本文件由中国计算机用户协会和北京软件造价评估技术创新联盟共同提出并归口。

本文件起草单位：北京软件造价评估技术创新联盟、中国计算机用户协会软件造价分会、北京中基数联科技有限公司、致同会计师事务所（特殊普通合伙）、北京明德微科科技有限公司、中国科学院软件研究所、中国进出口银行、国能信息技术有限公司、苏州市软件评测中心有限公司、中讯邮电咨询设计院有限公司、浙江天信咨询监理有限公司。

本文件主要起草人：王海青、王青、代寒玲、吴春雷、李诗豪、赵冉、李培圣、康培亮、刘丹琦、赵征、扈飞、朱拥军、刘俊、由志远、彭露瑶、吴刚。

系统与软件工程 软件开发造价评估平台 能力要求

1 范围

本文件确立了软件开发造价评估平台（以下简称“评估平台”）能力要求的框架，规定了评估平台能力等级划分及各能力域的能力项要求。

本文件适用于评估平台开发商的评估平台研发与测评；适用于第三方开展等级评估工作；也为用户选择评估平台提供参考。

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 4754—2017 国民经济行业分类
GB/T 36964—2018 软件工程 软件开发成本度量规范
GB/T 42449—2023 系统与软件工程 功能规模测量 IFPUG方法
GB/T 42588—2023 系统与软件工程 功能规模测量 NESMA方法
GB/T 45288.1—2025 人工智能 大模型 第1部分 通用要求
T/CCUA 005—2024 软件造价评估实施规程

3 术语和定义

下列术语和定义适用于本文件。

3.1

大模型 large-scale mode

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。

[GB/T 45288.1—2025, 定义3.1]

3.2

行业 industry

从事相同性质的经济活动的所有单位的集合。

[GB/T 4754—2017, 定义2.1]

3.3

基本功能组件 base functional component

功能规模测量方法中为了测量目标而定义和使用的用户功能需求的基本单位。

注：NESMAFPA方法中的基本功能组件有内部逻辑文件、外部逻辑文件、外部输入、外部输出和外部查询。通过这五种基本功能组件，功能用户需求的功能规模被测量。

[GB/T 42588—2023, 定义3.1.5]

3.4

微调 **fine-tuning**

为提升机器学习模型预测准确性,使用专门领域数据在大模型上继续训练的过程。

注1:专门领域数据一般是特定场景的生产数据或合成数据。

注2:常用的微调方法包括提示词微调、全参微调、参数高效微调等。

[GB/T 45288.1—2025, 定义3.4]

4 缩略语

下列缩略语适用于本文件。

CSBMK[®]: 中国软件行业基准数据 (China Software Benchmarking)

5 评估平台能力框架

评估平台能力框架由3个能力域和11个能力项组成(见图1)。其中:

- 核心功能能力域包括造价评估过程合规能力项、评估输入解析能力项、软件规模评估能力项、软件工作量成本评估能力项、造价评估过程管理能力项和学习进化能力项;
- 管理与支撑功能能力域包括系统集成能力项和系统管理能力项;
- 非功能能力域包括安全能力项、国产化能力项和性能能力项。

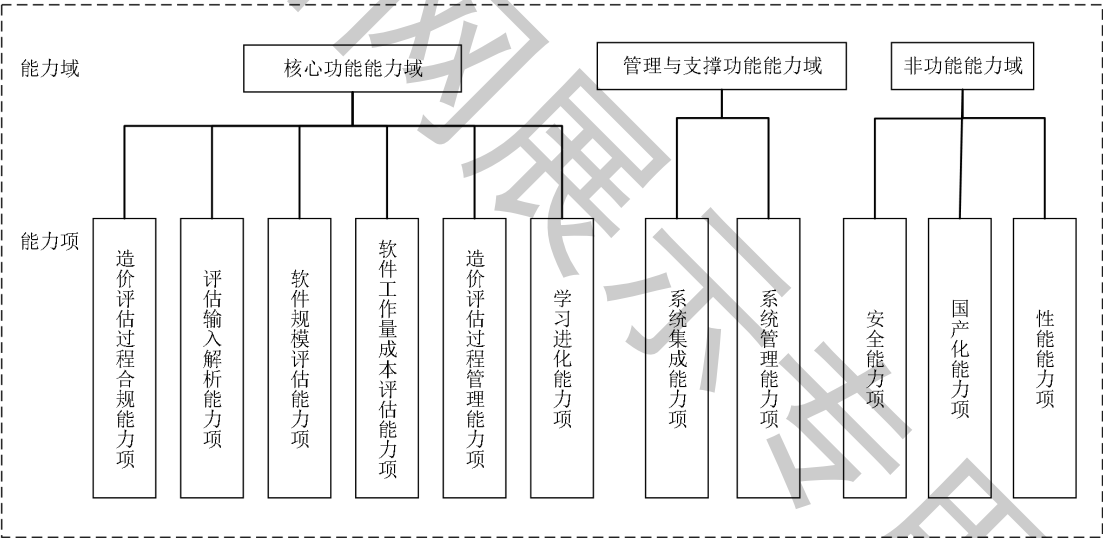


图1 评估平台能力框架

6 评估平台能力等级

评估平台能力等级由低到高为L1级至L4级。评估平台能力等级越高,智能化程度越高,评估平台综合能力越强,评估软件开发造价准确性越高。各等级说明如下:

- a) L1级评估平台基于GB/T 42588—2023和GB/T 42449—2023规定的功能规模测量方法审核特定格式评估结果,标识明显错误,并在线维护相关文档。

- b) L2级评估平台选用GB/T 42588—2023和GB/T 42449—2023规定的功能规模测量方法评估与审核大量送审项目的规模、工作量、成本，较准确识别出评估结果偏差较大的项目；同时，评估平台具备管理基准数据、评估结果、评估报告、评估流程、在线维护评估过程相关文档的能力。
- c) L3级评估平台选用GB/T 42588—2023和GB/T 42449—2023规定的功能规模测量方法较为准确地评估或审核特定行业的软件开发造价，评估结果仅需要少量人工调整；同时，评估平台应具备定制评估流程、评估报告能力，与第三方平台集成与被集成能力，私有化部署能力，安全使用大模型，运行环境国产化。
- d) L4级评估平台选用GB/T 42588—2023和GB/T 42449—2023规定的功能规模测量方法准确地评估一个以上行业软件开发造价，评估结果准确性不低于软件造价评估专家的通常水平；同时，评估平台应具备自主学习能力，定制评估流程和评估报告的能力，与第三方平台集成与被集成能力，私有化部署能力，安全使用大模型并自主可控，运行环境国产化。

评估平台各等级的典型应用场景例子见附录B。

注：L1、L2、L3、L4 的能力要求逐级提高，高等级要求覆盖低等级要求，差异性要求用粗体标注。评估平台能力等级各能力项主要差异见附录 A。

7 L1 级能力要求

7.1 核心功能能力域

7.1.1 造价评估过程合规能力项

评估平台应提供符合GB/T 36964—2018中7.1.1规定的软件开发造价评估基本流程。

7.1.2 评估输入解析能力项

评估平台应具备从特定格式评估结果中解析基本功能组件的能力。

7.1.3 软件规模评估能力项

评估平台软件规模评估能力项应具备如下能力：

- a) 基于GB/T 42588—2023和GB/T 42449—2023规定的功能规模测量方法，审核人工评估的软件规模，标识明显错误；
- b) 在线人工调整基本功能组件。

7.1.4 软件工作量成本评估能力项

评估平台软件工作量成本评估能力项应具备如下能力：

- a) 审核软件开发工作量和成本；
- b) 人工设置基准数据和调整因子。

7.1.5 造价评估过程管理能力项

评估平台应具备在线管理各审核项目的项目需求、人工评估结果、审核结果和审核报告的能力。

7.1.6 学习进化能力项

评估平台应具备构建、维护和使用软件规模审核规则库的能力。

7.2 管理与支撑功能能力域

7.2.1 系统集成能力项

评估平台应具备被第三方平台集成的能力。

7.2.2 系统管理能力项

评估平台系统管理能力项应具备如下能力：

- a) 管理用户、权限和菜单；
- b) 管理报表；
- c) 管理消息；
- d) 管理基准数据、调整因子；
- e) 自定义评估规则；
- f) 监控评估平台运行状态。

7.3 非功能能力域

7.3.1 安全能力项

评估平台应符合评估对象数据和应用场景安全分类分级的要求。

7.3.2 国产化能力项

评估平台应具备在国家认可的安全评测机构发布的国产化软硬件环境上运行的能力。

7.3.3 性能能力项

评估平台性能能力项应具备如下能力：

- a) 实时响应用户操作；
- b) 准实时审核人工评估结果。

8 L2 级能力要求

8.1 核心功能能力域

8.1.1 造价评估过程合规能力项

评估平台造价评估过程合规能力项要求如下：

- a) 应提供符合GB/T 36964—2018中7.1.1和T/CCUA 005—2024中5.1的规定的软件开发造价评估流程；
- b) 应具备根据T/CCUA 005—2024中5.7规定审核报告的能力。

8.1.2 评估输入解析能力项

评估平台应具备解析特定格式需求基本功能组件的能力。

8.1.3 软件规模评估能力项

评估平台软件规模评估能力项要求如下：

- a) 应具备基于需求和功能点分析法，审核人工评估的软件规模，标识明显错误，并给出修改建议的能力；
- b) 应选用GB/T 42588—2023和GB/T 42449—2023规定的功能规模测量方法作为软件规模估算和审核方法；
- c) 应具备基于需求和功能点分析法，评估软件规模的能力；
- d) 应确保评估结果一致性；
- e) 应具备在线人工调整基本功能组件的能力。

8.1.4 软件工作量成本评估能力项

评估平台软件工作量成本评估能力项应具备如下能力：

- a) 审核软件开发工作量和成本；
- b) 人工设置基准数据和调整因子；
- c) 评估软件开发工作量和成本。

8.1.5 造价评估过程管理能力项

评估平台造价评估过程管理能力项应具备如下能力：

- a) 在线管理各审核项目的项目需求、人工评估结果、审核结果和审核报告；
- b) 在线管理各评估项目的项目需求、评估方法、基准数据、调整因子、评估规则、评估结果和评估报告；
- c) 查询各项目流程节点；
- d) 查询各项目调用大模型信息。

8.1.6 学习进化能力项

评估平台学习进化能力项应具备如下能力：

- a) 构建、维护和使用软件规模规则库；
- b) 使用具备微调和优化能力的大模型。

8.2 管理与支撑功能能力域

8.2.1 系统集成能力项

评估平台应具备被第三方平台集成的能力。

8.2.2 系统管理能力项

评估平台系统管理能力项应具备如下能力：

- a) 管理用户、权限和菜单；
- b) 管理报表；
- c) 数据分析；
- d) 管理消息；
- e) 管理基准数据、调整因子；
- f) 自定义评估规则；
- g) 监控评估平台运行状态；
- h) 管理评估模板和评估报告模板。

T/CCUA 061-2025

T/BSCEA 006-2025

8.3 非功能能力域

8.3.1 安全能力项

评估平台安全能力项要求如下：

- a) 应符合评估对象数据和应用场景的安全分类分级的要求；
- b) 应具备安全存储大模型应用场景参数数据、训练数据的能力。

8.3.2 国产化能力项

评估平台应具备在国家认可的安全评测机构发布的国产化软硬件环境上运行的能力。

8.3.3 性能能力项

评估平台性能能力项应具备如下能力：

- a) 实时响应用户操作；
- b) 准实时审核人工评估结果；
- c) 准实时评估软件规模、工作量和成本。

9 L3 级能力要求

9.1 核心功能能力域

9.1.1 造价评估过程合规能力项

评估平台造价评估过程合规能力项要求如下：

- a) 应提供符合GB/T 36964—2018中7.1.1和T/CCUA 005—2024中5.1的规定的软件开发造价评估流程；
- b) 应具备根据T/CCUA 005—2024中5.7规定审核和自动生成报告的能力；
- c) 应具备查询软件规模、工作量和成本评估计算过程的能力。

9.1.2 评估输入解析能力项

评估平台应具备从长文本需求文件中自动抽取功能性需求，解析基本功能组件的能力。

9.1.3 软件规模评估能力项

评估平台软件规模评估能力项要求如下：

- a) 应高度适配特定行业；
- b) 应具备基于需求和功能点分析法，审核人工评估的软件规模，标识明显错误，并给出修改建议的能力；
- c) 应选用GB/T 42588—2023和GB/T 42449—2023规定的功能规模测量方法作为软件规模估算和审核方法；
- d) 应具备基于需求和功能点分析法，评估软件规模的能力；
- e) 软件规模评估准确率较高；
- f) 应确保基本功能组件命名规范性；
- g) 应确保评估结果一致性；
- h) 应具备在线人工调整基本功能组件的能力。

9.1.4 软件工作量成本评估能力项

评估平台软件工作量成本评估能力项要求如下：

- a) 应具备审核软件开发工作量和成本的能力；
- b) 应具备人工设置基准数据和调整因子的能力；
- c) 应具备评估软件开发工作量和成本的能力；
- d) 应适用于特定行业的预算、招投标、需求变更、结算和决算等场景。

9.1.5 造价评估过程管理能力项

评估平台造价评估过程管理能力项应具备如下能力：

- a) 在线管理各审核项目的项目需求、人工评估结果、审核结果和审核报告；
- b) 在线管理各评估项目的项目需求、评估方法、基准数据、调整因子、评估规则、评估结果和评估报告；
- c) 查询各项目流程节点；
- d) 查询各项目调用大模型信息；
- e) 查询历史项目评估的基本功能组件与需求对应关系；
- f) 定制软件造价评估流程。

9.1.6 学习进化能力项

评估平台学习进化能力项要求如下：

- a) 应具备根据特定行业特征构建、维护和使用软件规模规则库的能力；
- b) 应使用具备训练、微调和优化能力的大模型；
- c) 应具备查询包含准确率、损失、参数更新等信息的模型历史版本和微调迭代过程日志的能力；
- d) 应使用深度覆盖特定行业的大模型应用参数数据。

9.2 管理与支撑功能能力域

9.2.1 系统集成能力项

评估平台系统集成能力项应具备如下能力：

- a) 被第三方平台集成；
- b) 集成第三方平台。

9.2.2 系统管理能力项

评估平台系统管理能力项应具备如下能力：

- a) 管理用户、权限和菜单；
- b) 管理报表；
- c) 数据分析；
- d) 管理消息；
- e) 管理基准数据、调整因子；
- f) 管理规则库和知识库；
- g) 监控评估平台运行状态；
- h) 管理评估模板和评估报告模板；
- i) 管理系统备份；

T/CCUA 061-2025

T/BSCEA 006-2025

j) 管理大模型。

9.3 非功能能力域

9.3.1 安全能力项

评估平台安全能力项要求如下：

- a) 应符合评估对象数据和应用场景安全分类分级的安全要求；
- b) 应具备安全存储大模型应用场景参数数据、训练数据的能力；
- c) 应具备大模型训练数据隐私保护的能力；
- d) 应公开使用的大模型信息和使用用途；
- e) 应具备管理黑白名单、作弊用户和风险用户的能力；
- f) 应使用具备私有化部署能力的大模型；
- g) 应使用已备案的大模型。

9.3.2 国产化能力项

评估平台国产化能力项要求如下：

- a) 应具备在国家认可的安全评测机构发布的国产化软硬件环境上运行的能力；
- b) 使用的大多数组件应自主可控。

9.3.3 性能能力项

评估平台性能能力项应具备如下能力：

- a) 高并发情况下，实时响应用户操作；
- b) 高并发情况下，准实时审核人工评估结果；
- c) 准实时评估软件规模、工作量和成本。

10 L4 级能力要求

10.1 核心功能能力域

10.1.1 造价评估过程合规能力项

评估平台造价评估过程合规能力项要求如下：

- a) 应提供符合GB/T 36964—2018中7.1.1和T/CCUA 005—2024中5.1的规定的软件开发造价评估流程；
- b) 应具备根据T/CCUA 005—2024中5.7规定审核和自动生成报告的能力；
- c) 应具备查询软件规模、工作量和成本评估计算过程的能力。

10.1.2 评估输入解析能力项

评估平台应具备从文本、音频等不同模态组合的输入文件中自动抽取功能性需求，解析基本功能组件的能力。

10.1.3 软件规模评估能力项

评估平台软件规模评估能力项要求如下：

- a) 应高度适配一个以上行业；

- b) 应具备基于需求和功能点分析法,审核人工评估的软件规模,标识明显错误,并给出纠正建议的能力;
- c) 应选用GB/T 42588—2023和GB/T 42449—2023规定的功能规模测量标准作为软件规模估算和审核方法;
- d) 应具备整合一个以上文件为一份完整需求的能力;
- e) 应具备基于需求和功能点分析法,评估软件规模的能力;
- f) 软件规模评估准确率高;
- g) 应确保基本功能组件命名规范性;
- h) 应确保评估结果一致性;
- i) 应具备在线人工调整基本功能组件的能力。

10.1.4 软件工作量成本评估能力项

评估平台软件工作量成本评估能力项要求如下:

- a) 应具备审核软件开发工作量和成本的能力;
- b) 宜具备自动设置基准数据和调整因子的能力;
- c) 应具备评估软件开发工作量和成本的能力;
- d) 应适用于一个以上行业的预算、招投标、需求变更、结算和决算等场景。

10.1.5 造价评估过程管理能力项

评估平台造价评估过程管理能力项应具备如下能力:

- a) 在线管理各审核项目的项目需求、人工评估结果、审核结果和审核报告;
- b) 在线管理各评估项目的项目需求、评估方法、基准数据、调整因子、评估规则、评估结果和评估报告;
- c) 查询各项目流程节点;
- d) 查询各项目调用大模型信息;
- e) 查询历史各评估结果的基本功能组件与需求对应关系;
- f) 定制软件造价评估流程。

10.1.6 学习进化能力项

评估平台学习进化能力项要求如下:

- a) 应具备根据一个以上行业特征构建、维护和使用软件规模规则库的能力;
- b) 应使用具备训练、微调和优化能力的大模型;
- c) 应具备查询包含准确率、损失、参数更新等信息的模型历史版本和微调迭代过程日志的能力;
- d) 应使用深度覆盖一个以上行业的大模型应用参数数据;
- e) 应具备根据一个以上行业特征构建、维护和使用知识库和知识图谱的能力。

10.2 管理与支撑功能能力域

10.2.1 系统集成能力项

评估平台系统集成能力项应具备如下能力:

- a) 被第三方平台集成;
- b) 集成第三方平台。

10.2.2 系统管理能力项

评估平台系统管理能力项应具备如下能力：

- a) 管理用户、权限和菜单；
- b) 管理报表；
- c) 数据分析；
- d) 管理消息；
- e) 管理基准数据、调整因子；
- f) 管理规则库和知识库；
- g) 监控评估平台运行状态；
- h) 管理评估模板和评估报告模板；
- i) 管理系统备份；
- j) 管理大模型。

10.3 非功能能力域

10.3.1 安全能力项

评估平台安全能力项要求如下：

- a) 应符合评估对象数据和应用场景安全分类分级的安全要求；
- b) 应具备安全存储大模型应用场景参数数据、训练数据的能力；
- c) 应具备大模型训练数据隐私保护的能力；
- d) 应公开使用的大模型信息和使用用途；
- e) 应具备管理黑白名单、作弊用户和风险用户的能力；
- f) 应使用具备私有化部署能力的大模型；
- g) 应使用已备案的基础模型；
- h) 应具备内容安全审核机制，拒绝接受与生成不良信息的能力。

10.3.2 国产化能力项

评估平台国产化能力项要求如下：

- a) 应具备在国家认可的安全评测机构发布的国产化软硬件环境上运行的能力；
- b) 使用的所有组件、算法、大模型应自主可控。

10.3.3 性能能力项

评估平台性能能力项应具备如下能力：

- a) 高并发情况下，实时响应用户操作；
- b) 高并发情况下，准实时审核人工评估结果；
- c) 高并发情况下，准实时评估规模、工作量和成本。

附 录 A
(资料性)

评估平台能力等级各能力项主要差异

评估平台能力等级各能力项主要差异见表A.1。

表A.1 评估平台能力等级各能力项主要差异

能力域	能力项	L1级	L2级	L3级	L4级
核心功能能力域	造价评估过程合规能力项	软件开发造价评估基本流程应符合GB/T 36964—2018中7.1.1规定	1) 软件开发造价评估流程应符合GB/T 36964—2018中7.1.1和T/CCUA 005—2024中5.1的规定 2) 根据T/CCUA 005—2024中5.7规定审核报告	1) 根据T/CCUA 005—2024中5.7规定审核和自动生成报告 2) 查询软件规模、工作量和成本评估计算过程	同L3级一样
	评估输入解析能力项	从特定格式评估结果中解析基本功能组件	从经人工格式化后的需求中解析基本功能组件	从长文本需求文件中自动抽取功能性需求，解析基本功能组件	从文本、音频等不同模态组合的输入文件中自动抽取功能性需求，解析基本功能组件
	软件规模评估能力项	1) GB/T 42588—2023和GB/T 42449—2023规定的功能规模测量方法，审核人工评估的软件规模，标识明显错误 2) 在线人工调整基本功能组件	1) 基于需求和功能点分析法，审核人工评估的软件规模，标识明显错误，并给出修改建议 2) 选用GB/T 42588—2023和GB/T 42449—2023规定的功能规模测量方法作为软件规模估算和审核方法 2) 基于需求和功能点分析法，评估软件规模的能力 3) 确保评估结果一致性	1) 应高度适配特定行业 2) 软件规模评估准确率较高 3) 应确保基本功能组件命名规范性；	1) 高度适配一个以上行业 2) 整合一个以上文件为一份完整需求 3) 软件规模评估准确率高

表 A.1 评估平台能力等级各能力项主要差异（续）

能力域	能力项	L1级	L2级	L3级	L4级
核心功能能力域	软件工作量成本评估能力项	1) 审核软件开发工作量和成本 2) 人工设置基准数据和调整因子	评估软件开发工作量和成本	适用于特定行业的预算、招投标、需求变更、结算和决算等场景	1) 宜自动设置基准数据和调整因子 2) 适用于一个以上行业的预算、招投标、需求变更、结算和决算等场景
	造价评估过程管理能力项	在线管理各审核项目的项目需求、人工评估结果、审核结果和审核报告	1) 在线管理各评估项目的项目需求、评估方法、基准数据、调整因子、评估规则、评估结果和评估报告 2) 查询各项目流程节点 3) 查询各项目调用大模型信息	1) 查询历史项目评估的基本功能组件与需求对应关系 2) 定制软件造价评估流程	同L3级一样
	学习进化能力项	构建、维护和使用软件规模规则库	应使用具备微调和优化能力的大模型	1) 根据特定行业特征构建、维护和使用软件规模规则库 2) 应使用具备微调和优化能力的大模型 3) 查询包含准确率、损失、参数更新等信息的模型历史版本和微调迭代过程日志 4) 应使用深度覆盖特定行业的大模型应用参数数据	1) 根据一个以上行业特征构建、维护和使用软件规模规则库 2) 应使用深度覆盖一个以上行业的大模型应用参数数据 3) 根据一个以上行业特征构建、维护和使用知识库和知识图谱

表 A.1 评估平台能力等级各能力项主要差异（续）

能力域	能力项	L1级	L2级	L3级	L4级
管理与支撑功能域	系统集成能力项	被第三方平台集成	同L1级一样	集成第三方平台	同L3级一样
	系统管理能力项	1) 管理用户、权限和菜单 2) 管理报表 3) 管理消息 4) 管理基准数据、调整因子 5) 自定义评估规则 6) 监控评估平台运行状态	1) 数据分析 2) 管理评估模板和评估报告模板	1) 管理规则库和知识库 2) 管理系统备份 3) 管理大模型	同L3级一样
非功能功能域	安全能力项	符合评估对象数据和应用场景安全分类分级的安全要求	安全存储大模型的应用场景参数数据、训练数据	1) 大模型训练数据隐私保护 2) 公开使用的大模型信息和使用用途 3) 管理黑白名单、作弊用户和风险用户 4) 使用具备私有化部署能力的大模型 5) 使用已备案的大模型	具备内容安全审核机制，拒绝接受与生成不良信息
	国产化能力项	在国家认可的安全评测机构发布的国产化软硬件环境上运行	同L1级一样	使用的大多数组件应自主可控	使用的所有组件、算法、大模型应国内自主可控
	性能能力项	1) 实时响应用户操作 2) 准实时审核人工评估结果	准实时评估软件规模、工作量和成本	1) 高并发情况下，实时响应用户操作 2) 高并发情况下，准实时审核人工评估结果	高并发情况下，准实时评估规模、工作量和成本

注：表A.1中的L2级到L4级的每个能力项要求均包含了其前一个等级的能力项要求，此处仅列出其新增和差异部分。

附录 B (资料性) 典型应用场景

B.1 L1 级应用场景

L1级适用于功能点人工计数结果审核及评估文档在线维护场景。

某企业基于成品软件的定制化开发，项目有限。该企业有一名软件造价评估人员，负责评估定制开发软件的规模、工作量和成本，评估结果作为立项依据。为保证评估结果可用性，对评估结果进行审核。

达到L1级能力要求的评估平台适用于审核功能点人工计数结果及评估文档在线维护场景。该企业选择L1级评估平台审核人工计数的基本功能组件正确性，提示发现明显有问题的基本功能组件，并生成报告。

B.2 L2 级应用场景

L2级适用于项目审核及筛查人工评估结果场景。

某大型集团企业数字化部负责全集团各分子公司信息化大型项目立项审核，包括软件开发造价的审核。审核流程有一个软件开发造价合理性定性环节，初步筛查上报金额异常的项目。由于大数据局软件开发造价评估人员数量有限，无法在短时间内审核大量的送审项目，因此需要借助评估平台的能力筛查送审金额与需求偏离度非常大的项目。

达到L2级能力要求的评估平台能够评估并审核大量送审项目的规模、工作量、成本，并能够较准确识别出评估结果偏差较大的项目。该集团选择L2级评估平台，首先设置评估结果偏差阈值，然后人工设置行业调整因子基准值和人月单价，再通过评估平台评估每个项目的软件开发成本。比较评估结果与送审金额，筛查出超过阈值的项目，进一步重点审核。

B.3 L3 级应用场景

L3级适用于特定行业软件造价评估及审核场景。

某大型金融企业每年投入上亿元资金外包信息化项目软件开发，建设项目数量超过一百个，其中涉及关键信息基础设施系统。为了合理使用资金，该企业在招标阶段为每个项目设置最高限价。因金融行业的特殊性，不同系统有不同的规模评估要求。为了符合管理及审计要求，将评估流程嵌入到采购流程和项目管理流程中，实现评估输入材料和输出结果线上流转，每次评估过程和结果可追溯，评估平台本地化部署。该企业的软件开发造价评估人员少，用于评估的资金有限。

达到L3级能力要求的评估平台评估结果准确性较高，仅需要少量人工调整即可出具评估报告。该金融企业私有化部署L3级评估平台，并将评估流程各节点通过接口方式集成到企业的采购管理系统和项目管理系统中，完成一键登录和协同办公。在评估软件开发造价时，将需求文件输入评估平台，人工选择各调整因子和人月单价，待评估平台输出评估结果后，人工在线调整少量基本功能组件，并输出报告供后续采购流程使用。审计时，整体流程及输入输出文件可追溯。

B.4 L4 级应用场景

L4级适用于一个以上行业软件造价评估及审核场景。

某省级大数据局每年有大量维护、升级改造的项目需要结算，结算金额和规模、工作量密切相关。项目涉及医疗、交通、公安等多个行业。因各行业特征不同，规模评估规则也不同。为满足结算、决算要求，造价评估结果准确性要求不低于软件造价评估专家的通常水平。该大数据局的软件开发造价评估人员少，可用于评估的资金有限，无法聘请第三方专家，希望通过评估平台直接给出用于结算的软件开发造价，并将结算评估流程嵌入到结算流程和项目管理流程中。每次评估过程和结果可追溯，评估平台本地化部署。

达到L4级能力要求的评估平台评估结果准确性高，不低于软件造价评估专家水平。该大数据局私有化部署L4级评估平台，并将评估流程节点通过接口方式集成到企业的采购管理系统和项目审批管理系统中，完成一键登录和协同办公。当发起结算流程后，评估平台获取需求文件，根据不同行业选择不同规则，准实时输出评估结果和报告，供结算使用。

参 考 文 献

- [1] GB/T 5271.1—2000 信息技术 词汇 第1部分：基本术语
 - [2] GB/T 8566—2007 信息技术 软件生存周期过程
 - [3] GB/T 11457—2006 信息技术 软件工程术语
 - [4] GB/T 36964—2018 软件工程 软件开发成本度量规范
 - [5] GB/T 42449—2023 系统与软件工程 功能规模测量 IFPUG方法
 - [6] GB/T 42588—2023 系统与软件工程 功能规模测量 NESMA方法
 - [7] GB/T 45288.2—2025 人工智能 大模型 第2部份 评测指标与方法
 - [8] GB/T 45288.3—2025 人工智能 大模型 第3部份 服务能力成熟度评估
 - [9] T/CCUA 005—2024 软件造价评估实施规程
-