

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，避免因使用过于简单、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。