InfoWatch Data Discovery. Administration Guide

03/09/2026

© JSC “InfoWatch”

Tel/Fax +7 (495) 229-00-22

http://www.infowatch.com

Table of Contents

1 Who should read this guide 5

2 Set of documents 6

3 Technical support 7

4 Licensing 8

4.1 Trial license 8

4.2 Actions with licenses 8

5 System Settings 9

5.1 System Status 9

5.2 Integrations 11

5.2.1 Integration with Active Directory 12

Information about integrations 15

5.2.2 Integration with InfoWatch Traffic Monitor 16

5.2.3 Editing or Deleting Integration 19

5.3 Users 19

5.3.1 Creating New User 20

5.3.2 Adding New User from LDAP 21

5.3.3 Editing User Account 21

5.3.4 Viewing User Profile 22

5.3.5 Changing User Password 22

5.3.6 User Activation and Deactivation 23

5.3.7 Deleting Users 23

5.4 Roles 24

5.4.1 Creating and Configuring Role 24

5.4.2 Viewing and Editing Role 28

5.4.3 Revoking and Removing User Role 29

5.5 Audit 30

5.6 Password Settings 31

5.6.1 Requirements for the account name 33

5.7 Global Settings 33

5.7.1 Collect statistical data 33

5.7.2 Platform light version 33

5.8 Notification Services 34

6 Zones and Servers 36
6.1 Managing zones 36
6.1.1 Creating 36
6.1.2 Renaming 36
6.1.3 Deleting 37
6.1.4 Adding servers to a zone 37
6.1.5 Data export to This page has not yet been translated to English. 37
6.1.6 Security certificates for connecting to the XAPI service 39
Using This page has not yet been translated to English. certificates 40
Using your own certificates 40
6.2 Managing servers 41
6.2.1 Selecting a zone and deleting from a zone 42
6.2.2 Configuring storage 43
7 Authorization parameters 46
7.1 Yandex Disk 46
8 System Notifications 51
9 Configuring Autodeletion of Data Discovery Events in Traffic Monitor 52
10 OCR Extractor Selection and Default Operation Parameters 53
10.1 OCR extractor selection 53
10.2 Default OCR parameters 54
11 Log File Rotation 55
11.1 Scanning task services 55
11.2 Discoperator service 55
11.3 TPS Agent 55
11.4 Log file rotation 55
11.5 How to disable rotation and deletion of log files 56
11.6 How to enable rotation and deletion using logrotate and cron 56
12 Adding User Certificate 59
13 Receiving Events Tags from InfoWatch Traffic Monitor 60

This guide contains information on how to configure InfoWatch Data Discovery (also referred to as the System or Data Discovery), software designed to search for confidential information on shared network resources, workstations, servers, and in document storages.

1 Who should read this guide

This guide is intended for deployment engineers and security officers who are going to work with the System and administrate it. This guide is intended for users familiar with the basics of the Linux operating system and PostgreSQL DBMS.

2 Set of documents

The documentation includes:

- "InfoWatch Data Discovery. Installation Guide". The document describes how to install, update, and uninstall the System.
- "InfoWatch Data Discovery. Administration Guide". The document contains information on how to configure the System.
- "InfoWatch Data Discovery. User Guide". The document describes how to work with scanning tasks.

Related documentation for InfoWatch Traffic Monitor (also referred to as Traffic Monitor) includes:

- "InfoWatch Traffic Monitor. Installation Guide". The guide describes how to install, configure, update, and uninstall InfoWatch Traffic Monitor.
- "InfoWatch Traffic Monitor. Administration Guide". The guide describes how to administrate InfoWatch Traffic Monitor (database, back end).
- "InfoWatch Traffic Monitor. User Guide". The guide describes how to work with InfoWatch Traffic Monitor (set up configuration, export/import data, prepare policies for processing objects).
- "InfoWatch Traffic Monitor. Configuration Files Guide". Provides explanations for frequently used configuration files.

In addition to the primary System User Guide and Administration Guide, a Knowledge Base is available at lk.infowatch.ru, which includes troubleshooting materials and answers to questions. If you do not have access to the Knowledge Base, please contact the technical support service of InfoWatch at support@infowatch.com.

3 Technical support

If you have any problems or questions related to the operation of the System, please contact our technical support service:

- If you purchased the product from InfoWatch partner, please contact the technical support service of the partner
- If you purchased the product from InfoWatch directly, please contact InfoWatch technical support at support@infowatch.com.

You can also visit technical support section on our website: http://infowatch.com/support.

4 Licensing

In the **Settings** → **Licensing** section you can manage licenses.

Scanning tasks can be launched only if the System has an active license.

4.1 Trial license

By default, the System has a one-month trial license. The product installation date is the start date for this license.



4.2 Actions with licenses

To add a license:

1. Click + Add;
2. In the window that opens, add the license file.

To delete a license:

1. In the corner of the license tile, click ;
2. Select **Delete**;
3. Click **Ok**.

5 System Settings

The **Settings** → **System settings** section includes information on the following settings:

- **System Status**
- **Integrations**
- **Users**
- **Roles**
- **Audit**
- **Password Settings**
- **Global Settings**
- **Notification Services**

> **Note:**
>
> Data Discovery does not use the following sections:
>
> - **Data for analysis** – Data Discovery does not store and use data in partitions;
> - **Email notifications** – Data Discovery does not send email notifications;
> - **Filters** – Data Discovery uses separate filters which are not used by other products sections.

5.1 System Status

The monitoring subsystem checks the operation of all System elements. It analyses availability of the distributed network elements, information about the main hardware components and state of all System services.

The **Settings** → **System status** section contains panels with information about nodes of the Investigation Centre. Each panel contains the following information:

Field	Description
Label (name)	Label and name of the node in the format: `label(name)`
Status	Status of the node: • *Ready* – available and ready to work • *Not Ready* – not ready to work • *Unknown* – registered but not responding
Address	The IP address of Data Discovery node
Latest healthcheck status update	Date and time when information on the node status was received
Kernel version	OS kernel version
Operating System	The operating system of Data Discovery node

Field	Description
Memory	Server RAM: total and occupied size
CPU	Load of Data Discovery server CPU
Disk space (Clickhouse)	Total allocated and used disk space for storing Clickhouse DB data
Disk space (PostgreSQL)	Total allocated and used disk space for storing PostgreSQL DB data
Disk space (Root)	Total allocated and used space reserved for system data in the root directory
Disk space (NATS)	Total allocated and used disk space for storing NATS data
Disk space (Datastorage)	Total allocated and used disk space for storing binary data
The list of services of Data Discovery node	
Status	The service may have one of the following statuses: • *Running* – is running • *Stopped* – cannot get the status • *Not initialized* – is being launched • *Not found* – must be on the node, but is missing • *Error* – ended with an error
Name	Service name in the System (*containers*, *pods*)
Version	Service build version
View details	Displaying data on the service state, events, and internal messages

Note:

If other products of InfoWatch Investigation Centre are installed on the server, their services are also displayed in the list.

If Data Discovery is executing scanning tasks, then the list will also display services related to the running tasks.

Services of scanning tasks have a number of features:

- The service name includes the number of the scanning task the service belongs to. For example, for task number 42: `puller-42`, `sender-42`.

- Multiple services of the same type that belong to the same task will appear in the list with the same name.
  For example: If, to execute scanning task number 42, Data Discovery starts three `puller` services, then the list will display three `puller-42` entries.
- The list does not diplay services that has not started or stopped working.

Note:

* If one of the partitions has 10 GB or less free space left, then the Investigation Centre web interface will display a warning that the disk space limit has been exceeded. This warning cannot be disabled.
If the limit is exceeded:

- on the **Platform** `office` **node**, then integrations configured on this node will stop working;
- on the **Platform** `central` **node**, then all integrations in the cluster will be stopped.

To remove this restriction, in the **Settings** → **System Status** section, identify the node and partition that has little space left and free it up using the OS or product.

In Data Discovery, you can set a storage limit for scanned files (the default directory is `/opt/disctps`). For details, see "Zones and Servers".

If there is little free space left in the storage, you can speed up sending files to Traffic Monitor. For this, in the scanning task settings, reduce the number of scanners (`puller`) and increase the number of transmitters (`sender`).

All monitoring events in the System are logged.

To export the diagnostic data:

1. Click **Export system log files**. Log files of the running services and log files of the running scanning tasks from all Data Discovery nodes will be added to the diagnostic data. Exporting will begin once the diagnostic data has been generated.

   Note:

   If you need to get log files for a scanning task (e.g., log files of the Puller service instances), then before exporting the diagnostic data, make sure that the task is running.

2. Download the ZIP archive that contains TXT files with the logs.

To download diagnostic data for a separate service:

1. Click next to the required service. The log file will be generated and then loaded.
2. Download the ZIP archive that contains TXT files with the service logs.

5.2 Integrations

In the **Settings** → **Integrations** section, you can configure integrations with data sources:

- Active Directory – to get company's up-to-date directory tree Active Directory and to create dossiers;
- InfoWatch Traffic Monitor – to upload to the Platform:
  - the Traffic Monitor configuration (protected objects), which Data Discovery uses for content analysis, that is detecting protected objects in extracted text contents;
  - Traffic Monitor events, which you can add to investigations, as well as to receive data used by other products.

With Active Directory integration, you can:

- **Add new users to the System from LDAP directories**
  Users added from LDAP directories can log into the System using domain credentials. For details on adding users from LDAP directory, see "Adding New User from LDAP".
  Adding a new user from the LDAP directory does not require synchronization.

- **Add hosts from directories to scanning tasks** (see "InfoWatch Data Discovery. User Guide", "Adding and Configuring Hosts")

- **Add persons and groups from LDAP directories to the System**
  - The list of persons will be uploaded to the **Dossier** section where you will be able to view information about each person;
  - Persons added from LDAP directories will have a SID contact type;
  - The System will be able to identify persons and groups by their SIDs. Their names will be displayed in the table with information about files (see "InfoWatch Data Discovery. User Guide", "Table with File Information").

  To add the data to the System, synchronization is required.

For details on configuring integrations, see the following articles:

- "Integration with Active Directory";
- "Integration with InfoWatch Traffic Monitor";
- "Editing or Deleting Integration".

Warning!

Events received as a result of integration with InfoWatch Traffic Monitor are not used for Data Discovery functionality.

On how to configure data export to Traffic Monitor, see "Zones and Servers".

5.2.1 Integration with Active Directory

To add a new integration with Active Directory:

1. Go to **Settings** → **System settings** → **Integrations**.
2. Click + Add.
3. Select **Active Directory**.
4. Set parameter values:

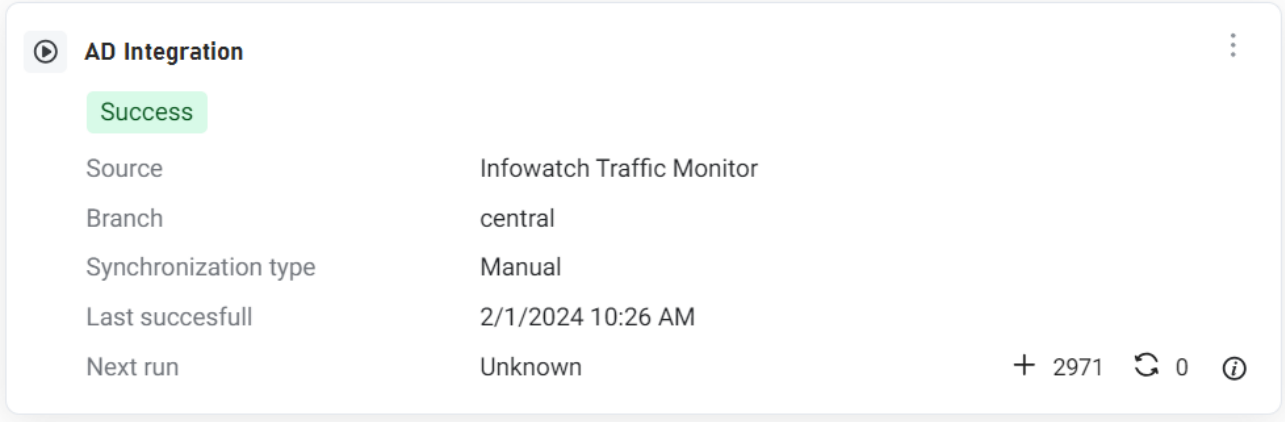
Parameter	Description
Integration name	Connection name which is used in the Console
Branch	Label of the Data Discovery node, which is used for connecting to Active Directory
Source address	The network address of Active Directory server
Connection type	Select the type of connection to be established: • Unsecure connection (Plain) • Secure LDAP connection (StartTLS) • Secure LDAPS connection (TLS) **Warning!** A connection without certificates is not secure.
Certificate	Upload the certificate file in PEM or DER format Parameter is used if secure connection type is selected
Global port	Port for connecting the global LDAP directory. By default the following values are used: • 3268: • for unsecure connection (Plain); • for secure LDAP connection (StartLTS); • 3269 for secure LDAPS connection (TLS). Click  to set default values.
Port	Port for connecting the domain local directory. By default the following values are used: • 389: • for unsecure connection (Plain) • for secure LDAP connection (StartLTS) • 636 for secure LDAPS connection (TLS). Click  to set default values.

LDAP query	Filtering attributes that are the full path to the specified directory. It is possible to specify only one directory in the organizational structure of Active Directory. You can use separate levels of the base hierarchy to optimize your search:
	`C – countryName`
	`O – organizationName`
	`OU – organizationalUnit`
	`DC – domainComponent`
	`CN – commonName`

Parameter	Description
	Example: To use the Users branch located in the company domain as a search base, enter: `cn=users,dc=company,dc=com` In case you need to exclude receiving specific groups from Active Directory, see how to do it in the article *Excluding groups from LDAP request using DN_Filter*
Login	Login to access the Active Directory server (user must have read access to the directory)
Password	Password to access the Active Directory server
Synchronization	Type of synchronization: manual or automatic. If automatic synchronization is selected, it is necessary to specify the values in the **Synchronization period** and **Repeat** parameters
Synchronization period	Interval of automatic synchronization. For example, if the **Synchronization period** parameter value is **Hourly** and the **Repeat** parameter value is **2**, then the synchronization starts every 2 hours.
Repeat	

5. Click **Test connection** to make sure that all parameter values are specified correctly. Wait for the **Successfully connected** message. If you see the following status:
   - **Could not connect to server** – check if the address and port are correct;
   - **Login or password is incorrect** – check if the login and password are correct.
6. Click:
   - **Save** – to save the integration without starting the synchronization;
   - **Save and start synchronization** – to save the integration and start the synchronization.

> **Warning!**
>
> Synchronization with Active Directory can create a high load on the domain servers and the network so we recommend running the first synchronization with Active Directory at night.



To start synchronization manually, click on the integration panel. The process may take some time.

Information about integrations

It is possible to synchronize with several LDAP directories. Each integration is displayed in a separate panel.

The bottom line of the panel shows:

- – number of persons, groups and workstations added during synchronizations;
- – number of persons, groups and workstations whose data was updated during the last synchronization;
- – detailed information on the number of persons added or updated during the last synchronization.

The synchronization status is displayed for each integration:

Status	Description
NOT SYNCHRONIZED	Integration is saved, synchronization has not been run yet
SYNCHRONIZING	Synchronization is in progress
SUCCESS	Synchronization completed successfully
DISCONNECTED	Synchronization has been terminated by user. Status also contains information on date and time of termination
ERROR	The status can indicate that: • an incorrect login, password, or address is specified; • an error occurred during synchronization.

5.2.2 Integration with InfoWatch Traffic Monitor

Integration with InfoWatch Traffic Monitor (also referred to as TM) allows you to load events, protected objects, as well as groups and persons into the Investigation Centre. For persons, you can create dossiers and initiate investigations (see "*InfoWatch Data Discovery. User Guide*", article "Creating, Editing and Deleting Dossier").

Data Discovery uses protected objects from Traffic Monitor for content analysis, that is detecting protected objects in extracted text contents.

> **Note:**
>
> For uploading protected objects into the Investigation Centre, we recommend using the same server with the Traffic Monitor database that was specified during installation of the Content Analysis Module. This ensures consistency of configurations in the module and Data Discovery. If the Investigation Centre does not have data about protected objects added to the module, then Data Discovery will not save content analysis results for those objects.
>
> Data Discovery does not use data on Traffic Monitor events. Other products installed in the Investigation Centre cluster can use this functionality.
>
> If a cluster does not include products, which use Traffic Monitor events, it is not necessary to store events in the Platform after configuring the integration.

When you log into the System for the first time, the integration list is empty.

To add a new synchronization with InfoWatch Traffic Monitor and start it:

1. Go to **Settings** → **System settings** → **Integrations**.
2. Click + Add.
3. Select **InfoWatch Traffic Monitor** and specify the following parameters:

Parameter	Description
Integration name	Connection name
Branch	A node for connection: main office or a branch
Source address	The network address of a working TM server
Token	A token required to receive data from the TM server (see the document "*InfoWatch Data Discovery. Installation Guide*", article "Traffic Monitor License and Plugin")
Data exchange protocol 2.0	In the current version, enabling the protocol disables loading of Traffic Monitor events. Data exchange protocol 2.0 does not affect Data Discovery functionality directly. Use the protocol to reduce the load, in cases when you need to receive only Traffic Monitor configuration, for example, to work with Content Analysis Module. Do not enable the protocol if some products in the cluster use Traffic Monitor events

Parameter	Description
Number of download threads	The number of active threads used to upload events and directories from Traffic Monitor. Increasing the number of threads will result in a higher load on Traffic Monitor database. We recommend specifying the number of threads not greater than the number of Traffic Monitor server cores.
Events loading interval (min.)	Time interval during which events occurred in Traffic Monitor. Data of these events will be loaded to the System as one thread. By default, 120 minutes. Minimizing this parameter allows you to reduce load on Traffic Monitor when sending data.
Use archiving options when downloading	When the parameter is enabled, events for the last N weeks (the N value is set in the archiving settings) up to the current date will be loaded. The default value is 12 weeks. If the parameter is disabled, events for all the time will be loaded
Not to download LDAP data	Enable the parameter, if you require downloading only data for analysis from Traffic Monitor. Enabling the parameter will exclude the download of data on persons, groups and workstations to Investigation centre from Traffic Monitor. In case the data has been already downloaded to Investigation centre, enabling the parameter will not remove them.
Download events details	When the parameter is enabled, all details will be further loaded to main data received from Traffic Monitor. The parameter allows you to enable a deferred loading.
Deferred loading	When the parameter is enabled, loading of detailed data is to be done at the time specified below
Time of deferred loading	Time to start loading of detailed data
Synchronization	*Manual* or *Automatic*. When selecting *Automatic,* you should specify *Frequency* and *Repetition*
Synchronization period	How often to start synchronization: *Minutely*, *Hourly*, *Daily*, *Weekly*
Repeat	Day of the week and time of the next synchronization

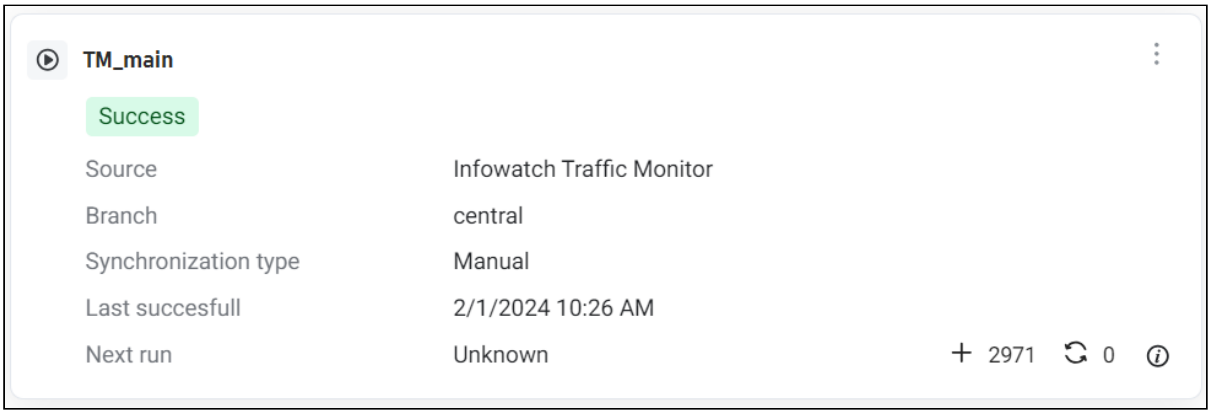
4. Click **Test connection** to check if you've specified all required parameters correctly. Wait until you see the message "*Successfully connected*".
5. Click:
    - **Save** – a new synchronization will wait for running with the NOT SYNCHRONIZED status or
    - **Save and start synchronization** to start the synchronization process.

In the central area, you can see the list of integrations with Traffic Monitor servers.

In the lower line, you will find:

- ⊞ – the number of entities added during the latest synchronization

- – the number of entities whose data was updated during the last synchronization
- – detailed information on synchronization results.



To set up auto synchronization on schedule (for example, at night in order to reduce time costs and network load):

1. Click and select **Edit**.
2. Select synchronization type: **Auto**.
3. Specify synchronization period: **Daily**.
4. Specify the time to repeat: **0:00**.
5. Click **Save**. The System will run synchronization in accordance with schedule.

> **Important!**
>
> When planning the initial synchronization with InfoWatch Traffic Monitor, keep in mind that a large number of events from the database can lower the channel capacity and general performance of the server and web console, so we recommend running synchronization with TM outside of working hours (at night).

If you require running one more synchronization with the same parameters:

1. Click .
2. Wait until synchronization is complete. The process can take a long time (e.g., if the data transfer rate over the network is low or a large amount of data is transferred).
3. After successful synchronization with TM, go to **Events** sections and make sure that the data is received and displayed in the table.

Possible synchronization statuses and their description:

Synchronization status	Description
NOT SYNCHRONIZED	Synchronization has been saved and can be started later
SYNCHRONIZING	Synchronization in process
SUCCESS	Synchronization completed successfully

Synchronization status	Description
DISCONNECTED	Synchronization has been terminated by user. Status also contains information on date and time of termination
ERROR	This status can indicate that: • You have entered an incorrect login name, password or address; • There is no plugin/license for using the product in TM; • An error occurred during synchronization.

5.2.3 Editing or Deleting Integration

To edit an existing integration:

1. In the upper-right corner, click  and select **Edit**.
2. Make the changes.
3. Click **Save**:
    - To create a new integration with the specified settings instead of the current one, click **Create a new one**. The current integration will be deleted.
    - If you need to change the settings of the current integration, click **Continue with this**.

To delete an integration:

1. In the upper-right corner, click .
2. Select **Delete** and click **Ok**.

> **Note:**
>
> The data received as a result of synchronizations will not be deleted if the integration is deleted.

5.3 Users

During the installation of the System, a pre-installed user is created. This user is called the Main Officer. To log into the System, use the default login and password:

Login	Password
officer	xxXX1234

When you first log into the System, you should change your password to a permanent one (see "Changing User Password").

The main page of the **Users** section contains a list of users who are allowed to work in the System:

- the Main Security Officer

- other Security Officers who report to the Main Officer

User's actions:

- Creating New User
- Adding New User from LDAP
- Editing User Account
- Viewing User Profile
- Changing User Password
- User Activation and Deactivation
- Deleting Users

5.3.1 Creating New User

The Main Officer can add new users who can also be Security Officers. For this:

1. Go to **Settings** → **System settings** → **Users**.
2. Click + Create.
3. Select **Create user**.
4. Specify values as described in the table below:

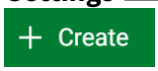
Parameter	Is required	Description
Photo	No	User photo
Username	Yes	Name of the new user (for username requirements, see " Password Settings").
Login	Yes	Name of the user account (for login requirements, see "Password Settings"). Login is not case-sensitive
Password	Yes	Account password (for password requirements, see "Password Settings")
Confirm password	Yes	Account password
Role	No	User role in the System. A user may have only one role.
Console language	No	Interface language for the new user: *Russian* or *English*
Display functions of installed products only	No	Switcher which makes sections and widgets of uninstalled products invisible
Contacts	No	Mobile phone number. You can add one or more values

Parameter	Is required	Description
	No	Email address. You can add one or more values

5. Click **Create**. The new user will be added to the System. When logging into the System for the first time, the user should change the account password (see "Changing User Password").

5.3.2 Adding New User from LDAP

The Main Officer can select new users from LDAP directory and grant them the rights of Security Officers. To add a new user from LDAP directory:

1. Go to **Settings**  -> **User settings** -> **Users**.
2. Click .
3. Select **Add from LDAP**.
4. Select an available LDAP server, integration with which is configured in the System (see "Integration with Active Directory").
5. In the search bar, start typing user name or login.
6. Select the user in search results and click **Add**.
7. Assign a role to the user (see "Editing User Account").

The user will be added to the list of the System users. The user can log into the System using domain credentials. Account login has the following format: `domain\login` or `login@domain` . Account login is not case-sensitive.

> **Note:**
>
> Security requirements (see "Password Settings") do not apply to users added from LDAP directory. Users from LDAP directory cannot change their passwords in the System.

> **Warning!**
>
> Changes to the account made in Active Directory will be available in the System only after the account is re-added: delete the account and then add it to the System again.

5.3.3 Editing User Account

To edit the user account:

1. Go to **Settings**  -> **System settings** -> **Users**.
2. Click  next to the user or in the upper-right corner of the user profile page.
3. Select **Edit**.
4. In the window that opens, edit one or more parameters. For parameter description, see "Creating New User".
5. Click **Save**. The new settings will be applied.

To assign a role to the user:

1. Go to **Settings** -> **System settings** -> **Users**.
2. Select a user.
3. In the user personal card, open the **Access Rights** tab and click the Apply role link.
4. In the **Role** field, select a role from a drop-down list. Users without roles cannot work in the System.
5. Click **Save**.
6. To apply changes, the user with this role should log out of the System and then log in again. To force the user session to end, select **End users active sessions**.
7. Click **Apply** to confirm. The new role will be assigned to the user.

5.3.4 Viewing User Profile

The profile page includes a user personal card:

- The **Personal info** tab displays the following data:
  - User photo;
  - Username;
  - User login name in the System;
  - User language in the System;
  - Contact data: phone numbers, email addresses;
  - Log of recent account activity;
  - *Display functions of installed products only* switcher, which makes sections and widgets for not installed products invisible.
- The **Access rights** tab displays the user rights in the System:
  - Role assigned to the user;
  - Available functions:
    - User access to products and settings;
    - Data the user may access.

  You can assign or revoke a role in the account editing window (see "Editing User Account").

User actions:

- Editing User Account
- Changing User Password
- User Activation and Deactivation

5.3.5 Changing User Password

To change your password that was provided for the first login or has been compromised:

1. Go to **Settings** → **System settings** → **Users** section.
2. Do either of the following:
   - Click next to the user;
   - Open the user's profile and click in the upper-right corner.
3. Select **Change Password**.
4. Fill in the following fields:

Field	Description
Old password	Enter your current password
New password	Enter a new password. The password must match the security settings (see "Password Settings")
Confirm password	Re-enter the new password

5. Click **Save**.

5.3.6 User Activation and Deactivation

The Main Officer can prevent the selected user from performing actions in the System. For this:

1. Go to **Settings** → **System settings** → **Users**.
2. Do either of the following:
   - Click next to the active user;
   - Click in the upper-right corner of the user's profile.
3. Select **Deactivate**. The user account will be blocked.

To unblock a previously blocked user account:

1. Go to **Settings** → **System settings** → **Users**.
2. Do either of the following:
   - Click next to the blocked user;
   - Click in the upper-right corner of the user's profile.
3. Select **Activate**. The unblocked user can now work in the System.

If other Security officers have access to user settings, they also can deactivate users.

5.3.7 Deleting Users

Main officer can delete user accounts. To do this:

1. Go to **Settings** → **System settings** → **Users**.
2. Do one of the following:
   - Click next to the active user;
   - Click in the upper-right corner of the user's profile.
3. Select **Delete**.
4. In the opened window, click **Ok**. User account and all its data will be permanently deleted from the System.

Note:

Main officer's account cannot be deleted.

5.4 Roles

System users can have different roles which define users' rights in the System. Users without a role cannot access the System.

A **User** working in the System can have only one **Role**. The **Role** includes **Functions** and **Data** with the following access types: ***full***, ***read-only*** and ***forbidden***.

Main page of **Settings** → **System settings** → **Roles** section contains all the roles which are used in the System. They include custom and two pre-installed roles:

1. **Main security officer** has full access to all System data, System settings and can mange roles including:
    - Creating new users;
    - Creating new roles and assigning roles to users.
2. **Security officer** has full access to System data except for the **Settings** section.

Find information about roles on the tabs:

- **Functions** – access of a user with this role to products, their functions and settings;
- **Data** – access of a user with this role to data;
- **Users** – on this tab, Main officer can assign roles to users or revoke roles from them.

Role can be assigned:

- When creating a new user;
- When creating a new role;
- By adding it from the list of existing roles in the user card in **Access rights** tab.

User actions:

- Creating and Configuring Role
- Viewing and Editing Role
- Revoking and Removing User Role

5.4.1 Creating and Configuring Role

The *Main Security Officer* role allows the user with this role to create new roles and assign them to users. If all users have different responsibilities, the Main Officer should create a separate role for each user. If user responsibilities change, you should either edit the role, or create a new role and assign it to the user instead of the obsolete one.

To create a role:

1. Go to **Settings** → **User Settings** → **Roles**.
2. Click **+ Create**.
3. Enter the role name. If necessary, add a description.
4. Click **Save**.
5. In the **Functions** tab, specify the access level to products and settings for users with this role:
    a. In the list of functions, select the checkmark as per products and sections the user should have full access to. By default, a user whose role does not contain open access can authorize in the System and work in **Main** and **Persons** sections.

Access to Investigation Centre functions

Access to *Events*, *Violation statistics* enables you to use functionality of **Events**, **Reports** and **Analytics** sections, including **Date**, **Senders**, **Recipients**, and **Event type** widgets.

Access to *Investigations* enables you to work in this section.

Access to *Event details* allows you to view additional event information in the **Events** section or navigate to Traffic Monitor to do this. With the *Event details* access, you can also use:

- **Attachment type**, **Top attachments**, **Event list (short)** widgets
- **Attachment name**, **Attachment existence**, **Email subject** filters

Access to *Violation details* enables you to unload data about details to the report and to use:

- **Level of violation**, **Verdict**, **Statistics**, **User decision widgets**, including corresponding filters
- Highlighting events and colouring widgets by violation type

Access to *Policies, Protected objects, Perimeters and Tags* enables you to use:

- **Policies**, **Protected objects**, **Statistics**, **DLP Tags** widgets, including corresponding filters, and also the filters **Perimeter** and **In perimeter**/**Out of perimeter**
- Additional columns in **Events** section

Access to *Dossier status* enables you to use the filter's attribute and **Status** widget in a person's dossier.

Access to functions of other products

Activity Monitor

Access to *Analytics* enables you to use:

- Statistics and activity events in **Analytics** and **Monitoring** sections
- Activity events filter attributes corresponding to available widgets
- Activity Monitor widgets in **Persons** section and on **Main** page
- Report functionality in **Analytics**, **Monitoring** and **Reports** sections

Access to *Screenshots* enables you to:

- Display screenshots in **Monitoring** and **Persons** sections
- Display and filter corresponding events

Access to *Audio records* enables you to:

- Display data on audio recordings in **Monitoring** and **Persons** sections
- Display and filter corresponding events

Access to *Videos records* enables you to:

- Display data on videos in **Monitoring** and **Persons** sections

Access to *Monitoring* provides you with monitoring and timeline in **Monitoring** and **Persons** sections.

Access to *Online monitoring* enables you to:

- View a person's workstation display
- Listen to audio on a workstation

- Lock a person's session
- Request a list of processes running on a person's workstation

Access to *Input text* enables you to:

- Display entered text in connected events,
- Filter events by presence of entered text

Access to *Web search queries* enables you to:

- Display search queries in connected events
- Filter events by presence of search queries

Access to *File operations* enables you to:

- Display file operations and connected events in **Monitoring** and **Dossier** sections
- Filter events by file operations
- Unload reports in **Monitoring** section

Data Access Tracker

Access to *Infrastructure* enables you to use the tabs of **Infrastructure** section:

- Data Access Tracker widgets on **Dashboard** tab (as well as on **Main** page);
- Reports on statuses and changes of directory service objects on **Directory audit** and **Group changes** tabs;
- Reports on access to mailboxes on **Access to mailboxes** tab;
- Reports on changes in access rights to filesystem objects on **File and folder access changes** tab.
- Reports on local administrator rights on **Local administrator rights** tab.

Data Discovery

Access to *File sharing* enables you to:

- Work in **File sharing section**:
  - On **Tasks** tab – to create and run scanning tasks;
  - On **File information**, **Duplicate files** and **Similar files** tabs – to analyze scanning results and grouping files by duplicates and similar ones;
- Work **in persons' dossiers on File sharing** tab, including the following sections:
  - **File owner**
  - **Available files**
- Unload reports with information from person's dossier subsections: **File owner** and **Available files**, as well as from the **File information**, **Duplicate files** and **Similar files** tabs in **File sharing** section.

Prediction

Access to *Dossier, Rating* enables you to:

- Work with persons rating in **Risks** section;
- Work in person's dossier on **Risk groups** and **Pattern map** tabs;
- Display widgets on the **Main** page on dashboards;
- Unload reports in **Risks** section and on **Risk groups** tab in person's dossier;
- Add Prediction block data in the report builder.

Vision

Access to *Graph* enables you to use graph functionality.

Access to *Violation statistics* enables you to work with:

- Statistics in **Analytics** and **Persons** sections, as well as to unload reports
- Widgets in Vision on **Main** page
- Report files in **Reports** section

Access to *Comments and files in the Dossier* enables you to:

- Use **Comments** and **Files** in a person's dossier
- Create an unloading or a screenshot of a widget or a graph and save it in the dossier

b. In **Settings** list, select:
   - ✓ to provide full access to settings. By default, access is denied;
   - 👁 to provide access to selected sections only;
   - ☐ to deny access to separate sections .

   Regardless of the selected access level, a user always has full access to personal information in their own profile.

6. In the **Data** tab, you can allow or restrict user access to data. To change user access rights to selected data, specify it in the following categories: *Persons*, *Groups*, *Branches*, *Contacts*, *Departments*. You can:
   a. ☐ – restrict access to selected data;
   b. ✓ – allow access to selected data;
   c. ✗ – deny access to selected data.

⚠ **Examples:**

If you set ✗ next to a person, the role will prevent the user from accessing any information on this person in the System.

If you set ✓ next to a person (or leave the default ☐ option for all persons), then the role will give the user full access to the person's data.

If you set ☐ next to a person and select other persons at the same time, this role will give the user access to information related to allowed participants (✓).

ⓘ **Note:**

The Main Officer should be careful when adding conditions as using different elements may result in narrowing of the available data instead of expanding it. This may happen when using logical AND between different categories.

Example:

a. The Main Officer configures access to *Group-1* OR *Group-2*.

b. The Main Officer configures access to *Person-1*, *Person-2*, and *Person-3*.
Now the user can access less data as the resulting condition is as follows:
(*Group-1* OR *Group-2*) AND (*Person-1* OR *Person-2* OR *Person-3*).

To configure a shared limitation for both groups and persons, create a cumulative group in Traffic Monitor, which would include all the groups and persons needed. Use it in the access settings.

To test the correctness of newly added visibility scopes, assign similar conditions in *Single filter* and then check results with the widgets used.

7. If you want to deny access to all data, select  **Prohibit access to data**.

> **Note:**
>
> For correct work with Data Discovery, do not prohibit access to data for a user.

8. In the **Users** tab, select a user without a role to whom the role should be assigned. For this:
   a. Click **+ Add**.
   b. From the **Users** field, select the required user.
   c. Click **Grant**. The role will be assigned to the user.
9. After you specify all the settings, click **Apply**.
10. For the changes to be applied, the user should log out of the System and then log in again. To forcibly end sessions for users with this role, select the  **End users active sessions** option and then click **Apply**.

5.4.2 Viewing and Editing Role

The Main Officer can view and edit roles assigned to users in their profiles. For this:

1. Go to  **Settings** → **System settings** → **Roles**.
2. Click the user whose profile you want to view. On the profile page, you can see the user card with the user's personal data.
3. Open the **Access rights** tab. The tab contains information about the user role. You can also find there:
   - Functions – if the user has access to the product settings;
   - Data (visibility scopes) – data to which the user has access.
4. To change the user role, click  and select **Edit**.
5. In the window that opens, remove the current role by clicking .
6. Assign a new role by clicking the field and selecting the required role from the list.
7. Click **Save** to apply changes.
8. For the changes to be applied, the user should log out of the System and get authorized again. To forcibly end the session for users with this role, select **End users active sessions**, and then click **Apply**.

You may need to copy a role, for example, when staff is increasing or for audit purposes:

1. Go to  **Settings** → **System settings** → **Roles**.

2. Click on the panel of the role you need to copy and select **Copy role**. The following role parameters will be copied:
    - **Name** – each copying adds the **Copy** word to the original name until the name of the role becomes unique:
      Example: After copying, the **Officer** name becomes **Officer copy**.
    - **Description**;
    - **Access settings** – all settings (function and data elements) are copied;
    - **Creation date** – the current date and time are specified.

The list of users to whom the role is assigned is not copied.

5.4.3 Revoking and Removing User Role

The user cannot have more than one role at the same time. So, if you need to assign a new role to a user, you should first revoke the current role.

To revoke a role in the role card:

1. Go to **Settings** → **System settings** → **Roles**.
2. Select a role you want to revoke.
3. Open the **Users** tab.
4. Select next to users from whom you want to revoke the role.
5. Click **Delete**. The role will be revoked from selected users. Now you can assign a new role to these users.
6. Click **Apply**.

To revoke a role in the user card:

1. Go to **Settings** → **System settings** → **Users**.
2. Do either of the following:
    - Click on the panel of the user from whom you want to revoke the role.
    - Select the user and then click in the user's card.
3. In the menu, select **Edit**.
4. In the **Roles** field, revoke the role by clicking .
5. Click **Save**.
6. For the changes to be applied, the user with this role should log out of the System and then log in again. To forcibly end sessions for this user, select **End users active sessions**.
7. Click **Apply** to confirm. The role will be revoked from this user.

The Main Officer can remove all roles except for preset ones.

To remove a role from the System:

1. Go to **Settings** → **System settings** → **Roles**.
2. Do either of the following:
    - Click on the panel of the role you want to remove.
    - Select the role and then click .
3. In the menu, select **Delete**.

4. For the changes to be applied, the user with this role should log out of the System and then log in again. To forcibly end sessions for this user, select ☑ **End users active sessions**.
5. Click **Remove** to confirm. If the role has been assigned to some users, it will be revoked from those users and removed from the list of roles.

5.5 Audit

The section displays the log of users' actions in the System. Only the user who has access to the section can view its data (for more information, see *Roles*).

The log contains the following information:

- **User** who performed an action;
- **Action in the system** performed with an object;
- **Object type**: user, integration, dossier, role;
- **Event object** with which the action was performed (for example, a role or integration);
- **Branch** – the name of the branch where the audit event was detected;
- **Date** and time when the audit event was created;
- **Event source IP address** – IP address from which the action was performed.

> **Note:**
>
> The System keeps audit logs indefinitely. They cannot be deleted.

The following actions are audited:

Users	• Log in • Log out • Create • Change • Delete • Password change • Block user
Roles	• Create • Change • Delete
Integrations	• Create • Change • Delete • Run manually
Dossier	• View • Create • Change • Delete • Combine

In **Audit** section, you can:

- Unload audit events in CSV format;
- Move columns in the table;
- Sort columns content (by default, events are sorted by date, from the latest to the earliest);
- Switch to a user card in **User** section, when you click on the user;
- Add a filter to the panel or remove it;
- Sort events by selecting filters;
- Copy content of cells.

By default all filters are added to the panel:

- *Date*;
- *Action in the system*;
- *Event object*;
- *User*;
- *Object type*;
- *Branch*.

To manage filters on the panel, use the **Add/delete filter** button . Select the checkbox , to add the filter to the panel. Or deselect the checkbox to remove the filter.

You can sort events by:

- selecting the filter attribute on the panel above the table;
- selecting the filter attribute in the table.

To select the filter attributes in the table:

1. Point cursor to the cell.
2. In the window that opens, press **Add to filter** . The table content will be sorted according to the selected filter.

To copy the content of table cell:

1. Point the cursor on the cell content.
2. In the window that opens, point the cursor on the content in the row, then press **Copy** .

In the **Date** column the **Copy** button is displayed when you point the cursor on a cell content.

To unload the report of audit events:

1. Press **Save report** .
2. Go to the **Downloads** folder of your computer.

If you moved columns in the table, the unloading will still have the columns order by default. To view the report in Excel, you need to enable the UTF-8 encoding. For more details, see the article "How to enable UTF-8 encoding in Excel" in our knowledge base.

5.6 Password Settings

The Security Officer can set rules for generating account passwords and password expiration dates. This is needed to prevent unauthorized access to the System and protect users from compromising their account data to third parties. To set a rule:

1. Go to **Settings**  → **System settings** → **Password settings**.

2. You can see default password settings. To set new requirements for passwords, modify the following parameters:

Parameter	Description
Letters	Uppercase and lowercase letters of the Russian and Latin alphabets that can be used in the password. You can set the following values: - required ☑ – the password must include letters - optional ☐ – letters in the password are not required
Digits	Arabic numerals that can be used in the password. You can set the following values: - required ☑ – the password must include digits - optional ☐ – digits in the password are not required
Special symbols	A list of special characters the password can include. You can set the following values: - required ☑ – the password must include special characters - optional ☐ – special characters in the password are not required
Minimal length	The minimum password length
Latest N passwords cannot be used	Select the option to forbid using previous passwords as they could be compromised
Time before password expires	How long the password is valid. The user should change the password before the expiration date
Failed login attempts period	For how many minutes the user can enter invalid passwords
Number of invalid input attempts	The number of unsuccessful input attempts after which the account is blocked
User blocking period	For how many minutes the user account is blocked

3. Click **Save** to apply new parameters. If you need to restore original settings, click **Reset to default settings**.

> **Example:**
>
> `Failed login attempts period – 5.`

`Number of invalid input attempts – 4.`

`User blocking period – 15.`

If you set these settings, then, after 4 unsuccessful attempts to enter the password within 5 minutes, the user account is blocked for 15 minutes. In 15 minutes the user can enter the password again.

5.6.1 Requirements for the account name

The username should contain at least four characters.

The user login name:

- Must start with a letter and not end with a dot
- May consist of Latin letters, Arabic numerals and contain special characters: "_", "-", "–", "."

5.7 Global Settings

In the **Settings** → **Global settings** section, you can manage the collection of statistical data.

Warning!

The other settings in this section are not used for Data Discovery.

5.7.1 Collect statistical data

The statistics collection system collects and aggregates data on user activity in the Management Console. By user's decision this data can be downloaded and sent to the InfoWatch company to be processed. This is done to improve the quality of the product, to assess the need for functionality and to create new work scenarios.

Statistical data collection is enabled by default. To disable data collection, uncheck **Enable** and click **Save**.

Note:

The statistics collection system does not collect, analyse or transfer any personal data. All users' and employees' full names are replaced with internal system IDs.

5.7.2 Platform light version

Operating mode of Investigation Centre products, in which they do not get events from Traffic Monitor. For the System to work correctly, save Traffic Monitor integration settings again after enabling/disabling the mode.

The mode does not affect Data Discovery functionality directly. Do not enable the mode if some products in the cluster use Traffic Monitor events.

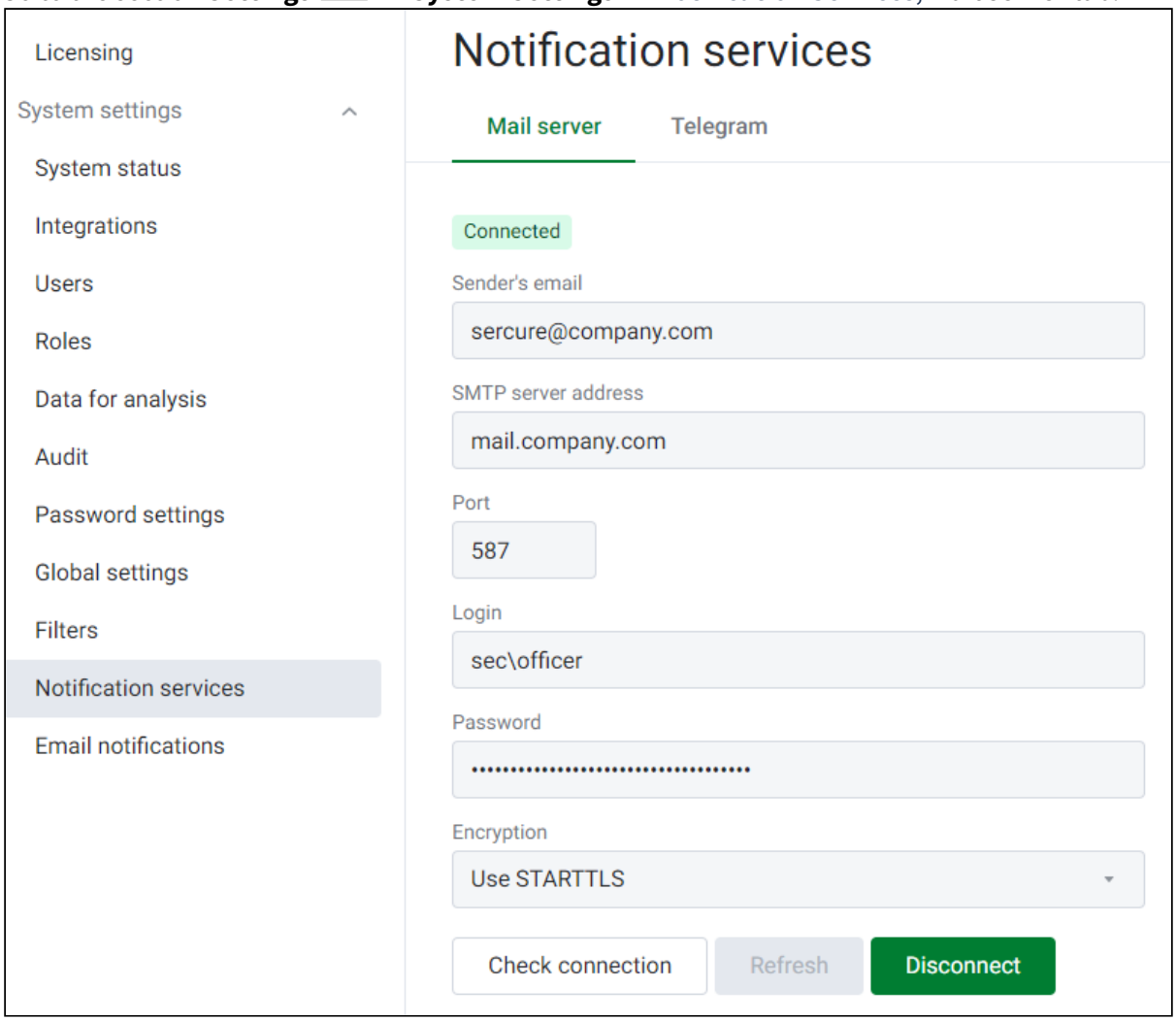
5.8 Notification Services

Data Discovery uses mail server for regular report sending.

> **Note:**
>
> Sending via **Telegram** is not supported in the current version.

To connect to the mail server:

1. Go to the section **Settings** → **System Settings** → **Notification Services**, **Mail server** tab.



2. Specify connection parameters:

Parameter	Description
Sender's email	The email address that will be used to send the notifications

Parameter	Description
SMTP server address	Mail server address that uses SMTP protocol. Specify the domain name or the IP address
Port	Port that is used to connect to the mail server
Login	Name of the account that has access to the mail server
Password	Password to log in to the mail server
Encryption	Specify the encryption protocol to connect to the mail server. Available protocols: • STARTTLS • TLS

All parameters are required.

3. After you have specified connection parameters, click **Connect** on the bottom.
   In case of a successful connection, the status will change its value to **Connected**.
   To check the availability of the server, you can click **Check connection** on the bottom.

To change connection settings or connect to a different mail server, make your changes and click **Refresh** on the bottom.

To disconnect from the mail server, click **Disconnect** on the bottom.

6 Zones and Servers

In the section **Settings** → **Products settings** → **Zones and servers**, you can manage

This page has not yet been translated to English.

zones and servers.

A zone is a group of

This page has not yet been translated to English.

servers. Dividing servers into zones allows you to:

- run task services only on specific servers. A zone can include several servers from the System cluster. Each server can be only in one zone.
- configure data export to
  This page has not yet been translated to English.
  for each server group. Each zone has its own settings for data export.

You can select the zone of a scanning task in the task settings (see

This page has not yet been translated to English.

. User Guide. Creating, Configuring and Running Tasks).

6.1 Managing zones

In the **Zones** tab:

- The list of zones is displayed on the left.
- The workspace on the right displays the settings for exporting data to
  This page has not yet been translated to English.
  and the list of zone servers.

6.1.1 Creating

To create a zone:

1. Go to the **Zones** tab.
2. Click + Create zone.
3. Enter the zone name.
4. Click **Create**.

6.1.2 Renaming

To rename a zone:

1. Go to the **Zones** tab.
2. The list of zones is displayed on the left. Click next to the zone name.
3. Select **Rename**.
4. Enter a new name.

5. Click **Save**.

6.1.3 Deleting

If you delete a zone, servers will be removed from the zone. The System on the servers will not be deleted.

You can't delete a zone if it is selected in tasks. First change the zone in the task settings or delete such tasks.

You cannot delete the preinstalled **Default** zone.

To delete a zone:

1. Go to the  **Zones** tab.
2. The list of zones is displayed on the left. Click  next to the zone name.
3. Select  **Delete**.
4. Click **Delete**.

6.1.4 Adding servers to a zone

To add servers to a zone:

1. Go to the  **Zones** tab.
2. The list of zones is displayed on the left. Select the zone.
3. Click  **Add server** in the workspace on the right.
4. The window that opens displays servers that are not added to any zone. Select the servers.
5. Click **Add**.
6. Restart the tasks in the zone so that the task services can run on the added servers.

6.1.5 Data export to

This page has not yet been translated to English.

To send file copies and information about files to

This page has not yet been translated to English.

for analysis while executing a scanning task, configure the data export.
This page has not yet been translated to English.

license and plugin are required to export data (see
This page has not yet been translated to English.

. *Installation Guide.* *Installing the License and Plugin*).

To configure data export:

1. Go to the  **Zones** tab.
2. The list of zones is displayed on the left. Select the zone.
3. Expand the **Data export to**
   This page has not yet been translated to English.

   panel.
4. Specify values for the parameters:

a. **XAPI Address:**
XAPI addresses of the
This page has not yet been translated to English.
servers. The servers should be in the same
This page has not yet been translated to English.
cluster.
This page has not yet been translated to English.
sends data to the servers from the list.
   - To add an address:
      i. Enter the address in IPv4 IP address format: "xxx.xxx.xxx.xxx".
      ii. Press Enter.
   - You can specify multiple addresses.

     > **Note:**
     >
     > To distribute the load and not to fill up only the first specified
     >
     > This page has not yet been translated to English.
     >
     > server,
     > This page has not yet been translated to English.
     >
     > uses the Round-robin distribution algorithm: the first file in the queue is sent to the first specified
     > This page has not yet been translated to English.
     >
     > server, the second file is sent to the next server, and so on.
     > The load distribution between multiple Content Analysis Modules is similar.

     If one address is unavailable, the System automatically starts using the next specified address.
   - This page has not yet been translated to English.
     version 7.7 or higher notifies if there is not enough space on the server to receive data from
     This page has not yet been translated to English.
     (see
     This page has not yet been translated to English.

     . *Configuration Files Guide*, the `FreeSpaceThresholdMountPoint` and `FreeSpaceThresholdGb` parameters in **xapi.conf**). In this case,
     This page has not yet been translated to English.
     switches to the next XAPI address or tries to send the data again in 1 minute.

b. **Token:**
Connection token.

> **Note:**
>
> To copy a connection token:
>
> i. In the
> This page has not yet been translated to English.
>
> Management Console, go to **Administration** -> **Plugins** section.

ii. At the top-left of the screen, select a preset plugin:
- *InfoWatch Crawler* for
This page has not yet been translated to English.

version 7.3 and lower;
- This page has not yet been translated to English.

for
This page has not yet been translated to English.

version higher than 7.3.

iii. Copy the token by clicking .

5. You can check the availability of
This page has not yet been translated to English.
servers by clicking the **Test connection** button.
6. Click **Apply**.
7. Restart the tasks in the zone. The tasks will get the new data export settings when started.

6.1.6 Security certificates for connecting to the XAPI service

This page has not yet been translated to English.

interacts with
This page has not yet been translated to English.

via the XAPI service. Both products use certificates and a key for SSL connection:

- **Trusted root certificate (CA)**
Default file name: `trusted_certificates`
- **Certificate signed by the root certificate (CA)**
Default file name: `server.pem`
- **Private key**
Default file name: `private_key.pem`

On the

This page has not yet been translated to English.

server, the certificate and key files **must** have the names from the list above and be located in the folder `/var/lib/iwplatform/data/current_config/discoperator/discovery/services_config/cert`.
On the
This page has not yet been translated to English.

server, the default file storage folder is `/opt/iw/tm5/etc/cert`.

> **Note:**
>
> In
>
> This page has not yet been translated to English.
>
> 7.14 the pre-installed certificates and key were updated. Take this into account if
> This page has not yet been translated to English.
>
> will interact with

This page has not yet been translated to English.

7.14 and newer versions.

Using
This page has not yet been translated to English.

certificates

To connect with

This page has not yet been translated to English.

, you can copy its certificates to the
This page has not yet been translated to English.

servers.
To do this:

1. On the
   This page has not yet been translated to English.
   server, go to the folder `/opt/iw/tm5/etc/cert`.
2. Copy the 2 security certificates and the private key to the
   This page has not yet been translated to English.
   server to `/var/lib/iwplatform/data/current_config/discoperator/discovery/services_config/cert`.
3. Rename the files as per the list above.
4. After that, go to the Investigation Centre web interface and restart all scanning tasks with the enabled mode **Send files to**
   This page has not yet been translated to English.

   for analysis.

Using your own certificates

You can use your own security certificates and key. In this case, **update them on all**

This page has not yet been translated to English.

and
This page has not yet been translated to English.

servers.
To do this:

1. Generate the certificates and key in a way that is convenient for you.
2. On the
   This page has not yet been translated to English.
   server:
   a. Copy the files to the folder `/opt/iw/tm5/etc/cert`.
   b. Specify the file paths in the configuration file `/opt/iw/tm5/etc/xapi.conf` in the section `"ThriftServers"`.
   c. Restart the service using the command:

```
iwtm restart xapi_xapi
```

3. On the
   This page has not yet been translated to English.
   server:
   a. Copy the files to the folder `/var/lib/iwplatform/data/current_config/discoperator/discovery/services_config/cert`.
   b. Rename the files as per the list above.
   c. After that, go to the Investigation Centre web interface and restart all scanning tasks with the enabled mode **Send files to**
      This page has not yet been translated to English.

      for analysis.

6.2 Managing servers

The  **Servers** tab displays a list of all

This page has not yet been translated to English.

servers. The list of servers for each zone is displayed on the  **Zones** tab.

Server panel displays the following information:

Server data	Description
 Name	Label and name of the server in the format: `label(name)`.
 Zone	The zone to which the server is added. There is no value if the server is not added to a zone. • Server with This page has not yet been translated to English. in the `central` mode is automatically added to the **Default** zone. • Servers with This page has not yet been translated to English. in the `office` mode are not automatically added to zones.
 Role	Roles describe the functionality and purposes of servers and let you know which components are installed: • **Scanner and sender** – a server where This page has not yet been translated to English. is installed. Used to scan hosts, obtain file metadata, create copies of files and send them to Traffic Monitor. If the zone does not include servers with this role, it is not possible to run scanning tasks in this zone; • **Text data extraction** – a server where This page has not yet been translated to English. is installed. Used to extract text data from files, as well as to recognize text in images using the OCR technology. The extracted texts will be sent for analysis to search for similar ones or to detect protected objects; • **Content analysis** – a server where the Content Analysis Module is installed. If the zone does not include servers with this role, the **Conduct content**

Server data	Description
	analysis of files mode is not available in the zone. This role cannot be disabled. If both This page has not yet been translated to English. and Content Analysis Module are installed on the server, then both roles of the server will be active.
 Storage	Data about the server storage space.

6.2.1 Selecting a zone and deleting from a zone

Note:

You cannot change the server zone or delete the server from a zone if the server storage has task files.

To get the number of task files in the server storage:

1. Connect to the server SSH console. Task files are stored in `PATH/NUMBER` directories, where:
   - `PATH` is the server storage path, by default: `/opt/disctpsd`;
   - `NUMBER` is the task number.
2. Run the `ls PATH/NUMBER -1 | wc -l` command for each task. For example: `ls /opt/disctps/7 -1 | wc -l`

To delete files from the storage, perform one of the actions for each task in the storage:

- Send the task files to
  This page has not yet been translated to English.
  :
  1. In the task settings:
     - enable the **Send files to**
       This page has not yet been translated to English.
       for analysis parameter;
     - specify the following values in the task settings:
       - **0** for the **Pullers count** parameter;
       - **N** for the **Senders count** parameter, where **N** is greater than or equal to the number of
         This page has not yet been translated to English.
         servers in the task zone.
  2. Start the task. The System sends files to
     This page has not yet been translated to English.
     and deletes them from storages, but does not download new files to zone storages.
  3. Wait until there are no task files left in the storage.
  4. Stop the task.
- Delete the task. Task files will be deleted and will not be sent to

This page has not yet been translated to English.

.

To add a server to a zone / to change the server zone:

1. Click  on the server panel.
2. Select  **Select zone**.
3. Select a zone from the list.
4. Click **Apply**.
5. Restart the tasks in the new server zone so that the task services can run on this server.

To remove a server from the zone:

1. Click  on the server panel.
2. Select  **Remove from zone**.
3. Click **Remove**. The System on the server will not be deleted.

6.2.2 Configuring storage

Each

This page has not yet been translated to English.

server has a storage for scanned files. All servers have the same storage path (by default, `/opt/disctps`). The storage path is specified when you install the product in the `central` mode.

> **Note:**
>
> We recommend that you place the server storage on a separate partition or disk.

> **Warning!**
>
> Only the files and directories that are added by the System should be in the storage. If you add other files or directories to the storage, the System may not function correctly.

During scanning, the System downloads a file to one of the zone storages. The System sends the file to

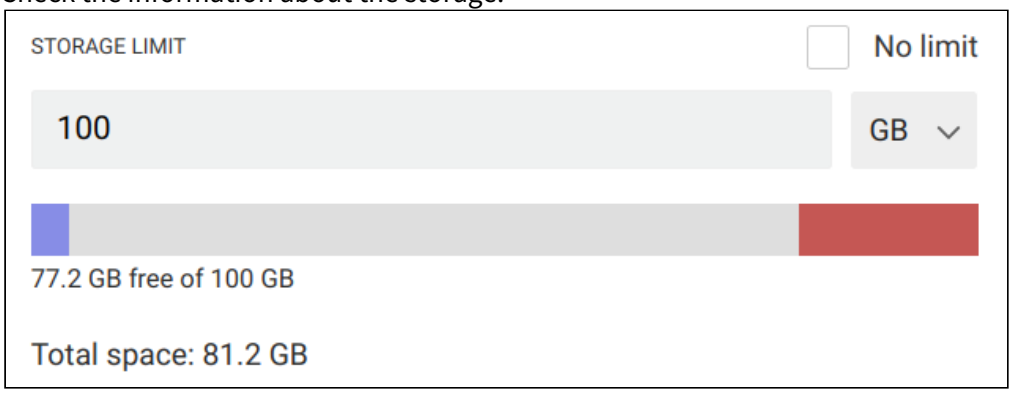
This page has not yet been translated to English.

and then deletes the file from the storage.

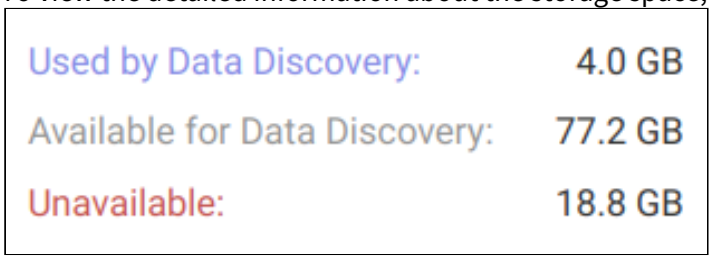
To configure a storage:

1. Click  on the server panel.
2. Select  **Set up roles**.
3. Select the role **Scanner and sender**.

4. Check the information about the storage:



   - the **Storage limit** field contains the current storage limit;
   - if the **No limit** setting is enabled, the storage limit is not used;
   - **Total space** = free disk space + space that is used by
     This page has not yet been translated to English.
     files.
5. To view the detailed information about the storage space, hover over the storage space indicator:



   - **Used by**
     This page has not yet been translated to English.

     is the space that is used by
     This page has not yet been translated to English.
     files;
   - **Available for**
     This page has not yet been translated to English.

     is the free disk space that
     This page has not yet been translated to English.
     can use;
   - **Unavailable** is displayed if the limit exceeds **Total space**.
6. Configure the limit:

> **Warning!**
>
> Set a storage limit if the System and storage are located in the same partition. We recommend that you set the limit so that 10 GB are always available in the partition.
>
> The System may not function correctly if there is not enough free space in the partition.

   - To set a limit:
     a. Disable the **No limit** setting.

b. Specify the limit in the **Storage limit** field and select the unit of measurement. The System stops downloading files to the storage if the used space in the storage reaches the limit. If space becomes available, the System continues downloading.

- If a limit is not required, enable the **No limit** setting. The System stops downloading files to the storage if the free space on the disk runs out. If space becomes available, the System continues downloading.

7. Click **Apply**. The System will start using the new storage settings.

7 Authorization parameters

In the section **Settings** → **Products settings** → **Authorization parameters**, you can add authorization parameters to connect to file storages. The added parameters will be available for selection when adding hosts with the specified connection type.

> **Note:**
>
> In the current version, adding authorization parameters is available only for the Yandex Disk storage.

7.1 Yandex Disk

To connect to the Yandex Disk storage,

This page has not yet been translated to English.

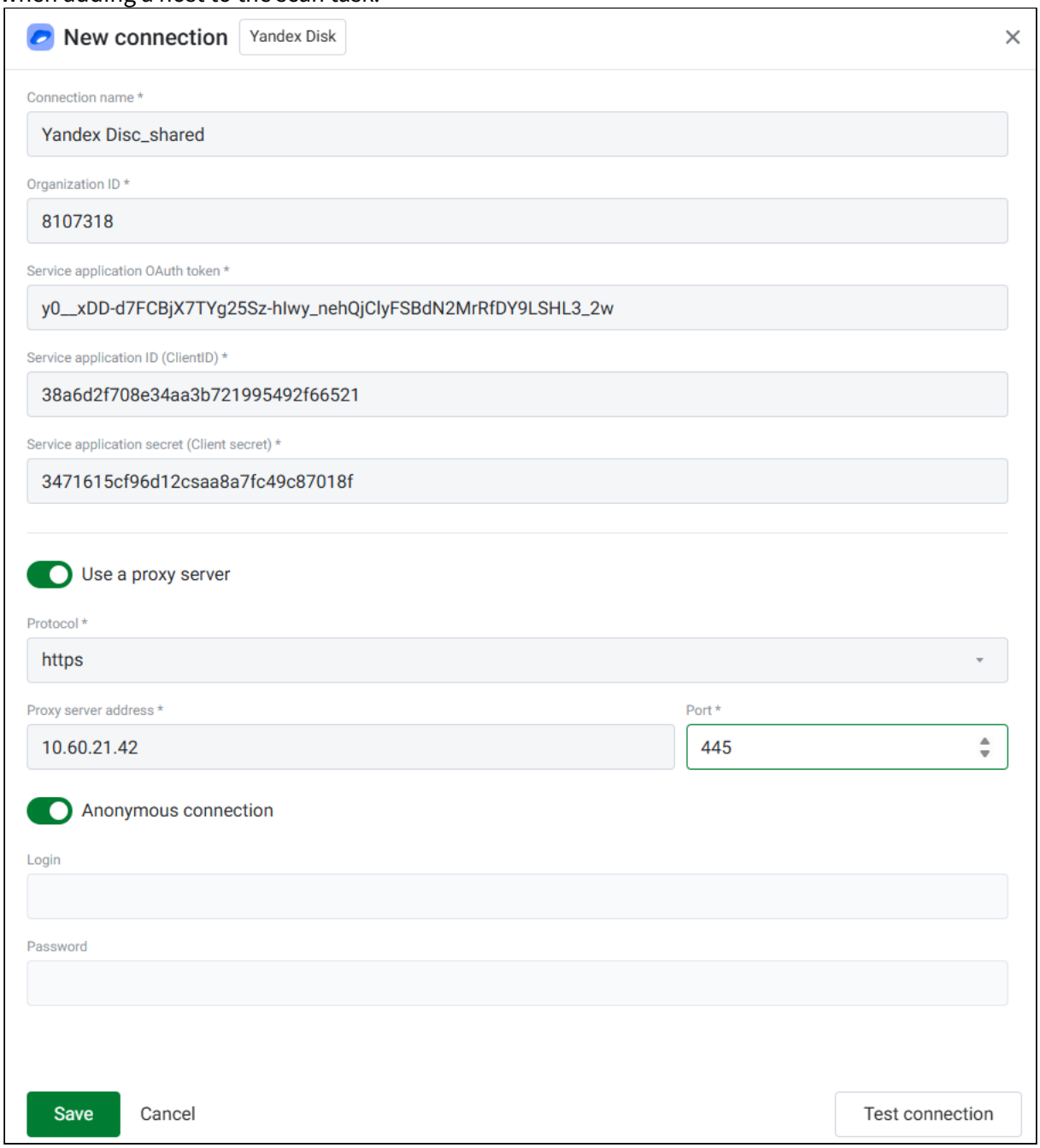
uses the Yandex Disk REST API.

> **Note:**
>
> Consider the following specifics of the Yandex Disk REST API:
>
> - You can only access user folders created inside the organization, in its domain.
> - There's no access to folders of third-party users added to the organization.
> - There's no access to the folders of the organization owner.

To add authorization parameters for connecting to the Yandex Disk storage:

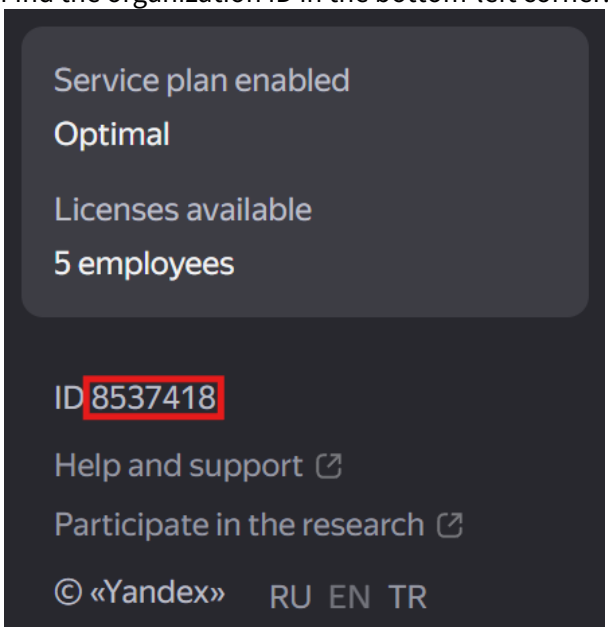
1. In the upper-right, click **Add**.

2. In the window that opens, specify the **Connection name** that will be displayed in the selection list when adding a host to the scan task.



3. In the field **Organization ID**, specify the required ID. To do this:
   a. In your browser, go to the Yandex 360 administrator panel.
   b. Log in as the organization owner.

c. Find the organization ID in the bottom-left corner. Use this value to fill out the field.



4. Create service applications in the Yandex OAuth web interface. To do this, follow the official instructions from Yandex.

Important!

By following the instructions, you'll create two applications:

- The main application that manages the list of service applications.
- The service application that
  This page has not yet been translated to English.
  will use for scanning.

When creating a service application, make sure to specify the following rights:

- `cloud_api:disk.app_folder`
- `cloud_api:disk.read`
- `cloud_api:disk.info`
- `directory:read_users`

Also, you need to specify the rights in the cURL request when registering the service application.

cURL request format:

```
curl -i --location \
--request POST 'https://api360.yandex.net/security/v1/org/<org_id>/
service_applications' \
--header 'Authorization: OAuth <owner_token_to_manage_service_app>' \
--header 'Content-Type: application/json' \
--data-raw '{ \
  "applications": [ \
    { \
      "id": "<OAuth_service_app_client_id>", \
```

```
      "scopes": [ \
        "cloud_api:disk.app_folder", \
        "cloud_api:disk.read", \
        "cloud_api:disk.info", \
        "directory:read_users" \
      ] \
    } \
  ] \
}'
```

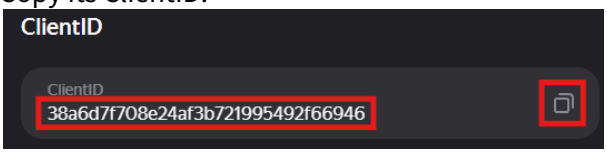
где:

- `<org_id>` – organization **ID**;
- `<owner_token_to_manage_service_app>` – **OAuth token**;
- `<OAuth_service_app_client_id>` – service application **ClientID**.

5. To fill out the field **Service application OAuth token**:
   a. Log in to the Yandex OAuth web interface as the organization owner.

   > **Note:**
   >
   > The access rights of the account on whose behalf the OAuth token was received determine which shared disks of the organization will be available for scanning.

   b. Select service application you've created before.
   c. Copy its ClientID.



   d. Follow the link that looks like this:

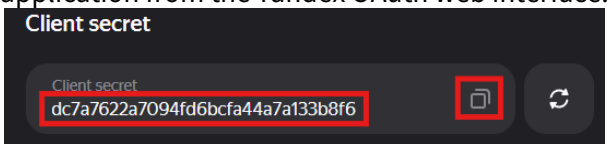
```
https://oauth.yandex.ru/authorize?response_type=token&client_id=<application_ID>
```

   In our example, it's:

```
https://oauth.yandex.ru/authorize?
response_type=token&client_id=38a6d7f708e24af3b721995492f66946
```

   e. If needed, log in as the organization owner or administrator. If the request is successful, you'll be redirected to the page with the token.
   f. In the authorization parameters window, copy the value from the field **Service application OAuth token**.
6. In the field **Service application ID (ClientID)**, specify the **ClientID** of the service application from the Yandex OAuth web interface.

7. In the field **Service application secret (Client secret)**, specify the **Client secret** of the service application from the Yandex OAuth web interface.

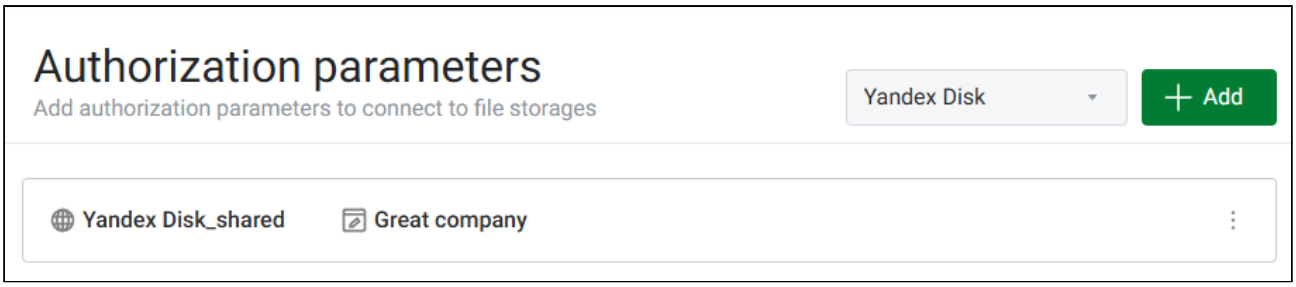


8. If you want
   This page has not yet been translated to English.
   to connect to the Yandex Disk storage via a proxy server:
   a. Enable the **Use a proxy server** toggle.
   b. Select **Protocol**.
   c. Specify the **proxy server address** and the **port number**.
   d. You can use the **Anonymous connection** by enabling the toggle or specify the **Login** and **Password**.
9. Before saving, you can test the connection with the storage by clicking the **Test connection** button in the bottom-right corner of the dialog box.
10. To complete adding the authorization parameters, click **Save**.

As a result, the connection will be added. You can select it in the field **Authorization parameters** when adding a host with the Yandex Disk connection type. To learn more about adding hosts, see

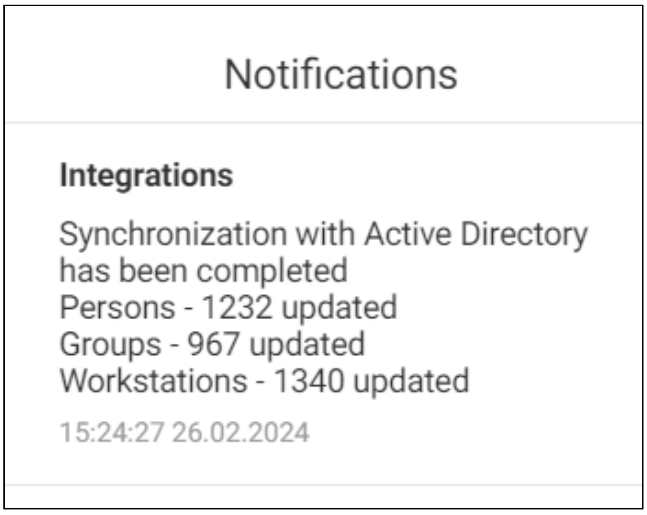
This page has not yet been translated to English.

. User Guide. Adding and Configuring Hosts).



8 System Notifications

To view Investigation Centre notifications, click .



Notification may contain the following information:

- The status of synchronization with sources initiated by user;
- Adding a license;
- Removing a license;
- The current number of dossiers loaded to the System;
- Possible disk space overflow;
- Soon password expiration.

9 Configuring Autodeletion of Data Discovery Events in Traffic Monitor

Traffic Monitor applies an atigarbage filter to Data Discovery events. **By default**, the filter removes Data Discovery events which have not triggered Traffic Monitor policies or which do not contain protected objects.

You can configure Data Discovery events deletion in `/opt/iw/tm5/etc/luaengined.conf` configuration file.

To define operation modes of the antigarbage filter, set a value for `DeleteDDEventsWithoutPolicy` parameter:

- `true` – to save only events which have triggered policies;
- `false` (default) – to save events which contain protected objects or which have triggered policies.

In case you want to disable deletion of Data Discovery events for certain servers, specify the servers in `DisableDeleteDDEventsForServers` parameter.

To apply changes:

1. Save the file.
2. Restart `iw_luaengined` service:

```
iwtm restart luaengined
```

Example:
You do not wish to apply the filter to events from *MainDD* and *DDcompany* servers, and, for the other servers, you want to save only events which have triggered policies.

Parameter values for luaengined.conf configuration file of the example:

```
"DeleteDDEventsWithoutPolicy": true,
"DisableDeleteDDEventsForServers": ["MainDD", "DDcompany"],
```

10 OCR Extractor Selection and Default Operation Parameters

This page has not yet been translated to English.

can recognize text in the extracted images using the OCR technology. To do this, you can use Tesseract or ABBYY FineReader. By default, ABBYY FineReader balanced mode is used. Tesseract and ABBYY FineReader are included in the package and don't require additional installation or license.

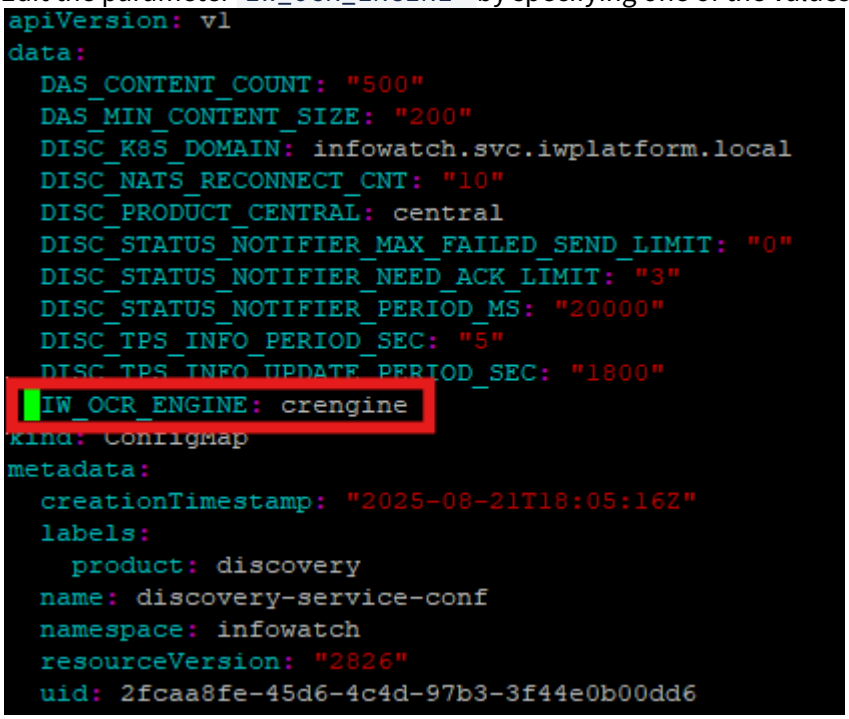
10.1 OCR extractor selection

To change the used OCR extractor or mode:

1. Execute the command:

   ```
   kubectl -n infowatch edit cm discovery-service-conf
   ```

   The command will open the configuration for editing in the text editor set as default in your OS.
2. Edit the parameter `IW_OCR_ENGINE` by specifying one of the values:



   - `tesseract` – to use Tesseract.
   - `crengine` – to use ABBYY FineReader in **balanced** recognition mode, that is the default option. The quality of text recognition in balanced mode is better that in fast mode.
   - `crenginefast` – to use ABBYY FineReader in **fast** recognition mode.
   - `none` – not to use OCR.
3. Save the changes.
   For example, to save the changes and exit the `vim` text editor, type:

   ```
   :wq
   ```

4. Relaunch the scanning tasks for the new settings to apply.

In case of distributed

This page has not yet been translated to English.

installation, you can perform these actions in any of the nodes.

10.2 Default OCR parameters

In the current version, you can't configure OCR parameters.

The following values are used by default:

Minimum image height in pixels	12
Minimum image length in pixels	140
Minimum image size in KB	40
Maximum image size in KB	1536

11 Log File Rotation

11.1 Scanning task services

To perform scanning tasks, the System uses the services: Puller, Sender, Ripper, CASI, Contents Processor, Task Execution Controller. Log files of service instances are stored in subdirectories of the `/var/log/infowatch/disc` directory on the server where the scanning task services were running.

Each time the task is started, the System adds a subdirectory to the `/var/log/infowatch/disc` directory for each service instance involved in the task execution. Log files of a service instance are added to the subdirectory. Each subdirectory has a unique identifier in its name.

11.2 Discoperator service

Log files of the Discoperator service are stored on the node with Data Discovery that is installed in the central mode, in the `/var/log/infowatch/disc` directory. This service is used for operation of the web interface.

Discoperator is constantly running and adds entries to the log file. If the service has been restarted, the System adds a subdirectory to the `/var/log/infowatch/disc` directory for the new service instance.

11.3 TPS Agent

The `/var/log/infowatch/disc` directory on each server stores the TPS Agent service log files. This service is used to manage the server storage.

The service is constantly running and adds entries to the log file. If the service has been restarted, the System adds a subdirectory to the `/var/log/infowatch/disc` directory for the new service instance.

11.4 Log file rotation

Services instances add a large number of entries to log files. The size of log files constantly increases, which leads to a decrease in free disk space.

You can free up disk space in the following ways:

- Stop all scanning tasks in Management Console, then delete subdirectories and log files manually from the `/var/log/infowatch/disc` directory. When you run the task again, the System:
  - creates new subdirectories with new identifiers;
  - creates log-files in subdirectories, to which entries about the task execution will be added.

> **Warning!**
>
> If you delete log files of a running task, entries will not be saved in the System until the task is stopped and started again.

> **Warning!**

It is not recommended to manually delete the directories and log files of the running Discoperator and TPS Agent services. The directories of these services have names like the following:

- `discoperator-central-5c5769b58d-77rq2`
- `tps-agent-f85x8`

- Use log file rotation (**enabled by default**).

For regular and automatic deletion of outdated entries, Data Discovery installer scripts enable log file rotation after the installation. Log file rotation is an automatic process for managing log files. During rotation, depending on the settings:

- New file is created that is a version of the original log file. Entries from the original log file are moved to the new file.
- Files that are versions of the original file are compressed.
- Outdated files are deleted.

11.5 How to disable rotation and deletion of log files

To disable log file rotation and deletion, remove:

- `/etc/logrotate.d/discovery` - a file with configuration for the `logrotate` utility;
- `/etc/cron.daily/discovery` - the `cron` planner task file, which is used for deleting log files.

Note:

The files listed above are created by the installer scripts by default. If you enabled rotation and deletion manually and created files with different names, then, in order to disable rotation and deletion, remove these files.

11.6 How to enable rotation and deletion using logrotate and cron

If rotation and deletion of log files are not enabled on the Data Discovery server, you can do it manually.

Warning!

You need to enable these settings on each Data Discovery server. Perform all actions as a user with root privileges.

To enable rotation of log files, configure the logrotate utility:

1. Check the contents of the `/etc/logrotate.conf` configuration file. The configuration file must contain the following line:

```
include /etc/logrotate.d
```

   If the line is missing, add the line to the configuration file.

2. In the **`/etc/logrotate.d`** directory, create a file for Data Discovery `logrotate` configuration (in our example, the file name is – `discovery`).
3. Open the created file for editing and enter the following configuration:

discovery

```
/var/log/infowatch/disc/*/*.log {
    daily
    copytruncate
    rotate 7
    notifempty
    compress
    missingok
}
```

Configuration description

Content	Description
/var/log/infowatch/disc/*/*.log {	rotate log files in the subdirectories of the **`/var/log/infowatch/disc`** directory
daily	perform daily rotation
copytruncate	at each rotation create a version of an original log file, then move all entries from the original file to the created version
rotate 7	keep the last 7 versions of each original log file, delete older versions of the log file
notifempty	do not rotate if there are no entries in the original log file
compress	compress each version of the original log file
missingok	do not report an error if the original log file was not found
}	

4. Save the changes. By default, `logrotate` runs automatically.
5. To check if the configuration is correct without performing a rotation, run the command:

```
logrotate -d /etc/logrotate.d/discovery
```

If the service instance was restarted, the new instance adds entries to a new log file in a new subdirectory. The old file will become empty after one rotation and will no longer be rotated. **Configure automatic deletion** of unused files and subdirectories. For this:

1. In the unpacked Data Discovery distribution kit (in our example, it is `/distr_dd`), go to the directory with the installer scripts: `/distr_dd/discovery/`**`installer_scripts`**

2. Copy the file with the uninstall settings to the directory with the `cron` planner daily tasks. For this use the command:

   ```
   cp delete_logs.cron /etc/cron.daily/discovery
   ```

   According to the task settings, `cron` will:
   - Delete log files and log file versions that were modified more than 7 days ago from subdirectories;
   - Delete empty subdirectories.

As a result of the configuration:

- `logrotate` regularly rotates Data Discovery log files and deletes outdated versions of log files;
- `cron` regularly deletes unused files and empty subdirectories.

12 Adding User Certificate

To prevent unauthorized access to the System and to increase the security of the data transmission channel, you can add a new user certificate for web service. You can add a certificate via command line at any time while using the System. To add a certificate:

1. Place two PEM files on the disk:
   - A certificate file (e.g., `test.crt`);
   - A private key file (e.g., `test.key`).
2. Navigate to the directory to which you unpacked the distribution archive during the installation.

   > **Note:**
   >
   > The Platform does not support using password-protected keys.
   >
   > Additionally, it is recommended to check hash sums of the certificate and the key by using the following commands:
   >
   > ```
   > openssl x509 -noout -modulus -in test.crt | openssl md5
   >
   > openssl rsa -noout -modulus -in test.key | openssl md5
   > ```
   >
   > The values in the outputs of the commands must be the same.

3. Run the command as a user with *root* privileges:

   ```
   ./setup.py setwebkeys --certificate=/folder/test.crt --privatekey=/folder/test.key
   ```

   In our example:
   - `/folder/test.crt` is the full path to the certificate file;
   - `/folder/test.key` is the full path to the private key file.

The last added certificate is considered valid. After adding or replacing a certificate, manually reload all browser tabs used by the System.

13 Receiving Events Tags from InfoWatch Traffic Monitor

Investigation Centre automatically receives events tags from Traffic Monitor when synchronizing. Thus the tags are available for viewing and their values cannot be changed in Investigation Centre. In the current version displaying tags in the web interface is enabled manually in the command line.

Important!

The tags from Traffic Monitor events are available **for the products on Platform version 2.4 and later only**, for example: Data Discovery 2.2 and compatible versions (see the information on the compatible versions in *InfoWatch Data Discovery. Installation guide. Installation/Update Types, Compatibility with Other Products, Hardware and Software Requirements*).

It is not recommended to enable the tags displaying in case there are products released **on Platform version 2.3 and earlier** in your Investigation Centre.

Take into consideration the events earlier uploaded, before updating the Platform to version 2.4 cannot contain tags. In this case you should re-synchronize with Traffic Monitor.

To enable tags:

1. On the *central* node, run:

```
kubectl exec -it -n infowatch $(kubectl get pods -n infowatch | grep postgresql-central |
awk '//{print $1;}') -- psql -U postgres -d configstorage -c "UPDATE key_values SET obj =
jsonb_set(Obj, '{Obj}', 'true'::jsonb) WHERE id = 'FT_tags'"
```

2. On *all* the nodes, run the service `factsstorage`:
   - On *central* node:

```
kubectl rollout restart deployment factsstorage-central -n infowatch
```

   - On *office* node:

```
kubectl rollout restart deployment factsstorage-office -n infowatch
```

3. Refresh the Investigation Centre browser page.

Then the following items will be available in Investigation Centre:

- New Single Filter parameter **DLP tags** (see *InfoWatch Data Discovery. User Guide. Single Filter)*;
- New widget **DLP tags** (see *InfoWatch Data Discovery. User Guide. Statistics Widgets*);
- New column in events table (see *InfoWatch Data Discovery. User Guide. Events Section*).