

《基于大模型的电力企业向量知识库及增强检索应用技术框架》编制

说明

(征求意见稿)

一、工作简况

1 主要工作过程

调研阶段:2024年5月初开始,国网信息通信产业集团有限公司牵头各单位成立标准编写组,讨论确定了标准的主要内容及分工,同时进行调研分析,收集资料,准备立项审查答辩;

标准立项阶段:2024年5月底,召开了第一次标准的专家立项评审会,经中国电工技术学会标准工作委员会能源智慧化工作组专家组审议,批准《基于大模型的电力企业向量知识库及增强检索应用技术框架》标准立项:

编写研制阶段:2024年6月标准编写组根据立项专家组意见和建议,标准编写组进行标准编写研制,形成了征求意见稿。

2 主要参加单位和起草工作组成员及其所做的工作

标准编写组收集了行业内基于大模型的电力企业向量知识库及增强检索应用技术框架方面的相关资料,通过对比整理分析确定了标准主要技术内容,由国网信息通信产业集团有限公司牵头完成标准初稿编制,其他参与单位配合并负责收集相关资料、提出建议。

主要参与单位有:国网信息通信产业集团有限公司、福建亿榕信息技术有限公司。

二、标准编制原则和主要内容

1、标准编制原则

本标准按照 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,遵循科学性、先进性、经济性:坚持实事求是,以基于大模型的电力企业向量知识库及增强检索应用技术框架为基础,遵守国家有关法律、法规,符合团体标准要求,目的在于指导企业如何利用大规模预训练模型来构建高效、精准的向量知识库,并通过先进的检索技术实现知识的快速定

位和应用，从而提升企业知识管理的智能化水平，实现对知识的深度挖掘和高效利用，为企业的决策支持、业务创新和服务提升提供有力支撑。

在标准编制过程中，主要依据《GB/T 41867—2022人工智能术语》、《GB/T 42755—2023 人工智能面向机器学习的数据标注规程》、《GB/T 43782—2024 人工智能 机器学习系统技术要求》、《GB/T 5271.28—2001 信息技术 词汇 第 28 部分：人工智能 基本概念与专家系统》、《GB/T 5271.34—2006 信息技术. 词汇. 第 34 部分：人工智能. 神经网络》、《T/CES 156—2022 电力智能交互文本训练语料标注规范》、《T/CES 129—2022 电力人工智能平台样本规范》标准。

2、标准主要内容

本标准分为 6 个章节，（1）范围；（2）规范性引用文件；（3）术语和定义；（4）符号、代号和缩略语；（5）企业向量知识库构建；（6）基于大模型的企业向量知识库增强检索应用。主要技术内容如下：

企业向量知识库构建：包括企业向量知识库技术框架规则、知识数据归集与预处理规则、向量模型训练规则、知识向量化规则、向量知识库构建规则、向量知识维护和更新规则。

基于大模型的企业向量知识库增强检索应用：包括基于大模型的企业向量知识库增强检索应用技术框架规则、检索意图识别规则、知识检索与整合规则、内容生成与优化规则、安全与隐私保护规则。

3、主要技术差异

本标准为新制度标准，无主要技术差异。

4、解决的主要问题

本标准主要解决传统电力知识应用缺少标准化知识表示与存储与缺少足够的语义理解能力两个主要问题：

（1）缺少标准化知识表示与存储：传统的知识库往往采用基于文本或关键词的存储方式，这种方式难以捕捉知识的深层语义信息。企业向量知识库构建技术框架利用向量表示技术，将知识转化为高维向量空间中的点，能够更准确地表示知识的语义信息，并支持基于向量的高效存储和检索。

（2）缺少足够的语义理解能力：传统的检索方式往往只能基于关键词进行匹配，难以捕捉查询的深层语义信息。基于大模型的增强检索应用框架利用大规

模预训练模型的强大语义理解能力，能够更准确地理解用户查询的意图，提高检索的准确性和相关性。

三、主要试验（或验证）情况

本标准不涉及试验（或研制）情况。

四、标准中涉及专利的情况

本标准不涉及专利问题。

五、预期达到的社会效益、对产业发展的作用等情况

- (1) 构建高效、精准的向量知识库。
- (2) 实现知识的快速定位和应用。
- (3) 提升企业知识管理的智能化水平，为企业的决策支持、业务创新和服务提升提供有力支撑。

六、与国际、国外对比情况

国际暂无相关参照标准。国内已有《GB/T 41867—2022 人工智能术语》、《GB/T 42755—2023 人工智能面向机器学习的数据标注规程》、《GB/T 43782—2024 人工智能 机器学习系统技术要求》、《GB/T 5271.28—2001 信息技术 词汇 第 28 部分：人工智能 基本概念与专家系统》、《GB/T 5271.34—2006 信息技术 词汇 第 34 部分：人工智能 神经网络》，可为标准制定提供借鉴与参考。目前，国内缺少面向电力企业向量知识库的检索应用技术框架标准。本标准与现行的相关法律、法规、规章与相关标准保持一致。

七、在标准体系中的位置，与现行相关法律、法规、规章及相关标准，特别是强制性标准的协调性

本标准与现行的相关法律、法规、规章与相关标准保持一致。

八、重大分歧意见的处理经过和依据

标准编制过程中广泛征集了专家意见，所有意见均按照标准编制程序进行了采纳，不存在重大分歧意见。

九、标准性质的建议说明

建议本标准的性质为团体标准。

十、贯彻标准的要求和措施建议

建议本标准批准发布 2 天后实施。

十一、废止现行相关标准的建议

无

十二、其他应予说明的事项

无