DataWalk

The DataWalk Analytics Platform

Database And Analytics Platform For Intelligence-Led Decision Making

Key Benefits

Get Answers 10X Faster
- All your data, connected in one place for search and analysis
- Fast response, even with billions of records
- Open system; no vendor lock-in
- Create analyses anyone can run with the click of a mouse

Rest Assured
- Produce repeatable results: instantly save and re-run analyses
- Make your best decisions by analyzing ALL available data
- Identify high value entities via easy risk scoring
- Military-grade security

A Fraction Of The Price
- Software license price 70-80% lower than alternative scalable systems
- Minimal professional services required to operate
- No forward-deployed engineers required
- Easily increase capacity just by adding commodity servers

The Problem

Large organizations in law enforcement, intelligence, insurance, and other sectors often face challenges analyzing large volumes of data across multiple data silos with various data structures and security permissions. Accessing, normalizing, combining, and analyzing this data can be highly problematic, impacting the organization's ability to fulfill their mission or meet their objectives.

The Solution

DataWalk is a commercial-grade Big Data software platform for connecting numerous large data sets, into a single repository for fast visual analysis.

Answers 10X Faster

DataWalk enables you to get answers up to 10X faster than traditional approaches, via the capabilities below.

All Your Data In One Place

DataWalk integrates all your your data from multiple sources into a universal model to deliver a big-picture view of all activities and connections, to support both tactical and strategic analyses. Data from internal database "silos", external databases, and Excel/CSV files are easily combined.

Fast Response With Big Data

Patented DataWalk technology is architected specifically for environments with many sources and large volumes of data, even if the data is structured or described in different ways. DataWalk identifies and stores connections between the data sets to deliver fast interactive analysis, even for billions of records.

Intuitive Visualizations And Graphical Displays

DataWalk is a visual analytics platform and does not require you to have expertise in SQL, programming, or a scripting language.

A core component of DataWalk is the Universe Viewer, which provides a visual representation of all your data. On the Universe Viewer you can easily structure and query your data through an intuitive visual interface.
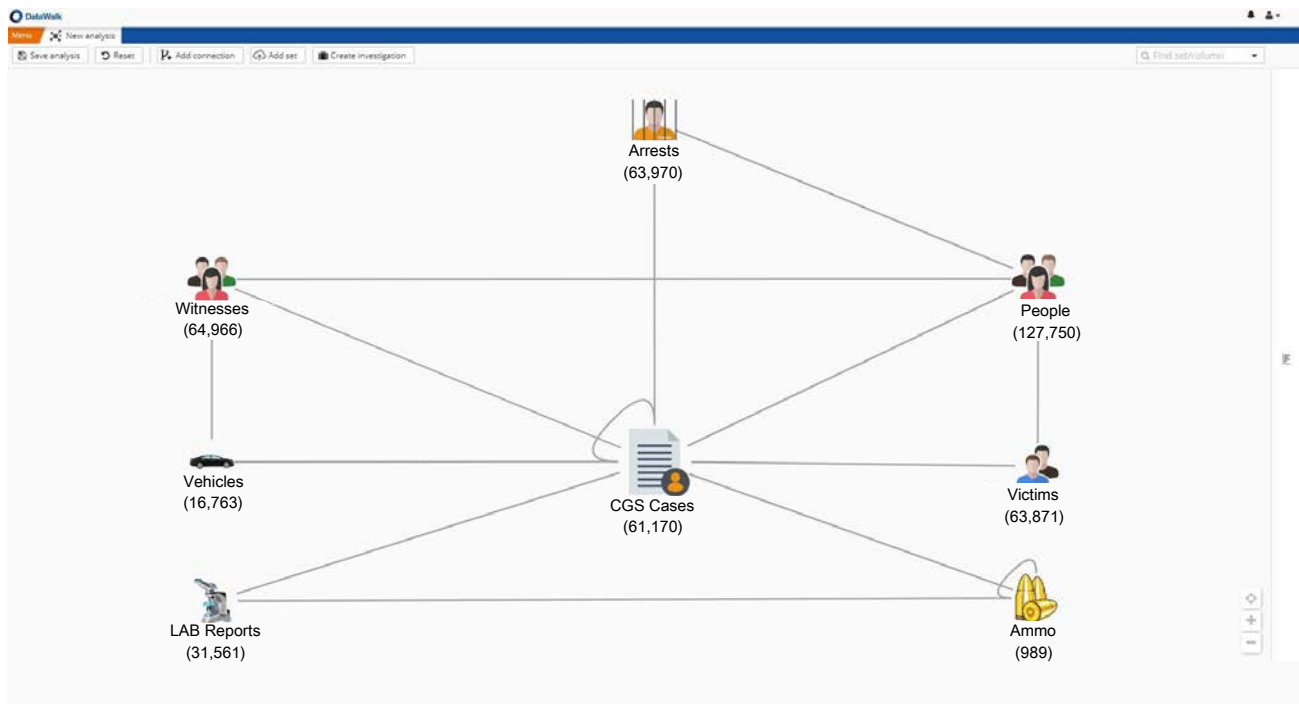
Fig. 1: The DataWalk Universe Viewer Provides A Highly Visual Analytics Environment

Powerful Link Charts

DataWalk link charts enable you to identify hidden relationships and view large networks of interconnected objects to quickly spot patterns or anomalies.

DataWalk link charts also integrate with maps and support Social Network Analysis heuristics. Also included is a graphical undo/redo capability.

Geospatial Analysis

DataWalk has embedded geospatial capabilities with various map providers. You can view heat maps, search polygons on maps, and link data based on geographical proximity, even with large amounts of data. You also can deploy a geocoding solution with or without internet access.

Dossiers

DataWalk includes a facility for customizable dossiers, which instantly provide any desired information you would consistently like to see about a person, address, vehicle, or anything else. Dossiers can be configured as Target Packages for law enforcement and intelligence applications.

Text Analysis

DataWalk provides a variety of capabilities to enable analysis of text content such as documents, emails, and social media excerpts. You can easily find words, phrases or patterns in text content. You can search text for specific terms using exact match, or to find similar, potentially matching terms or typos. You can also create custom dictionaries with specific terms of interest, enabling those text objects to be automatically tagged.

An Open System

DataWalk is an open system with published APIs to easily access data and analyses from other systems. For example, the popular program R can be called to deliver statistical analyses, machine learning, and other functionality with data provided by DataWalk.

Get Results, Even With Dirty Data

Traditional tools don't operate well on dirty, inconsistent, or incomplete data and often require long, expensive data cleanup projects. With DataWalk you can quickly profile your data, deduplicate content, and perform on-the-fly transformations, all without requiring any action from data owners or IT.

Instant Analyses

With DataWalk you can easily create and save your analyses and then easily re-run them with the click of a mouse. You can build a library of domain knowledge, share analyses with other users, and deliver more consistent and reliable results.

Correlate Unrelated Events

DataWalk provides you with the powerful capability to correlate events which are completely unrelated. For example, you could correlate bank robberies with data from license plate readers, to see if the same license plate is showing up in a nearby place and time as the crimes.
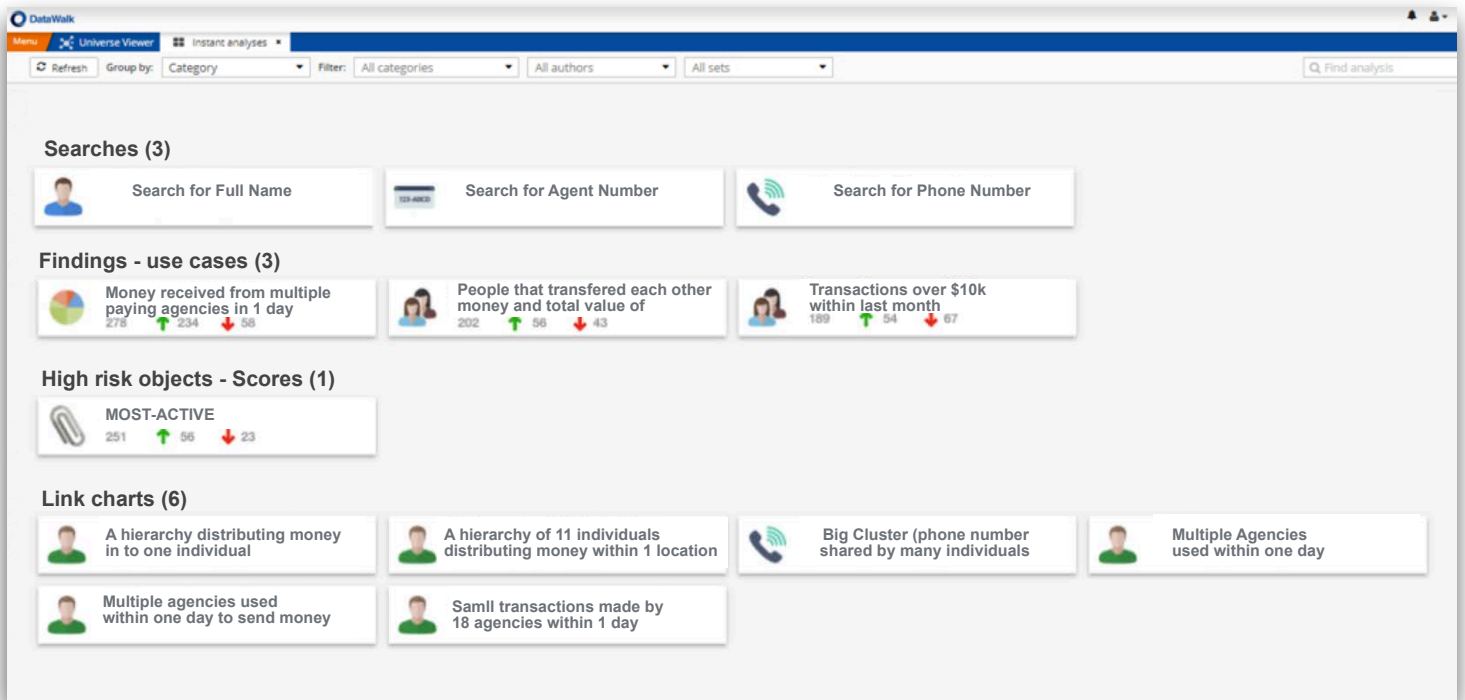
Fig 2. Non-technical users can easily run complex analyses via DataWalk Instant Analyses

Quickly Identify Clusters

DataWalk can scan your data and automatically identify clusters based on graph theory algorithms. This can greatly expand your capability to identify organized crime groups and other cluster patterns.

Reports

DataWalk also enables you to generate any desired reports via customizable templates.

Rest Assured

The work you do is important, and you want to know that you've done all the right things to get the best possible answers. DataWalk helps ensure you deliver the best, most reliable, and repeatable results.

Repeatable Results

With DataWalk you can instantly re-run a saved analysis, enabling easily repeatable results.

Use All Your Data To Make The Best Decisions

Traditional approaches often limit which data is available for you to analyze. In contrast, with DataWalk you can easily do analyses across ALL your data, enabling you to make better, more informed decisions.

Easy Risk Scoring

You can quickly generate or modify scores, across any number of analyses. You can score any objects (e.g., people, properties, activities, locations, etc.,) to effectively spot patterns across all your data.

Military-Grade Security

DataWalk provides highly granular permissions where users only see the data for which they are authorized. If desired, permissions can be implemented down to the level of an individual cell.

To further address needs for data security, DataWalk can be deployed on-premise behind your firewall, and supports secure communication (SSL) between your browser and the DataWalk application server.

Easily, Securely Collaborate

DataWalk is built for collaboration. You can easily share data, analyses, and investigation files with authorized colleagues and agencies.

Easy, Powerful Alerting

DataWalk offers a powerful alerting capability to monitor for user-defined conditions and value changes in the data. Alerts are easy to configure and data is constantly scanned to check against alert conditions. Alerts are managed in an alerts queue for easy review and follow-up.

Manage Your Work

The Investigation Workspace makes it easy to create and manage investigations in DataWalk. You can easily create and share a folder for a new investigation, and attach analyses, link charts, notes, and any related documents.

In addition to having folders for investigations, new DataWalk Object Folders enable you to easily create and share folders for any object, such as a person, address, vehicle, claim, or anything else.

A Fraction The Price

DataWalk is a fraction the price of alternative scalable systems, both at initial purchase, and over the lifetime of the solution.
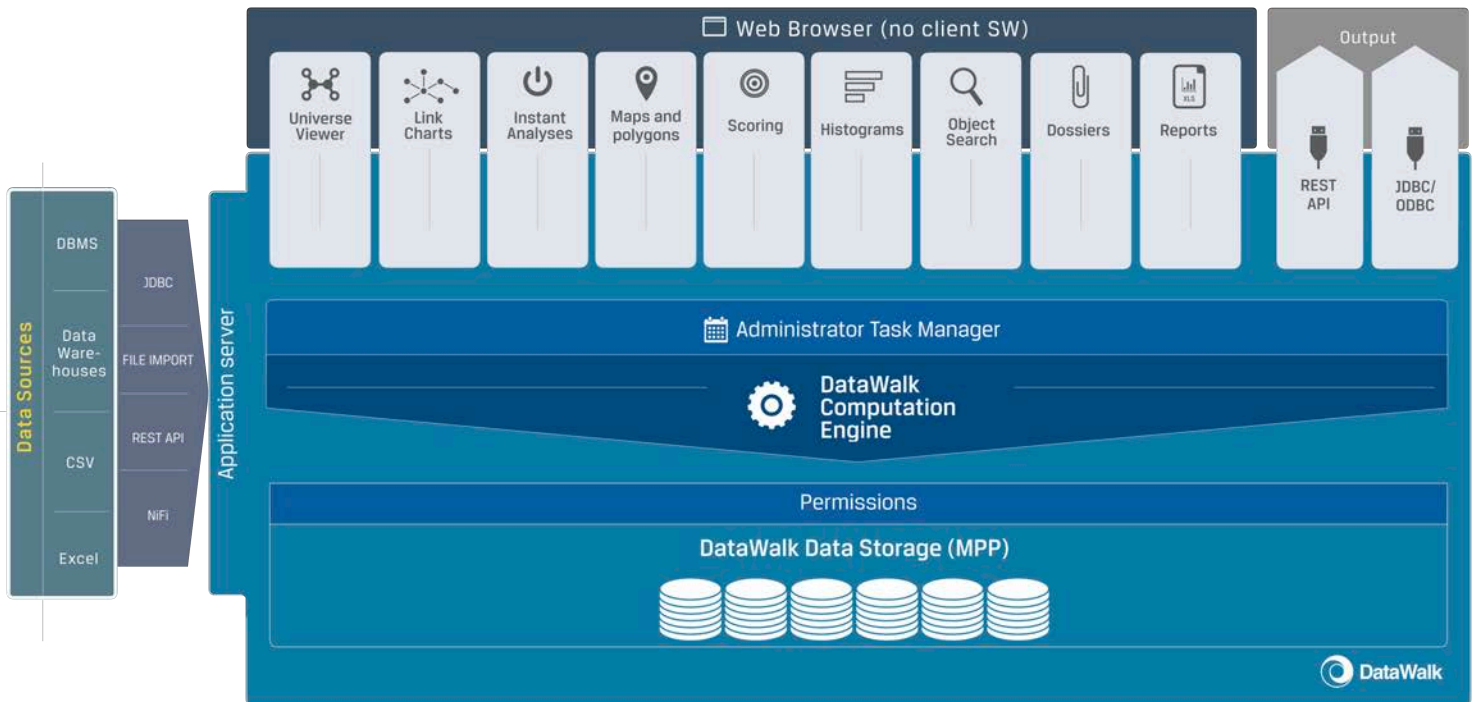
Figure 3: DataWalk System Diagram

Software Price 80% Lower

Depending on systems and configurations, the license price of DataWalk can be as much as 80% lower than other scalable Enterprise-class systems.

No Forward-Deployed Engineers

With DataWalk there is no need for extensive long-term professional consulting services. This further drives a dramatic cost advantage relative to alternative systems.

Minimal IT Investment

Once deployed, DataWalk typically requires minimal IT support. Users can instantly access all data they are authorized to see. Unlike alternative systems, administrative power-users can modify the analytic data structure on their own using simple graphical interfaces, without requiring professional services.

An "Off-The-Shelf" Scale-Out Platform

DataWalk is a Commercial Off-The-Shelf (COTS) software platform running on standard commodity servers, ensuring affordable economics.

If you need to expand your environment, with more data or more users, then you can increase capacity simply by adding more servers to the pool.

Fully Integrated Out-Of-The-Box Software Platform

DataWalk is a fully integrated analytics platform, so there's no need to cobble together multiple product components or separate modules to have a complete system for data analysis. DataWalk comes complete with all specified analytics as well as the software to enable scalable data storage.

No client-side software!

DataWalk is operated via a simple web browser. There are no client-side installs, so you can avoid the risk of data loss while deploying to large numbers of remote users.

Sample Applications:
- Anti-Money Laundering
- Contraband movement
- Counter-drug
- Counter-intelligence
- Counter-terrorism
- Counterfeiting
- Espionage
- External fraud
- Human trafficking
- Illegal immigration
- Internal fraud
- Lawful intercept
- Low intensity conflict
- Management of assets
- Message pattern tracking
- Organized crime
- Process monitoring
- Process exploration
- Targeting optimization

Key Specifications

System Architecture	• Shared multi-user system • Scalable, shared, single-instance, multi-node data repository • Two layers: UI and Computation Engine
Scalability	• Easily link dozens/hundreds of data sources • Interactive analysis of many millions/billons of records • Scale-out architecture: scale system capacity by adding commodity servers
Push-Button Analytics	• Instant Analyses enable you to encode organizational knowledge by easily generating and re-using searches and analysis paths on the Universe Viewer.
Link Analysis	• Enables network analysis and geospatial analysis. • Multiple layouts, including structural, radial, and hierarchical. • Social Network Analysis including betweenness, closeness, page rank, shortest path, and Eigenvector. • Save and retrieve link charts. • Undo/redo analysis steps on link charts as needed. • Easily visualize flows of any objects (e.g., money, material, etc.). • Quickly identify clusters across all data • Links can optionally be directional • Preview object details

Maps	• Link charts and flows can be presented on maps
	• Heatmaps visualize most frequent locations of specified objects or events
	• Create and search a polygon on a map
	• Maps can be integrated with GoogleStreetView
	• Geolocation translation is available via LocationIQ or an offline service
	• DataWalk supports:
	    ✓ OpenStreetMap Server (such that no request is sent off premise)
	    ✓ MapQuest
	    ✓ GoogleMaps

Text Analysis	• Extract metadata (including geocodes, if available) from unstructured binary files (e.g., images, video, etc.) • Pattern matching in unstructured text (e.g., finding/matching keywords or phrases), either via fixed or Regular Expression phrases. • Search text for terms based on exact match, Soundex, editorial distance, and distance between words. • Create user-defined dictionaries for specific terms of interest enabling text objects to be tagged.
Other Visualization and Analysis Capabilities	• Object search facility • Drill-down charts • Multi-dimensional scoring • Ability to create custom calculated columns on tabular data • Basic statistics on tabular data (min, max, sum, avg) • Customizable dossiers show all desired data about an object, on a single screen (can be used as target packages)
Investigation Workspace	• Easily create Investigation Folders for DataWalk analyses and link charts associated with an investigation • Add ad-hoc notes • Attach any other files to an Investigation Folder • Specify colleagues with whom a Folder is to be privately shared • Folders marked if any updates since last open
Object Folders	• Easily create and share folders for any object (e.g., person, place, vehicle, claim, etc.)

Data Import/Export	• Pull data from sources, and schedule automatic refresh • Your existing tools for ETL or data movement can drop data into DataWalk drop folders or use DataWalk RESTful APIs, and that data is then imported into DataWalk. • Drag and drop CSV or Excel XLSX files onto the Universe Viewer, or upload via the RESTful API. • Share data and analyses with other tools via RESTful access, JDBC, and ODBC. • Export results to Microsoft Excel, including to Excel templates for report generation.
Data Sources	• Data can imported into DataWalk from virtually any source with a defined interface • Structured or unstructured data • Data sources include any relational database, Microsoft Excel files, CSV files, web pages, Hadoop HDFS, and any other source with a JDBC, ODBC WSDL, or RESTful interface
Collaboration	• Single data instance shared by all users of the system • Easily share views, analyses, and investigation folders with colleagues who have appropriate permissions.
Security	• Highly granular (cell-level) rule-based permissions using predicates • Audit trail logs
Compliance	• 28-CFR Part 23 Compliance, i.e., data in DataWalk can be deleted on any desired schedule.
Alerts	• Specify alerts for effectively any condition. • Alert notifications delivered to user's view of DataWalk, or can be sent via email
Reporting	• Automatic report generation • Reports can be exported to Microsoft Excel, including Excel-based templates • Reports can be automatically distributed on a pre-set schedule
Platform	• DataWalk runs on commodity servers in a scale-out configuration • Supported operating system platforms are RedHat7 and CentOS7
Supported Browsers	• DataWalk is browser-based and there are no client-side software installs required • Chrome C38+ (highly recommended; enables highest DataWalk performance) • Firefox 33+ and higher • Microsoft IE11 • Other browsers supported as required
Deployment	• Software-only solution runs on commodity hardware • One-click deployment of base DataWalk software • Can be deployed on-premise, or in cloud