

INVENTA LAB GUIDE

VERSION 4.1.1

October 2024



TABLE OF CONTENTS

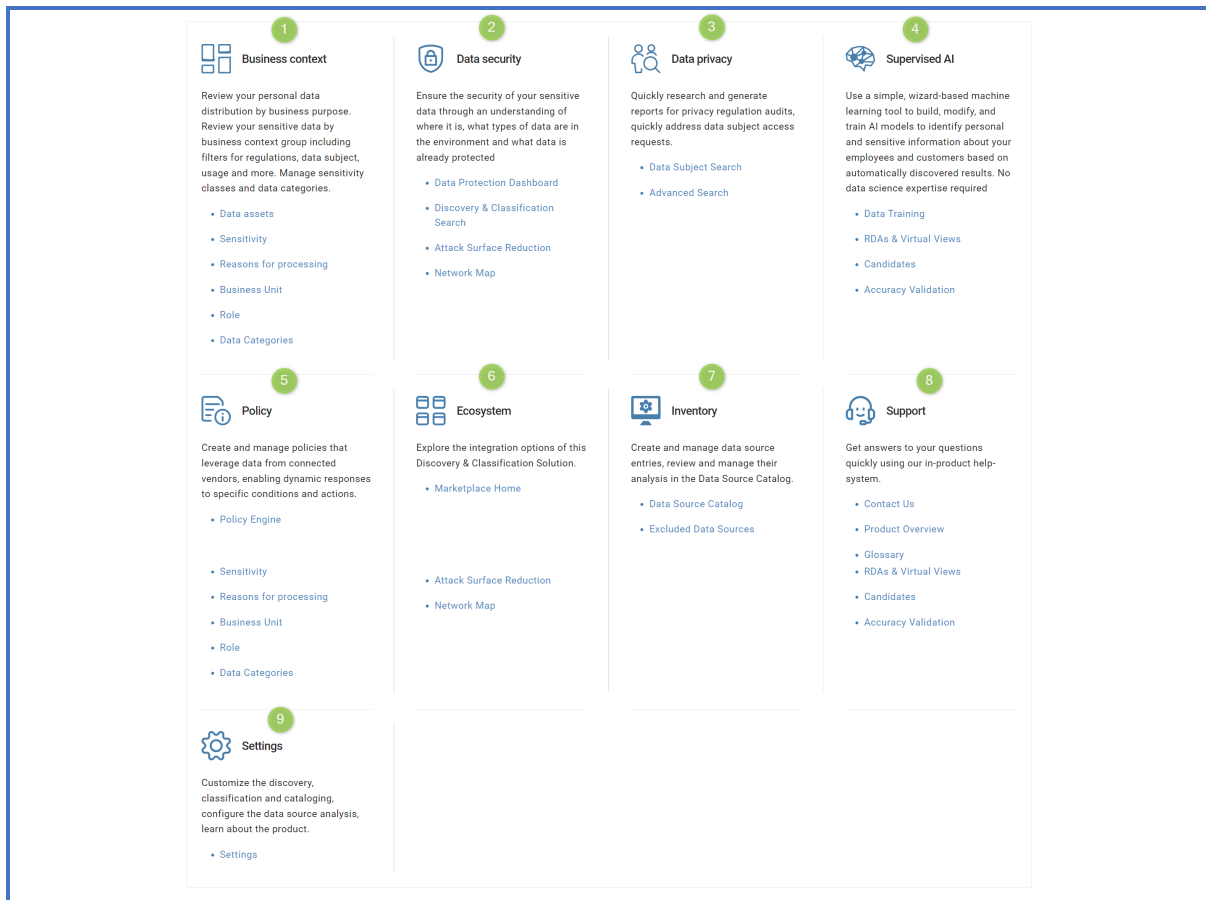
Table of Contents	2
Introduction	3
Module overview	3
Adding credentials	4
Adding data sources	5
Data subjects	6
Adding data elements	8
Adding data assets	10
Data subjects	12
Data subject search	14
Attack surface reduction dashboard	14
Discovery & classification search	16

INTRODUCTION

The following lab guide introduces the key features of Inventa, including the ability to add data sources, train the system on personally identifiable information (PII), create data assets and data subject catalogs, and search across structured and unstructured data to identify sensitive information and track data retention.

MODULE OVERVIEW

Inventa consists of 7 modules each responsible for separate functionality.



Business context module is used primarily to get insights into data after all the necessary scans have been completed.

Data security module is used to gain understanding of how sensitive data is being stored and processed.

Data privacy module is used for DSAR requests and personal data.

Supervised AI module is used for data training, it is an added layer of intelligence available after the initial classification.

Policy engine module is used to configure triggers, workflows, and other security and privacy tools.

Ecosystem module is used for integrations with other tools.

Inventory module is an inventory of sensitive data .

Support module is used for contacting 1touch.io support representatives.

Settings module is used for data source, credential, analytic engine, and other system elements configuration.

ADDING CREDENTIALS

To start off, add credentials for one of the desired data sources.

1. In the **Sidebar**, select **Settings > Data Source Catalog > Credentials (1)**. On the **Credentials** page, click the **Add** button (2).

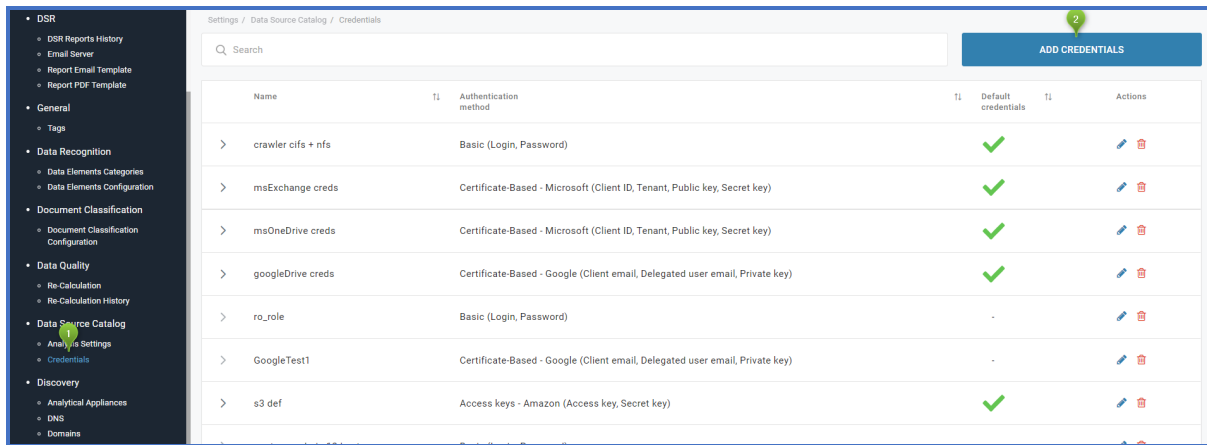


Figure 1: Selecting the Console Management Configuration page

2. In the new credentials form, enter the general information, for example, for Dropbox.

Table 2: General form for new credentials

PARAMETER	DESCRIPTION
Name (1)	Enter the credentials name to be used within the application. The name will be shown on the Credentials page and in the Credentials dropdown list while creating or editing a data source. For example, <i>Database credentials</i> .
Authentication method (2)	Select the desired method.

Figure 3: General form for adding a credential entry

3. Fill the credentials form according to the selected authentication method and click the **Save** button (3).

ADDING DATA SOURCES

1. To create a data source entry, go to **Sidebar > Inventory > Data Source Catalog**, then click the **Add Data Source** button. You will be redirected to the **Add Data Source** page.

2. In the **Data Source Type** dropdown list (1), select **Cloud Storage**. In the **Data Source Product** dropdown list (2), select **Dropbox**.

Figure 4: Selecting a Data Source Type and Product

3. In the **Data Source Details** form, enter the parameters required to connect to the data source, in this case, Dropbox.

Table 5: Data Source Details form

PARAMETER	DESCRIPTION
Analytic engine (1)	Name of the analytic engine that will analyze the data source. For example, core .
Analysis strategy (2)	Strategy the plugin will use to analyze the data source.
Credentials (3/4)	Select the credentials for access to the data source from the Credentials dropdown list (3) The list is available if the Use default credentials checkbox (4) is disabled.
Min timeout between analysis (5/6)	Set the interval of time to pass before the system can run data source analysis after the previous analysis cycle is completed. (Optional)

Figure 6: Data Source Details tab

5. To finish the data source record setup process, configure the analysis schedule in the **Individual Analysis Schedule** tab and assign the appropriate tags in the **Tags** tab.

6. To save the data source record, click **Save (8)**. This will add the data source to the list of data sources on the **Data Source Catalog** page with Status=New.

DATA SUBJECTS

After a couple of scans, the system is likely to identify a couple of data subject candidates. To manage them and further advance data recognition, proceed to the Supervised AI module.

1. Navigate to **Sidebar > Supervised AI > Data Training**. Review the candidate virtual views, explore the **Filters** panel on the left that allows for data sorting based on candidate group size, products, and analytic engine name.

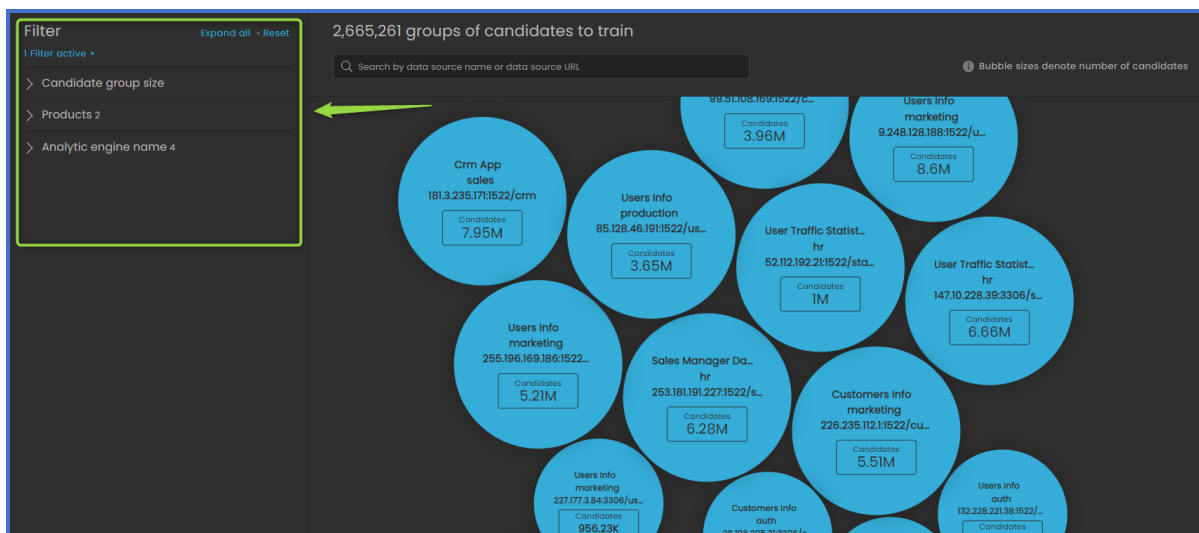


Figure 7: Filters panel

2. Hover over one of the bubbles and click the **Train** button to proceed with AI training.

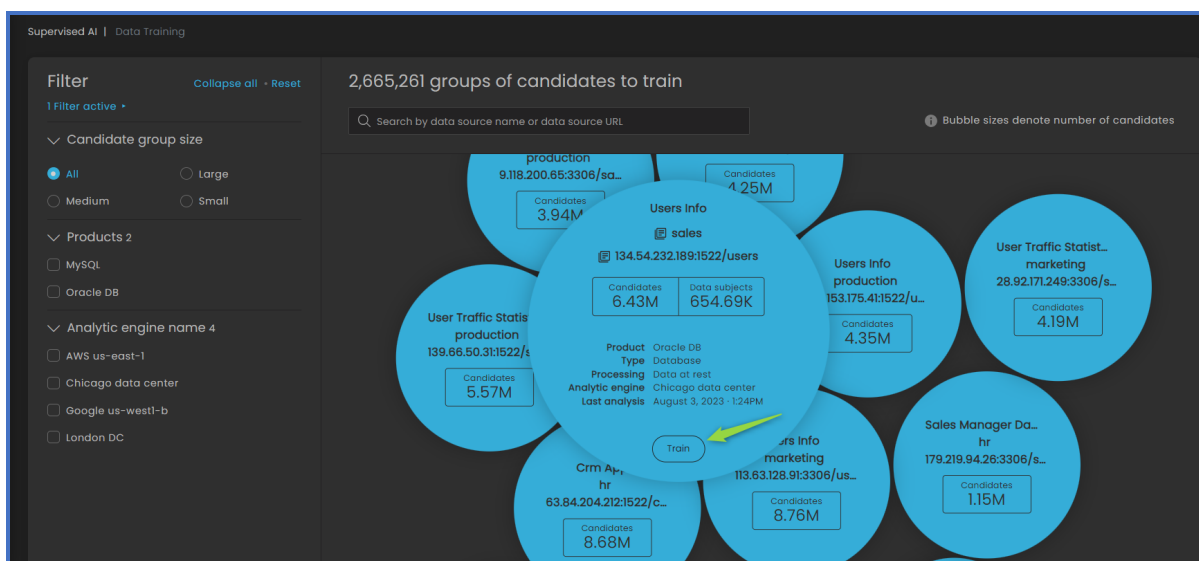


Figure 8: Candidate virtual view bubble

3. On the **Data training** page, you can explore the ability to edit the data fields and Supervised AI data elements.

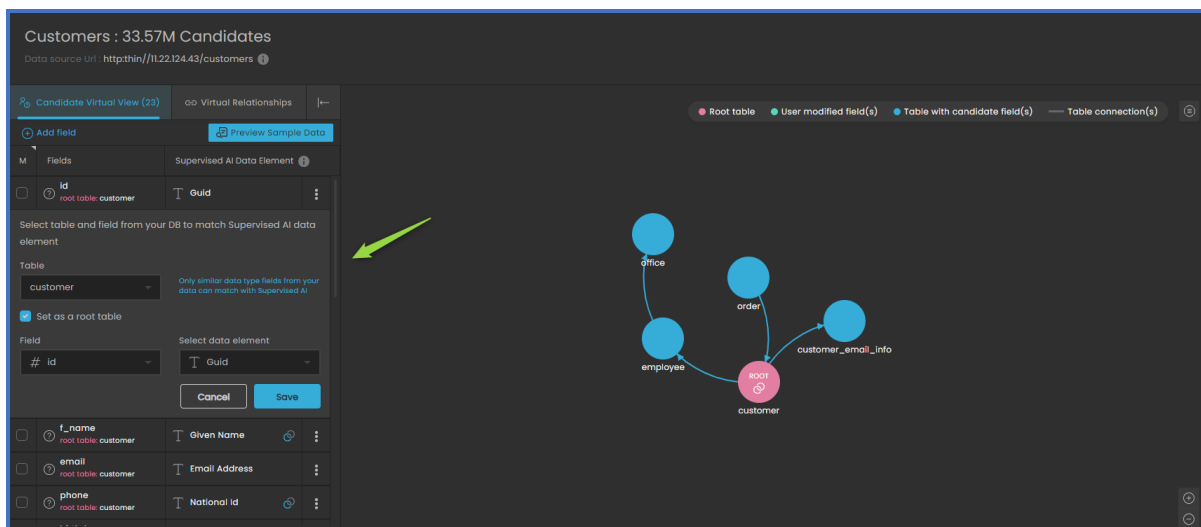


Figure 9: Data training page

4. Select one of the options in the bottom right corner of the page to either accept the virtual view (**Valid VV** button) or ignore it (**Ignore the VV** button). If needed, select the option to mark virtual view as root data asset.

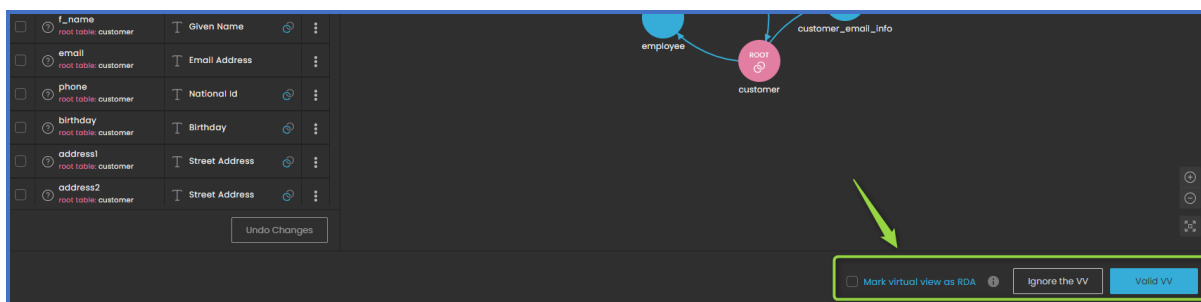


Figure 10: Virtual view menu

ADDING DATA ELEMENTS

To improve data recognition quality, you can add new or edit existing data elements.

1. Navigate to **Settings > Data Recognition > Data Elements Configuration**. On the **Data element configuration** page, click the **Add data element** button.

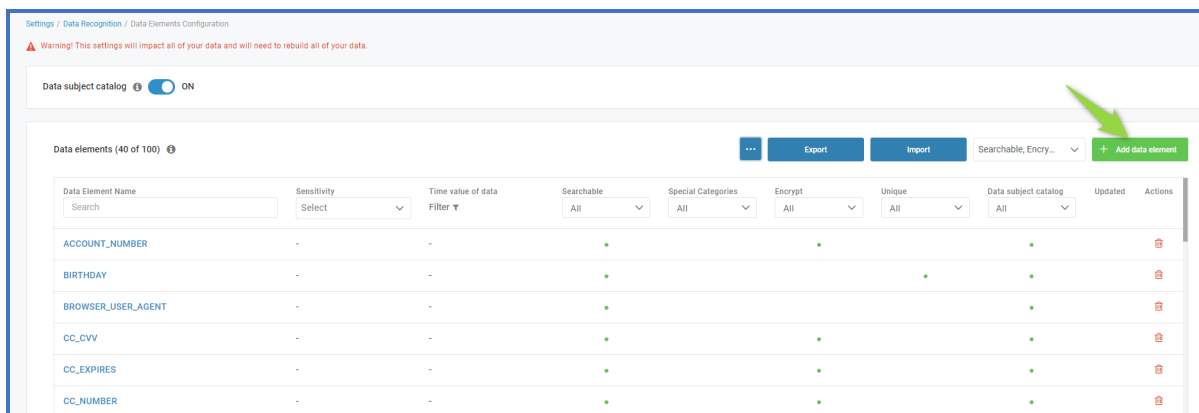


Figure 11: Adding a custom data element on the Data element configuration page

2. In the **Add/edit data element** popup window, fill the general information.

Table 12: General parameters of a data element

PARAMETER	DESCRIPTION
Name field (1)	Name of the data element defining the recognition rules for the Inventa analyzers.
Title field (2)	Title of the data element shown in other Inventa applications: Data Asset Manager, Personal Information Search, Supervised AI, Advanced Search, etc.
Sensitivity dropdown (3)	Level of confidence or sensitivity of the data element that is populated to other places in Inventa like data security dashboard, data subject profile, etc.
Time value of data parameter (4)	Time value of the data element defining how long the data is valid. This parameter is visualized as static information in other places in Inventa like data security dashboard, data subject profile, etc.
Description field (6)	Details of the data element. The description is shown when you hover over the data element name on the Data element configurations page.

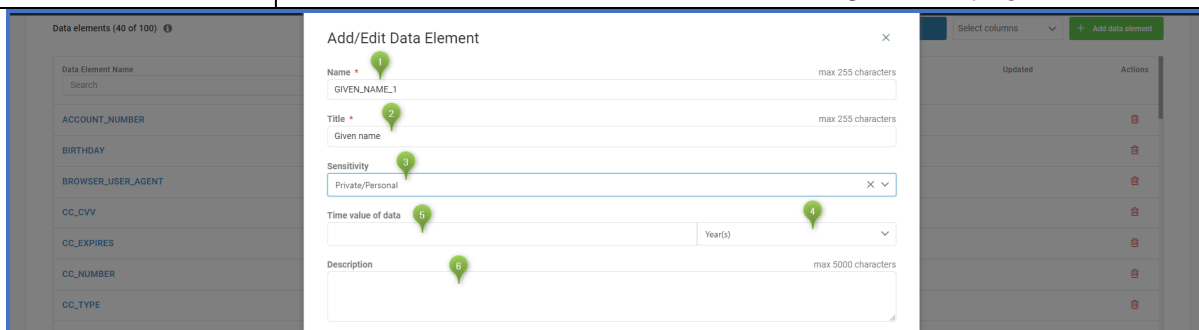


Figure 13: General settings of the data element in the Add/Edit Data Element popup window

3. Configure sensitive data type for the custom data element depending on the business use case (classification and cataloging options). See details on how to select a use case [above](#)

4. If necessary, configure the format for the discovered and stored values. From the **Preformatter** dropdown, select the formatting to be applied to values discovered by the

plugins. From the **Persistent formatter** dropdown, select the formatting to be applied to values when stored in the Inventa database.

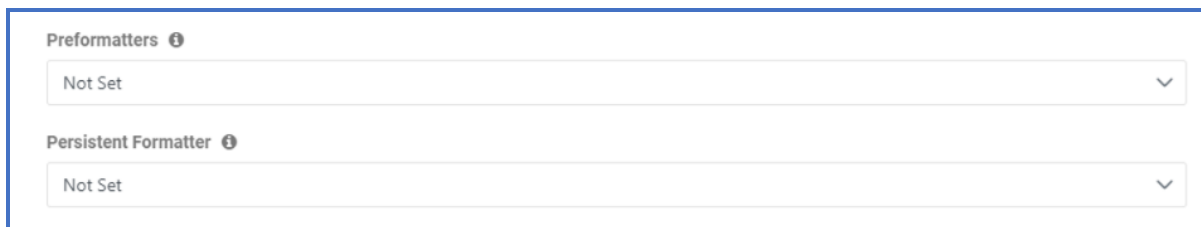


Figure 14: Configuring formatting for data element values

5. Select the **Language pack** for configuration and set the recognition rules – keywords, patterns, functions, and dictionaries for the data element. You can use the Inventa predefined keywords, patterns, functions and dictionaries and add your custom keywords, patterns and dictionary.

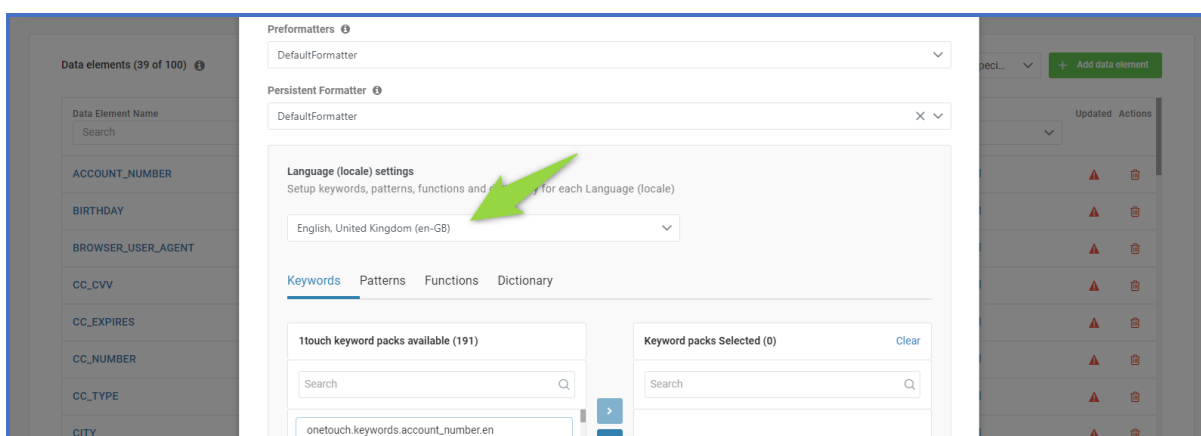


Figure 15: Selecting a language pattern for a data element

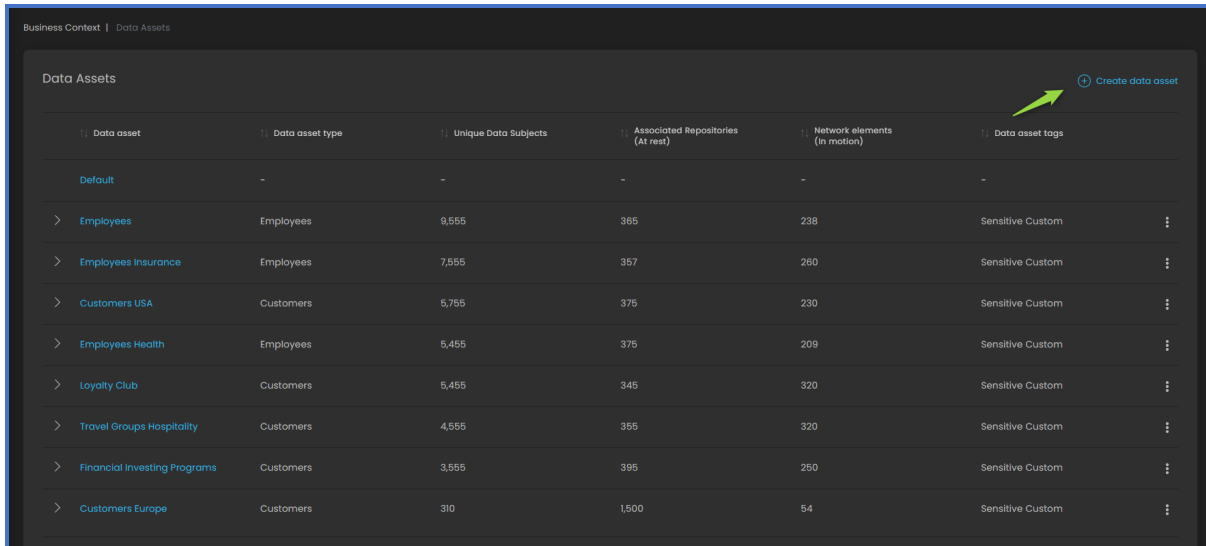
6. Once you set the recognition rules for all language packs, click the **Apply** button at the bottom of the popup. The system will close the popup and will add the new data element to the Data Elements pane.

7. Review the **Person identification constraints** section. A constraint is a data element (like National ID) or a combination of data elements (like Given Name + Vehicle Number) that defines a unique person within 1touch.io master catalog. It is used for the search for unique persons in the Data asset manager and for the root data asset validation.

ADDING DATA ASSETS

After the initial scans have been completed, you may find that the system has detected customer data from a couple of countries, for example, France and USA, states of California and New York. In this case, you can create corresponding data assets to map the new-found data.

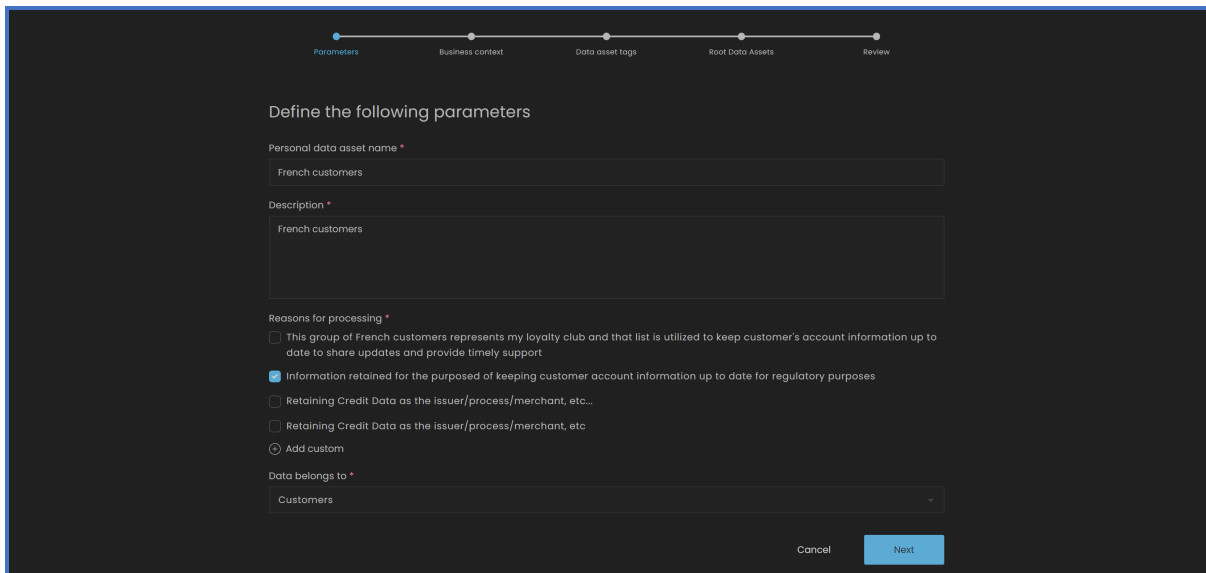
1. To create data asset, navigate to **Sidebar > Business Context > Data Asset** and click the **Create data asset** button.



Data asset	Data asset type	Unique Data Subjects	Associated Repositories (At rest)	Network elements (in motion)	Data asset tags
Default	-	-	-	-	-
> Employees	Employees	9,555	365	238	Sensitive Custom
> Employees Insurance	Employees	7,555	357	260	Sensitive Custom
> Customers USA	Customers	5,755	375	230	Sensitive Custom
> Employees Health	Employees	5,455	375	209	Sensitive Custom
> Loyalty Club	Customers	5,455	345	320	Sensitive Custom
> Travel Groups Hospitality	Customers	4,555	355	320	Sensitive Custom
> Financial Investing Programs	Customers	3,555	395	250	Sensitive Custom
> Customers Europe	Customers	310	1,500	54	Sensitive Custom

Figure 16: Data asset management page

2. Enter all the required details, add business context, assign tags, and select the desired root data assets.



Parameters Business context Data asset tags Root Data Assets Review

Define the following parameters

Personal data asset name *

French customers

Description *

French customers

Reasons for processing *

☐ This group of French customers represents my loyalty club and that list is utilized to keep customer's account information up to date to share updates and provide timely support

☒ Information retained for the purposed of keeping customer account information up to date for regulatory purposes

☐ Retaining Credit Data as the issuer/process/merchant, etc...

☐ Retaining Credit Data as the issuer/process/merchant, etc

☒ Add custom

Data belongs to *

Customers

Cancel Next

Figure 17: Data asset creation

3. After double-checking the preview, click the **Add data asset** button to save changes.

Reasons for processing:
Information retained for the purposed of keeping customer account information up to date for regulatory purposes

Data belongs to:
Customers

Business context

Personal data asset owner
Business unit:
Sales

Data asset tags
No tags were assigned to Data asset

Root Data Assets

Name	Appliance Name	RDA Records Inventorized	Last Analysis	Hostname	Source
France_Customer_Table	core	39.55K	Jul 24, 2024	10.192.191.71	DB: jdbc:sqlserver://10.192.191.71

Back Cancel Add Data Asset

Figure 18: Saving changes

4. You can create data assets for New York and California customers following the same steps. After initiating new scans in the Data Source Catalog, the system will start mapping RDAs onto the newly created data assets and updating data asset details and statistics on their respective pages.

DATA SUBJECTS

1. To view individual data subjects pertaining to a specific data asset, you can navigate to the **Files** section on the data asset's page and click on the number of files found.

Discovered data sources

Total: 11

o Total

@ Sensitive Data Insights

The Data Asset - NY_Partner Affiliate - is a centralized view for various data sources that store sensitive/personal information of your partner/3rd party organization. By identifying and grouping these data sources into a distinct asset, it becomes a holistic view of the data landscape and enable targeted actions for data protection, risk management, and compliance.

Note you have 4.92k files identified containing sensitive data. Please consider activating monitoring and protection policy on them (access controls and encryption). This will help in proactively addressing potential risks and ensuring ongoing compliance with data protection regulations.

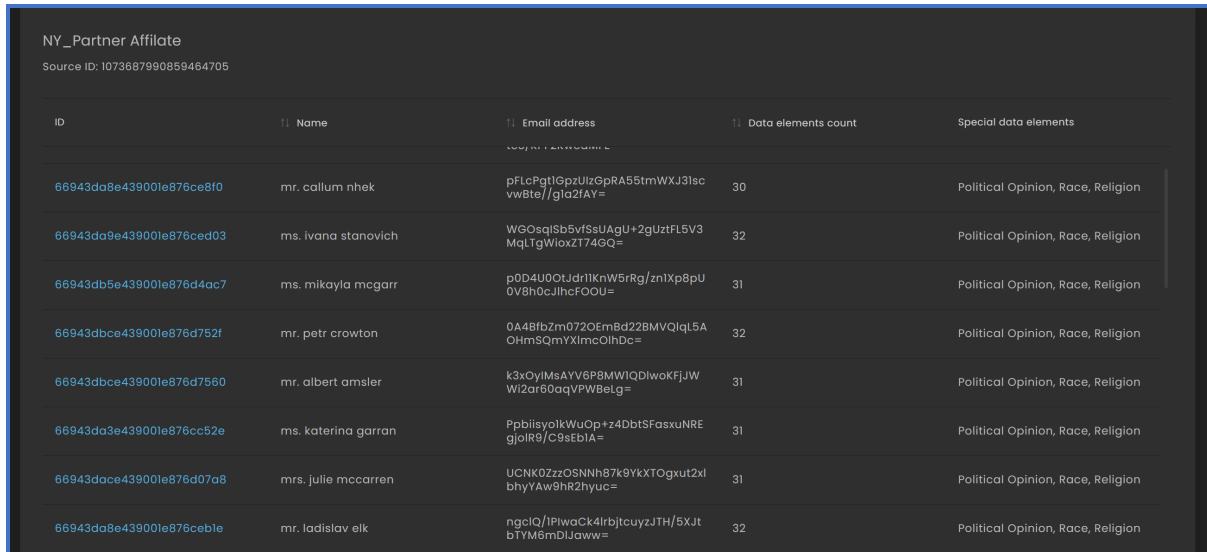
Data source type	Data sources
Central Storage	7
Database	4

Product name	Data sources
CIFS (SMB)	7
MariaDB	1
MongoDB	1
MS SQL Server	1
PostgreSQL	1

File type	Files found
txt	550
pdf	549
csv	546
doc	531
pptx	506
rtf	501
docx	501
xls	412

Figure 19: Files section

2. To review the data subject copies, click on the number in the corresponding column for one of the files and review the list of data subjects Inventa has found in it.



ID	Name	Email address	Data elements count	Special data elements
66943da8e439001e876ce8f0	mr. callum nhek	pFlcPgltGpzUizGpRA55tmWXJ3iscvw8te//gla2fAY=	30	Political Opinion, Race, Religion
66943da9e439001e876ced03	ms. ivana stanovich	WGosqISb5VfsUAgu+2gUztFL5V3MqLTgWioxZT74GQ=	32	Political Opinion, Race, Religion
66943db5e439001e876d4ac7	ms. mikayla mcgarr	p0D4U0OtJdr1KnW5rRg/zn1Xp8pU0V8h0cJlhcFOOU=	31	Political Opinion, Race, Religion
66943dbce439001e876d752f	mr. petr crowton	0A4BfbZm072OEmBd22BMVQlqL5A0HmSQmYXlmcOlhDc=	32	Political Opinion, Race, Religion
66943dbce439001e876d7560	mr. albert amsler	k3xOylMsAYV6P8MWlQDIwoKFJjWwI2ar60aqVPWBeLg=	31	Political Opinion, Race, Religion
66943da3e439001e876cc52e	ms. katerina garra	PpbliisYolkWuOp+z4DbtSFasxuNREgjoIR9/C9sEblA=	31	Political Opinion, Race, Religion
66943dace439001e876d07a8	mrs. julie mccarren	UCNK0ZzzOSNNh87k9YkXTogxuT2xlbhyYAw9hR2hyuc=	31	Political Opinion, Race, Religion
66943da8e439001e876ce8f0	mr. ladislav elk	ngclQ/1PlwaCk4lrjtcuyzJTH/5XJtbTYM6mDIJaww=	32	Political Opinion, Race, Religion

Figure 20: List of data subjects

3. Click on the data subject's unique ID to open its profile with all the information the system has found.

<

Figure 21: Data subject's info

4. You can export the data subject information location by clicking **Actions > Export for data portability**.

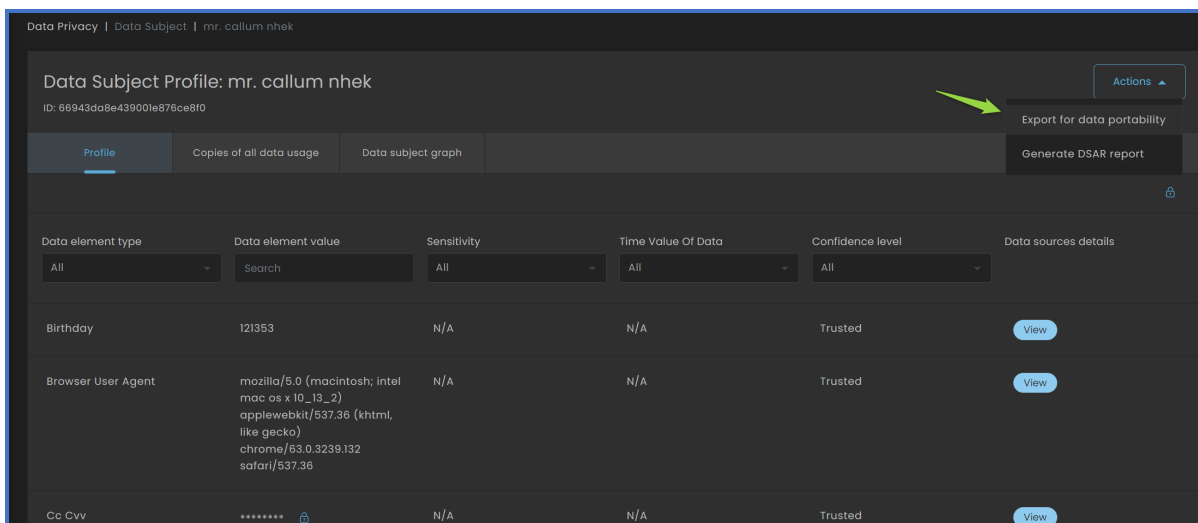


Figure 22: Export data subject information location

DATA SUBJECT SEARCH

1. If you want to find a specific data subject, navigate to **Data privacy > Data subject search** and fill in the desire fields, for example, name.

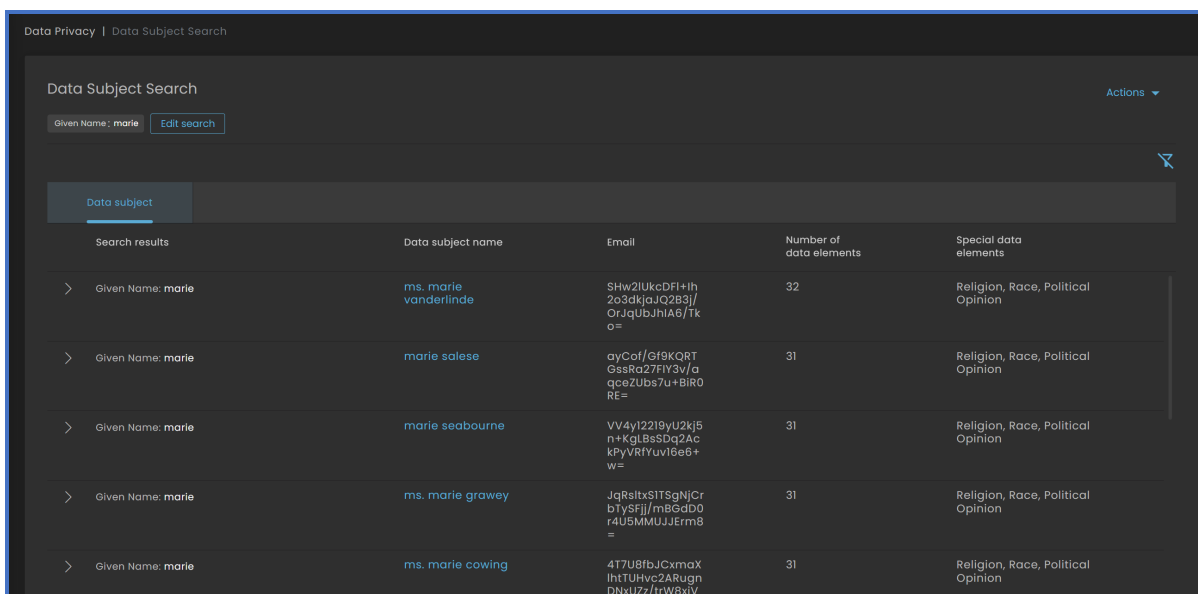


Figure 23: Data subject search results

2. Click on data subject's name to open its profile and see all the available information the system has found.

ATTACK SURFACE REDUCTION DASHBOARD

The Attack Surface Reduction Dashboard assists you in evaluating the number and age of files containing sensitive data in order to minimize the attack surface.

1. Navigate to **Data security** > **Attack surface reduction** and explore the information presented in its panes.



Figure 24: Attack surface reduction dashboard

2. Click on any of the columns to automatically launch a discovery & classification search with the corresponding parameters and find files of this kind.

Data Security | Discovery & Classification Search

Discovery & Classification Search

Search object (=) Files File created (=) 07/29/2021, 07/29/2022 Type of data (OR (=)) Sensitive personal Edit search

Filter Collapse all - Reset

0 Filters active

File type 12

Search

File source 1

Central Storage

File protocol 1

smb

Analytic engine 1

core

File name	Criticality	Document classification categories (tags)	File type
1touch-test-data-copies.zip	2.6	N/A	zip
3_XML_10kPII_single.xml	1.5	N/A	xml
CA_1000_1000_1.csv	2.5	N/A	csv
CA_1000_1000_2.xls	1.5	N/A	xls
CA_1000_100_1.csv	1.5	N/A	csv
CA_1000_100_2.xls	1.5	N/A	xls
CA_1000_101_1.csv	1.5	N/A	csv
CA_1000_101_2.xls	1.5	N/A	xls
CA_1000_102_1.csv	1.5	N/A	csv

Figure 25: Search results

3. Click on the file title to open its page and see all the available information the system has found.

The screenshot displays the 'File' page for a CSV file named 'CA_1000_1000_1.csv'. The interface is divided into two main sections: file metadata and data classification insights.

File Metadata:

- File type: csv
- File path: smb://10.192.191.70/ccpa/
- File owner: N/A
- Time created: May 09, 2022, 09:53:42 AM
- Last time modified: May 09, 2022, 09:53:42 AM
- Data categories: Contact information, Demographic information, Digital identification, Financial information, Government issued IDs [Show more](#)
- Special data elements: Political Opinion, Race, Religion
- Regulations: PCI DSS
- Data source tags: Credit Card Holders, Default, NY_Partner Affiliate
- Data assets: Default

Data Classification:

Sensitivity Data Insights
Please make sure that all the current copies of personal data are really needed.

Count	Description
32	Data elements types
N/A	Document classification categories (tags) types
11	Data categories types
3	Special data elements types
1	Regulations types

Data elements table:

Data elements	Sensitivity	Time value of data
Birthdate	N/A	N/A
Browser User Agent	N/A	N/A
Cc Cvv	N/A	N/A

Figure 26: File page

DISCOVERY & CLASSIFICATION SEARCH

The **Discovery & Classification Search** is a part of the search tool that allows you to query the structured data sources based on the classified sensitive data. You can use it to search for data sources and files.

1. Navigate to **Data security > Discovery & classification search**. Select the desired type of search for either files or data sources and fill in the required fields.

The screenshot shows the 'Discovery & Classification Search' interface. It features a search criteria table with columns for Search criteria, Operator, and Value. Below the table are sections for Granularity, Data source attributes, and Data attributes, each with input fields and operators. At the bottom, there are buttons for 'Save to favorites' and 'Search'.

Search criteria	Operator	Value
Search object	=	Data sources

Granularity

Report granularity level	=	Data source
--------------------------	---	-------------

Data source attributes

Data source type	OR (=)	
Data source URL	=	
Product	OR (=)	
Tags	OR (=)	
Analytic engine names	OR (=)	
Last analysis time	=	

Data attributes

Data elements	OR (=)	
Special categories	OR (=)	

Buttons: [Save to favorites](#) [Search](#)

Figure 27: Discovery & classification search

2. Click on a specific search result to see more details.

3. Explore the page with more detailed information about this data.

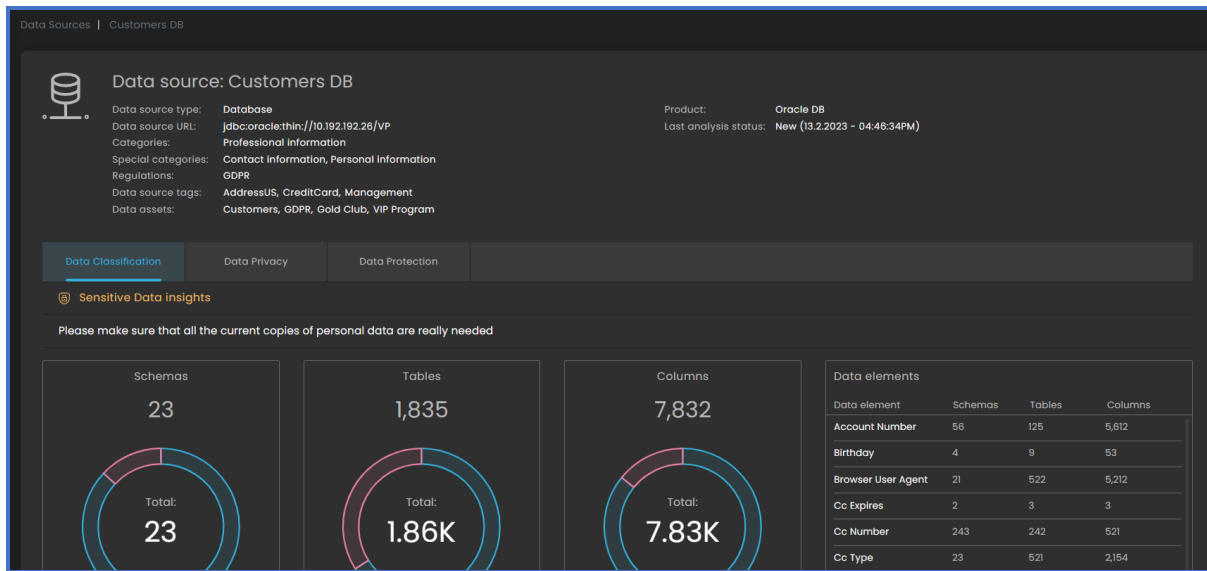


Figure 28: Data source info

1touch.io uniquely uses network analytics to help your company discover both sensitive data and its use, even the data you didn't know existed.

1touch.io's Inventa is a data privacy platform with unprecedented data lineage techniques for data discovery and classification. Inventa gives companies complete visibility into their unknown usage of customer data by automating the discovery process and providing them with a comprehensive, accurate, and up-to-date master catalog. This visibility enables you to easily meet regulatory, compliance, and security requirements.

