

人工智能大机企业文件管理系统设计

冯静乔

北京大型养路机械运用检修段 北京 100070

摘要: 针对大型养路机械企业文件管理的特点及提高文件管理查询效率需求, 系统性地构建基于人工智能AI的大机企业文件管理系统。介绍了大机企业文件管理的现状, 详细分析了人工智能AI的大机企业文件管理系统的系统结构模型、信息查询输入输出界面的结构模型、音频输入搜索的工作流程, 介绍了部分系统的软件架构, 并提出了系统建设与运行时的注意事项。

关键词: 人工智能; 大型养路机械; 文件管理

随着我国铁路高速、重载、轨道重型化的发展及客运专线、高速铁路的大规模建设及投产, 随着大型养路机械(以下简称大机)在铁道线路维修、大修施工中广泛的应用, 各铁路集团公司下属的大机企业都加强了相关的管理工作, 大机企业的技术文件、管理文件在管理中的作用愈发重要。人工智能(Artificial Intelligence)缩写为AI。它是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学, 它是计算机科学的一个分支, 该领域的研究包括机器人、语言识别处理、图像识别和专家系统等。将人工智能运用到大机企业的文件管理中, 是提高大机企业管理效率的发展方向 and 正确途径。

1 大机企业文件管理的现状

全路各大机企业使用的文件主要包括四大类, 第一类是铁总、局集团公司、段三级技术规章; 第二类是大机企业生产、管理中涉及的国家、铁道行业、大机行业、局集团公司、段的四级五类标准; 第三类是线路施工技术资料、大机设备技术资料; 第四类是企业日常人事、行政的管理文件。这些文件在各大机企业日常管理中的现状特点如下。

1.1 信息量巨大, 纸质管理向电子文档管理过渡

大机企业的各类技术、管理文件所涉及的线路施工、安全管理、大机设备管理等方面的知识结构多样, 信息量巨大, 仅第二类的国家标准、铁道行业标准就有300多项, 达140多万字。传统的纸质文件管理方式, 占用空间大, 不便保存、携带, 更不便查询。各大机企业的段、车间、工队三级文件管理都已经完成或正在向电子文档管理过渡, 一定程度上提高了文件管理的效率。

1.2 多部门管理, 文本内容不规范、格式不统一

作者简介: 冯静乔, 1978年7月, 工程师, 北京大型养路机械运用检修段

在大机企业中, 上述四类文件一般由不同的部门管辖。例如, 技术规章由指挥调度中心(线路科)管理, 企业标准由段行政办公室管理, 大机设备的技术资料由设备科管理等等。这种管理方式虽然责任清晰、专业性强, 但文件管理空间分散、文本内容不规范格式不统一, 不便于统一查询亦是明显缺点。例如, 同样的一个设备, 不同的部门写入本专业文件的名称却不同。安全部门的文件中所写的“无线防护电台”, 在行政办公室的文件中可能叫“电台”。

大部分文本类文件基本使用Word、Wps格式, 数据类的使用Excel格式, 图纸照片文件基本使用Jpg、Pdf格式。Word、WPS、Excel格式的数据使用简易查询还比较方便, 但大量的JPG、PDF格式的大机设备、线路施工技术图纸文件上包含的隐形信息, 数据读取非常困难。目前大部分大机企业没有专用的文件查询软件, 不论安全管理干部需要查询管理规章, 还是技术人员需要查询技术图纸, 信息查询的手段原始、效率低下。

2 人工智能的技术进展

人工智能从诞生至今, 应用领域不断扩大, 尤其是近十年来它在很多学科领域取得了丰硕的成果, 例如, 与海量存储器相适应的快速搜索算法、语音识别技术、图像识别核心算法。这三个领域都与大机企业文件管理密切相关。科大讯飞公司的语音识别软件能够将人的语音诵读(含方言)转化为文本, 准确度达到98%, 已是世界领先的水平。该技术能让查询者不再手动拼音或笔形输入文本, 大大节省了查询时间。2018年Yi+AI在国际计算机视觉竞赛中, 识别精度达到了90.7%。国内识别JPG、PDF等格式的软件已经比较成型, 识别精度可以达到96%。图像识别技术的进步, 大大提高了大机设备、线路施工的技术图纸中信息的查询准确度。人工智能AI软硬件技术的进步, 为建立大机企业文件管理的新模式创

造了必要的条件。

3 人工智能大机企业文件管理系统模式

3.1 运用人工智能管理大机企业文件的前提条件

针对目前大机企业文件管理的现状,要使用人工智能系统管理大机企业文件,必须提前实现文件的电子化、规范化、集中化、网络化。电子化是将文件内容按预定格式要求转化为可存储的电子版本;规范化是尽可能把文件的文件名、专业术语、文本格式,按预定要求或标准称谓进行统一处理;集中化是将所有文件的电子文档集中存储在一个服务器的数据库中,不能存在物理隔离;网络化是企业必须具备基本的互联网条件或企业内部网络条件。

3.2 人工智能的大机企业文件管理系统结构模型

人工智能的大机企业文件管理系统由搜索器、数据库、索引器、索引数据库、检索器及查询器等6部分组成,其结构模型如图1所示。该系统可以实现文件源头查找更新智能化、信息索引生成智能化、查询输入输出智能化。

搜索器即专用的搜索软件,计算机行业称之为搜索

蜘蛛。搜索蜘蛛定期通过互联网搜索大机行业适用的最新的国家标准等内容,定期通过铁路内网搜索最新的铁道行业标准、铁总技术规章、局集团公司技术规章、大机企业规章,将找到的最新文件与旧文件进行管理逻辑比对替换后存入数据库,实现文件源头查找更新智能化。

索引器即专用的索引软件,索引器可以自动将搜索来的文本中的信息切分成多个关键字,用关键字形成索引项,生成文件库的索引表。建立索引的目的就是快速地响应用户的查询,它也是搜索软件核心技术之一。智能化的索引器能够自主学习,自动将新索引与索引库中的索引项和文件内容比对,检查索引的正确性,实现信息索引的智能化。

查询器就是用户输入查询信息及接受输出信息的接口界面,使用智能化的音频识别软件,以语音输入替代笔画拼写输入。检索器即专用的检索软件,将用户欲查询内容的关键字与索引库内的索引项进行比对,并进行相关度评价。信息的输出界面一次性列出相关度靠前的信息,语音输出给查询者。这些步骤大大减少了查询者的任务量,提高了查询输入输出速度,实现查询输入输出智能化。

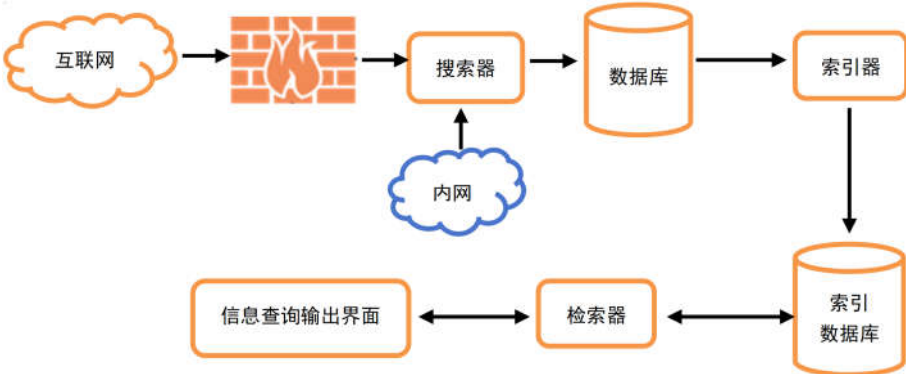


图1 人工智能的大机企业文件管理结构模型结构图

3.3 信息查询输入输出界面的结构模型

高效的文件管理的目的之一就是让使用者快速、准确地获取信息。受线路施工各车间远离办公基地特点的影响,要想让车间的管理人员、技术人员快速查询文件,可以设计一种基于移动通信和互联网技术的信息查询输入输出界面。具体如下图2所示。该方案采用前后端服务器分离+手机(移动PC设备)的结构模型。前端展示服务器负责手机页面展示、移动PC端页面展示、逻辑操作及业务文件服务器的数据交换工作,业务文件服务器负责文件有效性处理、文件索引自动生成、索引管理、文件管理等工作。前段展示服务器与业务管理服务器分离,从物理层面提高了系统安全性。

3.4 音频输入搜索的工作流程

音频输入搜索的工作流程为:(1)使用手机或PC完成语音诵读输入;(2)音频进入科大讯飞的语音软件;(3)经语音软件识别后的文本经索引软件形成检索关键字;(4)关键字进入索引数据库,与库中的索引表进行比对;(5)检查是否存在与关键字相符的文件;(6)如果没有相符文件则返回进行模糊关键字查询;如果有相符文件则继续检查是否有相符的具体信息;(7)存在相符的具体信息则在手机或PC显示或语音输出;不存在相符的具体信息则进行模糊查询。

4 人工智能大机企业文件管理系统的软件架构

人工智能大机企业文件管理系统的软件架构可分为三个层次,即展示层、业务层、数据层。以音频输入搜索、PDF识别为例,展示层采用Vue.js用户界面渐进

式框架,支持前后端分离架构。其中Vuex负责状态管理,Vue-router负责路由管理,Pdf.js负责Pdf的展示及操作处理,科大讯飞的XunfeiAPI负责提供语音转文字技术。业务层采用SpringMVC框架进行业务处理,采用

FileManager类进行文件管理,Regular-Expression对文件内容进行匹配处理,Dao及Clustered index进行文件及索引的关联管理。数据层支持Doc、Pdf形式的文件,数据库采用Mysql数据库。^[1]

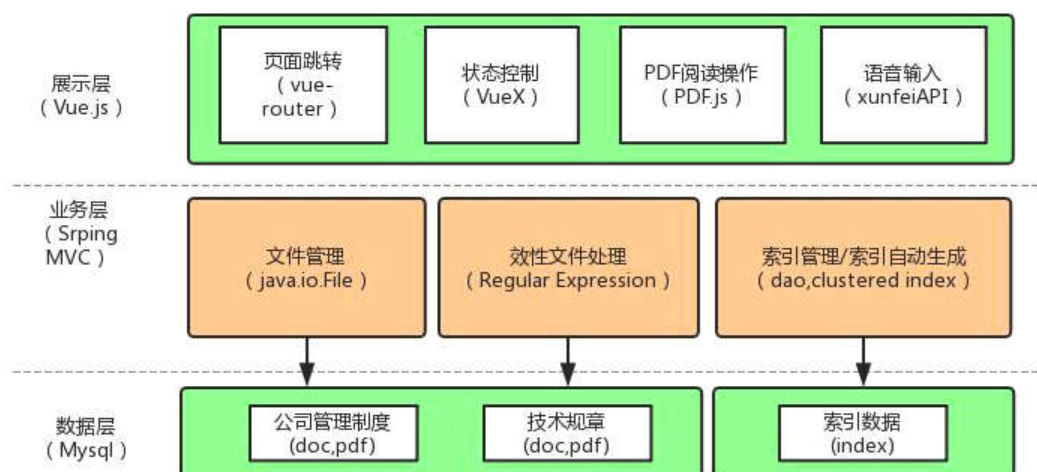


图2 软件架构图（以音频输入搜索、PDF识别为例）

5 系统建设与运行的注意事项

5.1 人工干预

在大机企业文件管理时,人工智能AI是一种完全可以得出准确的结论的解决方案。但建设过程中、使用过程中的一些环节也必须考虑人工干预。例如,在最初建立便于检索的索引表时,或是在索引器制作数据库中从未没有过的文件时,一些不是常见用语的大机企业的专业术语,需要AI系统工程师在相关大机专业技术人员的帮助下,人工干预确定进入索引表。

5.2 总体兼容与预留功能

人工智能的大机企业文件管理系统可以作为文件管理的单独系统使用,也可以作为企业管理的大系统中的一个部分使用。如果作为子系统使用,一定要提前设计与总体系统的兼容,在模型结构、软件架构等方面与总体系统协调。另外,随着网络移动通讯软硬件的进步,该系统要提前考虑与其他系统的互联互通、资源共享的要求,在软件开发、接口技术、数据格式等方面做出设计与安排。

5.3 系统的安全设计

大机企业文件中很多技术资料的数据是企业的核心秘密,加强系统的安全设计,确保文件在管理、使用、查询过程中的运行安全,保证数据、信息不超范围扩散,是关系到企业保密、发展的重要工作。在进行该系统安全设计时,要提前评估文件的涉密情况,要提前设计系统服务对象的查阅等级,合理规定文件等级对应的

查阅等级,不能越级查阅,高密级文件不能显示全部文件内容。

加强系统服务器间的物理隔离措施,前端展示服务器和后台管理服务器必须分开,一旦被黑客攻击也能保证后台管理服务器内容不泄露,系统可以尽快恢复运行。要合理运用各种现代网络安全技术,如防火墙技术、防病毒技术、入侵检测技术等,要定期请网络安全公司对系统进行安全等级保护评估审计,以确保系统可以平稳安全地运行。

6 结束语

大型养路机械是铁路工务系统先进生产力的代表,人工智能AI引入大机企业的管理后,必将产生显著的管理效益。人工智能大机企业文件管理系统的建立、运行要经过大量艰苦、细致的工作,需要计算机专家和大机专业管理工程师之间密切配合。相信通过不懈努力,一定能建好、用好它。企业文件存在的最大价值就是被按标准执行,使用人工智能大机企业文件管理系统,能够让企业的管理者和执行者最快地获取文件中的信息。凯文·凯利曾形象地总结说:“使用人工智能的人越多,它就越聪明。人工智能越聪明,使用它的人就越多。”只有如此,个人才有信息力、部门才有执行力,管理才有控制力,大机企业才有核心竞争力。

参考文献

- [1]邓磊.文档搜索智能化解决方案.北京崖柏科技有限公司