

易洗[®]数据整合清洗平台

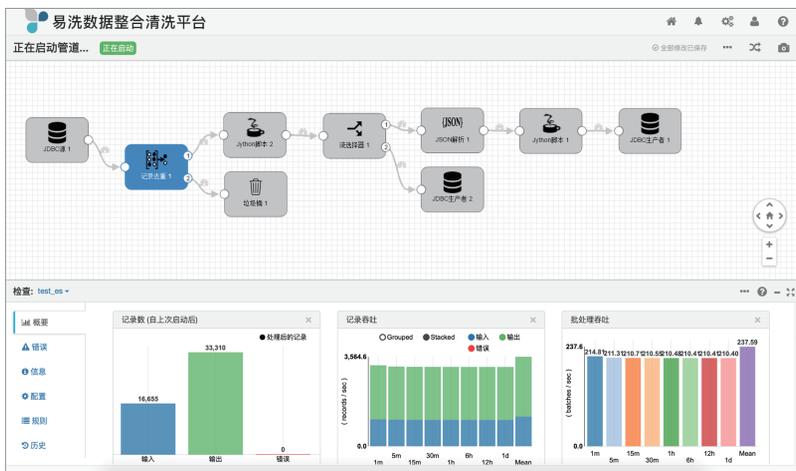
智能数据整合清洗专家

产品概述

易洗 (iDataIntergator) 数据整合清洗平台是一套专门对多源异构数据源的数据抽取、转换和维护过程进行简化和自动化，并将其输入数据集、数据仓库或大数据平台的集成工具。易洗平台可为数据抽取、转换和加载过程提供图形化的开发环境，已建立了130余个可复用组件，支持开发人员用最少的代码设计数据清洗作业。针对数据清洗融合的工作，易洗扩展了基于人工智能技术的多维智能匹配系列组件，可以在数据质量有限的环境中，实现整体数据的互相对比、更新、补全以及较为精确的主体识别，实践效果理想。

应用场景

- 开源情报数据处理** 易洗数据整合清洗平台将采集的数据进行合法化分析、数据格式匹配、数据正则化、数据格式转换、数据去重等专业的数据清洗，提供实时、高质量的情报数据，助力情报分析人员决策分析。
- 企业竞品数据处理** 企业市场分析需要实时准确地分析竞争对手的新闻、产品、价格、用户反馈等商业情报信息，易洗数据整合清洗平台能完成数据的实时预处理及清洗，为分析企业品牌以及竞品市场消费等需求提供准确的数据支撑。
- 企业征信底层数据处理** 企业征信业务需要汇集海量的发票数据、互联网工商数据、互联网司法数据等，在这种情况下，将面临数据融合的难题。易洗数据整合清洗平台采用人工智能主体识别技术，能解决同一家企业的信息相似但不一致引起的数据融合难题，且实践效果理想，能够为后期准确的数据分析提供有力保障。



产品优势

- 融合人工智能

集成了一系列人工智能组件，如社交账户信息的数据融合组件，工商数据智能匹配筛选组件（解决了同一家企业的信息相似但不一致引起的数据融合难题），以及自然语言处理的基础组件等。

- 支持多种数据源

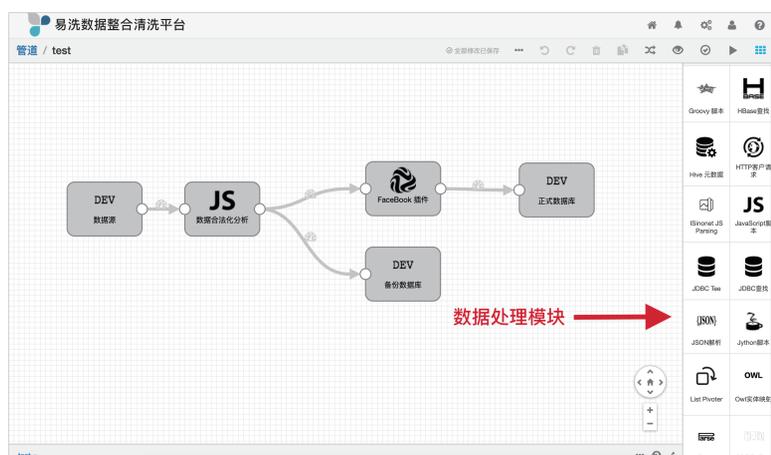
支持30多种数据源的一键接入，无需繁琐配置，数据源几乎涵盖所有数据应用场景。

- 低延迟实时融合

通过分布式并行计算，实时的数据融合与集成，不让延迟成为瓶颈。

- 清洗过程可视化

采用全导航交互式设计，通过管道化的图形界面建立数据处理节点，实时监控数据处理过程，一站式综合管控内部所有数据的交换进展。



主要功能

- 数据导入

平台支持将30多种数据源一键导入系统，包括Oracle、MySQL等主流关系型和HBase、Elastic Search等主流非关系型数据库连接器；本地文件系统与分布式文件系统HDFS及RabbitMQ、Kafka消息队列。

- 数据处理

已定义130余种数据处理组件，包括数据归并、数据拆分、数据去重、数据移除、数据脱敏、日志解析、自定义脚本等，覆盖了数据处理的全过程。

- 数据清洗

对无效数据、重复数据、异常数据等“脏”数据进行纠正和转换，并且支持错乱数据的匹配、纠错等。特色的数据清洗组件包括：社交账户信息的数据融合组件，工商数据智能匹配筛选组件等。

- 数据监控

实时监控数据集成、数据清洗状态。对异常任务处理情况，可实时报警，高质量数据监控体系保障数据处理的安全性及准确性。

合作客户

中国人民解放军某部，国家认证认可监督管理委员会，重庆发改委，国税，中财集团等。



北京市海淀区丰秀中路3号院7号楼中网数据大楼 100094
+86 (10) 58251700 +86 (10) 58251701
contact@isinonet.com
www.isinonet.com

