

PII Microsoft SQL

# Server Data Masking



 dataZense™ Whitepaper

# Objective



This paper provides an insight into Best Practices for protecting and securing the Microsoft SQL Server sensitive PII data by implementing the Data Profiling and Data Remediation solution, thereby assuring the companies to ensure compliance with the GDPR, SOX Compliance and similar data privacy laws worldwide and maximize the business value.



## Ganapathy Arumugam

Senior Enterprise Architect  
ChainSys Corporation

Senior Enterprise Architect in ChainSys with 20+ years of experience specialized in Data Security and Compliance. Heavily engaged with the Global Clients to secure their sensitive data. Published several ML Algorithms for sensitive data identification and remediation. My current focus areas are Data Security and Compliance, Data Migration, Data Integration, Big Data, Master Data Management, Business Intelligence and Artificial Intelligence.

PII Microsoft SQL

# Server Data Masking

## BUSINESS EXPECTATIONS

Personal information or Personally identifiable information (PII) is any data that could potentially identify a specific individual. No company can claim 100% compliance, but they have to make “Good Faith Effort” to cover most of the ground. The new rules grant people more rights regarding how companies handle their personally identifiable information (PII), and it imposes heavy fines for non-compliance and data breaches--up to 4 percent of a company’s yearly revenue.

For security teams, this means making sure that PII is adequately protected and that the proper reporting processes are in place. Companies are at high risk of significant fines, data breach, brand reputation, and potential loss of customers if they do not mitigate the risk of adhering to the proper compliance levels in all regions.

Making 'Zense' of data is a challenge for most companies. dataZense is an end-to-end analytics platform that integrates all of your company’s data, secures it, and brings it to life through easy-to-use Dashboards. The Data Catalog helps look for needed data in an organization. dataZense® provides an end-to-end capability for Data Profiling, Data Remediation (Masking, Scrambling, and Encryption), and Data Governance.



# PII Data Categories

Personal data are grouped in to two different categories Critical and Confidential Data Elements or Categories.

## Critical Data Elements/Categories

Critical data elements are those elements by which you can directly identify individual/person such as their social security number, National Identifier, Visa Number, Passport Number, Employee Number etc.

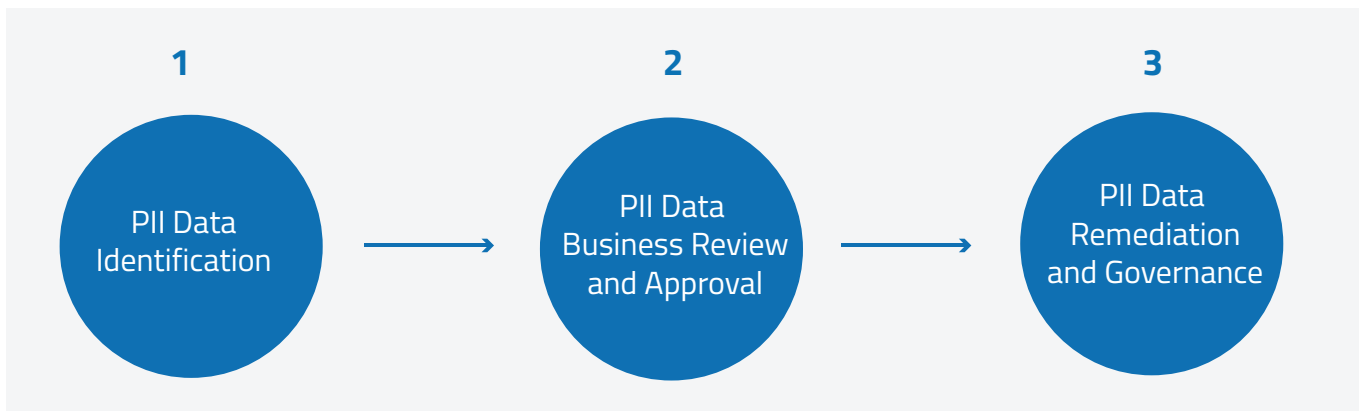
## Confidential Data Elements/Categories

Confidential data elements are information which when combined other personal data elements can able to identify individual. E.g.: First name, Last Name, Date of Birth.

## Microsoft SQL Server PII Data Profiling and Data Masking

The PII Data Profiling and Data Masking within the dataZense application is broken out into the following three high level process

- PII Data Identification
- PII Business Review and Approval
- PII Data Remediation and Governance



### PII Identification

Identify Systems considered for Sensitive Data Identification (PII) and run the Data profiling process to identify the PII Attributes. Provides the complete visibility of the identified PII data. i.e., Server, Instance, Database, Schema, Table, Column level details.

### PII Business Review and Approval

The identified PII attributes will go through False-Positive / False-Negative analysis and approved for remediation

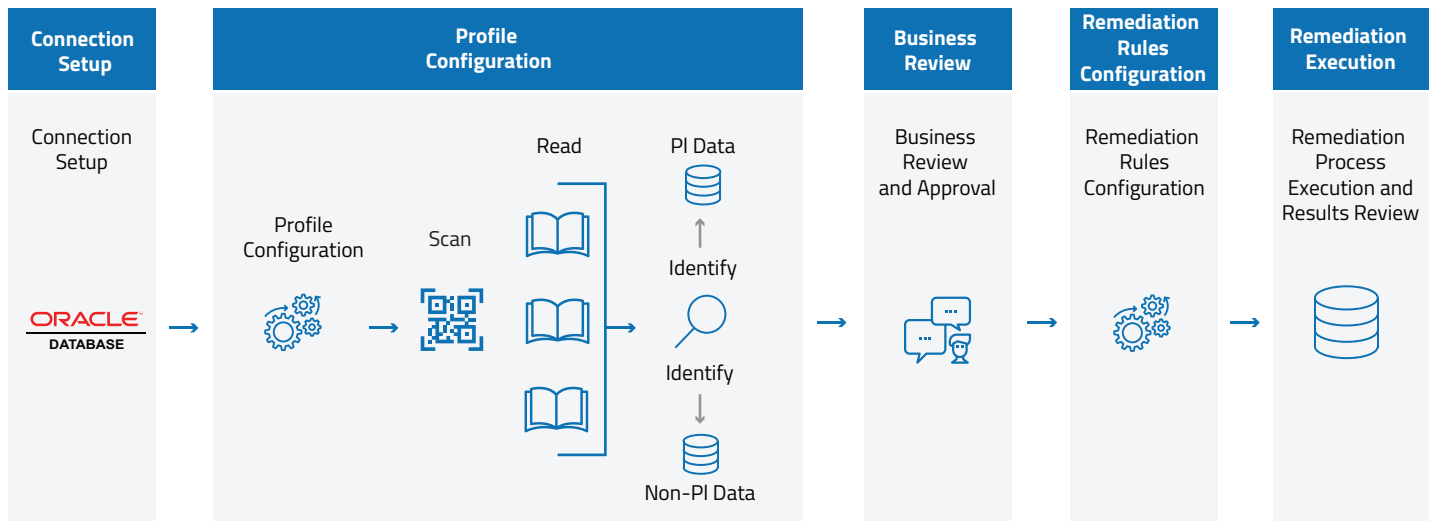
### PII Remediation and Governance

The business approved PII attributes will be considered for remediation.

## Remediation Options:

- Scramble the Data
- Mask the Data
- Remove Data (With capture of logs) – remove from all production and non-production environments
- Keep the Data (Encrypt in all Non-Production environments)

## High Level Architecture



## Data Masking Process Steps

### Steps to configure:

- Create a Connection to the Microsoft SQL Server Database.
- Create a Data Security profile for scanning the Microsoft SQL Server Database.
- Business Review of scanned results and complete the Business validation of Approve/Reject the PII Data.
- Create a Remediation Profile to configure the Remediation Rules for masking the data.
- Execute the configured Remediation Profile to mask the data and Review the masked results from the Microsoft SQL Server database.

## Connection

Verifying the Source Connection:

Ensure that the respective Microsoft SQL Server source connection is available. If not, create a new connection.

The screenshot displays the 'MS SQL Server Connection' configuration page in the Smart Data Platform. The interface is divided into several sections:

- General Information:** Includes fields for Sprint Name (DEV), Connection Name (MS SQL SERVER), Display Name (MS SQL SERVER), and Active status (Yes). It also has fields for Application Name and Description.
- Database Information:** Includes Database Version (Select Database Version), Host (192.168.57.102), Username (appmdm), Database (appmdm\_stg), Port (1433), and Password.
- Data Mart Information:** Includes a Data Mart field.
- Connect as:** Checkboxes for Source and Destination are both checked.

## Profile Creation and Execution

Create a profile for scanning the Microsoft SQL Server Database and execute.

The screenshot displays the 'Create New Profile' configuration page in the Smart Data Platform. The interface includes the following configuration options:

- Name:** CS\_MSSQL\_DEMO\_PROFILE
- Display Name:** CS\_MSSQL\_DEMO\_PROFILE
- Description:** CS\_MSSQL\_DEMO\_PROFILE
- Data Fetch Mode:** Sample
- Record Limit:** 1000
- Sampling Method:** Static
- Consider Previous Scanning Results:** Yes
- Auto Business Approval Required:** No
- Auto Remediation Required:** No
- Connection Parallel Process:** 5
- Table Parallel Process:** 10

On the right side, there is a 'Selected Connections' panel showing one connection: MS SQL SERVER (DEV) with an 'Edit' button.

# Business Review

Review the scanned results and complete the Business validation by Approving or Rejecting the PII Data.

**Business Validation and Approval**

Data Security Profile: CS\_MSSQL\_DEMO\_PROFILE | Execution Details: 01/22/2021 10:41:41 AM | Connections: MS SQL SERVER

Schema Name: dbo | Table Name: CP\_EMPLOYEES\_DETAILS | Status: Approved

Server Name: MS SQL SERVER | Instance Name: MYSQLSERVER | Database Name: appmdm\_stg

Table Name	Column Name	Possible Category	Match(%)	Category Name	Status	Sample
CP_EMPLOYEES_DETAILS	FULL_NAME	NAME	98	NAME	Approved	<a href="#">View Data</a>
CP_EMPLOYEES_DETAILS	IP_ADDRESSES	IP ADDRESS	74	IP ADDRESS	Approved	<a href="#">View Data</a>
CP_EMPLOYEES_DETAILS	PARENTS_NAME	NAME	87	PARENTS_NAME	Approved	<a href="#">View Data</a>
CP_EMPLOYEES_DETAILS	DEBIT_CARD_NUMBER	CREDIT_CARD	100	CREDIT_CARD	Approved	<a href="#">View Data</a>
CP_EMPLOYEES_DETAILS	NATIONAL_IDENTIFIER	NATIONAL IDENTIFIER	100	NATIONAL_IDENTIFIER	Approved	<a href="#">View Data</a>
CP_EMPLOYEES_DETAILS	AGE	AGE	83	AGE	Approved	<a href="#">View Data</a>
CP_EMPLOYEES_DETAILS	CURRENT_ADDRESS	LOCATION	29	LOCATION	Approved	<a href="#">View Data</a>
CP_EMPLOYEES_DETAILS	DATE_OF_BIRTH	DATE	99	DATE	Approved	<a href="#">View Data</a>
CP_EMPLOYEES_DETAILS	FIRST_NAME	NAME	94	NAME	Approved	<a href="#">View Data</a>

# Remediation Profile

Create a Remediation Profile to configure the Remediation Rules for masking the data.

**PII Remediation Profile Lookup Setup**

PII Remediation Name: CS\_PII\_REMEDIATION\_PROFILE | Apply To: Data Security

Category Rule: 99 Total

Category	Category Type	Security Level	Remediation Type	Mask/Scramble Type
<input checked="" type="checkbox"/> CLIENT APPLICATIONS AND OS	STRING	Confidential	Scramble	NA
<input checked="" type="checkbox"/> CONTACTS	STRING	Confidential	Scramble	NA
<input checked="" type="checkbox"/> CONTACTS COMMUNICATIONS	STRING	Confidential	Scramble	NA
<input checked="" type="checkbox"/> CORPORATIONS	STRING	Confidential	Scramble	NA
<input checked="" type="checkbox"/> COUNTRY	STRING	On the Record	Scramble	NA
<input checked="" type="checkbox"/> CREDIT CARD	CREDIT CARD	Critical	Mask	KEEP_LAST4
<input checked="" type="checkbox"/> CREDIT CARD NUMBER	CREDIT CARD	Critical	Mask	KEEP_LAST4
<input checked="" type="checkbox"/> CURRENT ADDRESS	STRING	Confidential	Scramble	NA
<input checked="" type="checkbox"/> DATE	DATE	Confidential	Mask	DAY
<input checked="" type="checkbox"/> DATE OF BIRTH	DATE	Confidential	Mask	DAY
<input checked="" type="checkbox"/> DATE OF DEATH	DATE	Confidential	Mask	DAY
<input checked="" type="checkbox"/> DEVICE IP	IP ADDRESSES	Confidential	Delete	NA
<input checked="" type="checkbox"/> DIGITAL RECORDS OF PHYSICAL	STRING	Confidential	Scramble	NA
<input checked="" type="checkbox"/> DISABILITY	STRING	Confidential	Scramble	NA

# Remediation Execution

Execute the configured Remediation Profile to mask the data.

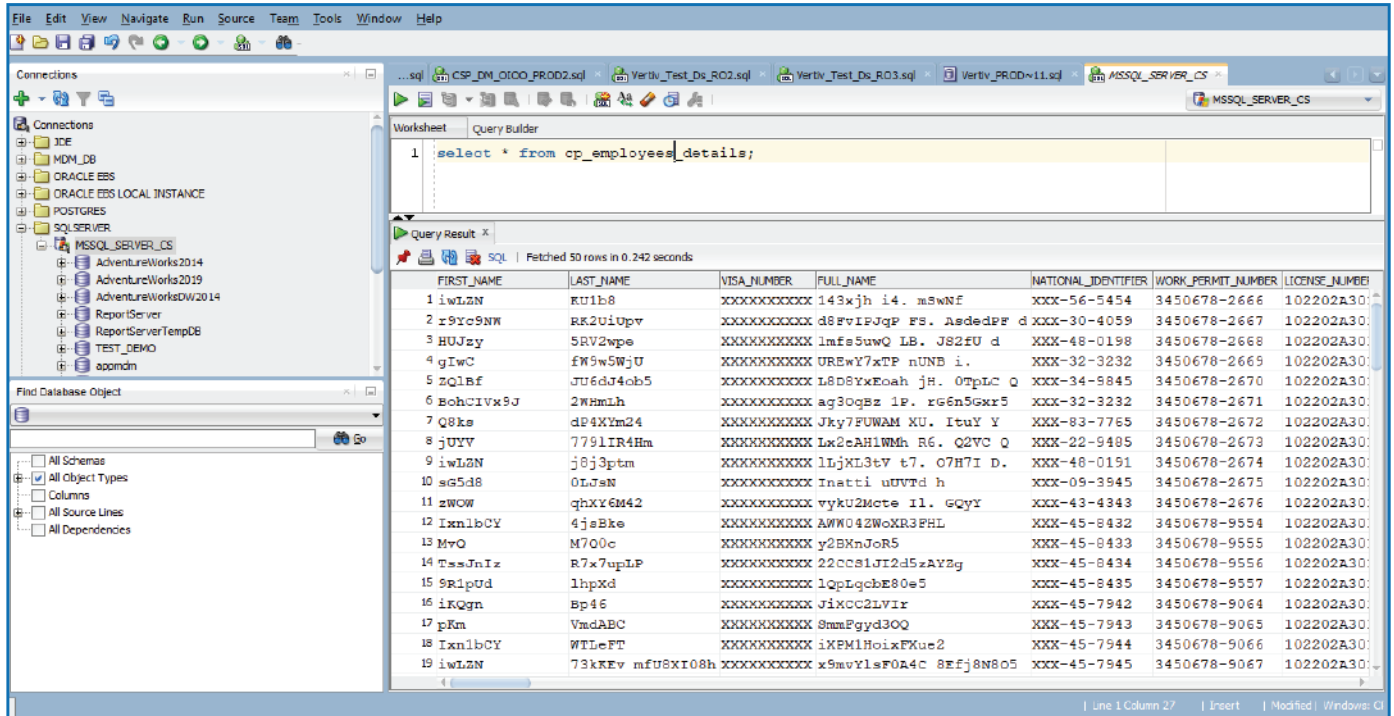
# Remediation Process Status

Remediation Profile Name	Server Name	Instance	Database	Schema Name	Table Name	Column Name	Category	Remediation	Status	Log
CS_PII_REMEDIATION_PROFILE	MS SQL SERV...	MYSQLSERVER	appmdm_stg	dbo	CP_EMPLOYE...	FAVORITES_U...	URL	Mask	Completed	
CS_PII_REMEDIATION_PROFILE	MS SQL SERV...	MYSQLSERVER	appmdm_stg	dbo	CP_EMPLOYE...	REGION_OF_...	STRING	Scramble	Completed	
CS_PII_REMEDIATION_PROFILE	MS SQL SERV...	MYSQLSERVER	appmdm_stg	dbo	CP_EMPLOYE...	VISA_NUMBER	STRING	Mask	Completed	
CS_PII_REMEDIATION_PROFILE	MS SQL SERV...	MYSQLSERVER	appmdm_stg	dbo	CP_EMPLOYE...	DEBIT_CARD...	CREDIT CARD	Mask	Completed	
CS_PII_REMEDIATION_PROFILE	MS SQL SERV...	MYSQLSERVER	appmdm_stg	dbo	CP_EMPLOYE...	COUNTRY	STRING	Scramble	Completed	
CS_PII_REMEDIATION_PROFILE	MS SQL SERV...	MYSQLSERVER	appmdm_stg	dbo	CP_EMPLOYE...	CREDIT_CAR...	CREDIT CARD	Mask	Completed	
CS_PII_REMEDIATION_PROFILE	MS SQL SERV...	MYSQLSERVER	appmdm_stg	dbo	CP_EMPLOYE...	ADDRESS_1	STRING	Scramble	Completed	
CS_PII_REMEDIATION_PROFILE	MS SQL SERV...	MYSQLSERVER	appmdm_stg	dbo	CP_EMPLOYE...	CHILDREN_N...	STRING	Scramble	Completed	
CS_PII_REMEDIATION_PROFILE	MS SQL SERV...	MYSQLSERVER	appmdm_stg	dbo	CP_EMPLOYE...	CURRENT_A...	STRING	Scramble	Completed	
CS_PII_REMEDIATION_PROFILE	MS SQL SERV...	MYSQLSERVER	appmdm_stg	dbo	CP_EMPLOYE...	SPOUSE_NA...	STRING	Scramble	Completed	
CS_PII_REMEDIATION_PROFILE	MS SQL SERV...	MYSQLSERVER	appmdm_stg	dbo	CP_EMPLOYE...	ITIN_NUMBER	NATIONAL ID...	Mask	Completed	
CS_PII_REMEDIATION_PROFILE	MS SQL SERV...	MYSQLSERVER	appmdm_stg	dbo	CP_EMPLOYE...	EMAIL_ADDR...	EMAIL	Mask	Completed	



# Review Masked Results

Review the masked results from the Microsoft SQL Server database.



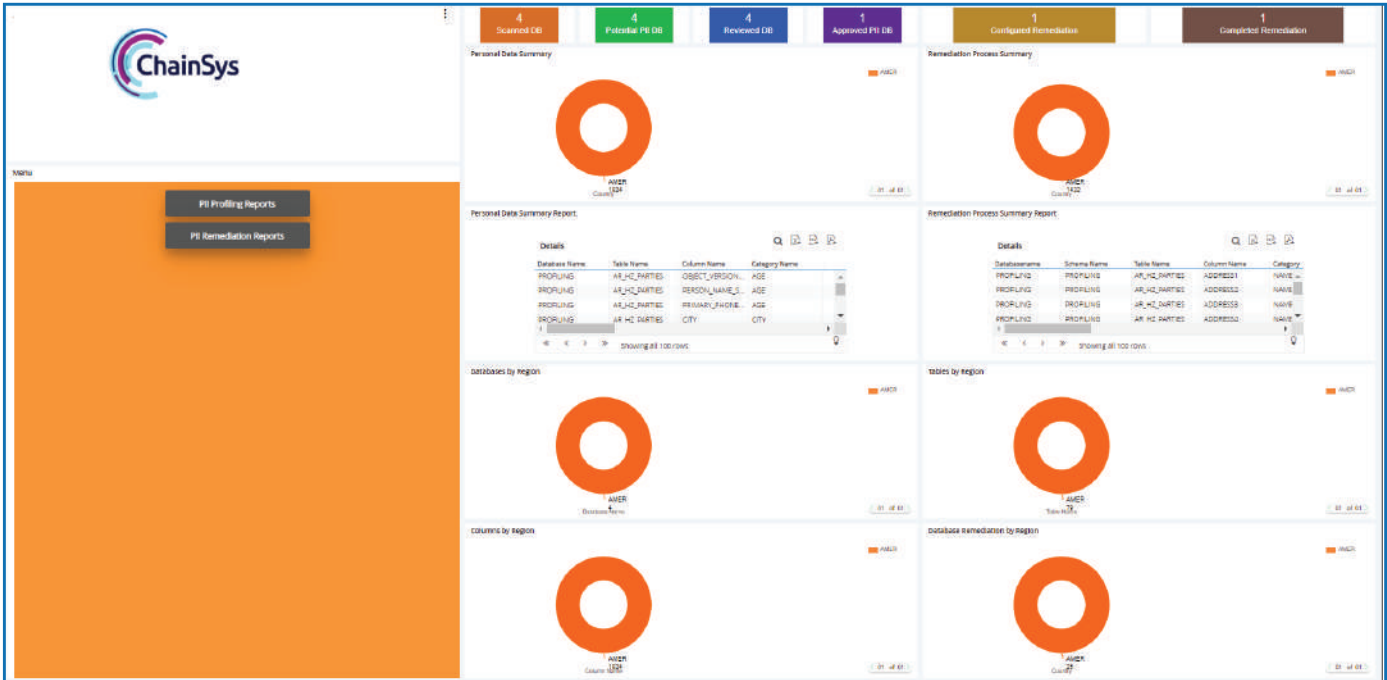
The screenshot shows the Microsoft SQL Server Enterprise Manager interface. The 'Connections' pane on the left shows a tree view of the server hierarchy, with 'MSSQL\_SERVER\_CS' selected. The 'Query Builder' pane shows a query: `select * from cp_employees_details;`. The 'Query Result' pane displays a table with 19 rows and 6 columns: FIRST\_NAME, LAST\_NAME, VISA\_NUMBER, FULL\_NAME, NATIONAL\_IDENTIFIER, WORK\_PERMIT\_NUMBER, and LICENSE\_NUMBER. The VISA\_NUMBER and FULL\_NAME columns are masked with 'XXXXXXXXXX'. The NATIONAL\_IDENTIFIER column contains values like 'XXX-56-5454', 'XXX-30-4059', etc. The WORK\_PERMIT\_NUMBER and LICENSE\_NUMBER columns contain values like '3450678-2666', '102202A30', etc.

FIRST_NAME	LAST_NAME	VISA_NUMBER	FULL_NAME	NATIONAL_IDENTIFIER	WORK_PERMIT_NUMBER	LICENSE_NUMBER
1 iwLzN	RU1b8	XXXXXXXXXX	143xjh i4. mSwNF	XXX-56-5454	3450678-2666	102202A30
2 r9Yc9NW	RK2UiUpv	XXXXXXXXXX	d8FvIFJgP FS. AsdedPF d	XXX-30-4059	3450678-2667	102202A30
3 HUJzy	SRV2wpe	XXXXXXXXXX	lmfe5wQ LB. JS2FU d	XXX-48-0198	3450678-2668	102202A30
4 gIwC	fW9w5WjU	XXXXXXXXXX	UREwY7xTP nUNB i.	XXX-32-3232	3450678-2669	102202A30
5 zQ1Bf	JU6dJ4ob5	XXXXXXXXXX	L8D8YkEoah jH. 0TptLC Q	XXX-34-5845	3450678-2670	102202A30
6 BohCIVx5J	2WHmLh	XXXXXXXXXX	ag3OqBz 1P. rG6n5Gxr5	XXX-32-3232	3450678-2671	102202A30
7 Q8ke	dP4XVm24	XXXXXXXXXX	Jky7FUWAM XU. ItuY Y	XXX-83-7765	3450678-2672	102202A30
8 jUYV	7791IR4Hm	XXXXXXXXXX	Lx2eAH1WMh R6. Q2VC Q	XXX-22-9485	3450678-2673	102202A30
9 iwLzN	j8j3ptm	XXXXXXXXXX	1LjXL3tv t7. O7H7I D.	XXX-48-0151	3450678-2674	102202A30
10 sg5d8	0LJsn	XXXXXXXXXX	Inatti uDUVtd h	XXX-09-3945	3450678-2675	102202A30
11 zwQW	qhXY6M42	XXXXXXXXXX	vykU2Mcte 1l. GQyY	XXX-43-4343	3450678-2676	102202A30
12 Ixn1bCY	4jaEke	XXXXXXXXXX	AWW042WoXR3PHL	XXX-45-8432	3450678-9554	102202A30
13 MvO	M700c	XXXXXXXXXX	y2BknJoR5	XXX-45-8433	3450678-9555	102202A30
14 TssJnIz	R7x7upLP	XXXXXXXXXX	22cc81JI2d5zAY2g	XXX-45-8434	3450678-9556	102202A30
15 9R1pUd	1hpXd	XXXXXXXXXX	lQpLqcbE80e5	XXX-45-8435	3450678-9557	102202A30
16 iKQgn	Bp46	XXXXXXXXXX	JixCC2LVir	XXX-45-7942	3450678-9064	102202A30
17 pKm	VndABC	XXXXXXXXXX	SmmPgyd3OQ	XXX-45-7943	3450678-9065	102202A30
18 Ixn1bCY	WTDeFT	XXXXXXXXXX	iXPM1HoixFKue2	XXX-45-7944	3450678-9066	102202A30
19 iwLzN	73kREv mFU8XI08h	XXXXXXXXXX	x5mV1sF0A4C 8Efj8N8o5	XXX-45-7945	3450678-9067	102202A30

# Dashboards

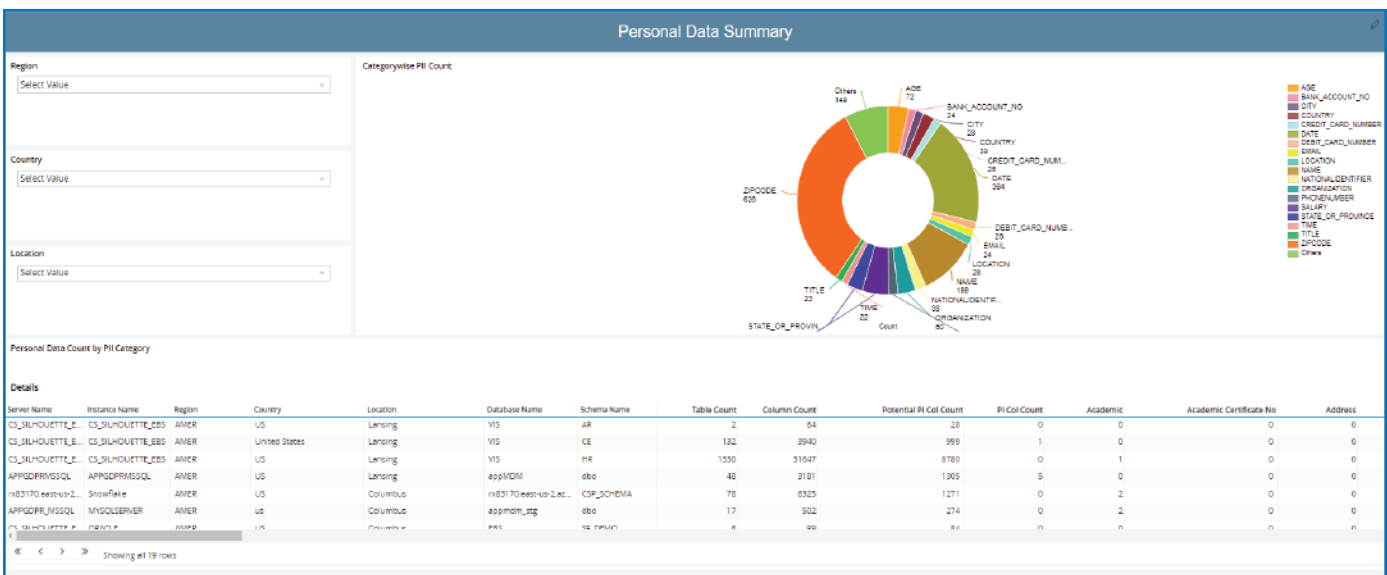
## PII Home Dashboard

Provides the high-level summary of Profiled and Remediated Databases



## Personal Data Count by PII Category

Provides the PII data count by each PII Category



## Supported Endpoints ( Partial )

Oracle Sales Cloud, Oracle Marketing Cloud, Oracle Engagement Cloud, Oracle CRM On Demand, SAP C/4HANA, SAP S/4HANA, SAP BW, SAP Concur, SAP SuccessFactors, Salesforce, Microsoft Dynamics 365, Workday, Infor Cloud, Procore, Planview Enterprise One

Cloud Applications

Oracle E-Business Suite, Oracle ERP Cloud, Oracle JD Edwards, Oracle PeopleSoft, SAP S/4HANA, SAP ECC, IBM Maximo, Workday, Microsoft Dynamics, Microsoft Dynamics GP, Microsoft Dynamics Nav, Microsoft Dynamics Ax, Smart ERP, Infor, BaaN, Mapics, BPICS

Enterprise Applications

Windchill PTC, Oracle Agile PLM, Oracle PLM Cloud, Teamcenter, SAP PLM, SAP Hybris, SAP C/4HANA, Enovia, Proficy, Honeywell OptiVision, Salesforce Sales, Salesforce Marketing, Salesforce CPQ, Salesforce Service, Oracle Engagement Cloud, Oracle Sales Cloud, Oracle CPQ Cloud, Oracle Service Cloud, Oracle Marketing Cloud, Microsoft Dynamics CRM

PLM, MES & CRM

Oracle HCM Cloud, SAP SuccessFactors, Workday, ICON, SAP APO and IBP, Oracle Taleo, Oracle Demantra, Oracle ASCP, Steelwedge

HCM & Supply Chain Planning

Oracle Primavera, Oracle Unifier, SAP PM, Procore, Ecosys, Oracle EAM Cloud, Oracle Maintenance Cloud, JD Edwards EAM, IBM Maximo

Project Management & EAM

OneDrive, Box, SharePoint, File Transfer Protocol (FTP), Oracle Webcenter, Amazon S3

Enterprise Storage Systems

HIVE, Apache Impala, Apache Hbase, Snowflake, mongoDB, Elasticsearch, SAP HANA, Hadoop, Teradata, Oracle Database, Redshift, BigQuery

Big Data

mangoDB, Solr, CouchDB, Elasticsearch

No SQL Databases

PostgreSQL, Oracle Database, SAP HANA, SYBASE, DB2, SQL Server, MySQL, memsql

Databases

IBM MQ, Active MQ

Message Broker

Java, .Net, Oracle PaaS, Force.com, IBM, ChainSys Platform

Development Platform

# One Platform for your

End to End

Data Management needs

## Smart Data Platform™



Data Migration  
Data Reconciliation  
Data Integration



Data Quality Management  
Data Governance  
Analytical MDM



Data Analytics  
Data Catalog  
Data Security & Compliance

[www.chainsys.com](http://www.chainsys.com)