



Datasheet

# DataMatch Enterprise™



68 Bridge, St. Suite 307  
Suffield, CT 06708



+1 888-779-6578



Sales@DataLadder.com



www.DataLadder.com

# DataMatch Enterprise:

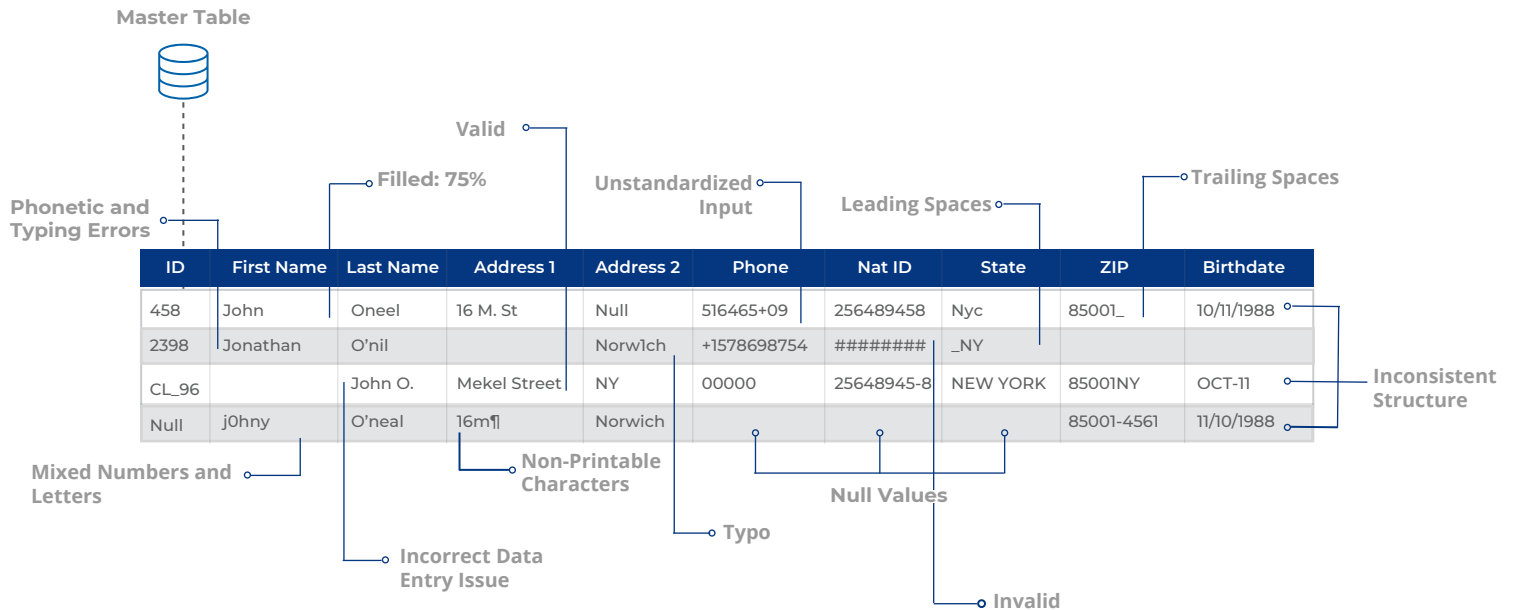
## Industry Leading Accuracy and Speed

We heavily invest in research and development improving the speed, accuracy, and usability for our products. Below, there are independent tests from a university which has evaluated various data matching companies on accuracy, speed, and cost.

Features of the Solution	Data Ladder	IBM Quality Stage	SAS Dataflux	In-House Solutions
Match Accuracy (Between 40K to 8M record samples)	96%	91%	84%	65-85%
Software Speed	Very Fast	Fast	Fast	Slow
Purchase/ Licensing Costs	80-95% below the competition	\$370K+	\$220K+	\$250K+
Time to First Result	15 Minutes	2 Months+	2 Months+	3 Months+
Comments	Above tests were completed on 15 different product comparisons with university, government, and private companies (80K to 8 million records). This includes the effect of false positives.	Need multi-threaded, in memory, no-sql processing to optimize for speed and accuracy. Speed is important, the more match iterations you can run, the more accurate your results will be.	Include base license costs, 2014 process or later, in-house, includes salary + benefits. Not in-house implementations had a 10% chance of losing in-house implementations had lost the core member who ran and understood the matching program.	A metric for ease of use. This is the time of first result, not necessarily full cleansing.

# DataMatch Enterprise™

## Standardization and Simplicity



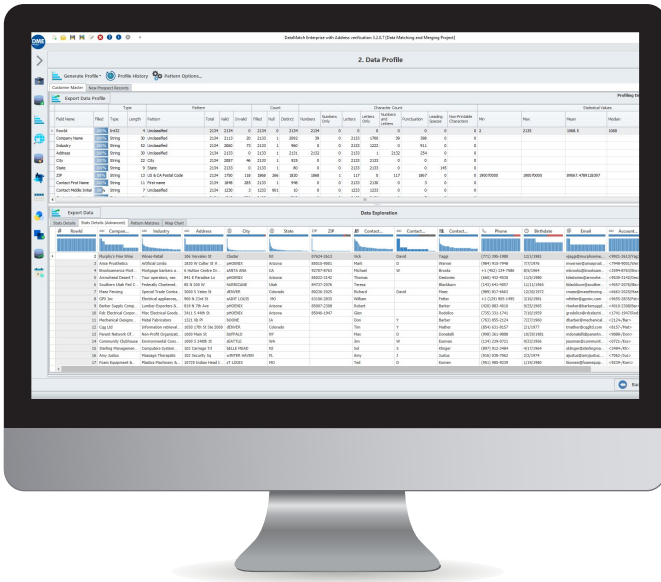
Oftentimes, data is inconsistent both in formatting and in wording. **Example:** Road Vs Rd, Sue Vs Susan, International Business Machines Vs, IBM, etc.

This is why we've included several standardization libraries with DataMatch Enterprise™, as well as allowed for quick and simple creation of your own libraries to be used repeatedly.

Data cleansing and fuzzy logic can be complicated. That is why we invested strongly in a world class visual interface that minimizes the number of clicks needed to complete a data cleansing and matching project, and it also provides a visual graphic user interface that can be filtered and customized. Whichever way works best for your project.



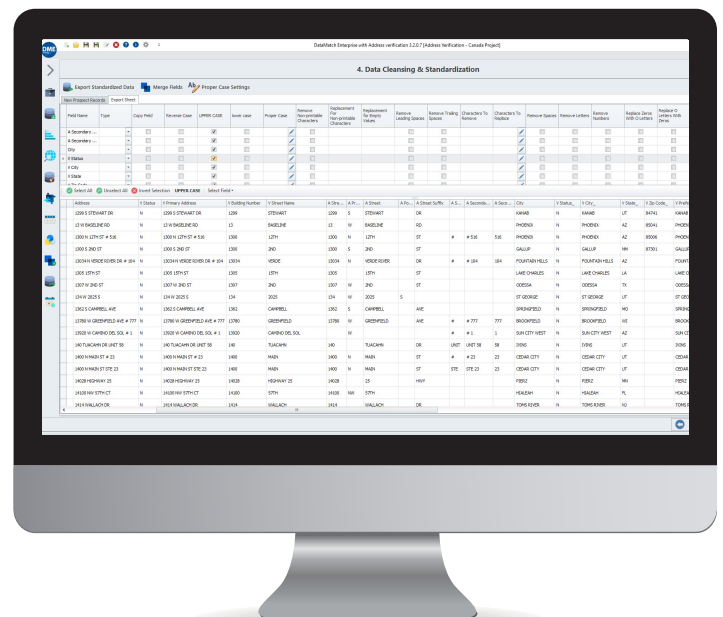
## Data Discovery and Profiling



Developing a deeper understanding of your data at the start of a project empowers users to make smarter, more informed decisions and prevent costly mistakes. Determine what data needs to be cleansed and standardized and what may be used as match criteria with our 'Quick Profile' feature and get the most out of your data. Findings are saved for future reference in a DataMatch Enterprise™ Project.

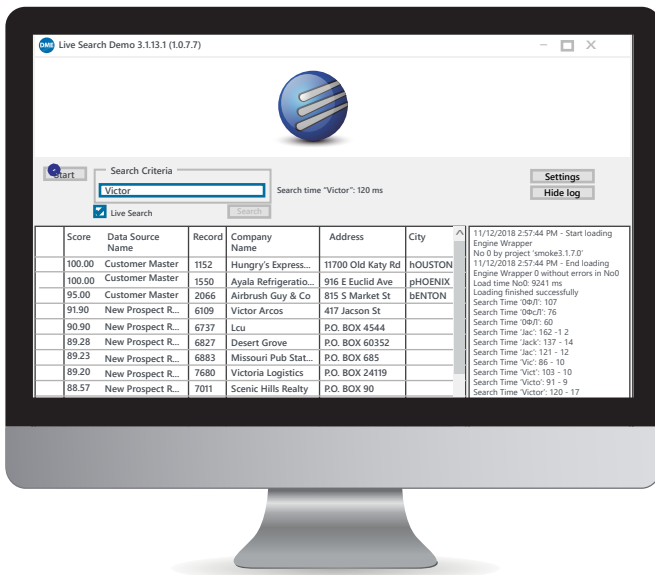
## Data Cleansing Done Right

Data cleansing and fuzzy logic can be complicated. Use our built-in libraries, proprietary matching capabilities, and sophisticated pattern recognition features to clean and standardize your data at scale. Our world-class visual interface further minimizes the number of clicks needed to complete a data cleansing and matching project, fully customizable according to your unique data quality project needs. Users can see their data as it changes with your data cleansing settings with our instant data preview feature.





## Data Quality Firewall

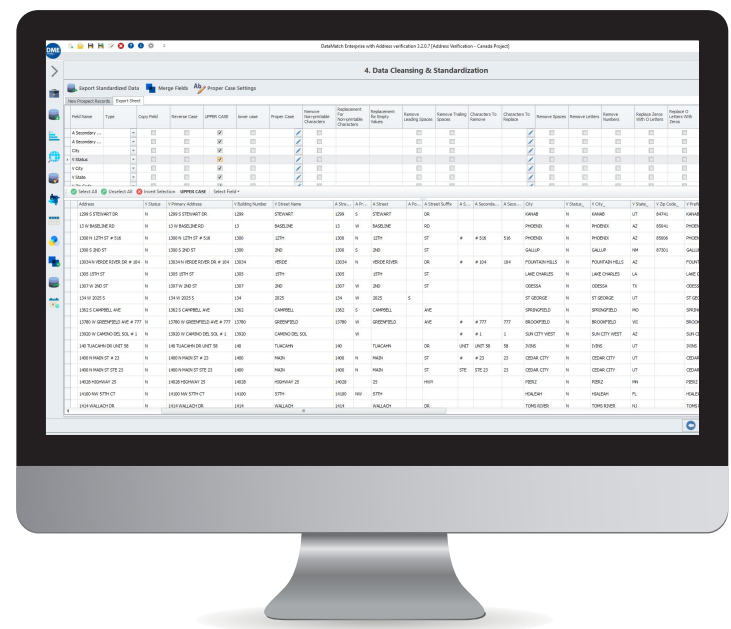


Prevent bad data from entering your systems with a powerful data quality firewall for perimeter protection across third-party and custom applications. The DataMatch Enterprise™ API splits and cases names and addresses, generates match keys for phonetic matching, generates 3-grams for more accurate fuzzy matching, and grades matching records - all in real-time - within or across systems and applications.

## Handling Difficult Search Problems

From spelling errors to redundancies, the platform works through many of the common issues found in large amounts of data.

- ✓ **Missing letters:** "Hammer" or "Hamer"
- ✓ **Variations:** "Vinnie Smith" or "Vinny Smith"
- ✓ **Extraneous letters:** "Folder" or "Foldwer"
- ✓ **Incomplete words:** "Cleaners" or "leaners"
- ✓ **Incorrect fielding in fielded data sets:** "Larry Jones" for "Jones Larry"
- ✓ **Incorrect or missing punctuation:** "World-class data" for "World class data"



# Sources and Integrations



Excel Files



Excel 2010



CSV/TSV



XML



Teradata



My SQL



PostgreSQL



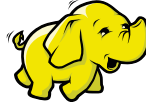
TXT Files



SQL Server



Oracle



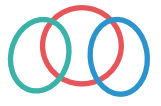
Hadoop



IBM DB2



MS Access



OLE DB



Facebook



Twitter



Salesforce



JSON



DBase - DBF



Microsoft Dynamics  
CRM



ODBC Driver



HBase

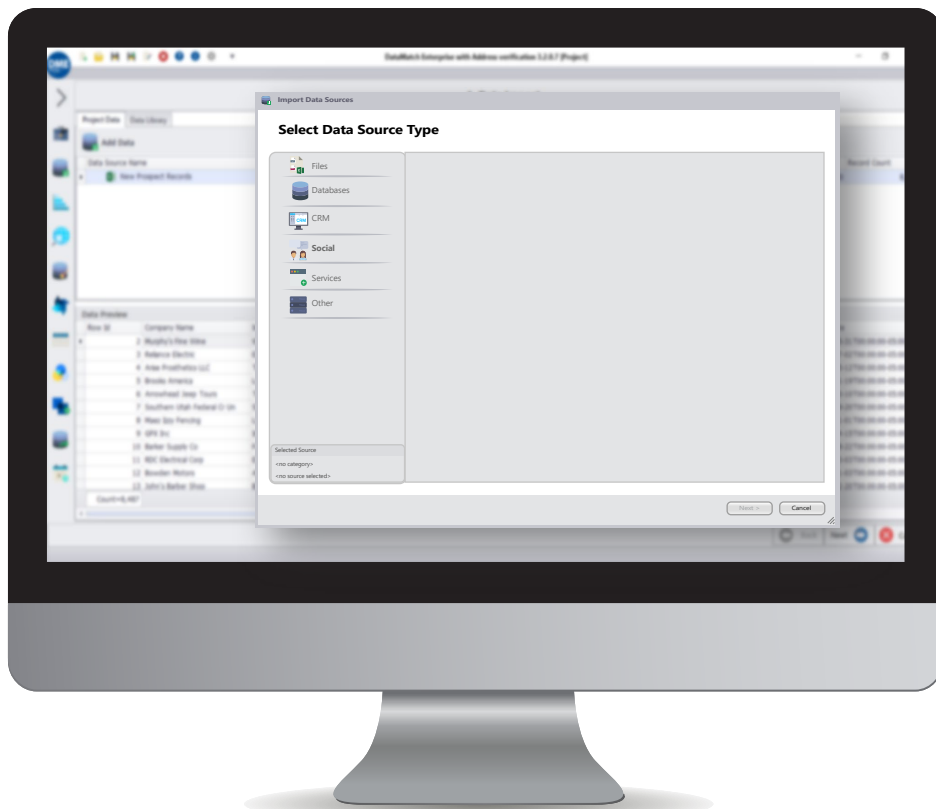


Sugar CRM



Mail

More Sources



# PC Specifications

DataMatch Enterprise™ has been designed for use on Windows based PC systems. To get the best results, you need to match the hardware to the job in hand. For smaller jobs where performance is not an issue, we would normally recommend a standard desktop machine that will no doubt be well above the minimum requirements (stated below). For larger data sets, you should consider upgrading the hardware according to the recommendations below. We have identified the main areas which affect the software's performance below.



**Processor Speed:** Processor speed or CPU speed (measured in GHz) controls the rate at which information can be processed. The higher the CPU speed, the faster the processing.

**Number of Cores:** The most critical parts of the matching process in DataMatch Enterprise™ are working in parallel; having more cores means better performances.

**Hard Disk Speed:** The faster the hard disk the better; DataMatch Enterprise™ is disk intensive which means the rate with which the data can be read and stored from and to the disk is an important factor. The higher the RPM the greater the data transfer rate is. Deployment of data and temporary files are covered by a separate document to optimize DataMatch Enterprise™ performance.

A SCSI interface will further increase the transfer rate from the disk thus being beneficial, especially, when processing large volumes of information (although SATA is sufficient for the great majority of organizations). Using a RAID optimized for performance can also be useful.

Our general recommendation is to use Solid State Drives (SSD). They have a proven speed superiority over traditional hard drives plus they do not suffer from regular disk fragmentation.

**Memory:** RAM is a key factor when processing large sets of data. Since DataMatch Enterprise™ is NoSQL software, a sufficient amount of RAM is critical.

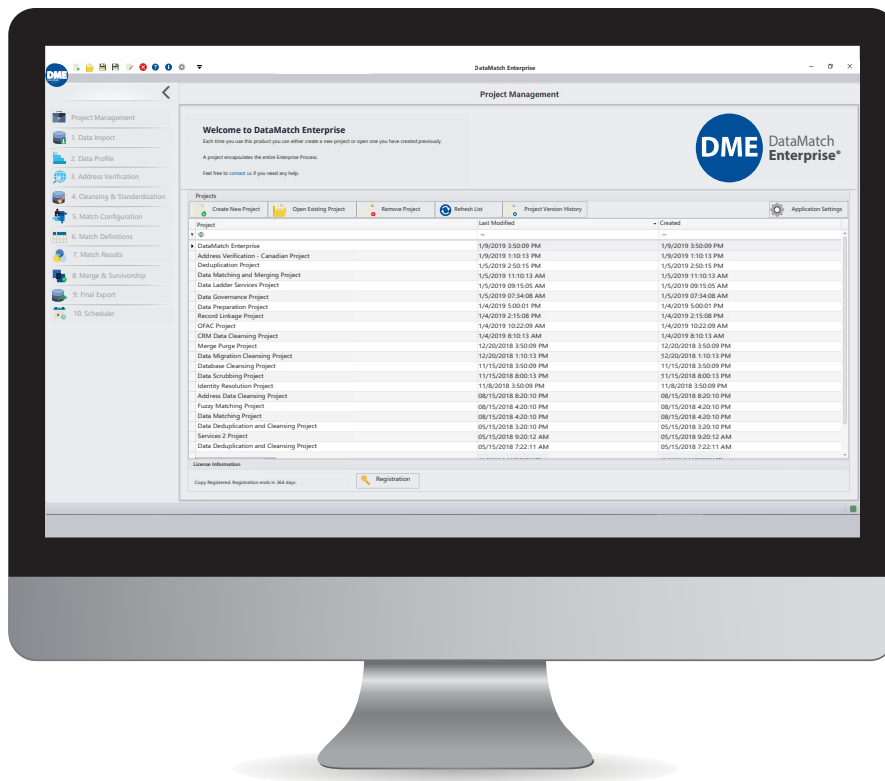
**Operating System:** DataMatch Enterprise™ is compatible with Windows operating systems from 7 onward (including server versions). We usually recommend using the most up-to-date version of a 64 bit operating system since this will benefit DataMatch Enterprise™ in terms of memory usage and robustness.



## Operating System

Optimum
4 or more cores
Intel Core i7
16 or more GB of RAM
SSD, or SCSI HDDs
SSD (Free space requirements depends on the size of input data)
1 TB Disk Space
Microsoft Windows 7 SP1 and above, 64 bit (Current Generation)

Minimum
Dual Core
X
4 GB of RAM
X
Hard Drive (free space requirements depends on the size of input data)
500 MB Disk Space
Microsoft Windows Sever 2003, Windows 7, 32 bit versions



## Our Customers





# ABOUT US

Data Ladder is a data quality software company dedicated to helping business users get the most out of their data through data matching, profiling, deduplication, and enrichment tools. Whether it's matching millions of records through our fuzzy matching algorithms, or transforming complex product data through semantic technology, Data Ladder's data quality tools provide a superior level of service unmatched in the industry.

## Why Data Ladder

It's simple: our user-friendly and powerful software helps business users across many industries manage their data more effectively and drive their bottom line. Our powerful software suite, DataMatch Enterprise, was proven to find approximately 5-12% more matches than leading software companies IBM and SAS in 15 different studies.

[Free Download](#)