# USER GUIDE

## Data Governance

Release 10.0-MP3 (October 2020)

StiboSystems

MASTER DATA MANAGEMENT

# Table of Contents

# Data Governance

To properly manage data, an MDM system provides support for information governance functionality such as policy evaluation, creation, and collaboration, as well as policy change management and impact analysis.

To meet these needs, the following functionality is available for use in STEP:

- Policies that monitor data and data streams
- Notifications in the UI and via email that inform users of deviations in the data quality
- Metrics to define data quality

## Initial Workbench Configuration

To set up Data Policies, on a CMDM-supported system, some objects must be created and some permissions must be added for the associated users. For information on configuring the workbench, see the Configuring Workbench to Create Data Policies section in the **Data Policies** topic.

## Configuring a Data Policy

For general guidelines on Data Policies, see the **Data Policies** topic in this documentation.

To begin, a completeness, value, or function metric must be configured. For information on configuring a metric, see the **Completeness Metrics** topic, the **Value Metrics** topic, or the **Business Function Metrics** topic of the **System Setup** documentation.

# Data Policies

Data policies allow users like data stewards to define thresholds and monitor breaches and deviations in the quality of the master data as well as incoming data streams. Prior to configuring these components, some system setup object types must be created.

## Configuring Workbench for Data Policies

To create data policies, the following system object types must be created, and a Users & Groups setting must be specified. For instructions on how to create new object types, see the **Creating an Object Type** topic in the **System Setup / Super User Guide** documentation.

---

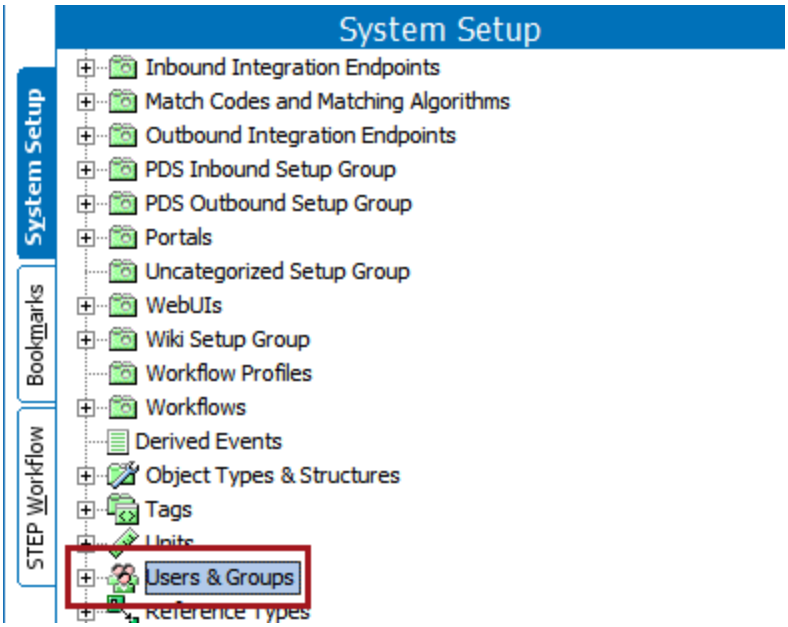
**Note:** These types may already exist in the system, but parents may need to be added.

---

### Selecting a Data Policies Default Context

To properly monitor data policies, a default context must be selected. To set this default:

1. In the workbench, navigate to the System Setup tab.
2. Select the Users & Groups node.

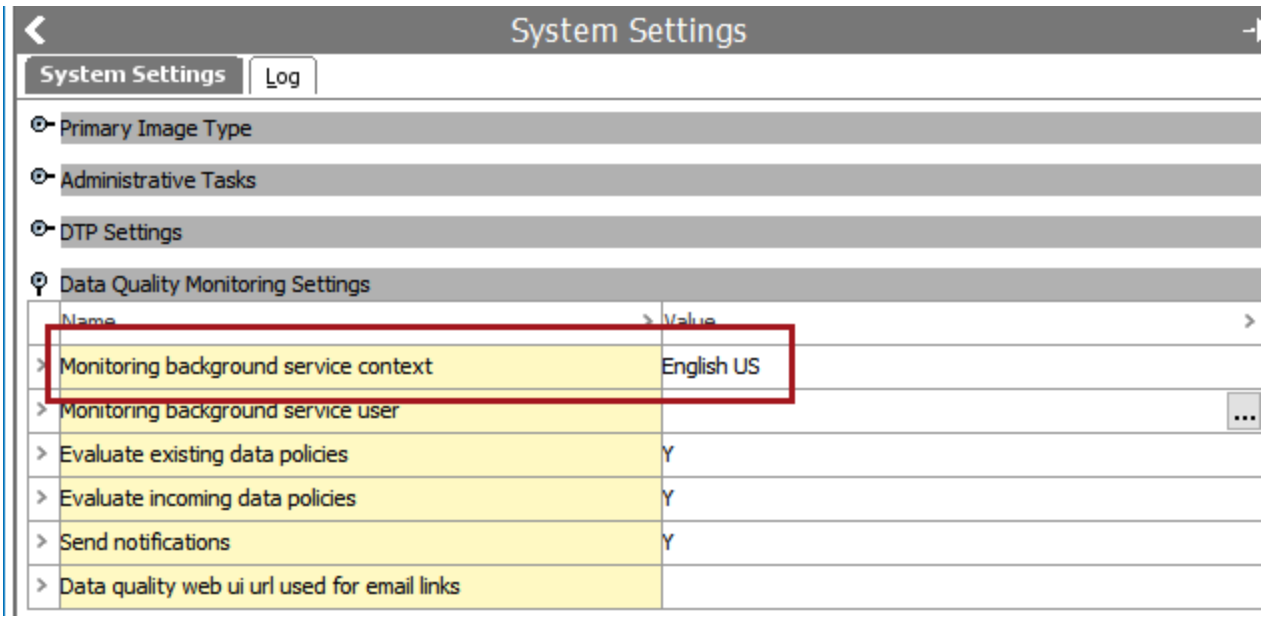



3. From the Users & Groups node, navigate to the Data Quality Monitoring Settings section. Under this section, select a context under the 'Monitoring background service context' as shown below.

---

**Important:** This section only displays if the cmdm-monitoring component has been activated. Contact Stibo Systems with assistance, if needed.

---




### System Setup Object Type

- Metrics root, containing the following types:
  - Business Function Type
  - Completeness Metric
  - Entity data quality metric
- Policy Type root, with the following types:
  - Entity Dataset Definition
  - Existing Entity Data Quality Policy
  - Incoming Entity Data Quality Policy

After adding these types, dataset definitions, metrics, and policies may be created.

## Creating Data Policies

A data policy comprises the following three elements:

- a Metric, such as a completeness metric or a function metric, that defines how to score each record.
- a Dataset Definition that defines which data to monitor.
- thresholds that defines when users must be notified.

To create meaningful policies, the following configurations are required:

1. Metrics must be created through the Workbench. For information on these metrics, see the **Completeness Metrics** topic, the **Value Metrics** topic, or the **Business Function Metrics** topic in the **System Setup** documentation.
2. Dataset definitions must be created through the Workbench. For more information, see the **Creating a Dataset Definition** topic of this guide.

3. a Web UI must be configured to manage the policies. For more information, see the **Web UI Configurations for Policies** topic of this guide.

---

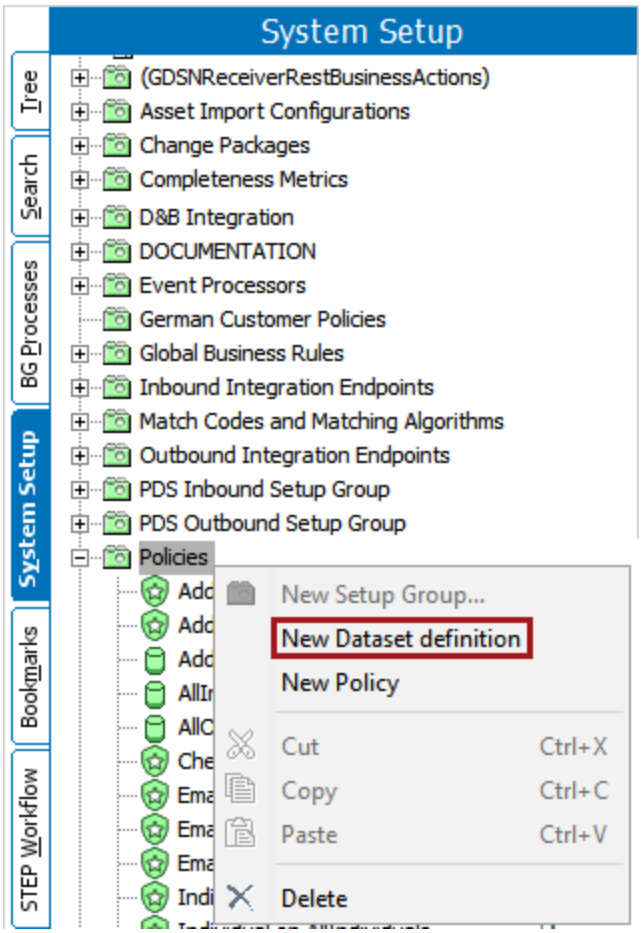
**Note:** This documentation will focus on the workbench elements of the data policies configuration. For information on setting up the Web UI screens for data policies, see the **Web UI Configurations for Policies** topic in this documentation.

---

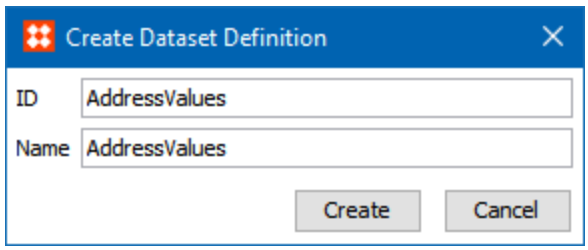
# Creating a Dataset Definition

The dataset definition specifies the STEP data to which a policy may apply. Follow the steps below to create an entity dataset definition.

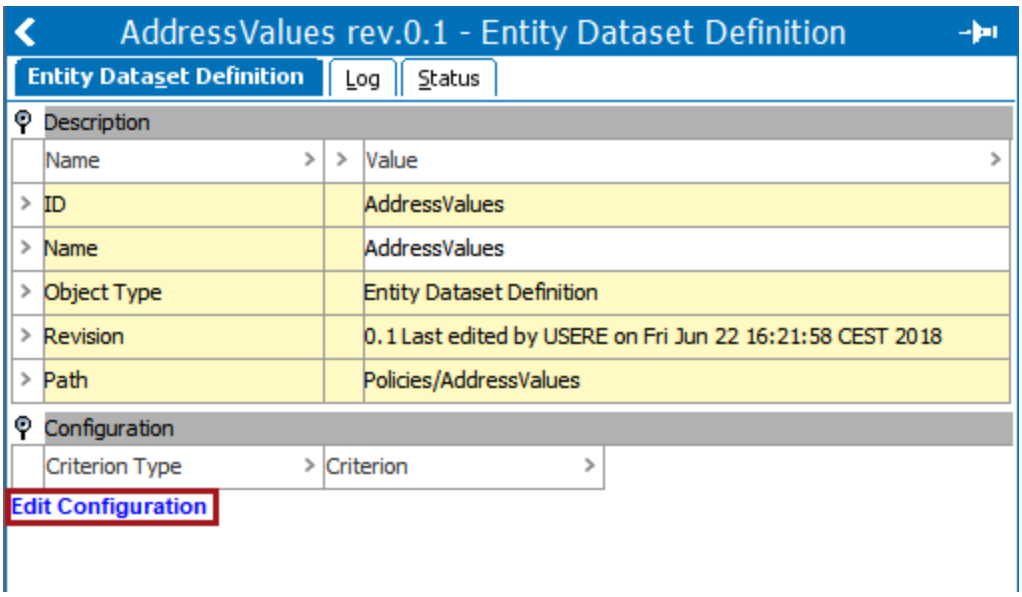
1. Navigate to System Setup > Policies. Right-click on the Policies folder and select 'New Dataset definition.'



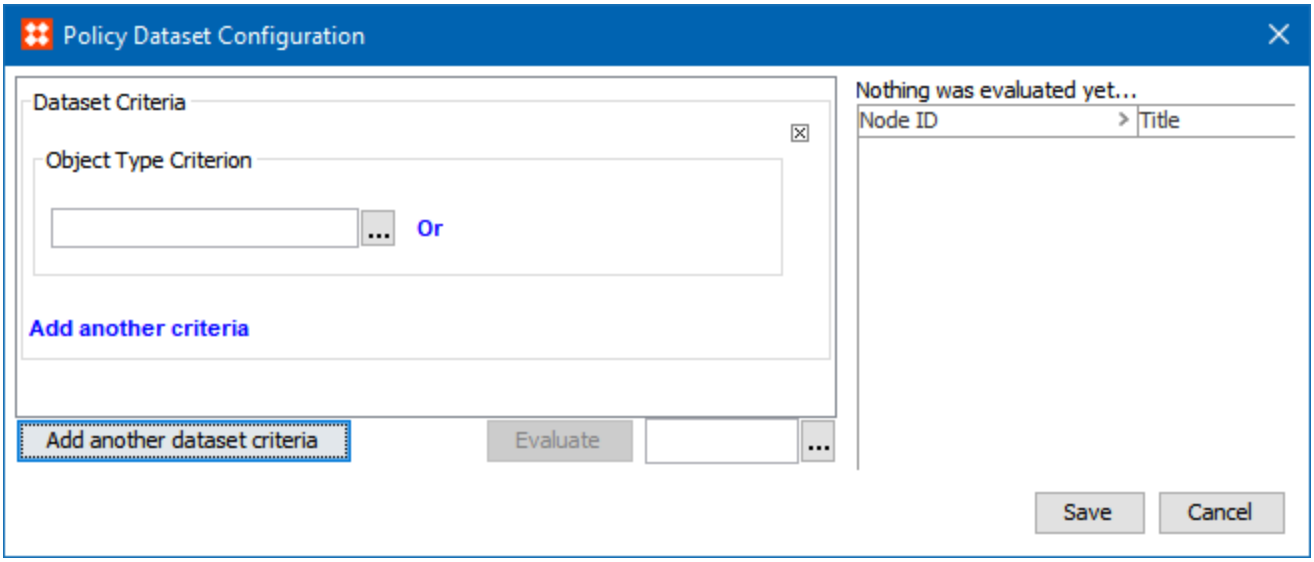

2. A 'Create Dataset Definition' dialog will display. Enter an ID and Name for the dataset definition, and click the Create button to display the Dataset Definition.



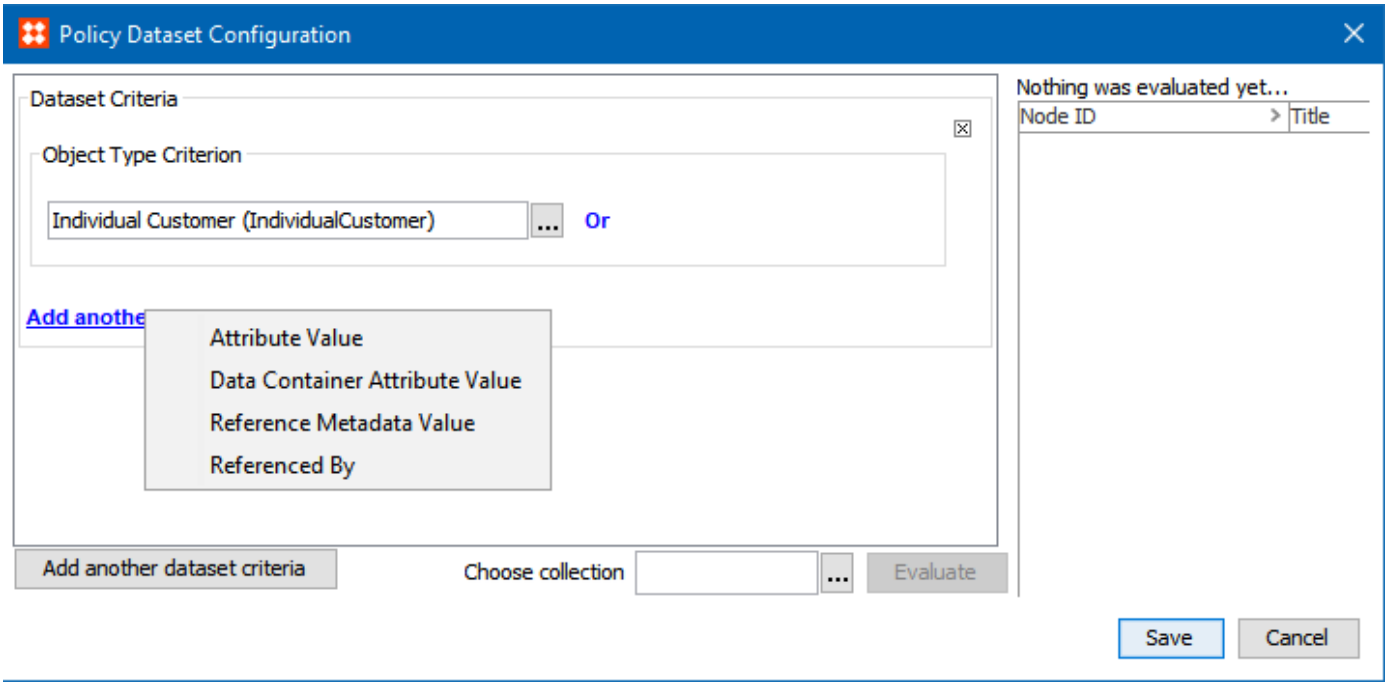

3. Click 'Edit Configuration' to add parameters to the dataset definition and open the Policy Dataset Configuration dialog.



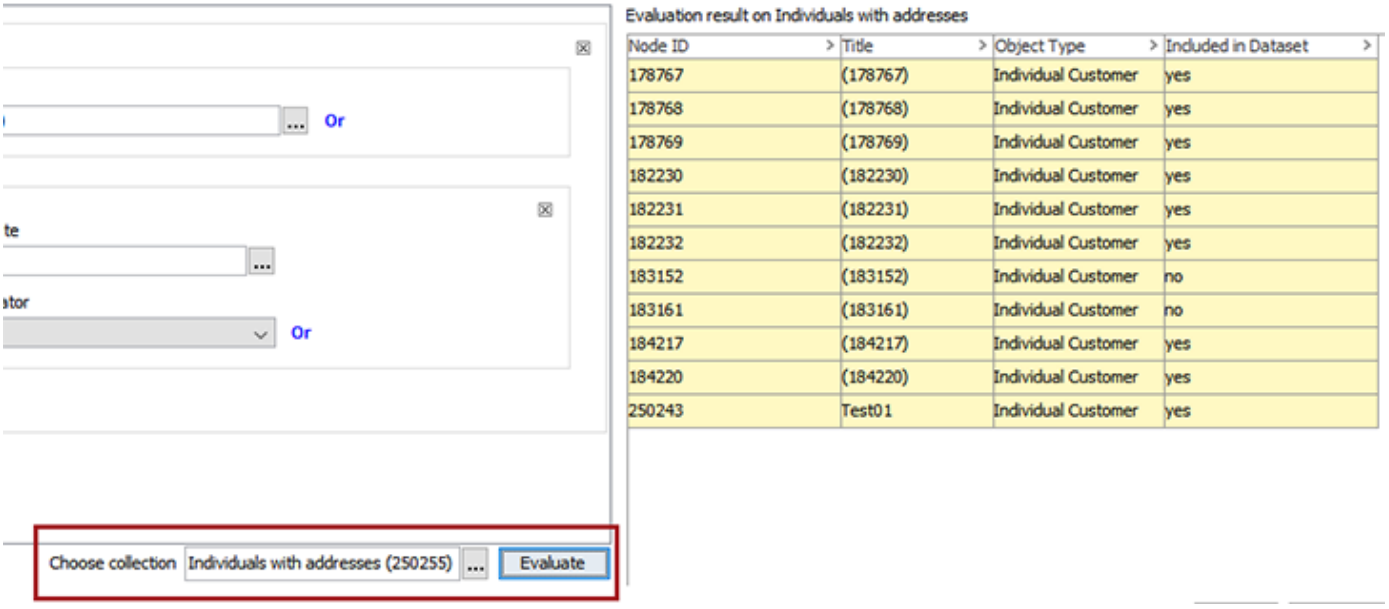

4. Click the 'Add another dataset criteria' button to display the Dataset Criteria group.




5. Click the ellipsis button (...) for the Object Type Criterion field. From the objectTypeSelectorField dialog, select the desired object type.
6. If another criterion should be evaluated, click the 'Add another criteria' link, select an option from the dropdown, and configure it. Select the desired criterion and configure it. For descriptions of all these criteria, see the **Explanations of Criteria** topic of this documentation.




7. Once all criteria have been configured, click Save to close the Policy Dataset Configuration dialog.
8. The Dataset Configuration allows for bulk evaluation of the dataset definition against any collection of nodes. To evaluate your dataset, click the ellipsis button (...) to choose a collection next to the Evaluate button in the bottom of the dialog.
9. Once a collection is selected, press 'Evaluate.' The first 100 nodes of that collection will be evaluated against the dataset, and results presented in the right hand side of the dataset editor.




| Node ID | Title | Object Type | Included in Dataset |
|---|---|---|---|
| 178767 | (178767) | Individual Customer | yes |
| 178768 | (178768) | Individual Customer | yes |
| 178769 | (178769) | Individual Customer | yes |
| 182230 | (182230) | Individual Customer | yes |
| 182231 | (182231) | Individual Customer | yes |
| 182232 | (182232) | Individual Customer | yes |
| 183152 | (183152) | Individual Customer | no |
| 183161 | (183161) | Individual Customer | no |
| 184217 | (184217) | Individual Customer | yes |
| 184220 | (184220) | Individual Customer | yes |
| 250243 | Test01 | Individual Customer | yes |

10. Click 'Save' to save the dataset definition.

Policies can now be applied to the newly created Dataset Definition. To add this dataset definition to a policy as described in the **Creating a Data Policy** topic in this documentation.

# Explanations of Criteria

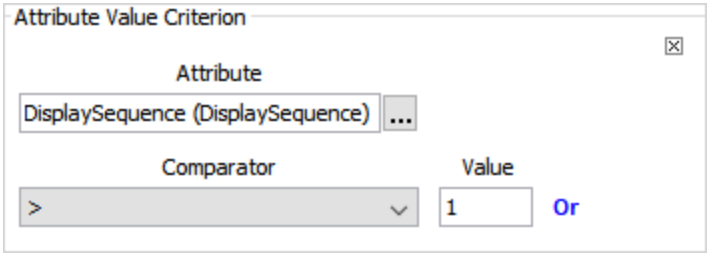
This section details the dataset configuration criteria. In the following sections, the 'Or' or 'And' buttons will add another criteria of the same object type criterion.

## Object Type Criterion

The object type criterion defines the type of object(s) to which the policy applies.


## Attribute Value Criterion

The attribute value criterion will allow the policy to assess attributes and their values. Then, using the comparator selector, users can select how to assess the value that is entered. The options for evaluating the value with the comparator are in the table below the example image.




| Comparator Symbol | Definition | Example of Use |
|---|---|---|
| > | Greater than | 'DisplaySequence' > 1; this is true if the associated object has a display sequence of any value higher than 1. |
| >= | Greater than OR equal to | 'DisplaySequence' >= 1; this is true if the associated object has a display sequence of 1 or greater. |
| < | Less Than | 'DisplaySequence' < 1; this is true if the associated object has a display sequence of any value lower than 1. |
| <= | Less than OR equal to | 'DisplaySequence' <= 1; this is true if the associated object has a display sequence of 1 or lower. |
| = | Equal to | 'City' = Annapolis; this is true whenever an object has 'Annapolis' in the City field. |

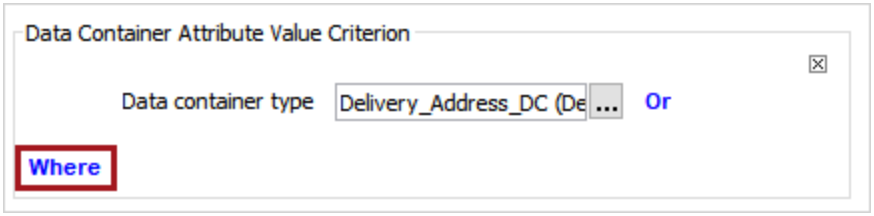
| Comparator Symbol | Definition | Example of Use |
|---|---|---|
| != | Not equal to | 'City' != Annapolis; this is true whenever an object has any value BUT 'Annapolis' in the City field. |
| has value | - | 'City' has value; this is true when there is any value in the City field. |
| has no value | - | 'City' has no value; this is true when there is no value in the City field. |
| starts with | - | 'City' starts with 'A'; this is true whenever a City field has a value starting with 'A' |

Select a comparator and then supply the value that should be used for evaluation.

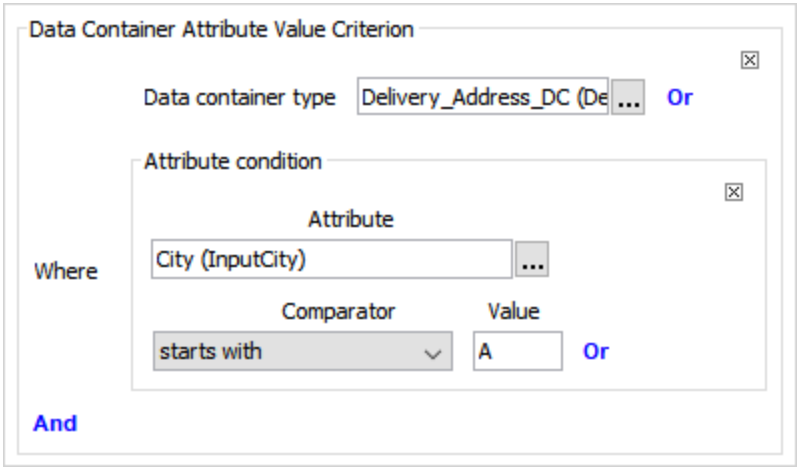
**Note:** The 'has no value' and 'has value' do not show a 'Value' field.

## Data Container Attribute Value Criterion

This criterion will isolate a data container on entity records and then evaluate an attribute under that data container. First, the data container type is specified. If only the data container type is specified, then any entity with at least one instance of this data container type will be returned.

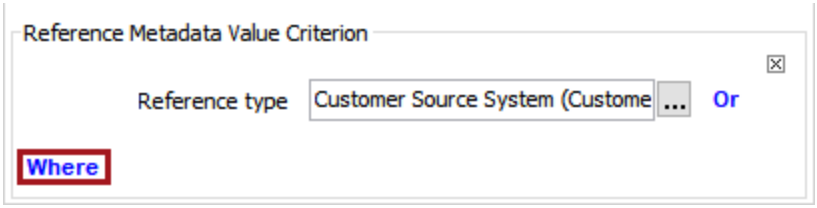



However, if the 'Where' addition is specified, then any entity with the data container type that has the specified criterion will be returned. Set the attribute, comparator, and value for evaluation. In the following example, any customer with a least one delivery address where the city starts with 'A.'

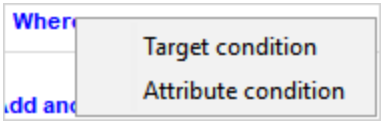


## Reference Metadata Value Criterion

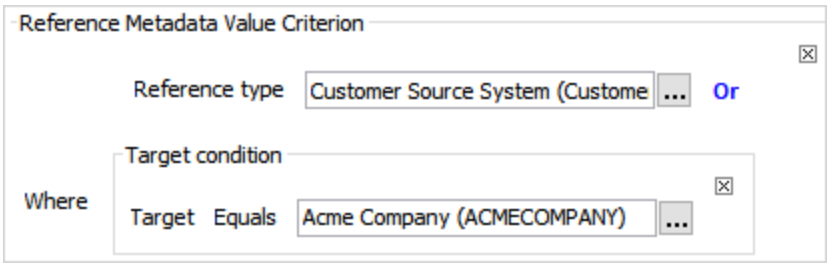
The reference metadata value criterion allows for specifying a reference type. Select a reference type, and then, select the 'Where' link to specify conditions for evaluation.



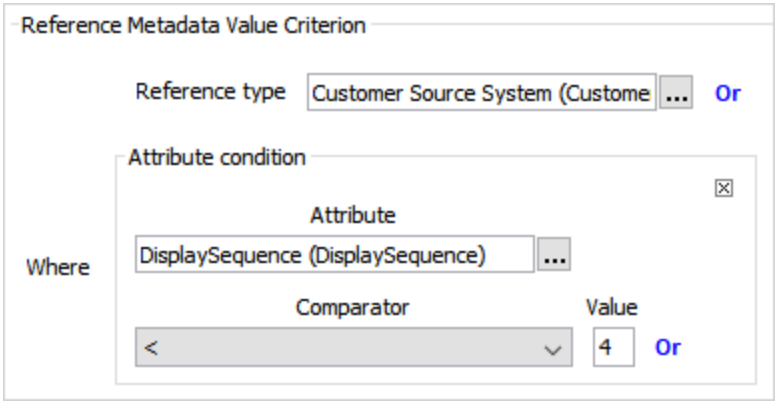

Choose whether this is for a target condition or an attribute condition.




- With the 'Target condition' selection, specify an object that, if it is the target, will satisfy this condition. This option works by limiting data that originates in the referenced target, in this case the ACME Company source system. In other words, using this example, users will only want to see data with a reference to a source system where the source system is ACME.

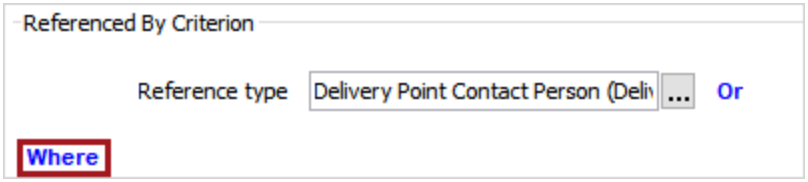



- With the Attribute condition, select 'Where' to set the attribute, comparator, and value for evaluation.

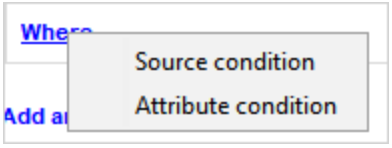


## Referenced By Criterion

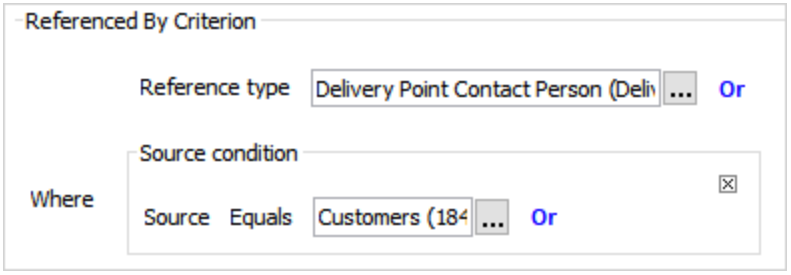
The referenced by criterion compares the source object of a reference. Select a reference type, and then, select the 'Where' link to specify conditions for comparing.



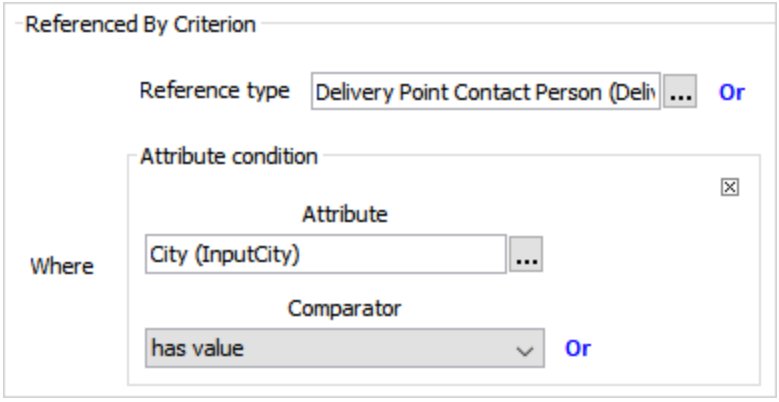

Choose whether this is for a source condition or an attribute condition.




- With the 'Source condition' selection, specify an object that, if it is the source object that is referencing the current node, will satisfy this condition.




- With the Attribute condition, select 'Where' to set the attribute, comparator, and value for evaluation.



# Creating a Data Policy

Within a data policy, metrics are used to measure data quality and express this evaluation numerically. There are three types of metrics:

- Completeness Metrics - Metrics that are based on groups of attributes. For information on creating and editing a completeness metric, see the **Completeness Metrics** topic in the **System Setup** documentation
- Value Metrics - Metrics used to take an attribute value and transform these results into a score. For more information, see **Value Metrics** topic of the **System Setup** documentation
- Function Metrics - Metrics based on the result of a an executed JavaScript function. For information on setting up a function metric, see the **Business Function Metrics** topic of the **System Setup** documentation

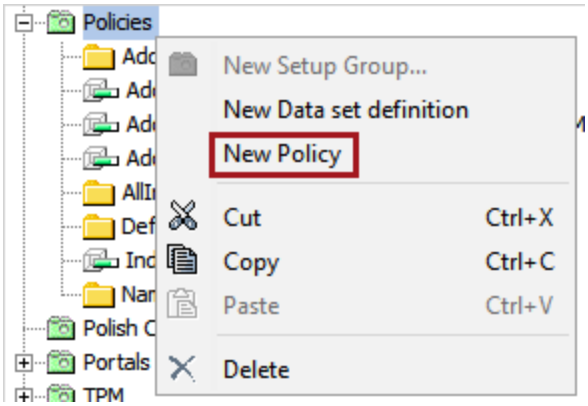
---

**Note:** This topic will provide sample metrics. These samples should only be considered as guidance for explanation and not a codified recommended practice.

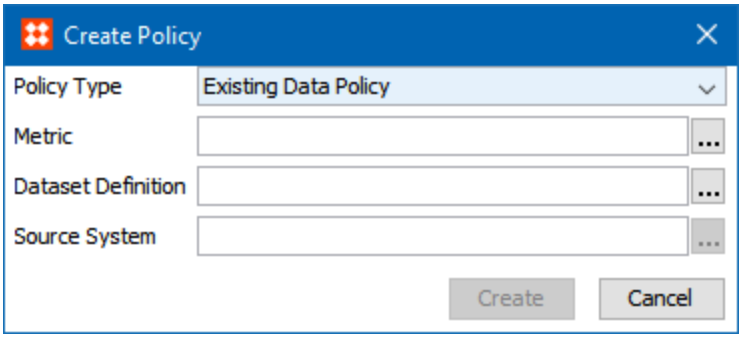
---

With at least one metric configured and one dataset definition configured, policies can be created to ensure that system data stays within the thresholds.

1. On System Setup, right-click the Policies node. Select 'New Policy' from this menu.




2. On the 'Create Policy' dialog supply information for the following parameters:

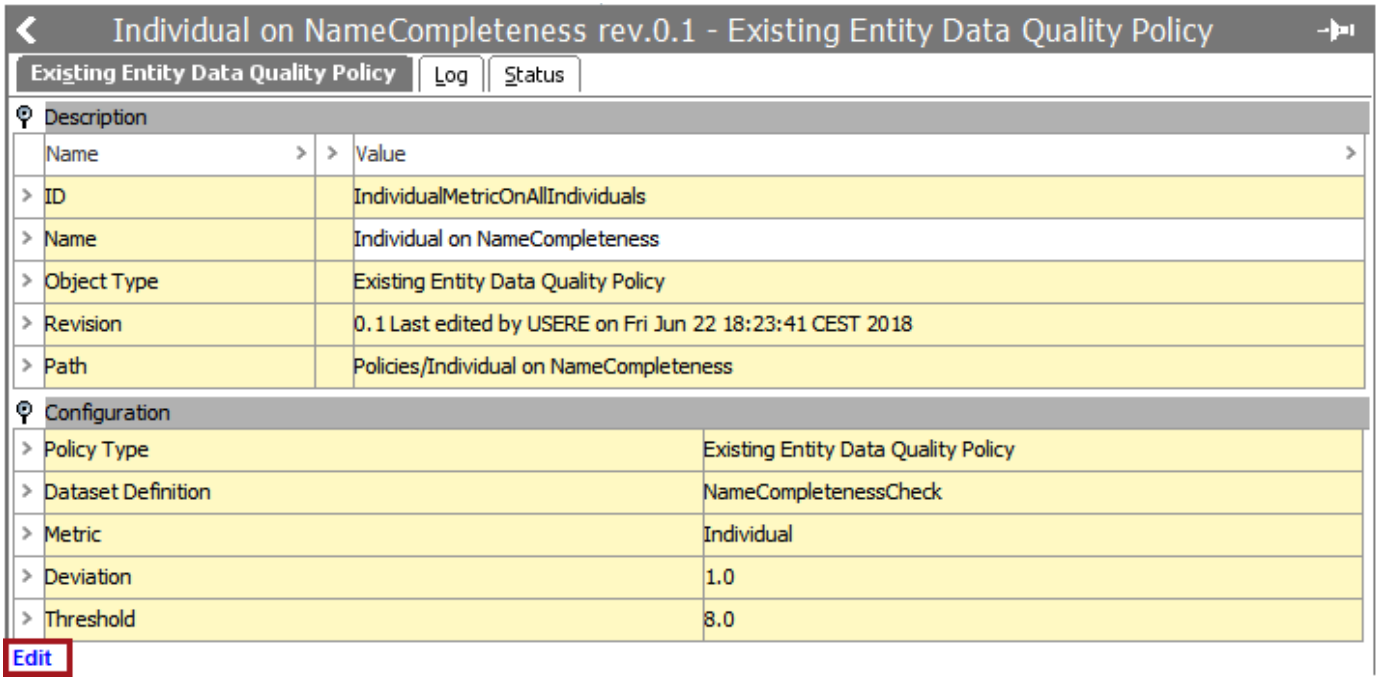



- For **Policy Type**, select whether this policy refers to for 'Existing Data' or 'Incoming Data.' Existing Data Policy is evaluating data that exists in STEP each night. Incoming Data Policy is evaluating only the incoming data from an Inbound Integration Endpoint of the Merge Golden Record type,

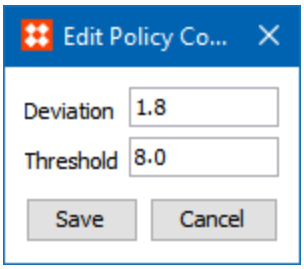
allowing early warnings if the source system starts sending bad data.

- Select the metric that was set up earlier in this topic.
- Select the Dataset Definition that was defined in the **Creating a Dataset Definition** topic of this documentation.
- If using the 'Incoming Data Policy' policy type, then a source system must be specified to limit the policy's focus.

3. Click the 'Create' button to confirm the configuration.
4. A policy will be created. Note that the deviation is set to 1.0 and the threshold is set to a default of 8.0. Click 'Edit' to change these values.

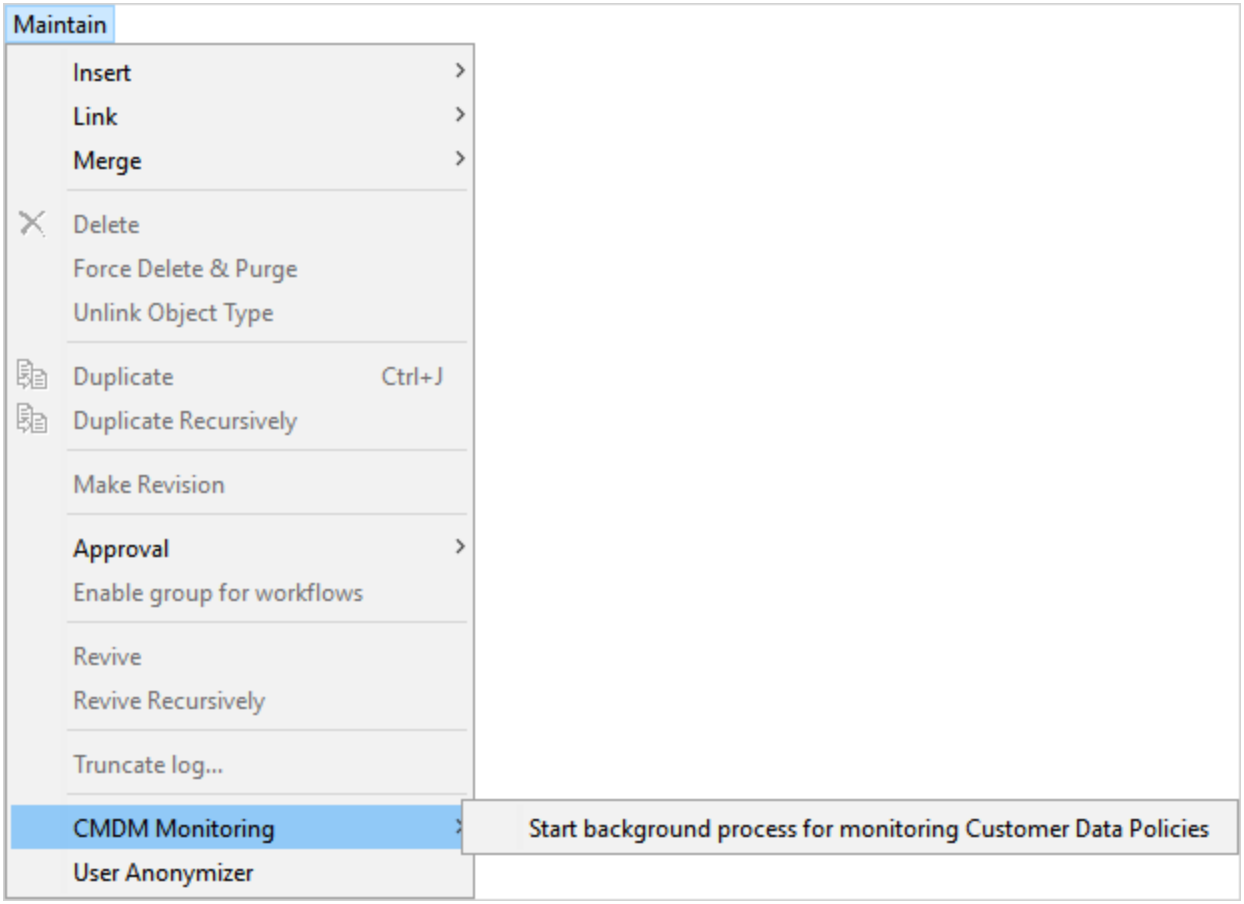



5. Enter the desired value with a decimal point format. Click 'Save' to change these values.




## Manually Starting a CMDM Monitoring Process

Should you wish to test a CMDM Monitoring Policy or to start a process as needed, the Maintain menu contains an option for starting the background process to monitor.




Once this option is selected, a background process will start. Once completed, it will show any notifications about this policy.

```
1  Logged on to server doc-trunk as User E
2  Using context English US and workspace Main
3  Policy evaluation started (reports every 10 minutes) (Thu Aug 09 16:24:06 EDT 2018)
4  Reading objects with the next Object Types: Customers (Customers) (Thu Aug 09 16
5  Evaluated policy "Address Validation on All Customers". Score: 6.1. Number of proces
6  Baseline and score updated on policy Address Validation on All Customers (Validate
7  Policy Address Validation on All Customers (ValidateIfCity) scores history cleanup: 0
8  Policy evaluation finished (Thu Aug 09 16:24:08 EDT 2018)
9  User notification started
10 1 policies found to notify about
11 0 notifications prepared to send
12 Sent 0 of 0 notifications. (Thu Aug 09 16:24:08 EDT 2018)
```

# Web UI Configurations for Policies

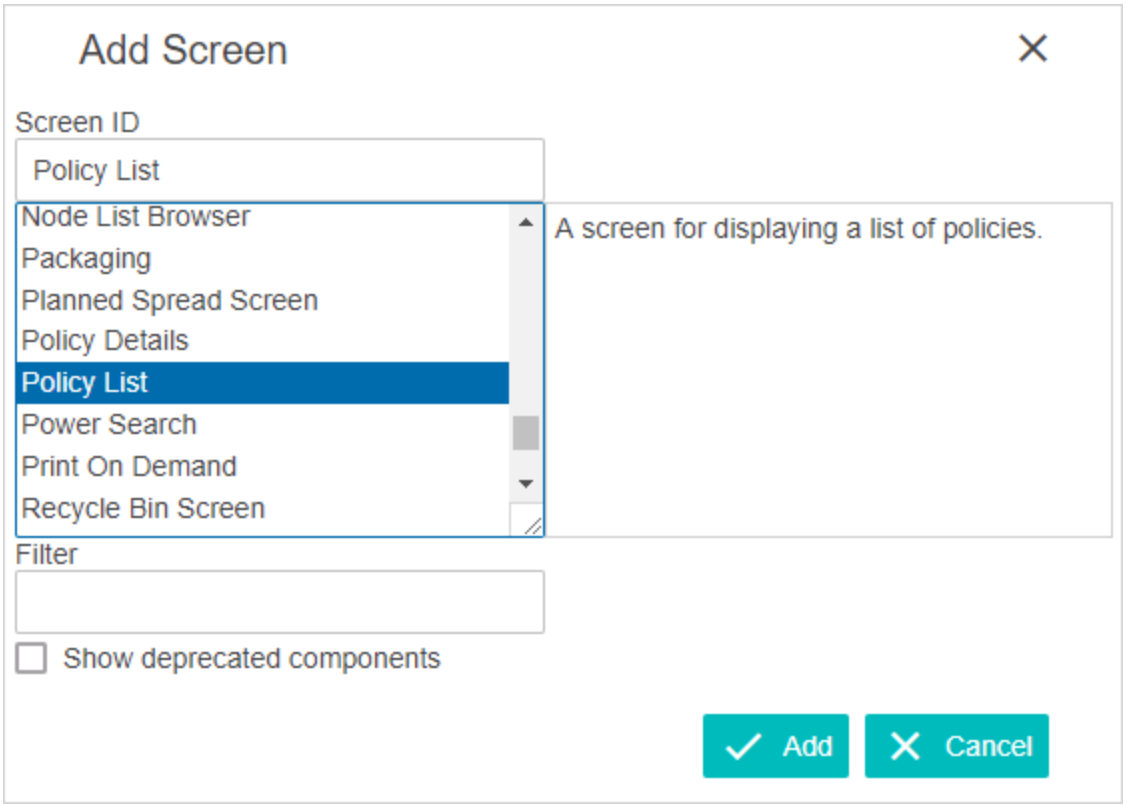
Data Quality Monitoring Policies are leveraged in the Web UI where they can be viewed and actioned on by a data steward.

To use the data policies within the Web UI, there are some minor pre-configurations required. First, see the **Data Policies** topic of the **Data Governance** documentation for all the configuration of data policies. From the Web UI, two screens will need to be created and a navigation component will need to be added. Finally, before using this topic for configuring the data policy Web UI, users should be familiar with the process of building out the Web UI using the Web UI Designer. For more information on configuring the Web UI, see the **Designer Mode Basics** topic in the **Web UI Getting Started** documentation.

## Configuring the Web UI

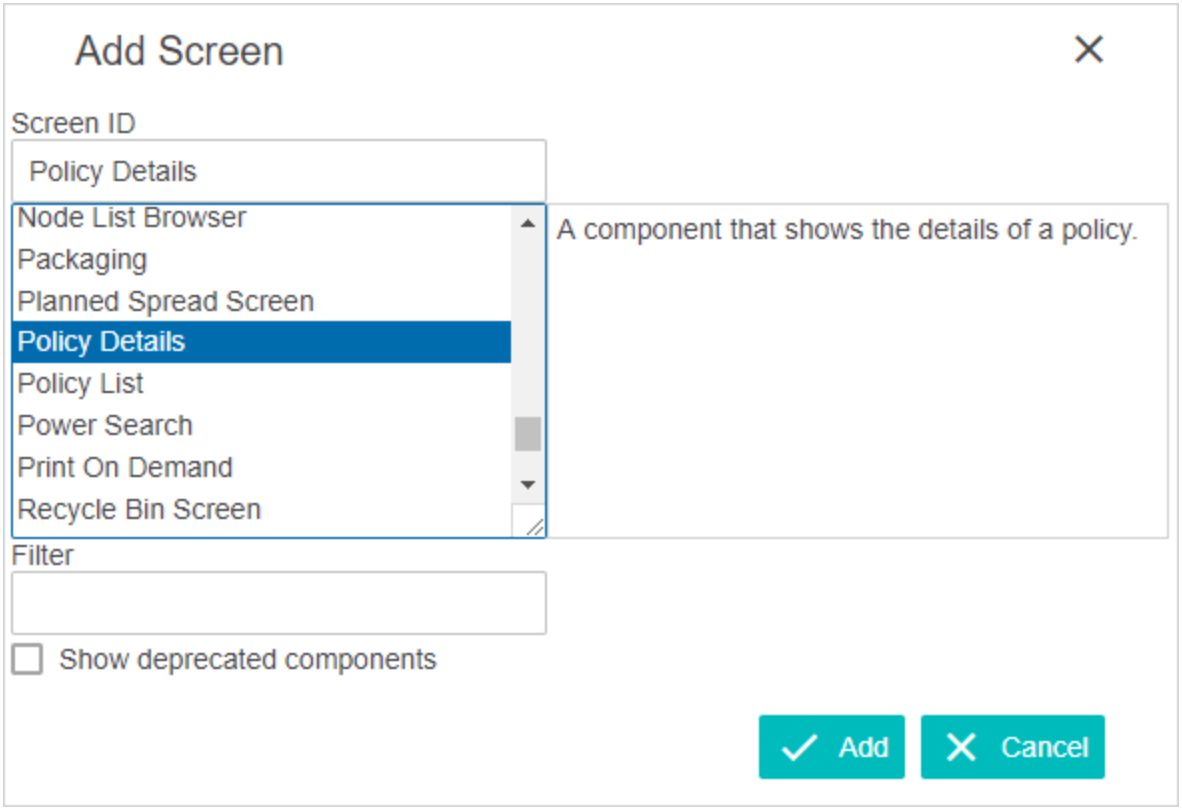
From the Web UI designer, create the following two screens:

- Policy List




  The Policy List will be the main screen where all the policies in the system are displayed.

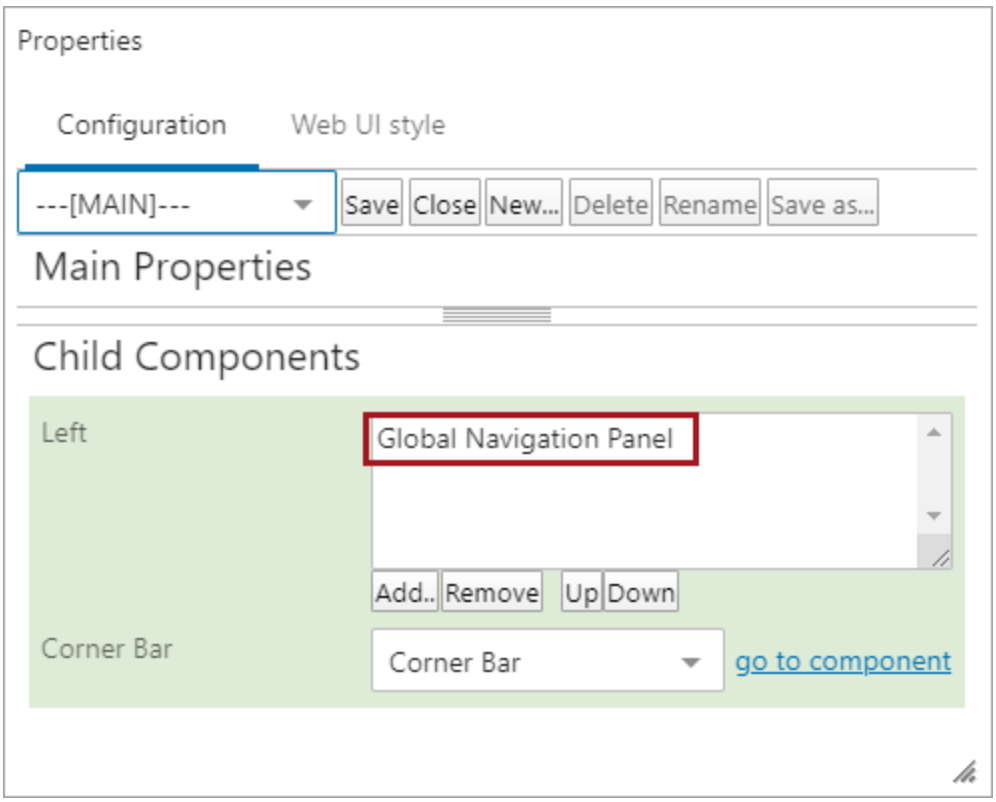
- Policy Details



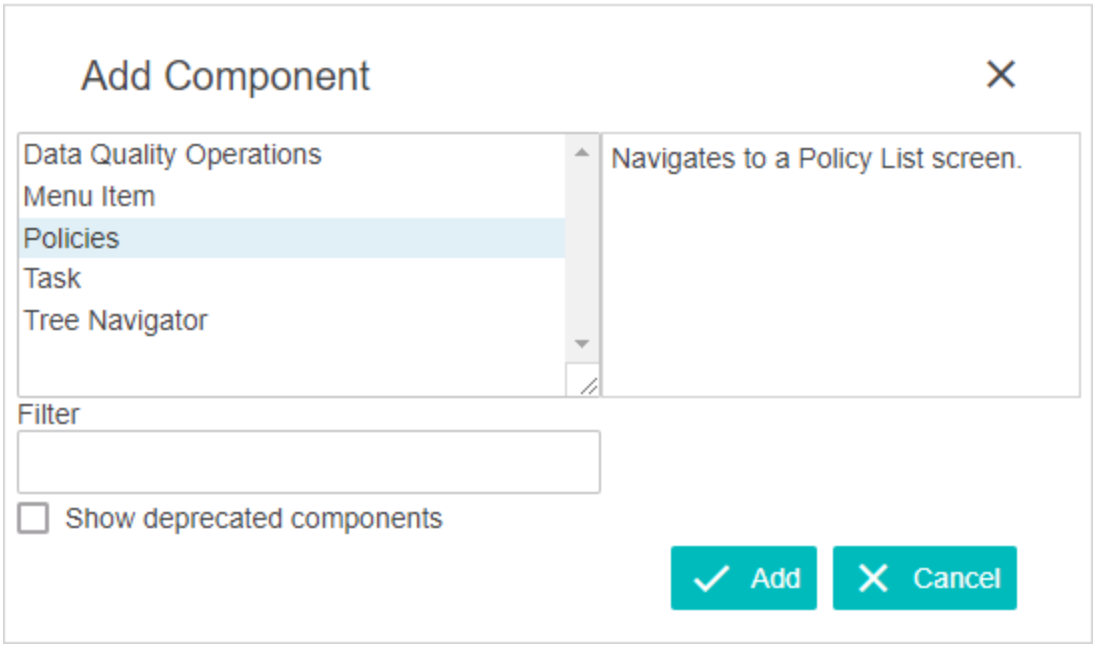

The Policy Details screen will show all the information on each policy after selecting from the policy list.

After the screens are created, from the ---[MAIN]--- screen of the Web UI configurations, under the 'Left' child components section, double click the Global Navigation Panel option.

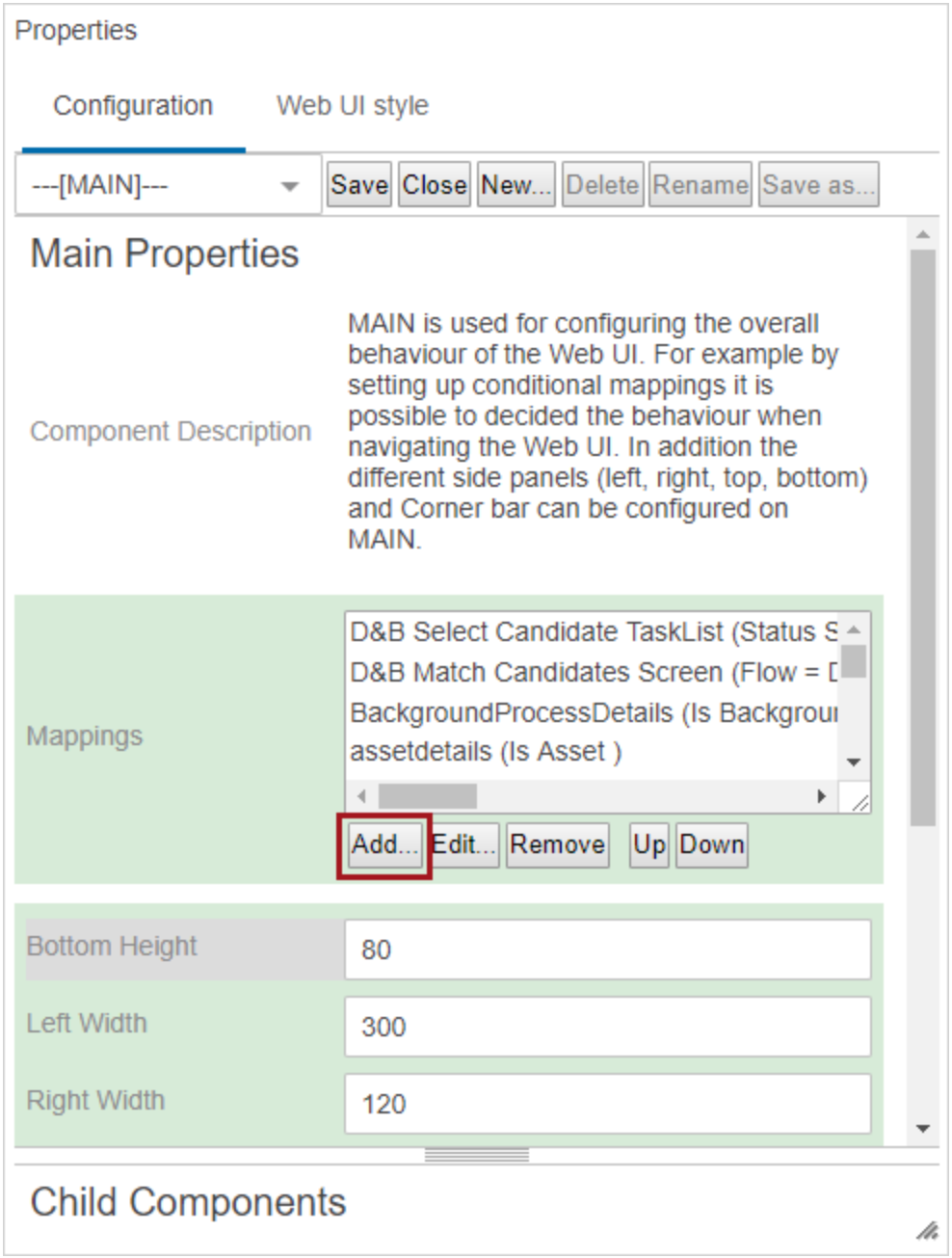
**Note:** Policies only works with the Global Navigation Panel component.




From the Menu Items section, add a Policies component.

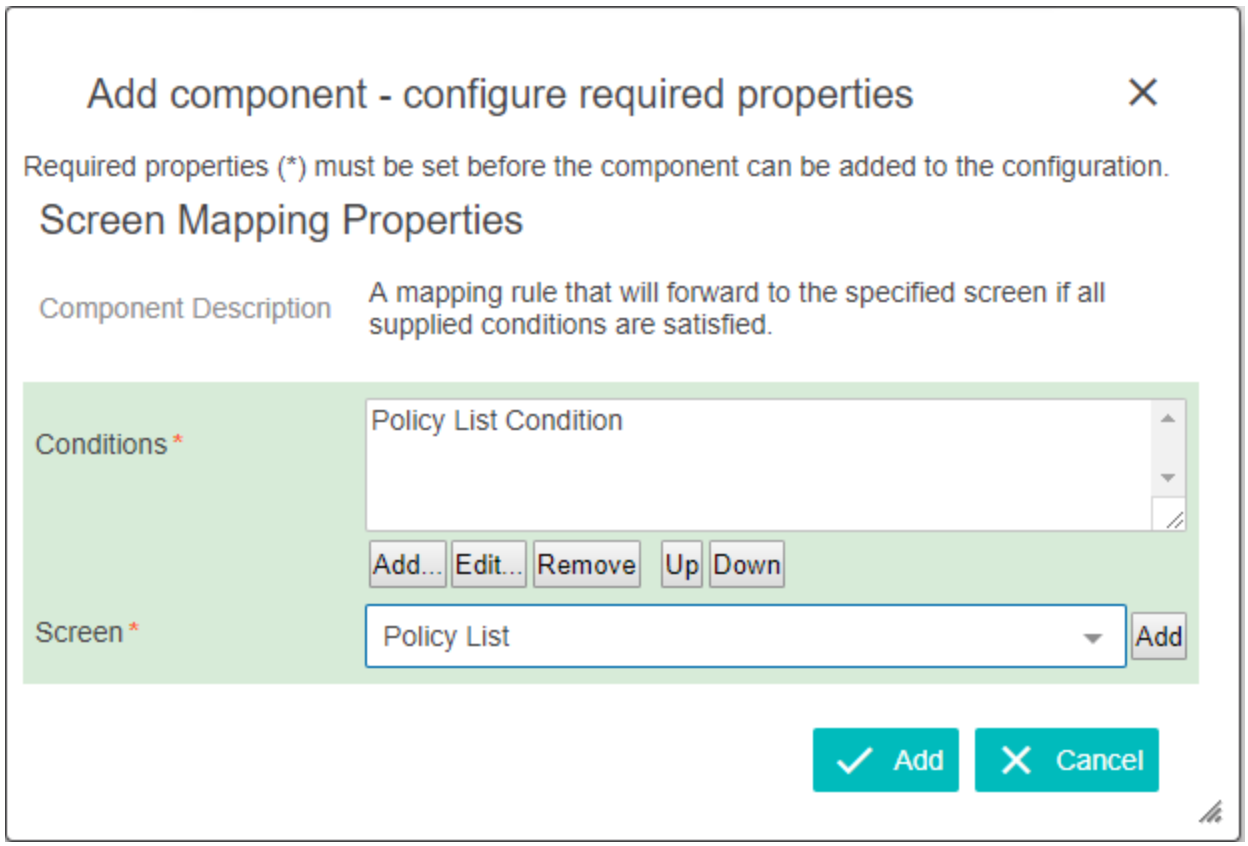



After the Policies component has been added to the Global Navigation Panel, return to the ---[MAIN]--- configuration. From the Mappings section, select the 'Add...' option to configure screen mappings.

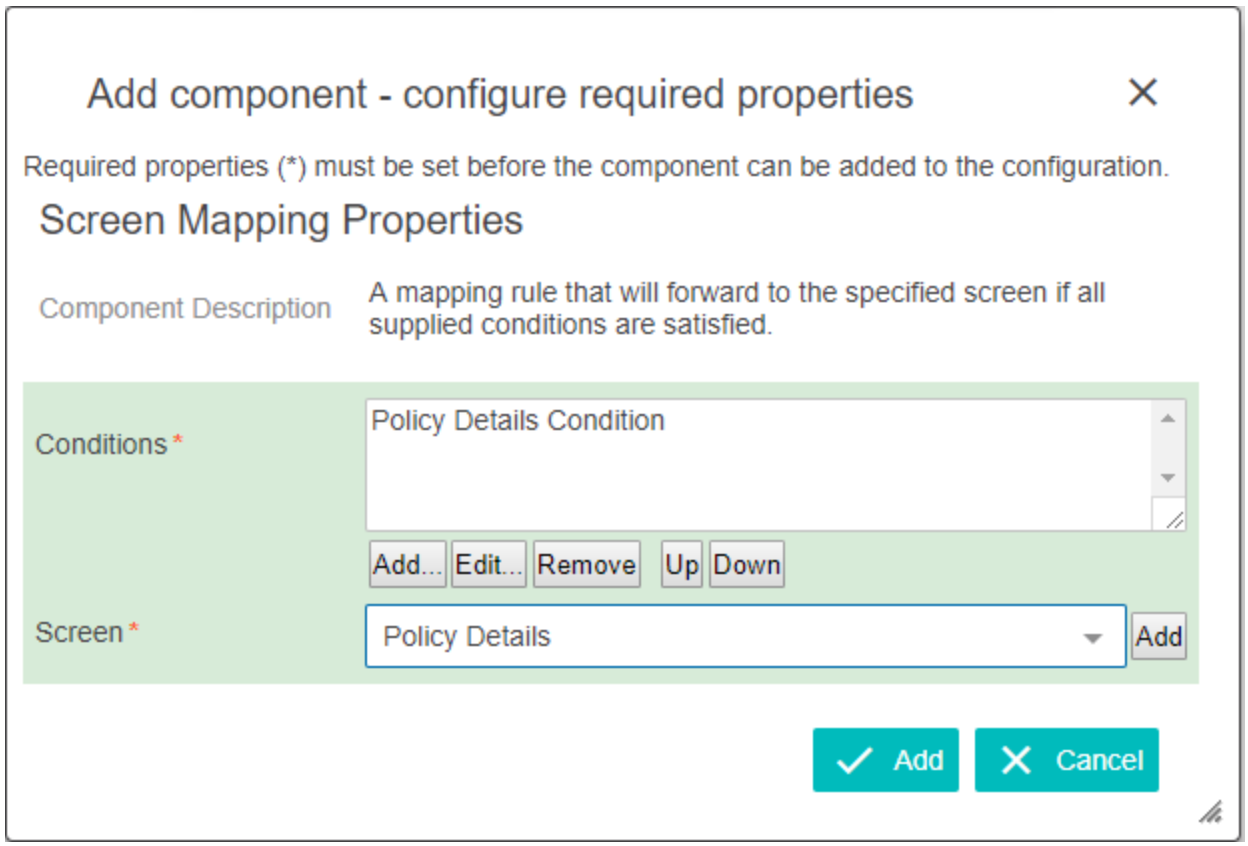



On the 'Add component' dialog, add the following two mapping conditions:

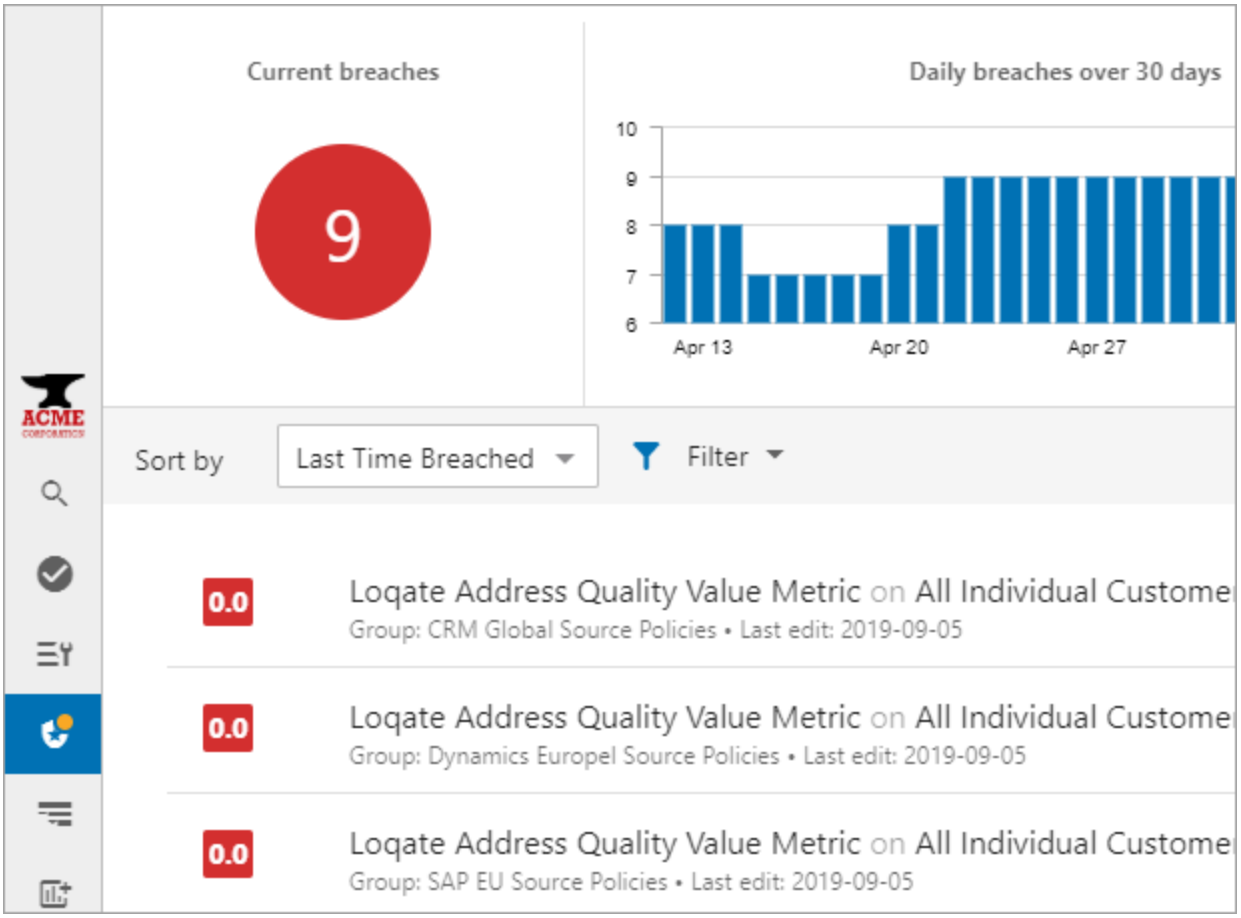
- Policy List:




- Policy Details:

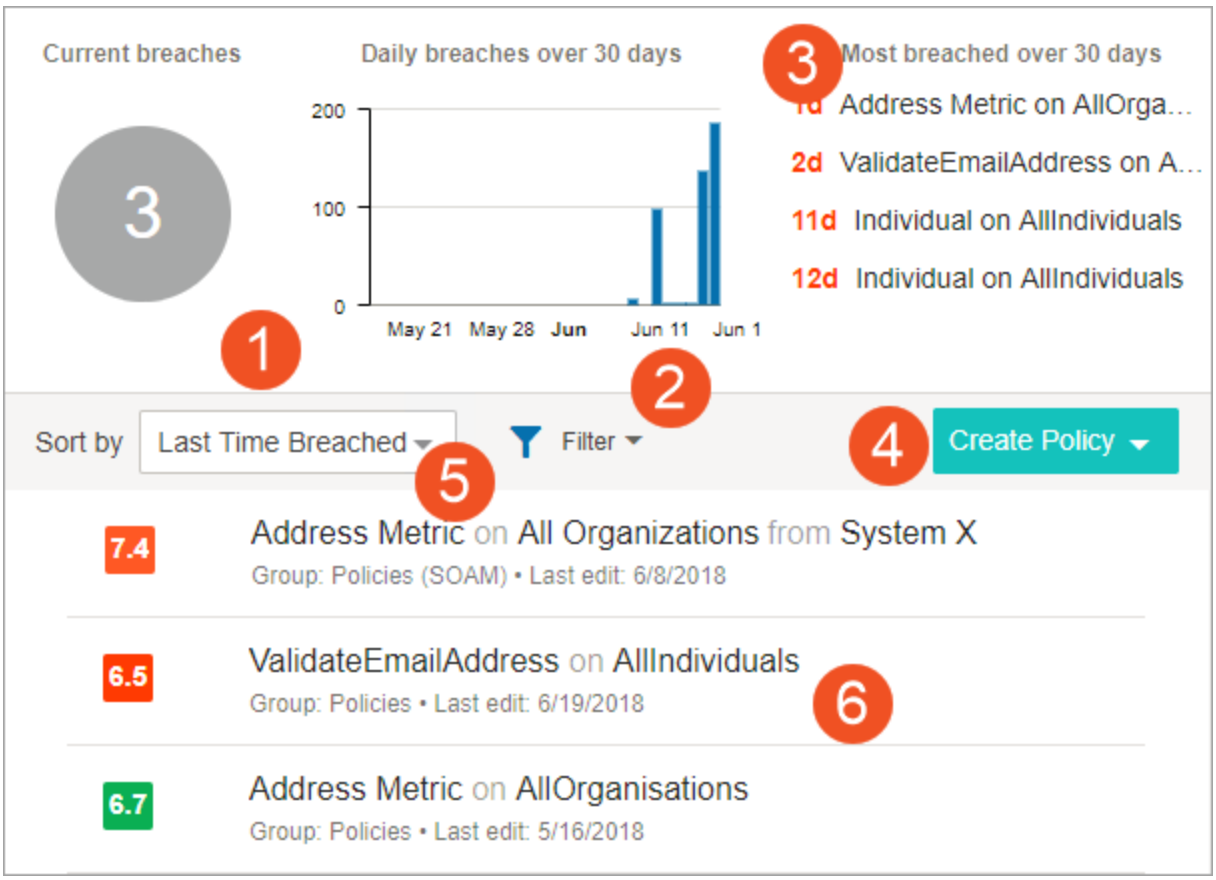



Select 'Add' then save the changes in the Web UI Designer. The Policies tab and each of the policies will now be functional.

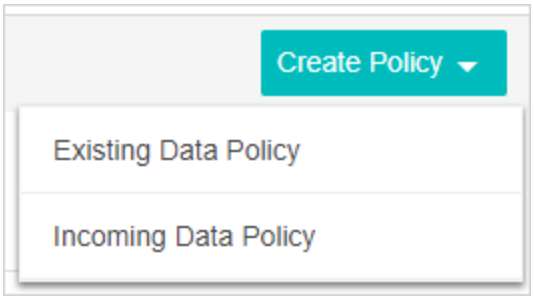



## Elements of the Data Policies Web UI

Select the Policies option from the navigational panel to see a list of all available policies.

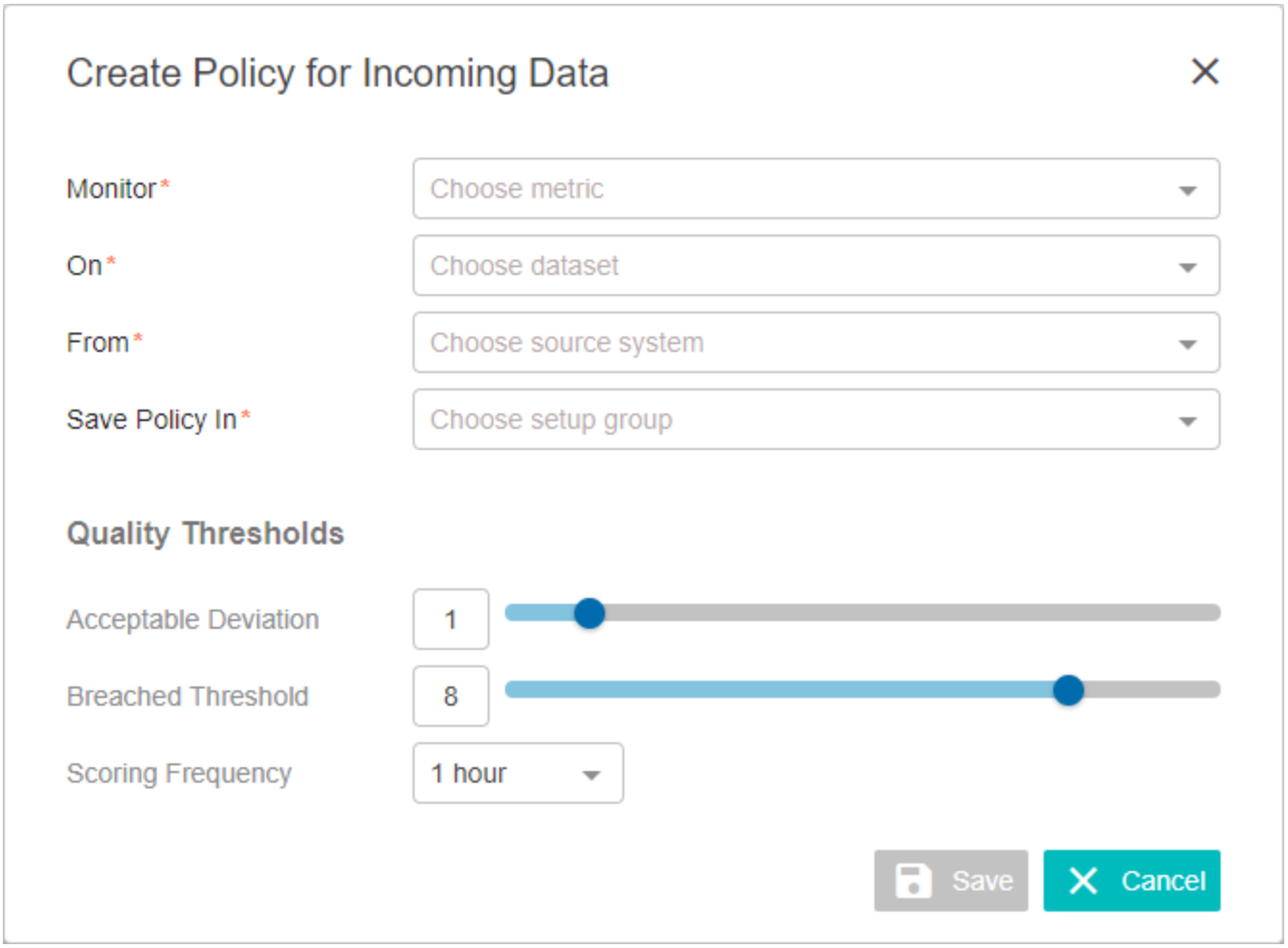


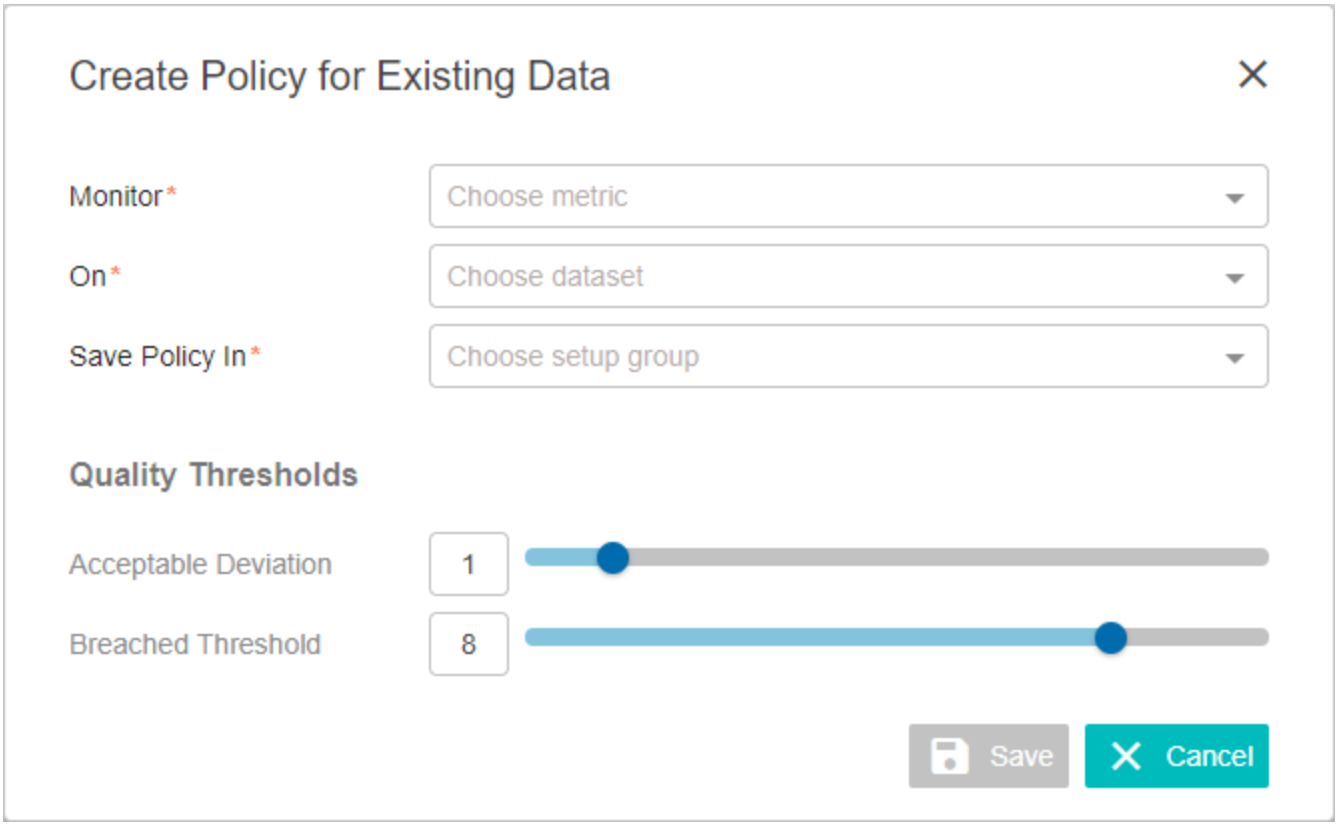

1. **Current Breaches** - This component shows the number of policies that the user has subscribed to and are active that are currently breaching the data policy threshold.
2. **Daily Breaches over 30 Days** - This component shows a breakdown of breaches per day over the last month.
3. **Most Breached in the last 30 days** - This section shows which policy has had the most policy breaches in the last 30 days.
4. **Create policy** - Click this option to create a new policy. Policies based on 'Incoming Data' is a policy for data that needs to be imported first, or a policy based on 'Existing Data' is data already on the system.



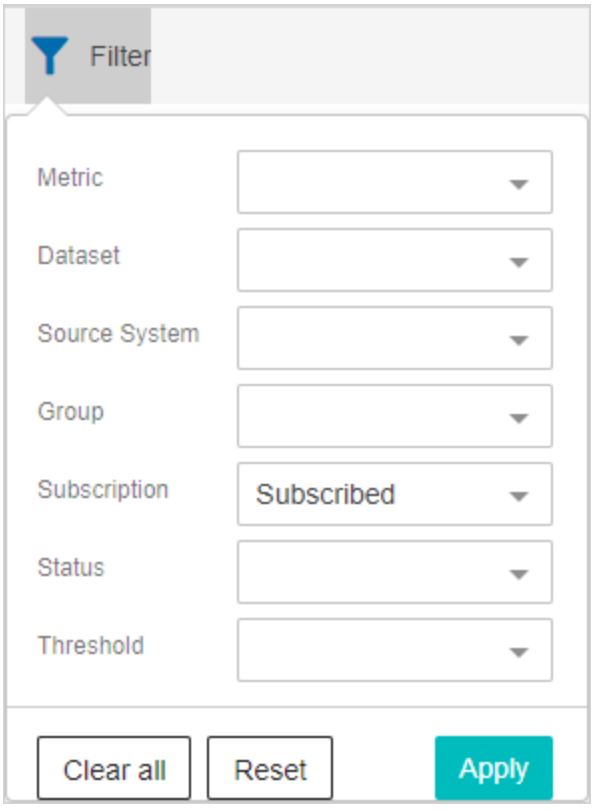

With either option, specify the data metric and data set definition and where to save the policy as well as the acceptable deviation and the breached threshold. Only with Incoming Data Policies will a source system need to be specified and the frequency at which the policy is evaluated.

This frequency determines the resolution of the score graph in the policy details screen, and the response time from bad data starts to come from a source system until the policy will be marked as breached and the user notified. A lower frequency will result in each evaluation being based on fewer incoming data. This lower frequency will cause the variation in scores may be larger.





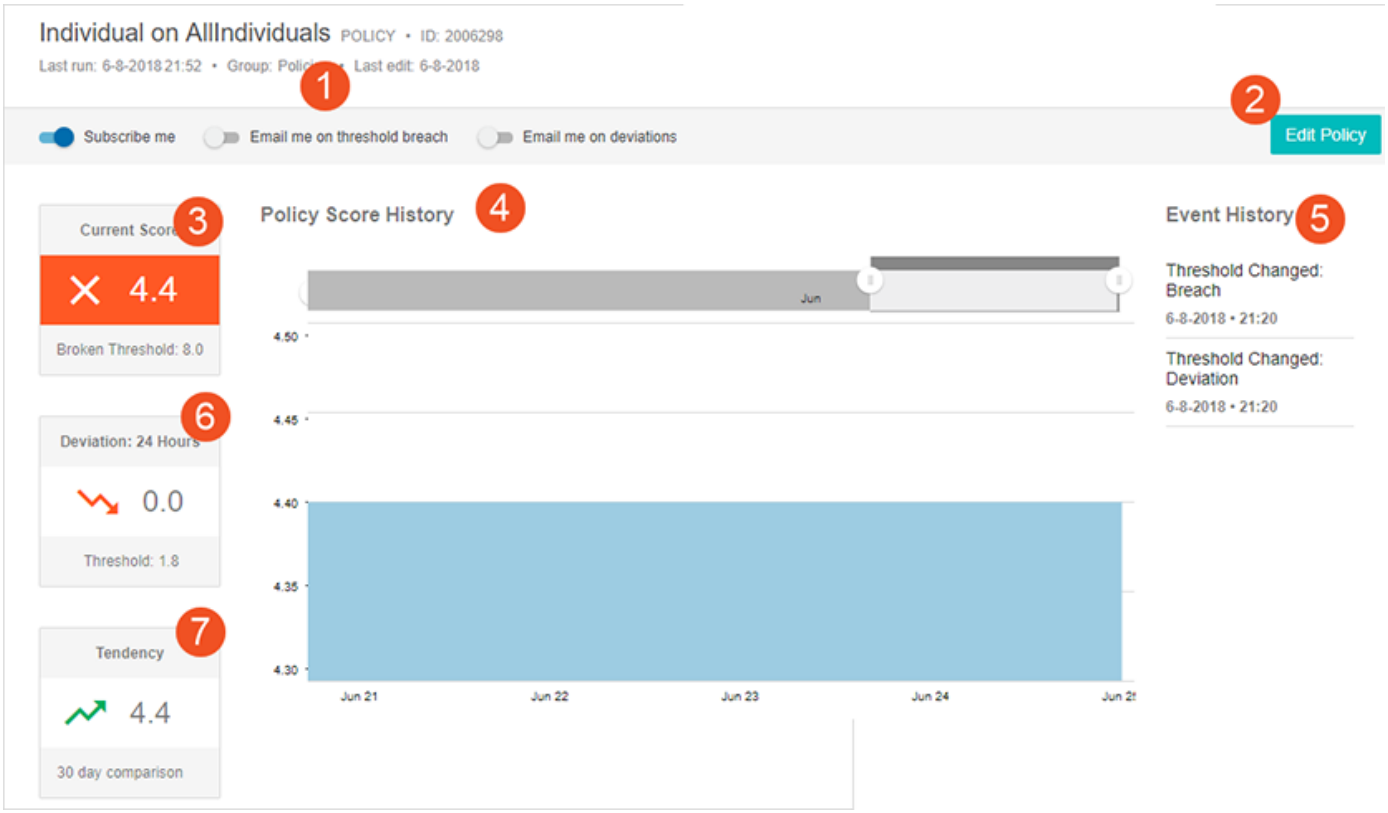

5. **Filter** - The filter allows users to remove undesired policies based on set criteria.



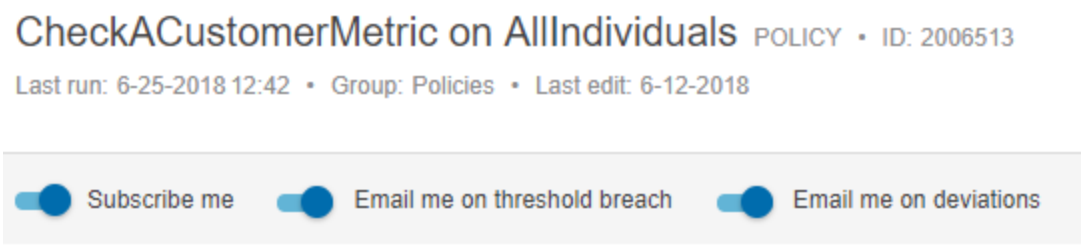
6. **Data Policy List** - This contains a list of all the data policies in your system. Selecting any will open the policy for modification.

## Elements of a Policy Details Screen

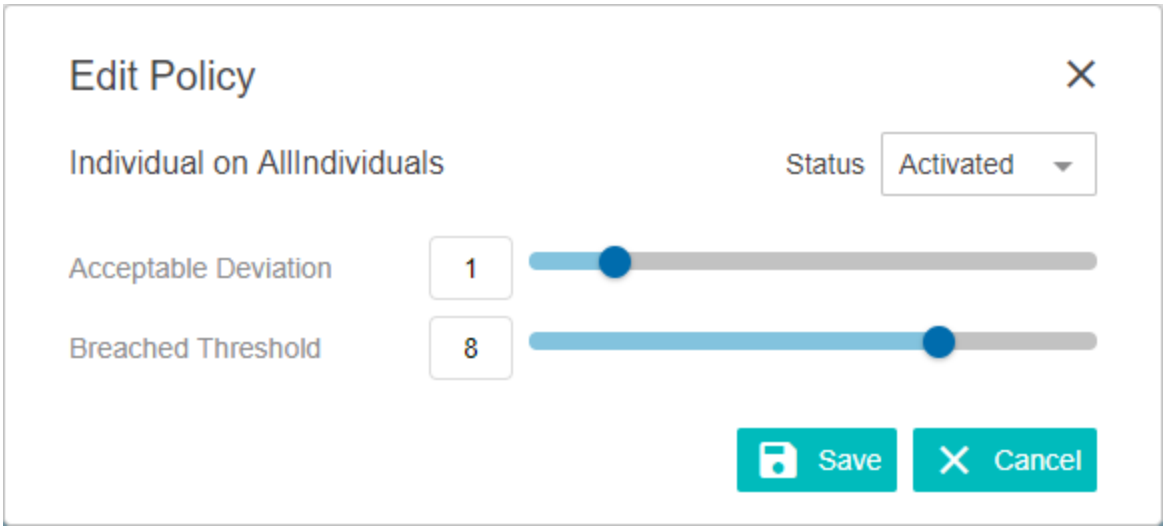
When a data policy is selected from the Data Policy List, the policy can be viewed and modified via the Policy Details screen.




1. **Policy Details Card** - This section includes all the information about the policy including name, ID, and creation date. Below the data policy details card, users may subscribe to the policy and/or receive email notifications when the policy has a threshold breach or deviation.



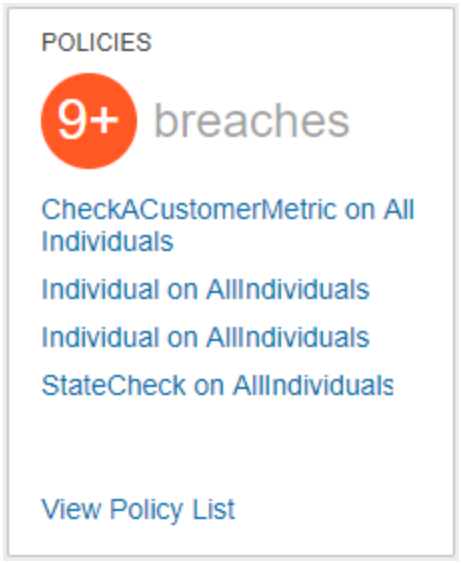

2. **Edit Policy Button** - This button opens the edit policy dialog which allows for changes to the threshold and deviation.




3. **Current Score** - This score is the last evaluation result of the policy.
4. **Policy History** - This section shows historic evaluation results of the policy. This allows users to see fluctuations in the policy's score.
5. **Event History** - This section shows all actions performed on this policy.
6. **Deviation: 24 Hours** - This section shows the deviation, if any, of the last 24 hours.
7. **Tendency** - This section is the comparison of the last seven days with an equivalent seven-day period from the previous month.

## Web UI Homepage Widget

A policy widget is available for use on the Web UI home screen. This widget is used to see the various breached policies currently active in the system.




For more information, see the **Policy Widget** topic of the **Web UI** documentation.