



**Comparison Sheet**

# Data Ladder Vs. Trifacta



68 Bridge, St. Suite 307  
Suffield, CT 06708



+1 888-779-6578



Sales@DataLadder.com



www.DataLadder.com

# Table of Content

## Introduction

<b>1. Feature Comparison: DataMatch Enterprise (DME) vs. Data Wrangler Enterprise</b>	4
- Data integration .....	4
- Data Profiling .....	4
- Data Cleansing .....	5
- Data Matching .....	6
<b>2. Features in DME that are Not Available in Trifacta</b>	7
- On-Premises Solution .....	7
- Address Verification .....	8
- Survivorship and Data Enrichment .....	9
- Final Export .....	9
- Scheduler for Automating Data Cleansing and Matching .....	9
<b>3. DME vs. Data Wrangler Product Comparison Chart</b>	10
<b>4. Making the Choice Between Data Ladder and Trifacta</b>	11
- On the other hand .....	11
<b>5. User Reviews: Data Ladder vs. Trifacta</b>	12
<b>6. Conclusion</b>	13

# Introduction

Data quality is a looming crisis for businesses. At a time when most businesses are aiming to implement digital transformation initiatives, **poor data** is a bottleneck stifling their efforts.

As businesses realize the competitive advantage of improving their data, the demand for a fast, accurate and easy-to-use solution is on the rise. Most of the issues in data quality are caused by the lack of data governance processes, human input errors, disparate data sources, data isolation and the lack of a unified customer view.

There is a frantic search for a data quality solution that can help businesses clean, sort, fix and unify their data. Moreover, the data quality solution should be flexible enough to allow the organization to maintain data quality in real-time while ensuring that security and data compliance standards are met.

Data Ladder's **DataMatch Enterprise (DME)** solution is a one-stop data quality solution that offers data profiling, data matching, data cleansing and much more.

This comparison sheet will compare Data Ladder's DataMatch Enterprise with Trifacta's Data Wrangler, a data cleaning and preparation solution.

Let's get started.

# Feature Comparison: DataMatch Enterprise (DME) vs. Data Wrangler Enterprise

Trifacta provides data transformation services through its Data Wrangler tool. Users can convert raw data into a clean and usable form, converting data types, removing duplicate data, and enriching the data.

Data Ladder also provides data transformation services but it is not limited to just data cleansing. It offers data matching as an essential solution that allows users to match data across disparate data sources to get a complete view of their information. Additionally, it also offers a full-fledged address verification feature that is not available in Trifacta.

We'll begin the comparison by listing down common features offered by both solutions.

## Data Integration

As businesses continue to use new CRMs and apps, they will want a quality solution that connects directly to the data source.

**DataMatch Enterprise (DME)** allows for easy integration with 150+ apps. This means businesses can directly clean, edit, match and fix data quality issues right within their data source without having to worry about importing or moving data. You can integrate:

- Files (Excel, Delimited Text File, Excel (97-2003), Fixed Width Text File)
- Databases (Direct integration with SQL Server, Oracle, Teradata, and many more)
- CRM (Integration with Salesforce, MS Dynamics CRM, Sugar CRM)
- Social (Integration with Facebook and Twitter)
- Services (XML and JSON)
- Email apps
- Other (ODBC-supported databases)

**Trifacta** supports most major Big Data and relational data sources, but **lacks connectivity with most CRMs, social media platforms, and enterprise applications.**

## Data Profiling

The primary goal of data profiling is to provide an overview of the problems within the data set. For example, incomplete addresses, invalid phone numbers, ZIP codes with letters etc are data quality issues that are highlighted during the profiling process.

**DataMatch Enterprise (DME)** lets you:

- Identify which of your lists have irregularities or oddities
- Gives a confidence score to each of your fields
- Statistics on your dataset's size, median, mean, average, standard deviation, etc
- Tailor-made expressions that let you sort data over 19 different fields
- Create custom expressions using a Pattern Builder option
- View all your profile history changes
- Visual data exploration to help you see your audience clusters
- Identify nicknames from a pre-defined dictionary of nicknames

**Trifacta**, on the other hand, provides:

- Statistics such as the dataset's size, distribution, quality, distinct values, median, mean, quartiles, average, and standard deviation, to name a few.
- Percentage of valid, mismatched, and empty values in your results file.
- Number of columns in your results file.
- Number of rows in your results file.
- The ID of your result file.
- The data source that was used to generate your results file.

Trifacta's data profiling is quite basic and does not have advanced features such as the ability to create customized expressions to sort your data according to your preference. For instance, with DME, you get pre-built expressions that let you sort your data on the basis of credit cards, gender, email addresses, websites, and other important fields.

## Data Cleansing

Data fields with inaccurate spellings, case settings, and other structural problems need to be cleansed by fixing the inconsistencies. Ideally, this needs to be done right within the data source so that errors are nipped in the bud.

DME has a full-fledged data cleansing and standardization system that allows users to:

- Change cases and ensure that each field follows a defined standardization
- Find and replace different characters
- Choose whether you want to preserve abbreviations
- Set your own exceptions
- Choose delimiter that separates your fields such as tabs, comma, semicolons etc
- Remove spaces, numbers, letters, non-printable characters, zeros and zeroes with letters

DME also has a unique WordSmith tool that allows you to standardize names and addresses. This is especially useful if you want to ensure that a certain word is automatically replaced or removed in a list.

For data cleansing, Trifacta follows a six-step process that sums up all the features discussed above.

- 1. Discovering:** Gives an overview of the data.
- 2. Structuring** Formats the data to be used in traditional applications.
- 3. Cleaning** Removes data that may distort analysis and standardizes it into a single format.
- 4. Enriching** augment the data with internal or third-party data to enhance the data for better analysis.
- 5. Validating** brings data quality and inconsistency issues to the surface so the appropriate transformations can be applied.

Trifacta, being a data preparation solution focuses on multiple aspects of data cleansing. It allows users to standardize data and remove or correct inaccurate records.

DME stands out in the cleansing part because it allows for identifying and removing non-printable characters, replacement of empty values and non-printable characters as well as removing spaces, zero characters with a dozen more options that are not possible in Trifacta.

## Data Matching

Data matching is an important function of almost every data cleansing and preparation solution. That said, none of the popular solutions including Trifacta, SAS or IBM offer a full-fledged data matching function. Data Ladder is the only data matching solution that has the highest matching accuracy rate at 96%.

Most data matching solutions rely on fuzzy logic to match data, but DME has a proprietary algorithm that out beats every other solution in the market. A study conducted by Curtin University proved DME's accuracy at 96%, while IBM and SAS were at 84% and 83% respectively.

Additionally, Data Ladder offers the ability to perform cross-jurisdictional matching where users can match multiple data sources instead of lists within a specific data source. **Cross-jurisdictional matching** is an important feature that government and public sector organizations need to create research reports, receive grants and optimize government processes.

**DataMatch Enterprise's** data matching allows you to:

- Define how you want to match data sources using four options – all, none, between and within. All looks for matches between each data source as well as matches within each data source. If you don't want to do all, you can just choose between or within to look for duplicates between data sources or within data sources.
- Choose the fields you want to map by the fields of your data list.
- Run multiple matching sessions simultaneously.
- Get matches based on exact, phonetic, numeric and fuzzy matching techniques
- View match results and create a Master Record that you can use as a surviving record.
- Save and export your results for later review.
- Use a merge and survivorship option to overwrite data based on your criteria.

**DME** is specifically dedicated to providing a flexible data matching solution that gives the user the ability to do much more than just basic matching.

# Features in DME that are Not Available in Trifacta

Being a data transformation solution, Trifacta is a great software for companies that want an easy-to-use, easy-to-access, web-based solution. However, Trifacta lacks some core features offered by Data Ladder.

Let's take a look at features that are not available in Trifacta:

## On-Premises Solution:

DataMatch Enterprise is primarily an on-premises solution. It can be deployed on a company's own server and can be used behind your firewall. This ensures safety, security and complete control over your data assets.

Trifacta is purely a cloud-based solution that caters to organizations using Google Cloud, AWS or Azure as their cloud platforms. Trifacta's Data Wrangler Enterprise is a web-based solution that uses an API framework enabling users to access live data without requiring them to pre-load or create a copy of the data separate from the source data system. This framework includes connecting to various Hadoop sources, Cloud services, Files (CSV, TXT, JSON, XML, etc.) and relational databases.

While it is true that some things are easier to administer through the cloud, security is not one of them. By its very nature, security is something most organizations will want to keep in-house rather than turning over confidential data (even if encrypted) to a cloud provider.

Being an on-premise solution, Data Ladder automatically bypasses the bulk of security concerns. Meaning, the solution is installed behind the client's own security measures. Since the Data Ladder software is installed in your own environment (desktop or server), nothing from our end travels in or out of the software itself. All data processing is done in-memory, effectively rendering any concerns regarding data-at-rest or data-in-transit encryption inapplicable.

Where customers require the software to be hosted on the cloud, Data Ladder uses Microsoft Azure, ensuring top-of-the-line security and compliance.

Compliance framework and offerings include but are not limited to:

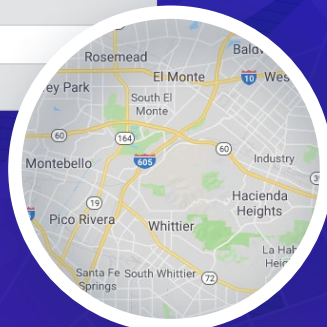


- CSA STAR Attestation based on independent audits
- CSA STAR Certification
- EU Model Clauses
- EU-US Privacy Shield
- EU-US Privacy Shield
- FISC
- HIPAA/HITECH
- UK G-Cloud
- ISO 9001
- ISO/IEC 27001
- PCI DSS
- SOC 1 and SOC Type 2 Reports
- SOC 3

## Address Verification

**DataMatch Enterprise** is a CASS certified address verification and validation solution that lets users verify address fields against government databases like the USPS and Canada Post among others. This feature is a "must-have" if you need to clean a mailing list. There are 54 fields according to which you can match, verify and validate your address. From checking 5 digits zip codes to identifying street names to longitude and latitude fields, you have a wide range of address validation field to use.

Input
<b>Address:</b> 11400 W Olympic BL # 200
<b>City:</b> LA
<b>State:</b> CA
<b>ZIP Code:</b> 90064



Output
<b>Address:</b> 11400 WEST OLYMPIC BLVD 200
<b>City:</b> LOS ANGELES
<b>State:</b> CA
<b>ZIP Code:</b> 90064
<b>Plus4:</b> 1584
<b>Delivery Point Check Digit:</b> 508
<b>Geocode Latitude:</b> 34.03
<b>Geocode Longitude:</b> -118.44

## Survivorship and Enrichment

Once you have performed the cleansing, matching, and fixing of your records, you might want to save that master record and keep it as the central record. DME allows for data survivorship and data enrichment where you can decide what record you will use to overwrite, determine master record or merge data. The merging of data allows you to create a complete master record. This is pretty handy if you have customer info dispersed over several platforms or channels. With this option, you can merge specific fields of those other data sources and create a complete picture of your customers.

## Final Export

What do you eventually do with the duplicates? With the final export option, you can export all of the records from your data sources and create a column for the match group ID in order to flag the matches and duplicates. The final export option lets you create a master record for each group of duplicates. Additionally, you can also choose to export duplicates only or use a cross-reference to export the results that have similar values in a cross-column format.

Whereas Trifacta allows you to Publish results on only major analytics and BI systems, rather than offering you the flexibility to choose exactly what part of the results to export, how to do so, and in which format.

## Scheduler for Automating Data Cleansing and Matching

A unique feature in DME, the scheduler allows the user to automatically run projects by creating time-based events. You can choose the date and time you want to run a data cleansing operation and the software will automatically start the process in the background. This is an important feature if you want to automate data cleansing and standardization without having to worry about leaving a project in the backburner.

# DME Vs. Data Wrangler Product Comparison Chart

## DataMatch Enterprise

Data matching at 96% accuracy for 40K records.

Offers data matching as a key feature along with other data quality operations.

Has on-premises and cloud-based versions.

Offers a complete data cleansing, data matching, data cleansing, data enrichment solution.

Has predefined rules based on experience working with 4500 clients which cover almost every major data quality issue.

Has a special WordSmith tool for parsing and standardization that is not offered by any other data quality solution.

Has a CASS certified Address Verification system that gives the user flexibility to add government databases and dictionaries for address verification.

Is designed for businesses with a focus on data matching, data cleansing, and data standardization.

Pricing the lowest in the industry, 90% less than SAS.

## Trifacta Data Wrangler

Data matching is not the primary function.

Offers basic text or pattern-based matching .

Does not have an on-premises solution.

Offers data preparation and data cleansing only.

Has very limited pre-defined rules even in the Wrangler Pro option.

Performs basic standardization and parsing operations.

Does not have standalone address verification and focuses only on the cleansing of inaccurate records.

Is designed for businesses that want a cloud-based solution.

High licensing costs with basic cost starting at \$419/Month.

# Making the Choice Between Data Ladder and Trifacta

**Trifacta's Data Wrangler** is primarily a cloud-based data preparation tool that allows users to prepare, clean and standardize data without the need for on-premises software. It is ideal for users who want a data preparation solution for improving their marketing or customer data. It is not an ideal solution for businesses that want a complete data matching solution along with data cleansing and data enrichment operations.

On the other hand:

**Data Ladder's** data quality products are designed to help business users get the most out of their data through data matching, profiling, deduplication, and enrichment tools. Whether it's matching hundreds of millions of records through our proprietary 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.

The company enables users to:

- Integrate multiple sources simultaneously from virtually any source system
- Profile their data within a couple of minutes to visually show data quality issues
- Clean it using point-and-click actions to transform their data visually
- Standardize it by creating custom libraries or using the 300,000+ pre-built rules
- Verify addresses and enrich existing records with geocoding data
- Deduplicate their data with a combination of proprietary and established matching algorithms
- Handle both structured and unstructured data with ease
- Leverage advanced semantic-based machine learning technology to process product data
- Link records across the enterprise for entity resolution and data enrichment purposes using proprietary fuzzy matching technology rated as the fastest and most accurate
- Deploy a data quality firewall to catch data quality issues before they make it into your systems

## Additional Features of Data Ladder

**Reviewers** claim that Data Ladder's tools are easy to use and offer unparalleled efficiency and productivity increases when the use was comparing disparate data sets, cleansing them, and implementing clean-data initiatives across the enterprise.

## User Reviews: Data Ladder vs. Trifacta

**"User-friendly, but lacks key features such as debugging, quick exploratory analysis/joins."**

**"Great tool, but hard to figure out new processes if you don't already know how to do them."**

**"Currently, Trifacta does not allow the ability to write data back to a postgre based RDBMS system. This causes us to need to perform additional steps outside of Trifacta to import and export data to their source and target systems."**

Data Ladder's biggest selling point is its ease of use, designed specifically for both business and IT users. Customers who have never used a data management tool before are able to create data quality projects easily within the modern, intuitive interface with basic training, as is evident from reviews:

**"DataMatch Enterprise is easy to learn and use. It's easy to review results. Saves us tons of time in manually checking records."**

**"I like that it is very straightforward and easy to use. With very little training, we had it up and running."**

**"This program is far more effective, user-friendly and reasonable than many other data cleansing programs I checked out prior to making my purchasing decision and their customer service is really great."**

**"I like the ease of use for this program when matching data from Access tables to confirm how successful our business practices are"**

### **When using Data Ladder software, customers especially liked the following:**

- The modern, highly intuitive, code-free interface.
- Extremely responsive and helpful support.
- Flexible creation of data quality rules for both business and IT users.
- Very easy to review result and see their data as it transforms.
- Straightforward setup - very little training required to get started.
- The ease of integrated data sources across the enterprise.

# Conclusion

The choice between Trifacta Data Wrangler vs Data Ladder's DataMatch Enterprise is one that is dependent upon your business requirement in terms of data initiatives. If you're looking for a solution that offers a low-cost, high-impact data cleansing and matching, then DME is the optimal choice. On the other hand, if you're looking for a cloud-based solution for data preparation and transformation purposes, Data Wrangler may be a better option provided, you don't have the need for address verification, export, merger or survivorship.

## Transform Dirty Data

Full Name	Address	Email
j smith	10 Main Street, springfield, 01234	Jsmith#JSLwaw.com
Jon Smyth	10 Main Street	NULL
jonathan smith law firm	10 Main Street, Apartment 304	NULL



## To Deduplicated, Cleaned and Merged Data

First Name	Last Name	Address	ZIP	City	Apt	Email	Industry
Jonathan	Smith	10 Main St.	01234	Springfield	304	Jsmith@JSLaw.com	Legal

# 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.

**DataMatch Enterprise, was proven to find approximately 5-12% more matches than leading software companies IBM and SAS in 15 different studies.**

- Unparalleled Matching Accuracy and Speed For Enterprise Level Data.
- Cleansing beating IBM and SAS.
- Proprietary Matching Algorithms with a high level of matching accuracy at blazing fast speeds on Desktop/Laptop Hardware.
- Big Data Capability with data sets up to 100 Million Records.
- Deduplication and Merge Purge within and across any number of files.
- Suppression of existing customers or Do Not Contact from marketing lists.
- Advanced record linking technology to create data warehouses.
- Quick Data Profile tool finds and fixes Data Quality issues within the first 5 minutes of setup to improve match quality.

[Free Download](#)