Lepide

DATA DISCOVERY AND CLASSIFICATION

Table of Contents

1 Introduction

The Data Discovery & Classification module from the Lepide Data Security Platform allows you to give more context around your most sensitive data. You can easily identify potential areas of exposure and apply the correct access controls. You will, for example, be able to see how many of your files contain sensitive data, where they are distributed, and what type of data you have.

This guide explains how data is classified within the module and helps to familiarize you with the configuration of Data Discovery & Classification.

If you have any questions at any point in the process, you can contact our Support Team. The contact details are listed at the end of this document.

2 Pre-Requisites

Before Setting Up Data Discovery & Classification the following requirements need to be met:

1. Minimum .netFramework 4.6.2 is required including Windows Updates.

2. If the 'Agent' server for Data Discovery is separate from the SQL server, TCP/IP connection must be enabled from the SQL Server Configuration Manager where SQL resides. Default port used is 1433.

3. DDC Agent server should have proper connectivity to SQL. Port being used should be whitelisted through a new Inbound Rule on the firewall for SQL to communicate between DDC Agent and SQL Server.

3 The Data Discovery & Classification Screen

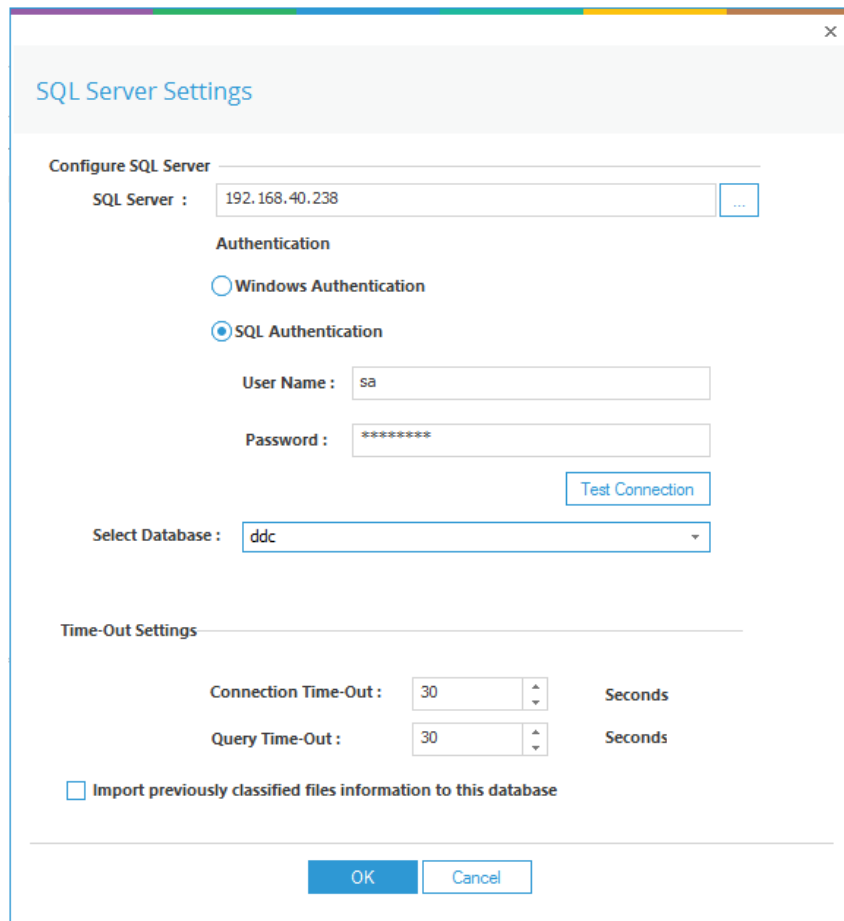- Click the ![icon] icon to go to the Data Discovery & Classification screen:
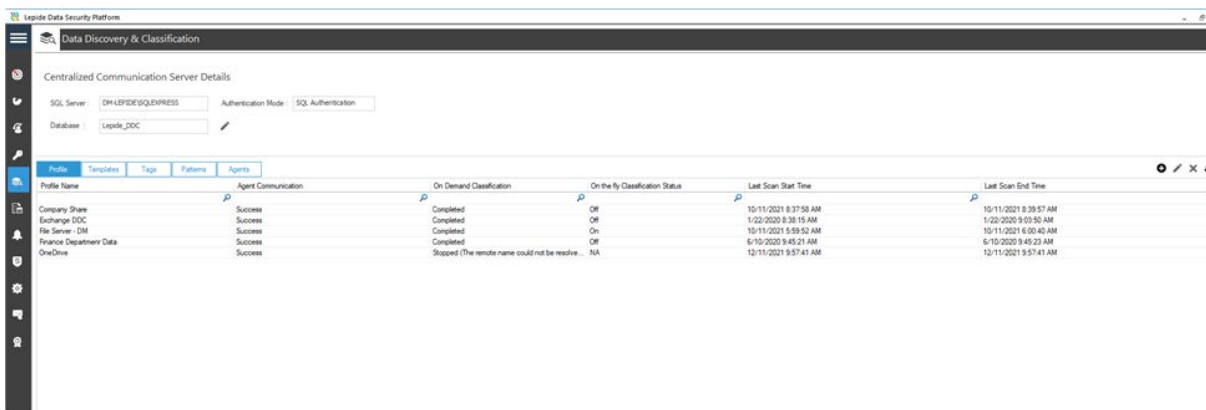
Figure 1: Data Discovery & Classification Screen

From this screen, you can add or modify the Centralized Communication Server Details. This SQL server will contain the file classification data and will be queried by the Data Discovery & Classification agent.

To Configure Data Discovery and Classification SQL server details:

- Click the ✏ icon to Modify the settings. The following dialog box will appear:

Figure 2: SQL Server Settings

- Add the IP address for the SQL server. Click the ⋯ icon to display a list of servers
- Choose **Windows Authentication** or **SQL Authentication**
- Click **Test Connection**
- The following dialog box will appear if the test is successful:
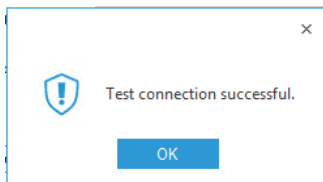
Figure 3: Successful Connection Message

- Click **OK**
- From the **Select Database** drop-down, type a new database name or select an existing database
- Change the **Time-Out Settings** if required
- Check the **Import previously classified files information to this database** if required
- Click **OK**

3.1 How Data is Classified

Once Data Discovery & Classification has been configured, your data will need to be scanned to highlight the most sensitive data and this can then be reported on.
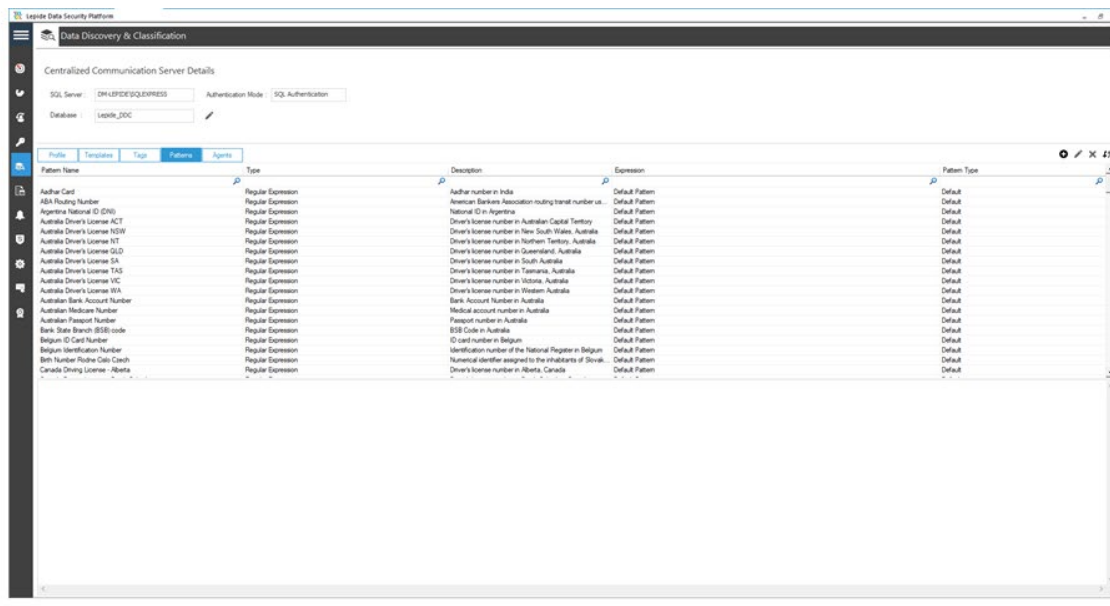
Sensitive data will vary for different organizations and can include anything you choose but typical examples are credit card numbers, passport numbers, social security numbers, personal data and so on. For the Solution to be able to detect your sensitive data, you will need to specify what it should look for. This is done in the form of patterns, tags, templates, and profiles which are explained as follows:

3.1.1 Patterns

Patterns are the strings or regular expressions which can be used to classify the data and are the lowest, most detailed level of the data classification structure.

Data patterns such as credit card numbers, passport numbers, postcodes etc. always follow a certain pattern and so can be detected within data. There are many patterns included by default within the Lepide Data Security Platform, but custom patterns can be created if necessary.

In the example below, the **Pattern Tab** has been selected. Information about the Patterns including Pattern Name, Type and Description is displayed on the screen:

Figure 4: Patterns Tab

3.1.2 Tags

A Tag contains one or more patterns so that you can group similar patterns together.

For example, with a credit card number, a pattern would be a regular expression containing a 16-digit number.

However, credit card numbers can be further subdivided as the first four digits of the number refer to a specific bank.

Therefore, we can have a pattern for say Credit Card Visa, another for Credit Card Amex and so on.

These patterns can then be grouped together within a Tag called Credit Card and it will contain all credit card numbers for the different banks.

In the example below, the **Tags Tab** has been selected. Information including Tag Name, Description, and the Patterns it contains is displayed on the screen:
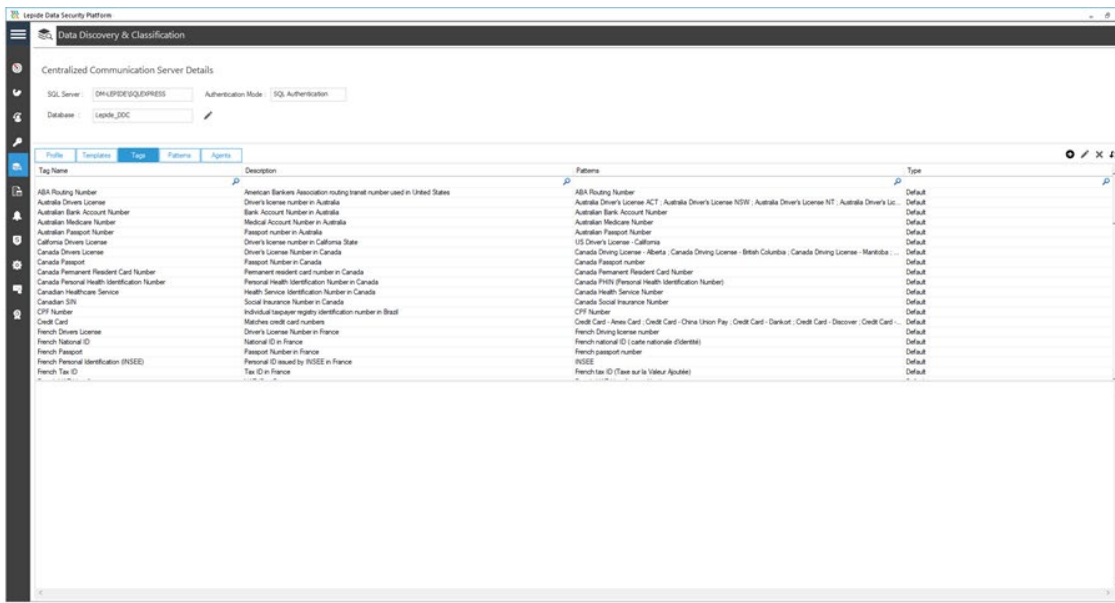
Figure 5: Tags Tab

If you click on a Tag, more information about the Tag is displayed at the bottom of the screen:
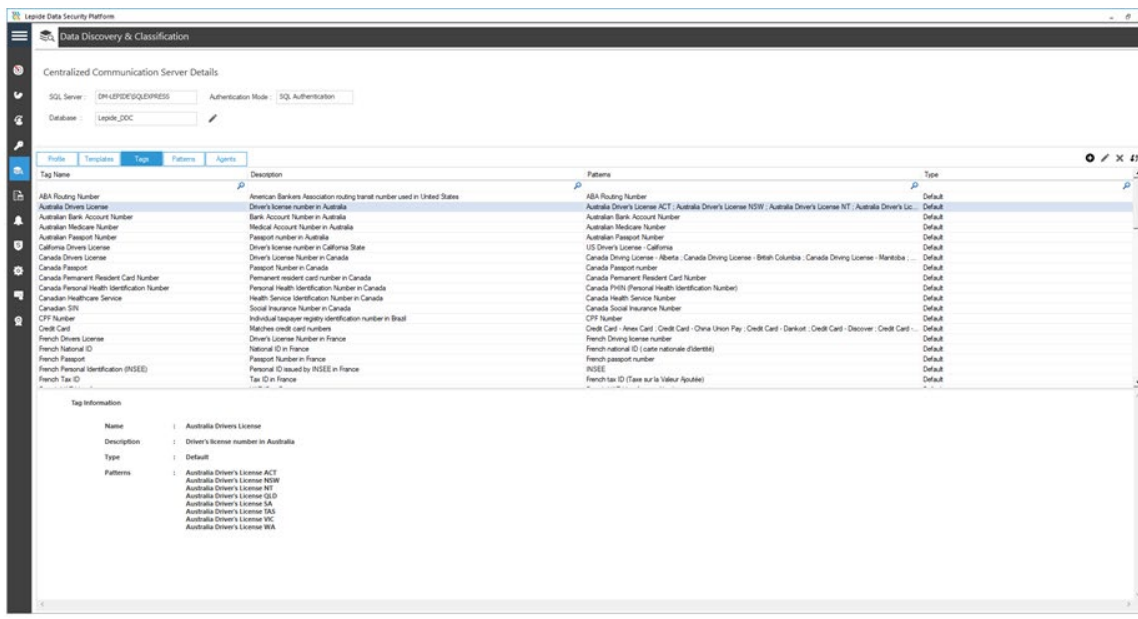

Figure 6: Tags Tab with Detail

3.1.3 Templates

A Template contains one or more Tags so that you can group similar Tags together.

In the credit card example, you may want to group credit card information together with other financial data and so you could create a Template called 'Financial Data'.

In the example below, the **Templates Tab** has been selected. Information including Template Name, Tag Name, Description, and the Patterns it contains is displayed on the screen.
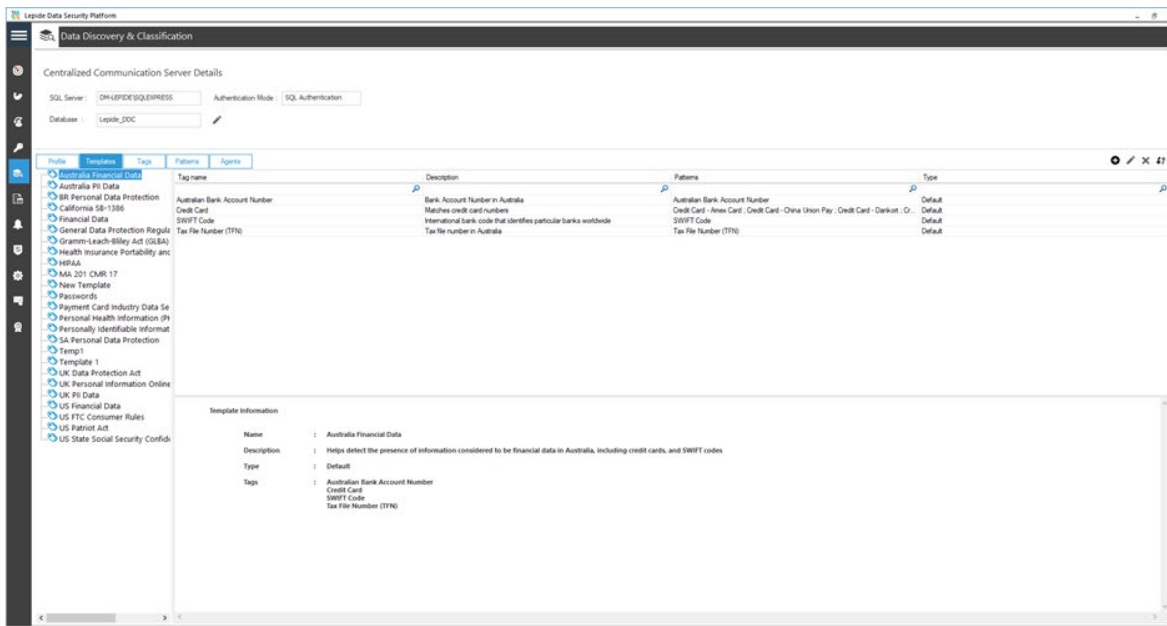

Figure 7: Templates Tab

3.1.4 Profile

A profile contains a group of Templates and is the highest level of the data classification structure.

The example below shows a Profile called **Company Share** and status information related to it. In the lower part of the screen is more detailed information relating to the Profile including the Component Type, Folder Information, and the Templates it contains.
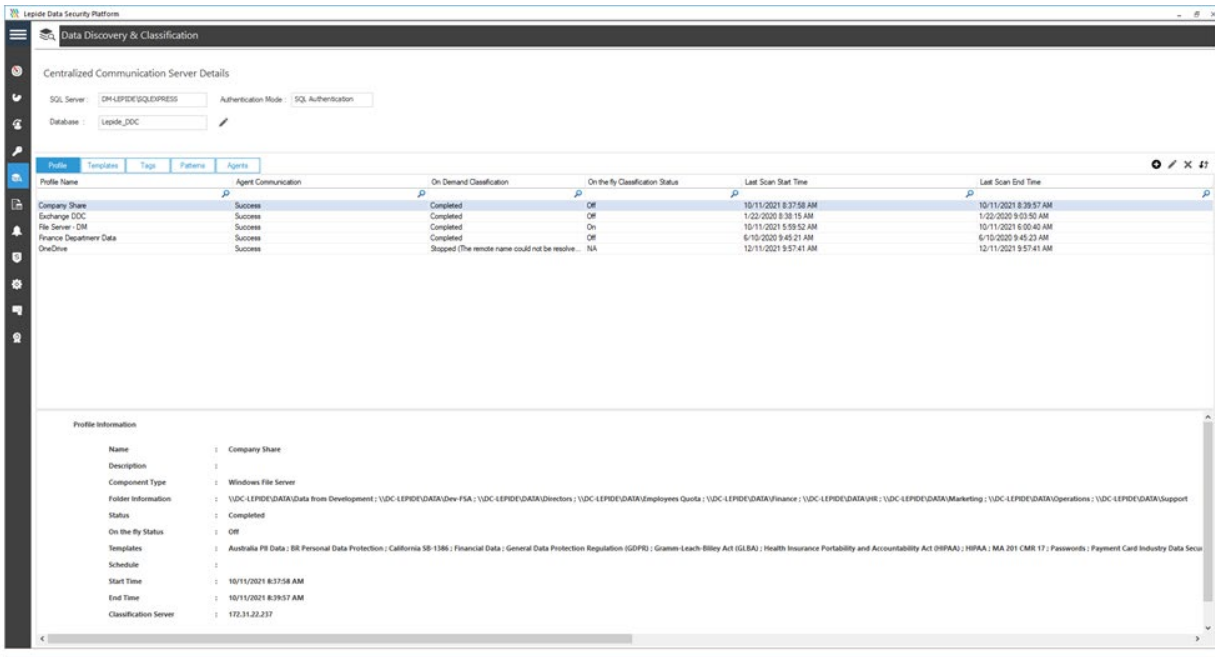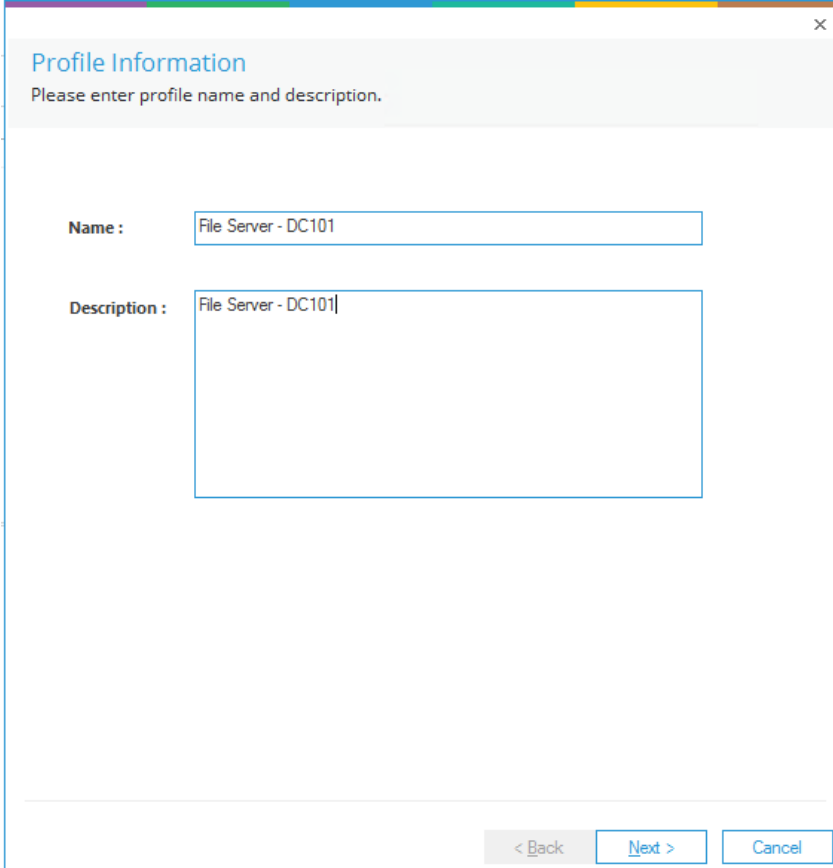

Figure 8: Profiles Tab

4 Create a New Profile for Windows File Server

The first step in configuring Data Discovery & Classification is to add a Profile. The Profile will contain Classification Templates which in turn will contain the Tags and Patterns for the data:

- From the Data Discovery & Classification screen
- Click on the **Profile** tab
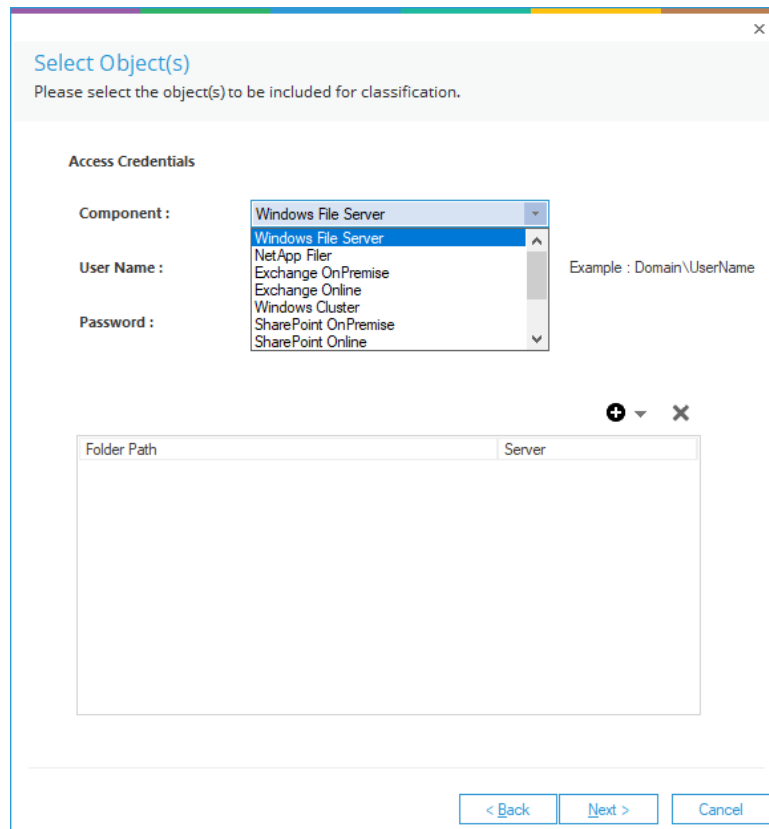- Click the ⊕ icon
- The **Profile Information** dialog box appears:

Figure 9:Profile Information Dialog Box

- Add a profile **Name** and optional **Description**
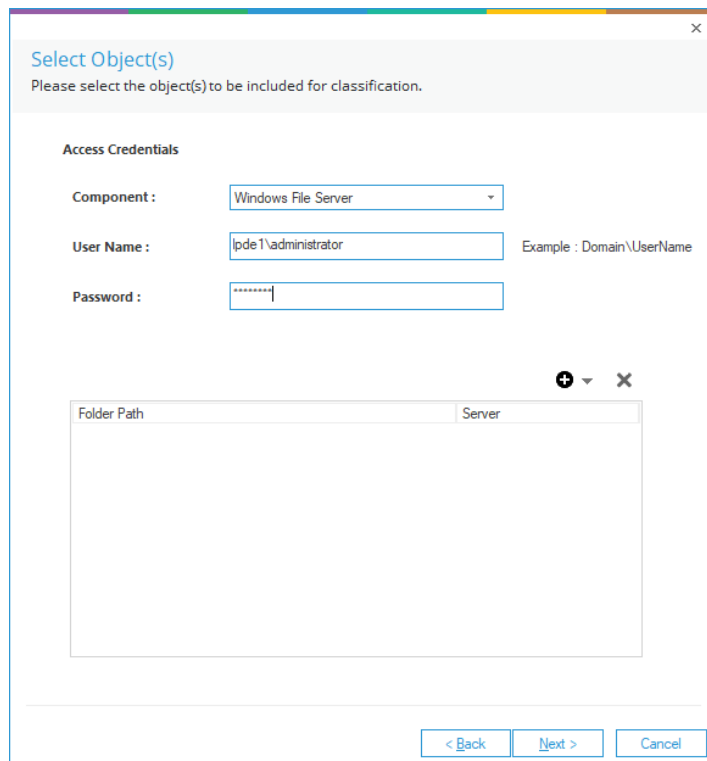- Click **Next**

The **Select Object(s)** dialog box is displayed:

Figure 10: Select Objects Dialog Box

This dialog box lets you add one of the following components:

- Windows File Server

- NetApp Filer

- Exchange OnPremise

- Exchange Online

- Windows Cluster

- SharePoint OnPremise

- SharePoint Online

- OneDrive

- Dropbox

- Linux

- Unix

The following steps will cover the **Windows File Server** component. The other components are described later in this document.

- Select the **Windows File Server** component
- Type in the **Username** and **Password**

Figure 12: Select Object - Windows File Server

4.1 Folder Options

You will need to specify the folder(s) where the data is stored. There can be multiple folders included within a profile. To add a Folder:

- Click the ⊕ icon. A menu is displayed with the following options:

  All Shares: This is all shared folders within file server components

  Import from Existing DataSet: This is an existing dataset within the Lepide Data Security Platform

  Import from CSV: This is a .CSV file containing folder location paths

  Set UNC Path: This is a path specified in the format: \\server-name\shared-resource-pathname
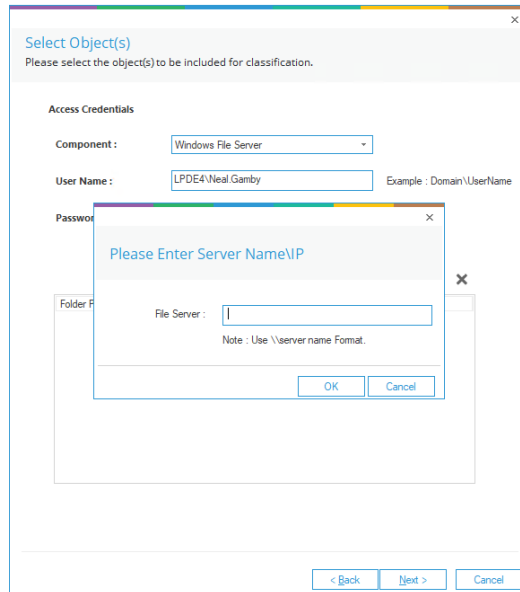
Scan Shares from audited File Servers: These are shared drives within file server components which have been added to the Lepide Data Security Platform.
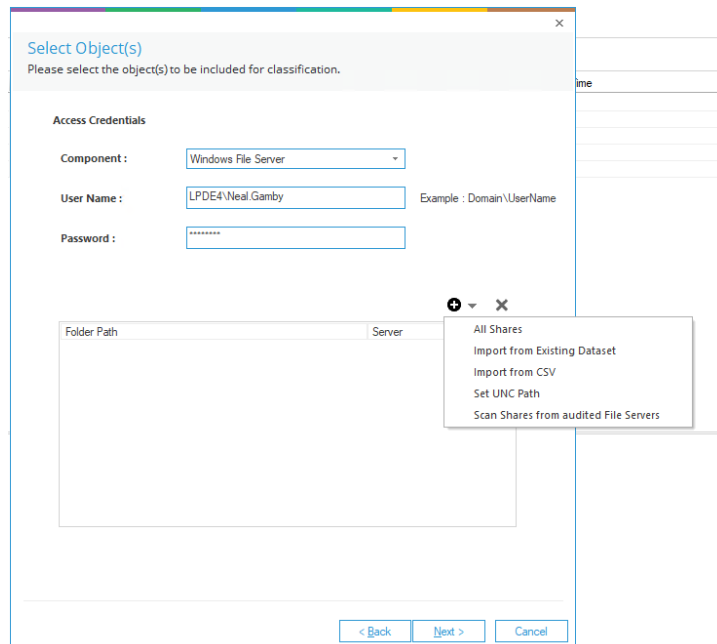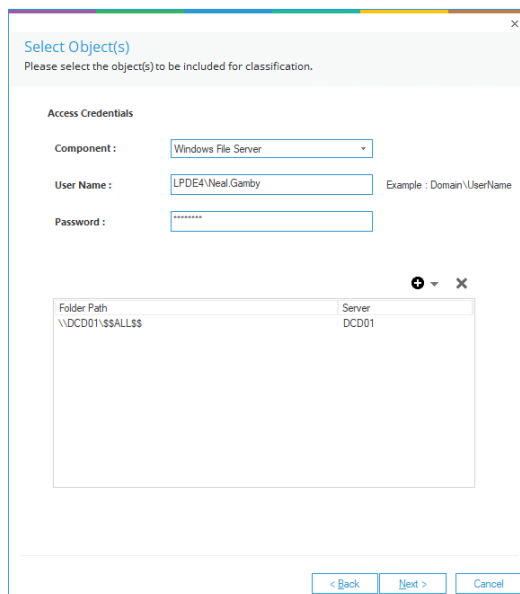

Figure 13: The Add Menu

- For this example, select **All Shares**. The following dialog box appears:

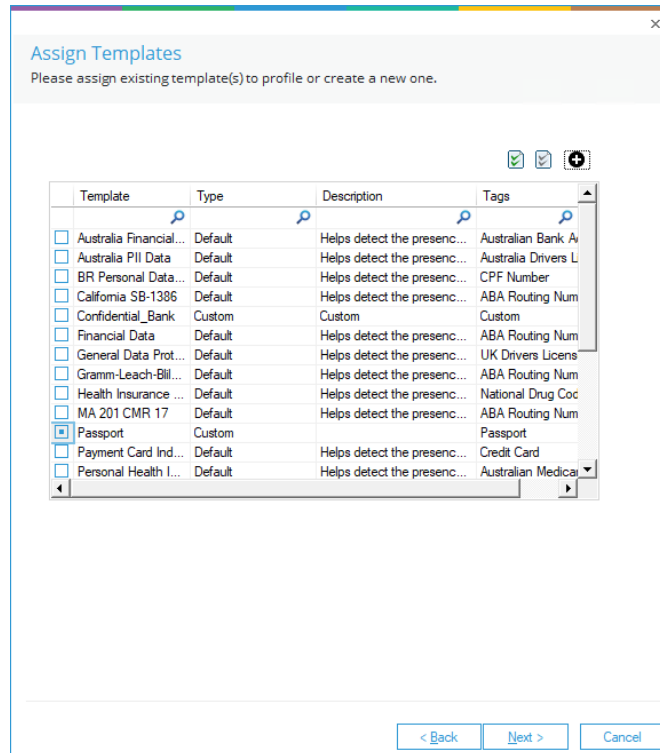- Specify the File Server path using \\server name format and click **OK**

Figure 14: Specify the File Server Path

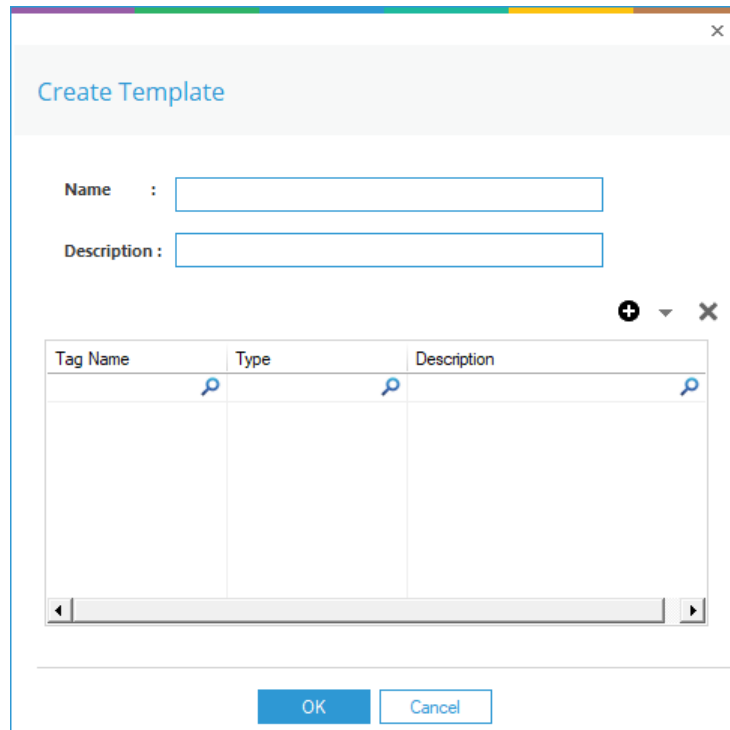*Figure 15: Folder Path Specified*

- Click **Next**

The **Assign Templates** dialog box is displayed:

Figure 16: Assign Templates dialog box

- From here you can:

– Select one or more existing Templates using the check boxes

– Click the ☑ icon to select all Templates

– Click the ☑ icon to remove the selection

– Click the ⊕ icon to add a new Template. The following dialog box is displayed:

4.2 Create a New Template

Figure 17: Create Template dialog box

- Add a **Name** and **Description** for the Template:

- Click the ⊕ icon to add a **Tag** to the template. A menu is displayed:
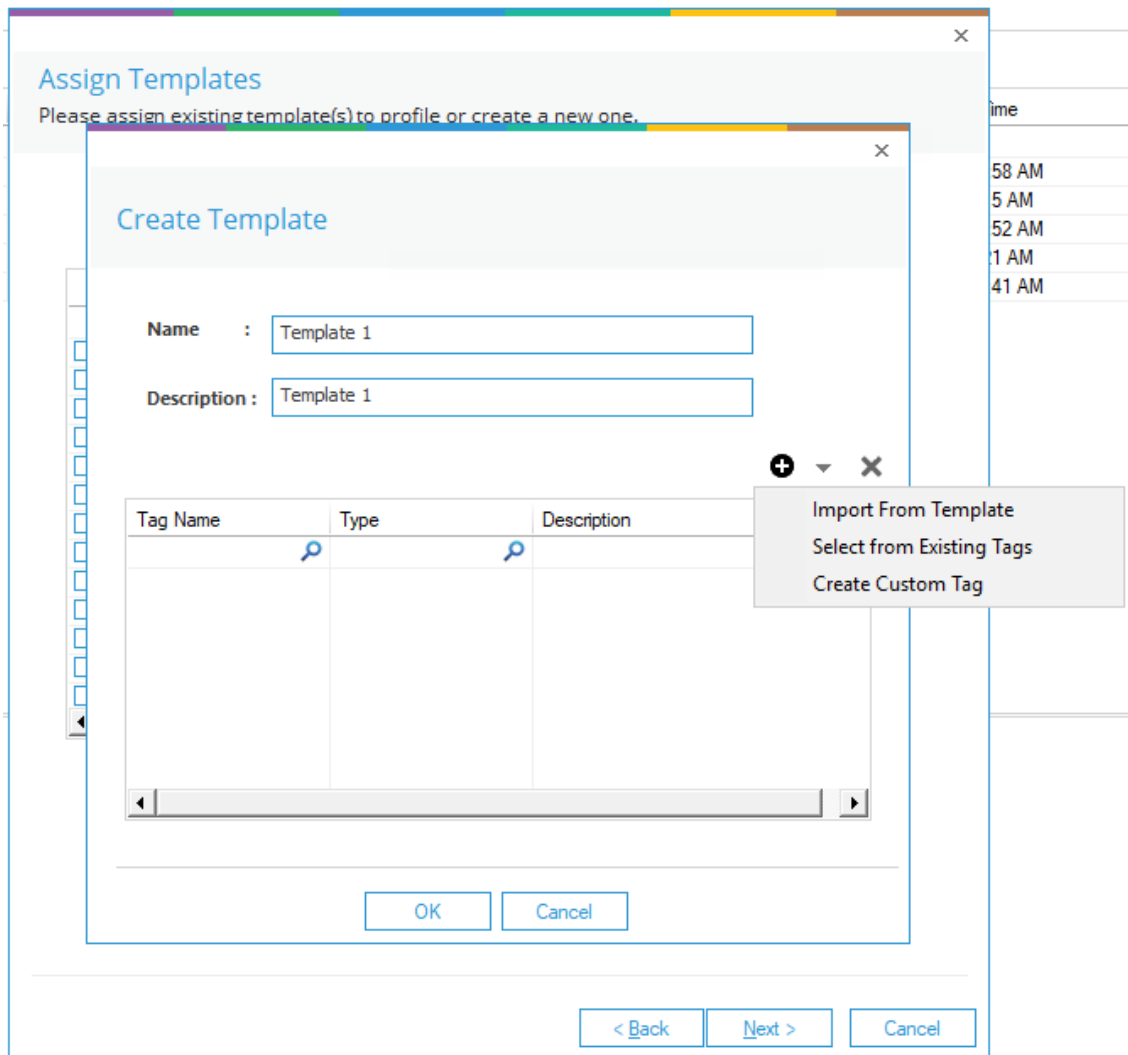
Figure 18: Add Tag Options

- You can do one of the following:

  – Import From Template

  – Select from Existing Tags

  – Create Custom Tag

- For this example, choose **Create Custom Tag** and the **Create Tag** dialog box will appear:
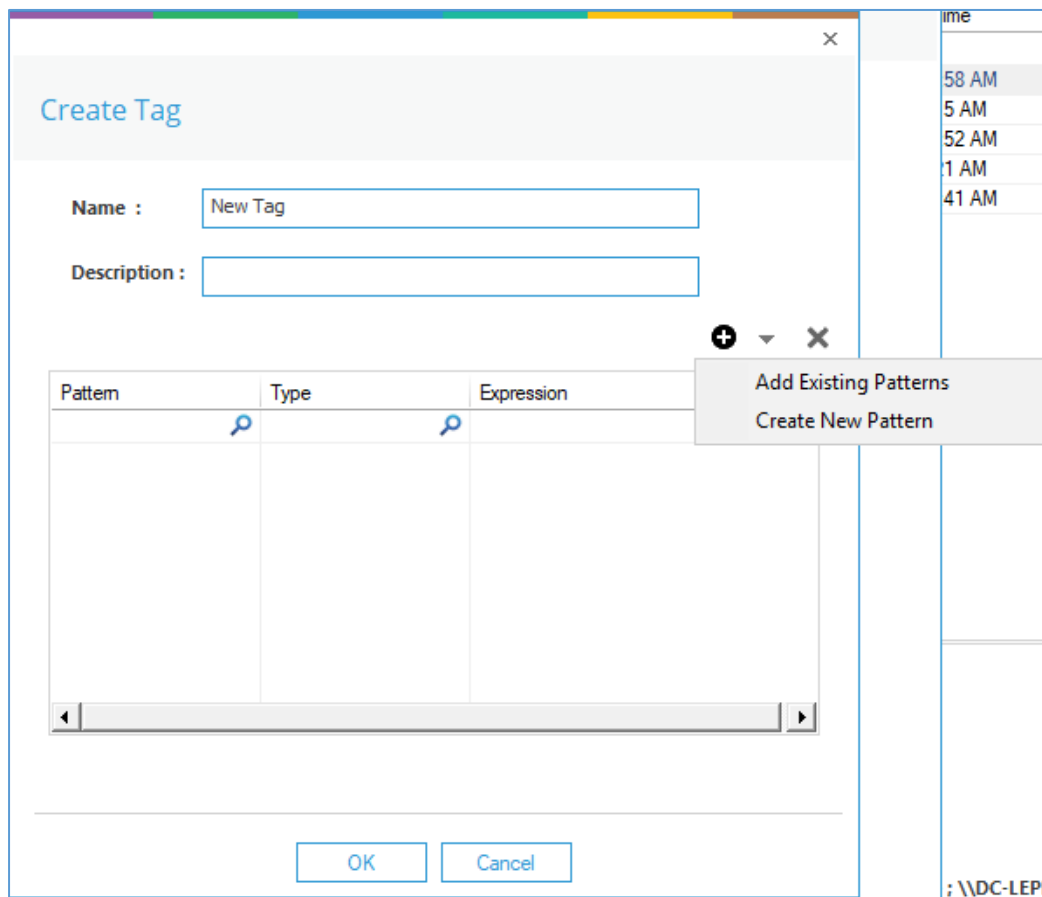
4.3 Create a New Tag

Figure 19: Create Tag

- Add a **Name** and **Description** for the Tag

- The next step is to add the pattern(s) to the Tag

Patterns are the strings or the regular expressions which can be used to classify the files. You can either choose from an existing pattern or create a new one.
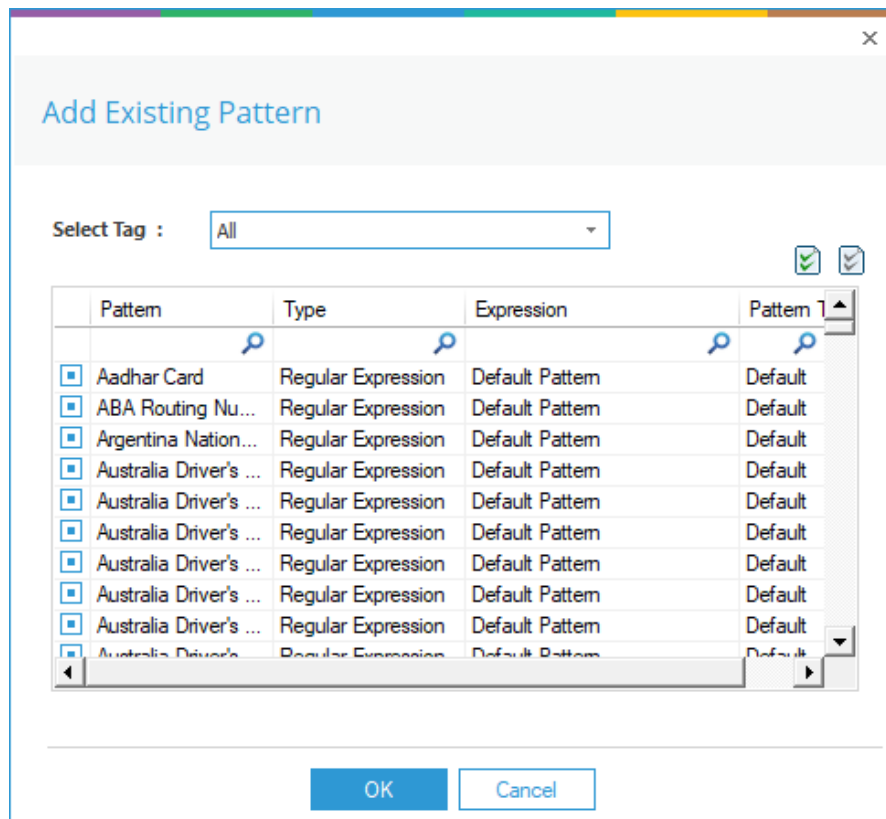
- Click the ⊕ icon and a menu will be displayed:

Figure 20: Pattern Options

4.4 Add an Existing Pattern to a Tag

- To add an existing pattern, choose **Add Existing Patterns** and the following dialog box will appear:

Figure 21: Add Existing Pattern

- From here you can:
    - Select one or more existing Patterns using the check boxes

    - Click the ☑ icon to select all Patterns

    - Click the ☑ icon to remove the selection
- Select which Patterns you require then click **OK**

4.5 Create a New Pattern

From the Create Tag dialog box:

Click the ➕ icon and select **Create New Pattern**. The following dialog box is displayed:

Figure 22: Create New Pattern

- Add a **Name** and **Description** for the Pattern
- Choose the **Type** which can be **Regular Expression**, **String** or **Case Sensitive String**
- Type in the **Regex/String**. This is in PowerShell code format
- Type a value for **Risk Level** and **Monetary Value**
- The **Supporting Evidence** check box will be checked by default and can be unchecked if required

NOTE: Default Supporting Evidence is included in version 24.0.1 for the US, UK, and Australia.

- Add **Proximity**: which is the number of characters around the Pattern
- **Keywords to include**:
- **Keywords to exclude**:
- **Regex to include**

- Click **OK** when finished

- Click **OK**

The File Type dialog box is displayed:
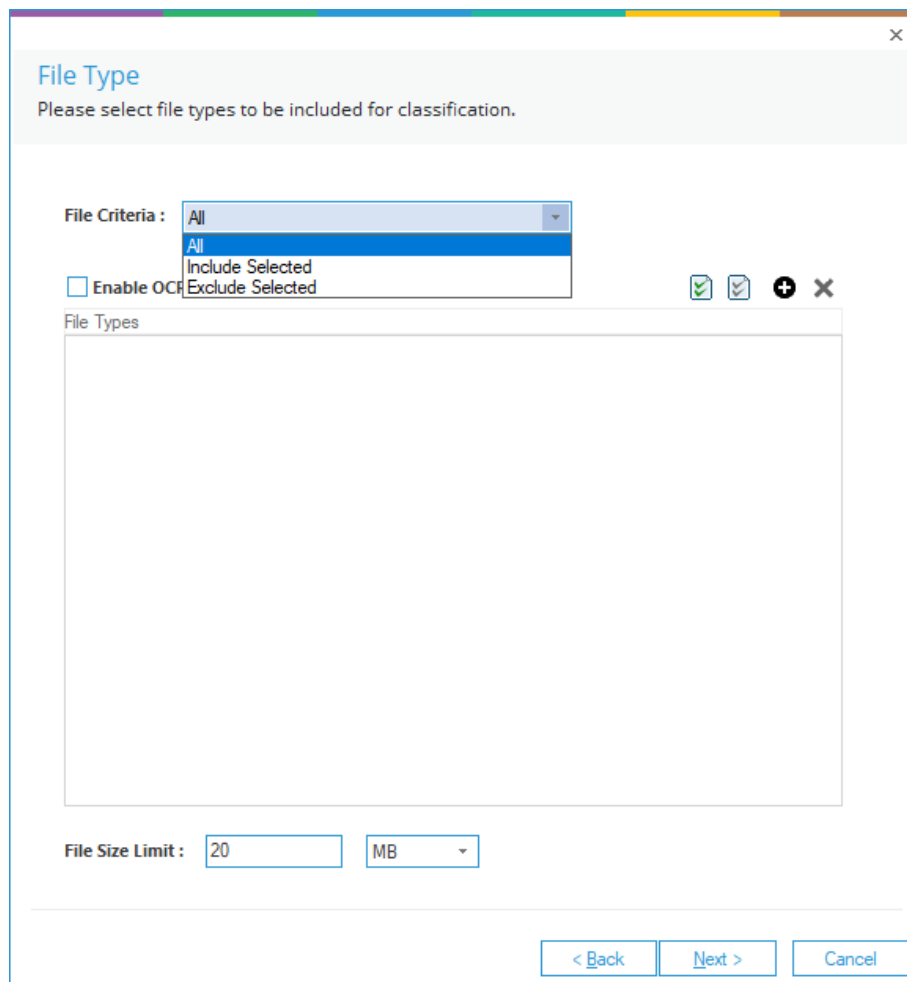
Figure 23: File Type dialog box

- Select the file types which you want to scan for the Data Classification. The Default is **All**.

- If you choose **Include** or **Exclude Selected** you can do the following:

    - Select one or more file types using the check boxes

    - Click the ☑ icon to select all File Types

    - Click the ☑ icon to remove the selection

    - Click the ➕ icon to add a new File Type. Check the **Enable OCR** box if required

    - Click the ✖ icon to remove a File Type

- Select the **File Size Limit**
- Click **Next**

- The **SQL Server Settings** dialog box is displayed:

Figure 24: SQL Server Settings

- Specify the **SQL Server Settings**
- Click **Next**

- The Classification Server dialog box appears:

Figure 25: Classification Server

- You will need to set the classification Method and specify the Classification Server.

- Classification Method has the following options:

  – **On the fly Classification** – Files will be scanned whenever they are changed or added

  – **Classify Now** – This will run a scan immediately

  – **Classification Schedule** – This allows you to specify a date and time to run a scan

- You can choose one or more of these options as required.

- If you choose Classification Schedule, the **Change Schedule** button is enabled. Click this button to define your schedule:

Figure 26: Define Schedule

- From the Classification Server dialog box, you will need to specify the Classification Server. This is the machine where the classification agent is installed. This can be either the Lepide Console, the Local File Server or any remote file server.

- Click **Finish** when you have completed the information

The Data Discovery & Classification window will be displayed along with the Profile Name, Profile Progress and related information:
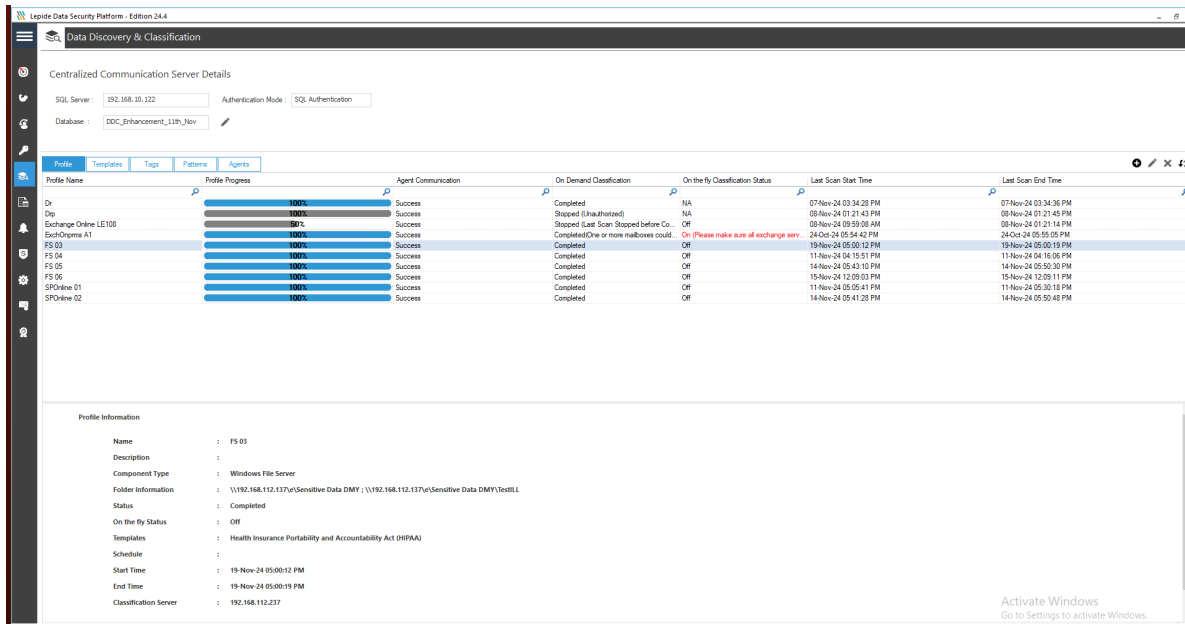
Figure 27: Data Discovery & Classification Window showing Profile Progress and Profile Information

> **NOTE**: The Lepide Web Console has a notification option to choose whether to send an alert when the scan starts and whether to notify of the scan status when it changes. This configuration can only be done in the Web Console so please refer to the Web Console Configuration Guide for more information about this.

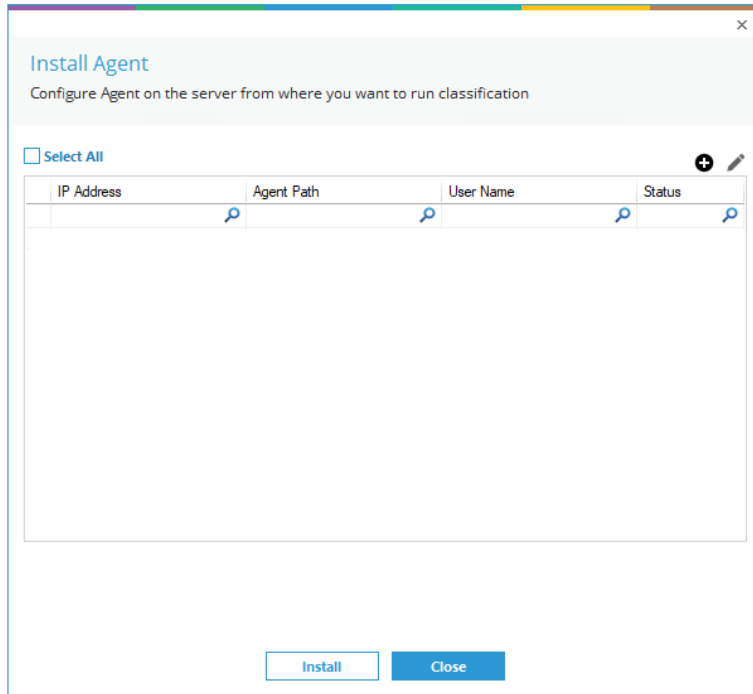If an Agent is not installed, the following message box will be displayed:
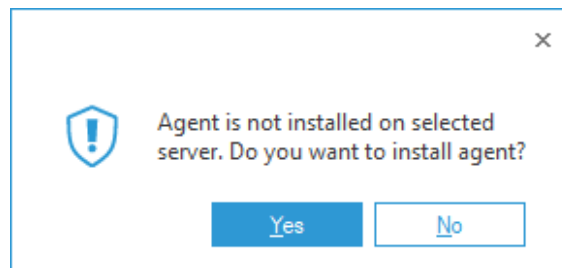
Figure 28: Message Box Asking to Install an Agent

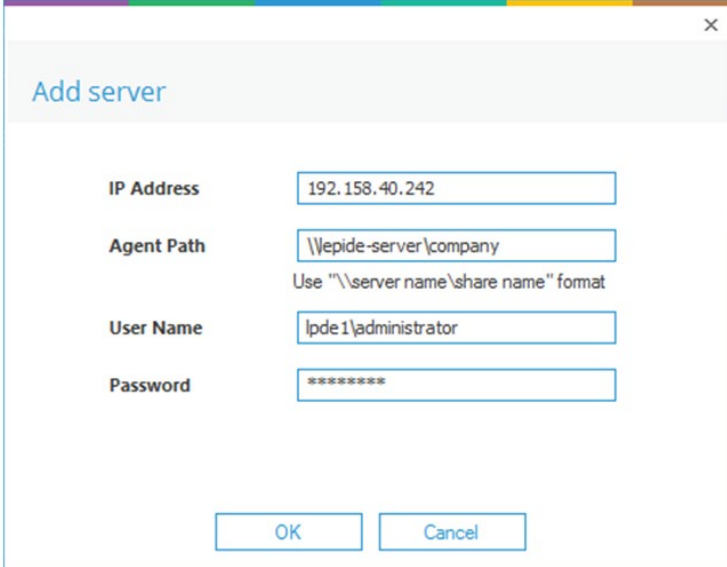- Choose **Yes** to install the Agent

4.6 Install an Agent

> **NOTE**:   The DDC Agent should be uninstalled after installing the latest setup.

The **Install Agent** dialog box is displayed:

Figure 29: Install Agent Dialog Box

- Click the ➕ icon to add a new server. The **Add Server** dialog box is displayed:
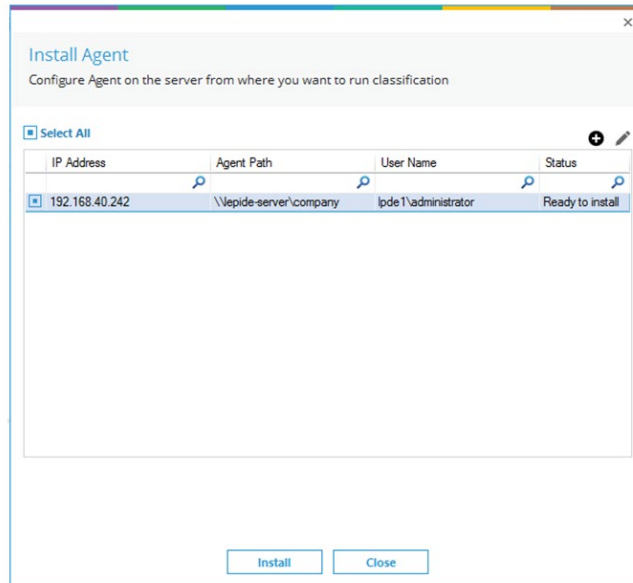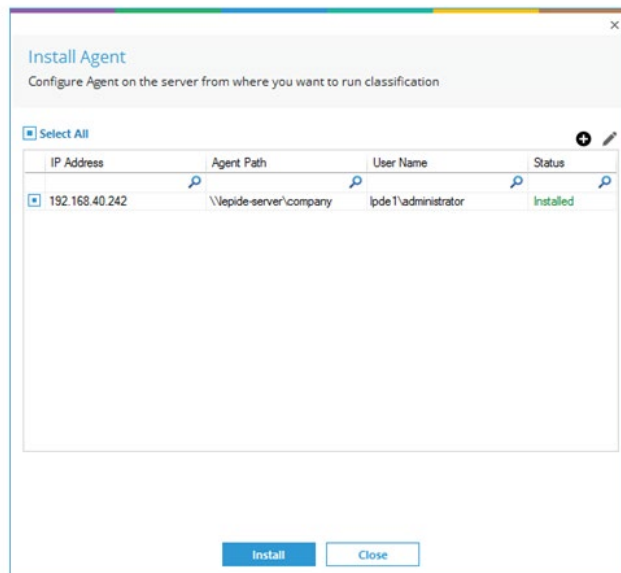

Figure 30: Add Server Dialog Box

- Enter the server information as follows:

  – **IP address**: The IP address of the server from where you want to run Classification, it can be the Lepide Data Security Platform server or the File server.

  – **Agent Path:** Create a shared folder on the Classification server. (\\server name or IP\Shared Folder) where the Agent will be installed.

  – **Username & Password**: Username with Read/Write access on the Shared folder.

- Click **OK** when finished to go back to the **Install Agent** dialog box:

Figure 31: Install Agent Dialog Box

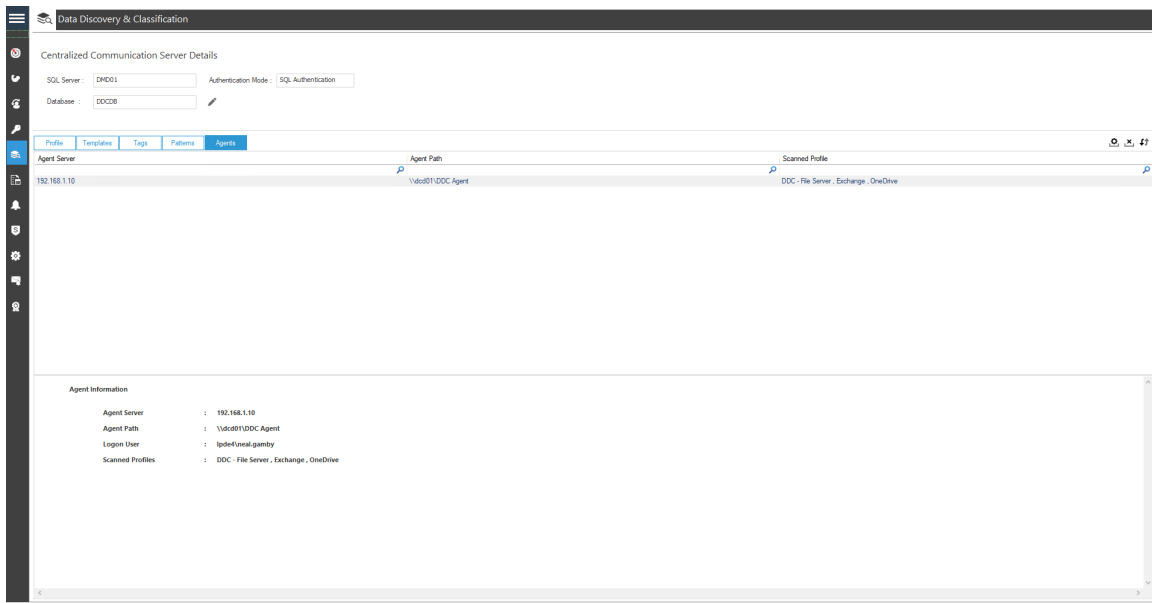- Select the server and click **Install** to install the agent:

Figure 32: Agent is Successfully Installed

- Once the installation is complete, click **Close** to return to the **Classification Server** dialog box
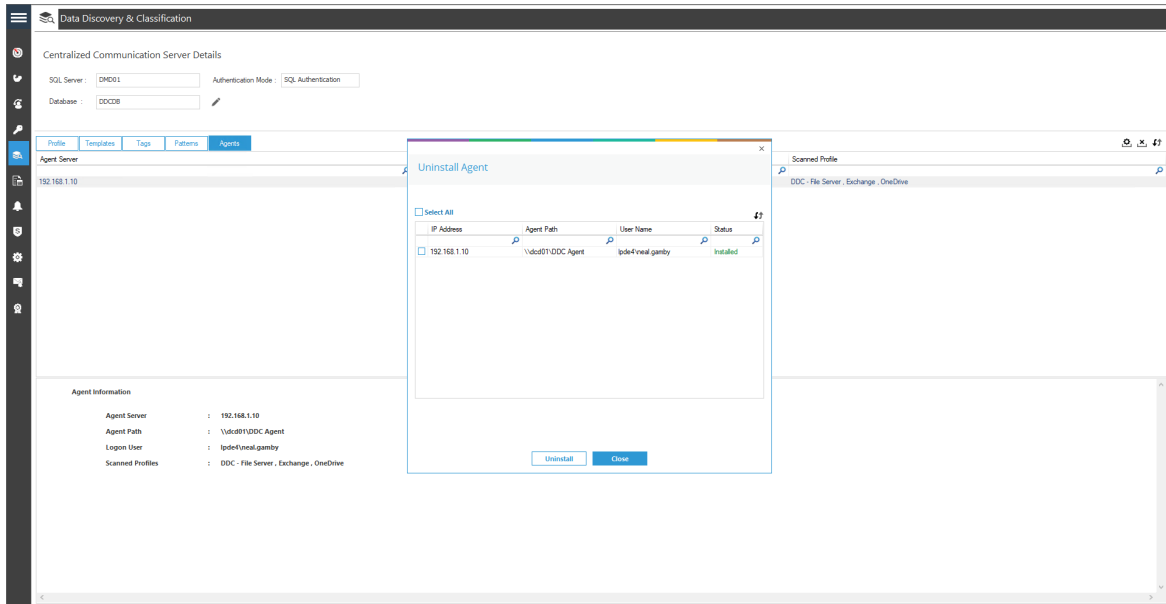- Click **Finish**

4.7 Uninstalling an Agent

- Click the Data Discovery & Classification icon to go to the Data Discovery & Classification screen and select the **Agents** tab:
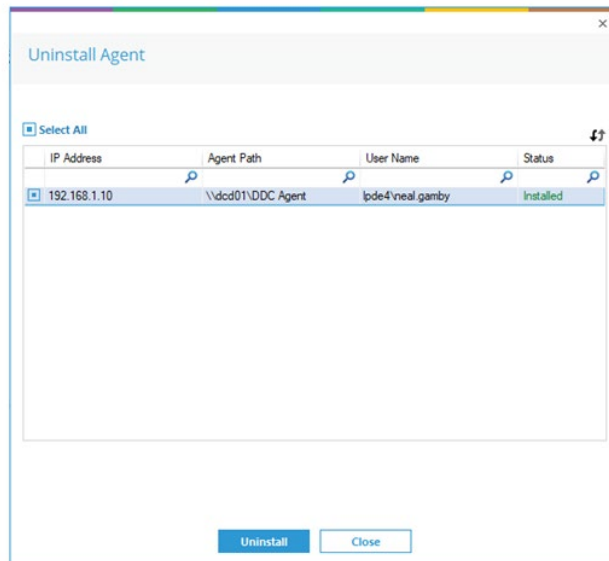

Figure 33: Data Discovery & Classification Screen

- Select the Agent you want to uninstall and click the Uninstall icon

- The Uninstall Agent dialog box is displayed:
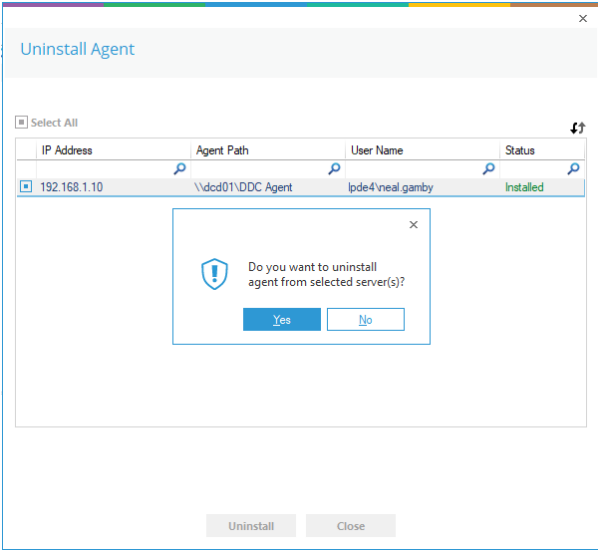
Figure 34: Uninstall Agent Dialog Box

- Choose the server to be uninstalled and click on **Uninstall**:

Figure 35: Uninstall Agent

- Select **Yes** to start uninstalling the Selected Server:
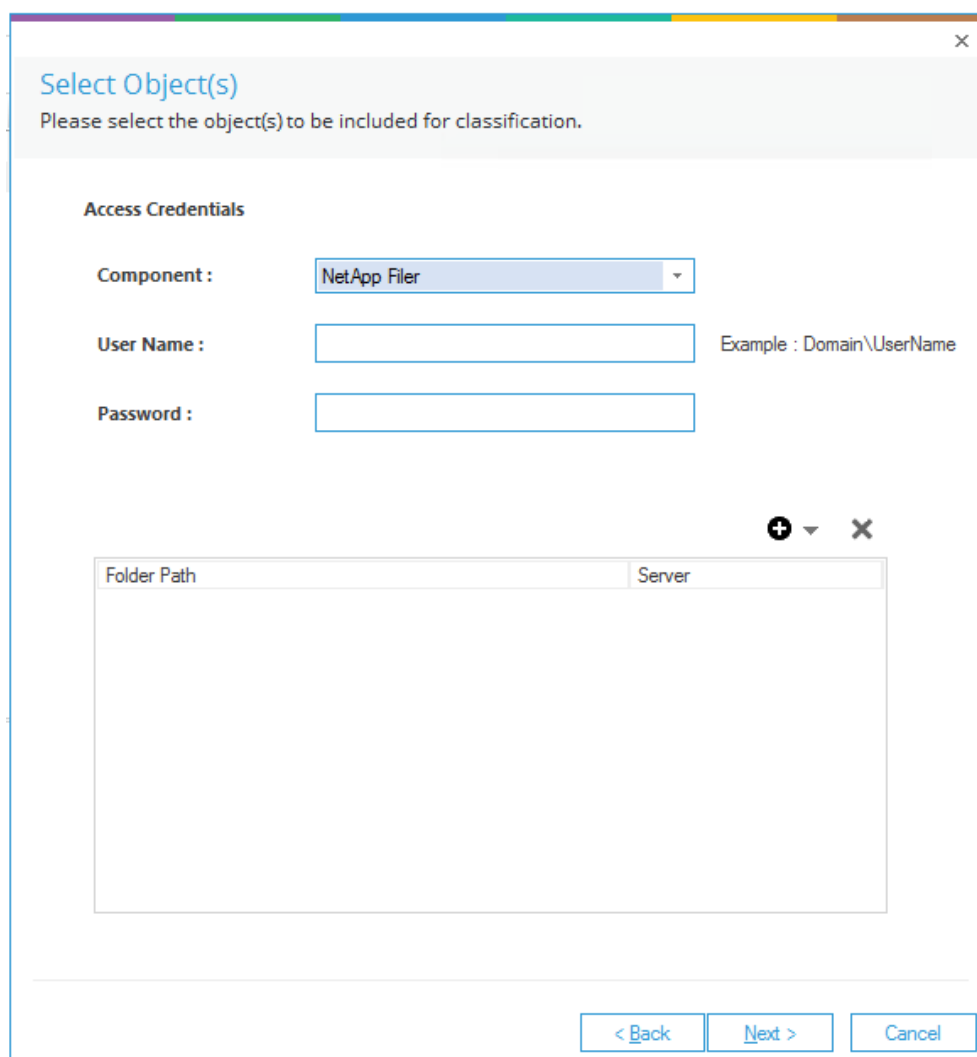

Figure 36: Confirm Uninstall of the Agent

- Click **Close**

5 Adding a Profile for Other Components

In Section 4 we added a new Profile for Windows File Server using the Wizard.

The **Select Objects** dialog box is the only step of the Wizard that changes depending on the component selected. Therefore, in the instructions that follow, only the **Select Objects** dialog box is shown, and the other steps are as described previously in Section 4.

5.1 Adding a Profile for NetApp Filer

Figure 37: Add a NetApp Filer Component

- Choose the **NetApp Filer** Component

- Add the **User Name** and **Password**

- To continue adding the Profile, follow Section 4.1 Folder Options given previously in this document.

5.2 Adding a Profile for Exchange OnPremise

Figure 38: Add an Exchange OnPremise Component

- Choose the **Exchange OnPremise** Component
- Add the **Server Name/IP**
- Add the **User Name** and **Password**
- To continue adding the Profile, follow Section 4.1 Folder Options given previously in this document (adding Mailbox Name rather than Folders)

5.3 Adding a Profile for Exchange Online

Figure 39: Select Objects

- Choose the **Exchange Online** Component
- Add the **Exchange Server Name**
- Add the **Tenant Name, Client ID** and **Secret Key**

  Click the ⃝? icon for information on how to generate the Client ID and Secret Key
- Click the Enumerate Mailboxes icon 🖳
- A list of Mailboxes will be displayed:

Figure 40: Select Objects with Mailboxes Enumerated

- Check the box next to the URL to select/deselect it

- To select all Mailboxes click the Select All Items icon

- To deselect all Mailboxes click the Remove Selection icon

- Click **Next** once the selections have been made

- To Add **On the fly Classification** select the check box:
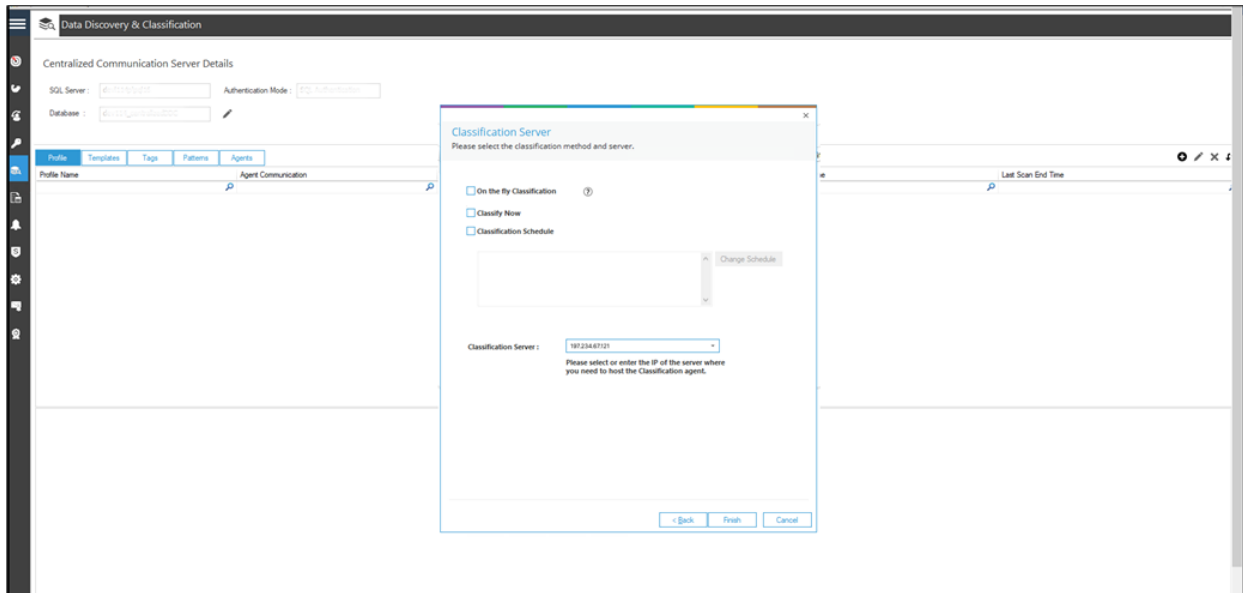
Figure 41: Selecting On the Fly Classification

- Enter the Thumbprint Value of the Tenant's Certificate in the text box:
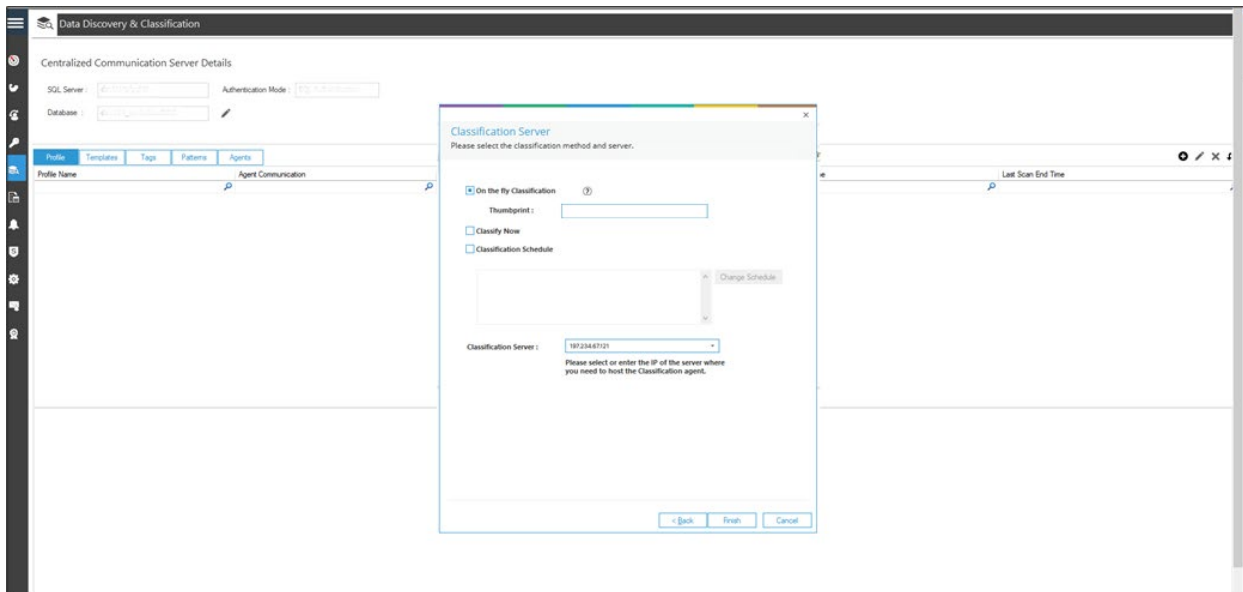
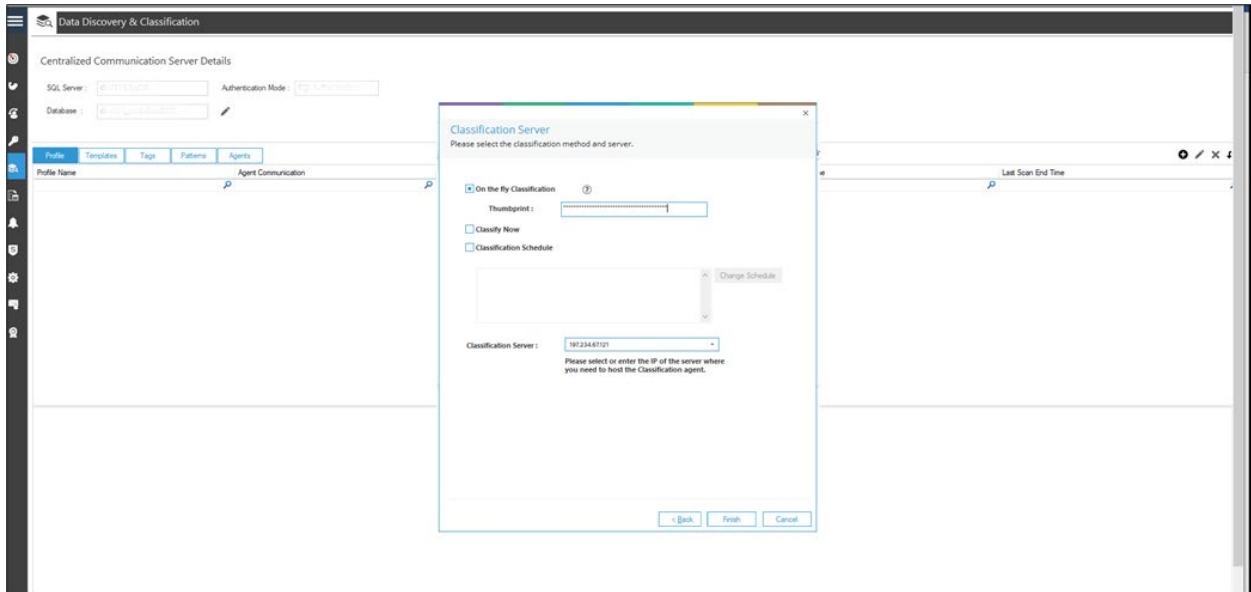*Figure 42: Add the Thumbprint Value*

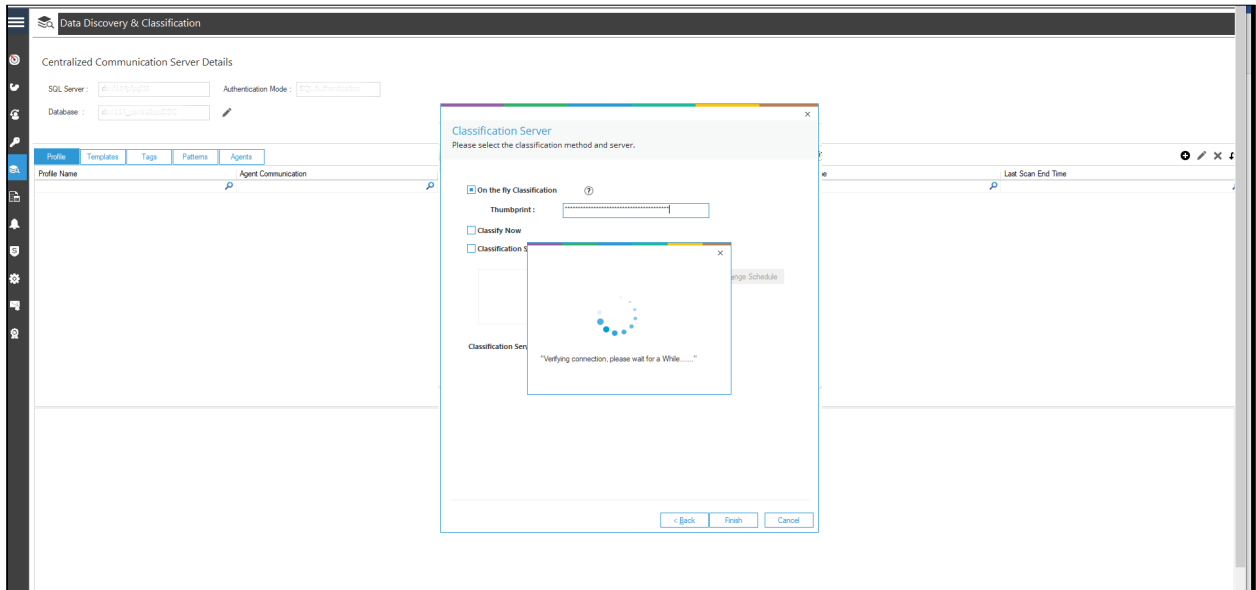*Figure 43: Thumbprint Added*

- Choose **Finish**

Figure 44: Credentials Validation

- The new DDC Profile will be set up

The Assign Templates dialog box is displayed

- To continue adding the Profile, follow the instructions given previously from page 15 of this document.

5.4 Adding a Profile for Windows Cluster

Figure 45: Add a Windows Cluster Component

- Choose the **Windows Cluster** Component
- Add the **User Name** and **Password**
- To continue adding the Profile, follow Section 4.1 Folder Options given previously in this document.

5.5 Adding a Profile for SharePoint OnPremise

Figure 46: Add a SharePoint OnPremise Component

- Choose the **SharePoint OnPremise** Component

- Add the **Central Admin URL**

- Add the **User Name** and **Password**

- To continue adding the Profile, follow Section 4.1 Folder Options given previously in this document (adding Site Collection URL rather than Folders)

5.6 Adding a Profile for SharePoint Online

Figure 47: Add a SharePoint Online Component

- Choose the **SharePoint Online** Component

- Add the **Central Admin URL**

- Add the **Client ID**

- Add the **Secret Key**

  Click **Info** for information on how to generate the Client ID and Secret Key

- Click the Enumerate Site Collections icon

- A list of Site Collection URL's will be displayed:
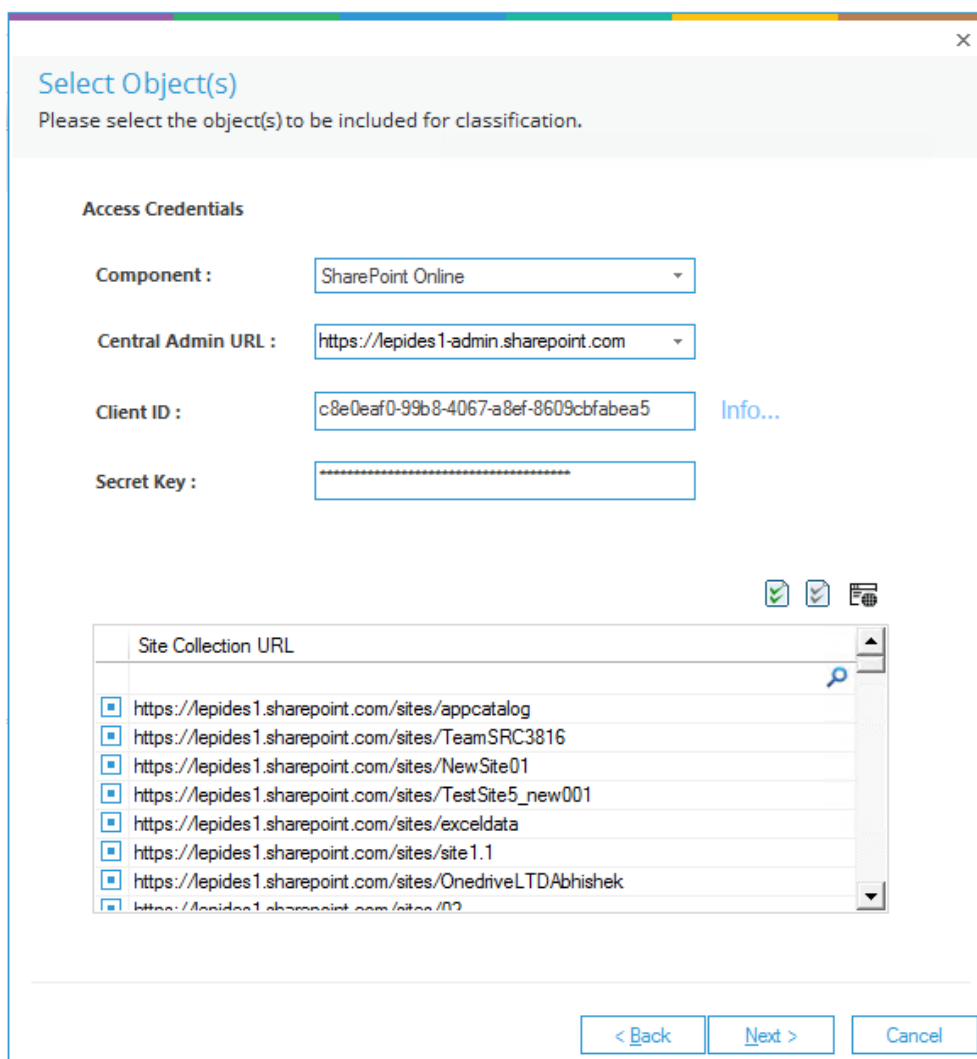
Figure 48: Select Objects with Site Collections Enumerated

- Check the box next to the URL to select/deselect it
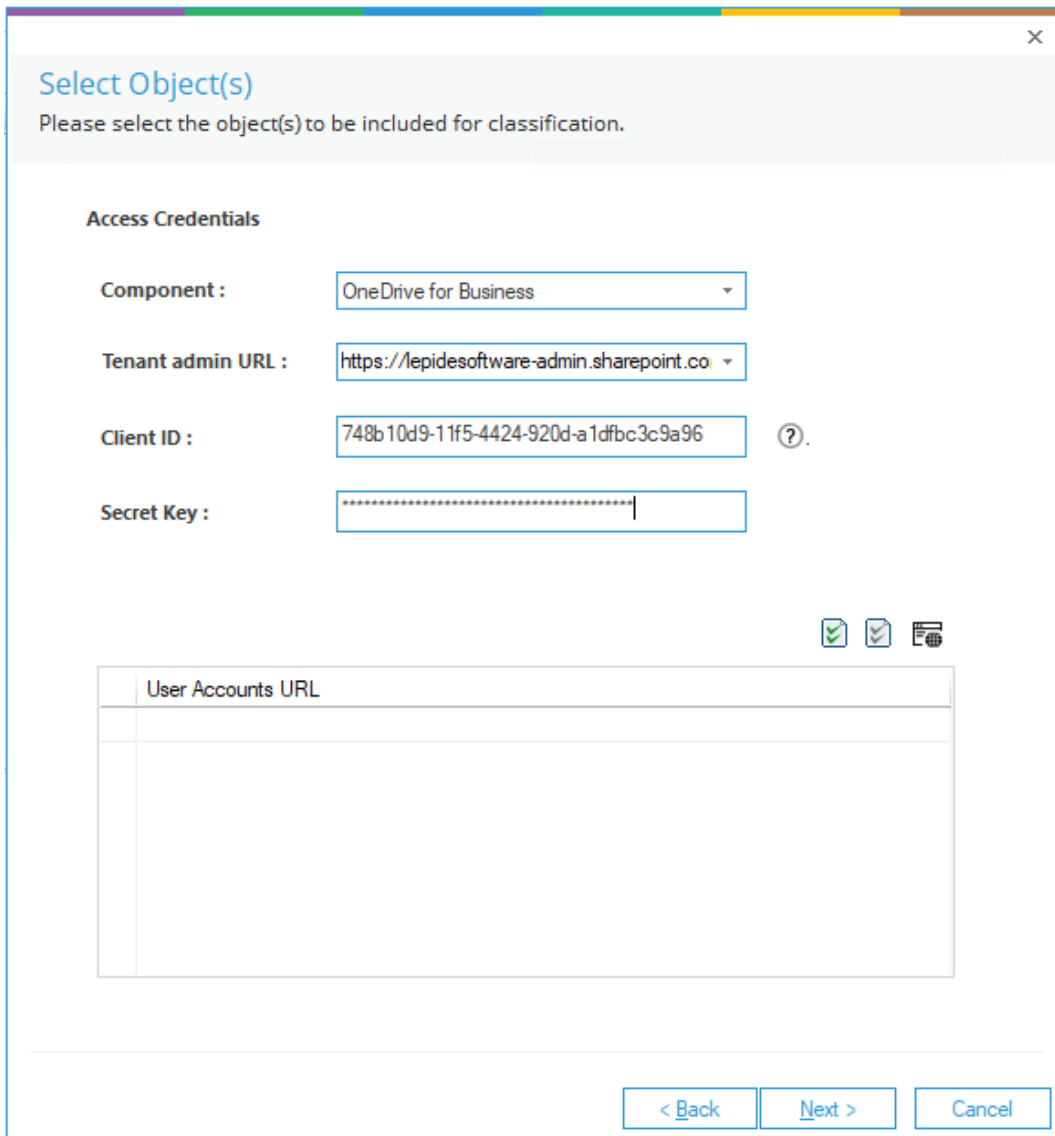
- To select all Site Collection URL's click the Select All Items icon

- To deselect all Site Collection URL's click the Remove Selection icon

- Click **Next** once the selections have been made

- The Assign Templates dialog box is displayed

- To continue adding the Profile, follow the instructions given previously from page 15 of this document.

5.7 Adding a Profile for OneDrive for Business

- Choose the **OneDrive for Business** Component

- Add the **Tenant admin URL**

- Add the **Client ID** and **Secret Key**

    Click the ⑦. for information on how to generate the Client ID and Secret Key

- Click the Enumerate User Accounts icon 🖳

Figure 49: Select Object(s)

- A list of User Account URL's will be displayed:

Figure 50: Select Objects with User Accounts Enumerated

- Check the box next to the URL to select/deselect it

- To select all User Account URL's click the Select All Items icon

- To deselect all User Account URL's click the Remove Selection icon

- Click **Next** once the selections have been made

- The Assign Templates dialog box is displayed
- To continue adding the Profile, follow the instructions given previously from page 15 of this document.

5.8 Adding a Profile for Dropbox

- Choose the **Dropbox** Component
- Add the **Server Name**
- Add the **Authentication**
- To continue adding the Profile, follow the Folder Options instructions given previously

Figure 51: Add a Dropbox Component

5.9 Adding a Profile for Linux

- Choose the **Linux** Component

- Add the **User Name** and **Password**

- To continue adding the Profile, follow Section 4.1 Folder Options given previously in this document

Figure 52: Add a Linux Component

5.10Adding a Profile for Unix

Figure 53: Add a Unix Component

- Choose the **Unix** Component

- Add the **User Name** and **Password**

- To continue adding the Profile, follow Section 4.1 Folder Options given previously in this document

5.11 Adding a Profile for EMC Isilon

- Choose the **EMC Isilon** Component

- Add the **User Name** and **Password**

- To continue adding the Profile, follow Section 4.1 Folder Options given previously in this document.

Figure 54: Add an EMC Isilon Component

6 Support

If you face any issues whilst installing, configuring, or using the solution, you can connect with our team using the contact information below.

Product experts

USA/Canada: +1(0)-800-814-0578

UK/Europe: +44 (0) -208-099-5403

Rest of the World: +91 (0) -991-004-9028

Technical gurus

USA/Canada: +1(0)-800-814-0578

UK/Europe: +44 (0) -208-099-5403

Rest of the World: +91(0)-991-085-4291

Alternatively, visit https://www.lepide.com/contactus.html to chat live with our team. You can also email your queries to the following addresses:

sales@Lepide.com, support@Lepide.com

To read more about the solution, visit https://www.lepide.com/data-security-platform/.

7 Trademarks

Lepide Data Security Platform, Lepide Data Security Platform App, Lepide Data Security Platform App Server, Lepide Data Security Platform (Web Console), Lepide Data Security Platform Logon/Logoff Audit Module, Lepide Data Security Platform for Active Directory, Lepide Data Security Platform for Group Policy Object, Lepide Data Security Platform for Exchange Server, Lepide Data Security Platform for SQL Server, Lepide Data Security Platform SharePoint, Lepide Object Restore Wizard, Lepide Active Directory Cleaner, Lepide User Password Expiration Reminder, and LiveFeed are registered trademarks of Lepide Software Pvt Ltd.

All other brand names, product names, logos, registered marks, service marks and trademarks (except above of Lepide Software Pvt. Ltd.) appearing in this document are the sole property of their respective owners. These are purely used for informational purposes only.

Microsoft®, Active Directory®, Group Policy Object®, Exchange Server®, Exchange Online®, SharePoint®, and SQL Server® are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

NetApp® is a trademark of NetApp, Inc., registered in the U.S. and/or other countries.