

# 学术大数据在企业专家对接中的应用

张永锋<sup>1</sup>, 霍东云<sup>2</sup>, 李振华<sup>1</sup>, 智强<sup>2,3</sup>, 李燕茜<sup>2</sup>

1. 清华大学软件学院, 北京 100084; 2. 北京赛时科技有限公司, 北京 100084;

3. 东莞理工学院科技资源大数据研究与成果转化中心, 广东 东莞 523808

## 摘要

基于创新驱动战略的大背景, 分析了企业和高校科研机构的专家在对接过程中面临的困难。以学术大数据技术为支撑, 搭建了产业领域专家数据集成系统, 并以此为基础, 建立了科技成果全链转化体系。在线上对企业需求进行收集与核实, 利用学者大数据画像技术进行精确匹配和推送; 在线下联系企业专家进行对接后, 持续提供项目管理、风险评估和科技保险服务。目前该系统已经收集了上万个企业的需求, 完成了多个企业专家对接项目, 提高了对接的精准度和效率, 实现了企业和专家的互利共赢。

## 关键词

学术大数据; 专家对接; 企业服务; 科技成果转化

中图分类号: TP311.13

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2019043

## *Application of academic big data in the connection of enterprises and experts*

ZHANG Yongfeng<sup>1</sup>, HUO Dongyun<sup>2</sup>, LI Zhenhua<sup>1</sup>, ZHI Qiang<sup>2,3</sup>, LI Yanxi<sup>2</sup>

1. School of Software, Tsinghua University, Beijing 100084, China

2. ScientistIn Co., Ltd., Beijing 100084, China

3. Science and Technology Resource Big Data Research and Achievement Transformation Center, Dongguan University of Technology, Dongguan 523808, China

## *Abstract*

Under the background of the innovation-driven strategy, the difficulties faced by enterprises and experts in universities in the process of connecting were analyzed. Based on academic big data technology, an expert data system for industrial fields was built and a full-chain transformation system for scientific and technological achievements was established. By collecting and verifying enterprise needs online, using scholars' portrait technology to accurately match and push, enterprise experts offline was contacted, project management, risk assessment and technology insurance services were continued to provide. At present, the system has collected tens of thousands of enterprise needs, completed a number of projects with connection of enterprises and experts, improved the accuracy and efficiency of connecting, and achieved mutual benefit and win-win between enterprises and experts.

## *Key words*

academic big data, expert connection, enterprise service, scientific achievement transformation

## 1 引言

随着大数据时代的到来,数据的重要性不断凸显,相比于传统数据库,大数据具有数据量大(volume)、数据种类繁多(variety)、处理速度快(velocity)、价值回报高(value)和数据真实性(veracity)的特点,即“5个V”<sup>[1]</sup>。随着全世界范围内科研活动的进行,每天都有数以万计的学术数据产生,这些海量学术数据(如专家学者信息、论文、专利等)被统称为学术大数据。

学术大数据在学术和科研领域有很多重要的应用。目前,企业和专家的对接存在许多问题和盲点,本文以学术大数据技术作为支撑,结合一对一的交易服务,解决了对接过程中企业技术转型、专家科研成果具体化、企业专家沟通等方面的难点和痛点。通过搭建产业领域专家数据集成系统、线上供需软件即服务(software as a service, SaaS)系统平台,建立科技成果转化全链体系,帮助专家和企业实现精准对接、企业的科研难题获得突破、专家的高新技术成功落地,实现双赢的局面。

## 2 学术大数据的研究方向和现状

近年来,学术数据呈现指数增长趋势,全球学术论文已超过3亿篇,学术工作者也已达1亿人。学术大数据技术的研究越来越受到高校和科研机构的重视。2015年10月,清华大学联合微软公司,成立了清华大学数据科学研究院·科技大数据研究中心,签署了清华-微软学术大数据项目合作备忘录。目前学术大数据的研究主要集中在学术图谱、学者画像和学术评价等方面。

### 2.1 学术图谱

知识图谱本质上是一种语义网络。其节点代表实体(entity)或者概念(concept),边代表实体/概念之间的各种语义关系。知识图谱融合了几种数据管理方式的特征,具体如下。

- 数据库:知识图谱可以像数据库一样进行结构化查询。

- 图:知识图谱可以像数据结构中的图一样进行分析。

- 知识库:知识图谱具有形式语义,并且可以用于解释数据和推断出新的事实<sup>[2]</sup>。

近年来,学术关系知识图谱(通常被称为学术图谱)的研究逐渐成为热点。2017年,微软公司和清华大学联合发布了学术图谱(open academic graph, OAG)。这个亿级图谱集成了微软学术图谱(Microsoft academic graph, MAG)和清华大学的AMiner学术图谱。具体来说,该学术图谱包含MAG中1.66亿篇学术论文和AMiner中1.55亿篇论文的元数据信息。经过集成这些数据信息,生成了两个学术图谱之间近6 500万链接(匹配)关系。2019年1月,在之前的基础上,2.0版本发布,OAG 2.0版本的更新在于它在论文数据的基础上,增加了与作者以及出版地点相关的数据,也就是说OAG 2.0版本包括论文、作者、出版地点3类实体以及其相应的匹配关系。

### 2.2 学者画像

学者画像是从学者学术研究信息中抽象出的一个标签化的模型。学者画像的基本目标是为每个学者建立档案。每个学者可以从多个维度进行描述,如个人信息、联系信息、教育经历等,通过学者画像可以更精确地描述一位学者的信息。

目前,对于学者画像的研究主要有以下几个方面。

- 构建高精度学者画像库:从异构数据源中提取科研数据,构建多维、立体的学术画像模型。

- 多维度标签化技术:解决科研人员、科研专家和科研行为的标签化方法,同时精确刻画其静态属性、动态行为和科研社群特征。

- 学者画像库的增量更新与溯源:通过增量更新,减少运算开销,同时保持数据的时效性。通过构建面向科研学者的溯源模型,在不断更新的过程中,始终保留数据溯源的机制,在时间层面保持数据的可靠性、可追溯性<sup>[3]</sup>。

## 2.3 学术评价

随着论文数据量的增加,单一的学术评价指标已经难以满足现今的学术评价要求。在大学排名上,就有QS世界大学排名、泰晤士高等教育世界大学排名、上海交通大学世界大学学术排名等多种方式;在个人的学术水平评价上,有论文数、被引数、H指数等一系列不同的指标。如何建立相对公正合理的学术评价体系,一直是国际学术界高度关注的话题<sup>[4-5]</sup>。

在大数据的背景下,涌现了很多新的学术评价的尝试。目前主要有以下几个方面。

- 基于学术图谱的评价方式:学术图谱绘制了巨大的关系网络,利用这些关系网络,结合数据挖掘和画像分析,得出一套评价方法。

- 借助深度学习的评价方式:深度学习相关技术的飞速发展使机器在很多领域表现出了比过去更优秀的性能,可以更好地帮助人们完成相应的工作。在对学者或大学的排序中,通过采集数据集,尝试不同方法训练神经网络模型,最终选出表现

最佳的模型作为评价方式,以求进一步优化学术评价。

## 3 专家和企业对接的问题和解决方案

### 3.1 问题描述

创新驱动战略是新时期我国全面深化供给侧改革的重要途径<sup>[6]</sup>。在创新驱动战略的时代背景下,我国科技成果产出大幅增加,但是质量和产业化的能力还较为欠缺;企业的创新能力不断增强,但是投入强度和技术能力还相对滞后;政府激励创新的体制机制和政策环境逐步改善,但仍难以满足创新驱动发展的需要<sup>[7]</sup>。在这样的大背景下,专家、企业、政府都面临着不同的发展困境。

- 在专家层面,专家普遍希望实现科技成果的转化和落地,但也存在困难:一方面,很多专家没有充足的科研经费,需要找一些横向课题作为支撑,难以把全部精力投入技术的研发与改进中;另一方面,专家精通技术,却不熟悉市场,很难在市场上找到有生产能力的合适企业进行合作,大部分科研机构的专家难以在企业的技术改造与创新过程中发挥作用,大量的科研成果被湮没。

- 在企业层面,企业面临着两大问题:一是原有的经营思维与当前需要的、适应新技术革命的研发思维之间的冲突,如何实现思维转化,迅速跟上时代技术潮流,是对企业的一个重要考验;二是创业者在新的环境下,要想实现企业的升级转型,必须有科技和创新作为支撑。

- 在政府层面,政府面临调结构、促改革的压力,需要主导经济结构转型升级,而经济结构转型升级、企业升级等都需要科技创新来驱动。

除了企业和专家自身存在的困难之外,

其沟通与合作同样有很多问题。第一,企业和专家之间难以架起相互沟通的桥梁。面对科技问题,若企业缺少专家技术的支持会陷入科技瓶颈,专家在筹备项目落地时也常因为没有对口企业而陷入僵局。第二,科研与生产的两套“话语体系”与运转机制的鸿沟使得某个生产场景下的具体技术问题难以被直接定义出来,需要有中间平台进行需求的深度加工。第三,技术交易的信用体系和定价机制的缺失使得企业和专家难以互相信任,在缺少中间人作为沟通桥梁和风险担保时,企业和专家难以实现互相信任,对推动进一步合作有一定的阻碍。

### 3.2 解决方案

针对以上困难和痛点,笔者利用学术大数据,通过打造科技资源数据库、搭建服务于一线科技部门的供需智能管理系统,提高对接工作的效率和效果,并提供全流程的在线服务及过程信息化。企业、专家对接系统总体可以分为以下几个部分。

#### (1) 学术大数据系统

笔者打造了专家数据集成系统作为数据支撑。搭建学术大数据系统是后续提供大数据科技服务的前提和保障。

#### (2) 供需SaaS系统平台

笔者打造了供需SaaS系统平台。通过线上收集企业需求,利用大数据技术进行线上匹配,实现企业和专家的高效对接。

#### (3) 科技成果全链转化体系

科技成果全链转化体系是连接产业研发需求与科技创新要素的服务体系,是基于产业发展的市场需求规律驱动科技成果研发的转化路径。它包括以企业为主体的产业技术需求、以转化平台为媒介的服务体系、以专家为核心的创新要素、以科技金融为保障的风险分担和投融资支持。

## 4 系统实现

通过分析我国政府、企业、科研人员、机构面临的科研创新困境,笔者构建了专家数据集成系统,与全国60多所高校和科研院所、50多个地方政府建立了合作关系,通过对接企业和专家的需求,为专家成果落地、企业技术需求、政府人才引进、科研绩效管理探索了一条有效的道路。企业专家对接系统的整体架构如图1所示,后文分别从专家数据集成系统、供需SaaS系统平台和科技成果全链转化体系3个方面讲述系统实现。

### 4.1 搭建专家数据集成系统

专家数据集成系统的研发目标是为后续平台提供面向科技领域应用示范环境的专家数据实时推送更新接口和产业领域专家精准推荐技术应用接口,为科技成果全链转化体系提供数据和关键技术支持。

系统主要通过数据爬取、数据清洗、画像和推荐,实现对专家的数据集成和推荐。其主要架构如图2所示,具体可以分为以下几个方面。

#### (1) 从多个来源整合专家数据

多个来源的数据包括:从互联网通过爬虫和清洗得到的大量互联网数据、通过对专利数据库的解析得到的大量专家数据、通过导入线下技术需求对接过程中的专家数据以及通过其他渠道(比如深度访谈、项目合作)获得的专家数据。

从互联网爬取专家数据是主要的数据获取方式。爬虫系统以Python、Java为基础开发语言,基于积累的一定数量的种子地址进行精准爬行和数据解析处理。爬虫的数据源非常广泛,包括百度学术、谷歌学术、ResearchGate、NSFC等。爬虫

首先由初始的统一资源定位符 (uniform resource locator, URL) 开始, 获取该 URL 的页面, 再通过该页面进一步获取后续的 URL, 并将其加入待爬取队列, 同时对已经爬取的页面进行存储, 直到满足结束条件为止。爬虫流程如图3所示。

在数据存储方面, 使用开源数据库 MySQL 作为数据存储数据库, 同时将阿里云对象存储服务 (object storage service, OSS) 作为大量网页文档资料的存储载体。MySQL 作为关系型数据库, 拥有数据库事务的原子性、一致性、隔离性、持久性 (atomicity, consistency, isolation, durability, ACID) 特性, 这些特性可以在海量数据的持久存储过程中保证数据的稳定和一致。阿里云 OSS 可以存放任何类型的文件, 而且有使用方便的应用程序接口 (application programming interface, API) 和很高的安全性, 成本低廉, 很适合在学术大数据存储中担任对象存储的角色。

#### (2) 对专家资源进行筛选和过滤

从互联网上爬取的专家数据量庞大, 涉及的专家科研水平参差不齐, 为了保证系统推荐专家的高质量和高准确度, 需要对专家数据进行筛选。筛选主要从成果数量、被引数量、H 指数、I10 指数等角度出发, 根据领域经验和规则, 对指标进行加权打分, 得到一个质量分。通过研究和分析质量分的分布, 可以将数据库中的所有专家分为初级、中级、高级、杰出4类。

在实际筛选和过滤的过程中, 将层次分析法 (analytic hierarchy process, AHP) 作为基础, 构建所需的数学模型。层次分析法是一种适用于具有复杂评价指标、目标质量分不便具体定量描述的评价系统的方法。层次分析法基于实际问题, 按层次分解总目标, 将总目标分解为多个层次进行评判。在同一层次中, 对筛选指标进行比较; 通过计算判断矩阵的特征向

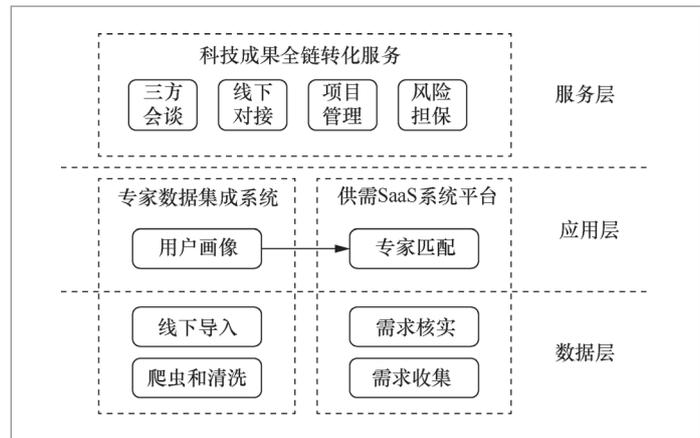


图1 系统整体架构

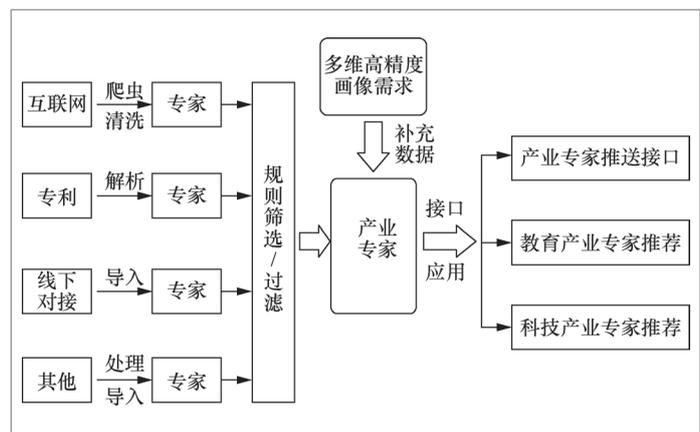


图2 专家数据集成系统架构

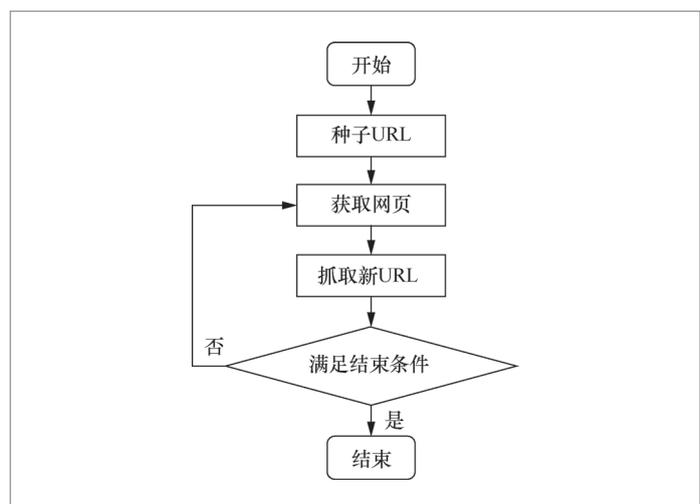


图3 爬虫流程

量, 确定每个因素相对于层次的权重; 最后通过加权和递归得到各个因素相对于总

目标的重要程度<sup>[8]</sup>。

具体而言,在对组织专家<sup>[9]</sup>、企业专家<sup>[10]</sup>的筛选过程中,对于特定领域的专家,有如下筛选过程:检索与专家相关的文档数据集;通过文档被引数量等指标计算其重要性得分,通过线性加和将所有文档得分相加,作为该专家得分;通过比较排序,确定专家的类别。由此保证推荐的专家质量更高,与需求契合得更加准确。

筛选和过滤可以凸显专家的类别和特征,针对不同类别和规模的企业实现不同的推荐方案,对专家推荐的准确性、企业的满意度、对接的成功率都有一定程度的提高。

#### (3) 形成多维度专家画像

依据多维、高精度专家画像的需求,对专家的数据进行各个维度的补全,建立基于标签的画像体系。以研究领域画像和合作网络画像为例,从项目课题、论文题目、论文关键词、论文正文词频等方面提取专家的主要研究方向和领域,对提出的所有领域信息进行去重、归类和筛选之后,形成研究领域画像;从论文合作者、合作机构等方面可以构造出专家的合作网络,进而提取出合作网络画像。

#### (4) 提供推荐和更新接口

推荐和更新接口是专家数据集成系统的主要功能,通过实现产业专家推送接口、教育产业专家推荐接口、科技产业专家推荐接口、专家信息获取接口和专家信息更新接口等,为供需SaaS系统和线下服务等提供数据支持。

## 4.2 建立供需SaaS系统平台

企业真实需求是科技成果转化的根本前提,科技成果转化最终要以市场需求为约束条件。企业真实的研发需求、技术发展瓶颈、行业共性难点都是促进科技研发的需求源。供需SaaS系统平台优化了企业

需求收集、技术点拆分、生产语言转化科研语言等一系列服务流程,解剖了研发需求结构,明确了产业行业需求列表,尤其是前沿重大领域攻关列表。

企业和专家供需对接的SaaS系统平台(以下简称平台)是基于海量学术大数据打造的,这些数据包括1 850万名专家、80万个科技类企业等。平台的主要运作流程包括需求征集、需求核实、智能匹配、精准推送,如图4所示。

平台通过与高校、科研机构、政府和企业进行合作,可以有效地扩充专家和企业数据库,助力专家企业高效对接。平台已覆盖清华大学、北京大学、西安交通大学、同济大学、中国科学技术协会、中国科学院计算技术研究所、中国科学院理化技术研究所等60多家高校和科研院所以及全国50多家政府机构。

#### (1) 需求征集

平台在线上征集企业需求,有专家对接需求的企业在系统中描述项目概况、技术需求、人才需求等信息,并明确给出技术研发与产品设计的瓶颈。为了吸引更多的企业尝试,需求征集系统的注册流程简单方便,门槛较低。

#### (2) 需求核实

由于系统向用户全面开放,企业注册门槛较低,因此需求的真伪性需要由专业团队进行核实和鉴别。专业团队利用企业提交的表单和常用关键字列表,提炼整理需求的关键技术信息和指标,将企业使用的“生产语言”进行简单易懂的“技术翻译”。同时,团队也利用企查查、天眼查等网站的接口完善企业资料。

#### (3) 智能匹配和精准推送

由于产业研发需求和高校院所的创新要素存在信息鸿沟,要把资源要素和解决技术问题的专家能力与企业研发难点精准匹配在一起,就要把产业研发需求和创新要素精

准对接作为突破口,平台中这一突破口就是智能匹配技术。借助大数据及人工智能的精准匹配功能,挖掘、吸引、整合全国乃至全球的科技资源,为企业链接技术专家,解决产业技术问题。通过对需求的分析和关键信息的抽取,在专家集成系统中寻找对应画像匹配的专家。平台对一个企业需求会匹配2 000名相关领域专家,并进行邮件、短信、站内信等形式的推送。一般地,推送后3天左右就会有专家响应对接。

### 4.3 提供科技成果全链转化服务

学术大数据技术实现了从“人找人”到“大数据找人”的转变,让企业和专家的高效对接成为可能。在大数据技术之外,平台提供额外的系统性服务,保障专家和企业的利益,减少技术交易风险。在平台的帮助下,企业、专家和平台可以进行线上三方会谈和线下对接,进而最终签署合同,完成项目交付。作为中间方,平台参与全流程,可起到保险和监督的作用。平台以市场化和专业化的服务支撑科技成果转化工作,需要从以下几点出发。

- 能够识别基于产业技术问题带来的潜在收益价值,为技术定价。
- 能够判别科技成果与技术研发在供给方和需求侧的风险状况,建立合理的风险共担体系,确定研发收益和知识产权保护的方案。
- 能够在科技成果转化过程中承担合同研发的项目管理工作,建立包括研发咨询管理、专家顾问和项目动态评估在内的服务支撑工作。
- 具备承担研发成品交付的管理能力,包括组织产品中试、寻找供应商进行方案的设备化管理、成品测试和交付产线等。
- 能够合理管控科技成果转化过程中的风险点,推动各方共同筹建市场导向

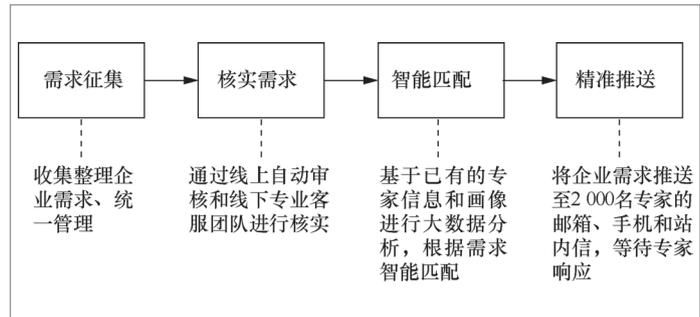


图4 供需 SaaS 系统平台

的研发资金池,组建研发风险创新担保基金,用于企业合同研发风险投资,开展合同研发风险担保。

以上的几点工作实现了成果转化、共享了知识产权、建立了全新的市场导向成果转化收益分配体系。提供科技成果全链转化服务需要从企业的需求中挖掘价值,并申请专利,持续运营,收益合理分配,建立以知识智力利益最大化的成果转化与收入分享系统,建立基于市场化的创新型的成果收益分享体系,激发调动积极性,将资源集聚到体系中,打造线上科技服务平台和线下研发机制,建立完整的成果转化全链服务中心。在科技成果转化服务的过程中,关键是要实现产业收益和研发增益,发挥市场需求在科技成果转化配置中发挥的决定性作用。在此基础上,形成具有社会效应和效益的创新驱动发展成效,完善研发合作后的项目管理,推动持续服务,追踪企业研发后的生产与销售情况,不断完善技术与升级,通过科技保险和担保基金直面交易风险,解决转化纠纷。

### 4.4 平台服务成果

在互联网和大数据的创新形式下,平台服务能力相较传统模式大大增强。截至2018年12月,平台已经为企业需求和多个政府人才项目引进需求服务近万次,匹配

和通知专家近百万人次,收到700余个创业项目需求,形成技术交易30余项,促成技术交易额近5 000万元。

在对2018年收集的9 266个需求进行线上核实的过程中,笔者发现一半以上的需求没有实际意义,或者没有实现的可能性,最终线上核实了4 432个需求。在线下联系阶段,经过电话、短信、邮件等方式与企业沟通后,得到真实需求659个,最终能够继续推进的有137个,占比仅为1.5%。在这137个需求中,每个需求会匹配2 000名专家。在这样一个信息不对称的背景下,供需SaaS系统平台能在短时间内寻找到对接技术供需的方法,通过“一对一”的技术咨询,为企业突破技术瓶颈、进行转型升级提供专项“问诊”,平台通过智力资源的优化和知识技术的溢出,解决生产技术瓶颈,助推企业创新,构建产业链上下游的全新生态。平台通过提高企业和专家对接的精准度和效率,实现企业和专家的互利共赢。

## 5 企业专家对接服务的应用案例

(1) 案例一:提高政府科技服务效率与效能

通过线上系统的精准匹配与线下对接活动的定向开展,笔者团队已多次成功与政府合作举办技术对接会,现场为企业解决问题类、研发类的技术需求。

2017年11月,笔者团队协助举办了院士专家集中对接台州市企业活动,50余位国内外专家为产业转型提供智力支持。在会议举办之前,系统收集、核实了当地62家企业的技术需求,并由专业团队进行需求的技术“问诊把脉”,核实出22个有效需求,共推送全国相关领域专家10 953位,最终31位专家积极响应,邀约到场12位专家。匹配专家推送完成后,系统同步推送信息

至科技供需系统平台,形成推送路径图。对接会举办后,笔者团队邀请专家分别前往11家对口企业进行现场考察,对其研发、生产线、物流等进行详细的技术诊断。在与企业深度对接后,为6家企业提供了8个技术方案,并签署了合作协议。

下面是其中一个成功对接案例。台州市某环保净化器有限公司在对接会前提交了技术难点。供需SaaS系统平台通过专业团队核实和实地走访调研,将需求指标进行量化处理,匹配出两位专家。在对接会上,该企业详细告知了现有的生产工艺、设备以及生产过程中的痛点。两位教授将企业实地数据与自身研发实力进行对比后,认为该技术难题可以解决,并先由专家方进行技术问题的“预研究”。目前,两位专家已为企业交付合格样品。

(2) 案例二:技术创新提升产品附加值

位于东莞市的某家企业有面膜生产的基膜分离设备,这套设备需配合自动折叠机使用。该企业的合作专家团队自研了自动折叠机,但在研制自动分离机时,遇到了技术瓶颈,由于基膜吸头技术要求高,造价昂贵,导致生产成本过高。

该企业向供需SaaS系统平台提交了需求,笔者团队做完市场调研后,明确了该技术创新应用的市场价值,决定开展此项目的研发,并通过科研大数据平台,为其筛选、匹配了专家。在专家团队的通力合作下,目前第二版原型机已研发成功。该技术还可以用于纺织行业,不仅可以提高生产的自动化程度,还可以大幅降低生产成本,因此市场前景非常广阔。

## 6 结束语

大数据时代,数据成为重要的生产要素和战略资源,数据的“5个V”特性,尤其是

数据价值显得越来越重要。企业和高校科研机构在学术活动中产生的海量数据成为学术大数据的一部分,而学术大数据的技术应用又可以促进企业和高校科研机构的学术创新和合作。目前,企业科研水平滞后、专家高新技术难落地的行业问题仍然广泛存在。笔者在应用学术大数据技术的同时,提供供需平台和服务,帮助专家企业对接,取得了很好的效果。未来,通过总结经验和积累技术,平台可以服务更多的政府科技管理部门、高新技术园区、项目孵化园区,形成更加精准、有效的科研人才及成果供给,帮助企业解决问题、实现创新驱动发展。

## 参考文献:

- [1] DEMCHENKO Y, DE LAAT C, MEMBREY P. Defining architecture components of the big data ecosystem[C]//2014 International Conference on Collaboration Technologies and Systems (CTS), May 19-23, 2014, Minneapolis, USA. Piscataway: IEEE Press, 2014: 104-112.
- [2] 刘峤, 李杨, 段宏, 等. 知识图谱构建技术综述[J]. 计算机研究与发展, 2016, 53(3): 582-600.  
LIU Q, LI Y, DUAN H, et al. Knowledge graph construction techniques[J]. Journal of Computer Research and Development, 2016, 53(3): 582-600.
- [3] 袁莎, 唐杰, 顾晓韬. 开放互联网中的学者画像技术综述[J]. 计算机研究与发展, 2018, 55(9): 1903-1919.  
YUAN S, TANG J, GU X T. A survey on scholar profiling techniques in the open Internet[J]. Journal of Computer Research and Development, 2018, 55(9): 1903-1919.
- [4] 周光礼, 武建鑫. 什么是学术评价的全球标准——基于四个全球大学排行榜的实证分析[J]. 中国高教研究, 2016(4): 51-56.  
ZHOU G L, WU J X. What is the global standard for academic evaluation, empirical analysis based on the four global university rankings[J]. China Higher Education Research, 2016(4): 51-56.
- [5] 叶鹰. 国际学术评价指标研究现状及发展综述[J]. 情报学报, 2014, 33(2): 215-224.  
YE Y. The research progress and developing perspective of assessment indicators[J]. Journal of the China Society for Scientific and Technical Information, 2014, 33(2): 215-224.
- [6] 黄志雄. 科技创新补贴供给侧改革与企业研发策略研究[J]. 经济理论与经济管理, 2018(12): 57-69.  
HUANG Z X. Technology innovation subsidy supply-side reform and firms' R&D strategy[J]. Economic Theory and Business Management, 2018(12): 57-69.
- [7] 吕薇, 马名杰, 戴建军, 等. 转型期我国创新发展的现状、问题及政策建议[J]. 中国软科学, 2018(3): 10-17.  
LYU W, MA M J, DAI J J, et al. Research on present situation, problems and policy suggestions of China's innovation development in the context of transition[J]. China Soft Science, 2018(3): 10-17.
- [8] 李光红, 杨晨. 高层次人才评价指标体系研究[J]. 科技进步与对策, 2007, 24(4): 186-189.  
LI G H, YANG C. Research on evaluation index system of high-level talents[J]. Science & Technology Progress and Policy, 2007, 24(4): 186-189.
- [9] 陆伟, 韩曙光. 组织专家的检索系统设计与实现[J]. 情报学报, 2008, 27(5): 657-663.  
LU W, HAN S G. Design and implementation of organization expert search system[J]. Journal of the China Society for Scientific and Technical Information, 2008, 27(5): 657-663.
- [10] 陆伟, 赵浩镇. 基于文档权重归并法的企业专家检索[J]. 现代图书情报技术, 2008, 24(7): 38-42.  
LU W, ZHAO H Z. Enterprise expert search based on combing document weight[J]. New Technology of Library and Information Service, 2008, 24(7): 38-42.

## 作者简介



张永锋 (1994-), 男, 清华大学软件学院硕士生, 主要研究方向为云存储、网络信息爬取等。



霍东云 (1981-), 男, 北京赛时科技有限公司联合创始人兼首席技术官 (CTO), 主要研究方向为大数据、云计算、移动互联网等。



李振华 (1983-), 男, 博士, 清华大学软件学院副教授、博士生导师, 主要研究方向为云计算、云存储、移动互联网等。



智强 (1983-), 男, 北京赛时科技有限公司创始人兼首席执行官 (CEO), 东莞理工学院科技资源大数据研究与成果转化中心主任, 主要研究方向为创新过程、科研管理与科技政策。



李燕茜 (1989-), 女, 北京赛时科技有限公司联合创始人兼首席运营官 (COO), 主要研究方向为技术创新、政府管理、产业政策等。

收稿日期: 2019-06-10

通信作者: 霍东云, hdy@scientistin.com

基金项目: 国家重点研发计划基金资助项目 (No.2018YFB1004700); 东莞理工学院高层次人才 (创新团队) 基金资助项目 (No.KCYCXPT2017001, No.KCYXM2018001)

Foundation Items: National Key Research and Development Program of China(No.2018YFB1004700), High Level Talent (Innovation Team) Project of Dongguan University of Technology (No.KCYCXPT2017001, No.KCYXM2018001)

2019043-10