Big Data for Investment Research Management

Discover how IDT Partners helps Financial Services, Market Research, and Investment Management firms turn big data into actionable research insights.
Executive Summary

IDT Partners was retained by a financial client to investigate a legacy Research Management System (RMS) and provide its team with a more robust research dataset. The client’s investment research process includes supply chain [1] research and ongoing investment portfolio monitoring. The research information is collected and aggregated through proprietary supply chain data and by conducting channel checks. [2]

The IDT Partners’ team worked with the client to understand the business function dependencies on the legacy RMS, explored what business value potential system improvements could bring, and identified what functionality, such as reporting, researchers and executive management needed. A proposal offering options for a new system featuring improved data management, business process automation, and an intuitive user interface was put forth by IDT Partners and approved by the client.

IDT Partners Solution

The IDT Partners’ team worked with the client to develop and deploy a new system. The new system featured a much needed intuitive User Interface (UI), improved Workflows, Business Process Automation (BPA), Supply Chain Trend Spotting Tools, Dashboards, Reporting, and many other improvements. Additionally, dozens of new data sources were added and a scalable Data Warehouse was created to store and manage historical data.
IDT Partners increased the client’s research dataset eightyfold, from 50GB to over 4TB, while decreasing the data processing time from ten hours to less than two hours. The IDT Partner’s solution has provided the client’s research team with a technology-enabled competitive edge by improving the client’s operational efficiency by 20%, significantly improving its business processes management and automation capabilities, and using big data processing to provide actionable research insights needed to support its investment research process.

**IDT Partners Solution Highlights**

- 8,000% Larger Dataset
- 1,000% Increase in number of Data Sources
- 500% Faster Data Processing and Up-to-date Information
- 100% ROI within 12 months
- 20% Increase in Operational Efficiency
- Automated Trend Spotting & Monitoring with Big Data
- Business Process Management & Automation (BPM, BPA)
- Consistent Data with Automated Data De-duplication
- Intuitive and Streamlined User Interface (UI)
- Highly Customized Automated Reports
- Executive Summary Dashboards
Solution Overview:
Research Management System (RMS)

The following is an overview of a Research Management System (RMS) built by IDT Partners for a financial client. The system facilitates investment research, provides actionable insights, and helps support investment decisions by analyzing the supply chain \[1\] for companies in the client’s investment portfolio.

System Architecture

A proprietary framework was used to build the RMS as a **Service-Oriented Architecture (SOA)**. This has allowed for an increased flexibility in tool choice and made it easy to exchange information with client’s internal systems (Business Intelligence, Compliance, Vendor Management, and Billing) via **XML-based Web Services**.

Data Sources

IDT Partners assisted the client with data vendor selection, custom data collection, data aggregation, data processing, and Master Data Management (MDM) strategy for the following multi-terabyte datasets.

- Supply Chain Data
- Company Data
- Product Data
- Company Taxonomies
- News
- Market Data
- Industry Contacts
- Supply Chain Taxonomies
- Social Media
- International Trade Data
- Custom Research
- Proprietary Datasets
The following diagram provides an overview of the data management process as it relates to extracting research insights from raw data sources.
**Big Data**

Hadoop and MapReduce technology was used to convert multi-terabyte supply-chain-related datasets into actionable research data and **Business Intelligence (BI)** tools used by client to help facilitate investment research and decisions. The solution analyzes over a billion supply chain data points on ongoing basis to provide insights into historical supply chain changes.

The **Hadoop** infrastructure breaks up large datasets into "chunks" and coordinates the processing of the data out into the distributed, clustered environment. The Hadoop infrastructure included the following core components: HDFS containers, **MapReduce** processing infrastructure, Hive for immutable table data store, and HBase for mutable table data store.
The following diagram illustrates a simplified version of a data workflow that utilizes big data processing technology to provide the RMS with actionable data via an integrated BI platform.
Key Technologies

The following are some of the key technologies used for this solution.

- PHP
- LAMP
- JAVA
- jQuery
- Hadoop
- MapReduce
- HDFS
- HBase
- Pig
- Hive
- Linux
- Apache
- Hadoop
- HDFS
- HBase
- Pig
- Hive
- Lucene
- Solr
- ETL
- SSIS
- MarkLogic
- Pentaho

User Interface (UI)

The solution leveraged a combination of Rich Internet Application (RIA) technology and intuitive User Interface (UI) design to simplify the supply chain navigation process, seamlessly manage the underlying research processes, and to provide a way to easily review large quantities of data using data visualization and dashboards.
The following are some of the UI features of the RMS.

- Innovative Data Visualization interfaces
- Intuitive upstream and downstream supply chain navigation with rollover information popups.
- Analysis of companies within each node in the supply chain.
- Supplier, vendor, manufacturer, customer, and distributor information.
- Competitor information for companies within the supply chain.
- Comparison matrix for any two entities in the supply chain.
- Supply chain company filter with research-focused criteria.
- Supply chain search using hundreds of criteria including historical trends.
- Sharing of screens, graphs, charts, dashboards, and research feedback.

**Automated Trend Spotting & Monitoring**

IDT Partners developed a data warehouse and tools to keep track of historical supply chain changes to help identify supply chain trends. Automated trend spotting and monitoring automated and streamlined what had previously been a highly time consuming and error-prone manual process. This has proved to be invaluable for the client’s research team.

**Reporting**

The solution features dozens of highly customized reports leveraging a much larger dataset and the newly built data warehouse.
Some of the features include Custom Dashboards, Trend and Portfolio Monitoring Alerts, dozens of Highly Customized Reports, and Data Export (Excel, PDF, Email).

**Definitions**

The following are definitions for some domain-specific terminology referenced in this white paper.

**[1] Supply Chain**

A supply chain is a system of organizations, people, technology, activities, information and resources involved in moving a product or service from supplier to customer. Supply chain activities transform natural resources, raw materials and components into a finished product that is delivered to the end customer. The following is an overview of an electronics manufacturing supply chain.

The following simplified supply chain diagram is an example of how NVIDIA, a global technology company and GPU manufacturer, supplies
directly to the customer, to PC vendors, and retailers. Each point in the supply chain may include hundreds of companies across many tiers.

[2] Channel Check

In financial analysis, a channel check is third-party research on a company's business based on collecting information from the distribution channels of the company. It may be conducted in order to value the company or to perform due diligence in various contexts. Industries where channel checks are more often conducted include retail, technology, commodities, etc. Channel checks can give insights complementary to balance sheet analysis, such as distributor and retailer attitudes towards a product and its competitors, seasonal and geographic variation, and inventory levels (notably channel stuffing).
About IDT Partners

IDT Partners is a New York City based technology solutions and web application development firm specializing in enterprise-level web application development, custom product development, and technology solutions that help customers accelerate growth, capitalize on new market opportunities, and optimize operational efficiency by leveraging the latest technology.

Contact IDT Partners at info@idtpartners.com