

Copyright Notice

"The contents of this document are proprietary and confidential to Infosys Ltd. and may not be disclosed in whole or in part at any time, to any third party without the prior written consent of Infosys Ltd."

© 2020 Infosys Limited, Bangalore, India. All rights reserved. Infosys believes the information in this presentation is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of other companies to the trademarks, product names and such other intellectual property rights mentioned in this document. Except as expressly permitted, neither this presentation nor any part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, printing, photocopying, recording or otherwise, without the prior permission of Infosys Limited and/or any named intellectual property rights holders under this presentation.

EXTERNAL

Applies to:	Salesforce practitioners
Objective:	This document provides instructions for installing the Infosys Data Masking application
Policy or Handbook Reference	N/A
Prerequisites	N/A

Introduction

This document provides step-by-step instructions for installing the Infosys Data Masking application.

Purpose of application

The main purpose of the Data Masking app is to protect the actual data by using a functional substitute for occasions when the real data is not required.

Scope

Salesforce CRM platform

Business Case

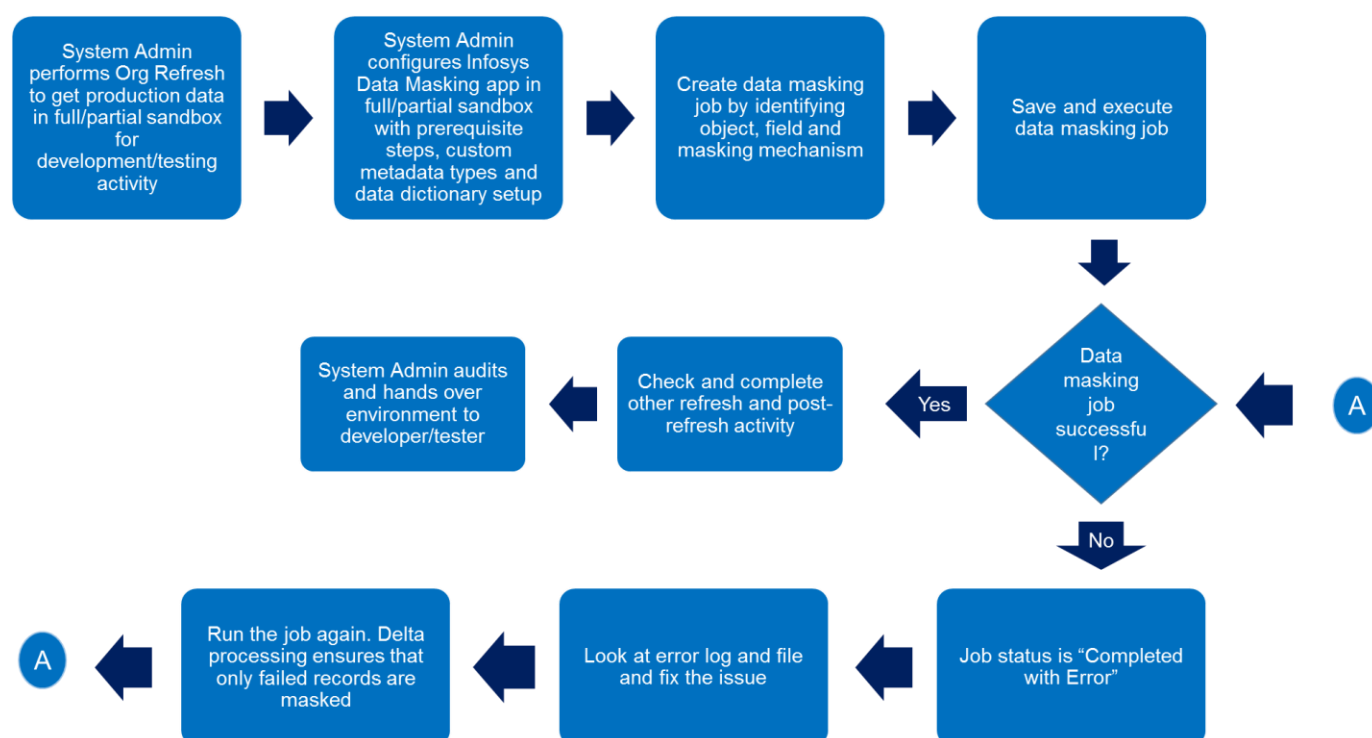
Data quality is of utmost importance while testing applications. It takes significant effort and time on the part of the testing team to create testing data in a sandbox. A possible alternative is to refresh the sandbox with production data, but such a move is often accompanied by data security risks as sensitive customer data could fall in the hands of unintended recipients. Hence, enormous resources and time are spent on masking data in the test environment (partial/full sandbox) from the production environment to ensure protection of sensitive personally identifiable information or PII. There are risks involved in the manual masking process too as data could be accessible during masking and uploading, thus leading to data security issues.

There was need for a Salesforce application to eliminate manual data masking and ensure that test data is masked/encrypted during data refresh from production to sandbox/testing environment. This would ensure that PII and other sensitive information are unavailable to the testing team, with no impact on their functioning.

Solution

The solution was to develop a Salesforce application to mask sensitive data in the lower testing environment to reduce manual effort and data security concerns. The Infosys Data Masking App ensures that sensitive data is being masked with minimal setup in the testing environment, resulting in significant saving of time and effort.

Process Flow



Features

1. An easy-to-use intuitive UI for masking of data
2. Multiple masking techniques to mask data based on testing requirements:
 - a. **Substitution mechanism:** This technique masks data by replacing or substituting the existing content of a column with similar data. For instance, names can be sanitized by substituting real customer names with names from a customized list of names. The substitute data is maintained in a data dictionary
 - b. **Shuffling mechanism:** This technique masks data by shuffling the existing content of a column. Instead of generating substitution data from separate files, the data is generated from the column that needs to be masked. The data in the column is shuffled between rows to make it entirely unrelated to the real data
 - c. **Encryption mechanism:** This technique masks data by encrypting the existing content of a column using an encryption algorithm along with a key. For instance, names can be encrypted by using AES256 algorithm into a binary element
 - d. **Masked-out mechanism:** Also known as the dynamic data masking method, this technique involves character scrambling or masking out of certain fields. It is another simple yet highly effective method to prevent leakage of sensitive information
 - e. **Randomization mechanism:** Data is masked using random method as below:
 - i. Alphabetic randomization
 - ii. Numeric randomization
 - iii. Alphanumeric randomization
 - iv. Random date
3. Robust error handling functionality:
 - a. Error log to capture job failure detail
 - b. Error file to capture all failure records
4. Delta processing of failed job to ensure that only failed records are processed:
 - a. In case of data masking job failure, the app allows for delta processing so that only failed records are processed
 - b. Improves job efficiency and data masking process
5. Flexibility to exclude object and fields that are not required for data masking. Given that only a certain set of object and fields may be required for data masking activity, the app provides functionality to exclude objects and fields using custom metadata type "DataMasking Exclude Object" and "DataMasking Exclude Field".
6. Export the data masking job to another testing environment. This functionality will ensure that the administrator would not need to create the jobs from scratch and is able to reuse existing data masking jobs from another environment.

Use Cases

- Testing team needs to test an application in the sandbox and needs to create test data similar to production data for optimal testing while ensuring no defects in production due to data once application moves to production environment. However, creating test data similar to that in the production environment is challenging and requires substantial time and effort from the testing team
 - Infosys Data Masking app helps the testing team use masked data from production for testing purpose without worrying about data security or affecting any existing application
 - It enables significant time and effort savings on test data creation and ensures that test quality is as per production environment
- Production data has been refreshed in sandbox, but sensitive data needs to be masked to prevent unauthorized access
 - Infosys Data Masking app helps mask data in the sandbox environment as per business requirements, thus preventing unauthorized access to sensitive production data

Prerequisites

- Infosys Data Masking app should only be used in partial or full sandbox after it is refreshed with production data and post refresh activity completion during which System Admin might have access to production data with prior approval
- Processes must be defined to identify PII and other sensitive data that need to be masked
- Separate jobs should be created for the identified object with PII and sensitive data while ensuring that all the required data are masked before allowing data to be loaded into the sandbox for testing purpose. Certain processes need to be defined and adhered to so that sensitive data is not viewed by unauthorized personnel:
 - Ensure only designated users and profiles that are authorized have access to the Data Masking App. Assigning required profile and permission sets will ensure that only designated users have access to Data Masking tool and authorized data
 - Disable the field history tracking for all objects that need to be masked. This will ensure that masked data is not revealed due to field tracking and will remove unnecessary filling up of audit details.
 - Disable triggers, approval flows, and validation rules on object to be masked to ensure smooth data masking process

User Manual

Step	Instruction	Illustration	Expected Result
1	Install the package from AppExchange listing	Please refer to Figures 1 and 2	Installation is successfully completed

Figure 1

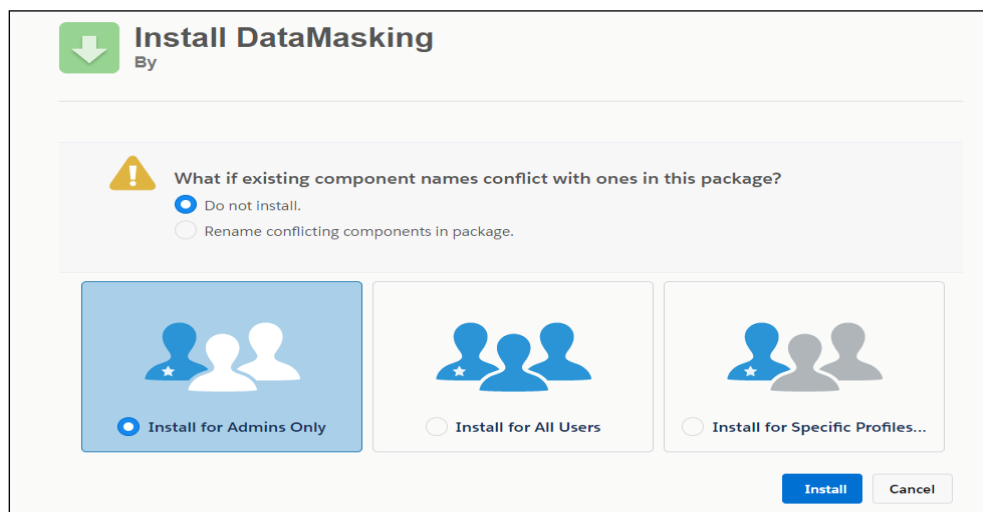


Figure 2



2	Prerequisites Configure pre-requisite steps as below: • Provide access at App, Tab and Field level for respective user profile • Disable Field History Tracking for objects to be masked • Disable triggers, validation rules, approval flow on object to be masked		
---	---	--	--

Step	Instruction	Illustration	Expected Result
3	Custom Metadatatypes - Go to Custom Metadatatypes “DataMasking Exclude Object” to exclude objects from the data masking - Go to Custom Metadatatypes “DataMasking Exclude Field” to exclude fields from the data masking - Click on “Manage” Action and add object be excluded by clicking on “New” button	Please refer to Figure 5	Object and fields selected in custom metadata type settings will not reflect in Object and Field drop-down list in Object and Field configuration page, respectively

Figure 5

All Custom Metadata Types

Custom metadata types enable you to create your own setup objects whose records are metadata rather than data. These are typically used to configurations that need to be migrated from one environment to another, or packaged and installed.

Rather than building apps from data records in custom objects or custom settings, you can create custom metadata types and add metadata re manageability that comes with metadata: package, deploy, and upgrade. Querying custom metadata records doesn't count against SOQL limits.

New Custom Metadata Type				
Action	Label	Namespace Prefix	Visibility	Api Name
Manage Records 	DataMasking_Exclude Field	infy_dm	Public	infy_dm__DataMasking_Exclude_Filed__mdt
Manage Records 	DataMasking_Exclude Object	infy_dm	Public	infy_dm__DataMasking_Exclude_sObject__mdt

Step	Instruction	Illustration	Expected Result
4	Data Dictionary - Data to be substituted needs to be maintained in a separate object. Data Dictionary is an object where all data for different data types is to be maintained - Create data in Data Dictionary object to be used for substitution mechanism	Please refer to Figure 6	Data Dictionary is created

Figure 6

Data Dictionary Detail		Edit	Delete	Clone
NAME	TEST2			
STRING	Sample string			
PHONE	999999999			
EMAIL	test2@gmail.com			
TEXTAREA	TEST DATA2			
CURRENCY	\$44,565,255			
URL	http://yyyy.com			
NUMBER	43,543,543,545			
DOUBLE	20.00000			
DATE	6/5/2018			

Step	Instruction	Illustration	Expected Result
5	Select the Data Masking app from the available apps		Data Masking app with its tab opens
6	Click on the “ Data Masking ” Tab to initiate data masking for any object		Data Masking Welcome page opens

Step	Instruction	Illustration	Expected Result
7	Welcome Page - In the Welcome page, overview of the tool with different masking techniques is shown - Click on next to go to Jobs Details page to configure data masking job	Please refer to Figures 3 and 4	

Figure 3

Infosys® | Data Masking

Hello Amit Kumar Agarwal!

What is Data Masking?

Data Masking refers to the process of protecting, sanitizing or changing the certain elements in the live production data, to prevent sensitive information from misuse. It creates unauthenticated version of data which can be used for testing and user training.

Next Exit

Data Masking Mechanisms

Substitution Mechanism
Substitution technique masks data by replacing or substituting the existing content of a column with the data that appears to be similar but is entirely unrelated to the real data. For instance, the Names can be sanitized by substituting the real names with the names from a customized list of names.

Shuffling Mechanism
Shuffling technique masks data by shuffling the existing content of a column. Instead of generating the substitution data from separate files, the data is generated from the same column that needs to be masked. The data in a column is shuffled between rows to make the data entirely unrelated to the real data.

Encryption Mechanism
Encryption technique masks data by encrypting the existing content of a column using an Encryption algorithm along with a key. For instance, the Names can be encrypted by using AES256 algorithm into a binary element.

Masked-Out Mechanism
Masking out technique also known as dynamic data masking method involves Character scrambling or masking out of certain fields. It is another simplistic yet very effective method of preventing sensitive information to be viewed.

Randomization Mechanism
Randomization technique masks data in four ways:
1. Alphabetic Randomization
2. Numeric Randomization
3. AlphaNumeric Randomization
4. Random Date Mechanism

Figure 4

Infosys® | Data Masking

Hello Amit Kumar Agarwal!

No Jobs Found !!

Please create new Job using below "Create New Job" Button

Create New Job Home

Step	Instruction	Illustration	Expected Result
8	Job Details Page - Click on “Create New Job” button and provide object configuration details - or - If change to existing job is required, then click on Edit icon for selected job - or - If jobs needs to be deleted, then click on Delete icon “Delete Jobs” to delete selected jobs		Data masking job gets created/edited/deleted
9	Object Configuration Page - Select object to be masked from the drop down “Object Name” Job Name will be auto generated - Optionally provide Job Description - If certain workflow rules are associated with the object, the same can be de-activated during masking by checking “Deactivate Workflow Rule” - Click on “Next” button to proceed to next screen for selecting fields to be masked	Please refer to Figure 7	Object to be masked is selected

Figure 7

Infosys | Data Masking

Object Configuration
Data Masking: Step 1 of 3 (Object Level Configuration)

Select an Object !

Object Name: -None- ?

Job Name: ?

Job Description: ?

Deactivate Workflow Rule: ☐ ?

Previous Next Cancel Home

Step	Instruction	Illustration	Expected Result
10	Field Configuration Page Fields that can be masked for the object are shown with their data type and masking mechanism and rule to be	Please refer to Figure 8	Field masking mechanism is selected

	selected • Select the masking mechanism for the field to be masked from the drop down. Each field will have only certain masking mechanisms allowable from below: ○ Substitution – Masks data from Data Dictionary created ○ Shuffling – Masks data by shuffling the existing content of a column ○ Encryption – Masks data by standard encryption mechanism ○ Masked Out – Masking out of certain field based on configuration rule ○ Randomization – Masks data by randomizing data of the field to be masked • Select suitable masking rule against the masking mechanism selected: ○ Substitution mechanism will show field from data dictionary in masking rule drop down based on the field's data type ○ Encryption mechanism will have standard AES128, AES192 and AES256 encryption technique as its masking rule ○ Masked out rule requires masked out character and length in configuration setup ○ Alphabetic and alphanumeric randomization masking rule for string data type and numeric randomization masking rule for double data type can be selected ○ Random date masking mechanism and rule would be configured for date data type • Click on “Next” button to proceed to Configuration View page		
--	--	--	--

Figure 8

Infosys® | Data Masking Field Configuration
Data Masking: Step 2 of 3 (Field Level Configuration)

Job Details:

Job Name	Object Name	Description
Job_Lead_3	Lead	Lead Object Masking

Field Configurations: Data Masking Information ?

Field Name	Data Type	Masking Mechanism	Masking Rule
title	STRING	Shuffling	Shuffling
street	TEXTAREA	Substitution	TEXTAREA
city	STRING	Encryption	AES128 Encryption
state	STRING	Masked Out	Masked Out
postalcode	STRING	Randomization	Alphabetic
country	STRING	-None-	
phone	PHONE	-None-	
mobilephone	PHONE	-None-	
fax	PHONE	-None-	

[Configuration](#)

[Previous](#)
[Next](#)
[Cancel](#)
[Home](#)

Step	Instruction	Illustration	Expected Result
11	Configuration View Page - The page shows all the data masking configuration done for the selected object - Review the configuration and click on “Save Configuration” button and Click on “Yes” in pop-up if everything is as per the requirement Else click on “Previous” button to go back to the previous screen or - Click on “Cancel” button to go back to the Job Details page or - Click on “Home” button to go back to the Welcome page	Please refer to Figure 9	Job will be created/edited and show up in the Job Details page

Figure 9

Infosys® | Data Masking Preview Configuration
Data Masking: Step 3 of 3 (Configuration View Page)

Job Details:

Job Name	Object Name	Description
Job_Lead_3	Lead	Lead Object Masking

Configuration:

Field Name	Data Type	Masking Mechanism	Masking Rule
title	STRING	Shuffling	Shuffling
street	TEXTAREA	Substitution	infy_dm__Mailing_Address__c
city	STRING	Encryption	AES128 Encryption
state	STRING	Masked Out	Masked Out
postalcode	STRING	Randomization	Alphabetic

Step	Instruction	Illustration	Expected Result
12	Data Masking Job Execution Select “Data Masking Job” and click on “Execute Data Masking” button	Please refer to Figures 10 and 11	- Only one job should be selected. - Select one job and execute. Job will be executed and would complete with either “Complete” Status if data masking is successful else will show “Complete with Error” if some error occurs during data masking

Figure 10

Infosys[®] | Data Masking

List of all Data Masking Jobs existing in the Application
Data Masking:Jobs Details

Select Jobs to Mask Data

Job Name represents the grouping of an Object and it's fields to be Masked.

Create New Job Delete Jobs Execute Data Masking Home Export Jobs

Select Jobs	Job Name	Object Name	Records Processed	Description	Data Masking Status	Last Masked On	Created Date	Error Log File
☐	Job_Lead_3	Lead	0	Lead Object Masking	Open		7/18/2018 3:22 AM	☐

Delete Jobs Execute Data Masking Home

Figure 11

Infosys[®] | Data Masking

List of all Data Masking Jobs existing in the Application
Data Masking:Jobs Details

Select Jobs to Mask Data

Success:
The Job is under execution and will be completed Shortly..

Job Name represents the grouping of an Object and it's fields to be Masked.

Create New Job Delete Jobs Execute Data Masking Home Export Jobs

Select Jobs	Job Name	Object Name	Records Processed	Description	Data Masking Status	Last Masked On	Created Date	Error Log File
☒	Job_Lead_3	Lead	0	Lead Object Masking	Open		7/18/2018 3:22 AM	☐

Delete Jobs Execute Data Masking Home

Step	Instruction	Illustration	Expected Result
13	Error Handling - Error logs will be created under the Error Logs tab of Data Masking app to provide detail of exception/error during data masking job run - An error file will be generated with failed record IDs under Data Masking Object Config Object 's attachment related list - Once the errors are resolved based on error log and file, select the failed job from the Job Details page and execute it again	Please refer to Figures 12,13 and 14	- Only failed records of the data masking job will be picked for data masking purpose - No processing of records successfully masked in earlier job will take place - If all the failed records are successfully masked, then the job status will change to “Complete”

Figure 12

Select Jobs to Mask Data							
Job Name represents the grouping of an Object and it's fields to be Masked.							
<a>Create New Job <a>Delete Jobs <a>Execute Data Masking <a>Home <a>Export Jobs							
Select Jobs ☐	Job Name	Object Name	Records Processed	Description	Data Masking Status	Last Masked On	Created Date
☐  	Job_Lead_1	Lead	164	This is test	Completed with Error	7/18/2018 11:18 AM	7/18/2018 11:18 AM
<a>Delete Jobs <a>Execute Data Masking <a>Home							

Figure 13

<a>New Error Log <a>Change Owner 		
☐ Action	Error L Name ↑	Job Name
☐ <a>Edit <a>Del	a056A00000GsPbl	Job_Lead_1

Figure 14

Error Log Detail		<a>Edit <a>Delete <a>Clone
Error L Name	a056A00000GsPbl	
Line Number		
Message	NUMBER_OUTSIDE_VALID_RANGE: Number of Locations: value outside of valid range on numeric field: 200000.0	
Stack Trace		
Type Name		
Time Stamp	6/4/2018 2:39 AM	
Job Name	Job_Lead_1	

Step	Instruction	Illustration	Expected Result
14	Export Job - Select Job to be exported and click on “Export Jobs” button. - Credentials for the target org needs to be provided and connection established with click of “Export Jobs” button. - Data Masking jobs are exported to the target org and reflects in Job page. The exported job name has ‘imp’ appended). Please note that there should be at least one job in target org before export is initiated.	Please refer to Figures 15 , 16 and 17	Selected jobs are exported to the target org.

Figure 15

Job Name	Object Name	Records Processed	Description	Data Masking Status	Last Masked On	Created Date	Error Log File
☐ Job_Account_34	Account	137		Completed	13/7/2020 4:12 pm	13/7/2020 4:12 pm	☐
☐ Job_Account_33	Account	0		Open	25/6/2020 5:12 pm		☐
☐ Job_Lead_11	Lead	0		Open	15/6/2020 5:41 pm		☐
☐ Job_Opportunity_3	Opportunity	32		Completed	28/4/2020 6:02 pm	28/4/2020 6:00 pm	☐
☐ Job_infy_dm_DM_Test_Data__c_1	infy_dm_DM_Test_Data__c	1,024		In Progress	1/1/2019 11:33 am	1/1/2019 11:32 am	☐
☒ Job_Case_3	Case	25		Completed	28/4/2020 10:32 am	22/5/2019 12:33 pm	☐
☒ Job_Lead_9	Lead	22		Completed with Error	28/4/2020 10:28 am	2/7/2019 12:13 pm	☒

Figure 16

Success:
Data Masking Jobs has been successfully Exported

Enter Credentials

Enter Username:

Enter Password:

Enter Security Token:

Is Sandbox? ☐

Job Name	Object Name	Description	Data Masking Status	Last Masked On	Created Date
Job_Case_3	Case		Open	28/4/2020 10:32 am	22/5/2019 12:33 pm
Job_Lead_9	Lead		Open	28/4/2020 10:28 am	2/7/2019 12:13 pm
Job_Lead_10	Lead		Open		27/9/2019 4:30 pm
Job_Campaign_3	Campaign		Open	22/5/2019 12:27 pm	22/5/2019 12:22 pm
Job_Opportunity_2	Opportunity		Open	22/5/2019 12:17 pm	22/5/2019 12:13 pm
Job_Account_31	Account		Open	22/5/2019 11:51 am	22/5/2019 11:50 am

Infosys Data Masking

Work Instructions

Infosys®
Navigate your next

Figure 17

Infosys® | Data Masking

List of all Data Masking Jobs existing in the Application
Data Masking:Jobs Details

Select Jobs to Mask Data

Job Name represents the grouping of an Object and it's fields to be Masked.

Create New Job

Delete Jobs

Execute Data Masking

Home

Export Jobs

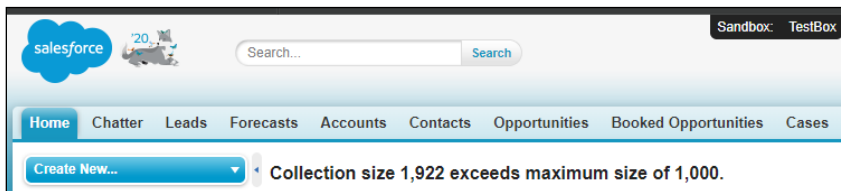
	Job Name	Object Name	Records Processed	Description	Data Masking Status	Last Masked On	Created Date	Error Log File
☐  	Job_Lead_10[imp4]	Lead	0		Open		8/17/2020 2:08 AM	☐
☐  	Job_Lead_9[imp5]	Lead	0		Open		8/17/2020 2:08 AM	☐
☐  	Job_Account_31[imp6]	Account	0		Open		8/17/2020 2:08 AM	☐
☐  	Job_Opportunity_2[imp7]	Opportunity	0		Open		8/17/2020 2:08 AM	☐
☐  	Job_Campaign_3[imp8]	Campaign	0		Open		8/17/2020 2:08 AM	☐
☐  	Job_Case_3[imp9]	Case	0		Open		8/17/2020 2:08 AM	☐
☐  	Job_Account_3	Account	0		Open		8/7/2020 5:31 AM	☐

Limitations

- The following data types are supported for data masking by this app:
 - STRING
 - TEXTAREA
 - PHONE
 - DOUBLE
 - CURRENCY
 - DATE
 - DATETIME
 - EMAIL
 - URL
- Unmasking of data is not possible
- Shuffling mechanism can shuffle data back to original status if very small amount of data is to be shuffled
- If a data masking job fails due to data issue, then the data needs to be corrected and job run again.
- Only one job should be executed at a time.

Trouble Shooting

- Error: Collection size xxx exceeds maximum size of 1000.



This error will come if number of objects is greater than 1000. In such scenario, the number of objects to be masked should be brought down to number less than 1000 using **DataMasking Exclude Object Custom Metadatatypes**.

- Converted Leads data cannot be updated unless the permission 'View and Edit Converted Leads' is assigned for the profile. After enabling the permission, one can update the converted leads.
- While exporting the data, there is a chance of getting some errors due to unauthorized endpoint in remote site settings. Check the error logs to find the accurate error which occurred during job export.

In order to rectify unauthorized endpoint error; go to setup->security Controls->Remote site settings. Click on New Remote site. Copy the URL from error logs and enter in Remote Site URL. Click on save.