技术标准知识图谱精准检索与智能问答

验证评测任务书

二零二一年四月

1. 评测背景

电网设备技术标准是指导基层安全生产、推进业务高质量发展的必要保障，是设备管理技术工作的前提条件。基层人员在开展运维、检修等作业时，须严格按照技术标准的流程规范开展作业，存在大量技术标准的检索和问答需求。公司已有上千项电网设备技术标准，数量庞大、内容复杂，主要以书籍、文档形式线下进行管理，基层人员不能及时查阅和充分运用，未充分发挥技术标准对各类作业的指导作用。亟需采用新技术手段，实现技术标准的结构化、知识化、智能化，改进基层人员对于技术标准的使用形式，提高检索效率，解决技术答案难找到，甚至找不到的问题，提升设备管理作业质效。

本次验证分为精准检索和智能问答两方面，选取变压器相关技术标准作为原始语料数据，构建基于变压器技术标准文档的检索和问答模型，验证知识图谱、自然语言处理等相关人工智能技术和工程实用水平。精准检索主要验证算法模型对于特定内容和知识点的定位准确性、完整性和效率，智能问答验证算法模型对于问题理解、答案生成及逻辑推理的能力。

根据业务特性和基层人员需求，由业务专家设计样例和验证用例，同时按照场景差异、难易程度划分等级，考察参与单位的业务问题理解能力和人工智能水平。另外，由人工智能领域专家参考业界通用评价方法，结合业务特点制定反映算法模型准确性和有效性的综合评价指标。最终基于国网公司人工智能两库一平台实现集中部署、统一验证，由评测工具调用各参与单位的算法模型接口，按照指定格式获取模型运行结果，自动化校验答案后生成评测结果。

1. 任务描述
2. **子任务一：电网主设备技术标准精准检索**

精准检索考察算法模型对设备技术标准知识点的定位准确性，需根据输入的关键词组、短语返回相关知识点和标准名称等内容并进行排序。检索内容包括但不限于定义、范围、规定、条件、要求、流程、方法、工艺。为客观、全面评价算法有效性、稳定性和先进性，根据检索词的不同，在设备技术标准精准检索赛道下设单一信息、多信息和关联信息三类检索题目，题目的详细说明如下：

1. 问题类型
   * + 1. **单一信息检索**

检索词是由单一词或单一短语表示的电网设备领域专业术语，一般是不可再分的语义单元，例如：“巡视周期”、“温升限值”、“有载分接开关”等。

* + - 1. **多信息检索**

检索词为两个及以上的词、短语的组合，且词、短语间无语义关系，例如：“变压器 噪声水平”、“额定容量 温升限值”、“油浸式变压器 检修类别”等。

* + - 1. **关联信息检索**

检索词同样为两个及以上的词、短语的组合，但词、短语存在明显的语义逻辑关系，可理解为句法结构简单的短句，其中一个词或短语为单一信息检索词，另一词或短语不做具体规定，例如：“有载分接开关的型式试验项目”、“分接选择器触头过热原因”。

1. 问题设置

问题共计100道，其中单一信息检索类问题20道，分数权重为1，多信息检索类60道，分数权重为2，关联信息检索20道，分数权重为3。

1. 评价原则
2. 针对一个检索词，算法需从参与单位提供的技术标准中定位到对应条款，并形成条款摘要，按照给定的答案格式输出。条款是文档内有标号的细分层次，由高到低、由概括到具体，如正文的1、1.1、1.1.1、1.1.1.1和附录的A、A.1、A.1.1等。一般条款含条款号、标题和内容，但部分条款也存在无标题的情况。关于“条款”的定义详见（GBT 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 第7.3条）。精准检索目的是找到最（尽量）完整反映检索词语义的条款中最（尽量）低层级条款。条款摘要指某个技术标准文档下条款中，能够反映检索词意图的那部分最精简的原文内容，可以是条款原文字中的一句话、一段话或是几句话、几段话。
3. 一个检索词可能对应一个或多个条款，算法因此也需要反馈关联度最高的几个答案，并按照顺序进行排列，形成对应的检索结果。相应的，针对每个精准检索问题，专家会提前给出一个或多个最佳参考答案。
4. 检索结果的文本编码须符合UTF-8格式，答案中的条款摘要不得出现无实际语义的转义字符。
5. 若检索目标为技术标准中的表格，仅需将表格的标题作为条款摘要进行输出，表格内容不在考察范围之列。另外，技术标准中的图片和公式也不在本次验证工作的检索范围之中。
6. 三类问题的结果形式一致，评测函数相同（详见评价细则）。
7. 评价细则

对应每个检索问题，算法给出的结果需要包含“任务执行状态”、“检索词”和“问题类型”等标识性的信息以及五项检索答案，并按照答案的正确程度或关联性自上而下进行排序。每项检索答案包含答案序号ID（1、2、3、4或5）、条款摘要（为方便后期的自动评分，条款摘要中还需额外注明条款在标准中的标号及标题（如有））、条款对应文档的标准号、条款对应文档的标准名称、条款在标准中的标号。算法最后还需在结果中标明参与单位的名称信息。

算法验证阶段，每个问题均包含一个或多个由专家给出的参考答案，评测脚本通过将参与队伍的检索结果与参考答案进行比对，并综合条款定位准确性、答案排序合理性、摘要抽取准确性三方面，评价算法对于这个问题的检索准确度。

1. **条款定位准确性**

针对一个检索题，当检索结果中的一个候选答案与专家提供的一个参考答案在标准号、标准名称和具体条款号上完全匹配时，称此候选答案为一个“命中答案”。针对任意的检索问题，假设专家给出的参考答案数量为，而检索结果中的前五个答案里有个为命中答案，则算法模型针对这一题的命中率为：

|  |  |  |
| --- | --- | --- |
|  |  | （1） |

式中α为根据问题难度设定的权重，对于个检索问题，平均命中率为：

|  |  |  |
| --- | --- | --- |
|  |  | （2） |

1. **答案排序合理性**

针对一个检索题，考察检索结果中第一个命中答案在候选答案中的排序的优先位置。针对任意的检索问题，假设算法模型的第一个命中答案在候选答案中的位置为，则这一题的排序精准率为：

|  |  |  |
| --- | --- | --- |
|  |  | （3） |

式中为根据问题难度设定的权重，如果返回的候选答案中无一命中，则。对于个检索问题，平均排序精准率为：

|  |  |  |
| --- | --- | --- |
|  |  | （4） |

1. **摘要抽取准确性**

针对一个检索题，评估每个命中答案与参考答案中的对应答案在文本内容上的相似度，以评价算法模型在这个问题上的摘要抽取准确率。

|  |  |  |
| --- | --- | --- |
|  |  | （5） |
|  |  | （6） |
|  |  | （7） |
|  |  | （8） |
|  |  | （9） |
|  |  | （10） |
|  |  | （11） |

其中为元词组，指一元词组（即单字），为二元词组（即含有两个字的字符串），为命中答案与对应参考答案中能够匹配的元词组数量，为对应的参考答案中元词组的总数量，为最长公共子序列，为命中答案的摘要字数，为对应参考答案的条款摘要总字数，指第个问题的第个答案的摘要抽取准确率，未命中的答案的设置为,式中为根据问题难度设定的权重。

1. 接口样例
2. **单信息检索接口样例**

输入：

|  |  |
| --- | --- |
|  | POST(json) / GET(parameters) /query |
| POST | {      "id": "IR01001",      "query": "专业巡检"  } |
| GET | http://ip:port?id=IR01001&query=专业巡检 |

输出:

|  |
| --- |
| {      "id": "IR01001",      "status":"OK", // ok or error      "query":"专业巡检",      "type":"单信息检索",      "results":[          {              "rank":1,              "snippet":"3.3 专业巡检 specialty inspection 为由专业技术人员对设备开展常规巡视以及借助各种先进检测仪器、装备对设备开展的带电检测与 诊断。",              "technology\_standard\_code":"Q/GDW 11085-2013",              "technology\_standard\_name":"油浸式电力变压器（电抗器）技术监督导则",              "technology\_standard\_subtitle":"3.3"          },          {              "rank":2,              "snippet":"8.6 变压器例行/诊断性试验、检修（包括专业巡检、带电检测）、改造等记录应及时录入生产管理信息系统并归档，例行/诊断性试验报告、检修报告必须严格履行审批手续。",              "technology\_standard\_code":"Q/GDW 11085-2013",              "technology\_standard\_name":"油浸式电力变压器（电抗器）技术监督导则",              "technology\_standard\_subtitle":"8.6"          },          {              "rank":3,              "snippet":"7 检修类别              表2 油浸式变压器 (电抗器)检修类别与检修内容对应表",              "technology\_standard\_code":"DL/T 1684-2017",              "technology\_standard\_name":"油浸式变压器（电抗器）状态检修导则",              "technology\_standard\_subtitle":"7"          },      ],      "company":"XXX公司"  } |

1. **多信息检索接口样例**

输入：

|  |  |
| --- | --- |
|  | POST(json) / GET(parameters) /query |
| POST | {      "id": "IR02001",      "query": "额定容量 温升限值"  } |
| GET | http://ip:port?id=IR02001&query=额定容量 温升限值 |

输出:

|  |
| --- |
| {      "id": "IR02001",      "status":"OK", // ok or error      "query":"额定容量 温升限值",      "type":"多信息检索",      "results":[          {              "rank":1,              "snippet":"6.2 额定容量下的温升限值              对于分接范围不超过士5%，且额定容量不超过2500kVA（单相833kVA）的变压器，温升限值适用于与额定电压对应的主分接(见GB 1094.1)。",              "technology\_standard\_code":"GB 1094.2-2013",              "technology\_standard\_name":"电力变压器 第2部分：液浸式变压器的温升",              "technology\_standard\_subtitle":"6.2"          },          {              "rank":2,              "snippet":"10.6.1 在额定电压、额定持续中性点电流和额定二次绕组容量下的温升              如果变压器有二次绕组，则变压器在二次绕组额定容量下不应超过温升限值。或当规定了持续中性点电流时，在二次绕组额定容量和持续中性点电流组合作用下不应超过温升限值。",              "technology\_standard\_code":"GB/T 1094.6-2011",              "technology\_standard\_name":"电力变压器 第6部分：电抗器",              "technology\_standard\_subtitle":"10.6.1"          }      ],      "company":"XXX公司"  } |

1. **关联信息检索接口样例**

输入：

|  |  |
| --- | --- |
|  | POST(json) / GET(parameters) /query |
| POST | {      "id": "IR03001",      "query": "有载分接开关的型式试验项目"  } |
| GET | http://ip:port?id=IR03001&query=有载分接开关的型式试验项目 |

输出:

|  |
| --- |
| {      "status":"OK", // ok or error      "query":"有载分接开关的型式试验项目",      "type":"关联信息检索",      "id": "IR03001",      "results":[          {              "rank":1,              "snippet":"5.2 型式试验              下列型式试验是在相关的有载分接开关的最终研制的样品上或等效的组成部件上进行的。所谓等效的组成部件是指制造单位已表明只用这些组成部件替代完整的有载分接开关后，其试验条件及试验结果不受影响。              —触头温升试验(见5.2.1);              —切换试验(见5.2.2);              —短路电流试验(见5.2.3);              —过渡阻抗试验(见5.2.4);              —机械试验(见5.2.5);              —绝缘试验(见5.2.6)。",              "technology\_standard\_code":"GB 10230.1-2007",              "technology\_standard\_name":"分接开关 第1部分：性能要求和试验方法",              "technology\_standard\_subtitle":"5.2"          }      ],      "company":"XXX公司"  } |

1. **子任务二：电网主设备技术标准智能问答**

问答任务考察算法模型对于设备管理技术标准“分析、理解及归纳”的智能化水平。根据提出的技术标准相关问题，智能问答算法模型需从文本中抽取若干段落、句子、实体或属性值作为该问题的答案。答案应来自于给定的文档，且尽量满足唯一性、正确性、完整性、简洁性和客观性等要求。主办方从验证集提取一定量技术标准问答问题，按照问答模型的输入、输出接口要求，调用参与单位设备技术标准问答模型，输出问答结果，由工具对模型结果和标准答案进行自动化对比，给出评测结果。

按照问题理解和答案组织的难易程度，智能问答赛道将设备运检过程中基层人员比较关注的问题划分为基础、统计和推理三类。基础类题目主要考察算法对于技术标准原文内容的抽取能力，如问题“散热器的油道厚度不应小于多少”和答案“散热器的油道厚度不应小于9mm”。统计类题目主要考察算法对于纲要的提取能力，如问题“变压器分为哪几类”和答案“3，[配电变压器，中型变压器，大型变压器]”。推理类题目主要考察算法对文本上下文语义的理解和逻辑判断，如问题“750kV油浸式变压器换油后静止72h是否满足要求？”和答案“否，750kV 96h”。

1. 答案类型

出于后期自动化评分的需要，根据问题答案不同的表示形式，将问题另分为：数字类、抽取类、统计类和判断类四大类别，并分别设置判分方法。

1. **数字类**

针对数字类，算法需返回准确的数字答案，表示为以下形式：

|  |
| --- |
| {“num”:“数字”，“unit”:“单位”} |

其中，“数字”是必填项，“单位”则取决于问题表达，问题中有单位出现，则可以选择不填，若问题中无单位出现，则需要填写。

1. **抽取类**

答案为某一个技术标准中一个或多个条款下原文语句、段落的组成，不存在语义的加工（与精准检索赛道的条款摘要类似，但不限定在一个条款下抽取内容）。

1. **统计类**

答案由统计数字以及各统计项的纲要信息组成，具体形式如下所示：

|  |
| --- |
| {“num”:“数量”，“argument”:“核心判据文本描述”} |

例如，问题“片式散热器按冷却方式可以分为几类？”的反馈答案应为：{"num": "3","argument": "自冷式（ONAN）,风冷式（ONAF）,强油风冷式（OFAF）"}。

1. **判断类**

答案由正误结论和核心判据组成。不仅需要给出“是”或“否”，同时必须给出相应判断的核心判据。具体接口格式定义如下：

|  |
| --- |
| {"judgement": "是（否）", "argument": "核心判据文本描述"} |

1. 问题设置

智能问答赛道题目100道，其中基础类问题20道，分数权重为1，统计类问答60道，分数权重为2，推理类问答20道，分数权重为3。

1. 评价原则
2. 针对每个问题，参与单位反馈的结果中仅能包含一个答案。通常情况下，专家组会为每个问题提供一个参考答案，用于比对评价。不过鉴于部分问题的特殊性以及语言的多样性，某些问题的参考答案可能不止一个。在计算每个问题的得分时，会将选手答案和每个参考答案进行分别计算，取最高的分数作为得分。
3. 在实际验证中，不给出答案类型的具体提示，需要参与单位模型算法自行判断并在提交结果中给出答案类型标签字段。答案类型标签判别失败，将导致该题得分为0分。
4. 检索结果的文本编码须符合UTF-8格式，答案中的文本内容中不得出现无实际语义的转义字符。
5. 抽取类答案、统计类答案的每项纲要、判断类答案的判据均应来自技术标准原文。
6. 评价细则

对应每个问答题目，算法给出的结果需要首先包含“任务执行状态”和“问题文本”等标识信息，再按照具体的结构要求给出一项答案，包含“答案类型”、“答案内容”、“答案对应文档的标准号”、“答案对应文档的标准名称”、“答案相关条款的条款号”和“答案相关条款的内容（包含条款号、名称和具体内容）”，其中“答案相关条款的内容”不在评价范围内，仅用于后期结果展示。如果答案涉及到多个条款，需要以数组形式将全部条款输出。算法最后还需在结果中标明参与单位的名称信息。

1. **答案类型的判别**

答案类型标签判别成功得1分，否则得0分并不再进一步评分。

1. **参考标准号、参考标准名称、具体条款的评价**

对于提交结果中的参考标准号、参考标准名称、具体条款号等标识信息，采用完全匹配的方式。

|  |  |  |
| --- | --- | --- |
|  |  | （12） |

其中为提交答案的标识，ref为参考答案的标识，I为指示函数，当给出答案和参考答案完全一致时，得1分，否则为0分。若答案对应多个条款，须对应多个条款的标识。

1. **具体答案的评价**
   1. 数字类

数字类答案，采用全匹配的方式。

|  |  |  |
| --- | --- | --- |
|  |  | （13） |

表示提交答案中的数字（根据问题，选择是否额外填写单位），表示参考答案中的数字（根据问题，选择是否额外填写单位），当提交的答案和标准答案完全一致得1分，否则得0分。

* 1. 抽取类

抽取类答案的评价分两部分：

首先，判断专家预选的关键词（短语、短句）中有多少能够被算法输出的答案所覆盖。假如专家给出了n个关键词，而算法输出的答案仅覆盖了其中的m个，这部分对应的得分为m/n。下面是一个具体实例，“分接选择器触头过热的原因是什么”的参考答案为：

|  |
| --- |
| “9.1.3 当分接选择器在固定分接位置上运行时的触头过热分接选择器发生触头过热的原因有：  ——分接选择器触头弹簧压力小；  ——长时间(数月)在一个位置上运行；  ——周围油的温度高。  ...” |

假设专家挑选的关键词为“弹簧压力”、“长时间”、“油”。算法模型给出的答案“分接选择器触头弹簧压力小；长时间(数月)在一个位置上运行”仅包含“弹簧压力”和“长时间”。由于缺少了关键词“油”，这部分得分为2/3。

其次，使用Rouge判断专家给出的参考答案与算法的输出答案的文本匹配程度，Rouge评分采用Rouge-1，Rouge-2，Rouge-L的综合评分，具体计算方法见检索赛道。

综上，抽取类答案由语句相似度得分和关键词得分综合计算得出，具体如下。

|  |  |  |
| --- | --- | --- |
|  |  | （14） |
|  |  | （15） |

其中arg为算法输出的答案，g\_arg为专家给出的参考答案，包括关键词和摘要答案两部分。Skwd表示关键词（短语、短句等）评分，Rouge表示文本匹配得分。

* 1. 判断类

判断类答案由两部分构成，包括回答“是”或“否”的正确性，以及核心判断依据，具体计算公式如下。

|  |  |  |
| --- | --- | --- |
|  |  | （16） |

其中表示对于“是”或“否”的判断正误，当判断正确时为1、错误时为0。g(arg, g\_arg)是对核心判据的评测，具体方法同抽取类。

* 1. 统计类

答案由统计数字以及各统计项的纲要信息组成，其中数字采用完全匹配的方法，纲要文字采用抽取类相同的方法进行评测。具体计算规则如下。

|  |  |  |
| --- | --- | --- |
|  |  | （17） |

其中表示数字的完全匹配，完全匹配得1分，不匹配得0分。

* 1. 综合评价

如前所述，最终答案由三个部分的得分组成：

1. 参考标准名称、参考标准号和具体条款号得分Score1，这一部分的得分只有0、1两个取值；
2. 具体答案得分Sq2，该部分得分取值范围为[0, 1]。

问答部分的最终得分计算如下：

|  |  |  |
| --- | --- | --- |
|  |  | （18） |
|  |  | （19） |

其中当问题为数字类时为1，其他情形为0；当答案为抽取类时为1，其他情形为0；当答案为判断类时为1，其他情形为0；当答案为统计类时为1，其他情形为0。α为根据问题设定的权重。

1. 接口样例
2. **基础类问答接口样例**

**样例1：**

输入：

|  |  |
| --- | --- |
|  | POST(json) / GET(parameters) /query |
| POST | {      "id": "QA01001",      "question":"新投运的220kV变压器，在施加电压前静止时间应不少于多少小时？",      "answer\_type":"数字类"  } |
| GET | http://ip:port?id=QA01001&anwser\_type=数字类&question=新投运的220kV变压器，在施加电压前静止时间应不少于多少小时？ |

输出：

|  |
| --- |
| {      "id": "QA01001",      "status":"OK", //ok or error      "question":"新投运的220kV变压器，在施加电压前静止时间应不少于多少小时？",      "results":      [          {              "answer\_type":"数字类",              "answer": {                  "num": 48,                  "unit": "小时",              },              "technology\_standard\_code":"GB 50150-2016",              "technology\_standard\_name":"电气装置安装工程电气设备交接试验标准",              "technology\_standard\_subtitle":["3.0.3"],              "technology\_standard\_ref":["3.0.3  油浸式变压器及电抗器的绝缘试验应在充满合格油，静置一定时间，待气泡消除后方可进行。静置时间应按制造厂规定执行，当制造厂无规定时，油浸式变压器及电抗器电压等级与充油后静置时间关系应按表 3.0.3 确定。              表 3. 0. 3  油浸式变压器及电抗器电压等级与充油后静置时间关系              ≥48"]          }      ],      "company":"XXX公司"  } |

**样例2**

输入：

|  |  |
| --- | --- |
|  | POST(json) / GET(parameters) /query |
| POST | {      "id": "QA01002",      "question":"易损件的更换周期要求是什么？",      "answer\_type":"抽取类"  } |
| GET | http://ip:port?id=QA01002&answer\_type=抽取类&question=易损件的更换周期要求是什么？ |

输出：

|  |
| --- |
| {      "id": "QA01002",      "status":"OK", //ok or error      "question":"易损件的更换周期要求是什么？",      "results":      [          {              "answer\_type":"抽取类",              "answer":"使用10年以上的油泵轴承或冷却风扇轴承运行中发岀不正常的噪声时,在变压器或并联电抗器退出 运行时予以更换：使用15年以上时，根据具体情况更换所有密封垫。",              "technology\_standard\_code":"DL/T 1176-2012",              "technology\_standard\_name":"1000kV油浸式变压器、并联电抗器运行及维护规程",              "technology\_standard\_subtitle":["5.16"],              "technology\_standard\_ref":["5.16易损件的更换周期              使用10年以上的油泵轴承或冷却风扇轴承运行中发岀不正常的噪声时,在变压器或并联电抗器退出运行时予以更换；使用15年以上时，根据具体情况更换所有密封垫。"]          }      ],      "company":"XXX公司"  } |

1. **统计类问答接口样例**

输入:

|  |  |
| --- | --- |
|  | POST(json) / GET(parameters) /query |
| POST | {      "id": "QA02001",      "question":"片式散热器按冷却方式可以分为几类？",      "answer\_type":"统计类"  } |
| GET | http://ip:port?id=QA02001&answer\_type=统计类&question=片式散热器按冷却方式可以分为几类？ |

输出:

|  |
| --- |
| {      "id": "QA02001",      "status":"OK", //ok or error      "question":"片式散热器按冷却方式可以分为几类？",      "results":[          {              "answer\_type":"统计类",              "answer":{                  "num": "3",                  "argument": "自冷式（ONAN）,风冷式（ONAF）,强油风冷式（OFAF）"},  // 答案为一个数组，数组的第一个元素为表示数量的数字，第二个元素为一个数组，其中枚举全部的结果。              "technology\_standard\_code":"JB/T 5347-2013",              "technology\_standard\_name":"变压器用片式散热器",              "technology\_standard\_subtitle":["4.1.2"],              "technology\_standard\_ref":["4.1.2 按冷却方式分为：              a)自冷式（ONAN）；              b)风冷式（ONAF）；              c)强油风冷式（OFAF）。"]          }      ],      "company":"XXX公司"  } |

1. **推理类问答接口样例**

**样例1**

输入:

|  |  |
| --- | --- |
|  | POST(json) / GET(parameters) /query |
| POST | {      "id": "QA03001",      "question":"750kV油浸式变压器换油后静止72h是否满足要求？",      "answer\_type":"判断类"  } |
| GET | http://ip:port?id=QA03001&answer\_type=判断类&question=750kV油浸式变压器换油后静止72h是否满足要求？ |

输出:

|  |
| --- |
| {      "id": "QA03001",      "status":"OK", //ok or error      "question":"750kV油浸式变压器换油后静止72h是否满足要求？",      "results":[          {              "answer\_type":"判断类",              "answer": {                  "judgement": "否",                  "argument": "96h"                  },  // 答案为一个数组，第一个元素表示正误的判断，只能回答“是”或“否”，第二个元素为判断依据              "technology\_standard\_code":"DL/T 572-2010",              "technology\_standard\_name":"电力变压器运行规程",              "technology\_standard\_subtitle":["5.2.6"],              "technology\_standard\_ref":["新装、大修、事故检修或换油后的变压器，在施加电压前静止时间不应少于以下规定：              a）110kV 24h              b）220kV 48h              c）500（330）kV 72h              d）750kV 96h"]          }      ],      "company":"XXX公司"  } |

**样例2**

输入:

|  |  |
| --- | --- |
|  | POST(json) / GET(parameters) /query |
| POST | {      "id": "QA03002",      "question":"容量为1600kVA及以下油浸式电力变压器交接试验时开展哪些试验项目？",      "answer\_type":"抽取类"  } |
| GET | http://ip:port?id=QA03002&answer\_type=抽取类question=容量为1600kVA及以下油浸式电力变压器交接试验时开展哪些试验项目？ |

输出:

|  |
| --- |
| {      "id": "QA03002",      "status":"OK", //ok or error      "question":"容量为1600kVA及以下油浸式电力变压器交接试验时开展哪些试验项目？",      "results":[          {              "answer\_type":"抽取类",              "answer": "1 绝缘油试验或SF6 气体试验;2 测量绕组连同套管的直流电阻;3 检查所有分接的电压比;4 检查变压器的二相接线组别和单相变压器引出线的极性;5 测量铁心及夹件的绝缘电阻;6 非纯瓷套管的试验;7 有载调压切换装置的检查和试验;8 测量绕组连同套管的绝缘电阻、吸收比或极化指数;9 绕组连同套管的交流耐压试验;10 额定电压下的冲击合闸试验;11 检查相位;",              "technology\_standard\_code":"GB 50150-2016",              "technology\_standard\_name":"电气装置安装工程电气设备交接试验标准",              "technology\_standard\_subtitle":["8.0.1","8.0.2"],              "technology\_standard\_ref":["8.0.1 电力变压器的试验项目，应包括下列内容:              1 绝缘油试验或SF6 气体试验;              2 测量绕组连同套管的直流电阻;              3 检查所有分接的电压比;              4 检查变压器的二相接线组别和单相变压器引出线的极性;              5 测量铁心及夹件的绝缘电阻;              6 非纯瓷套管的试验;              7 有载调压切换装置的检查和试验;              8 测量绕组连同套管的绝缘电阻、吸收比或极化指数;              9 测量绕组连同套管的介质损耗因数( tanO')与电容量;              10 变压器绕组变形试验;              11 绕组连同套管的交流耐压试验;              12 绕组连同套管的长时感应耐压试验带局部放电测量;              13 额定电压下的冲击合闸试验;              14 检查相位;              15 测量噪音。              ","8. 0. 2 各类变压器试验项目应符合下列规定:              1 容量为1600kVA 及以下油浸式电力变压器，可按本标准第8. 0. 1 条第1、2、3、4、5、6、 7、8、11、13 和14 款进行试验;"]          }      ],      "company":"XXX公司"  } |

1. 任务提交

本次评测任务将采取容器化部署方式。在验证阶段，允许参与单位在正式评测前预先进行模型部署及相关调试工作，组织方指导并配合参与单位完成模型部署和接口测试。正式验证阶段，组织方统一向各参与单位发布精准检索和智能问答的输入web请求，并在国网人工智能云平台指定目录下输出验证结果文件。

接口命名参考电网主设备技术标准精准检索和智能问答子任务“接口格式”说明。格式与任务描述中的示例输出相同，验证结果输出和队伍排名会在验证可视化大屏实时显示。

最终提交文件时，参与队伍需提交的材料如下：

1.精准检索和智能问答任务验证结果接口，详见各子任务的“接口样例”的格式要求。

2.参与单位需准备方法描述文档。

3.参与单位在评测开始阶段展示讲解构建的知识图谱规模和设计思路，评测结束后，将邀请成绩优秀的参赛队伍参与路演。

1. 部署方式

验证工作将依托国家电网有限公司电力系统人工智能联合实验室提供容器化部署环境。依据不同评测任务，代码及模型需分别打包为容器文件（建议tar格式），用“厂商.子任务.版本.tar”命名，要求提交所有的程序代码及相关模型服务的配置说明。

容器化部署详见附件1《电力人工智能实验及公共服务平台》。

1. 验证环境软硬件配置信息

评测阶段统一使用国网公司人工智能两库一平台作为部署环境进行算法模型的验证，主办方为参与单位配置同等计算资源，信息如下：

操作系统版本：Ubuntu 18.04.4 LTS

系统内存：64GB

显卡：NVIDIA QUADRO RTX 6000（2张）

容器外挂空间：1TB

NVIDIA Driver Version: 450.57

Docker Version: 19.03.6

NVIDIA Docker: 2.4.0

CPU MHz：2775.550

1. 时间安排

4月10日——5月23日，各单位依据评测任务书要求开展模型构建和训练。

5月24日——5月31日，参与单位分批次开展现场预调试。

6月1日，参与单位入场，开放技术验证平台，各参与单位分批次进行算法模型现场部署。

6月2日——6月3日，依据验证评分规则，对各参与单位算法进行技术验证。每家参与单位测试完毕后，给出评测分数，待全部参与单位验证结束后，公布验证结果。

6月7日，择优选择部分参与单位进行路演，组织技术交流活动。

1. 评测规则

以下通用规则适用于本评测当中的所有子任务。在此基础上，参与单位还需遵循具体任务的特有规则（参见章节2）。如有冲突，以后者为准。

1. 参与单位需要提交“参与单位名称，联系人（姓名，邮箱，联系电话），单位名称”等信息。
2. 参与单位可以选择只参加某个（某些）子任务，也可以选择全部参加，以在比赛平台上的结果提交为准。每个子任务单独评价。
3. 每支队伍需指定一名联系人。
4. 每名选手只能参加一支队伍，一旦发现某选手参加多支队伍，将取消所有相关队伍的验证资格。
5. 允许使用开源代码或工具，但不允许使用任何未公开发布或需要授权的代码或工具。
6. 参与单位最终需要提交可运行的代码和方法描述文档，**若在排行榜上的结果无法复现，将取消验证资格**。
7. 联系方式

联研院计算所：

陈其鹏，15210253513

中国电科院高压所：

张博文，18101241735

联系邮箱：

tianbao-wu@sgcc.com.cn（内网）

jsyz@geiri.sgcc.com.cn（外网）