

中国特种设备安全与节能促进会

中特促函〔2026〕36号

关于开展特种设备法规标准智能知识库系统 测评工作的函

各有关单位：

为推进特种设备法规标准数字化、智能化应用能力建设，支撑特种设备安全与节能技术委员会软科学研究课题《特种设备法规标准知识体系构建与智能语义分析研究》（课题编号：AJW-2025-04）成果评价工作，中国特种设备安全与节能促进会拟组织开展特种设备法规标准智能知识库系统测评。

本次测评对象为基于该课题研发的“特种设备法规标准智能知识库”系统（特规通）。该系统面向特种设备法规标准使用场景，提供法规标准检索、原文阅读、条款定位、智能问答、关联分析和依据溯源等功能。

本次测评坚持问题导向、场景导向和依据核验导向，重点考察系统在法规标准知识组织、语义理解、现行有效性识别、条款引用准确性、依据可追溯性、专业适用边界和交互体验等方面的表现。

请贵单位结合本单位业务实际，组织熟悉特种设备法规标准应用的相关人员参与试用测评，围绕日常工作中具有代

表性的法规标准查询、理解、适用和核验问题开展测试，并按照附件要求反馈测评意见。

本次测评仅用于研究课题成果评价和系统完善分析，不作为特种设备监管、检验、评审、咨询等实际业务办理的直接依据。涉及内部管理信息、业务数据或敏感信息的内容，可作脱敏处理。

测评时间：2026年7月1日至2026年7月31日。

反馈截止时间：2026年8月6日。

反馈方式：请按附件2格式填写测评反馈表，经签字盖章后提交电子扫描件。反馈邮箱：yinlj@cpase.org.cn。

联系人：

课题组：张会朝 13810700689

中特促进会：尹立军 18600107403

附件：1. 特种设备法规标准智能知识库系统测评工作要求
2. 特种设备法规标准智能知识库系统测评反馈表



附件 1

特种设备法规标准智能知识库系统测评工作要求

一、测评目标

围绕特种设备法规标准查询、理解、适用和核验等实际业务需求，评价“特种设备法规标准智能知识库”系统（特规通）在知识覆盖、检索定位、智能问答、依据溯源和交互体验等方面的能力。

通过参评单位结合真实业务场景开展测评，形成对系统功能、适用边界、存在问题及改进方向的评价意见，为课题成果评价、系统优化完善和后续应用研究提供依据。

二、测评对象与登录方式

测评系统：“特种设备法规标准智能知识库”系统（特规通）。

登录地址：<https://co.teguitong.com/>。

操作指引：<https://co.teguitong.com/guide>。

系统账号可直接在线申请，或联系技术支持申请。技术支持联系人：王凯 18911620150。

三、参评人员建议

建议各参评单位结合自身业务特点，组织 1-3 名熟悉特种设备法规标准使用的人员参与测评，并由本单位统一汇总反馈。

参评人员可优先覆盖检验检测、设计制造、使用管理、运维服务、监督管理、无损检测、标准化管理、质量安全管

理等岗位，并结合本人日常工作中真实、典型、可复核的问题开展测评，避免仅使用过于简单、脱离业务背景或无法核验的问题。

四、测评方式

建议参评单位按照“提出业务问题—查看系统回答—核验引用依据—记录评价意见—提出完善建议”的流程进行。

对系统回答涉及法规、规章、安全技术规范、标准条款或现行有效状态的，应尽量通过原文阅读、条款定位和外部资料核对等方式进行复核。

五、建议测评场景

序号	场景类别	测评内容建议
1	文件查找	检验系统能否准确查找特定法规、规章、安全技术规范、标准或相关文件，并区分文件名称相近、版本不同的情况。
2	条款定位	检验系统能否根据业务问题定位到具体条款、章节、附录或表格，并支持原文核验。
3	条文理解	检验系统能否对专业条文作出准确、审慎、不过度扩展的解释。
4	适用条件判断	检验系统能否结合设备类别、参数、使用场景、检验环节等条件判断相关法规标准的适用范围。
5	多文件关联	检验系统能否呈现法规、规章、安全技术规范、标准之间的引用、替代、配套或上下位关系。
6	现行状态确认	检验系统能否识别文件或条款的现行、废止、替代、修订等状态，并提示核验依据。
7	依据溯源	检验系统回答是否能够列明依据来源，引用内容是否与原文一致，是否便于追溯核验。
8	交互体验	检验系统在提问、检索、结果展示、原文跳转、二次追问等方面是否便捷、清晰、稳定。

六、重点评价维度

序号	评价维度	评价要点
1	结果准确性	回答内容是否符合相关法规标准原文及专业理解，是否存在错误解释、错误适用或明显遗漏。
2	依据可追溯性	是否提供可核验的法规标准依据，引用文件、章节、条款、原文内容是否准确。
3	现行有效性	是否能够识别法规标准的现行、废止、替代、修订等状态，避免引用失效依据。
4	覆盖范围	是否覆盖特种设备主要类别、关键法规标准和常见业务场景。
5	垂直专业深度	是否能够处理具有特种设备专业背景的问题，回答是否体现行业语境和业务逻辑。
6	适用边界	是否能对不确定、需人工判断或超出系统能力的问题作出边界提示，避免过度推断。
7	操作便捷性	检索、问答、原文查看、依据核验、结果阅读等操作是否顺畅。
8	提问门槛	对非信息化专业人员是否友好，是否支持自然语言提问和逐步追问。

七、反馈材料要求

请参评单位按照附件二填写测评反馈表。原则上每家单位记录不少于 5 个具有代表性的测评问题或场景；如发现明显错误、引用不准、无法核验、现行状态错误、体验障碍等情况，请重点记录。

反馈表经填表人签字并加盖单位公章后提交电子扫描件。确需补充截图、问题清单或说明材料的，可作为附件一并提交。

附件 2

特种设备法规标准智能知识库系统测评反馈表

填表说明：请结合真实业务场景填写。评价结果仅用于课题成果评价和系统完善分析，不作为产品认证、等级评定、业务办理或商业推广依据。涉及敏感信息的，可脱敏处理。

一、基本信息

项目	填写内容	项目	填写内容
课题名称	《特种设备法规标准知识体系构建与智能语义分析研究》（课题编号：AJW-2025-04）	测评系统	特种设备法规标准智能知识库系统（特规通）
单位名称		单位类型	☐ 设计单位 ☐ 制造企业 ☐ 使用单位 ☐ 运维单位 ☐ 检验机构 ☐ 监督机构 ☐ 检测机构 ☐ 其他：-----
参评人员		职务/岗位	
专业领域	☐ 锅炉 ☐ 压力容器 ☐ 压力管道 ☐ 电梯 ☐ 起重机械 ☐ 场内车辆 ☐ 检验检测 ☐ 其他：-----	联系电话	
主要使用场景	☐ 法规查询 ☐ 标准查询 ☐ 条款定位 ☐ 条文理解 ☐ 依据核验 ☐ 多文件关联 ☐ 业务适用判断 ☐ 其他：-----	测评日期	年 月 日

二、测评问题记录

序号	问题类型	业务背景/ 设备类别	提问内容	系统回答摘要	引用依据	依据核验结果	问题严重程度	评价结论	改进建议
1	☐ 文件查找 ☐ 条款定位 ☐ 条文理解 ☐ 适用判断 ☐ 多文件关联 ☐ 现行状态 ☐ 依据溯源 ☐ 其他					☐ 准确 ☐ 基本准确 ☐ 部分不准 ☐ 无法核验 ☐ 错误	☐ 一般 ☐ 重要 ☐ 严重	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	
2	☐ 文件查找 ☐ 条款定位 ☐ 条文理解 ☐ 适用判断 ☐ 多文件关联 ☐ 现行状态 ☐ 依据溯源 ☐ 其他					☐ 准确 ☐ 基本准确 ☐ 部分不准 ☐ 无法核验 ☐ 错误	☐ 一般 ☐ 重要 ☐ 严重	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	
3	☐ 文件查找 ☐ 条款定位 ☐ 条文理解 ☐ 适用判断 ☐ 多文件关联 ☐ 现行状态 ☐ 依据溯源 ☐ 其他					☐ 准确 ☐ 基本准确 ☐ 部分不准 ☐ 无法核验 ☐ 错误	☐ 一般 ☐ 重要 ☐ 严重	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	

4	☐ 文件查找 ☐ 条款定位 ☐ 条文理解 ☐ 适用判断 ☐ 多文件关联 ☐ 现行状态 ☐ 依据溯源 ☐ 其他					☐ 准确 ☐ 基本准确 ☐ 部分不准 ☐ 无法核验 ☐ 错误	☐ 一般 ☐ 重要 ☐ 严重	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	
5	☐ 文件查找 ☐ 条款定位 ☐ 条文理解 ☐ 适用判断 ☐ 多文件关联 ☐ 现行状态 ☐ 依据溯源 ☐ 其他					☐ 准确 ☐ 基本准确 ☐ 部分不准 ☐ 无法核验 ☐ 错误	☐ 一般 ☐ 重要 ☐ 严重	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	
6	☐ 文件查找 ☐ 条款定位 ☐ 条文理解 ☐ 适用判断 ☐ 多文件关联 ☐ 现行状态 ☐ 依据溯源 ☐ 其他					☐ 准确 ☐ 基本准确 ☐ 部分不准 ☐ 无法核验 ☐ 错误	☐ 一般 ☐ 重要 ☐ 严重	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	

三、测评维度评价

序号	评价维度	评价说明	评价选项	意见建议
1	结果准确性	回答是否符合相关法规标准原文和专业理解，是否存在错误解释、错误适用或明显遗漏。	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	
2	依据可追溯性	是否列明可核验依据，引用文件、章节、条款和原文内容是否准确，是否便于跳转核验。	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	
3	现行有效性	是否能够识别文件或条款的现行、废止、替代、修订状态，是否避免引用失效依据。	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	
4	覆盖范围	是否覆盖本单位关注的设备类别、法规标准文件、常见业务问题和典型应用场景。	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	
5	垂直专业深度	对特种设备专业术语、业务环节、检验检测场景和标准逻辑的理解是否充分。	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	
6	适用边界	对需人工判断、证据不足、超出能力范围或存在不确定性的事项是否有明确提示。	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	
7	操作便捷性	检索、问答、原文阅读、条款定位、结果展示和二次追问是否便捷、清晰、稳定。	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	
8	提问门槛	自然语言提问是否友好，非信息化专业人员是否容易理解和使用。	☐ 优秀 ☐ 良好 ☐ 一般 ☐ 较差 ☐ 无法评价	

四、典型问题与错误类型记录

问题类型	是否发现	典型表现或示例	建议完善方向
引用依据不准确	☐ 是 ☐ 否		
依据缺失或无法追溯	☐ 是 ☐ 否		
现行状态识别错误	☐ 是 ☐ 否		
答非所问或理解偏差	☐ 是 ☐ 否		
过度推断或边界提示不足	☐ 是 ☐ 否		
文件覆盖不足	☐ 是 ☐ 否		
条款定位不准确	☐ 是 ☐ 否		
使用体验问题	☐ 是 ☐ 否		

五、总体专家意见

项目	填写内容
主要优点	
主要不足	
完善建议	
适用场景判断	☐ 适合用于法规标准初步查询 ☐ 适合用于条款定位辅助 ☐ 适合用于依据核验辅助 ☐ 适合用于培训学习辅助 ☐ 暂不建议用于复杂业务判断
是否建议继续扩大试用	☐ 是 ☐ 否 ☐ 需进一步完善后试用
其他意见	

填表人：

填表日期： 年 月 日

单位盖章：