

“蜂鸟”数字化业务文档快速搜录系统的设计与实现

李源灏, 王晓宇, 张霖, 施加轮, 俞敏杰

(国网福建省电力有限公司信息通信分公司, 福建 福州 350000)

摘要 电力公司项目不断增多, 其工程项目管理的复杂程度也在成倍地增加, 而当前的电力公司对工程项目管理的质量要求颇高, 如何有效地提高电力公司工程项目管理的质量与效率成为当前的关键。因此, 电力公司的项目管理者迫切需要一套完善的项目文档管理系统来支持项目的工作。本文基于企业文档管理的数字化需求, 研究并实现了一种工具化的智能文档管理系统——“蜂鸟”数字化业务文档快速搜录系统。该系统支持文档快速搜录与表单关键字段对比合并, 采用分权限管理、数据加密和操作留痕等技术保障数据安全。通过开发友好的用户界面与便捷的导入导出功能, 项目成果显著提高了文档管理效率, 并为个人与小型团队提供了便捷的文档管理解决方案。

关键词 文档管理; 数字化需求; 办公自动化

基金项目: 国网福建省电力有限公司信息通信分公司“蜂鸟”数字化业务文档快速搜录应用研究项目资助(项目编号: 52130M24000N)。

中图分类号: TP391.3

文献标志码: A

DOI: 10.3969/j.issn.2097-3365.2025.06.003

0 引言

在当今高度信息化和计算机技术的迅猛发展时代, 有效管理庞大而复杂的信息对于各行各业的重要性应该被人们深刻地认识到。信息技术的广泛应用已经成为提升工作效率、降低人力成本的普遍且实用的手段。同时, 借助现代计算机强大的计算能力和全面的网络覆盖, 信息的处理和传递变得更加便捷高效^[1-2]。随着电力行业企业数字化转型加速, 文档管理需求从传统存储逐步转向智能化管理。特别是在小型企业和个人工作中, 数据搜录、快速查询、信息保护面临诸多挑战。“蜂鸟”数字化业务文档快速搜录系统以此为背景, 旨在实现一套适用于个人和小型团队的文档管理解决方案。同时, 为了环保和资源节约, 电力行业应逐步减少对纸质卷宗的依赖, 实现信息管理的绿色化^[3]。

1 研究概述

1.1 研究意义

在当今信息化迅速发展的时代, 大型企业无不面临着数据的管理与利用挑战。针对这一问题, “蜂鸟”数字化业务文档快速搜录系统的设计与实现, 展现了其在提升企业运营效率、优化信息管理及增强决策支持等方面的显著应用价值。该系统通过先进的文档识别与处理技术, 能够快速对各类业务文档进行数字化转化, 极大地减少了人工录入的时间和错误率。原有

的文档管理方式主要是基于资料员以及其他人员所记录的纸质档案, 不仅信息检索效率低下, 且容易造成信息的遗失与混乱。而“蜂鸟”系统通过建立高效的索引机制, 客观上将所有的纸质档案信息化, 能够使得用户能够在短时间内精准找到所需信息, 从而提升了工作效率。该系统的智能搜索功能, 结合自然语言处理技术, 用户在搜索过程中能够自动地联想其他有可能的内容, 尽可能地理解用户的查询意图, 从而提供大量有关的信息, 降低了信息检索的难度。这一特性不仅提高了信息获取的便捷性, 也为企业员工的日常工作提供了强有力的支持, 还能够便于造价及审计人员从系统及时获得他们所需要的信息。

此外, 系统还具备强大的数据分析能力, 能够对文档内容进行深度挖掘, 提取出有价值的信息, 为企业的战略规划与市场分析提供数据支撑。“蜂鸟”系统的应用还促进了企业内部知识的积累与共享, 企业能够形成一个知识库, 员工可以随时访问和更新相关信息, 增强了团队协作的效率。这种知识的沉淀与共享, 不仅提升了员工的工作积极性, 也为企业的创新能力提供了保障。而且, 随着企业规模的扩大和业务的多样化, 信息管理的复杂性也随之增加。“蜂鸟”数字化业务文档快速搜录系统的实施, 能够有效应对这一挑战, 帮助企业在信息化转型过程中建立起高效、灵活的文档管理体系。通过持续优化与升级, 该系统

将不断适应企业发展的需求，推动企业在数字化浪潮中立于不败之地。

1.2 研究内容

本文聚焦于以下研究方向：开发数字化文档快速检索与录入技术；设计友好界面；提供安全保障，包括分权限管理、数据加密及操作记录功能；支持文档版本迭代和联动，提高数据利用率与管理规范。

1.3 数字化文档管理系统的概念及作用

数字化文档是指存放于电子设备中的文档，包括文本、图片、音频、视频、绘图、电子邮件、传真、计算机软件等各种可以存储在计算机系统里的信息^[4]。数字化文档管理系统是一种以信息技术为基础，对文档的全生命周期进行管理的解决方案。

2 需求分析

2.1 需求概述

随着企业对数字化文档管理需求的增长，系统需满足高效、便捷、安全的文档处理要求。本项目以“蜂鸟”系统为基础，针对以下需求展开：

1. 高效性：支持海量文档快速检索、提取和索引。
2. 安全性：数据加密存储，分权限访问控制，操作留痕。
3. 用户友好性：界面简洁直观，便于非技术用户快速上手。

2.2 功能需求

1. 主要功能描述。(1) 文档检索：支持 PDF、Word、Excel 等多种格式文件的导入。自动提取关键字段并生成索引，支持字段合并对比。(2) 检索功能：提供关键字检索、全文检索和语义搜索，定位相关内容。(3) 安全管理：数据加密，采用 AES-256 对存储文档加密。(4) 权限管理：RBAC 权限模型，确保不同用户访问不同文档。(5) 操作留痕：记录文档导入、修改、删除等操作。

2. 系统角色及功能描述。(1) 管理员。管理权限：分配或调整用户权限。审核日志：查看用户操作记录，监控系统活动。生成统计报表：根据文档使用情况提供可视化分析。上传文档：批量上传并构建索引。(2) 普通用户。上传文档：提交文档进行关键字段提取。检索文档：通过关键词、全文、语义搜索功能查找文档。字段合并：对比多个文档字段并生成合并结果。管理版本：查看或恢复文档的版本历史。(3) 访客。搜索文档：只能浏览公开文档的信息。查看文档：无需权限操作，仅限查看。

2.3 用例设计

在“蜂鸟”系统中，主要用户角色包括普通用户、

管理员和系统维护人员。普通用户主要负责文档的上传、搜索和下载；管理员则负责用户管理、权限分配以及系统配置；系统维护人员则负责系统的日常维护和故障排除。在明确用户角色后，我们需要为每个角色设计相应的用例。对于普通用户，关键用例包括文档上传、文档搜索、文档下载和文档分享。文档上传用例需要支持多种格式的文档，并提供批量上传功能；文档搜索用例需要支持关键词搜索、全文检索以及高级搜索功能，以提高搜索的准确性和效率；文档下载用例需要确保下载的便捷性和安全性；文档分享用例则需要支持多种分享方式，如链接分享和邮件分享。对于管理员角色，关键用例包括用户管理、权限管理和系统配置。用户管理用例需要支持用户的添加、删除和信息修改；权限管理用例需要支持不同用户角色的权限分配和调整；系统配置用例则需要支持系统参数的设置和调整，以满足不同业务场景的需求。

系统维护人员的用例设计主要集中在系统监控和故障处理。系统监控用例需要提供实时的系统状态监控和日志记录功能；故障处理用例则需要支持故障的快速定位和修复，以确保系统的稳定运行。在用例设计过程中，我们采用 UML（统一建模语言）进行用例图的绘制，以直观地展示系统的功能结构和用户交互。每个用例都需要详细描述其前置条件、后置条件、基本流程和备选流程，以确保用例的完整性和准确性^[5]。

2.4 非功能需求

1. 性能需求：单次搜索响应时间 < 2 秒。系统支持 1 000+ 并发访问。

2. 兼容性需求：支持主流浏览器（Chrome、Edge）。界面支持多分辨率显示。

3 系统设计及实现

3.1 系统总体设计

1. 系统架构。采用三层架构：表示层（前端 UI）、业务逻辑层（Flask API）、数据层（MySQL 数据库）。支持分布式部署，提升并发性能。

2. 技术选型。(1) 前端框架：Vue.js 实现动态交互。(2) 后端技术：Python Flask 负责核心业务逻辑。(3) 数据存储：使用 MySQL 进行字段和索引管理。

3.2 功能模块详细设计与实现

针对系统的几个重要功能点进行分析和解剖，采取活动图的方式展现主要功能如何得以实现。

1. 功能点 1：文档上传。文档上传是系统的核心功能之一，用于支持用户将文档导入系统，并自动完成解析和索引^[6]。实现步骤：(1) 用户选择上传文件。(2) 系统验证文件格式和大小。(3) 成功后，系统解析文

件内容,提取关键字段。(4)生成文档索引并存储到数据库。(5)上传完成,返回上传成功提示。

2. 功能点 2: 文档检索。文档检索功能支持用户通过多种方式(关键词、全文、语义)快速查找相关文档。实现步骤:(1)用户输入查询条件(关键词或语义)。(2)系统检查查询条件的合法性。(3)根据条件从索引表中匹配文档。(4)返回文档列表,用户可选择查看或下载文档。

3. 功能点 3: 文档字段合并与比对。该功能点允许用户选择多个文档进行字段比对,并生成合并结果。实现步骤:(1)用户选择多个文档。(2)系统提取选定文档的关键字段。(3)用户选择需要比对的字段。(4)系统生成比对结果,并允许用户确认。(5)用户选择合并操作,生成新的文档版本。

4. 功能点 4: 文档版本管理。文档版本管理支持用户查看历史版本、进行版本比对以及恢复到指定版本。实现步骤:(1)用户选择查看文档版本。(2)系统查询文档的版本历史。(3)用户选择需要对比的版本。(4)系统展示版本间的差异。(5)用户可选择恢复到某一版本。

5. 功能点 5: 安全管理。安全管理包括用户权限控制、操作记录、以及数据加密。实现步骤:(1)用户登录并验证身份。(2)系统根据用户角色分配权限。(3)用户执行操作,系统实时记录操作日志。(4)数据在存储或传输时加密。

4 结论与展望

4.1 总结

本文设计并实现了一种高效、安全的小型化文档管理系统,旨在解决个人及小型团队对离线文档管理的需求。系统具备以下核心创新点:

1. 快速检索功能:结合倒排索引技术和 NLP 方法,系统支持高效的关键字段检索。通过解析文档内容并构建索引,优化了检索速度与精度,同时实现了文档的结构化存储和多版本联动功能。

2. 高安全性设计:系统采用分权限管理机制,支持字段级自定义权限控制;通过加密存储、录入留痕和数据分批导出功能,提升数据安全性,防止信息泄露,满足用户的隐私需求。

3. 用户友好性:系统通过模块化界面设计与可视化功能布局,实现便捷操作,用户可以轻松完成文档上传、解析、检索及管理任务。同时支持跨平台使用,为文档管理提供灵活性。

4. 小型化与便携性:系统优化了存储和运行效率,适合于资源受限的设备,特别适用于移动办公和临时项目管理需求。

通过本文的研究,所提出的文档管理系统不仅提升了文档的利用效率,还在安全性与便捷性上具有显著优势,为文档管理领域提供了新的解决方案。

4.2 展望

未来,系统可进一步优化智能化水平,以满足更多行业和应用场景的需求。具体优化方向包括:

1. 语音转录功能:引入语音识别技术,将语音信息自动转录为文本文档,并同步构建关键字段索引。这将极大地简化文档录入过程,提升非结构化数据的处理效率,适应移动办公场景。

2. 自动归类功能:通过机器学习算法,系统可基于文档内容的上下文信息进行智能分类。利用聚类分析、主题建模等技术,实现多维度的文档归类,减少人工干预,提升分类准确性。

3. 上下文关联推荐:引入知识图谱或推荐算法,根据用户上传或检索的文档内容,智能推荐关联文档或参考资料。结合用户行为分析,系统能够提供精准、个性化的文档资源,帮助用户快速获取相关信息。

4. 多语言支持与翻译功能:面向国际化需求,集成自然语言处理技术,支持文档的多语言自动翻译和跨语言检索功能,扩展系统的应用范围。

5. 行业专属优化:根据不同行业需求,开发特定的文档模板和优化功能。例如,法律行业的条款自动识别、医疗行业的病例标签化等,进一步提高行业文档的使用效率。

6. 云端协同与分布式支持:将系统升级为云端协同版本,支持团队实时协作和跨设备操作。通过分布式存储和处理能力,满足大规模数据场景下的高并发需求。

通过以上智能化优化,系统将具备更强的功能和更广泛的适应性,进一步提升文档管理效率,推动多领域数字化转型。

参考文献:

- [1] 邱俊. 电力公司安全工器具管理系统的设计与实现[D]. 成都: 电子科技大学, 2020.
- [2] 邓权伦. 地市电力公司安全工器具管理系统的设计与实现[D]. 西安: 西安电子科技大学, 2016.
- [3] 钱辉敏, 姚胜红, 管磊. 电力安全工器具管理系统的研究设计[J]. 自动化应用, 2023, 64(01): 25-27.
- [4] 孙为琴. 精通 Struts 基于 MVC JavaWeb 设计与开发[M]. 北京: 电子工业出版社, 2004.
- [5] 齐红伟, 邹德旋. 基于改良和声搜索的电力系统动态经济调度[J]. 科学技术创新, 2024(01): 212-216.
- [6] 康晓华. 基于改进布谷鸟搜索算法的电力系统多目标运行优化研究[J]. 微型电脑应用, 2024, 40(04): 218-220, 225.