基于知识管理系统的数字员工业务辅助平台构建与实践——以核电生产文件管理为例

陈莹，胡丹丹，王少杰
福建福清核电有限公司，福建 福州 350300

DOI: 10.61369/TACS.2026080040

摘　　要： 大模型技术正在加速进入各个行业，但核电生产文件管理这一领域的智能化应用尚处于起步阶段。本文基于核电厂自有的知识管理系统，面向核电生产文件管理的实际场景，设计并实现了一套数字员工业务辅助系统。在已有知识管理平台之上，系统拓展了文档结构化解析、领域大模型微调、知识图谱构建、智能推荐与关联、智能写作与纠错、多人在线协同编辑等功能。在协同效率提升、编制经验沉淀、文件自动关联等方面均取得了可衡量的效果，核电生产文件的编制效率和规范化程度有了明显进步。文中所呈现的技术路线与实践经验，可为同类核电企业的文档智能化转型提供参考。

关 键 词： 数字员工；知识管理系统；生产文件管理；大模型；核电

Construction and Implementation of a Digital Employee Business Support Platform Based on Knowledge Management Systems: A Case Study of Nuclear Power Production Document Management

Chen Ying, Hu Dandan, Wang Shaojie
Fujian Fuqing Nuclear Power Co., Ltd., Fuzhou, Fujian 350300

Abstract: While AI large language model technologies are rapidly penetrating various industries, intelligent applications in nuclear power production document management remain limited. Taking the knowledge management system as the platform foundation, this paper constructs a digital employee business assistance system for production document management. The system extends AI functionalities on the existing knowledge management platform, integrating document structuring, domain-specific LLM fine-tuning, knowledge graph construction, intelligent recommendation and association, intelligent writing and error correction, and multi-user collaborative editing. Practical results show that the platform effectively addresses pain points including low collaboration efficiency, experience-dependent authoring, lack of automatic association, and heavy manual proofreading. This research provides a replicable pathway for intelligent document management transformation in nuclear power enterprises.

Keywords: digital employee; knowledge management system; production document management; large language model; nuclear power

一、研究背景

在数字化转型背景下，以大模型为代表的新兴技术正在逐渐以“数字员工”的形式参与到各个行业的一线业务当中。对于企业来说，采用数字员工参与业务工作，能够加快数字转型，实现降本增效，开启智慧流程，洞察企业痛点，增加员工体验，减少重复劳动，提升工作体验。

核电工程通常周期长达几十年，从核电站前期准备、选址、建设、运行到退役的全生命过程[1]，产生了大量技术文件，是保障电厂人员、设备及机组安全的首道防线，是技术积累和知识沉淀与传承的重要载体。虽然大部分核电厂实现了生产文件纳入数字化流程管理，但在实际业务中，大量跨信息系统的业务流程、存在线上线下两条线重复操作的场景。如下问题逐渐显现出来：

（1）协同机制缺失：文件编制环节仍以单机离线操作为主，跨岗位协作效率偏低，版本难以管控。

（2）知识管理滞后：新文件编写依赖个人经验，缺少智能化的检索、引用溯源和内容规范手段。

（3）智能检索缺位：现有功能仅支持关键词检索，无法实现语义层面的搜索和问答。

（4）动态关联不足：知识的复用基本靠人工记忆来维系，系统无法主动推送关联文件。

综上所述，已有的信息化设施在智能支撑上已难以匹配核电生产业务对高效、规范和安全的诉求[2]。核电厂亟需引入数字员工技术，将大模型、知识图谱、自然语言处理等技术能力[3]与文档管理业务做深层次融合，搭建覆盖生产文件收集、编制、校验、管理、关联的全流程智能辅助体系的数字员工业务辅助平台。通过该平台的建设与运行，实现文件编制协同化、内容辅助智能化、 版本管控规范化，有效提质增效、降低重复劳动与作业风险，为企业数字化智慧运营赋能，推动生产文件管理从传统的人工模式向智能化、协同化、规范化的方向演进，为核电安全生产提供技术保障[4]。

 Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.

二、构建思路

（一）国内外技术相关研究

国外技术整体呈现起步早、体系全、与业务深度融合、大模型应用成熟等特点，在知识沉淀、智能检索、文档辅助生成等方面具备先发优势，但多以通用场景为主，在中文语境、国内核电规程标准、内网私有化部署、等保合规等方面适配性不足，难以直接满足国内生产文件编制协同、版本联动、智能关联等定制化需求。

国内随着大模型技术普及，百度文心、科大讯飞星火、腾讯混元等在中文语义理解、长文档处理、内容纠错与知识问答上能力持续提升，为生产文件智能辅助提供坚实底座。但仍存在短板：多数系统侧重通用协同或单一检索，缺少面向具体业务全流程的一体化智能辅助，核心业务深度融合的方案较为缺乏。

（二）总体架构

该系统主要包含三个方面：一是对文件的基础处理，包含多源数据收集与预处理，二是文档管理，包含文件内容分析、编制模板管理、提供数据管理能力等；三是文件在线协同编制的业务辅助能力，包括文件协同编制、智能编写、自动关联、智能纠错等应用。

系统在架构上采用分层模块化设计，自下而上贯通基础设施、数据治理、模型服务、业务中台及应用门户五大层级，以模块化设计支撑高可用、可扩展、智能化与安全可控的运行基座。

三、具体实施路径

（一）文档基础处理

数据预处理是平台的基础，将各类异构、非结构化生产文档转化为大模型可精准识别的高质量结构化知识单元，为上层智能问答、检索、推荐等应用提供可靠的数据底座。

经过工具识别得到的文本会送入基于 LayoutLMv3架构的版面分析模型，借助视觉和语义两路信息的融合来还原文档原本的逻辑结构[6]。数据处理还需依据文件类型或用户需求灵活实现元数据抽取，如编制部门、分发范围、核心观点这些需要语义理解的内容，则交由基于领域微调的中文预训练模型，结合 BERT 来做实体抽取[5]。另外，生产文件附录模板较多，也可对大量文件的附录模板进行提取。

整条处理链路由统一的调度引擎做异步编排，最终输出标准化数据包。



（二）文档管理

根据核电厂业务流程及文档管理双生命周期为指引，结合现场业务需求和特点，构建文件库和模板库。文件库支持构建多分类知识库，实现各类文件的有序存储与精细化管理。模板库自动提取管理程序附件中的标准表单，生成标准化模板并集中存储，覆盖多个核电业务场景，实现模板版本统一维护、快速调用与高效复用。

另外，为保证系统数据安全可靠，还需配以具备数据管理及统计能力的功能模块，使该平台数据为上层业务功能提供数据治理支撑，也保障数据使用可管、可控、可追溯。

（三）文档智能协调编辑

模块的核心建设目标为构建多用户协同编辑能力，针对性解决核电生产文件跨部门协作中存在的效率低下、版本管控困难、修改留痕不足等痛点问题，支撑团队知识共创过程高效、安全、合规。主要分为协同编辑、智能写作、智能关联和智能校核4项功能。

1. 协同编辑

协同编辑模块基于与 WPS 的深度集成构建在线协作框架，实现实时与异步协同编辑及规范化文档管理。系统支持多用户实时或异步编辑同一文档的多模式协同编辑，无需依赖邮件传递与人工合稿，实现跨部门协作的无缝衔接；同时，通过分级权限控制、全链路版本管理与智能冲突解决技术设置合规性安全管控机制，确保所有文档修改操作可追溯、历史版本可回退。

通过上述能力，实现了核电文件编制模式从传统线下传递向“线上协同、集中管控”的升级。

2. 智能写作与纠错

智能写作模块利用大语言模型[7]的语义理解能力，解析用户输入的标题或主题并结合领域语料特征自动生成章节大纲；同时支持基于素材库的检索增强生成（RAG）[3]，当检测到特定标引符号时，自动匹配知识库中的文件片段、标准术语及典型段落，并依据选定模板生成规程文档。

3. 智能关联

采用语义关联分析与协同过滤相结合的混合推荐策略。如基于同岗位的文档推荐、基于部门维度的协同兴趣推荐、基于部门维度的协同兴趣推荐在首页推荐场景下。在搜索结果推荐的实际应用场景下，系统根据用户输入的搜索词检索与当前查询文件相关的参考文件、引用文件、上下游文件及同类主题文件并予以呈现。

智能关联模块旨在构建动态的核电领域知识网络，其核心方法涵盖多维度关系挖掘、图谱构建与可视化展示。

4. 智能校核

构建适配核电专业文档体系的智能校核能力，解决规程、技术文件编制过程中文字疏漏、术语不规范等问题，强化文档质量的事前管控。采用“通用纠错为基础，领域知识为增强”的双层架构。基础纠错层校验及预训练模型上下文语义判别技术。领域增强层基于半自动化流程构建核电领域术语库，并针对术语使用不一致或未定义缩写提供规范性提示。

四、实施成效

（一）管理效益

一是生产文件的编制效率呈现多维度提升。编制人无需在多

Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.

个系统间反复切换即可完成从大纲到成稿的全流程编写；支持多人在线实时同步编辑，跨岗位、跨部门协同编制效率大幅提升；智能纠错模块使人工校对工作明显减轻；智能推荐模块实现了从被动到主动推送的模式转变。二是文件质量管控取得了系统性改善。智能关联模块协助编制人在浏览任文档时，可逐层探索关联文档内部信息，降低遗漏引用或引用过时文件的风险。智能推荐与关联功能通过主动推送关联文件和上下游文件，盘活了沉淀已久的历史文件资源，知识复用的门槛降低。

（二）经济效益

智能化辅助与在线协同覆盖文件处理全流程，人工投入明显缩减。智能推荐与依据关联减少对人工查阅比对的依赖，智能纠错将人工事后抽查转变为实时自动校验，审核返工率下降。统一模板与规范格式管控减少因文件不规范带来的整改成本。项目对接现有应用系统集成落地，盘活既有信息化资产，投入成本可控，效益长效可观。

（三）社会效益

该平台的构建推动核电文件管理从人工单机、分散作业，向智能化、在线化、协同化全面升级，有效提升企业数字化运营与知识管理水平。通过规范文件编制标准、精准关联技术依据、严控内容合规与版本时效，全面提升生产文件质量，为机组安全稳定、规范运维提供坚实保障。项目形成安全合规、可复制、可推广的核电文档智能化落地方案，为行业智能建设提供示范路径与实践经验，为核电安全生产和规范执行提供了技术保障。

五、未来思考

（一）技术局限性

大模型进入核电垂直领域后，幻觉现象仍然是最核心的挑战之一。核电场景对安全性的要求又特别苛刻，这意味着人工审核校验机制必须更加严格。核电专业术语体系本身庞大且复杂，领域大模型的微调高度依赖大批量高质量的标注数据，标注工作的成本不容低估。底层大模型的推理[8]吞吐量直接约束了系统的并发处理能力，多用户同时访问时偶尔会出现响应延迟加大的情况。

（二）未来方向

从已有的实践积累出发，后续还需在以下方面继续深耕：

一是多模态大模型的深度应用，图可把纸识别、设备故障诊断这类视觉理解能力也纳入业务辅助的范畴。如通过对不同训练方法如高校参数微调、迁移学习、提示学习、强化学习等来提高模型的范化能力和表现能力。[9]

二是隐私计算技术的探索，目标是在不牺牲数据安全的前提下实现跨基地的知识共享和模型协同训练。

三是需要着手建立核电领域大模型应用的标准规范和评估体系，为行业的规模化推广打好技术标准的基础。

六、结论

国内核电企业普遍将数字化转型作为企业战略转型升级的重要方向。[10]本文从核电生产文件管理与编制的实际业务痛点切入，在已有知识管理系统之上进行 AI 应用功能的拓展研发，建设成了一套集数据预处理与训练、智能推荐、智能关联、智能写作、协同编辑、智能纠错、文档管理于一体的数字员工业务辅助平台，以满足生产活动需要，形成了一套与核电生产场景相匹配的完整技术方案。项目成果贴近核电一线实际需求，技术路径成熟可靠且可复制，在行业内的先进性和引领性也较为突出，为核电企业推进生产文件智能化管理提供了一条可落地的实践路径。

参考文献

[1] 韩季红 . 核电文档管理的地位与特点研究 [J]. 北京档案 ,2010（5）：33-34

[2]Lewis P, Perez E, Piktus A, et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks[C]. NeurIPS, 2020: 9459-9474.

[3]Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]. NeurIPS, 2017: 5998-6008.

[4]Brown T B, Mann B, Ryder N, et al. Language models are few-shot learners[C]. NeurIPS, 2020: 1877-1901.

[5]Devlin J, Chang M W, Lee K, et al. BERT: Pre-training of Deep Bidirectional Transformers[C]. NAACL-HLT, 2019: 4171-4186.

[6]Huang Y, Lv T, Cui L, et al. LayoutLMv3: Pre-training for Document AI with Unified Text and Image Masking[C]. ACM Multimedia, 2022: 4083-4091.

[7]Touvron H, Martin L, Stone K, et al. Llama 2: Open Foundation and Fine-Tuned Chat Models[EB/OL]. arXiv:2307.09288, 2023.

[8]Wei J, Wang X, Schuurmans D, et al. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models[C]. NeurIPS, 2022.

[9] 文森，钱力，胡懋也，等 . 基于大语言模型的问答技术研究进展综述 [J/OL]. 数据分析与知识发现 ,1-17.[2023-11-30].http://kns.cnki.net/kcms/detail/10.1478.G2.20231110.1612.002.html.

[10] 徐建军，王浩，张翼 . 核电厂全生命周期数字化转型研究与设计 [J]. 仪器仪表用户 ,2022（8）:63-68.

 Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.