MatchUp®

DATA MATCHING
& RECORD
LINKAGE ENGINE

800.MELISSA (635.4772)
www.Melissa.com

melissa™

MatchUp – Affordable, Powerful, and Easy-to-Use Data Matching Software

On average, a database contains 8-10% duplicate records. These duplicates result in waste and inefficiencies that cloud your ability to get a single, accurate view of the customer.

Use Melissa MatchUp® to quickly find and link customer data, consolidate data across multiple sources and remove unwanted business and customer records – quickly and easily improving master data management, data warehousing and business intelligence, as well as marketing and mailing efficiency.

- **Combines data parsing, matching and consolidation to eliminate clutter and duplicates** that prevent a clear view of your customers.
- **Provides probabilistic matching and deterministic matching engines** for best-in-class identity resolution.
- **Can match structured and semi-structured data** with advanced record linking technology.
- **Reduces postage and mailing costs** by eliminating duplicates using advanced name and address matching + fuzzy algorithms.

MatchUp was designed around a heuristic-based approach with the aim to resolve customer and contact data quality issues by increasing the identification of duplicate records while minimizing false matches or minimizing matches between genuinely unique records. MatchUp includes proprietary and standard algorithms to detect phonetic, fuzzy, miskeyed, and abbreviated variations, combined with deep domain knowledge of names, including nicknames and multicultural name variations, and international addresses.

Technical Information

MatchUp provides easy integration into your applications and databases.

- **Technology:** on-premise API for Windows® and Linux®, RESTful web API, component for Microsoft® SQL Server® and Excel®, Pentaho® PDI, Salesforce®, Dynamics® CRM and desktop software.
- **Fast performance:** millions of records per hour for batch processing or real-time duplicate identification.

MatchUp Versions

MatchUp is a versatile data matching solution with many different options available to meet your individual business needs.

	MATCHUP DESKTOP	MATCHUP OBJECT	MATCHUP FOR ETL	MATCHUP FOR CRM	MATCHUP WEB
MatchCode Editor	✔	✔	✔		
Programming Required		✔			✔
Real-time Deduping		✔			
Global Processing	US, CA, UK	✔	✔		
Output Record Consolidation	Gathering		Survivorship		
Output Record Priority	Priority		Golden Record		
Direct File/Data Handling	✔		✔	✔	
CASS Certified Option	✔				
Automatic Reports	18 reports			✔	

MatchUp Desktop: ideal for data stewards and business users with an intuitive interface for fast, easy record deduping. **See Page 16 - 17.**

MatchUp Object: multi-platform, on-premise API that provides the ability for real-time and batch deduping and matching. **See Page 14.**

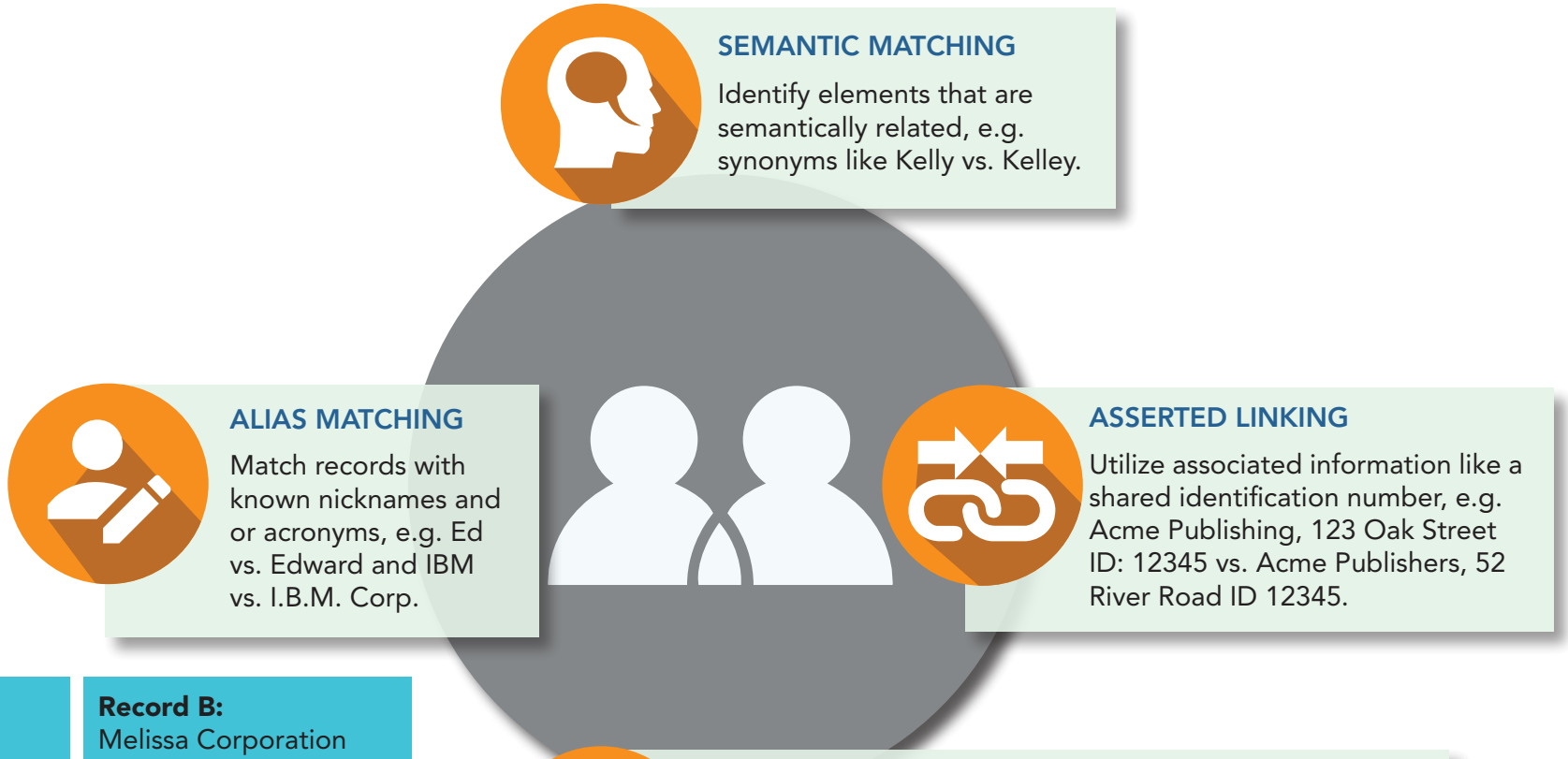
MatchUp for ETL: native integrations for leading ETL platforms including Microsoft® SSIS, Pentaho® PDI and Melissa Contact Zone®, with advanced golden record and survivorship features. **See Page 12 – 13.**

MatchUp for CRM: plugins for popular CRM programs including Salesforce®, Microsoft Excel® and Dynamics® CRM that provide clean data needed for effective sales and marketing.

MatchUp Web: versatile web service to identify and remove duplicates with no extensive installation, dedicated processors, or maintenance required. **See Page 15.**

Matching Systems

MatchUp supports a set of deterministic and probabilistic matching strategies to identify matches. MatchUp leverages Melissa's deep domain knowledge of many domains of data (address & non-address), custom component matching, and custom data dictionaries to identify hidden relationships in your data. MatchUp can examine records by many fields including, but not limited to: First Name, Last Name, Company Name, Address, Phone, Email Domain, and Website, and utilize any or all of these matching strategies to identify duplicates.



SEMANTIC MATCHING

Identify elements that are semantically related, e.g. synonyms like Kelly vs. Kelley.

ALIAS MATCHING

Match records with known nicknames and or acronyms, e.g. Ed vs. Edward and IBM vs. I.B.M. Corp.

ASSERTED LINKING

Utilize associated information like a shared identification number, e.g. Acme Publishing, 123 Oak Street ID: 12345 vs. Acme Publishers, 52 River Road ID 12345.

INFERRED LINKING

Infer connections between individual records from various data sets. In this example, by showing that Phil Smith and Melissa Corp are related – now you know Philip Smith works at Melissa, you have his phone number, alternate phone number, and email address.

Record A:
Phil Smith
949-858-3000
22382 Avenida Empresa
Rancho Santa Margarita

Record B:
Melissa Corporation
800-635-4772
22382 Avenida Empresa
Rancho Santa Margarita

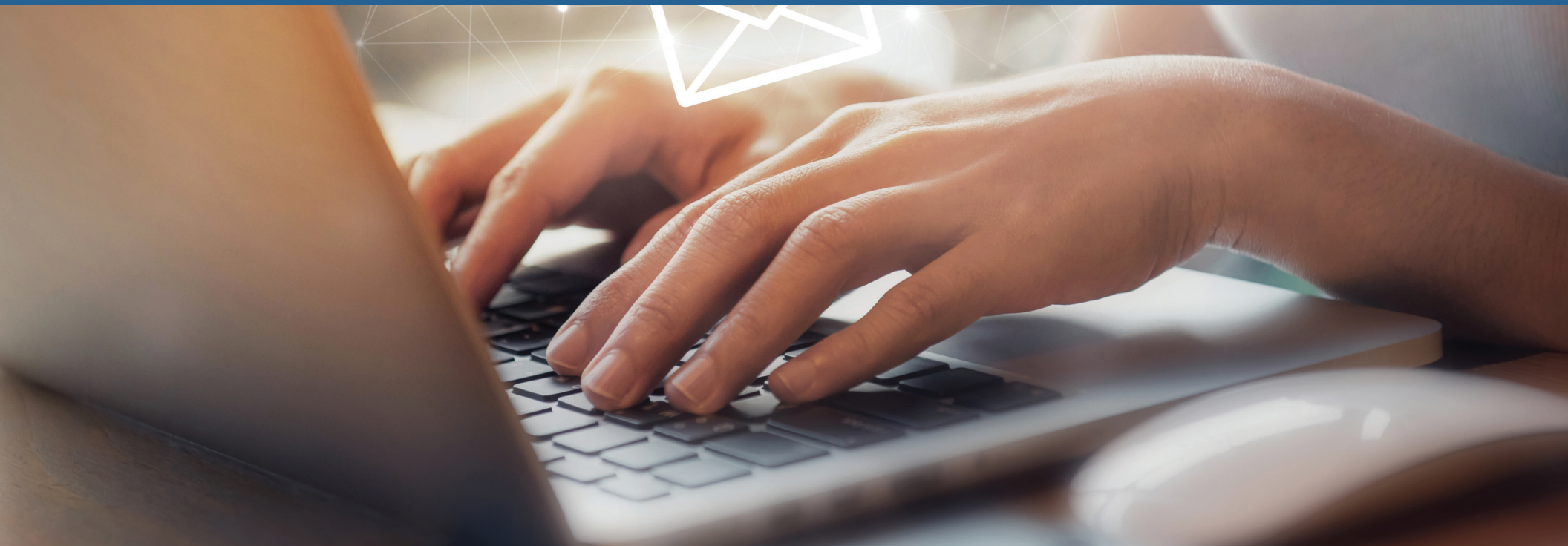
Record C:
Smith Phil
949-858-3000 x1110
22382 Avenida Empresa
Rancho Santa Margarita

Record D:
Smith Philip
949-858-3000
p.smith@melissa.com
RSM, CA 92688

Industrial Strength Fuzzy Matching

MatchUp combines Melissa's deep domain knowledge of contact data with more than 20 Approximate String Matching (ASM) algorithms to match similar records and quickly dedupe your database. MatchUp can employ the following ASM algorithms (including Melissa proprietary ones*) to identify "non-exact matching" duplicate records:

- Phonetex
- Soundex
- Containment
- Frequency
- Fast Near
- Accurate Near
- Frequency Near
- UTF-8 Near
- Vowels Only
- Consonants Only
- Alphas Only
- Numerics Only
- Jaro Distance
- Jaro-Winkler Distance
- n-Gram
- Needleman-Wunsch
- Smith-Waterman-Gotoh
- Dice's Coefficient
- Jaccard Index
- Overlap Coefficient
- Longest Common Substring
- Double Metaphone
- MD Keyboard *
- Proximity Matching*



Domestic and International Matching

Domestic (US & Canada) Matching

MatchUp leverages Melissa's 30+ years of experience working with numerous U.S. and Canada contact data idiosyncrasies. Based on this knowledge base, MatchUp has built-in rules and logic to handle Address Obscurities, Nicknames and Abbreviations, Acronyms, Different Formatting, and more.

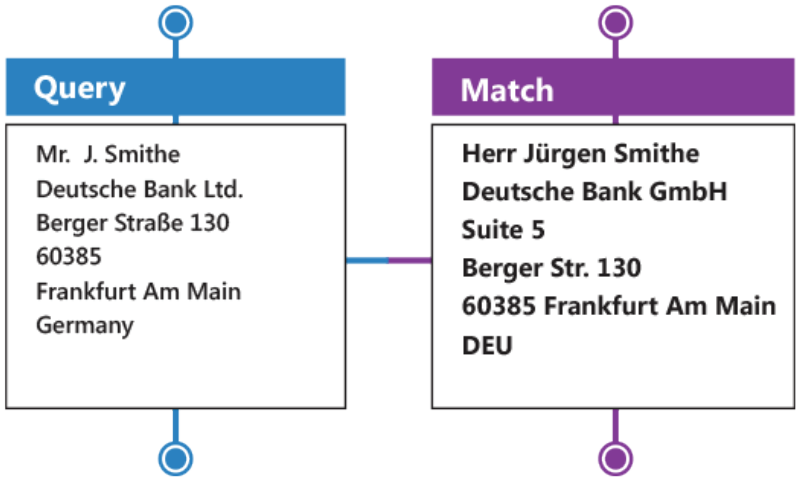
John Smith	United Data Machines	12 Main St	Boston, MA 02134
John Smith	United Data Machines Co.	Twelve Main St	Boston, Massachusetts 02134
John Smith	UDM	12 Main Street	Boston, MA 02134

International Matching

What sets MatchUp apart from the rest is its ability to intelligently understand and parse the various components of international addresses with support for 17 countries, including Germany, U.K., Australia, Netherlands, and new countries added every quarter.

MatchUp handles:

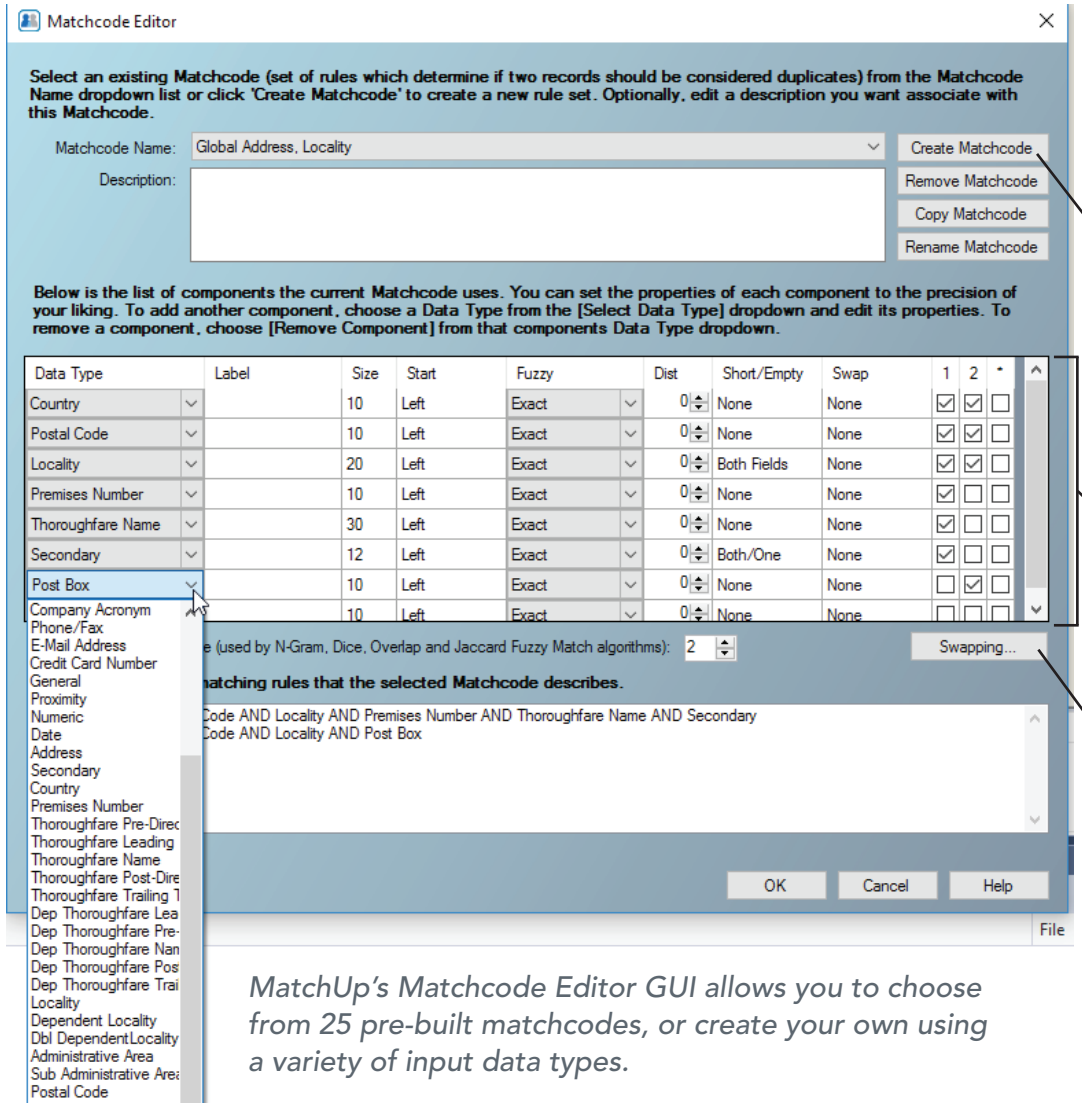
- Extended character sets
- Different languages and scripts
- International style address fields



MatchUp's advanced deduping can see through diacritic equivalents to Latin characters and interpret keywords that are the same but spelled differently (i.e. Germany and DEU).

Matchcode Editor

MatchUp's Matchcode Editor GUI allows you to choose from 25 pre-built matchcodes, or create your own using a variety of input data types. Choose from more than 50 distinct component datatypes (and a general catch all for proprietary data types), each of which can be configured individually – size to use, exact or one of 20+ fuzzy algorithms, blank matching and up to 16 simultaneous combinations of components in a single run.



MatchUp's Matchcode Editor GUI allows you to choose from 25 pre-built matchcodes, or create your own using a variety of input data types.

Mapping Your Data

The data you are processing does not need to have the same structure as the Matchcode components. Once you have selected your matchcode, MatchUp will extract the necessary data to satisfy your matchcode and build the matchkey correctly.

MATCHCODE COMPONENT	DATABASE CONTAINS	FIELD MAPPING NEEDED	KEY BUILT
Zip5	RSM, CA 92688-2112	CityStateZip	92688
Last Name	Smith Jr., John L.	Inverse Name	Smith
First Name	Smith Jr., John L.	Inverse Name	John
Street Number	12 North Main Street Apt. 2	Address	12
Street Name			Main
PO Box			

92688SMITH	JOHN	12	MAIN

Select one of MatchUp's supplied match codes or create your own

Process up to 16 match codes simultaneously

Swap matches like "John Smith" to "Smith, John"

Matchcode Logic

MatchUp utilizes matchcodes, or sets of rules, to determine if two records are considered duplicates. These matchcodes contain data types, size, order and combinations, and include logic for handling empty comparisons and swapping.

EXAMPLE 1
Global Address, Last Name, First Name - Exact

Matchcode Data Type	Input Column	Input Data Type
Country	COUNTRY	Country
Last Name	NAME	Full Name
First Name	NAME	Full Name
Address Line 1	ADDRESS1	Address
Address Line 2	ADDRESS2	Address
Address Line 3	ADDRESS3	Address
Address Line 4	[Select Ma...	[Select Data T...
Address Line 5	[Select Ma...	[Select Data T...
Address Line 6	[Select Ma...	[Select Data T...
Address Line 7	[Select Ma...	[Select Data T...
Address Line 8	[Select Ma...	[Select Data T...

Using the Global Address matchcode with these input data types, MatchUp returns these duplicate records in group 2:

mu_RESULTS	mu_GROUP	mu_COUNT	NAME	COMPANY	ADDRESS1	ADDRESS2	ADDRESS3	ADDRESS4	COUNTRY	ACCT	DATE
MS01	1	1	Dr. Grace Johnson		1001 High Street		London	England	UK	60	18/4/2014
MS02,MS06	2	3	Ms. Anna Jones	AGT Healthcare	61 Wellfleet Road	Apartment 2	CF24 3DG	Cardiff	UK	400	13/2/2016
MS03,MS06	2	3	Ms. A. Jones	AGT	61 Wellfleet	Cardiff	CF24 3DG		United Kingdom	200	1/12/2005
MS03,MS06	2	3	Anna Jones	AGT Healthcare Associates	Apartment 2	61 Wellfleet Road	Cardiff	CF24 3DG	England	700	
MS01	3	1	Leslie Rogers	RGNT Consulting	91 Western Road	Brighton	BN1 2NW	East Sussex	United Kingdom	880	6/5/2011
MS02,MS06	4	2	JOHN SMITH	First Bank Sussex	91 Western Road	Brighton	East Sussex	BN1 2NW	UK		20/7/2008
MS03,MS06	4	2	J. SMITH	First Bank	91 Western Rd.	Brighton	BN1 2NW	England	Great Britain	100	18/3/2012
MS01	5	1	Leslie Smith	RGNT Consulting	91 Western	BN1 2NW	England		United Kingdom	300	22/9/2014
MS02,MS06	6	2	Annabell Johnson		456 High Street	London		England	UK	735	12/12/2011
MS03,MS06	6	2	Annabelle Johnson		456 High St.		London	England	UK	950	6/11/2014

EXAMPLE 2
Global Address, Last Name, First Nickname

Data Type	Label	Size	Short/Empty	1	2
Country		10	None	☑	☑
Last Name		10	None	☑	☑
First Nickname		4	Initial	☑	☑
Postal Code		10	None	☑	☑
Premises Number		10	Both Fields	☑	☐
Thoroughfare Name		30	None	☑	☐
Secondary		12	Both/One	☑	☐
Post Box		10	None	☐	☑
[Select Data Type]		10	None	☐	☐

By replacing First Name with First Nickname and allowing for an initial, MatchUp will catch names like 'A.Jones' to 'Anna Jones,' and 'J. Smith' to 'John Smith.'

EXAMPLE 3
Global Address, Last Name, First Name - Fuzzy

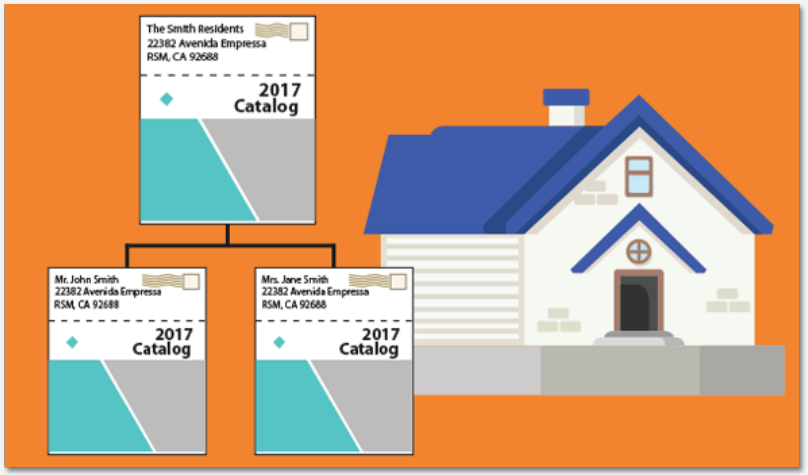
mu_RESULTS	mu_GROUP	mu_COUNT	NAME	COMPANY	ADDRESS1	ADDRESS2	ADDRESS3	ADDRESS4	COUNTRY	ACCT	DATE
MS01	1	1	Dr. Grace Johnson		1001 High Street		London	England	UK	60	18/4/2014
MS02,MS06	2	3	Ms. Anna Jones	AGT Healthcare	61 Wellfleet Road	Apartment 2	CF24 3DG	Cardiff	UK	400	13/2/2016
MS03,MS06	2	3	Ms. A. Jones	AGT	61 Wellfleet	Cardiff	CF24 3DG		United Kingdom	200	1/12/2005
MS03,MS06	2	3	Anna Jones	AGT Healthcare Associates	Apartment 2	61 Wellfleet Road	Cardiff	CF24 3DG	England	700	
MS01	3	1	Leslie Rogers	RGNT Consulting	91 Western Road	Brighton	BN1 2NW	East Sussex	United Kingdom	880	6/5/2011
MS02,MS06	4	2	JOHN SMITH	First Bank Sussex	91 Western Road	Brighton	East Sussex	BN1 2NW	UK		20/7/2008
MS03,MS06	4	2	J. SMITH	First Bank	91 Western Rd.	Brighton	BN1 2NW	England	Great Britain	100	18/3/2012
MS01	5	1	Leslie Smith	RGNT Consulting	91 Western	BN1 2NW	England		United Kingdom	300	22/9/2014
MS02,MS06	6	2	Annabell Johnson		456 High Street	London		England	UK	735	12/12/2011
MS03,MS06	6	2	Annabelle Johnson		456 High St.		London	England	UK	950	6/11/2014

Change the matchcode for First Name from exact to a fuzzy algorithm, and MatchUp will match Annabell Johnson and Annabelle Johnson as duplicates. Fuzzy algorithms can be used on many different data types.

Matching Techniques

Householding

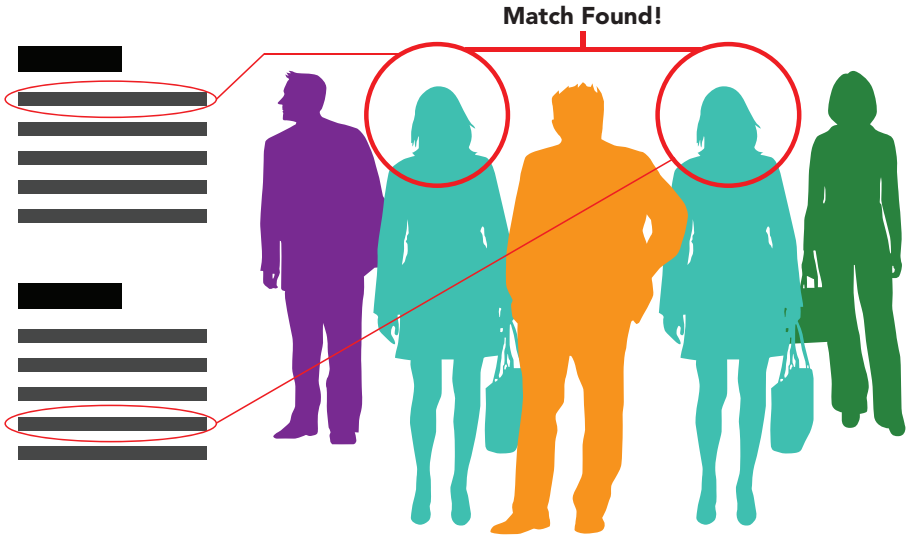
MatchUp can identify and consolidate records that are members of the same household to better understand customer relationships, lifecycle, and needs. Use MatchUp to bring together multiple business accounts into "corporate families" to build insight and better evaluate the total sales relationship.



Householding can also be used to eliminate unnecessary multiple mailings to the same household to cut down on wasted print, production, and postage costs. For instance, if you only want to mail to the oldest female in the household, set the priority for the records at that household with a custom expression for sex and DOB.

List Intersection

MatchUp saves you from the boring task of comparing two or more lists and performing logical operations such as difference or intersect operations. MatchUp does this by automatically comparing two or more lists to identify similar lines and to save them into a single output list.



Matching Techniques

Proximity Matching

MatchUp's patented distance algorithm enables distance criteria to be used in matching customer records, capitalizing on latitude, longitude, and proximity thresholds to help data managers eliminate duplicate records. This allows for the detection of matching records at different addresses but within a specified distance from each other.

Using location attributes, MatchUp can confirm that a home address and post office box actually belong to the same individual, or match buildings with different addresses by identifying different entrances common to large campus-style facilities.

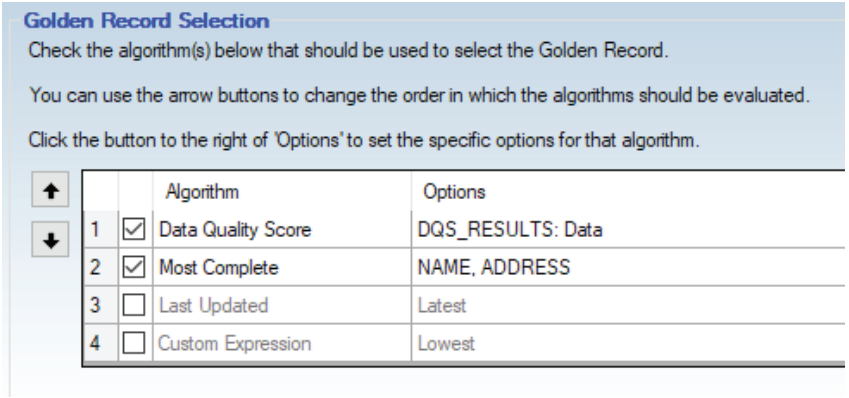
MatchUp for ETL - Record Management for Matching

Golden Record

MatchUp's Golden Record Selection options allow for intelligent selection of the Golden or Master Record from a group of duplicate records, using different logic and algorithms (based on a hierarchy you provide). Using the Golden Records function helps you achieve a single, accurate, and complete version of each customer record.

MatchUp can identify the best record of a matched group based on virtually unlimited criteria including:

- Data Quality Score (most accurate address, name, phone and email info)
- Last Updated
- Most Complete
- Custom Expression



The Data Quality Score criterion is unique to MatchUp. It leverages reference data to determine the best record based on the most accurate address, name, phone and/or email information. This same technique can be applied to product data, or virtually any other kind of data.

NAME	ADDRESS	ZIP	DQS_RESULTS	MU_RESULTS
John Doe	21907 64th Ave. W Ste 360	98043-6202	AC03, AS01, NS01, NS05, NS06	MS02, MS06
Mickey Mouse	21907 64th Ave. W Ste 360	98043-6202	AC03, AS01, NE04, NS02, NS05, NS06	MS03, MS06
Doe	21907 64th Ave. W	98043	AC02, AC03, AE09, AS02, NS01, NS06	MS03, MS06
J Doe	21907 64th Ave. W Ste 4	98043	AC03, AE08, AS02, NS01, NS06	MS03, MS06
John Doe	21907 64th Ave. W	98043	AC03, AE09, AS02, NS01, NS05, NS06	MS03, MS06

MatchUp will match partially verified addresses to completely verified addresses, but gives priority to good addresses. This makes for a better decision in the survivorship schema as opposed to selecting the most frequent. Taking Name Quality into account will prevent a fully verified address with a nuisance name from being chosen as the Golden Record.

Survivorship (Available for Microsoft SSIS, Pentaho PDI and Melissa Contact Zone only)

MatchUp's Survivorship functionality allows you to consolidate all your duplicate records into one master record that "survives" the merge process. This allows you to establish criteria for each data column individually instead of using the default method of writing full record content to output.

NAME	ADDRESS	PHONE	EMAIL	AMOUNT	DATE
John Doe	21907 64th Ave. W	949-777-HOME		420	10/25/2013
J Doe	21907 64th Ave. W Ste 4	949-700-CELL	john@xyz.com		10/25/2013
John Doe	21907 64th Ave. W Ste 360		john@xyz.com	10	8/6/2010
Doe	21907 64th Ave. W	949-444-WORK			8/6/2010
Mickey Mouse	21907 64th Ave. W Ste 360			200	4/16/2017

Using the Golden Record selection by latest update, the selected remaining record would be:

NAME	ADDRESS	PHONE	EMAIL	AMOUNT	DATE
Mickey Mouse	21907 64th Ave. W Ste 360			200	4/16/2017

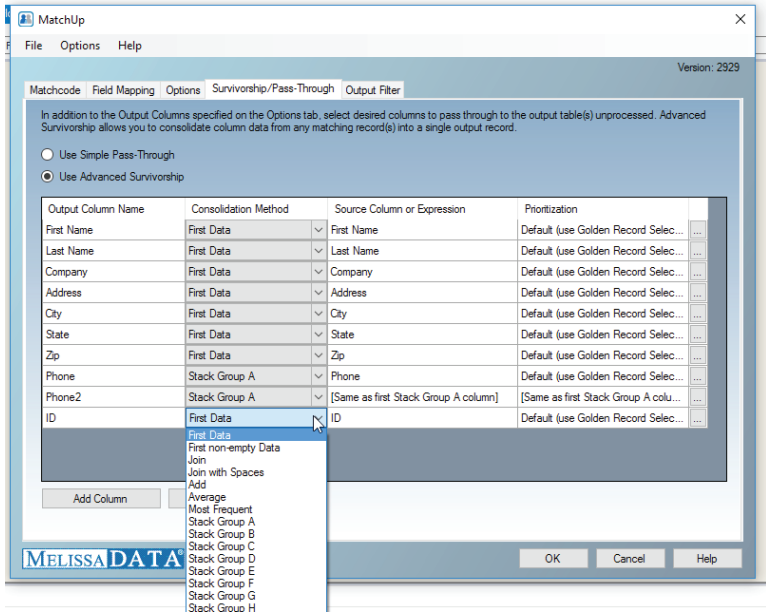
If we based Golden Record selection by largest Amount, the selected remaining record would be:

NAME	ADDRESS	PHONE	EMAIL	AMOUNT	DATE
John Doe	21907 64th Ave. W	949-777-HOME		420	10/25/2013

Using just Golden Record to determine output would result in loss of data, but, if we configure Survivorship rules, we can roll up data from other matching records into the Golden Record.

NAME	ADDRESS	PHONE	PHONE2	PHONE3	EMAIL	TOTAL
John Doe	21907 64th Ave. W Ste 360	949-777-HOME	949-700-CELL	949-444-WORK	john@xyz.com	630

Here we've used Advanced Survivorship consolidation methods to gather matching phone numbers, select the first non-blank email, and add the 'Amount' fields of matching records – all in a single output record.



Combining Golden Record Selection with Survivorship on a field-by-field basis gives you unmatched flexibility in enterprise-level record selection and consolidation, removing the tradeoff between concise and complete in data storage.

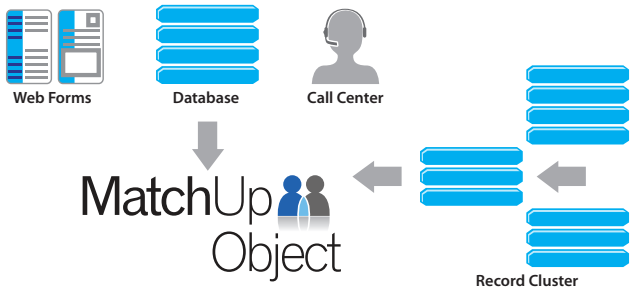
MatchUp Object - Real-time Matching for Developers

Easily integrate MatchUp Object (on-prem API) into your website or applications so as a new record is created, MatchUp instantly identifies whether it is a duplicate. Advanced name matching eliminates typos, nicknames, extra and missing information, acronyms and suffixes.

In fact, MatchUp Object is so flexible you can choose from three different methods of operation (or ways to match records) based on your business needs: match "one to "one, "many to one", or "many to many" for intersection, merge, and supposition techniques.

Read/Write Deduping – compares records and one or more databases at once. Each unique group will have one record that receives an "output" status; the other matching records receive a "duplicate" status - ideal for matching entire databases at one time.

Incremental Deduping – enables real-time matching by comparing each record as it comes in (like from a web form or call center) against the existing master database. If the incoming record is not a duplicate, it can be added.



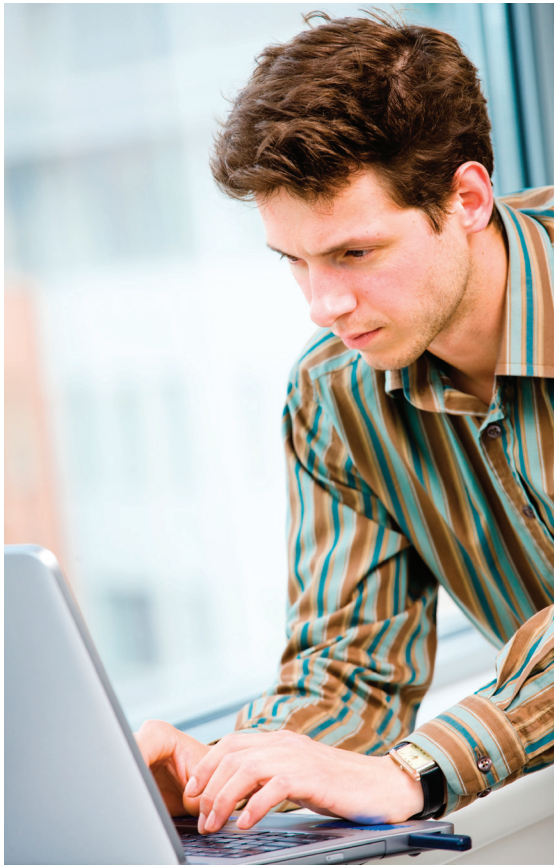
Hybrid Deduping – provides a combination of the first two methods with the flexibility to customize the process to match an incoming record against a small cluster of potential matches. With hybrid deduping, you can store the match keys in a proprietary manner - ideal for real-time data entry or batch processing of entire lists.

MatchUp Web – Deduping Made Fast and Easy for Developers

MatchUp Web is an easy to implement service. There are no libraries or data files to maintain and it can be as easy as sending sample requests through your browser for batch match processing.

Rapid Application Development

Rapid application development is made easy for any language-platform-protocol combination that can call a web service.



MatchUp Web – What is it Good For?

✔ When you need to call from an environment or language which Melissa doesn't support.

Examples: nodeJs customers – no easy way to call C++ library from JavaScript.

✔ Where you need to integrate into an application where a service can be called, but linking to the Melissa library and data files is not easy.

Examples: CRM integrations, updated OS, DBMS, or IDE, or where everything you do is in the Cloud.

✔ Where advanced MatchUp options are not important but you want to link matching records.

Examples: You want to group records and the default matchcodes available fit your needs.

Available Matchcodes

MatchUp Web provides the following matchcodes:

- Address
- Address, Company
- Address, Company AND Last Name, First Name
- Address, Company OR Last Name, First Name
- Address, Fuzzy Last Name, Fuzzy First Name
- Address, Last Name
- Address, Last Name, First Name
- Name AND Address OR Phone OR Email
- Phone OR Email

MatchUp Desktop

MatchUp Desktop can be used by anyone, not just IT personnel. No programming is required, just employ MatchUp Desktop's straightforward interface and powerful matching capabilities to:

- **Merge Duplicates**
  Easily find, merge & purge duplicates without losing data
- **Validate and Standardize Addresses**
  Stop losing money by mailing to undeliverable addresses
- **Improve Sales Productivity**
  Prevent multiple teams from interacting with the same contact unknowingly
- **Get a Clearer Picture**
  Produce clean, reliable data for a better view of your pipeline, accurate reporting and better analyses

Enterprise-Class Matching Available to Everyone

MatchUp Desktop provides complete address quality, matching and de-duplication in one easy-to-use software. MatchUp Desktop cleans your mailing lists, database, and Excel spreadsheets, then dedupes your data with advanced technology previously available only in enterprise level solutions.

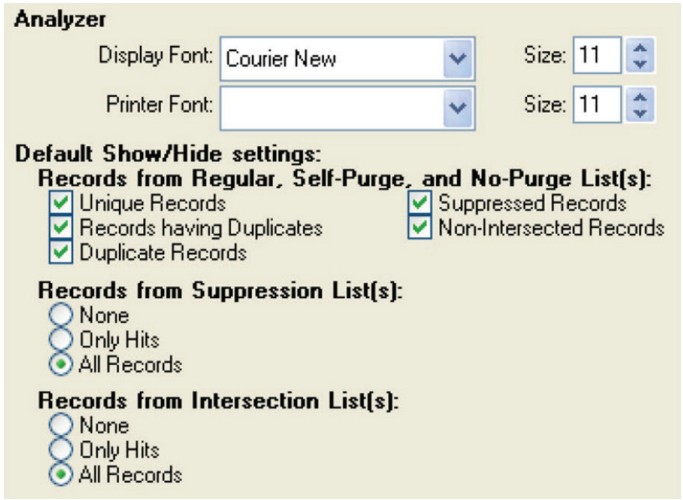
- Verify and standardize addresses with a CASS Certified™ engine
- Suppress records that match a suppression list for exclusion
- Intersect records that do not match an intersection list for exclusion
- Self-purge by finding matches in self, intersection, and suppression tables only
- No Purge to match a table to intersection and suppression lists but not between tables
- Custom rules using valid expression to hone advanced matches

Fast Data Matching in Just 5 Steps

MatchUp Desktop makes the process of deduping your database exceedingly simple and visually intuitive. It takes only 5 simple steps for you to start and finish your project - in minutes

1. Select a Setup
2. Select a Matchcode
3. Add Your Data (Source Table)
4. Process
5. Review Results

MatchUp Analyzer



MatchUp Desktop's Analyzer allows you to change processing results, which records are output, which are considered dupes – and record groups can easily be shown or hidden so 'clutter' records don't distract you.

Mail List Processing – CASS Certified™

MatchUp Desktop is CASS Certified by the USPS®. Coding Accuracy Support System (CASS) is a certification system for address validation. A CASS-certified address validation solution like MatchUp Desktop will standardize your mailing list, update outdated addresses, and verify that addresses are valid and complete. Using CASS with merge/purge process in MatchUp Desktop means in just one pass, deduplication is more accurate, you'll spend less time processing, you'll get more records corrected and validated to the ZIP+4® level, and ensure your mailing reaches more recipients.

Real World Solutions for Real World Problems

Nothing matches up to MatchUp. But don't take our word for it. Here's what some of our customers have to say.

> *It de-duplicates our customer data in an effective way so that we are able to reduce marketing costs and increase the quality of communication with customers.*

'We are using it for daily 1) direct matching and 2) column-level survivorship/ golden record generation for millions of customer records and 3) mail household-ing. We started with B2C customers and later added B2B customers. The tool supports unique matching specific to organization names and individual names (as well as a variety of other specialized types of data values) and works well in both cases.'

Gary M, Data Architect

> *Match process provided reliable single view of data.*

'Very straightforward installation & configuration steps. We configured the Match process as part of our daily SSIS ETL load process. Match process provided reliable single view of data. We used it for de-duplication, matching & mastering customer records.'

Mohan P, Director of Business Intelligence

> *MatchUp has been a huge help.*

'We are often tasked with merging many contact lists from a variety of formats. Due to human involvement there are often slight differences so we needed a package of programs that could standardize all the information and then efficiently compare records. One of our projects had over 4 million records going in and several thousand unique ones at the end. Developing this process on our own would have been cost prohibitive.'

David M, Consultant

> *MatchUp is reliable!*

'We use MatchUp daily with our client databases. We use it to standardize addresses for merge/purge projects, to determine which addresses will not DPV and should be removed from a direct mail file, etc.'

G2 Crowd User in Marketing and Advertising

Check out the full reviews and other customer reviews for Melissa products and solutions on G2Crowd, ITCentralStation, and Gartner Peer Insights.

WHY MELISSA?

Data is one of the most valuable assets you have—accurate, high quality customer data empowers you to do business with anyone, anywhere in the world and deliver exactly the experience necessary not only to help drive sales, but to improve loyalty, trade, business intelligence, technology, logistics, and more. Since 1985, over 10,000 companies around the world have counted on Melissa to harness the value of their Big Data, legacy data, and people data (names, addresses, phone numbers, and emails) to drive insight, maintain data quality, and support global intelligence.

Here's what sets us apart:

- Free unlimited worldwide tech support
- 99% uptime on our servers
- Free trials with absolutely no risk

Time Inc. · BAE SYSTEMS · SIEMENS · NORDSTROM · Bank of America · verizon · AAA · Ford · HYUNDAI · FedEx · GOV.UK

MatchUp Data Matching Software – Try it for Free: www.melissa.com/matchup

US

22382 Avenida Empresa
Rancho Santa Margarita, CA 92688-2112

800.MELISSA (635.4772)

www.Melissa.com

UK

+442030510140

GERMANY

+493079788829

INDIA

+918065603282

AUSTRALIA

+61280917500