

Transforming Ideas into Solutions... Transforming Data into Information...

## Twister Data Integrator Features

### Automation

- Dynamic data flow
- Flexible scheduling

### Ingestion

- Pull via FTP, HTTP, CIFS
- Query via ODBC / JDBC
- File system / web crawlers
- Search engine APIs

### Scalability

- Handle large data sets in a clustered parallel processing environment
- Distribute processing load over 1-500 computers
- Data streaming for near-real time analysis

### Data Normalization & Transformation

- Entity extraction
- Geo-tagging
- Address/Date standardization
- Upper/Lowercase
- Soundex/Metaphone
- Concatenation
- Tokenizing
- Regular expressions
- XML transformations
- XPath

### Data Encryption

- Encryption at rest
- Encryption in transit

### Cataloging

- Data pedigree
- Data statistics

### Delivery

- E-mail notification of new data availability
- Watch lists/alert generation
- Output to any file format via templates and XSLT
- Output to any JDBC database

### User Level Authentication

- Strict LDAP security model

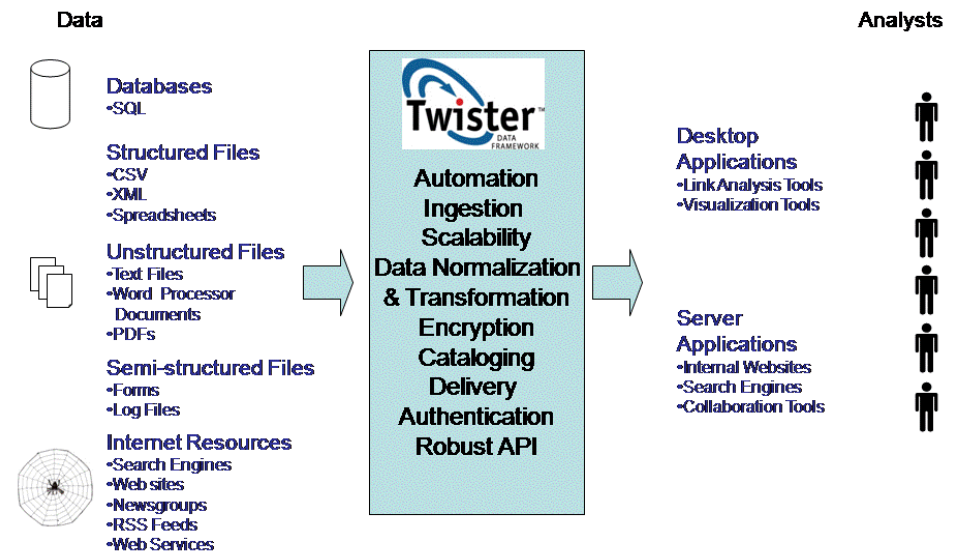
### Robust API & Component Architecture

- Reusable components for user-developed software
- Integration of third party COTS / GOTS tools

Solutions Made Simple, Inc. (SMSi) is solving data management problems across the Intelligence Community (IC) and DoD. If you can identify a data source and gain permission to use it, we can quickly make it accessible and usable for your analysts. Our expertise and software product, Twister, transforms data into information and enables sharing of data sources within communities of interest (COIs).

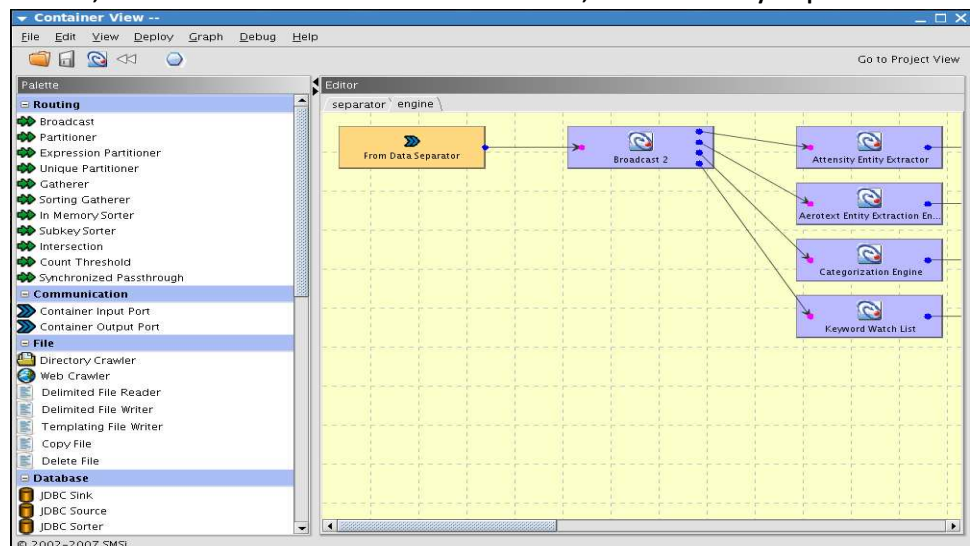
## Twister Data Integrator

The Twister Data Integrator ingests, normalizes, extracts, and delivers massive amounts of data, and integrates data for use by third party analysis tools, all in parallel. Twister Data Integrator performs the Extract / Transform / Load (ETL) process to transform unstructured or semi-structured data into structured data for querying. The extracted data can then be used by third party tools or to populate other databases through the Data Enrichment process.



## Twister Graph Designer

The Twister Graph Designer allows analysts to define and change the data flow through an intuitive Graphical User Interface (GUI). Using Graph Designer, analysts can specify data sources, data normalization and transformation routines, and data delivery output formats.



## Twister Data Server Features

### Federated Query

- Proxies data from multiple sources for centralized access to disparate data sources and obtaining query results from search engines, databases, and web services across networks
- Adapts / translates between standard JDBC / ODBC, XQuery, and SOAP interfaces to adapt XML data for use by SQL-based tools

### Database Clustering

- Distribute inserts across multiple database servers
- Combine query results from multiple database servers into one results set

### Supported Platforms

- Windows 2003 Server (x86)
- Windows 2003 Server 64-bit Edition (x64)
- Solaris 10 (SPARC 32-bit, SPARC 64-bit, x86)
- Linux (kernel 2.6 or higher)

### Java Runtime

- JRE 1.5 (Sun JRE recommended on all platforms)

### Supported Browsers

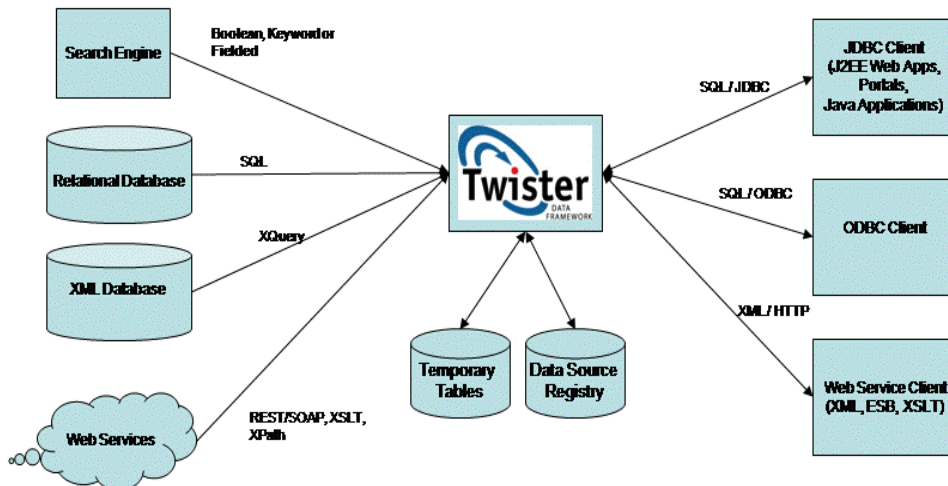
- Internet Explorer 6.x or higher
- Firefox 2.0



11720 Sunrise Valley Drive  
Suite 320  
Reston, VA 20191  
Deb Yeagle, Director, DoD Programs  
Cell: 540-907-6789  
Fax: 571-323-5735  
Desk: 571-323-6540  
E-mail: dyeagle@sms-fed.com

## Twister Data Server

The Twister Data Server performs federated queries across data sets and networks. Access to different data sets is mediated and all credentials are managed by the Twister Data Server. The query results are collected from each source and combined into either a single correlated result set or multiple results sets. Since Twister Data Server operates in a parallel environment, it is scalable to meet the demands of queries on large data sets and/or increasing analyst loads.



Analyst Tools

## Systems of Record

### Twister Return on Investment (ROI)

The Return on Investment (ROI) for Twister is high, with the software paying for itself in less than one year in cost savings from reduced custom software development. Twister is operational in one third of the time.

(based on 8 data sources, 2 months for integration of each, @\$20K per month in development costs)

