Deduplicate and consolidate your data from multiple data sources using match and merge in Informatica MDM SaaS

Data Sheet | LumenData -

Compare your data to identify similar records and combine them into a single, reliable representation of the entity.

PRODUCTS EXPLAINED:

1.Informatica MDM SaaS (IICS account)2.Business 360 Console and3.Customer 360

(The user must have the required privileges to access the mentioned services in IICS)



STEP BY STEP GUIDE TO IDENTIFY AND MERGE DUPLICATE OR SIMILAR RECORDS WITHIN THE MASTER DATA

1.CREATING RECORDS

After logging into our IICS account, we navigate to Customer 360.



Click on '**new**' and select the record type you want to create for the selected business entity.

	1EN <mark>DA</mark>	TA Cu	stomer 360 🗸	Person	Search	Q	35 Partner - Lumen Data 🗸 📔 🎅 🙎 📍
+ New	🐴 H Ne	ew					⊗
Home Search File Import Workflow Inbox Hierarchies My Jobs Reports	IN R	Record dierarchy Report	Select the record type you war Person Person Details Person_AMPS Person_AMR Person_Ashok Person_Ashok	nt to create		Find	
×.					C	eate Cancel	

A new window will pop up after you click on create as shown below picture including all the data fields. Fill in all the necessary information and submit the changes which you have made.



N LUM	IENDATA Customer 360	✓ Person	• Search Q	35 Partner - Lumen Data ∨ 🏳 🖉 ?
+ New	Sew Person			
Mome				Ť
O Search	 General Information 			
File Import	First Name*	Middle Name:	Last Name*	Eull Name: 00
📢 Workflow Inbox]		alar Ree
Hierarchies	Prefix Name:	Suffix Name:	Title:	Designation:
My Jobs	•	•		
Peports	Birthplace:	Birth Date:	Gender:	Image:
💄 New Person 🛛 📀		MM/dd/yyyy 🗎		•
	Marital Status:			
	•			
	 Address (0) 			+
«	Phone (0)			+

After submitting and approving the changes the following window will show up with the new record with all the data fields entered by the user.

■ LUM	ENDATA Customer 360 ¥			Person	Search Q	35 Partner - Lumen Data 🗸 🕴 🏱 🔒 🤗
New Mome	Vaibhav Singh Negi 28/13 Aspire Heights, SSJ Layout, Greater 7 9987615432	Noida, Uttar Pradesh, India				• 8
O Search	Record Details Source Records Hierarchy	History Relationship Graph				
- File Import						
• 🐔 Workflow Inbox	 General Information 					0
Hierarchies	First Name:*	Middle Name:	Last Name:*		Full Name:	
In My Jobs	Vaibhav	Singh	Negi		Valbhav Singh Negi	
Preports	Pretix Name: Mr.	Suffix Name:	Engineer		Designation: Developer	
💄 Valbhav Singh 💿	Birthplace:	Birth Date:	Gender:		Image:	
	Dehradun	11/20/1998	Male			
	Marital Status: Single					
	Address (1)					+
	Home 🛛 (Primary)					
	Home					
	28/13 Aspire Heights SSJ Layout Kunj Vihar Greater Noida Uttar Pradesh India 445610					
	Phone (1)					+
	Cell (Primary)					
	9987615432					
<	Prome					



NOTE:

Data can also be entered without manual practices using an ingress job that loads the data from source file to the desired business entity using the required connections. This process can be divided into four steps.

• **Mapping**: Define the source and target connections using the appropriate transformations in data integration.

• Mapping Task: To process the data based on the data flow logic defined in the mapping in data integration.

• Task Flow: Control the execution of the mapping task in data integration.

• Ingress Job: Define a job in the Business 360 Console to run the task flow and load the data into MDM.

2. DEFINING MATCH AND MERGE RULES FOR THE BUSINESS ENTITY IN USE:

In Business 360 Console, go to explore and search for the business entity for which you're creating the record.

	🔶 Informatica Bu	siness 360 Cons	ole 🗸				35 P	artner - Lumen Data 🗸 🛛 🏳	요 ?
+	New	📑 Person	✓ Valid					Save 🗹 🗄	8
	Home	Model	Data Flow		Consumption				
-	Explore	Attributes	Data Quality	Match Survivorship	Events				
•	My Import/Exp					â %		Show: Enabled Fields	•
	Business Events	Basic F	Fields 🛛	T First Name *	T Middle Name	T Last Name *	T Full Name	= Prefix Name	
₿	Translations		-					:=	
	My Jobs	1	Text	Suffix Name	T Tile	Designation	T Birthplace	31 Birth Date	
•	Security	1	23						
£\$3	Global Settings		100	Gender	T Image	Marital Status	Address	C Phone	
		De	puble				θ	0	
	Person 😵	1	.23	Email	Social Media	Employment	Education	Specialization	
		De	cimal	0	0	,	0	•	
			/=						1
		T Prop	erties: First Nar	ne					

For this case, we are using the business entity as shown above. After creating all the records in Customer 360, define the match rules which we need to work on to match the records. Navigate to the match tab and select the name of the match model which is to be used for the match and merge process.



In case there is no suitable match model available to perform the match and merge operation, we will define the model as per our requirement as shown below.

1. Click on the + (plus) symbol to add to a new model.

😑 🔶 Informatica Bu	usiness 360 Console 🗸	
+ New	Person Volid	Save 🗹 🗄 🔇
🏠 Home	Add Match Model 🛞	
Explore	Attributes Dat Configure General Properties Add Candidate Selection Criteria	
My Import/Exp	Match Mode	J↑• 🛛 Find
Business Events	Name Model Name:* Baseline Match Model for person 1	nc
Translations	Baseline : Model Description: Match model to find the duplicate data in Person Dusiness entity	e set for person BE.
My Jobs	Model Objective:" Resolve Duplicates *	
🔶 Security	Default Population 📀 : Usa 🔻	
Global Settings	About Population: Created by: (c)2000-2021 Informatica LLC. All rights reserved, Last change: 28 Jul 2021 USA Standard - v10.2.0HF1-28Jul/21 This population uses a character set that surport (SD-LATINL). The population uses a	
📑 Person 🛛 😒	typical identity data found in the United States.	
	() Next > Cancel	

The window shown above will appear where we need to provide the details of the match model. Fill the field as per requirement.

2. Add the candidate selection criteria required for our matching conditions. You can provide multiple criteria for a single match model.

😑 🔶 Informatica [:] Bu	usiness 360 Console 🗸	
+ New	📑 Person 🔽 Valid	Save 🔽 🗄 😵
🏠 Home	Add Match Model 🛞	
늘 Explore	Attributes Dat Configure General Properties Add Candidate Selection Criteria	
My Import/Exp	Match Model	
Business Events	Criteria 1	nc
Translations	Baseline t person Dr Field Name:	e set for person BE.
My Jobs	Field Type:"	
🖤 Security	Filter Candidates By 🤡 :	
Global Settings	Key Generation Level: Standard 💌	
	Candidate Search Level: Typical 🔹	
Person 😵	Immediately open the model	
	() < Back Add Cancel	



3. Define declarative rules to configure the match properties. Select the match strategy intended to use according to the requirement.

😑 🔶 Informatica Bu	isiness 360 Conso	ie V		35 Partner - Lumen Data 🗸 🛛 🏳 🧕 ?
+ New	Person	> Baseline Match Model for person Deduplication 1		🔹 Draft 🛛 Publish 🛛 😣
🏠 Home	Model Config	Add Declarative Rule	\otimes	
Explore Explore	Declarativ	Configure Properties and Thresholds		Save Cancel
My Import/Exp				
Business Events	Rank	Properties		
Translations		Match Strategy:*	•	
My Jobs		EXACT		
Security		FUZZY		
Global Settings				
Parron				
- reison 🖉			Next > Add Cancel	

We have two match strategies named Fuzzy match and Exact match.

Exact Match: A match strategy that matches only records that are identical. You can perform exact matching on any data type.

Fuzzy Match: A match strategy that uses probabilistic matching, which considers spelling variations, possible misspellings, and other differences that can make matching records non-identical. You can perform fuzzy matching on any data type.

After configuring the match strategy, select the merge strategy we need to implement.

😑 🔶 Informatica Bu	isiness 360 Consol	e V			
+ New	Person	> Baseline Match	Model for person Deduplication 1		Draft Publish
🏠 Home	Model Config	Add Declarative R	λule	8	
ig Explore	Declarativ	Configure Propert	ties and Thresholds 2 Add Required Fields 3 Coni	figure Exact Match Field Properties	Save Cancel
My Import/Exp			(), ()		+
Business Events	Rank	Properties			
Translations		Match Strategy:*	EXACT	•	
My Jobs		Merge Strategy:*		•	
🜒 Security		Description:	Automated		
Global Settings			Skip Manual		
Parron (0)					
reison 🖉		0		Next > Cancel	



The three merge strategies available are as follows:

- Automated: Merge the record automatically after successful matching.
- Skip: Skip the merging of record.
- Manual: Manually merge the records after applying the match strategy.

4. Add the required fields for comparison: **Note: At least one exact field match field is required.** Define all the field names for the exact match fields and select the comparison outcome as identical for exact match as shown below.

😑 🔶 Informatica Bu	usiness 360 Console	\sim			35 Partner - Lumen Data ∨ P ?
🛨 New	📑 Person	> Baseline Match Model for pers	on Deduplication		Published
🏠 Home	Model Config	Declarative Rule		\otimes	
늘 Explore	Declarativ	Configure Properties and Thresholds	Add Required Fields Add Optional Fields	-	Edir
👔 My Import/Exp	Rank	compore repenses and meshods			
Business Events	1	Exact Match Fields (3) ()		I	and exactly matching
Translations		Field Name	Comparison Outcome	— I	
My Jobs	2	PostalAddress > country	Identical		address and exactly
🔰 Security		PostalAddress > state	Identical		email
Global Settings		birthDate	Identical		
·()·					s phone number
Person 🛞		Fuzzy Match - Required Fields (2)			
		1		Close	

Configure the properties and thresholds to be used in case of a fuzzy match. Define the match strategy, merge strategy, match criteria, match level, and description if required. Define the merge threshold for the selected merge strategy, the merge strategy value can be anything from 0 to 100. However, keep the merge threshold between 90 to 100 for better performance.

😑 🔶 Informatica Bu	isiness 360 Console 🗸	35 Partner - Lumen Data ∨ 🖓 🖓 🤶
+ New	Person > Baseline Match Model for person Deduplication	Published
🏠 Home	Model Config Declarative Rule	
Explore	Declarative Configure Properties and Thresholds Add Required Fields Add Optional Fields	Edit
My Import/Exp	Rank Properties	
Business Events	Match Rule ID: MR01	ind exactly matching
My Jobs	2 Match Strategy: FUZZY	address and exactly
V Security	Merge Strategy: Automated	email
🛞 Global Settings	Merge Threshold	phone number
Person 🛞	Automated Merge: 90 - 100	
	Close	



Add the required fields for exact and fuzzy matches as shown below,

😑 🔶 Informatica Bu	usiness 360 Consol	le 🗸		3	15 Partner - Lumen Data 🗸 🛛 🏳 🧕 🔶
+ New	Person	> Baseline Match Model for perso	on Deduplication		• Published
🏠 Home	Model Config	Declarative Rule		\otimes	
Explore Explore	Declarativ	Configure Properties and Thresholds	Add Required Fields Add Optional Fields		
My Import/Exp	Rank	5			
Business Events	1	Field Name	Comparison Outcome	ind	exactly matching
Translations		PostalAddress > country	Identical		
🚺 My Jobs	2	Partal Address > state	Identical		aress and exactly
🜗 Security		FosidiAddress > sidre	idenncai		mail
Global Settings		birthDate	Identical	a pi	hone number
Person 🚳		Fuzzy Match - Required Fields (2)			
		0		Close	

= 🔶 Informatica Bu	siness 360 Console	\checkmark					35 Partner - Lumen Data 🗸	P Q ?
+ New	Person	> Baseline Match	Model for person Ded	uplication				• Published 🔀
🏠 Home	Model Config	Declarative Rule				\otimes		
늘 Explore	Declarativ	Configure Propertie	s and Thresholds Ad	Required Fields	Add Ontional Fields	_		
My Import/Exp	Rank	comgete ropene						
Business Events	1	E	in d Fields (2)				and exactly matching	
Translations		Field Type	Field Name		Comparison Outcome		address and exactly	
My Jobs	2	Person_Name	fullName		Identical		address and exactly	
🖤 Security		Address_Part1	PostalAddress > addre:	sLine 1	Identical		email	
Global Settings							phone number	
Person 🛞								
		0				Close		

You can add optional field for fuzzy match as per requirements. After configuring all the match and merge strategies, click on save and publish your business entity to use the match model.



3. CREATE A JOB TO PERFORM THE MATCH AND MERGE TO OPERATION:

Navigate to Business 360 Console, click on new, select match and merge from the jobs section as shown below,



Provide the display name, location and description of the match and merge job and click on OK as shown below,

📑 Nev	w Match and Merge	e Job Definition				Save Run	8
Proce	Match and Merg	e : matchandmerge_Deduplication_	Person		\otimes		
Desc Busin	Display Name:* Internal ID:* @ Description: Location:*	matchandmerge_Deduplication_Person matchandmerge_deduplication_person Default	Browse				
	?			ОК	Cancel		



Enter all the required input for the fields as shown in the below picture which specify the business entity upon which we are applying the match and merge process.

			Save Run
cess:*	MDMMatchMerge		
cess Type: •	Match and Merge	•	
cription:	Identifies matching records and automat	ically merges them.	
iness Entity:	Person	•	
tch Model Name:*	Baseline Match Model for person Dec	duplication 💌	
tch Model Version:*	1	•	

Make sure the business entity, match model name, match model version are selected correctly as per the previous configuration. Save the job and run it. Navigate to my jobs in business 360 console to check the status of the running job. Once, the job has been completed successfully download the output file from the output tab from the match process.

Stort		End
		Merge index for Se
lob Instance Loc	k The Job Match Merge In	ıdex for Search Release The Job
lob Instance Loc	k The Job <u>Match</u> Merge Ir	ndex for Search Release The Job
lob Instance Loc Overview	k The Job <u>Match</u> Merge In Record Pairs	ndex for Search Release The Job Count
ob Instance Loc Dverview Runtime Param	k The Job Match Merge In Record Pairs Total Record Pairs Generated	Count
lob Instance Loc Overview Runtime Param Vietrics	k The Job <u>Match</u> Merge Ir Record Pairs Total Record Pairs Generated	Count 20 Download



The output file would be a CSV file which will give us all the details about the operation and the records undergoing that operation as below and the match records thus obtained,

Record ID Record So Matched FMatched FM	Match Sco Match Ou Matched [Match Mo Mat	tch Mo Match Mo	Aatch Rul Populatio	Descriptio Match CritRecord fir	Matched	FRecord ful Matched FR
6450c0335c360.defai6450c12d.c360.defai	100 AUTO_ME Declarativ 65984184- Bas	seline N 1	1 usa	ABHINAV	ABHINAV	ABHINAV ABHINAV K
6450c12d;c360.defai6450c0335c360.defai	100 AUTO_ME Declarativ 65984184-Bas	seline N 1	1 usa	ABHINAV	ABHINAV	ABHINAV IABHINAV IK
6450c3377c360.defai6450c7185c360.defai	100 AUTO_ME Declarativ 65984184-Bas	seline N 1	2 usa	Person_N(VIGNESH	VIGNESH	VIGNESH FVIGNESH FK
6450c3377c360.defai6450c7be(c360.defai	100 AUTO_ME Declarativ 65984184-Bas	seline N 1	2 usa	Person_N; VIGNESH	VIGNESH	VIGNESH FVIGNESH FK
6450c7185c360.defai6450c3377c360.defai	100 AUTO_ME Declarativ 65984184-Bas	seline N 1	2 usa	Person_N; VIGNESH	VIGNESH	VIGNESH FVIGNESH FK
6450c7185c360.defai6450c7be6c360.defai	100 AUTO_ME Declarativ 65984184-Bas	seline N 1	2 usa	Person_N; VIGNESH	VIGNESH	VIGNESH FVIGNESH FK
6450c7betc360.defai6450c3377c360.defai	100 AUTO_ME Declarativ 65984184-Bas	seline N 1	2 usa	Person_N: VIGNESH	VIGNESH	VIGNESH FVIGNESH FK
6450c7betc360.defai6450c7185c360.defai	100 AUTO_ME Declarativ 65984184-Bas	seline N 1	2 usa	Person_N; VIGNESH	VIGNESH	VIGNESH FVIGNESH FK
6450c97eEc360.defai6450ca9e7c360.defai	100 AUTO_ME Declarativ 65984184-Bas	seline N 1	2 usa	Person_N; NIHARIKA	NIHARIKA	NIHARIKA NIHARIKA S
6450ca9e7c360.defai6450c97e€c360.defai	100 AUTO_ME Declarativ 65984184-Bas	seline N 1	2 usa	Person_N: NIHARIKA	NIHARIKA	NIHARIKA NIHARIKA S

Analysis: After running the match and merge job we navigate to customer 360 to check for the merged record in the MDM. Select the previously used business entity and search for the required record. Once the record is found, check for the source records.

In this case, two different records with different IDs as shown above have been merged into one single surviving record with the details of both the merged records.

Vaibh 28/13 A 9987611	av Singh Negi spire Heights, SSJ Layo 5432	out, Greater No	ida, Uttar P	radesh, India				•
Record Details	Source Records	Hierarchy	History	Relationship Graph				
 General Info 	rmation							0
First Name:*				Middle Name:	La:	st Name:*	Full Name:	
Prefix Name:				Suffix Name:	Titl	9° le: Ninaar	Designation:	
Birthplace:				Birth Date:	Ge	nder:	Image:	
Marital Status:						-		
 Address (2) 								+
Home 🗢	Primary				Home 🗢 (Pimary)			
Home 28/13 Aspin Greater Noid	e Heights SSJ Layout 5a Uttar Pradesh Ind	Kunj Vihar ŝia 460108			Home 28/13 Aspire Heights SSJ Layout Kunj Vihar Greater Noida Uttar Pradesh India 460108	1		
Phone (2)								+
Cell (Prime	7				Cell (Primary)			
9987615432 Home	2				9987615432 Home			

Upon searching for the above record in Customer 360, we only found the surviving record in the MDM with the two merged records under the source record section as previously discussed. These records were merged as the ID and Full Name were same for both records and this was our declarative rule defined in the match strategy with the highest priority. The details of the match and merge operation can be found in the output file.



Vaibhav Sing	h Negi Vaibhav Si	• Vaibhav Singh Negi Vaibhav Sing Source System: Informatica Customer 360	 Vaibhav Singh Negi Vaibhav Sing Source System: Informatica Customer 360
 System Fields 	S		
Business ID	: MDM00000B		
ID	: 64660731d3be	ex62873c19f03d912957697764	ex62873c19f03d912957697761
Updated On	: 05/18/2023	05/18/2023	05/18/2023
Updated By	: Adarsh.ld	Adarsh.ld	Adarsh.ld
Record State	: ACTIVE	ACTIVE	ACTIVE
Population	:		
Match Pair De	tails 🕞		
 General Information 	mation		
First Name	: Vaibhav	Vaibhav	Vaibhav
Middle Name	: Singh	Singh	Singh
Last Name	: Negi	Negi	Negi
Full Name	: Vaibhav Singh	Valbhav Singh Negi	Vaibhav Singh Negi
Prefix Name	: Mr.	Mr.	Mr.

ABOUT LUMENDATA:

LumenData is a leading provider of Enterprise Data Management, Cloud & Analytics solutions provider. We enable enterprises to modernize their legacy data infrastructure and derive greater insights into their business. Founded in 2008, with locations in multiple countries, LumenData is privileged to serve over 100 of the world's leading companies, including KwikTrip, Versant Health, US Food & Drug Administration, US Department of Labor, Cummins Engine, BCG, and others. LumenData is SOC2 certified and has instituted extensive controls to protect client data including adherence to GDPR and CCPA regulations.

To discuss how we can facilitate data-driven transformation for your organization.



Adarsh Sinha Informatica/Data ETL Engineer

MEET OUR AUTHORS



Kousalya S Associate Consultant



Shadwal Srivastava Technical Lead



Contact us +1 (855) 695-8636 info@lumendata.com

lumendata.com