

科技数据统计分析平台,采用分布式网络数据采集系统为某省科技厅实现企业库、政策库、项目库的建设,从而建设科技数据统计分析平台。

首页 / 项目案例 / 正文



科技数据统计分析平台

采用分布式网络数据采集系统为某省科技厅实现企业库、政策库、项目库的建设,从而建设科技数据统计分析平台。



案例展示

项目背景

随着“大数据”的发展,科技管理工作呈现出管理主体多样性和科研需求个性化的趋势。在“数据”本身成为一种公共资源的情况下,科技管理需要一种创新的模式来满足时代发展的需求,“数据”将在这种发展中起到支撑作用,为科技管理工作和社会科研工作提供分析、决策与评估依据,进一步提升全社会的科技创新能力。同时,“数据”在实现政府决策科学化,推动政府治理能力现代化方面也发挥着重要作用。

该平台的建设将加快信息化与科技业务的深度融合,有利于全区经济社会发展与改革要求,有利于科技管理工作更加高效,有利于优化科技服务水平。平台建成后推动形成基于大数据的科学决策,实现精细化科技管理,增强科技创新引领能力。

政务数据资源是国家数据资源的重要组成。加快推动政务数据共享工作,是转变政府职能、创新社会管理的重要举措,有利于全面释放数据红利。《自治区加快推进“数字政府”建设工作方案》中明确要求加强数据资源规划、采集、存储、共享、开放,加强数据治理,提升数据质量和价值,构建全区共建共享的大数据资源体系。

需求分析

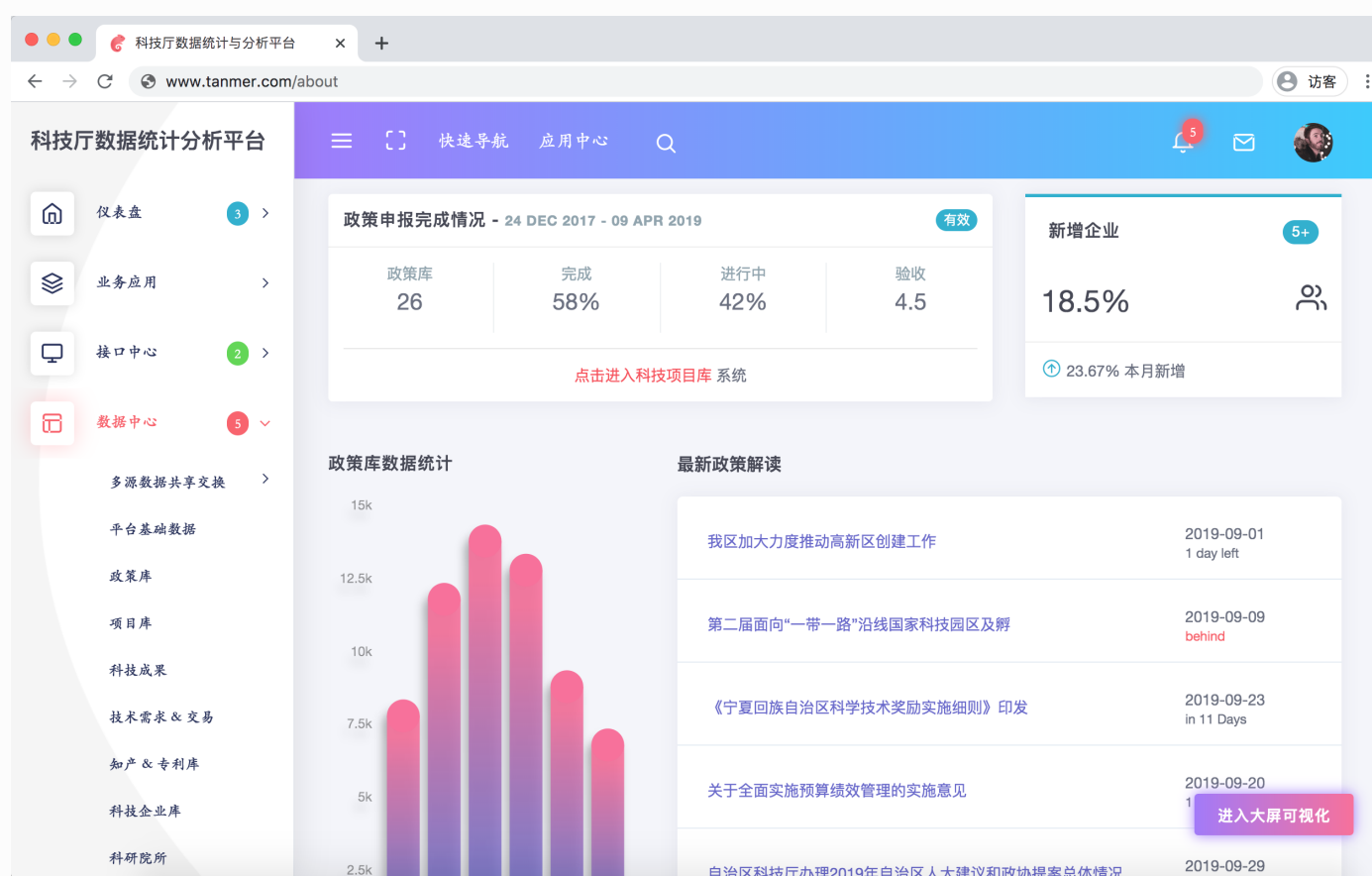
需求对象:各省市州区县科技信息化中心、经济信息中心、双创办、科技成果转化处

需求目标:

- 深入了解本省市州区的科技管理信息系统和科技统计工作的业务特性,全面梳理基础数据,整理全套规范样表;
- 分析各项指标,理清数据关联性,制定数据清洗规范;
- 建设科技数据平台框架,整合不同类型数据源,打通业务系统数据接口,建立数据汇聚通道;
- 通过构建统一的科技数据交换、存储、共享、应用接口等标准体系,构建汇聚式一体化的科技数据仓库。

项目展示

探码结合自身优势,充分发挥在大数据领域数据采集、数据分析、数据治理方面的数据积累和产品应用能力,调研100多个数据源平台,采集上万条科技政策数据,聚合6万多企业工商基础信息,成功完成了科技数据统计与分析平台。



科技厅统计与分析平台Demo截图

本次项目深入了解科技管理信息系统和科技统计工作的业务特性,全面梳理基础数据,整理全套规范样表。分析各项指标,理清数据关联性,制定数据清洗规范。建设科技数据平台框架,整合不同类型数据源,打通业务系统数据接口,建立数据汇聚通道。通过构建统一的科技数据交换、存储、共享、应用接口等标准体系,构建汇聚式一体化的科技数据仓库。



六大子系统,十四个功能模块

本项目涉及到科技厅内外部各种多源数据,需通过“数据共享交换系统”实现:数据共享交换、目录管理服务、共享数据管理、共享业务管理。具体有以下几个方式实现:

- 基于SQL接口、XML、Web Service、消息中间件等技术,实现与关联系统的数据交换接口开发。
- 实现与科技厅内部已建成的信息管理系统或数据库系统的接口开发和互联互通。
- 平台还需预留与自治区政府(如区政府信息中心、区数据资源局)、省科技厅、市州区科技管理部门的业务系统互联互通接口(待条件成熟时,便于实现各级科技业务系统实现纵向的数据交换、资源共享、互联互通)。
- 后期根据发展需要,还需与横向其他委办局(如财政、工商、税务、监察、人社等政府部门)的相关系统实现数据对接。
- 通过爬虫技术,实时爬取和关联外部数据,保证科技数据仓库数据的实时性。

技术解析

探码分布式网络数据采集系统介绍:系统实现数据从采集,处理到应用的全生命周期管理;另类数据,网页解析及采集自动化。针对互联网数据抓取、处理、分析、挖掘,并灵活迅速地抓取网页上散乱分布的信息,通过智能数据中心提供存储与计算,利用网页应用服务器和开放平台服务器进行大数据存储、管理以及挖掘服务,平台服务器居中调节,实现大数据的智能化分析,准确挖掘出所需数据。

通过详细的需求分析,该省科技厅数据统计分析平台其核心目的是建设科技数据仓库,采用探码分布式网络数据采集平台,建设了政策库、企业库、项目库等科技数据仓库,基于科技数据仓库的数据实现以下建设内容:

- 政策库建设

通过探码采集系统采集本省各类的科技申报政策,储存建立庞大的本省科技政策库。

- 企业库建设

通过对企业公开数据的采集,以及调用权威工商平台的API接口数据,建设含有企业工商信息、经营状况、知识产权信息、法律纠纷、品牌传播等信息的本地区企业库。

- 政策分类筛选解读管理

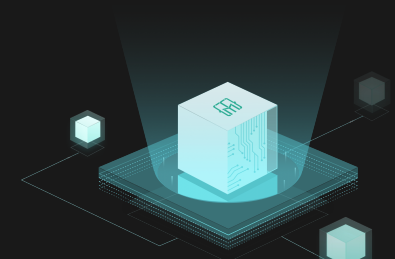
通过人工的初步分析,制定相应的算法对政策数据进行分类。同时结合平台智能文字处理系统对采集到的政策信息逐条进行解读,保存政策的基本条件(包括企业资产、所在地区,法人,所属产业,经营范围),用于企业和政策的基本匹配。分级保存政策的一级,二级,三级条件,逐步引导用户完善企业信息,更准确地匹配相关政策,给企业的决策提供最可靠的帮助。

- 政策和企业的匹配管理

企业与政策基本属性关系表,保存公司通用的基本属性,如注册资金,注册地址,注册时间等关键属性。基本关系表在公司或政策信息进入数据库的同时自动完成填写,并根据公司或政策的相应变动。

- 系统控制台管理模块

系统控制台管理模块完成该科技厅数据统计分析平台的启动、停止、备份、恢复,以及重新启动、注销登录用户等功能,实现对企业库、政策库、项目库的管理。



快速导航

大数据业务

大数据知识

合作案例

项目导览

信息流

机器学习

数据模型

社区招聘

官方博客

正在招聘

联系我们

信息检索

探码之名,源自探码,指先锋。探码网络大数据采集,SaaS软件开发,企业信息化智能化解决方案!

输入关键词

GO