

科研平台合作共建

项目建议书

项目名称： 电力大数据算法及应用研究

建 议 人： 胡飞虎

日期： 2018 年 4 月 28 日

一、大数据的背景简介

随着互联网的迅速发展以及人工智能等新兴数据挖掘与分析技术的不断创新发展，目前，大数据以爆炸式的发展速度迅速蔓延至各行各业。随着各国抢抓战略布局，不断加大扶持力度，全球大数据市场规模保持了高速增长态势。大数据是指通过对大量的、种类和来源复杂的数据进行高速地捕捉、发现和分析，用经济的方法提取其价值的技术体系或技术架构。所以，广义上讲，大数据不仅是指大数据所涉及的数据，还包含了对这些数据进行处理和分析的理论、方法和技术。

20 世纪 90 年代，“大数据”概念被提出，最初只是对一些在一定时间内无法用传统方法进行抓取、管理和处理的数据统称。2009 年以来，互联网数据以每年 50% 的速度增长，整个互联网领域开始意识到大数据时代来临。2011 年 5 月，麦肯锡全球研究机构发布了《大数据：创新、竞争和生产力的下一个前沿领域》报告，大数据概念在计算机行业迅速火热起来。2012 年 1 月，瑞士达沃斯论坛上《大数据，大影响》(Big Data, Big Impact) 报告称，数据如同货币或黄金一样，已成为一种新经济资产类别。显然，人们对大数据的关注度正在逐步提升。而近年来随着互联网、人工智能、云计算技术的发展，已经进入了大数据发展的新阶段。

据 IDC 预测，全球大数据市场规模年增长率达 40%，在 2017 年将达到 530 亿美元。全球许多国家已经将大数据列为国家发展的主要任务，美国于 2012 年便启动了“大数据研究和发展计划”，英国政府自 2013 年开始就注重对大数据技术的研发投入，日本在《日本再兴战略》中提出开放数据，将实施数据开放、大数据技术开发与运用作为 2013-2020 年的重要国家战略之一。

面对迅猛发展的大数据技术，我国也早已将大数据提升到国家战略层面。为指导和推动大数据产业发展，先后出台了一系列的文件和政策：

2015 年 9 月，国务院印发《促进大数据发展行动纲要》，工业和信息化部为全面部署“十三五”时期大数据发展，组织编写了《大数据产业发展规划（2016-2020 年）》。

2015 年国务院印发的《关于积极推进“互联网+”行动的指导意见》针对“互联网+智慧能源”专项中指出，“推进能源生产智能化”，“鼓励能源企业运用大数据技术对设备状态、电能负载等数据进行分析挖掘与预测，开展精准调度、故障判断和预测性维护，提高能源利用效率和安全稳定运行水平”。

2016 年 4 月国家发改委召开促进大数据发展部际联席第一次会议，会议强调发挥数据资源的战略作用，把握工业经济和信息经济交汇发展的关键时期，推动产业创新发展，做好大数据产业发展的规划，推动工业大数据、互联网与制造业的融合发展。

2016 年 12 月国务院印发的《“十三五”国家信息化规划》中多处提及大数据及其相关内容，其中，在第四部分重大任务和重点工程及第五部分优先行动中，强调了“建立统一开放的大数据体系”及“数据资源共享开放”的任务和行动。

2017 年 12 月习近平在主持中共中央政治局实施国家大数据战略的第二次集

体学习时强调，大数据是信息化发展的新阶段，我们应该深入了解大数据发展现状和趋势及其对经济社会发展的影响，推动实施国家大数据战略。

同时，各地政府也高度重视大数据产业发展，相关政策密集出台。据不完全统计，已有 20 多个省级地方和 10 余个部委出台了本地区、本行业大数据发展规划。

目前，我国大数据产业发展在经历初期探索、市场启动等阶段后，大数据的技术、应用以及社会公众的接受度等方面逐步趋于成熟，整个产业开始步入快速发展阶段，行业规模增长迅速。从大数据产业规模来看，根据中国信息通信研究院《中国大数据发展调查报告》统计显示，2015 年 1 月-2016 年 7 月，大数据领域公开招标项目为 1276 个，项目平均额度为 300 万，总采购金额约为 38 亿元；2016 年大数据项目增长呈加速之势，仅 2016 年上半年启动的项目就与 2015 年全年启动的项目数相当；2016 年我国大数据市场规模为 168.0 亿元，增速达到 45%，预计 2017-2020 年增速保持在 30%以上。

未来，随着市场对数据的重视，大数据市场规模还会继续增长，大数据技术还将为国家发展带来更大的前景和空间。

二、电力大数据的发展概况

随着大数据、人工智能等新兴数据挖掘与分析技术的不断创新发展，为电力行业业务创新、智能化辅助决策、服务能力提升、市场竞争力增强等方面的发展提供了无限空间。

2013 年 3 月中国电机工程学会信息化专委会发布《中国电力大数据发展白皮书》，将 2013 年定为“中国大数据元年”，掀起了电力大数据的研究热潮。作为正向能源互联网转型的传统电力行业，大数据及云计算时代的到来将为传统电力行业的发展注入新的活力，传统电力行业即将发生革命性的变化。

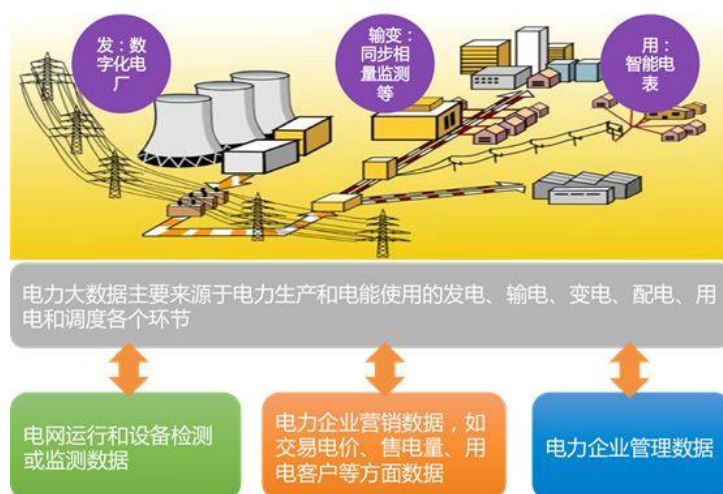
2015 年 11 月，国家能源局，江苏省政府和国际能源署在苏州举办了“国际能源变革论坛”，并联合发表了《苏州宣言》。宣言中提到“探索能源互联网发展对能源变革的作用，推进两者之间的协同发展，对能源大数据的管理与利用进行前瞻性研究”。

根据麦肯锡报告预测，在全球范围内，大数据分析方案的广泛使用能够带来每年 3000 亿美元的电费削减。电力大数据的有效应用可以面向行业内外提供大量的高附加值的增值服务业务，对于电力企业盈利与控制水平的提升有很高的价值。有电网专家分析称，每当数据利用率调高 10%，便可使电网提高 20%~49% 的利润。

在能源互联网+新电改的背景下，在大数据+云计算的新时代，依托电力大数据的电网将迈进全景实时电网的时代。电力作为二次能源，其高效利用、优化配置和便捷服务，对节能减排、建立资源节约型和环境友好型社会的意义重大。电力大数据贯穿电力工业生产及管理等环节，对电力工业发展将起着巨大的作用，是解决中国电力工业未来有效应用资源有限、环境压力大等问题的良好途径。大数据必然会助力电力工业的发展，成为电力工业发展的驱动器。因此，研究电力大数据具有重要意义。

2.1 电力大数据概述

电力大数据是指通过传感器、智能设备、视频监控设备、音频通信设备和移动中断等各种数据采集渠道收集到的，结构化、半结构化和非结构化的海量业务数据的集合。电力大数据不仅是大数据技术在电力行业的深入应用，也是电力生产、消费及相关技术革命与大数据理念的深度融合，将加速推进电力产业发展及商业模式创新，如下图所示。



电力大数据主要来源于电力生产和电能使用的发电、输电、变电、配电、用电和调度各个环节，可大致分为三类：一是电网运行和设备检测或监测数据；二是电力企业营销数据，如交易电价、售电量、用电客户等方面数据；三是电力企业管理数据。

电力大数据具有四个特点：

（1）数据体量大：PB 级；常规的调度自动化系统包含数十万个采集点；配用电、数据中心将达到千万级；

（2）数据类型繁多：实时数据历史数据文本数据多媒体数据时间序列数据等各类结构化、半结构化数据以及非结构化数据；

（3）价值密度低：所采集的绝大部分数据都是正常数据，只有极少量的异常数据，而异常数据是状态检修的最重要依据；

（4）处理速度快：在几分之一秒内对大量数据进行分析，以支持决策制定。

云计算、大数据分析等信息新技术必将激活电力大数据中蕴含的价值，也必将释放电力大数据的市场潜力。根据 GTM Research 的研究分析，到 2020 年，全世界电力大数据管理系统市场将达到 38 亿美元的规模，电力大数据的采集、管理、分析与服务行业将迎来前所未有的发展机遇。

2.2 电力大数据应用研究现状

2012 年以来，国内外在电力大数据技术研究和工程应用方面做了一些有益的尝试，奠定了一定的基础，电力大数据理论研究和应用已取得了一些成效。在此起步阶段，充分借鉴国内外研究成果，分享知识经验，对于促进开展电力大数

据的研究和应用，具有重要的意义。

国外应用研究举例：

（1）法国电力公司基于大数据的用电采集应用

法国目前已安装 3500 万个智能电表，智能电表采集的主要是个体家庭的用电负荷数据。以每个电表每 10 min 抄表一次计算，3500 万智能电表每年产生 1.8 万亿次抄表记录和 600 TB 压缩前数据，电表产生的数据量 5~10 年内达到 PB 级。

针对这一情况，法国电力公司的研发部门成立了 Big Data 项目组，对数据进行挖掘分析，从而实现负荷曲线数据能够进行高速处理，短期用户用电趋势能被预测。通过借助大数据技术研究海量数据的处理架构，形成能够支撑在规定延迟内的复杂、并行处理能力；可以在不同尺度上进行处理，某些应用实现了实时处理；实现了电网调度等高级应用（电网状态监测、电网自动愈合；对电网调度进行局部优化；用电需求侧管理，实现了实时电价及电网的可再生能源的接入。

（2）丹麦维斯塔斯基于大数据的数据实时处理平台

在过去十年中，丹麦维斯塔斯已安装的风力发电机及其他收集到的环境信息，遍及全世界，已累积 2.6 PB 气象数据。面对这一庞大数据，通过现有方案无法及时处理。丹麦维斯塔斯采用了 IBM 的 BigInsights 大数据平台解决了海量数据分析与处理问题，优化风力涡轮机配置方案，从而实现最高效的能量输出。该项目对天气建模以优化风力发电机的放置，最大限度提高发电量并延长设备使用寿命；将确定风力发电机的位置所需的时间从几周缩短为几小时；纳入 2.5 PB 的结构化和半结构化信息流，预计数据量将增长到 6 PB。

国内应用研究举例：

在国内，中国电机工程学会 2012 年发布了电力大数据白皮书；国家科技部 2014 年下达了 3 项 863 项目，支持智能电网大数据研究；自 2012 年以来，国家电网公司启动了多项智能电网大数据研究项目。

（1）国网福建省电力有限公司建立的电网运营在线监测分析系统，实现了配网运营效率由线下定性评价向在线定量监测分析的转变。系统将大数据分析技术与实际业务需求有机结合，在配网供电能力、电网优质服务、防范企业经营风险等方面展开积极探索与实践，快速定位电网运营薄弱环节，为该公司领导决策、专业部门提升管理、基层单位改进工作提供支撑。另外，国网福建电力依托单体配变的配电网重过载风险预警模型，开展城网和农网配变重过载预警分析，根据预警的周期和时效性，识别可能出现重过载的配变，为企业制定每周的配变巡视维护计划及年度配网技改项目安排提供了参考，有效减少了低效运行的设备。目前，该公司重过载配变预测准确率可以达到 80%~90%。

（2）国网公司通过大数据挖掘、聚类分析客户基本信息、用电量信息、业务办理信息等用电行为特征，建立客户用电行为分析模型，区分不同属性和行为特征的客户群，对不同特征的大客户提出差异化服务策略。

电力大数据理论研究和应用已取得了一些成效，但总的来看，研究成果仍比

较粗糙，不成体系，研究和应用尚处于起步和探索阶段。电力大数据的研究和应用是一个长期而复杂的工作。客观上，大数据的理论尚未成熟，大数据的相关技术仍在快速发展中，还没有进入稳定时期；同时，电网通信信息系统的互操作问题仍然存在，数据模型尚未统一，给数据的获取和应用带来实际困难；主观上，电力公司在大数据的基本概念、研究方法、应用价值方面认识不足，没有达成共识，在思想认识和技术准备上存在不足，也给大数据在智能电网中的应用造成一定障碍。即使有些电力公司和电力研究者对大数据表示出很大兴趣，但由于缺少战略性研究和顶层设计的指导，影响了电力大数据研究和应用的有序推进。

3、建立电力大数据算法及应用研究平台的必要性

3.1 大数据对电力的重要性

对于电力领域来说，要实现电力设备的数字化和智能化，就需要利用计算机软件技术、计算机网络技术、远程实时监测技术、远程诊断技术、通信技术等，建立起一套高效、稳定的电力大数据采集、监测、管理、分析与服务系统，从而为电网安全、可靠、经济、高效地运行提供保障。并且在大数据及云计算技术的支撑下，电能双向传输才能更有针对性，形成供需的动态平衡，目前常见的应用模式如下图，可以根据数据来源，将电力大数据管理系统分成用户管理系统、电网运行管理系统、企业端；相应地，电力大数据也有三种商业模式可以挖掘，即智能化节能产品、电力大数据服务平台和优化需求侧响应。



3.2 电力大数据的应用领域

参考近年来各国的研究，结合中国电科院各研究所开展的大数据应用需求分析，电力大数据目前重点在三个方面开展：一是为社会、政府部门和相关行业服务；二是为电力用户服务；三是支持电网自身的发展和运营，每个方向包含了若干技术领域，概括如下表所示。

方向	重点领域
服务社会、 政府部门和相关行业	社会经济状况分析和预测
	相关政策制定依据和效果分析
	风电、光伏、储能设备技术性能分析
面向电力用户服务	需求侧管理/需求响应
	用户能效分析
	客户服务质量分析与优化
	业扩报装等营销业务辅助分析
	供电服务舆情监测预警分析
	电动汽车充电设施建设部署
支持公司运营和发展	电力系统暂态稳定性分析和控制
	基于电网设备在线监测数据的故障诊断与
	状态检修
	短期/超短期负荷预测
	配电网故障定位
	防窃电管理
	电网设备资产管理
	储能技术应用
	风电功率预测
	城市电网规划

3.3 西安交通大学电气工程学院的基础条件和优势

西安交通大学是教育部直属全国重点综合性研究型大学，是国家“七五”、“八五”首批重点建设高校之一，“211工程”首批重点建设的七所大学之一，“985工程”首批重点建设的九所高校之一，“双一流”战略的36所一流大学A类建设高校之一。是中国九校联盟（C9）、中国大学校长联谊会、“111计划”成员，“珠峰计划”首批11所名校之一，教育部首批“卓越工程师教育培养计划”高校，国家2011计划“高端制造装备协同创新中心”牵头高校，“中俄交通大学联盟”成员高校。

西安交通大学电气工程学院是我国高等教育创办最早的电工学科。学院师资力量雄厚、实验设备先进，二级学科设置最为齐全。学院目前主体学科为电气工程国家一级重点学科，并涵盖控制科学与工程、仪器科学与技术两个一级学科。拥有电力设备电气绝缘国家重点实验室、国家工科基础课程电工电子教学基地、特种电气技术教育部重点实验室（B类）、智能电器教育部工程研究中心、国家能源局先进电网与装备可靠性及寿命评估技术重点实验室、陕西省智能电器及CAD工程研究中心、陕西省智能电网重点实验室、陕西省高压放电与等离子体工程技术研究中心、陕西省高电压大电流测试技术及装备工程实验室、平高集团-西安交通大学电力装备技术研究院、西电-交大电气技术研究院、大全集团-西安交通大学先进电气技术研究院等。

成立电力大数据算法及应用研究平台，响应了西安交通大学电气工程学院“电气+”和做增量的发展思路，也响应了电气工程学院创新港内涵建设工作的要求。电力大数据算法及应用研究平台的项目团队骨干成员，在诸多领域中颇有建树，主持或参与的纵向项目包括国家高技术研究发展863计划项目6项、国家自然科学基金项目（含重大、重点、面上项目）12项、国家科技支撑计划项目2

项、国家重点基础研究发展计划（973 计划）子项目 1 项、国家社科基金项目 2 项、教育部新世纪优秀人才计划项目 2 项、教育部博士点基金项目 2 项、中国工程院重大咨询课题 4 项、各类省级科技计划 10 余项。项目团队还承担了多项与大数据、数据挖掘、深度学习等有关的企业横向课题，涉及电力、铁路、烟草、电信、政府等行业和部门。项目团队还获得省部级科技进步奖 3 项，申请发明专利 30 余项，申请软件著作权 40 余项，发表各类学术文章近 200 篇，其中 SCI、EI 收录 80 余篇。

项目组还与大数据算法与分析技术国家工程实验室建立了合作关系，该实验室由西安交通大学于 2017 年 3 月联合清华大学、百度公司、国家电网全球能源互联网研究院、河南中原大数据研究院有限公司等组织机构共同成立，地点位于西安交通大学校内，实验室旨在解决我国大数据分析基础算法薄弱、大数据核心算法缺乏和大数据应用产品开发效率低等实际问题，凝聚了大量大数据分析技术研发与应用的高端人才，有着一定的人才和技术优势。同时，实验室围绕提升大数据分析能力、算法设计与应用水平的迫切需求，构建大数据分析算法与分布式计算试验与研发平台，在异构数据融合、大数据计算模式、大数据统计建模与分析方法、面向大数据的优化技术、大数据核心算法库和大数据算法评测与工程化等方面形成了可持续的产学研协同创新机制，有着独特的机制优势。

4、平台的研究目标

限于篇幅，此处仅结合本项目组研究方向，针对电力节能、配电、售电、变电、电力市场、知识管理等，介绍平台的研究目标。

（1） 电力节能

在电力节能方面，主要围绕基于大数据的绿色数据中心建设开展研究工作。根据《国家电网公司关于印发 2017 年信息通信新技术创新发展行动计划实施意见的通知》（国家电信通〔2017〕26 号）（以下简称“2017 实施意见”）安排，新技术应用为智能电网生产方式、企业运营模式变革带来新的契机，项目组与国网陕西电力数据灾备中心承担了基于深度学习的机房分区监控及节能研究项目，该项目入选了国网新技术创新发展行动试点项目，并在 4 月 20 日国网公司举行的论证会中得到了一致好评。该项目的研究目标是建立现有机房分区温度监控，并根据监控数据，利用大数据和深度学习神经网络进行建模仿真，为机房的空调系统运行提供控制策略，提高空调系统制冷效率，从而降低机房空调系统用电量。该项目积极响应 2017 实施意见和国网建设绿色数据中心的要求，也符合国家提倡的节能减排的要求，结合大数据和深度学习技术，提高数据中心空调系统制冷效率，降低数据中心用电量。



(2) 配电

电力系统中，为提高配电变压器利用率，降低线损，需要平衡各台区负载及台区三相负载。此外，为保证线损计算的准确性，供电管理部门需要经常普查台区用户信息，特别是用户台区和相位信息。然而，随着用电量的增加，供电公司的一些地区新增了变压器，电工在接线改动或均衡分配负荷线路改造时，连线信息混乱，造成采集器或电表与变压器从属关系记录不准确，用户信息与实际不符；另外，一些居民社区、沿街门面和部分商业街等场所，由于配电网台区信息更新不及时等原因，容易造成台区用户信息混乱、不准确甚至缺失，严重制约了智能电网建设进程；一些低压用户还私自搭接供电线，使得无法判定用户电能表属于哪台变压器。这些都带来了台区负载不均衡和相位不平衡的问题，而相位不平衡很容易造成用电高峰期间的风险，增加线路损耗和配电变压器的电能损耗，影响用电设备的安全运行，极易引发配变发热，严重时甚至会造成配变烧损，进而引发安全事故。以西安市为例，据调研西安市电力公司有近一半的居民用户的台区及相位无法识别，基于目前的紧迫形势，国家电网运检部 2017 年专门发布第 68 号文，开展配电台区三相负荷不平衡问题治理工作。

本项目组去年已开始与有关部门进行业务交流和技术试验，开展了基于大数据的台区及相位识别研究，并正在着手开展基于大数据的用户相位调整研究工作，以缓解基层部门在变压器负责不均衡方面的考核压力。



此外，项目组还正在开展基于大数据的配电网故障定位预研工作，利用大数据技术，配合故障投诉系统，融合 SCADA、EMS、DMS、D-SCADA 等系统中的数据作出最优判断，建立新型配电网故障管理系统，可以快速定位故障，应对故障停电问题，提高供电可靠性。

(3) 售电

目前与相关业务部门的业务沟通主要围绕基于大数据对电力用户提供精准服务，利用聚类模型等挖掘手段，开展用电行为特征深入分析，针对不用用户实施精准服务。主要思路是通过电表大数据等对电力用户进行精准画像，负荷曲线是客户消费行为的直观显示，智能电表数据能详细的记录客户消费细节，如间隔为 15min 的电量、功率、电压等参数。由于具有相似消费习惯的客户具有相似的负荷分布形状，因此可以根据其负荷分布的相似度来对客户加以分类，从而更加方便企业对客户的分类和精细化管理，并针对某类用户制定更加有针对性的运营策略。把用户实际的负荷曲线叠加到电网峰谷时段上，可以展现出用户更多的用电细节，计算其峰值时段的电能费用，估算用户错峰的潜力。因此利用智能电表数据对客户行为进行分析，能更有效地促进需求侧管理，合理抑制负荷峰值，提

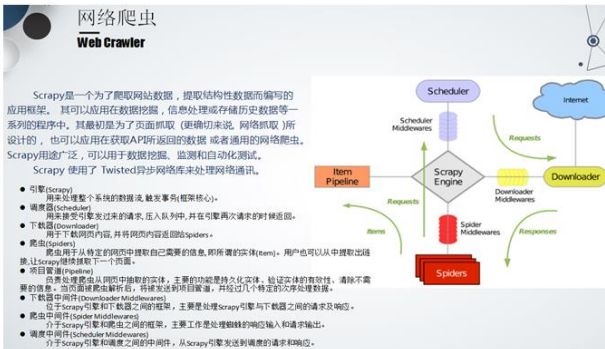
高电网资产的利用率。

项目组围绕售电舆情开展了一些预研工作。通过与微博、微信等互联网新媒体的服务对接机制收集海量用电信息、用户信息以及互联网舆论信息，利用大数据采集、存储、分析、挖掘技术，从互联网海量数据中挖掘、提炼关键信息，建立负面信息关联分析监测模型，及时洞察和响应客户行为，拓展互联网营销服务渠道，提升企业精益营销管理和优质服务水平。

项目组围绕防窃电管理也开展了一些预研工作，电力公司通过电量差动越限、断相、线损率超标、异常告警信息、电表开盖事件等数据的综合分析，建立窃电行为分析模型，对用户窃电行为进行预警；通过营配系统数据融合，可比较用户负荷曲线、电表电流、电压和功率因数数据和变压器负载，结合电网运行数据，实现具体线路的线损日结算，通过线损管理功能不仅可以知道实施窃电用户所在的具体线路，并且可以定位至某一具体用户，克服目前检查范围广，查处难度大的问题。

(4) 变电

变电环节的研究拟主要针对变压器等变电设备的运行及故障诊断开展大数据研究。变压器项目组目前正在西电集团变压器厂开展大数据项目的研究工作，主要目标是对变压器的市场需求做出精准预测，以便于企业进行精准的市场开拓。该项目目前进展顺利，变压器厂的相关负责人主动提出，下一步的合作方向希望在变压器的运行及故障诊断方面开展大数据研究的合作，与省电力公司合作，在实现 GIS、PMS、在线监测系统等各类历史数据和实时数据融合的基础上，应用大数据技术进行故障诊断，并为状态检修提供决策，可实现对电网设备关键性能的动态评估与基于复杂相关关系识别的故障诊断，为解决现有状态维修问题提供技术支撑。

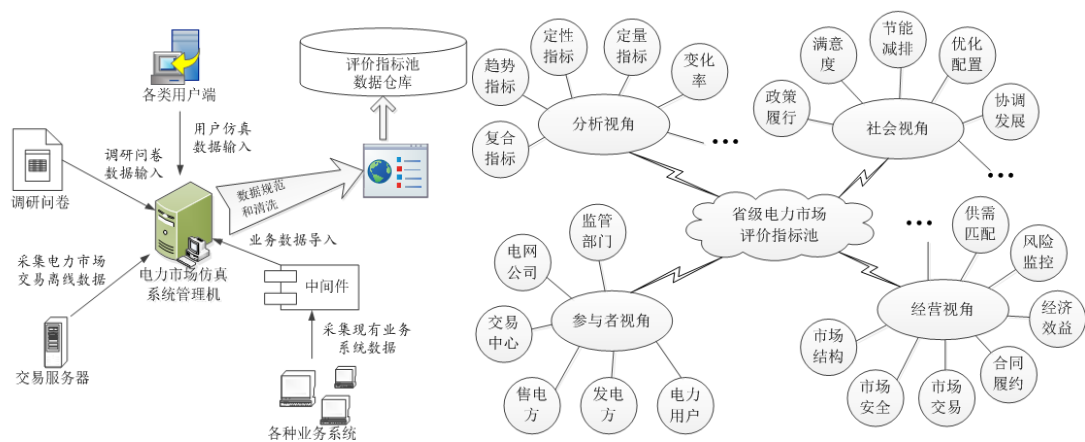


(5) 电力市场

随着我国电力市场化改革的不断推进，不少省份和地区已经开展了电力用户与发电企业直接交易、大用户与发电企业双(多)边交易、竞价上网等试点工作。

2016 年 5 月陕西电力交易中心有限公司成立，为陕西电力市场主体提供交易服务。国内外已有不少关于电力市场交易评价指标和评价体系的相关研究。由于国情不同，国外相关研究成果并不适用于国内。国内虽然有不少关于电力市场交易评价指标和评价体系的研究，但国内绝大多数研究成果提出的评价指标并不完备，不能满足各类参与者的分析评价要求，并且现有评价体系也局限于理论研究，距离实操还有一定的距离。此外，由于环境和地域差异，省级电力市场的差异化不断增大，基于其它省份的研究成果很难直接在陕西应用。

目前，与陕西电力交易中心有限公司的业务沟通主要集中在仿真系统的构建和指标体系的建设两个方面。通过仿真系统的开发，建立电力市场交易仿真系统，采集交易数据、相关业务系统数据和用户仿真数据，建立电力市场的大数据集，建立电力市场评价指标池，开发相关评价算法，为电力市场的各项业务提供技术支撑。



(6) 知识管理

课程组正在陕西省烟草公司进行基于大数据的知识管理算法、模型及应用研究，相关研究成果也可以应用于电力行业。

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。

