

TRS SMAS聚网大数据分析平台

TRS SMAS BigData Integrating & Analyzing Platform

基于云服务模式实现对互联网信息的智能洞察

目 录

CONTENTS



- 一、大数据时代信息管理的机遇 1
- 二、大数据时代信息管理面临的挑战 2
- 三、大数据时代选择信息类产品的三大标准 3
- 四、拓尔思大数据方面的优势 4
- 五、TRS全生命周期信息管理 6
- 六、典型案例 12
- 七、客户名录 16
- 八、功能架构及特性 17
- 九、系统架构及关键技术 23
- 十、TRS SMAS舆情事件库 24

一. 大数据时代信息管理的机遇



每天，6.88亿中国网民通过人和人的互动，人和平台的互动，平台与平台的互动，实时生产海量数据。这些数据汇聚在一起，就能够获取到网民当下的情绪、行为、关注点和兴趣点、归属地、移动路径、社会关系链等一系列有价值的信息。

数亿网民实时留下的痕迹，可以真实反映当下的世界。原本分散、孤立的信息通过分析、挖掘具有了关联性，激发了智慧感知，感知用户真实的态度和需求，辅助政府在智慧城市，企业在品牌传播、产品口碑、营销分析等方面的工作。



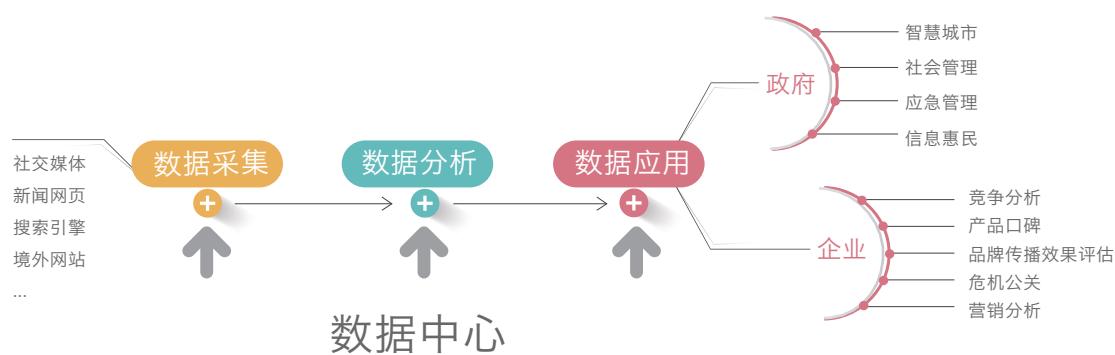
大数据时代的舆情管理不再局限于危机解决，而是梳理出危机可能产生的各种条件和因素，以及从负面信息转化成舆情事件的关键节点和衡量指标，增强我们对同类型舆情事件的认知和理解，帮助我们更加精准的预测未来。

2012年，拓尔思基于TRS SMAS聚网大数据分析平台，对2013年厄瓜多尔的总统选举做了预测，采集多方西语数据，以及各方面的数据结果，成功预测到了当选人为拉斐尔·科雷亚（Rafael Correa），且预测结果与最终的选举结果仅差1.3%。



无论是政府的公共事务管理还是企业的管理决策都要用数据说话。政府部门在出台社会规范和政策时，采用大数据进行分析，可以避免个人意志带来的主观性、片面性和局限性，可以减少因缺少数据支撑而带来的偏差，降低决策风险。通过大数据挖掘和分析技术，可以有针对性地解决社会治理难题；针对不同社会细分人群，提供精细化的服务和管理。

大数据时代信息管理生态图



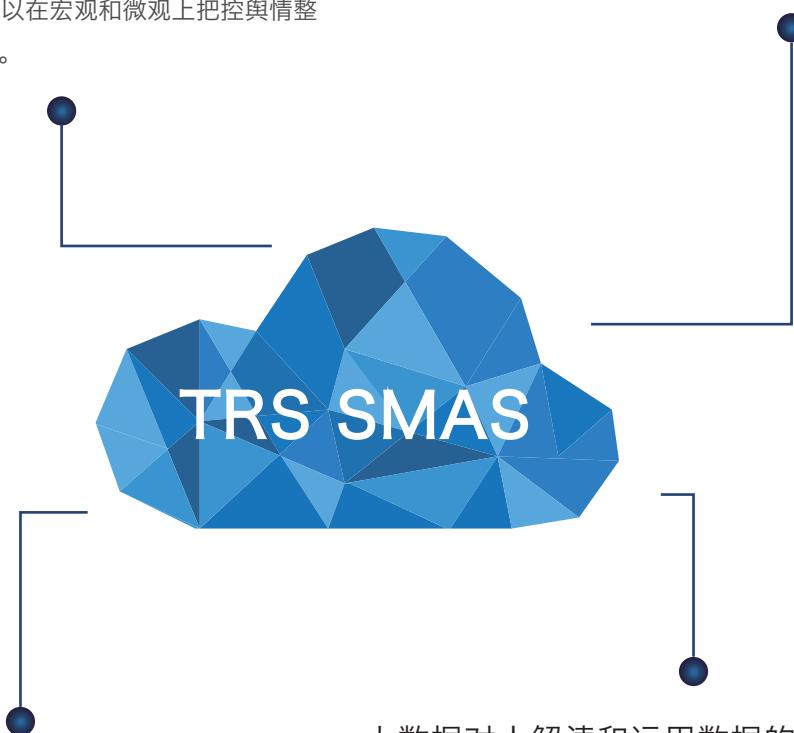
二. 大数据时代信息管理面临的挑战

海量非结构化数据，分析难度大

大数据时代数据的结构正在发生变化，包括网页数据、文档数据、图片数据、音频数据和视频数据在内的非结构化数据的增长远远超过结构化数据。传统的数据挖掘算法多是基于封闭的结构化数据进行挖掘，对于半结构化或者非结构化数据无能为力，难以在宏观和微观上把控舆情整体规律、变化趋势。

自建大数据中心，成本高昂

数据中心是大数据采集、处理、分析的“工厂”，舆情管理离不开数据中心的支持。但是数据中心动辄千万甚至数十亿的高昂投资额，使得政府和企业用户望而却步。没有数据中心的支持，大数据舆情分析也就无从谈起，面对这一挑战，政府和企业将何去何从？



秒传播时代，及时响应要求更高

随着微博、微信、新闻客户端等基于移动互联网技术开发的应用的普及，舆情传播速度更快，已经进入秒传播时代。而且，在“人人都有麦克风”的时代，任何一个的关注，都可能借由网民、意见领袖、媒体呈高速病毒式传播，从而演变成重大舆情事件。这就对及时捕捉和发现舆情的弱信号提出了更高的要求。

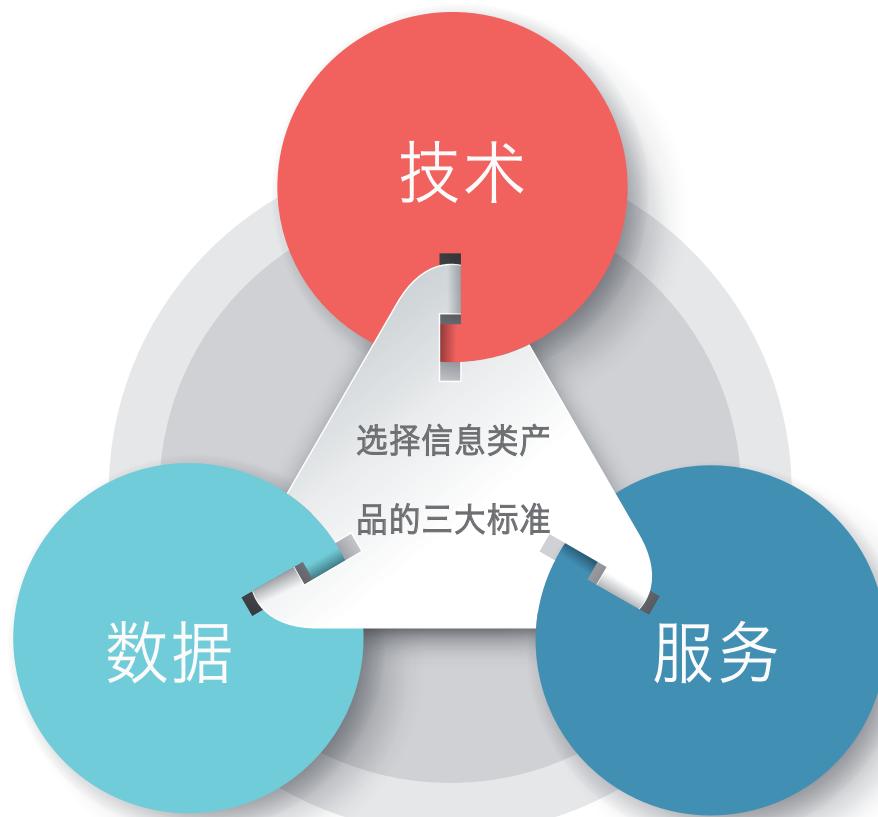
大数据对人解读和运用数据的能力要求更高

再智能的计算机软件也难以匹敌经验丰富的专家人脑。一方面，舆情分析、研判、预警属于前沿探索技术，软件的准确率需要一定的时间和突发事件检验；另一方面，舆情隶属于社会科学范畴，诸如社会心理、网络情绪等纯主观舆情指标，很难分解为具体可量化指标。这些因素导致目前阶段舆情系统必须辅以一定的专家人工服务，尤其是在舆情预警、研判方面，能够提高舆情管理的效率和水平。大数据时代，大数据对人解读和运用数据的能力要求更高，人的能力和价值将更加凸显。

三. 大数据时代选择信息类产品的三大标准

面对机遇与挑战，拓尔思提出大数据时代选择信息类产品的三大标准

——数据、技术、服务



无大数据中心，难以看清世界的真相和细节。



无大数据采集、分析、存储的核心技术，难以看清数据之间的关联及背后的价值。



无大数据服务的实施经验和成功案例，难以做到跨界融合，易流于纸上谈兵。

四. 拓尔思大数据方面的优势

1、数据优势

拓尔思拥有自建独立的大数据中心，数百台服务器对新闻、论坛、微博等多种类型互联网数据进行7*24小时不间断实时采集，具备上千万数据量的数据索引、挖掘分析和存储能力，可支撑政府、企业、媒体、金融、公安等多行业用户的舆情分析云服务。TRS数据中心监测采集源上百万个，日采集和处理数据超过5000万条。数据中心采集的覆盖范围每天还在不断加大。



2、技术优势

拓尔思作为国内领先的大数据技术厂商，长期专注于非结构化信息处理，坚持自主研发，致力于大数据领域的技术创新和产品研发。

拓尔思凭借非结构化信息管理和海量数据处理方面的突出优势，推出TRS海贝大数据管理系统。TRS海贝大数据管理系统提供了一体化的数据仓库构建手段：实现大数据存储、管理和检索的集成平台，可以实现结构化数据、半结构化数据、非结构化数据的统一管理和检索，满足了“非结构化数据的结构化处理、结构化数据的非结构化处理”的技术趋势。以TRS海贝大数据管理系统为支撑，拓尔思还研发了机器日志挖掘引擎，大数据关联关系挖掘引擎等多款产品，配合TRS网络信息雷达，TRS数据处理工具（类似ETL），构成完整的大数据解决方案。

拓尔思是国家认定的高新技术企业、软件企业和中关村国家自主创新示范区创新型试点企业，连续7年被认定为国家重点规划布局内软件企业，并被中国软件行业协会评为中国十大创新软件企业，产品获得国家科技进步二等奖、上海市科技进步一等奖等多项奖励。

拓尔思拥有北京市级企业技术中心、北京市知名品牌和北京市著名商标、公司获得的主要资质还包括ISO9001质量体系认证、CMMI L3级认证、系统集成三级资质、总装备部武器装备生产许可证、北京市纳税信用A级企业、软件行业协会AAA企业信用等级等。

拓尔思研发实力强大，承担多项国家级重大项目，如核高基、863、国家支撑计划、工业和信息化部电子发展基金、发改委信息安全专项等。

2011年，TRS云搜索平台、TRS舆情云服务平台被评为20项中关村云计算重大创新技术产品。

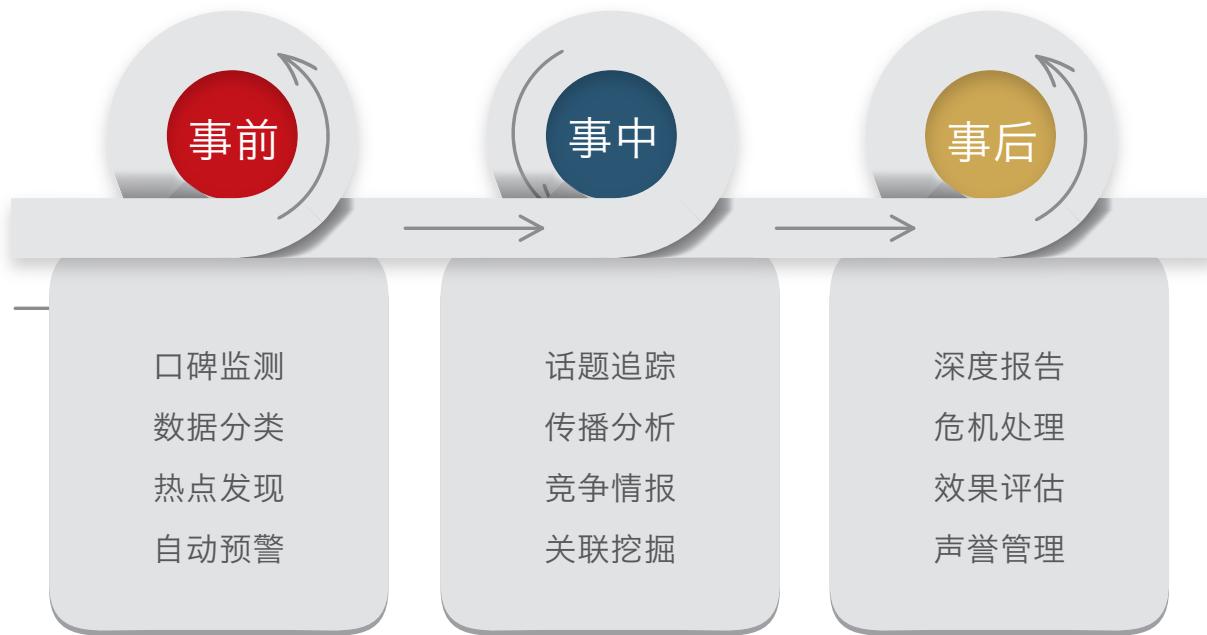
3、服务优势

拓尔思拥有舆情管理全生命周期产品线，在大数据采集获取、企业搜索、数据交换、数据分析挖掘和身份管理等各节点拥有底层技术和成熟的软件产品。TRS产品被国内外各行各业约4000多家企业级机构客户广泛使用，覆盖了众多国家部委和地方政府部门、国内主要新闻媒体、大型企业集团等。

拓尔思在大数据和舆情领域实践已久，具有成熟、成功经验，对各行业的用户需求均有深入理解。拓尔思能够为用户提供：大数据私有云和舆情服务私有云解决方案，包括数据中心机房建设、数据采集私有云、大数据管理系统、数据挖掘分析私有云、舆情管理平台等模块方案和技术，帮助行业用户自建数据中心和舆情私有云服务。

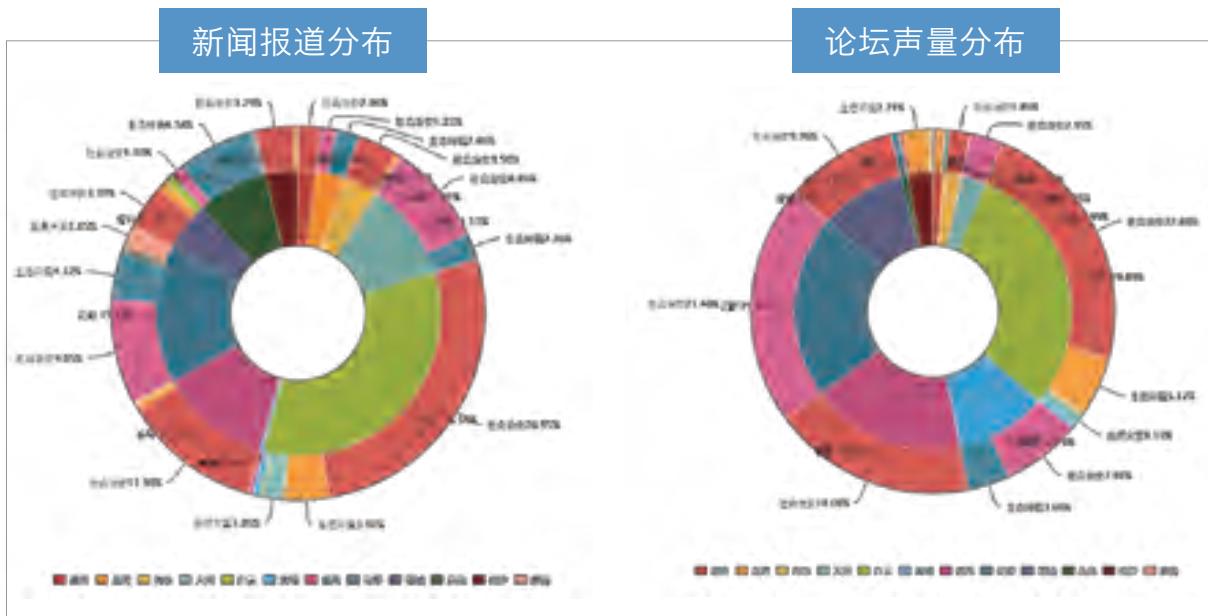
拓尔思拥有一支成熟稳定的分析师团队，来自北京大学、天津大学等国内重点高校，以及留学美国、法国的归国人才，业务能力突出、善于学习、年轻实干。

五. TRS全生命周期信息管理

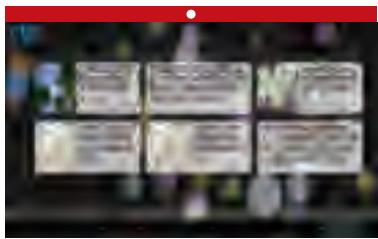


事前 | 及时发现弱信号，为应对处置赢得时间

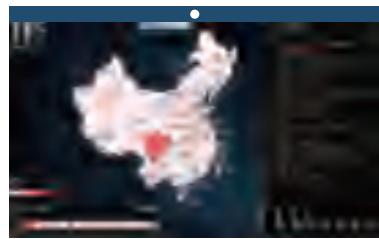
通过对不同媒体阵地的对比监测,发现潜在问题



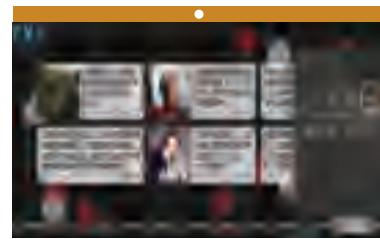
全国热点，一览无余



地方热点，明察秋毫



行业热点，捕捉异常



从行业到地方乃至全国，每一部分的热点事件都可单独显示。通过对相关信息的热点排序及自动关联，发现潜在问题，并进行层层深入分析，帮助用户将所有关联信息挖掘出来。



五. TRS全生命周期信息管理

预警案例1：及时为国家某总局发现舆情信息，避免演化成舆情事件

两会期间，通过TRS SMAS聚网大数据分析平台TRS运营团队及时捕捉到一条涉及国家某总局的重要负面信息。经过分析研判之后，TRS运营团队将此情况报给相关单位，国家某总局及时应对处置，当天下午事情即得到圆满解决。

(1) 及时捕捉到舆情信息 (2) 分析研判过程（略） (3) 总局应对处置，迅速化解舆情



预警案例2：Saas服务模式下的自动预警功能

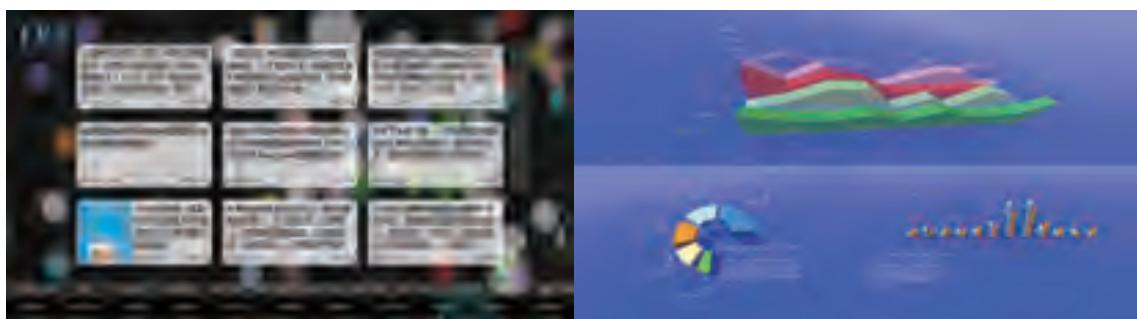
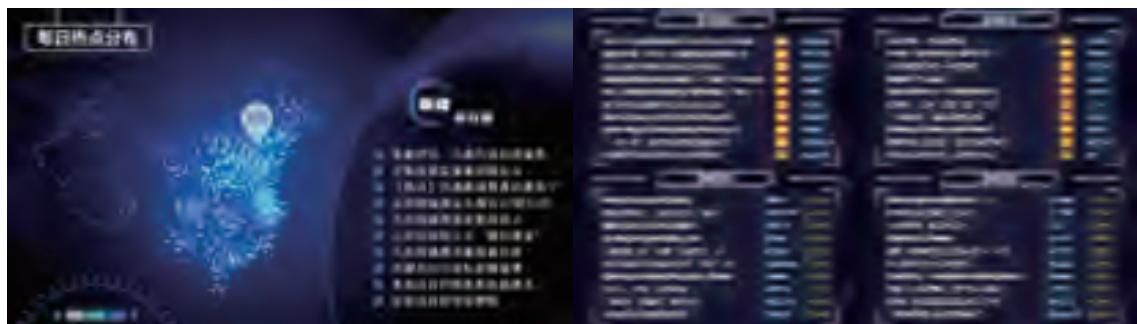
(1) 及时捕捉到客户需要的预警信息 (2) 自动发送到微信/手机APP上



事中 | 丰富的分析工具，透过舆情表面看本质，辅助决策

清晰的事件迁徙以及多维度的事件组合

让你多维度/全方位的了解事件的发展情况



五. TRS全生命周期信息管理

快速检索，快速建立事件分析页面

The screenshot displays the TRS event analysis platform. On the left, there is a search bar with fields for '搜索词' (Search term) and '时间' (Time), followed by a red-bordered '搜索' (Search) button. Below the search bar are several filter checkboxes for location ('地区') and category ('类别'). A large list of search results is shown on the right, each with a title, date, and a small preview icon. At the bottom, there is a navigation bar with various links and a red-bordered '分析' (Analysis) button.

事后 | 依托平台、专业分析师团队为客户定制各类分析报告

依托TRS SMAS聚网大数据分析平台、专业分析师团队，TRS运营团队能够为客户定制各类分析报告，包括事件分析报告、年度分析报告、行业研究报告、课题研究、大数据新闻、大数据出书等。



大数据新闻

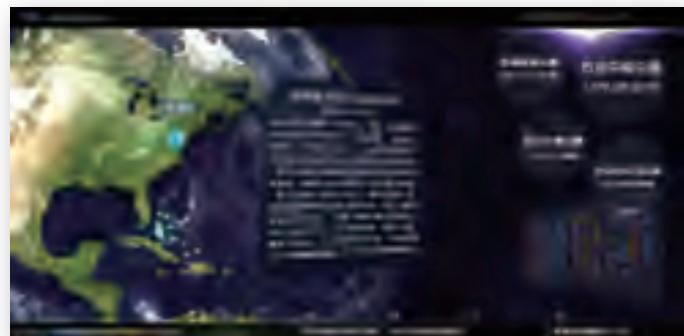


大数据出书

大数据不仅在新闻媒体上大放光彩，更在图书领域有自己的一席之位。

丰富的可视化工具

大屏显示是TRS SMAS的另一个特色，可根据客户需求，将数据、图表等设置在便于观看的大屏幕上，为客户提供最为专业的数据展现。



六. 典型案例

政府案例1：大数据舆情分析案例



全网同步采集涉及某城市的负面、敏感信息，囊括微博、新闻、论坛、博客、搜索引擎等领域。及时捕捉舆情弱信号，为客户应对处置赢得宝贵时间。



强大的数据运算和处理能力，丰富的分析工具、为客户提供多角度、多维度的分析利器，迅速为客户拨云见日，看清真相。

政府案例2：大数据打击非法集资案例



打击非法集资监测预警云平台是拓尔思旗下金信网银推出的全国第一款大数据打非产品，通过不同维度的数据源和大数据处理技术甄选出有非法集资风险的企业，发布专业风险分析报告，辅助金融监管部门监测金融风险。

打击非法集资监测预警云平台融合数十大风险数据源、国内顶尖的数据处理分析技术和财经专业背景的人才队伍于一体，深入分析非法集资行为，实现主动、全面、智能发现预警金融风险。

政府案例3：大数据+城市管理

兰州市大数据项目通过大数据技术手段，利用兰州市数据中心的四大基础库数据资源建设成果，形成全社会关注的环境、交通拥堵、网络舆情、经济决策等专题服务，面向各级领导、政府部门、市民及企业以“全城一张图”的形式提供综合性信息服务及决策支持。其业务需求包括：智能交通、智能环保、智能决策、舆情研判及大数据基础平台五部分。



政府案例4：厄瓜多尔总统大选



2012年，拓尔思基于自主研发的TRS SMAS聚网大数据分析平台对2013年厄瓜多尔的总统选举做了预测，拓尔思共采集西语系超过1000家媒体，涵盖了厄瓜多尔国内主流媒体、西语主流媒体、欧美主要媒体涉厄报道、Facebook、Twitter等社交媒体。结合各方面的数据结果，成功预测到了当选人为拉斐尔·科雷亚（Rafael Correa），而且预测结果与最终的选举结果仅相差1.3个百分点。

六. 典型案例

政府案例5：大数据传播分析案例



基于拓尔思大数据中心媒体分类划分，建立相关数据模型，传播指数体系，并自动计算某总局每条主动发布信息在互联网上的传播影响力。为总局提高传播影响力、围绕传播范围、传播深度、传播速度提供相应的技术解决方案和建议。

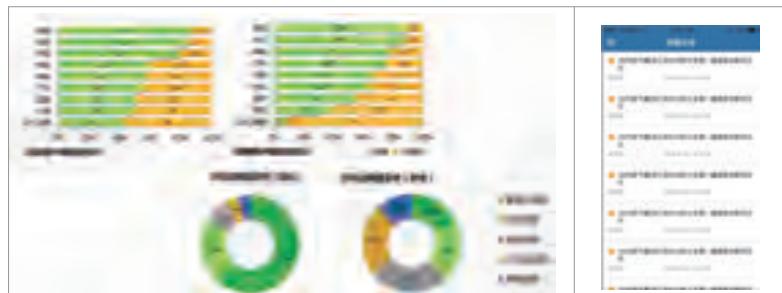
企业案例1：某世界500强汽车企业口碑监测案例



通过对企业及其竞争对手对比分析，了解企业口碑情况；通过对比企业的售后活动了解企业在售后活动推广方面的影响力，正负面情况，为决策提供分析依据。

企业案例2：某知名世界500强电子、电气设备品牌产品问题挖掘

根据产品的主要功能特点将信息进行具体分类，在此基础上了解各项指标的口碑情况，也可通过故障提及率及时通过异常数据发现问题，并将相关问题以预警的形式推送至手机APP。



企业案例3：某上市集团项目个性化定制实现PC端双微榜单排名

自媒体发达的今天，网民的发声越来越受到企业的重视。榜单的形式则最为直观地呈现出了大家热议的观点，进而可对当下舆论环境有准确的把握。榜单中有不少贴心的设置，更是便于分析不同地区、不同行业的受众关注点的差异。



企业案例4：个性化日报内容自动分类 —某大型全球连锁餐饮集团

通过对某大型餐饮集团信息的拆分与重组，小到集团旗下每个品牌大到整个餐饮行业，多维度、精分类地进行舆情监测。全网采集海量信息，满足客户不同时期的不同需求，内容自动生成的个性化报告为企业打造私人订制的尊享体验。



企业案例5：超级意见领袖千万粉丝构成分析

通过大数据挖掘制作明星个人画像，对相应的明星效应有一定的预判。再针对其粉丝群体进行深入挖掘，以此获得更有针对性的营销策略参考值，帮助企业选择性价比更高的产品代言。



七. 客户名录



政府客户名录

中央纪律检查委员会、中共中央党校、外交部、国家质量监督检验检疫总局、陕西省人民政府、乌鲁木齐市人民政府、海南省委宣传部、山西农厅、黑龙江省国土资源厅、广东省国家税务局、广东省环保局、北京市科学技术委员会、山东省政法委员会、上海市工商行政管理局、内蒙古自治区公安厅

人民日报社、西藏新华社、正义网

中国检验检疫科学研究院、教育部考试中心、中国教育研究院、北京外国语大学、中国海事服务中心、国家环境与健康数据中心、北京市科技信息中心

厄瓜多尔政府



企业客户名录

中国联通、中建国际、中石油建筑总公司、山东核电、河南中烟、山东航空、四川长虹、珠海华发地产集团、上海期货交易所、华兴银行、新华人寿

用友集团、广联达、奥美广告、艾瑞咨询、友拓公关、去哪儿网、酒仙网、亿赞普科技集团、美即面膜
索尼（中国）、宝马（中国）、三星（中国）手机研究院、日本佳能、英国航空



八. 功能架构及特性

平台功能框架图



八. 功能架构及特性

全网监测

TRS SMAS聚网大数据分析平台的监测范围覆盖全网，不仅包括新闻、论坛、博客、评论、平媒等传统网络媒体，也覆盖了新浪微博、腾讯微博、微信等新媒体，TRS SMAS聚网大数据分析平台同样可对百度、谷歌、搜狗等主流搜索引擎监控，还有港澳台、境外等外媒，实现了全网监测。TRS SMAS聚网大数据分析平台从数据上实现了舆情监测信息与大数据的融合，在应用实现方面不仅仅停留在信息的采集和舆情监测，发展了数据的深度分析和挖掘。

The figure consists of three separate screenshots of the TRS SMAS platform's user interface, each titled with its respective monitoring function:

- 传统网络媒体、新媒体监测**: Shows a list of news items from various media sources, including titles and small thumbnail images.
- 搜索引擎监测**: Shows search results for specific queries, with filters for search engines like Baidu, Google, and Sogou.
- 港澳台、境外网站监测**: Displays a pie chart and a map showing the distribution of monitored websites across different regions.

机构关系挖掘

SMAS系统可对微博中的事件、官微、人物等关系进行解析，根据数据的关联性，支持人与人、事与事、人与事之间的深度关系挖掘。支持无限层次关系挖掘，并自动生成关系图，所有关联一目了然。



单条微博传播分析

单条微博传播分析，旨在对每一条原发微博的传播情况进行分析。提供转发关系图，实现对转发层级的分析。提供地域分布图和柱状图，按转发者所属地域分析各省份的转发人数并进行排名。支持转发时间趋势图，分析转发微博的时间分布。支持微博认证分析和微博来源分析。支持情感分析和关键词分析。



八. 功能架构及特性

“意见领袖”挖掘

支持对微博意见领袖的分布区域、认证情况和性别、转发量和评论量的排行榜等分析。并支持分析图表。意见领袖指在某事件中参与的所有意见领袖的动态进行综合分析。包括对意见领袖的分布区域、认证情况和性别、转发量和评论量的排行榜等分析。



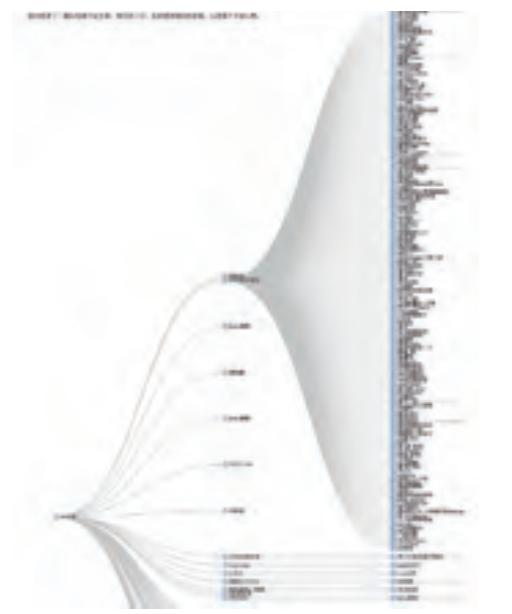
简报

TRS SMAS聚网大数据分析平台具有简报功能，可以通过各类word编辑器进行日报、周报、月报、及时报的生成。且简报可以以PDF、DOC格式导出以及直接发送到微信公众号或手机APP，另外，生成简报时可以引用系统内的任意统计图表。

TRS SMAS聚网大数据分析平台可以根据用户需求定制报表模版。使用过程中可以将感兴趣的新闻、图片作为简报素材，并随时制作日期报、周报等，依托TRS运营团队，可为客户提供各种形式的报告。

传播路径分析

传播路径指某篇微博传播过程的路径。可对某篇微博传播过程自动生成路径图。传播路径图支持环状、树状两种模式。传播路径指某篇微博传播过程的路径，可对某篇微博传播过程自动生成路径图。传播路径图支持环状、树状两种模式。



话题事件分析

TRS SMAS聚网大数据分析平台支持微博中各类事件的信息，可按话题量、转发量、评论量、意见领袖、参与用户、传播路径等分析，并自动生成各类图表。

也可以从事件的角度去进行分析，即考察舆情信息、广告营销等事件在微博中传播的情况，从中找出最核心的人物、最热门的微博、最全面的传播等等，全方面、多层面的了解事件的发展与内容。



系统提供的查询入口，通过关键字和条件的组合，可以获取所查询事件在微博中的传播情况、影响力等。对持续的话题事件分析还可以设置成专题。



八. 功能架构及特性

专题统计分析功能



用户可自行按需搭建专题，积累专属事件库。建专题时按照事件属性贴上相应标签，系统便可根据标签数量提供针对专题库的事件分类分析，便于用户对自己的数据库进行分析和管理。

手机APP

随着移动端的崛起，人们已经习惯随时随地拿起手机获取信息资源，TRS已经率先实现移动端信息监测。下载SMAS手机客户端后，用户登陆后可通过手机实时查看图表和列表类数据，监测内容完全自定义。舆情事件随身行，工作生活两不误。



九. 系统架构及关键技术

系统架构

TRS SMAS聚网大数据分析平台是基于云服务模式的互联网舆情信息监测平台，提供用户舆情信息的采集、分析，敏感事件预测和预警，监测信息全文搜索等服务，囊括事前预警、事中分析、事后处理全方位的舆情服务。



TRS SMAS聚网大数据分析平台架构

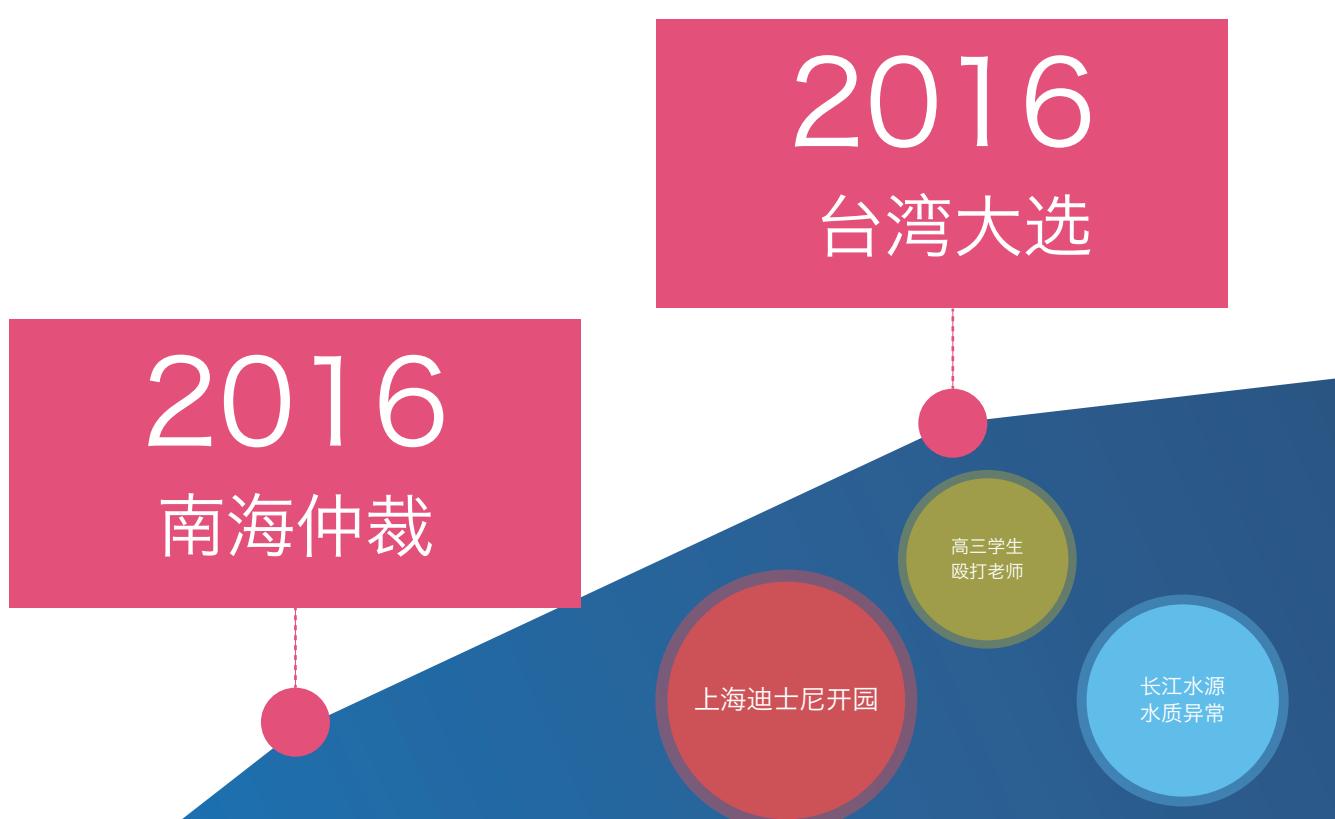
关键技术

高效全面采集技术	采集重点技术	智能挖掘分析		海量数据存储和检索	完善的数据服务
监测超过10万家网站	分布式采集技术	智能分词技术	智能提取	千亿微博数据存储和索引	7*24小时服务
新闻门户	多链路采集	文本自动分类	自然语言检索	千万微博用户关系存储	精准的行业知识库
论坛、博客	VPN隧道采集	自动过滤	情感分析	百亿网页数据存储和索引	各类图表的自动展现
主流微博	智能调度	关联关系挖掘	趋势分析	支持复杂检索条件	数据可视化技术
微信	图片表格采集	自动聚类	热词自动发现	秒级检索响应	专题跟踪
地区媒体	多语种支持	未知探索	实体自动识别	精准信息查询	事件深度分析
每日数据量超过数千万条	及时报				日报周报月报服务
聚网搜索采集					邮件短信

十. TRS SMAS舆论事件库

自2013年以来，TRS SMAS建立了自己的舆论事件库。根据不同事件的特性，总结分为社会、政府形象、文娱等8类主要信息，积累了1000多个专题，涵盖了近四年的热点事件，全面打造最全的舆论事件库。信息库可满足每个专题的需求，完成搜索目标匹配，呈现多媒体全方位的信息监测效果。庞大的数据采集工程，精确的信息匹配程度，多元化的数据展现形式，灵活的信息搜索工具，都为用户了解事件概况及其发展提供了必要的支持。

2013—2016年间，共监测搭建1052个舆论事件专题，其中62%为负面事件，10%为正面事件，28%为中性事件。从类别上来看，社会类事件占了53%，30%的事件涉及政府形象，文娱和教育方面分别以5%和4%的概率次之，自然灾害、生态环境、经济以及医疗卫生均占据2%的席位。



2013

恒大亚洲夺冠

2014

昆明暴恐案

2014

马航失联

2015

上海外滩踩踏事件

2015

天津塘沽爆炸

复旦投毒案

屠呦呦获
诺贝尔医学奖

恒大亚洲夺冠

玉林狗肉节

XP停止服务

美国总统
见达赖喇嘛

内地小孩
在香港街头小便

英国脱欧公投

《穹顶之下》

令计划接受调查

南海仲裁案

优衣库不雅视频

纽交所暂停交易

真功夫、肯德基
冰块超标

尼泊尔
8.1级地震

马航失联

克洛泽踩着巴西
超越罗纳尔多

天津塘沽爆炸

雾霾

手术室医生自拍

房祖名吸毒被抓

公务员调薪

高考改革

上海外滩踩踏事件

云南鲁甸地震

青岛天价虾

学校发现毒跑道

汪峰头条

香港旺角骚乱



从数据到智慧

Data · Knowledge · Intelligence



扫一扫，关注拓尔思



北京拓尔思信息技术股份有限公司
Beijing TRS Information Technology Co., Ltd.

总部地址：北京市朝阳区大屯路风西奥中心B座16层 邮编：100101
电话：010-64859900/64848899 传真：010-64879084 服务热线：4006 300229
E-mail: trs@trs.com.cn 官微: <http://t.sina.com.cn/bjtrs>
网址: www.trs.cm.cn