

# 拓尔思网站搜索云服务平台

## 产品背景

长期以来,互联网搜索平台占据了互联网搜索的绝大部分流量,但无法为政府、媒体、银行、企业等机构提供垂直化和行业化的搜索服务,而自建的搜索系统也普遍存在一些问题:

- 搜索系统建设的成本投入高、建设周期长。
- 搜索效果不佳,包括搜不全、搜不准、搜索结果排序不合理、搜索结果包含敏感内容等问题。
- 缺乏运营,导致搜索效果无法持续改进。

针对以上问题,拓尔思信息技术股份有限公司推出了网站搜索云服务平台(以下简称“网搜”)。

## 产品概述

网搜以云服务的方式提供搜索服务能力,集成到客户的官方网站上,为公众提供一站式、行业化、安全可靠、智能易用的网站搜索服务。



## 产品功能

### ● 底座能力

#### 数据采集：

爬取全站网页内容，包括文字、图片、附件、音视频等，智能识别数据结构，清洗干扰因素，持久化存储。

#### 基础能力：

NLP分词确保搜索结果全且准，智能权重计算确保搜索效果排序合理。

#### 后台管理：

租户+站点模式，利于数据隔离和数据安全；结构化的库表数据管理，便于搜索能力的灵活配置、扩展和调优。

### ● 运营能力

#### 词库管理：

词库种类丰富，包括敏感词、同/近义词、百姓体、自动纠错、拼音转换等；能力强大，实现字词联想、自动纠错、自动补齐、权重设置、新词自动发现等实用特性。

#### 搜索运营：

包括内容置顶、敏感内容屏蔽、设定搜索暗文、设定热搜词、用户反馈、搜索轨迹分析等特性。

#### 框应用：

默认内置行业化的框应用，包括领导人名片、相关领导、机构名片、相关机构、地图服务等，也可自定义扩展框应用，实现自定义的静态内容和动态服务的搜索。

#### 统计分析：

包括搜索次数和趋势、点击次数和趋势、搜索排行、词条使用统计等。

### ● 搜索服务

搜索前，提供热搜词、搜索暗文、自动纠错、近义词推荐、拼音转化等引导和辅助特性；搜索时，提供常规搜索、高级搜索、附件搜索等搜索能力；搜索后，系统自动搜集用户的搜索行为和点击行为，同时也支持用户主动提交留言反馈。

## 产品优势

### ● 一键采集

用户只需要输入要采集的网站首页地址，系统将自动开始采集数据、分析数据结构、清洗入库。

### ● 快速发布搜索服务

自动开启各项初始化操作，无需人工干预，默认内置经典配置、规则及算法，快速对外提供搜索服务。

### ● 持续的运营能力

云服务模式的好处是可集中资源，使搜索运营可持续、更高效、合理共享，最终促使搜索效果持续改进。

### ● 个性化扩展能力

包括搜索页面扩展（网站 Logo 设置，网站背景设置，字体字号颜色设置，搜索结果样式设置等），数据分类扩展（除常规数据分类外，还可扩展音/视频、图片、微信、微博等），以及框应用。

### ● 多端适配

兼容PC端、移动端，适应各种大屏、小屏。

## 用户价值

### ● 降低成本

大幅缩短建设周期，同时降低采购和运维成本。

### ● 提升效用

向公众持续提供高效率和高质量的网站搜索服务。