

附件一

国网陕西省电力公司科技项目 可行性研究报告

项目名称： 基于大数据的 IT 部门知识管理工具

申请单位： 国网陕西省电力公司信息通信公司

起止时间：

项目负责人：

通信地址：

邮政编码： 710065

联系电话： 029-81003211

传 真：

申请日期： 2018 年 4 月

一、目的和意义

随着互联网时代信息化水平的不断提高以及大数据的发展,陕西省电力公司面临的业务数据量、服务专业性、处理时效性等方面的要求也愈加严峻,对员工的能力素质提出了极高的要求。尤其是陕西电力信通公司作为陕西电力通信网、信息系统及西安信息灾备中心的运行维护管理职责的承担者,主要负责信息化建设、为省公司企业和电网发展提供技术支撑和优质服务,因此,员工的知识熟练程度和业务水平显得尤为重要。特别是由于近两年技术更新的速度越来越快、知识容量越来越大,员工需要掌握的知识和技能也越来越多,不仅需要掌握办公自动化系统积累的基本操作、安全常识、行业规范等技术应用方面的知识,还包括大量的企业规章制度、办事指南、党政理论学习等管理方面的知识,还有一些在业务实践中积累完善的经验及解决方法,这些知识不仅数量巨大,而且极为琐碎,单独让员工采取传统的记忆、考试等方式进行知识的学习转化,不仅费时费力而且效率很低,还会削减员工技术更新的积极性,从而降低工作效率。因此通过适当的知识管理工具对部门知识进行整理、转化,并且采取简单、有趣的方法促进员工的吸收理解、将这些知识内化为员工的业务水平,是既有必要性又极为重要的。另一方面,大数据时代所要求的知识管理工具的更新速度与处理速度,已经超过了以往知识管理工具传统的提速方式的逻辑负荷,必须以新兴的大数据技术来重塑整个业务流程。

现有的知识管理工具往往更注重信息的搜集和整理,却极少考虑如何促进员工对于知识的消化和吸收,忽视了知识的转化。传统的知识管理中,知识的转移主要通过公司组织集中学习、会议讨论、专题讲座等方式实现,依赖员工被动式地接收和学习,缺少一定的激励机制和趣味性;或者通过构建知识共享学习平台,但现有的平台提供的学习是单方面的,只是员工向数据库中的知识学习,而不是针对员工自身情况对其进行推荐、建议,也无法实现员工与系统的反馈交流,员工的学习积极性和粘性并不高。另一方面,现有知识管理工具缺乏对知识库的更新、筛选和迭代,也存在一些过时或者错误的知识,会对员工产生误导;而且现有知识管理方式下,员工多采用纸质文件或者电脑网页学习,极为不方便,存在时间和地点的限制;管理者也无法掌握员工的学习情况,更无法针对性地对员工提出指导,这些导致员工学习过于被动、枯燥,知识转化为技能和业务的效果甚微。

因此，目前的知识管理工具缺少一种促进员工学习常态化、时间碎片化、操作简单化、学习趣味化的方式和方法。为解决这些问题，本项目将大数据和机器学习算法结合起来，并采用微信小程序开发作为员工知识管理工具，在整理、构建各领域知识库的基础上，设计基于员工评分与统计检测的机器学习动态迭代题库，并在题库的基础上建立以好友对战答题、答题练习、学习排行、知识回顾、积分奖励等多种应用，提升员工的兴趣，并且利用大数据技术对员工答题情况进行数据挖掘和分析，为其制定符合自身情况和学习梯度的练习题目和知识路线，同时利用机器学习技术学习员工的学习方式和答题情况，训练答题机器人模拟与其水平相当的题目与其对战，从而引领其更快地进步，并增强学习乐趣。

需要指出的是，本项目知识管理工具的建设思路是不取代现有应用系统（如办公自动化系统、培训系统等）的数据存储功能，而是从原有的应用系统中有选择地提取各方面有价值的信息，结合用户录入的显性知识和社区沉淀下来的隐性知识搭建知识库。而且，考虑到智能终端的便捷性和使用频率，选择采用依托于微信平台的微信小程序作为前端入口，这样用户无需长时间对着电脑或者书本，学习时间也更加便捷，可以大大利用碎片化时间，而且无需安装APP。

该项目虽然立项于信通部门，但是不仅适用于信通部门，实际上对于陕西省电力公司全体部门适用。如果在信通部门应用成功，可以方便地推广到其他部门，仅仅需要对知识库的设计进行改变，推广便捷且成本极低。

本项目在推广之后的效益是极大的，实施知识管理不仅整合了公司内部的知识，促进了知识的转化，增强了员工的业务水平和知识熟练程度，降低了安全隐患，提高了业务水平，而且通过业务水平的提升和整体素质的提高可以为省电力公司带来可观的业绩和经济收益，而且还能促进省电力公司的文化建设与可持续发展，对于电力企业实现管理创新和制度创新、增强企业竞争力都具有重要意义。

二、国内外研究水平综述

传统的知识管理工具，不论是知识库的建设还是扩展上，很大程度上依靠人的智力与智慧。而大数据时代的知识管理系统，不仅能够替代人们自动地建设知识库，还能像人类一样具有智力与智慧去思考和推理，并生产和组织知识，必须以新兴的大数据技术来重塑整个业务流程，用大数据的方法来指导知识管理工具的建设，而且必须充分考虑员工主体的积极性和创造性，促进知识的转化与吸收。目前，国内外一些企业已经进行了知识管理工具的革新，并且取得了不错的效果。

（1）谷歌智能化知识管理工具

谷歌公司于 2014 年 8 月末公布了在建的“知识库”（Knowledge Vault），这是谷歌继 2012 年开始的“知识图谱”（Knowledge Graph）之后又一大作。“知识库”通过算法自动搜集网上信息，通过机器学习把数据变成可用知识。目前，Knowledge Vault 已经收集了 16 亿件事实，其中，2.71 亿件是“可信的事实”。据专家估计，大型企业特别是大型的科技企业，都在建设类似的知识库。当机器系统掌握了全人类的知识，它的智力将远远超过现有的智能助手。知识管理系统的智能化发展，将把人类从知识管理流程的低端阶段中解放出来，从而更专注于高级知识的生产和应用。这将是大数据时代中的知识管理系统对传统的知识管理系统的一大革命。

（2）广州供电局基于知识库的主动式学习电网运行经验库

2014 年广州供电局有限公司提出了运用面向对象技术与产生式规则知识表达方式，将电网中的运行经验按工作性质分类，建立一个拥有主动学习能力的电网运行经验知识库。运用基于权重的贝叶斯分类模型来对统计数据进行分类，从而识别语义，以此达到电网运行经验知识与用户的反馈交流的目的。利用知识库强大的学习能力，使电网运行经验知识库在与用户反馈交流中优化自身知识体系结构，有地区针对性地进一步扩充与完善运行经验库。该方法可以有效地为工作人员提供技术支持，提高工作人员的业务素质，加强电网的安全稳定运行。

（3）广西电网电气设备知识库管理工具

2010 年广西运用先进的计算机网络技术和数据库技术构建广西电网电气设备知识库，通过系统与员工互动、员工与员工互动来沉淀企业隐性知识，使知识库体现企业的集体智慧，实现了知识的主动个性化推送，提高知识的使用效率，通过知识奖励，激励员工，创建学习型企业，该知识库为员工提供电气设备知识检索、学习与沉淀的平台，为电气设备故障处理、设备危险点排查、设备预警等提供辅助决策服务。

这些知识管理工具的革新取得了不错的效果，但也存在一些不足之处，例如对大数据的应用还不足，对员工知识转化规则设计还不够，缺少对于用户的粘度和激励机制。

三、项目的理论和实践依据

基于现阶段知识管理存在的问题和改进之处，本项目提出基于大数据的 IT

部门知识管理工具，在整理、构建各领域知识库的基础上，设计基于员工评分与统计检测的机器学习动态迭代题库，并在题库的基础上建立以好友对战答题、答题练习、学习排行、知识回顾等多种应用，并利用数据挖掘和机器学习算法引导和辅助员工进行知识管理。下面将按照知识管理、知识库构建、动态迭代题库设计、员工学习情况挖掘、微信小程序等方面对原理进行简单介绍。

1. 知识管理原理简介

知识管理是指企业对知识资源(包括企业无形资产、信息资源、智力资源)进行有效管理和利用，以提高企业的创新能力，进而提高企业创造财富的能力。知识管理的前提条件是企业内成员相互信任，最大限度地贡献出其积累的知识，共享企业内的知识资源。其直接目的是提高企业的创新能力，终极目的则是为了提高企业创造财富的能力。

知识可分为显性知识和隐性知识。显性知识是指可以用规范语言清晰表达的，且易于在个体之间同步和异步传播的知识；而隐性知识是难以形式化、难以编码和交流的知识，表现为个人的经验教训、洞察力等。相比之下，隐性知识能为企业创造更大价值，也更难获取。

知识管理工具一般具有很强的针对性和实用性，具有以下特点：

- (1)对公司知识进行全面梳理，建立跨平台、跨数据库的知识库。
- (2)有效管理部门级知识，实现部门文件的存储和共享。
- (3)通过知识实践社区的建立，来沉淀企业隐形知识，使知识库不仅仅是企业的文档集合，更是企业集体智慧的体现。
- (4)实现知识的主动个性化推送，如知识推荐、桌面推送。

2. 知识库构建原理简介

知识库是建立知识管理系统的重要支撑，任何领域的专家系统、智能化服务都需要可靠的领域知识。

电力企业进行知识库建设需要有丰富的信息资源作为基础。在进行电力企业知识库建设时，首先要利用组织保障体系完善电力信息资源系统的开发，形成集中、统一、稳定的信息采集渠道。电力企业知识库可分为两部分：一是关于企业自身的知识，主要由企业的政务信息、业务信息、综合服务信息和辅助决策信息构成，开发利用的重点是业务信息；一是关于企业外部资源的信息，获取这部分知识的主要途径是互联网。据统计，70%的信息分散在互联网的各个页面上。所

以说,这又是目前我们获取有用信息效率最低的途径。因此,电力企业必须建立网上专业信息抓取、查询系统,建成包括文字、图形在内的多媒体分布式和集中式综合信息数据库。

随着信息化建设的不断提升,知识库的建设方向,应纳入丰富多样的知识源:包括日常运作的核心业务流程的主要流程数据(人力资源管理、科技报告、项目管理、设备管理、采购管理等等);企业内部网站数据(企业规章制度、质量管理文件、办事指南、工作表格等),此外在网站虚拟社区中包含了用户主动上传的可供共享的知识。以上这些都是将知识从隐性化转化成显性化的良好途径。

在上述知识源中,既有公文等国网公司统一管理的数据,也包含了许多员工在职务服务范围内创造或履行职务产生、具有自主知识的试验方法、作业规范、质量控制、实验数据等技术知识,而正是这些知识,构成企业不断创新、不断提高的重要基础。

3. 基于员工评分与统计检测的机器学习动态迭代题库原理简介 题库设计原理简介

题库设计理论的涵义是广泛的、全面的,重要包括如下几个部分:学科体系与教育、目标层次分类理论、题型与命题技术理论、题目分析理论、参数等值理论,以及试题与试卷统计性质关系的理论等。题库设计的指导理论主要有2种:经典测量理论(CTT)和项目反应理论(IRT)。

题库在设计时有许多要求,一般包括以下几点:

(1)题库中的题目要按领域分类,每一领域的题目又细分为若干个题目组合,这些题目组合可用于特定的目标,测验特定的概念或能力。

(2)题库中的每一题目由两部分组成,即题目属性指标和题目本身。题目属性指标有多项,反映有关题目的信息,这些信息将某一题目与库中的其他题目区别开,同时标记着题目使用的历史和使用情况(如使用次数,被高分员工答对的次数和被低分员工答对的次数等)。对于这些题目属性,入库前都应一一查明、核实,且质量参数的值都应表达在统一度量的基础之上。

(3)题库的题量必须很大,比例合理。如果题量太小,就难以生成让员工满意的试卷,也不能长时间地多次地为考试服务。同时题目的比例要合理科学,主要是内容范围、能力层次、难易比例、题目类型等方面。

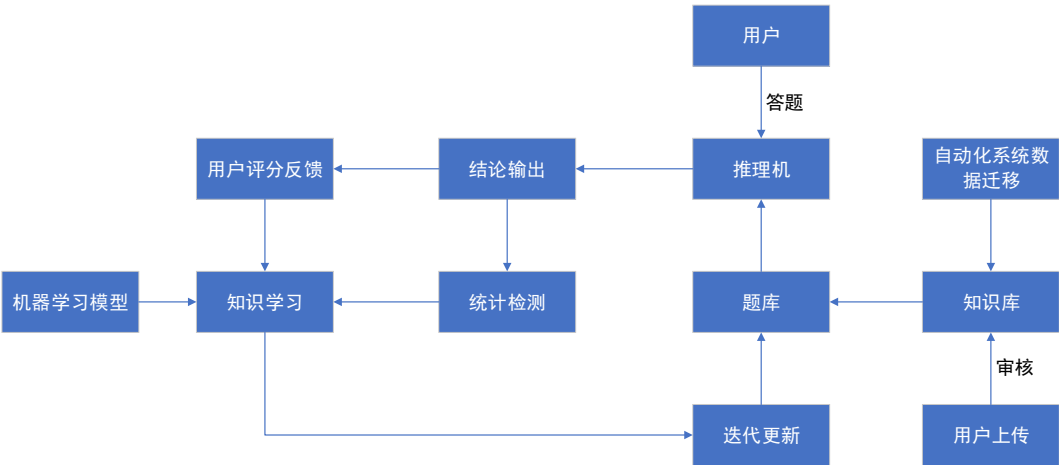
(4)试题库可以动态地维护。这是因为各个领域都在发展变化,因此领域的

知识也应不断更新。此外，员工水平也在不断地变化，这样题目参数及考核能力层次也会有所变化。因此要使试题库能长时间地服务，就应该在题目的更新、修改、删除、添加上下功夫，即要求题库必须是动态的。

(5)题库具有构成等价形式测验的能力。等价形式的测验是由不同的题目组合构成的，但每个测验的统计结果具有等效性，即这些不同的题目组合去测验相同的目标或概念，并用相同的统计方法来区分学生学习的好与差。

基于员工评分与统计检测的机器学习动态迭代题库原理简介

基于员工评分与统计检测的机器学习动态迭代题库的主要步骤是，在整理、归纳得到不同领域知识库的基础上，通过抽样和题库设计理论得到原始题库（同时，题库可以对知识库进行反馈修正），利用员工的答题情况以及员工的评分反馈，采用机器学习方法对题库进行迭代更新，原理结构图如下图。



实现题库的动态迭代，要求系统能根据员工的评分对题目进行校正与更新，并能根据全体员工答题情况的统计检测，主动学习员工对各类题目的掌握情况，从而对题库进行修正。例如，员工答错某一类问题的频数越大，该信息的权重也就越大，系统推荐给员工的题目类型和次数也会根据权重大小进行调整；如果大多数员工都没有答错某一类题目，则题目的权重应该很小。同时，如果员工对题目的答案有问题，或者觉得这类问题没有意义或设置不合理，都可以通过对题目评分等方式反馈给系统，在系统获得这种反馈后，针对性对题库进行调整更新。

机器学习(Machine Learning, ML)是一门多领域交叉学科，涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能。它是人工智能的核心，是使计算机具有智能的根本途径，

对能通过经验自动改进的计算机算法的研究，其应用遍及人工智能的各个领域，

4. 员工学习情况挖掘原理简介

利用大数据技术对员工学习情况进行数据挖掘和分析，为其制定符合自身情况和学习梯度的练习题目和知识路线，同时利用机器学习技术学习员工的学习方式和答题情况，训练答题机器人模拟与其水平相当的题目与其对战，从而引领其更快地进步，并增强学习兴趣。

数据挖掘（Data mining）一般是指从大量的数据中通过算法搜索隐藏于其中信息的过程。数据挖掘通常与计算机科学有关，并通过统计、在线分析处理、情报检索、机器学习、专家系统（依靠过去的经验法则）和模式识别等诸多方法来实现上述目标。

5. 微信小程序原理简介

微信小程序是一种用户无需下载，无需安装就可以直接在微信内打开的简单应用。对于用户来说方便，有很好的体验；对于单位来说，做一款小程序，可以降低开发费用，同时小程序方便进行运营推广，无需高额的推广费用。

从技术而言，小程序是一种需要部署到腾讯服务器上的网站“前端”应用，通过微信里面的各种入口，打开小程序后通过JSON 数据格式，连接到小程序的“后台”做数据交互。

小程序的“后台”是小程序的开发商或者客户自身部署在自己的服务器上的“后台”程序，这些后台程序一般可以由各种现有的Web 后端编程语言编写，例如PHP，Python，Java，Go等等。

而小程序本身就是我们说的 Web 前端语言，是由腾讯对JS/H5/CSS/XHTML定制的各种不同组件的整合，考虑了对iOS和Android 以及各种手机屏幕大小以及手机类型的兼容后的一种Web前端框架，里面包含了各种组件，并针对手机的不同功能开发了各类API。

小程序基于月活跃用户9亿人次的微信生态中，流量大，入口多，功能简单却便捷。对于企业而言，无论是在企业宣传、企业营销、售后分析等方面存在巨大的潜力。有业内人士认为：小程序功能快速迭代，意味着围绕小程序的开发和生态工具建设将会是移动互联网的有一个巨大机会。目前各行业内诸多企业单位加入了小程序开发，开通了小程序功能，且成果喜人。

本项目实施的关键在于两部分，第一是基于用户评分反馈的题库更新的实

现；第二是运用大数据技术和机器学习算法对员工答题情况进行指导、辅助。

本项目实施存在以下难点：题库的组织如何兼顾到难度系数和区分度，如何用机器学习模型对题库进行动态更新，以及选择合适的模型对员工的答题情况进行学习从而训练出与之水平相适的答题机器人进行辅助答题等。

四、项目研究内容和实施方案

本项目将大数据和机器学习算法结合起来，在整理、构建各领域知识库的基础上，设计基于员工评分与统计检测的机器学习动态迭代题库，并在题库的基础上建立以好友对战答题、答题练习、学习排行、知识回顾、积分奖励等多种应用，采用微信小程序和Java语言开发知识管理工具。

本项目的系统设计和实施方案如下：

1 系统架构

采用微信小程序作为前端，Java作为服务器端开发语言，使用Tomcat作为服务器，数据库采用MySQL，前端采用wepy框架，后台采用MVC架构，采用了传统的三层开发模型，业务层与表现层之间用了IOC组件；表现层采用了组件进行模型与控制、视图的分离。

2 开发平台

后台开发工具为IntelliJ IDEA，小程序开发工具为微信web开发者工具。

3 功能模块

功能模块可以分为系统知识管理功能和个人知识管理功能两方面，其中系统知识管理主要处理题库的动态迭代和更新、总体答题情况统计分析，个人知识管理功能包括好友对战答题、答题练习、错题回顾、上传题目、题目评分、答题排行以及机器人模拟答题等功能。

系统知识管理功能：

（1）题库设计和更新。将部门专业知识和业务规范、业务平台操作、规章制度等知识以及实践经验等整理归纳形成知识库，并按照题库设计理论形成题库，利用员工的答题情况以及员工的评分反馈，采用机器学习方法对题库进行迭代更新。

（2）总体答题情况统计分析。答题情况的统计分析能帮助管理者掌握各类题目被员工学习和掌握的情况，一方面有助于及时补充和更新系统的知识，另一方面为制定适合员工学习情况的答题路线提供参考。

个人知识管理功能：

(1) 好友对战答题。员工可以向部门的其他员工或者其微信好友发起答题对战，两人同时在小程序界面对一定数量的题目进行答题，最后根据每个人答题的情况对其进行积分增减，积分达到一定程度可以购买道具，道具可以获得额外的上传题目、题目问答等次数，也可以实际兑换奖品，或当做考核的额外参考。

(2) 答题排行。不同的积分分段对应不同的“段位”，积分越高，“段位”越高。系统会根据员工答题的积分授予其不同的“段位”，同时在部门内部会有排行，按照“段位”从高到低进行排行，对员工进行激励。

(3) 题目收藏，错题回顾。提供用户个人收藏夹功能，员工可将自己感兴趣的题目或不确定的题目收藏到个人收藏夹，下次登录系统时不必再查找，在收藏夹里可直接点击浏览。同时，员工的错题会自动放在个人收藏夹之中。这样，员工可以更快地找到需要的信息，及时复习。

(4) 题目反馈。员工可以对答过的题目进行评价，比如当认为题目有错误或者题目不必要等，都可以对题目进行反馈。

(5) 上传题目。员工可以对自己总结的有意义的且题库中没有的题目进行上传，员工上传的题目需要经过审核才会更新到题库。

(6) 积分奖励。系统通过积分奖励来调动员工学习知识的积极性。系统自动记录员工的答题次数、时长，以及答题情况，根据奖惩机制进行积分奖励，所获得的积分可以购买道具，道具可以获得额外的上传题目、题目问答等次数，也可以实际兑换奖品，或当做考核的额外参考。

(7) 答题路线推荐。帮助员工根据自身的实际情况和相同水平的员工的答题水平制定适合他的答题路线，系统会全程、自动跟踪学习完成进度。

(8) 机器人辅助答题练习。利用机器学习技术学习员工的答题情况，并通过大数据分析相同答题水平的员工的答题情况，训练答题机器人模拟与其水平相当的题目与其对战，从而引领其更快地进步，并增强学习兴趣。

4 实施步骤：

整体设计遵循软件开发的流程，确保系统具有较高的安全性、稳定性和易用性，主要包括以下步骤：

(1) 需求分析，包括数据模型分析、功能模块分析等；

(2) 总体设计；

(3) 产品原型设计；

(4) 美工设计；

(5) 微信小程序前端实现，负责具体的页面展现、页面跳转、动态数据的绑定和交互；

(6) Java服务器后端实现，实现数据处理、业务逻辑处理、数据库处理等；

(7) 模块测试、试运行、上线。