

苏杨:如何运用大数据进行精准高考志愿填报

□文/白翎

编者:随着云计算、大数据、人工智能等新一代信息技术迅猛发展和广泛应用,数据资源在国家安全与发展各领域的地位作用日益凸显,大国之间博弈和较量正在由传统领域向数据空间扩展。我国大数据战略持续推进,聚焦数据价值释放,数据正成为重组要素资源、重塑经济结构、改变竞争格局的关键。近日,《秦智》杂志有关人员与西北大学中国西部发展研究中心大数据研究院院长苏杨就大数据发展和鸡维数据的建设及应用等相关问题进行了深度交流,苏杨的探索和实践中有许多闪光点,值得学习和借鉴。

《秦智》:西北大学中国西部发展研究中心大数据研究院成立已经两年多了,请介绍一下研究院建立的初衷和发展的宗旨?

苏杨:自信息技术、互联网产业兴起以来,“数据”在生产要素中的地位越来越重要,目前大数据已经演变为经济社会发展和变革的重要引擎,特别是在国家《“十四五”数字经济发展规划》和《“十四五”大数据产业发展规划》的大方向指引下,我们认为大数据研究这个方向大有可为,因此在2021年底成立了大数据研究院。成立初期主要是为了做数据研究和数据服务,为我们西部发展建设乃至全社会的各项研究提供强大的数据支持。

大数据研究院希望推动大数据产业不断走向规范化、标准化、科学化,还需要拥有赋能传统产业数字化转型升级发展的思路。凭借多年来对科技产业发展的深刻洞察,以技术和资源为支撑,大数据研究院从早先驱动中国研究院升级转型中完成全面数据研究的拓展,并围绕数据挖掘、数据存储、数据分析、数据应用与数据安全等环节,全力打造大数据全链能力,构建完整的数据应用能力,为政务、教

育、医疗乃至全产业等各领域用户带来全维度服务解决方案。

在2021年大数据研究院成立的发布会上,我们用“决策赋能·预见未来”八个字对大数据研究院的宗旨进行了概括。这八个字代表我们研究院的愿景,就是要通过打造可靠的数据源服务平台,为各领域提供决策支撑,从而让我们的服务对象能够从数据结果中找到各领域的变化规律和趋势,从而达到“预见未来”的目的与效果。

《秦智》:大数据研究院在2022年初推出的鸡维数据平台,除了提供公共数据收录查询服务,我们看到鸡维数据发布的一些榜单和研究报告在业内外引发了不小反响,请问大数据研究院对鸡维数据平台的定位和发展方向是怎样的?

苏杨:大数据研究院存在的意义就是要通过数据能力为我们的受众提供和创造价值。鸡维数据就是我们对外能够承载更广泛服务的商业化品牌。

鸡维数据是专注于数据查询、数据挖掘、数据分析研究和数据算法设计等综合型大数据服务平台。在平台上,用户可以通过在线阅读等方式进行数据的查询和对比,同时平台还支持数据分析、交易及下载等服务。

目前鸡维数据平台囊括了众多类型的数据。其中最具特色的是地方数据,无论是从时间年限和指标体系上来讲,鸡维数据的地方数据涵盖面几乎是目前各平台中最全面的。在社科类数据方面,平台还包括了我国省(市)、地级市、县区和乡镇等各个行政等级的数据,同时也涵盖了全球各国的近百项主要指标数据。

另外,鸡维数据在不断更新海量的行业数据,比如教育数据、企业数据、医疗数据、汽车数据等等众多领域

的产业信息数据,其数据体量已经达到了百亿级,并为企业级客户提供定制化数据采集与分析、API(应用程序编程接口)接口服务等多方面的服务。

除了丰富的数据品种,鸡维数据还提供强大的各种数据工具,便于用户进行数据对比和处理。比如在我们的高校数据库里,用户能够使用鸡维数据的工具进行学科专业的排名筛选。今年我们上线的高考志愿填报助手,这套系统融合了先进的大数据和AI技术,助力高考学生少走弯路,以最佳策略和超高的效率轻松实现高考志愿的精准填报。

目前,我们依托鸡维数据平台也同样获得了不少研究成果。比如《西部城市数字竞争力报告》这种社科类专业学术报告,还有基于创新算法的全国城市综合竞争力指数排名。近年来我们在网络热度监测研究上做出了一些开创性的成果,陆续发布了网红城市、网红县城或者名人热度等多方面的热度监测榜单,一些榜单成果还被地方政府做了引用,这些榜单在文旅经济、流量经济研究盛行的当下,对提升研究效率以及提升产业效益都带来较大裨益。另外,前不久我们还发布了全新的鸡维数据版2024中国大学排名,融合了我们在大数据研究方面的创新性成果,在教研等方面赢得显著关注。

总的来说,鸡维数据是开展数据服务和科学研究的载体平台,在发展方向上鸡维数据会一直聚焦到数据服务和数据研究这两个层面上。

《秦智》:已经有很多机构在做高等院校评价的相关排名,据了解这也是一项非常复杂的系统研究工程,鸡维数据为什么也会切入这个课题做中国大学排名?鸡维数据版的中国大学排名有什么特点?

苏杨:首先需要说明的是,鸡维数据在高校评价排名方面并不是跟随者,去年我们率先推出全国民办高校百强榜就是一个很好的例证。

事实上,我国目前高等教育的发展已经进入了新的阶段,也需要更多数据研究作为支持,因此我们非常重视高等教育数据。在大数据研究院成立之初,我们就把高校数据作为一个重点领域进行了部署。其中,鸡维数据版中国大学排名的相关数据采集研究工作在前2年就已经立项。

我们在对高校数据进行采集的初期,就发现此前不少大学排名榜单存在缺乏实用性和参考性等问题。比如有的机构在进行普通本科高校的综合排名时将理工类和文史类高校混排,导致排名失之偏颇;有些看似已经分类的排名榜单,但受众并不能直观获悉学科的对标情况,导致排名意义不足。

为了更加客观的研究与评价,我们研究过程中首创引入对同一所高校的“文理”两大核心学科实施独立评估的机制,分编入“理工百强”与“文科百强”两大榜单。

除了在体系上的创新,我们在研究方法上也进行了创新。在这次高校评价数据研究中,我们引入了人工智能大模型工具,辅助进行数据清洗和分析,并构建了多模型反复验证,不仅大大提升了研究效率,也提升了数据质量和评价结果的客观性。

另外,作为高校数据研究的延伸,鸡维数据还重磅推出了高考志愿填报助手,可以通过大数据和AI手段,帮



苏杨

助高考考生快速分析大量院校信息、专业信息以及历史录取数据等,并通过构建预测模型,预测各院校的录取分数线、录取概率等,为考生提供有价值的参考。

同时,我们的高考志愿助手还可以对考生的兴趣爱好、职业规划以及高校信息对比分析等方面进行综合评估,定制最适合考生自己的志愿方案,大大减轻了考生和家長收集研究信息的负担,显著提升填报效率和质量。

从服务属性不难看出,大数据不仅仅是为了服务于教育产业从业者,同时也应该潜心服务于广大受教育的群体,所以我们认为,这是鸡维数据在推出高校评价、志愿填报指引等服务的出发点上,相较于其他机构最为显著的不同之处。

《秦智》:“十四五”规划纲要中明确提出要加快数字发展,建设数字中国,迎接数字时代,激活数据要素潜能。在进一步促进大数据领域新技术理论成果应用、推动形成群智众创的大数据建设生态方面,您还有哪些见解?大数据研究院还将在哪些重点领域关注和突破?

苏杨:数据要素被视为21世纪的“黄金”和“石油”,是新时代的重要生产要素。《“十四五”大数据产业发展规划》中明确,要释放数据要素价值,培育数据要素市场,并作为产业基础能力提升的着力点。

我认为,大数据研究院通过鸡维数据,恰恰正在实践和探索数据要素的价值释放。现在,用户在鸡维数据平台实现了规范可控的交易流通,而数据采集和数据安全则是我们的基础。在这些基础上,我们去挖掘数据要素更广泛的价值,进行大数据研究提供赋能是核心。

作为科技行业的资深从业者,我一直看好数据作为生产力要素的大方向,不断更新的大数据产业发展政策给了我们更大的动力和勇气。也是基于此,我认为大数据研究院将重点要在三个方面突破:

首先是在数据科学基础研究方面。大数据研究院要进行大数据领域的基础科学问题和关键技术问题的探索性研究,提升大数据技术的基础研究能力。比如,进行更有价值和影响力的开创性数据科学研究,或者通过与国内外顶级机构合作,发表高水平的研究成果,提高在国际上的学术影响力。按照十四五规划方向,大数据研究院要通过提升基础科学研究能力,成为基础制度或标准规范指定的参与者或示范者。

(下转4版)

鸡维数据
oweidata.com

2024中国大学排名系列百强榜

理工类·文科类·民办类·高职类



突破性首次按照理工类和文科类两大方向分别排名,针对两类学科采用差异化考评算法和指标权重,让结果更具参考性和实用性



引入AI大模型辅助计算推演
采集分析超亿条基础数据
全类型评价变量涉及指标超百个



查看榜单排名欢迎登录鸡维数据官网(www.oweidata.com)或搜索关注鸡维数据微信公众号/百度/抖音/快手/微博等平台账号

2024 鸡维数据中国大学排名
RESEARCH BY www.oweidata.com