



译见 · 大数据平台

GTCOM (中国)

总机: 010-53223600

电话: 010-53223646

邮箱: marketing@gtcom.com.cn

地址: 北京市石景山区石景山路 20 号

网址: www.yeesight.com



GTCOM

译见·大数据平台

汇聚 海量数据 智能多语 技术 透析 数据价值



采集平台
YEESIGHT Spider



治理平台
YEESIGHT Governor



分析平台
YEESIGHT Miner



数据平台
YEESIGHT Express



数据地图
YEESIGHT DataMap



管理平台
YEESIGHT Manager

译见 | 采集平台 Spider

全网数据采集
全面 覆盖 精准 高效

平台架构

数据中心

清洗引擎 数据仓库 搜索引擎 系统管理
治理引擎 索引库 服务管理

采集集群

新闻网站 社交媒体 论坛网站 点评网站 流媒体网站 图片网站
博客网站 平面媒体 电商平台

采集源

新闻网站 社交媒体 论坛网站 点评网站 流媒体网站
图片网站 博客网站 平面媒体 电商平台



译见 | 采集平台 Spider

[特性]

强大数据采集集群

全球分布数据中心，3000+台高性能服务器

高效数据采集算法，智能采集调度策略

秒级数据采集实时性，高扩展分布式集群架构



精准数据源目录

多领域智能数据源发现

自动和人工标注相结合，精准标注全球数据源

千万级行业分类语料库，数据源智能行业分类



全面数据采集覆盖

65+语言、200+国家和地区、千万级数据源

新闻、社交、论坛、博客、电商、APP、点评、问答、
平媒、图片、音/视频等全网数据覆盖

日更新3000万+篇新闻数据

日更新5亿+条社交数据



领先数据采集技术

多语言数据采集、数据源自动发现

数据源自动行业分类、文章位置识别

全球元搜索采集技术、网站探针技术

自动网页结构化技术、智能摘要提取

精细数据 **清洗 & 治理**

“无效”数据变 “有效”



数据清洗引擎

精准数据去重处理

数据源归一化处理

智能文本质量分析



数据治理引擎

关键词提取

文本要素提取

命名实体识别

情感分析

敏感性分析

文本摘要提取

文本权重分析

文本聚类

文本分类

知识图谱构建

文本相似度计算

机器翻译

[特性]



数据引擎并行治理
实时数据流式处理，毫秒级延迟

自定义治理流水线工序
弹性计算资源调度



基于深度学习神经网络的
算法模型

译见 | 分析平台 Miner

敏捷 数据挖掘

多彩 数据呈现



[特性]

智能数据分析

内置百种数据模型及算法
独特跨语言文本分析算法
快速自助数据建模
灵活扩展第三方算法

数据可视化

仪表盘、大屏幕
图形可视化
地图可视化
企业报表
分析报告

数据能力

多类型数据接入，如RDBMS、
OLAP Server、Hadoop

元数据引擎

数据查询分析引擎

分布式计算

人工智能引擎

数据即服务
专属数据 个性定制



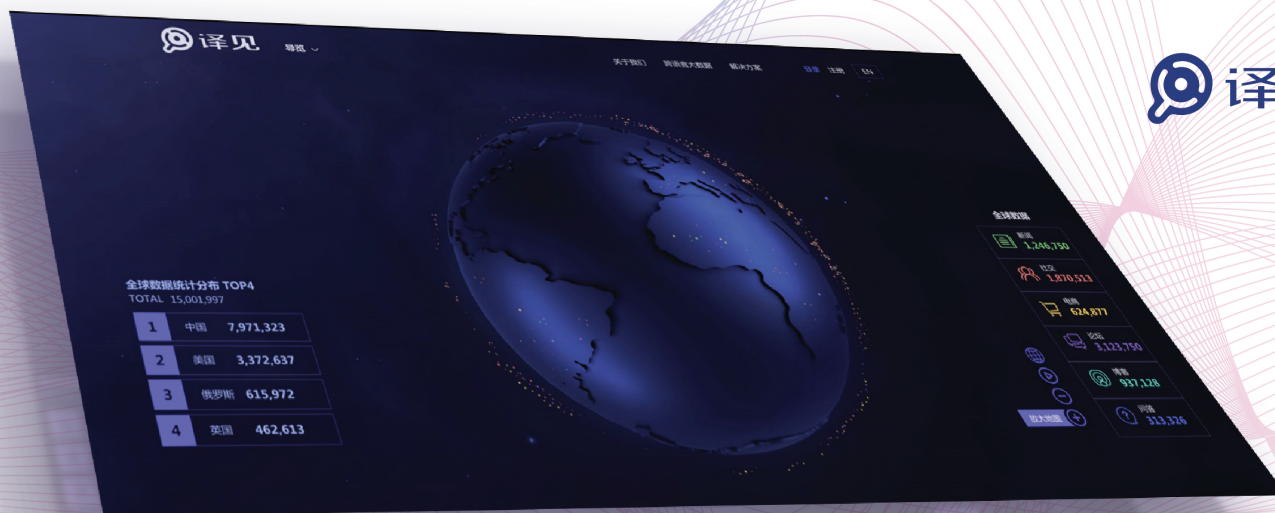
 译见 | 数据平台 Express
[特性]

 数据集市

千亿级历史数据存量
秒级实时数据更新
百种行业主题数据库
如政府舆情、法律服务数据库
用户个性化数据自定义
用户私有数据接入及集成

 数据服务

便捷的云端数据服务
基于Miner平台的在线数据分析
多样数据交付，如实时API、批量下载
自助数据管理
精准数据使用计量
第三方应用云端集成



译见 | 数据地图 DataMap

数据在地图上舞动

[特性]

数据动态分析

传播路径

生命周期

技术曲线

数据分布分析

热力图

闪点图

星光图

地图效果

2D/3D地图相互切换

多层矢量缩放

多样化地图样式定义

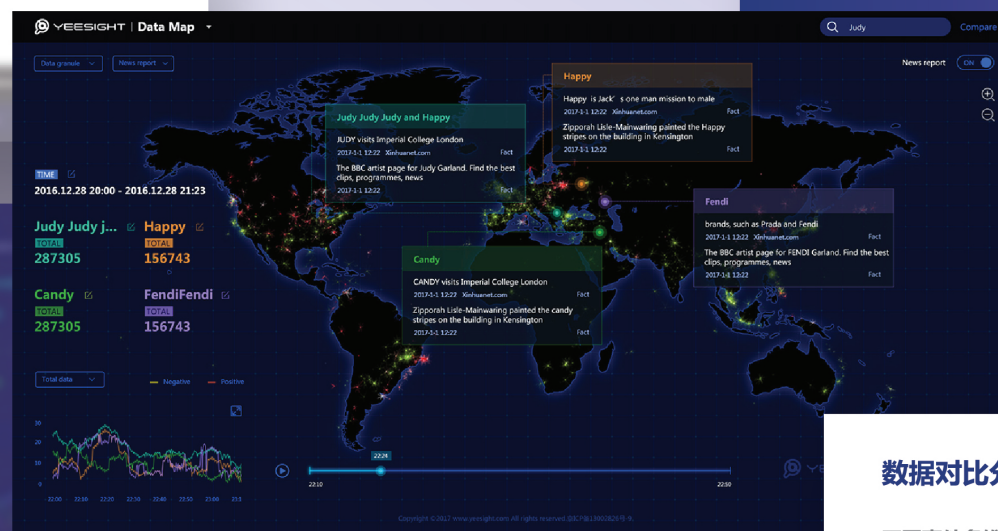
多样化数据呈现

地图数据交互

实时动态数据展示

数据心电图

简易的数据接口

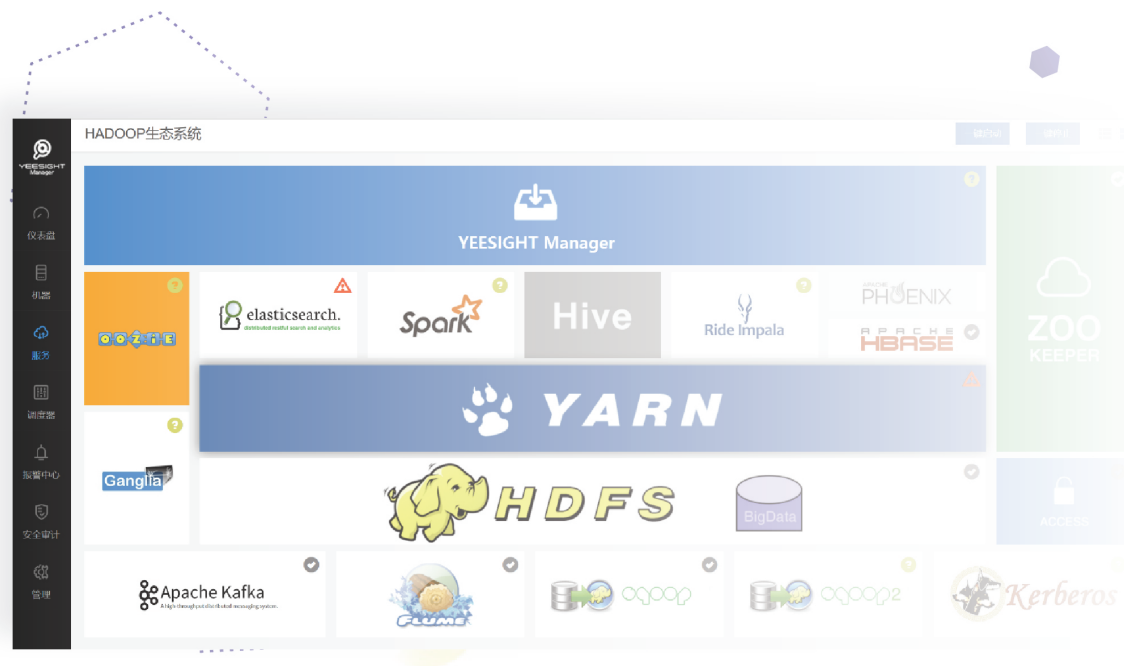


数据对比分析

不同事件多维对比

事件、时间、地点相关性分析

高效 数据支撑
安全 可靠管理



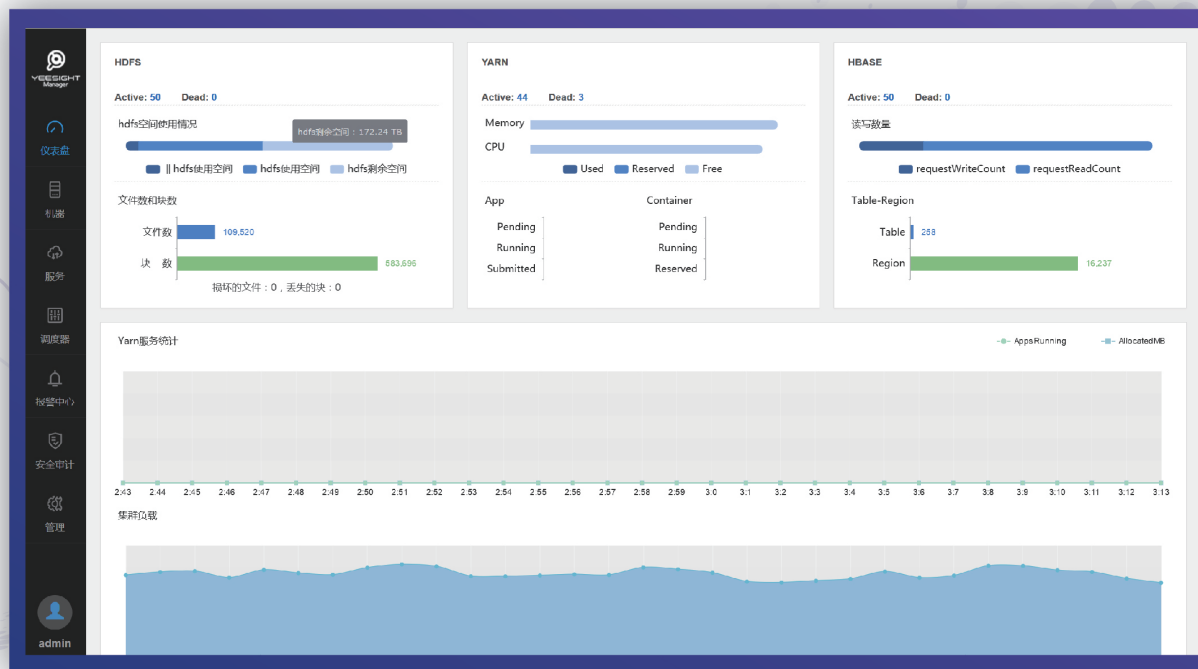
数据管理

资源管理

服务管理

安全管理

系统管理



[特性]

安全性

多层次安全体系
全平台细粒度SQL权限控制
数据加密保证数据安全
安全审计探查可疑行为

高可用

全服务双机热备
网络高可用机制
数据多备份保障

高性能

秒级响应
支持百亿数据OLAP
千亿数据实时查询

易用性

图形化自动化运维管理
一站式数据作业
无忧数据接入

开放性

Apache纯净开源版本
开放接口、开放协作
自主升级、不加绑定

[行业解决方案]

医疗大数据

旅游大数据

教育大数据

电商大数据

媒体大数据

工业大数据

电力大数据