

Snowflake

Data Masking



 dataZense™ Whitepaper

Objective



This paper provides an insight into Best Practices for protecting and securing the Snowflake 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 Snowflake

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.

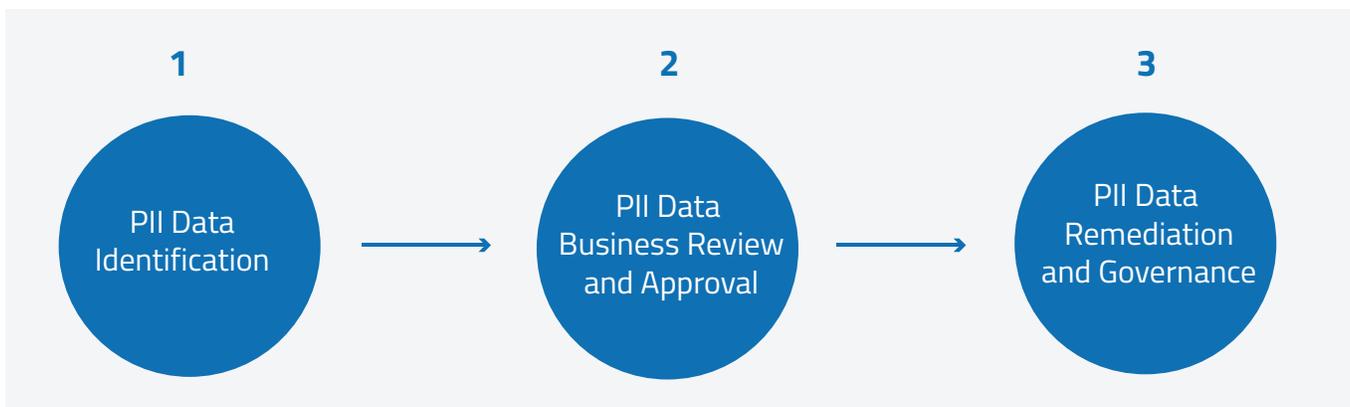
Confidential Data Elements/Categories

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.

Snowflake 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

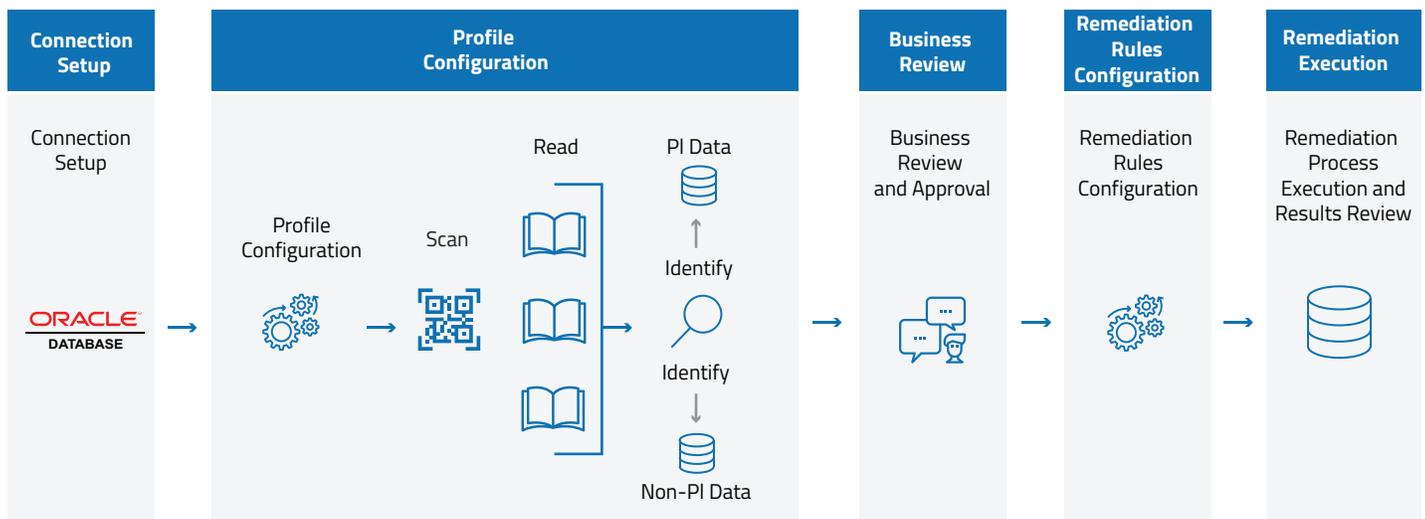
PII Remediation and Governance

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

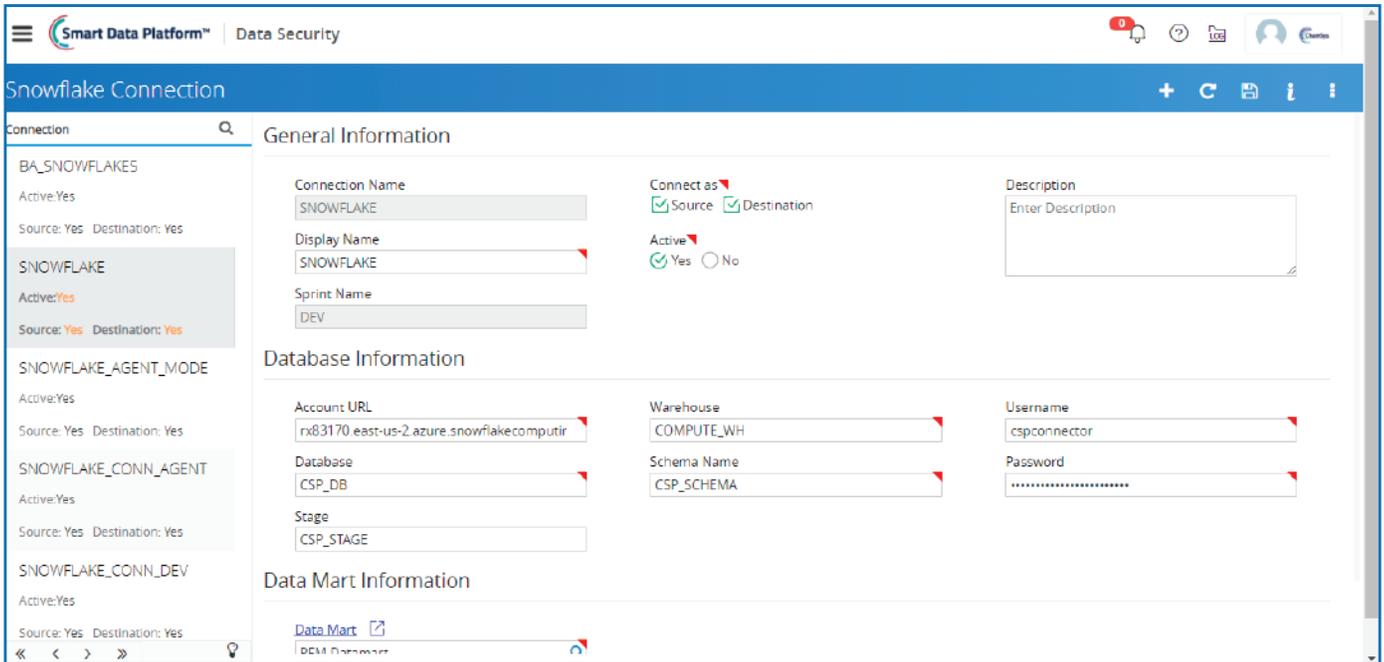
Steps to configure:

- Create a Connection to the Salesforce instance.
- Create a Data Security profile for scanning the Salesforce Instance.
- 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 Salesforce Application and database.

Connection

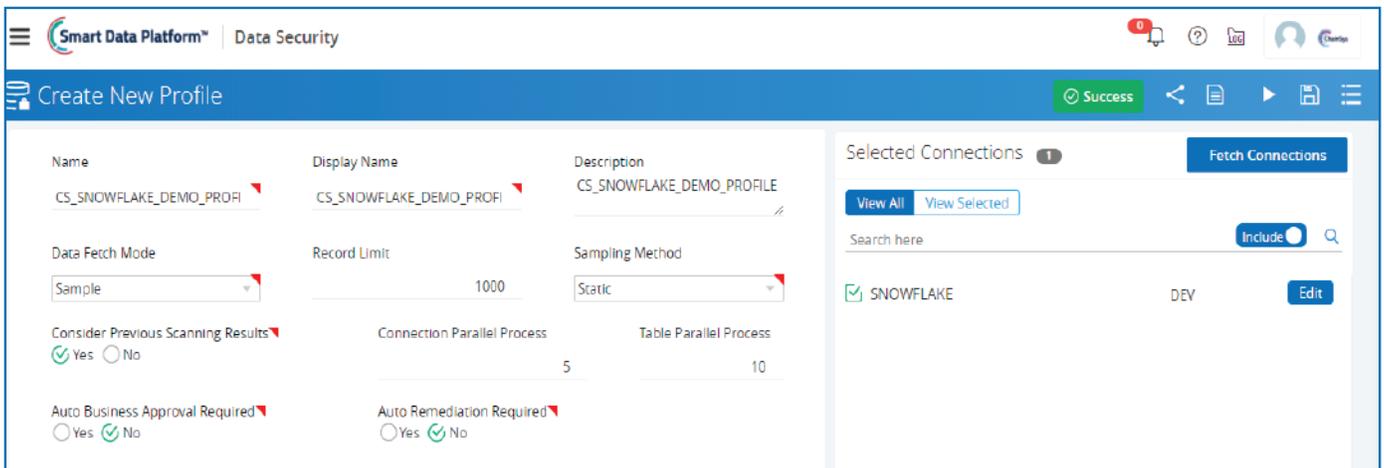
Verifying the Source Connection:

Ensure that the respective Snowflake source connection is available. If not, create a new connection.



Profile Creation and Execution

Create a profile for scanning the Snowflake Database and execute



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_SALESFORCE_DEMO_PROFILE

Execution Details: 12/07/2020 22:32:13 PM

Connections: Salesforce

Schema Name: SF_DEMO

Table Name: Table Name

Status: Approved

Server Name: SALESFORCE Instance Name: SALESFORCE Database Name: SALESFORCE

Table Name	Column Name	Possible Category	Match(%)	Category Name	Status	Sample
CP_EMPLOYEE_DETAILS	REGION_OF_BIRTH	COUNTRY	100	COUNTRY	Approved	View Data
CP_EMPLOYEE_DETAILS	CREDIT_CARD_NUMB...	CREDIT_CARD	100	CREDIT_CARD	Approved	View Data
CP_EMPLOYEE_DETAILS	NATIONAL_IDENTIFIER	NATIONALIDENTIFIER	100	NATIONAL_IDENTIFIER	Approved	View Data
CP_EMPLOYEE_DETAILS	EMAIL_ADDRESSES	EMAIL	100	EMAIL	Approved	View Data
CP_EMPLOYEE_DETAILS	LAST_NAME	NAME	87	NAME	Approved	View Data
CP_EMPLOYEE_DETAILS	LEGAL_NAMES	NAME	88	NAME	Approved	View Data
CP_EMPLOYEE_DETAILS	VISA_NUMBER	VISA	100	VISA	Approved	View Data
CP_EMPLOYEE_DETAILS	COUNTRY	COUNTRY	100	COUNTRY	Approved	View Data
CP_EMPLOYEE_DETAILS	ADDRESS_1	LOCATION	28	ADDRESS	Approved	View Data

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

Description: CS_PII_REMEDIATION_PROFILE

Apply To: Data Security

Category Rule: 99 Total

Category	Category Type	Security Level	Remediation Type	Mask/Scramble Type
<input 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

Buttons: Cancel, Save Profile

Remediation Execution

Execute the configured Remediation Profile to mask the data.

Remediate by Configured Profiles

Configuration Profile Name: Contains CS_PII_REMEDIATION_PROFILE
 PII Category Type: Equals --Select--

Search Clear Minimize

Results 1 Records Remediate Selected

Configuration Profile Name	Description	PII Categories	Databases
CS_PII_REMEDIATION_PROFILE	CS_PII_REMEDIATION_PROFILE	22	7

Remediation Execution

Remediation Process Status

Category Name: ACADEMIC, ACADEMIC CERTIFICATE N
 Region: Region
 Country: Country
 Location: Location
 Remediation Status: Completed
 Instance: Instance
 Database: Database
 Schema Name: Schema Name

Search Reset

Remediation Profile Name	Server Name	Instance	Database	Schema Name	Table Name	Column Name	Category	Remediation	Status	Log
CS_SNOWFLAKE_PII_REMEDIA...	https://rx831...	Snowflake	CSP_SCHEMA	PROFILING	CP_CONTACT...	EMAIL_ADDR...	EMAIL	Mask	Completed	Log
CS_SNOWFLAKE_PII_REMEDIA...	https://rx831...	Snowflake	CSP_SCHEMA	PROFILING	CP_CONTACT...	NATIONAL_I...	NATIONAL ID...	Mask	Completed	Log
CS_SNOWFLAKE_PII_REMEDIA...	https://rx831...	Snowflake	CSP_SCHEMA	PROFILING	CP_CONTACT...	FIRST_NAME	STRING	Scramble	Completed	Log
CS_SNOWFLAKE_PII_REMEDIA...	https://rx831...	Snowflake	CSP_SCHEMA	PROFILING	CP_CONTACT...	LAST_NAME	STRING	Scramble	Completed	Log

Review Masked Results

Review the masked results from the Snowflake database.

The screenshot shows the Snowflake web interface. On the left is a sidebar with a tree view of database objects. The main area displays a SQL query and its results. The query is:

```
1 SELECT CONTACT_ID,  
2 FIRST_NAME,  
3 LAST_NAME,  
4 NATIONAL_IDENTIFIER,  
5 BANK_ACCOUNT_NO,  
6 CREDIT_CARD_NO,  
7 DEBIT_CARD_NO,  
8 EMAIL_ADDRESSES  
9 FROM CP_CONTACT_DETAILS;
```

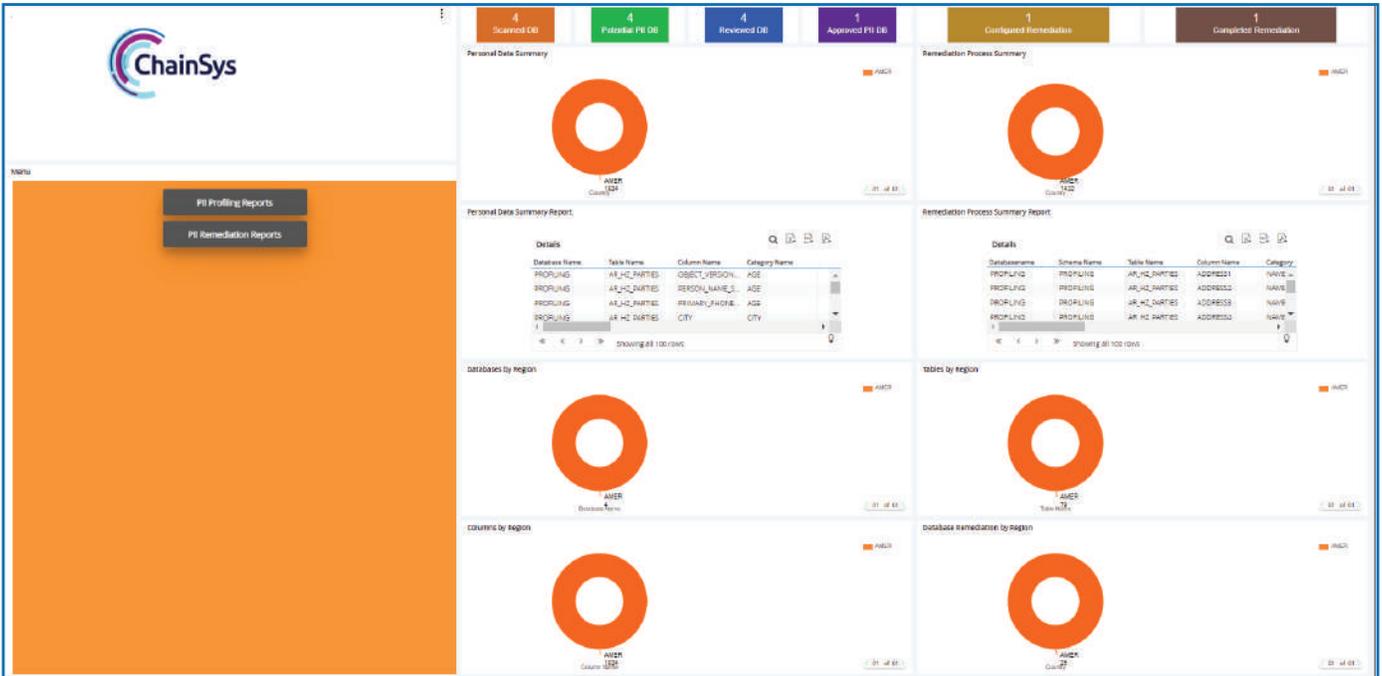
The results are shown in a table with 7 rows and 8 columns. The columns are CONTACT_ID, FIRST_NAME, LAST_NAME, NATIONAL_IDENTIFIER, BANK_ACCOUNT_NO, CREDIT_CARD_NO, DEBIT_CARD_NO, and EMAIL_ADDRESSES. The data is masked with 'X' characters.

Row	CONTACT_ID	FIRST_NAME	LAST_NAME	NATIONAL_IDENTIFIER	BANK_ACCOUNT_NO	CREDIT_CARD_NO	DEBIT_CARD_NO
1	13092486	E6Rk9H	alkGy7	XXX-30-1409	0005-501-1-1200907	XXXX-XXXX-XXXX-1804	XXXX-
2	13092488	GvfG3L	z3uk	XXX-30-1430	0005-501-1-1200909	XXXX-XXXX-XXXX-1806	XXXX-
3	13092489	IYK	aAMwhJd	XXX-30-1382	0005-501-1-1200910	XXXX-XXXX-XXXX-1807	XXXX-
4	13092197	p5Ow	SmpXSRvU	XXX-43-4343	0005-501-1-1200911	XXXX-XXXX-XXXX-1515	XXXX-
5	13092198	cMVz2K	zNeJ7x	XXX-55-5555	0005-501-1-1200912	XXXX-XXXX-XXXX-1516	XXXX-
6	13092200	wmOE	2bl9jbO7R	XXX-76-0000	0005-501-1-1200914	XXXX-XXXX-XXXX-1518	XXXX-
7	13092201	eCwiOc	5zefcS	XXX-00-9987	0005-501-1-1200915	XXXX-XXXX-XXXX-1519	XXXX-

Dashboards

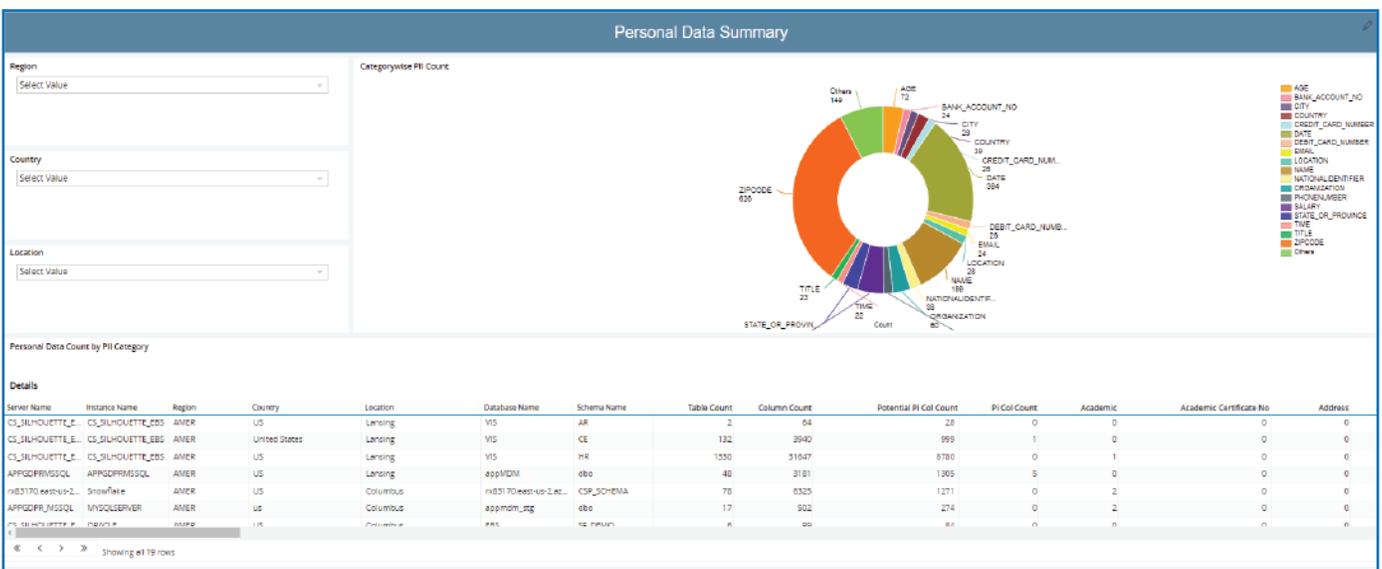
PII Home Dashboard

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