



Entity Resolution

GUIDE

Entity resolution (ER), also traditionally known as record linkage or record matching plays a fundamental role in resolving real-world challenges, e.g., in a fraud prevention system where a robust entity resolution system can detect multiple records describe the same real-world entity (in this case, a person or an identity) and consolidate them into a single consistent record, thereby preventing fraudulent accounts from causing real-world harm.

In this guide, we'll walk you through the basics of entity resolution and also introduce you to WinPure, a data quality management tool equipped with Sensing Inside™ an AI powered engine purpose-built for entity resolution.

P.S: This guide is for business & tech users who are exploring entity resolution.

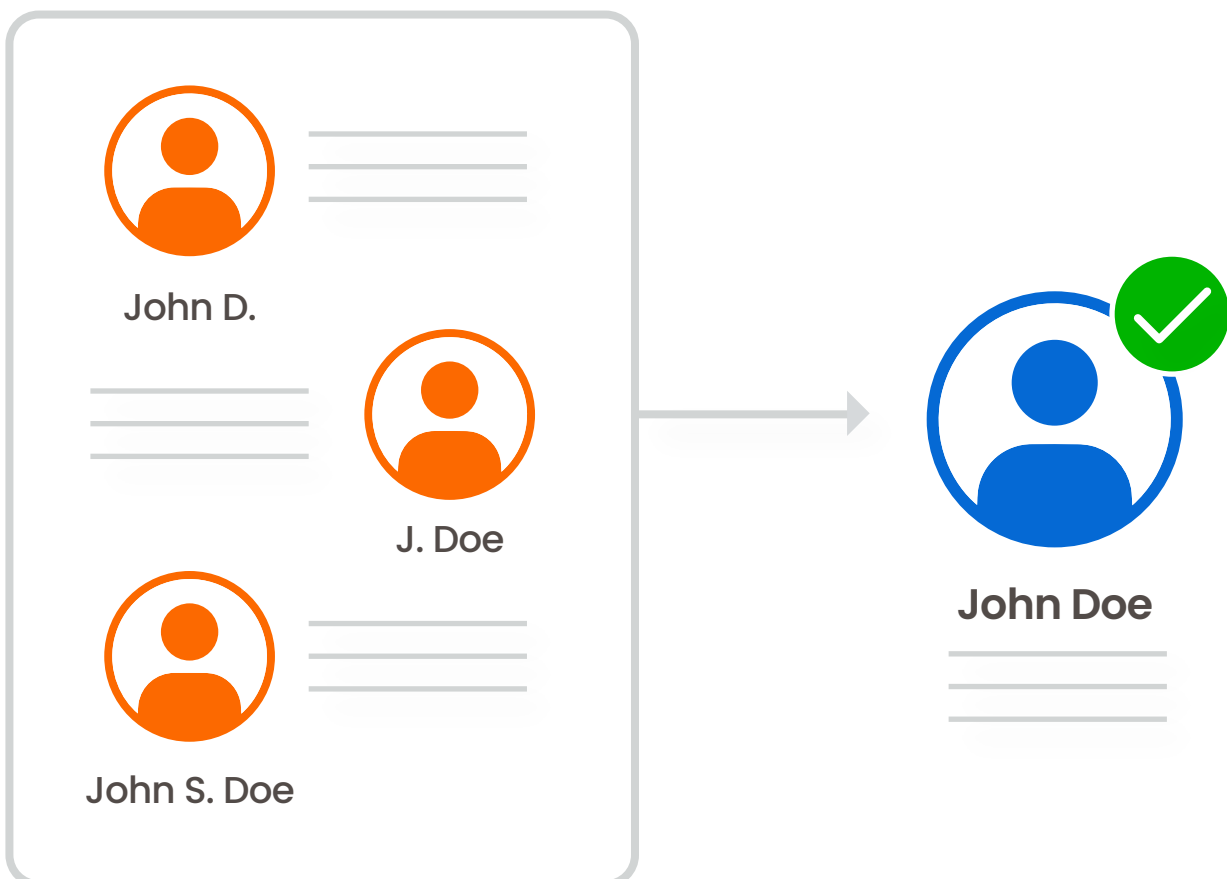
A Quick Introduction to Entity Resolution

It's not an exaggeration when companies claim data is their new lifeline – especially since they rely on this data to derive valuable insights, make critical business decisions, and even attempt to predict the future.

But unlike a few decades ago, the structure and nature of data has changed drastically. Companies are now collecting **overlapping information** about the same real-world

entity from multiple sources at a time.

When this data is fed into systems and platforms designed to draw insights or extract meaning, it needs to meet a key quality criterion – it must not be redundant or duplicate.



Entity resolution is the process of treating and linking heterogeneous data through the use of techniques like fuzzy matching and deterministic matching – and now – , with artificial intelligence.

Here's an example of entity resolution in action:

A long time customer of ABC bank, Mary has used several of the bank's services and has also interacted with several touchpoints where she has submitted varying information. Over the past few months though, Mary has been receiving her bank statement at her old address instead of her new address.

Real-world Entity



Full Name	Mary Jane Smith
Address	High St, Standlake, Oxfordshire
Phone Number	07424868565
Email Address	maryjs@gomail.com
Credit Card Number	5466 XXXX XXXX 9741
Membership Number	874521548
Loyalty Card Number	125880
Instagram User ID	@maryjanes

Digital Manifestations



Source System	Transactions
Credit Card Number	5466 XXXX XXXX 9741
Credit Card Name	Mary Jane



Source System	Loyalty
Loyalty Card Number	125880
Full Name	Mary Jane Smith
Email	maryjs@gomail.com
Phone	07424868565



Source System	E-commerce
Membership Number	874521548
First Name	Mary
Last Name	Jane Smith
Address	High St, Standlake, Oxfordshire
Phone	07424868565
Credit Card Number	5466 XXXX XXXX 9741



Source System	Social Media
Full Name	Mary Jane Smith
User ID	@maryjanes
Email	maryjs@gomail.com

Why so?

Because the bank may have failed to update and consolidate Mary's information!

Entity	First Name	Last Name	Phone	DoB	Address	Product
Record 1	Mary	J.Smith	1865300239	10.06.1975	High St, Standlake, Oxfordshire	Platinum Business Card
Record 2	Mary	Jane	7424868565	10/01/1975	High St, Standlake, Oxfordshire	Loan Application
Record 3	Mary	Jane Smith	7424324869	10.06.1975	29 Gosden Rd, Sussex	3rd Party Platform

Multiple Phone Numbers

Date Error

New Address

Not only does Mary have varying information but her records also have typographical errors – not to mention, she goes by different variations of her name. When the bank wants to extract this data for downstream applications, it will first need to 'clean', 'standardize,' and finally, 'deduplicate' the data before it can be used.

Entity resolution has one end goal:

Clean, duplicate-free, unified data that you can trust.

Understanding the importance of entity resolution

Data is textual or numeric information about an individual.

We are all identified in some database as a number – for example, we are all citizens assigned a unique ID number in a government database, or we are unique patient IDs in a hospital database.

Why all these numbers?

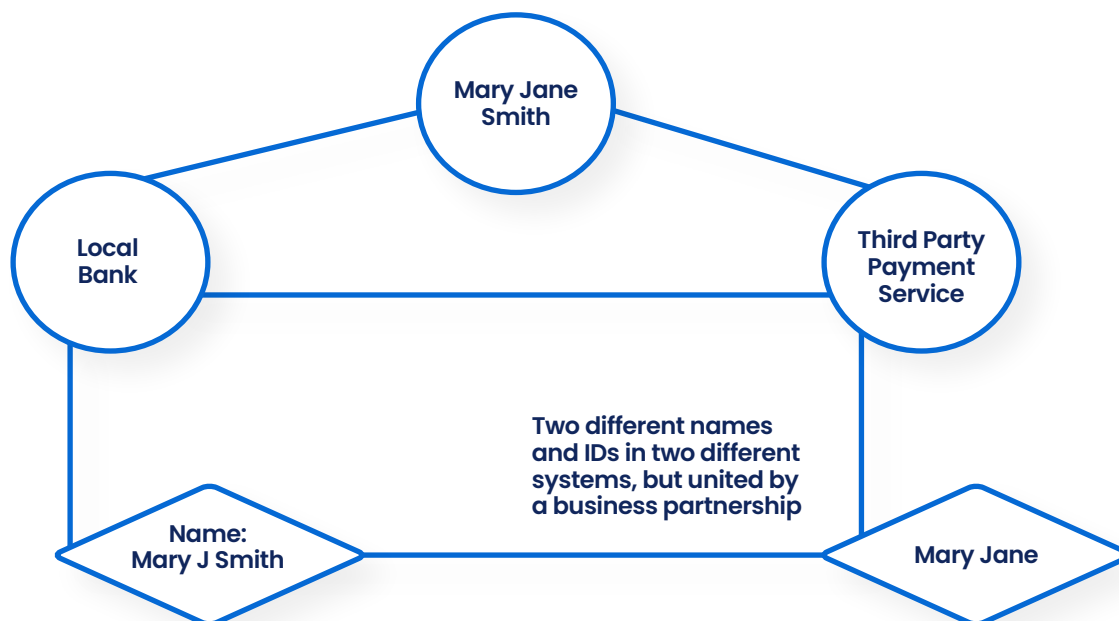
Because unlike olden times, business and services are being provided at scale and are depersonalized. For example, bank services in the early 1900s was community-led. Tellers would know each of their customers on a face-to-face basis, however, as services became more nationalized and offered at scale, names were no longer enough.

A unique identifier would make it easy for service providers to identify customers and ensure they didn't sign up multiple times for a single service.

However, fast forward to the 21st century, each customer has dozens of records streaming in from different sources.

Like in the case of Mary, there are two overlapping records, one at the local bank and one with a third-party payment service provider she just signed up to send money abroad.

An entity resolution exercise would mean the bank has to consolidate data from the third-party payment provider with its internal data sources to get a complete picture of Mary.

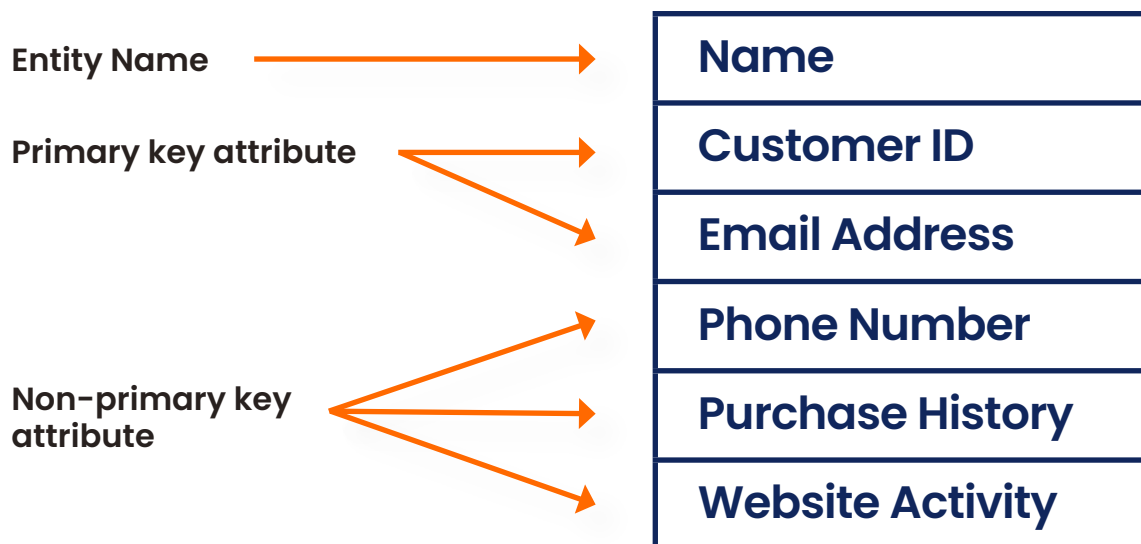


This is a typical situation that arises during platform migrations, and mergers and acquisitions, where overlapping customer information needs to be amalgamated and brought to a single source of truth.

Because there is typically no universally adopted standard (since organizations have their unique ways of collecting and processing data), entity resolution is a complex exercise that many businesses shy away from.

Breaking down the components of entity resolution

Broadly speaking, entity is a noun and refers to any object, thing, person, or place that has physical existence. In technical terms, however, entity refers to an object on which data is collected and stored. So for example, a person is an entity, but in a database, a group of people will be classified as an entity type called, 'customers' sharing a set of attributes.



These attributes are data points about the entity that you've collected. They are divided into primary and non-primary keys, where primary means an identifier that is typically unique to an individual and which can be used as the main identifier. In the case of a customer, their customer ID can act as a primary key.

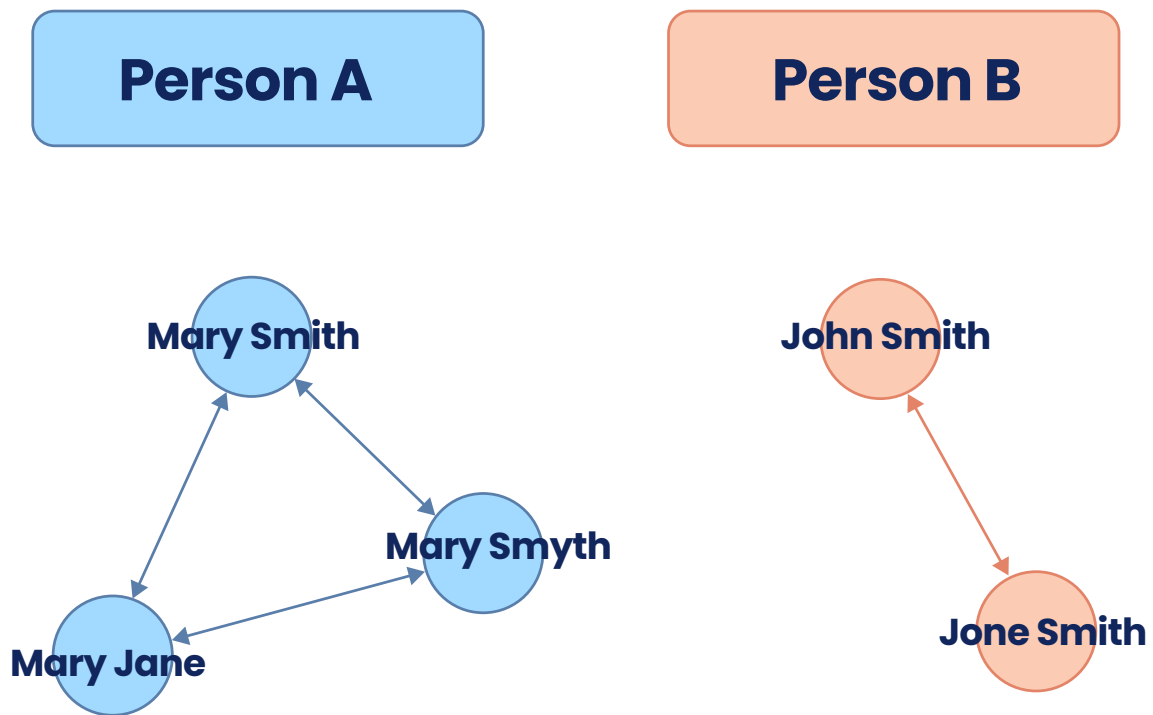
Secondary or non-primary keys are corresponding information about the customer.

Most of the work in entity resolution revolves around fixing, identifying, deduplicating, or combining non-primary attributes to get a bigger picture of customer data.

Given below is a list of common attributes for customers as entities.

Category	Description	Examples
Identifying Attributes	Uniquely define a customer within your system	Name, Email Address Phone Number Customer ID Loyalty Program ID
Descriptive Attributes	Provide additional details about a customer	Purchase History, Product Preferences, Communication Preferences, Account Information, Payment Information
Behavioral Attributes	Capture how a customer interacts with your business	Website Activity (Clickstream data, pages visited), App Usage (Features used, frequency), Customer Service Interactions
Additional Attributes (Optional)	May be relevant depending on your business type	Company Affiliation (B2B customers), Occupation, Marital Status, Social Security Number (hashed for privacy)

Entity resolution is not limited to finding duplicates. A larger goal is also to identify **relationships between entities**. Using the example of Mary Jane Smith:

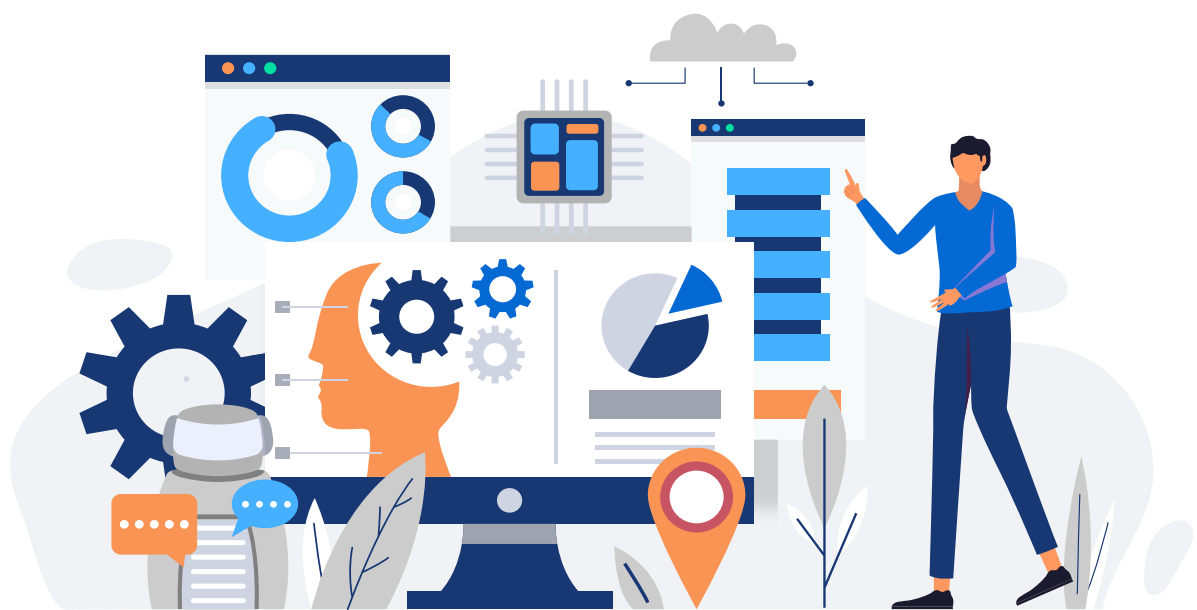


Mary Jane Smith’s husband is John Smith, but in the system he is written as Jone Smith because of a clerical error. Entity resolution looks at the commonality between Mary and John and indicates a potential match, based on which analysts can determine if these two records represent a family.

This finding can help the bank target this family with a more customized service or

offer. For example, they may be added to a yearly discount offer on travel deals for two!

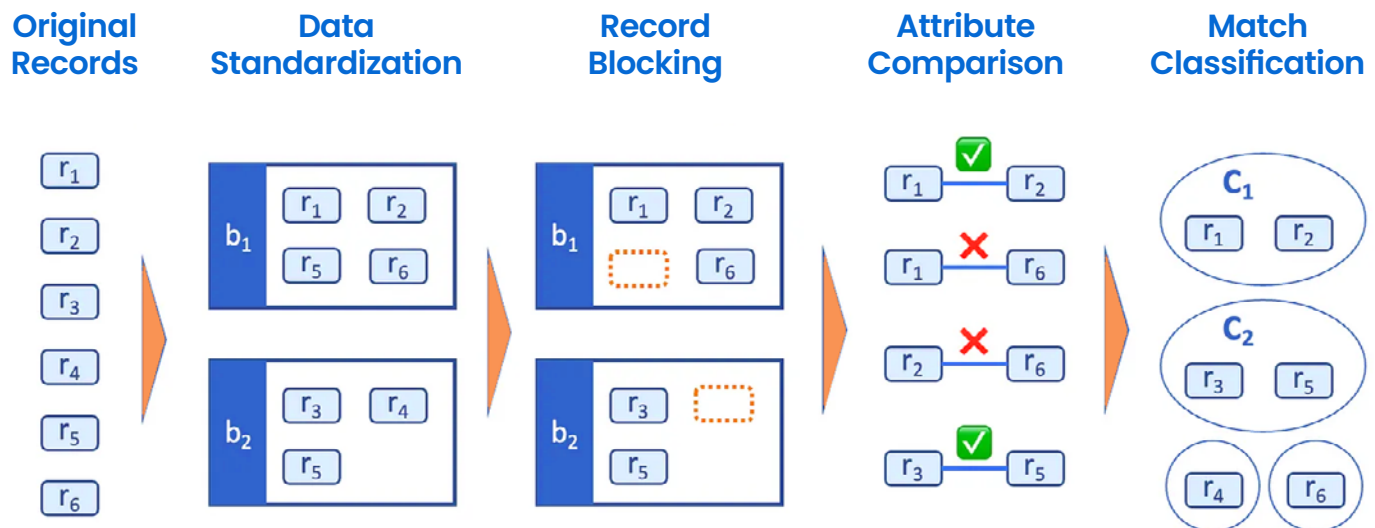
Entity resolution not only consolidates fragmented data but also uncovers relationships between entities, providing a comprehensive understanding that drives better customer service, targeted marketing, and effective risk management.



How is entity resolution traditionally done?

Until a few years ago it was a luxury only businesses with large budgets could afford because entity resolution is a resource intensive exercise. Additionally, different companies have different ways to resolving entities.

For this guide, we'll touch on the basic framework suggested by Michael Shearer, based on which many ER processes have been documented.



The first step of the process is:

● Data standardization

which is also known as data cleansing or data preparation. Before you can compare records to identify duplicates, you need to be sure that you have consistent data structures so you can test for equivalence between attributes. For example, if you don't clean or standardize Mary J. Smith as Mary Jane Smith, you might trigger a false negative (where the system fails to classify

as duplicates because of the difference in a component of a name).

Data standardization by itself requires trained analysts to build complex codes for fixing poor data quality. This is where data analysts are also sometimes called as data janitors – because they spend 80% of their time cleaning/wrangling/preparing data.

The second step of the process is:

2

● **Record blocking**

more easily understood as a filtering process where you select records you want to compare. Instead of comparing high volumes of records, which is practically impossible at a large scale, record blocking is a nifty technique allowing analysts to concentrate the resolution process on the

records with the highest possibility of a match. For example, you would want to match on phone numbers because these have the highest probability of being unique – and therefore easier to use in a match process.

The third step is where all the magic lies! Here you will use data matching techniques like fuzzy match or exact match to compare two attributes.

3

● **Attribute comparison**

here you decide on comparing individual attributes between the pairs of records selected by the blocking process. For example you might compare names by checking for exact matches, common variations, or phonetic similarities.

Similarly, email addresses can be compared by domain and username variations. This detailed attribute comparison helps in determining whether the two records represent the same entity.

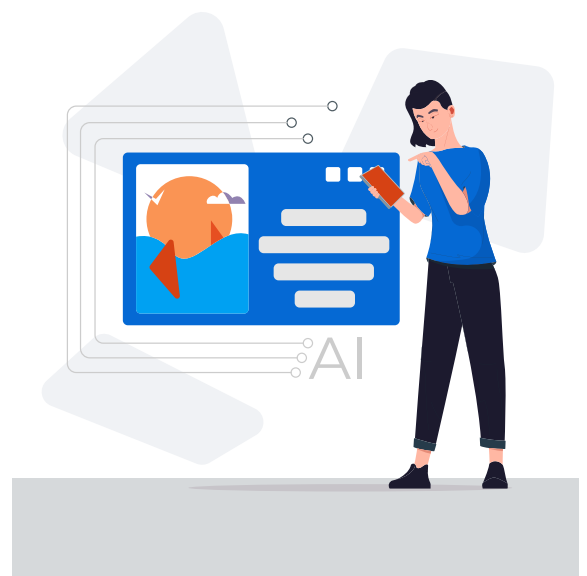
Finally, making a choice:

4

● **Match classification**

after the algorithm or script that you've used to compare attributes results a match, you need to conclude whether the similarity between the attributes is sufficient to declare two records a match, i.e., determining if they refer to the same real-

world entity. This decision can be made on the basis of defined rules or in the case of modern data management, on the basis of an artificial intelligence approach.



Let's see this with an example:

Step 1: Before standardization

First Name	Last Name	Date of Birth	Place of Birth	Mobile Number
mary	J smith	10.06.1975	High St, Standlake, Oxfordshire	07424-868565
M. Jane	Smith	10/07/1976	High St, Standlake, Oxfordshire	+447424696969

Step 1: After standardization

First Name	Last Name	Date of Birth	Place of Birth	Mobile Number
MARY	JANE SMITH	10/06/1975	High Street, Standlake, Oxfordshire, UK	+447424868565
MARYLENE	JANE SMITH	10/06/1976	High Street, Standlake, Oxfordshire, UK	+44742423583

Step 2: Compare attributes on the basis of a fuzzy match

(we touch on fuzzy match later!)

Attribute	Value (record 1)	Value (record 2)	Fuzzy Match?
First Name	MARY	MARYLENE	No Match
Last Name	SMITH	SMITH	Match
Date of Birth	10/06/1975	10/06/1976	No Match
Address	High Street, Standlake, Oxfordshire, UK	High Street, Standlake, Oxfordshire, UK	Match
Mobile Number	+447424868565	Not Available	No Match

First Name: "MARY" and "MARYLENE" are different names. Fuzzy matching typically wouldn't consider them a match unless the threshold for allowable edits is very high – for example at 80%.

Last Name: "SMITH" matches exactly in both records.

Date of Birth: The dates differ by one year. Fuzzy matching might consider this a match depending on the specific algorithm and its tolerance for date variations.

Address: The addresses are identical.

Mobile Number: Only one record has a mobile number, so it's not a fuzzy match.

The final step involves deciding whether the records represent the same entity based on the fuzzy matching results .

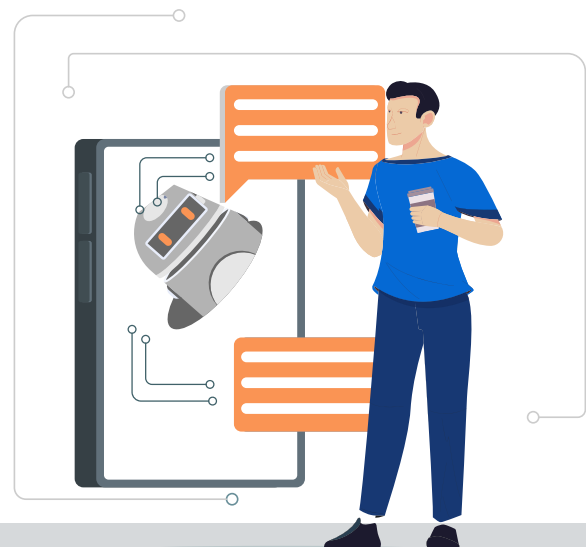
As you can see, the traditional method of 'comparison' between two attributes is time consuming. Additionally, data matching is only one part of the process.

Once the matches are classified, users would have to select final records, cluster them into consolidated records, and present a single source of truth. Such a task is impossible to do at scale and therefore users resort to partially resolving the data only to meet a very specific business demand.

Majority Rule:

A straightforward approach is to consider the records a match if a majority of the attributes exhibit a match (including both exact and fuzzy matches identified in the previous step). In our example with the "MARY" and "MARYLENE" scenario, since most attributes (Last Name, Address) are matches, we might conclude it's the same person under the majority rule.

It is practically impossible to resolve entity data at scale, without the use of some in-house machine learning system, requiring heavy resource involvement. Yet, not all users are capable of configuring and using these ER systems. As a result, the potential user base of such systems is restricted to experts, and even in that case, their capabilities and scope are rather limited.

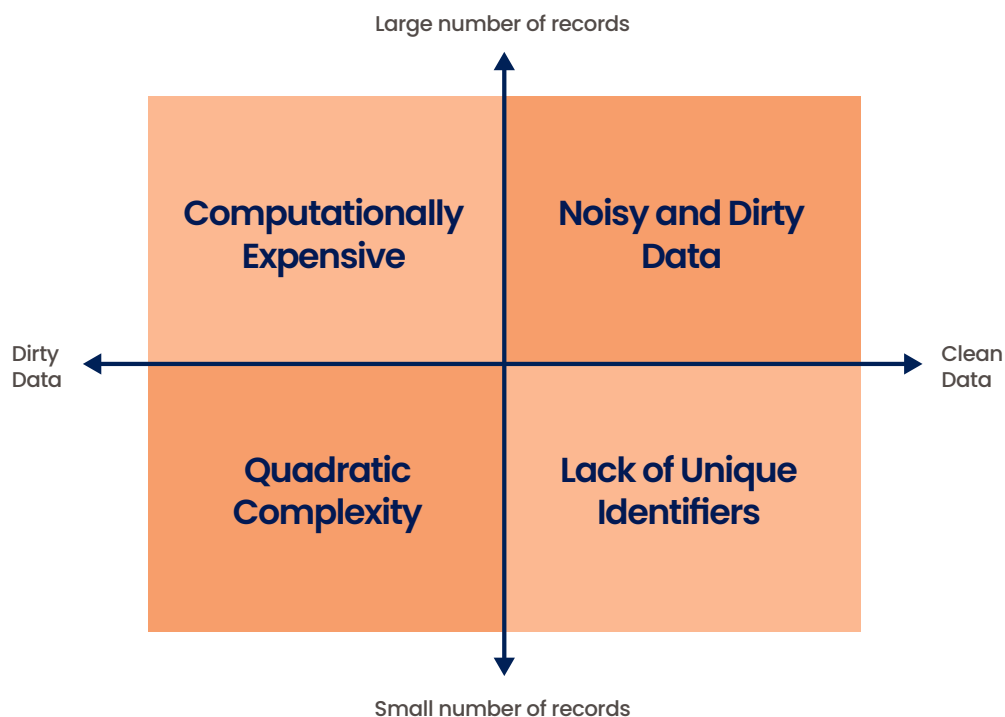


Four major bottlenecks that makes traditional ER a challenge

We've established through multiple examples that entity resolution is a critical need that directly impacts business operations, but its complexity makes it a challenge for businesses to take on with confidence.

Some of the bottlenecks that businesses struggle to resolve include:

Entity Resolution Challenges



● **Computationally expensive:**

Data matching is computationally expensive and application specific. For example, a modern CRM holds thousands of nicknames instead of actual names. A customer can register as **Kathy, Katherine, Catherine, Kath**, to name a few. Businesses would need to build nickname algorithms and trainable engines to identify these

nicknames when matching customer records. Even on the latest hardware, the speed of matching a single record can take up to 0.36ms, **provided there is no other complexity involved such as adjusting match thresholds**. This can lead to hours of runtime for matching thousands of records.

● **Noisy and dirty data:**

Entity resolution struggles with noisy and dirty data. Inconsistent formats, typos, missing fields, and other inaccuracies complicate the matching process. Traditional data processing platforms rely on clear comparisons between

data points, and these errors introduce ambiguity. As a result, the system spends more time calculating similarities and may produce inaccurate matches, hindering the efficiency and effectiveness of entity resolution.

● **Quadratic complexity:**

Entity resolution is hindered by its quadratic complexity. For example, to compare a customer database with 100 records, we'd need to compare each record with every other record to identify duplicates. This translates to $100 \times 99 = 9900$ comparisons!

Now, consider a larger database with 10,000 records. The number of comparisons

explodes to a staggering $10,000 \times 9,999 = 99,990,000$ comparisons! This becomes incredibly time-consuming and resource-intensive for large datasets.

This is where quadratic complexity comes in. As the data size increases, the number of comparisons grows exponentially, making entity resolution impractical for real-world applications with massive datasets.

● **Lack of unique identifiers:**

A unique identity is an ID or a number assigned to each entity. For example, SSNs are unique identifiers, however, these IDs are protected by privacy laws and are not used in data processing. The lack of a unique identifier means users have to resolve entities based on other secondary data – such as phone numbers or email addresses which susceptible to errors and inconsistencies.

The absence of a perfect unique identifier makes entity resolution more challenging, but techniques like fuzzy matching and blocking help us navigate these complexities and link records across datasets.

Existing ER systems attempt to tackle the above challenges in a uni-dimensional way. They extract a block of data, process it based on heavily defined rules (of which no one else other than the original author would know of), and rely on domain knowledge to resolve the data for business use.

They apply the same four-step method only to end up struggling with limitations and erroneous data that demands more attention.

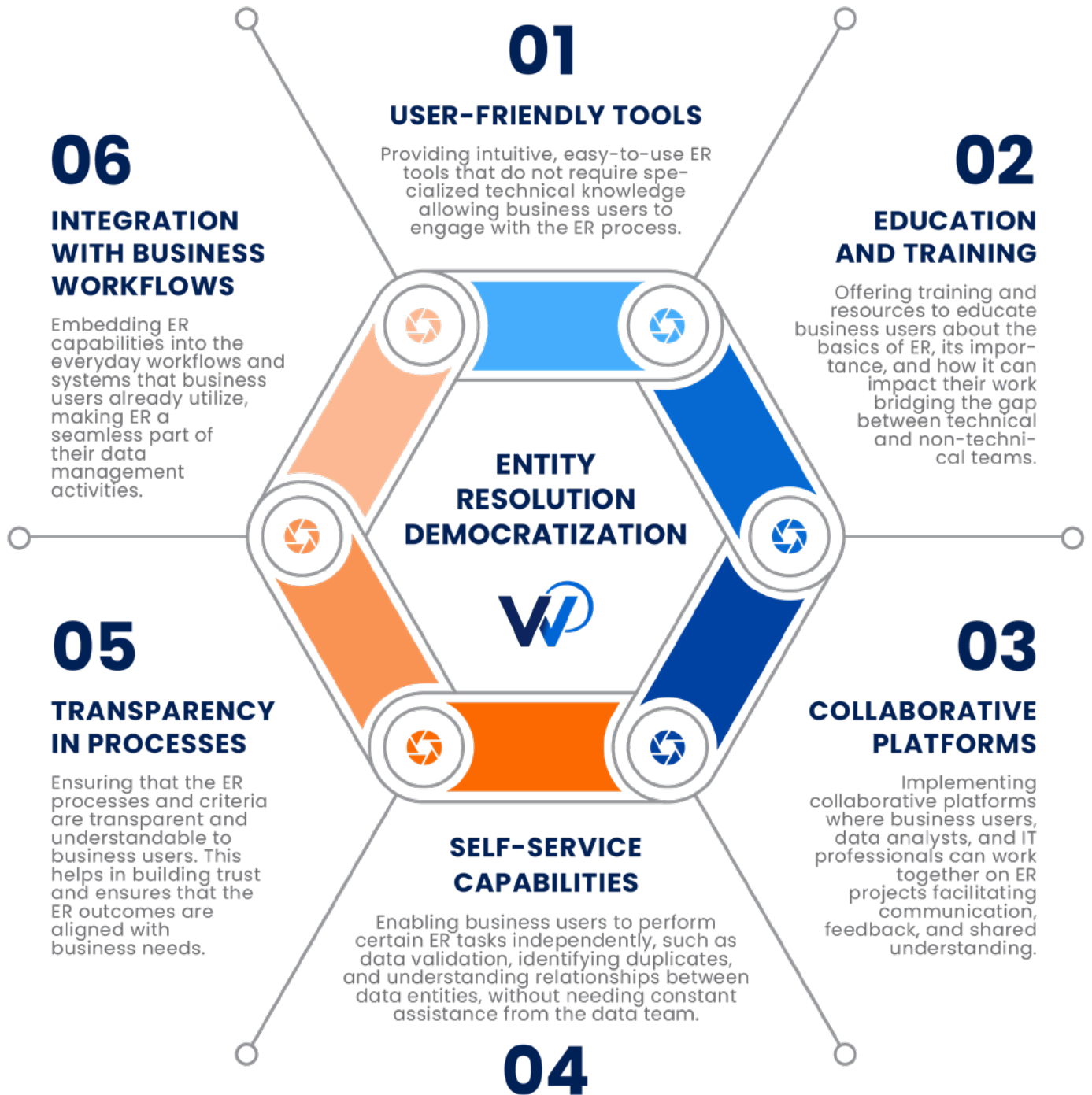
A more efficient solution has never been so direly needed.

WinPure and **Senzing**'s partnership attempts to resolve these challenges and democratize entity resolution. WinPure is the industry's most powerful do-it-yourself data quality management platform equipped with Senzing Inside TM's AI engine. With this end-to-end no-code platform, analysts can resolve complex entities without requiring a single line of code.

What does it mean to democratize ER?

To democratize entity resolution (ER) means to make the tools, processes, and knowledge required for ER accessible to

a wider range of stakeholders within an organization, beyond just the data and IT teams. This involves several key elements:

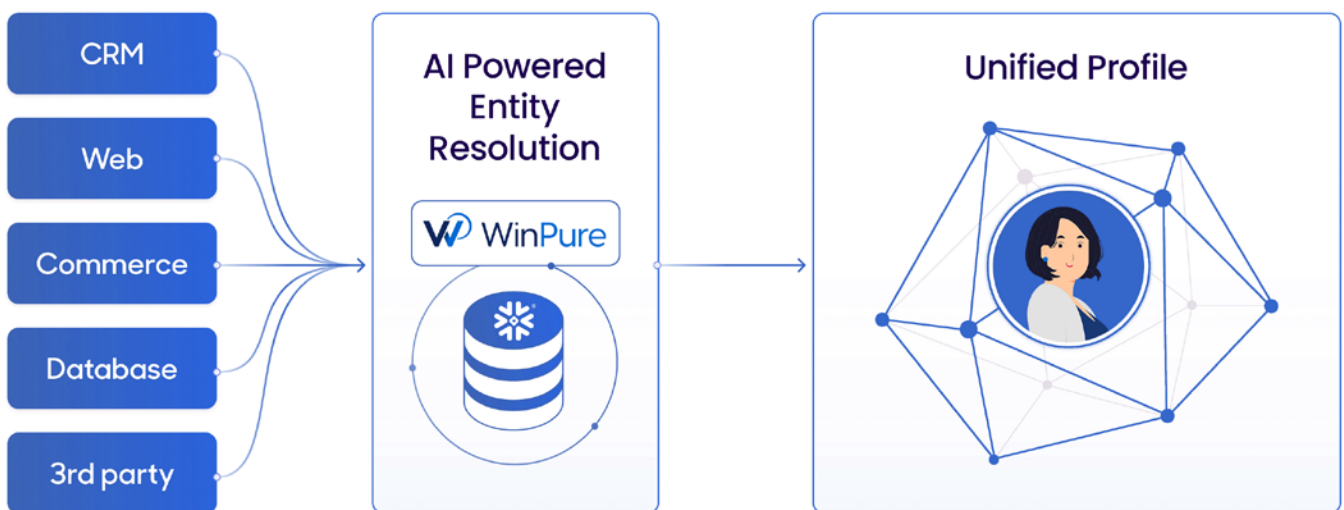


By democratizing ER, organizations empower a broader set of users to contribute to and benefit from the ER process. This leads

to better data quality, more informed decision-making, and ultimately, a more data-driven organizational culture.

Resolving ER challenges with a no-code, AI-powered approach

WinPure industry's first AI-powered entity resolution platform that incorporates the complete entity resolution process in a secure on-premises software. With WinPure, users can integrate data from multiple sources to clean, standardize, and use a combination of fuzzy and exact match algorithms to perform a controlled entity resolution process.



Because it is a no-code solution, businesses do not require a programming expert or a trained specialist. Anyone in the organization with an understanding of the data can use WinPure to achieve entity resolution.



1

Time to Value Ratio

Users can start seeing value within 2-4 weeks after implementation.

Efficiency Improvement

Data processing tasks can become up to 70% faster due to automation.

2



3

Cost Cut Down

Organizations can reduce data management costs by up to 50% through reduced manual effort and improved data quality.

Accuracy Level

WinPure can achieve data matching accuracy levels of over 95%.

4



5

Data Handling Capacity

The tool can process and deduplicate millions of records per hour, significantly improving data management efficiency.

But this is not all.

WinPure's unique proposition lies in its multi-facet purpose.

You could perform entity resolution yourself

OR

You could use our AI-powered entity resolution module to

detect not just duplicates but also hidden relationships.

The screenshot displays the WinPure software interface, which is used for data analysis and entity resolution. The interface is divided into several sections:

- Configuration:** A sidebar on the left with options for Configuration, Result, Data, Clean, Match, EntityAI, Verify, and Automation.
- Step 1 - Select Tables:** A table listing data sources: Sample File 1, Sample File 2, customers, reference, and watchlist. The 'reference' table is selected.
- Step 2 - Define Column Types then click "Review/Analyze":** A table defining column types for the selected data. Columns include EMPLOYER_NAME, PRIMARY_NAME_FULL, NATIVE_NAME_FULL, PRIMARY_NAME_LAST, PRIMARY_NAME_FIRST, DATE_OF_BIRTH, ADDR_TYPE, ADDR_FULL, ADDR_LINE1, ADDR_CITY, ADDR_STATE, ADDR_POSTAL_CODE, ADDR_COUNTRY, and PHONE_TYPE. Checkmarks indicate which columns are included in the result.
- Mapping Columns:** A section on the right with an 'Auto Map' button and a 'Clear' button.
- Review & Analyse:** A section showing 'Licence Limitation: 1 million records' and 'Records to Analyze: 142'. A 'Review / Analyze' button is present.
- System Status:** A section showing 'System status good. Click for details.' with a green checkmark icon.
- Data Table:** A large table at the bottom displaying data from the 'reference' table. The table has columns: DATASOURCE, RECORDID, RECORDTYPE, PRIMARY_NAME_ORG, SECONDARY_NAME_ORG, EMPLOYER_NAME, PRIMARY_NAME_FULL, NATIVE_NAME_FULL, PRIMARY_NAME_LAST, and PRIMARY_NAME_FIRST. The table contains 22 rows of data.

Here's how you can resolve entities using WinPure.

- 1. Import data from multiple sources with a one-click integration and save dozens of working hours in file extractions
- 2. Profile data for hidden errors, and save hours of manual inspection for errors
- 3. Use 30+ built-in data cleansing and standardization options, saving weeks of manual effort
- 4. Match using a combination of fuzzy, exact, and numeric algorithms for all types of datasets
- 5. Verify physical addresses of over 250+ countries and resolve address data challenges

What can AI-powered entity resolution do?

WinPure's entity resolution module is a partnership with Senzing Inside™ the world's first AI engine for record matching.

Senzing's AI engine comes pre-built with common sense that includes principle-based entity resolution and advanced knowledge.

Common sense allows WinPure to be smart on day one, even with data sets as small as two records and ensure its learning is not fooled by newly introduced anomalies e.g., mis-mapped fields or other errors.

Principles are a special form of generalized knowledge that draw on common attribute behaviors. The use of principles is a key reason WinPure software does not need training, tuning or experts to deploy into new domains or to add new data sets, new languages, etc.

WinPure's Entity AI leverages the inherent characteristics of data attributes to achieve accurate entity resolution.

The ER module in WinPure includes more than 10 pre-built comparison routines containing deep knowledge about specific attributes such as phone numbers, SSNs, dates, etc. Since culturally-aware name recognition and global address matching are most critical for achieving high quality ER, the comparators WinPure software uses for these attributes are particularly advanced.

Why involve business users in ER?

Because business users are the custodian of their data. Only a business user can understand the purpose, context, use, and structure of their data. However, in large organizations, business users are often left out of the entity resolution process, completely sidelined by the data team.

Including business users in the ER process ensures that the data is accurately interpreted and matched according to its real-world applications. Moreover, a business user can provide critical insights into the nuances and business rules that govern data usage, helping to define matching criteria that are both relevant and effective. Their involvement reduces the risk of misinterpretation and ensures that the resolved entities truly reflect the business realities, leading to more reliable and actionable data outcomes.



WE RECOMMEND READING THE ADVANCE AI ENTITY RESOLUTION BUYER'S GUIDE TO GET AN IN-DEPTH UNDERSTANDING OF HOW THE MODULE WORKS.

With a no-code data quality framework + an AI-powered entity resolution module, WinPure empowers users to meet ER goals at a pace and efficiency level unlike any other existing solution.

And most importantly – the solution is business-user friendly.

Step 2 - Define Column Types then click "Review/Analyze"

Column name	Type	Label	Include in result only	Ignore column
ADDR_CITY	Addr City	BUSINESS	<input type="checkbox"/>	<input type="checkbox"/>
ADDR_COUNTRY	Addr Country	BUSINESS	<input type="checkbox"/>	<input type="checkbox"/>
ADDR_FULL	Addr Full	BUSINESS	<input type="checkbox"/>	<input type="checkbox"/>
ADDR_LINE1	Addr Line1	BUSINESS	<input type="checkbox"/>	<input type="checkbox"/>
ADDR_POSTAL_CODE	Addr Postal Code	BUSINESS	<input type="checkbox"/>	<input type="checkbox"/>
ADDR_STATE	Addr State	BUSINESS	<input type="checkbox"/>	<input type="checkbox"/>
ADDR_TYPE			<input checked="" type="checkbox"/>	<input type="checkbox"/>
CATEGORY			<input checked="" type="checkbox"/>	<input type="checkbox"/>
DATASOURCE			<input checked="" type="checkbox"/>	<input type="checkbox"/>
DATE			<input checked="" type="checkbox"/>	<input type="checkbox"/>
DATE_OF_BIRTH	Date Of Birth		<input type="checkbox"/>	<input type="checkbox"/>
EMAIL_ADDRESS	Email Address		<input type="checkbox"/>	<input type="checkbox"/>
EMPLOYER_NAME	Employer Name		<input type="checkbox"/>	<input type="checkbox"/>
NATIVE_NAME_FULL			<input checked="" type="checkbox"/>	<input type="checkbox"/>
PHONE_NUMBER	Phone Number		<input type="checkbox"/>	<input type="checkbox"/>

Mapping Columns

AB Auto Map Clear

Review & Analyse

Licence Limitation: 1 million records

Records to Analyze: 142

Review / Analyze

System Status

System status good. Click for details.

Matching Records Options: Get Master Records, Update / Overwrite, Not duplicate, Delete, Merge, Process possibilities, Export

All Matches: Matching time 00:00:00.21, Group count 35, Matches 88, Matches % 61.97%

Viewing Options: Auto Filter Panel Off, Show System Fields

Source name	Is Master record	Is Selected	Match Key	DATASOURCE	RECORDID	RECORDTYPE	PRIMARY_NAME_ORG	SECONDARY_NAME_ORG	NAME_FULL	NATIVE_NAME_FULL	NAME_LAST	NAME_FIRST	PRIMARY_NAME_MIDDLE	GENDER
Group ID: 2														
customers	<input type="checkbox"/>	<input type="checkbox"/>		CUSTOMERS	1001	PERSON								
customers	<input checked="" type="checkbox"/>	<input type="checkbox"/>	+NAME+DOB+ADDRESS	CUSTOMERS	1002	PERSON					Smith	Bob	J	
customers	<input type="checkbox"/>	<input type="checkbox"/>	+NAME+DOB	CUSTOMERS	1003	PERSON					Smith	Bob		
customers	<input type="checkbox"/>	<input type="checkbox"/>	+NAME+EMAIL	CUSTOMERS	1004	PERSON					Smith	B		
customers	<input type="checkbox"/>	<input type="checkbox"/>	+NAME+ADDRESS	CUSTOMERS	1005	PERSON					Smith	Robbie		
Group ID: 1														
reference	<input type="checkbox"/>	<input type="checkbox"/>		REFERENCE	2018	PERSON			Wang Jie					
customers	<input checked="" type="checkbox"/>	<input type="checkbox"/>	+NAME+DOB	CUSTOMERS	1070	PERSON					Wang	Jie		Male
Group ID: 5														
customers	<input type="checkbox"/>	<input type="checkbox"/>		CUSTOMERS	1071	PERSON				王杰				F
reference	<input type="checkbox"/>	<input type="checkbox"/>	+NAME+DOB	REFERENCE	2014	PERSON			Wang Wei					
customers	<input checked="" type="checkbox"/>	<input type="checkbox"/>	+DOB+ADDRESS+NATIONAL_ID	CUSTOMERS	1072	PERSON					Wang	Wei		Female
Group ID: 10														
reference	<input type="checkbox"/>	<input type="checkbox"/>		REFERENCE	2071	ORGANIZATION	Universal Exports	USA						
reference	<input type="checkbox"/>	<input type="checkbox"/>	+NAME+EMPLOYER	REFERENCE	2101	PERSON		USA		Keeley Jones				
reference	<input type="checkbox"/>	<input type="checkbox"/>	+NAME+EMPLOYER	REFERENCE	2111	PERSON		USA		Susan Meyer				
reference	<input checked="" type="checkbox"/>	<input type="checkbox"/>	+NAME+EMPLOYER	REFERENCE	2121	PERSON		USA		Kristen Salinger				
reference	<input type="checkbox"/>	<input type="checkbox"/>	+NAME+PHONE	REFERENCE	2122	PERSON								
reference	<input type="checkbox"/>	<input type="checkbox"/>	+NAME+EMPLOYER	REFERENCE	2131	PERSON								
reference	<input type="checkbox"/>	<input type="checkbox"/>	+NAME+EMPLOYER	REFERENCE	2161	PERSON								
Group ID: 11														
reference	<input checked="" type="checkbox"/>	<input type="checkbox"/>		REFERENCE	2074	ORGANIZATION	Universal Exports							
customers	<input type="checkbox"/>	<input type="checkbox"/>	+NAME+ADDRESS	CUSTOMERS	2073	ORGANIZATION	Worldwide							
Group ID: 17														
customers	<input checked="" type="checkbox"/>	<input type="checkbox"/>		CUSTOMERS	1009	PERSON					Kusha	Edward		
customers	<input type="checkbox"/>	<input type="checkbox"/>	+NAME+DOB+ADDRESS	CUSTOMERS	1010	PERSON					Kusha	Eddie		
Group ID: 23														
customers	<input type="checkbox"/>	<input type="checkbox"/>		CUSTOMERS	1016	PERSON					Kusha	Marie		
customers	<input checked="" type="checkbox"/>	<input type="checkbox"/>	+NAME+DOB+ADDRESS	CUSTOMERS	1015	PERSON					Kusha	Mary		
customers	<input type="checkbox"/>	<input type="checkbox"/>	+NAME+SSN	CUSTOMERS	1017	PERSON					Kusha	Mary		
customers	<input type="checkbox"/>	<input type="checkbox"/>	+NAME+DOB+PHONE	CUSTOMERS	1018	PERSON					Kusha	Marie		

AI generates possible related match keys based on the knowledge it holds about data structures

Let's get you started with entity resolution!

Organizations using WinPure have reported:

- 100% increased efficiency
- 40% reduction in time spent on data prep
- 97% unmatched accuracy with the lowest false positive rates

Our customers are using both our DQM module and ER module to achieve goals like:

1. Improving big data analytics
2. Improving fraud detection capabilities
3. Improving marketing intelligence
4. Meeting compliance regulations
5. Consolidating disparate data
6. Creating golden records and single sources of truth

... and much more.

You can be rest assured that your data will now be able to support your most demanding needs.

And the best part?

You don't need to spend a million dollars on new infrastructure or expensive ER experts, or a large team. WinPure's entity resolution module is affordable, scalable, and most importantly – reliable.

For more information, book a demo with our solution specialist

[Book a Demo](#)

