InfoWatch Data Discovery. Руководство пользователя

03/09/2026

© АО “ИнфоВотч”

Тел./Факс +7 (495) 229-00-22

http://www.infowatch.ru

СОДЕРЖАНИЕ

1 Аудитория 6

2 Комплект документов 7

3 Техническая поддержка пользователей 8

4 Обзор 9

5 Интерфейс Центра расследований 11

6 Главная страница 14

6.1 Дашборды и виджеты 14

6.1.1 Дашборды 14

6.1.2 Виджеты 15

Виджет Нарушения 16

Виджет Сотрудник под наблюдением 18

7 Единый фильтр 19

7.1 Фильтры 19

7.1.1 Фильтры в разделах Хранение файлов и Аналитика → Файлы 19

7.1.2 Фильтры в разделах События и Аналитика → Статистика нарушений 22

7.2 Работа с Единым фильтром 29

7.2.1 Добавление и удаление фильтров 29

7.2.2 Значения фильтров 30

Условие с несколькими фильтрами 30

Тип значения "поиск по шаблону" 30

Включение и исключение значений фильтра 31

Сброс значений фильтра 32

7.2.3 Сохранение и загрузка наборов фильтров 33

8 Раздел "События" 35

8.1 Добавление событий в расследование 37

8.2 Просмотр и выгрузка данных событий 38

8.2.1 Просмотр используемых контактов 38

8.2.2 Просмотр данных о вложениях 38

8.2.3 Выгрузка данных события 39

8.3 Фильтрация событий по персонам и контактам 40

9 Раздел "Аналитика" 42

9.1 Виджеты статистики ... 42
9.2 Выгрузка отчетов с данными статистики ... 44
9.3 Вкладка "Файлы" ... 46
10 Раздел "Хранение файлов" ... 49
10.1 Вкладка "Задачи" ... 49
10.1.1 Задача сканирования ... 49
Сканирование ресурсов ... 50
Событие Traffic Monitor ... 52
10.1.2 Создание, запуск и настройка задачи ... 54
Создание задачи ... 54
Запуск задачи сканирования ... 55
Настройки задачи ... 55
Добавление и настройка хостов ... 74
Добавление путей сканирования ... 92
Примеры настроек задачи, хоста и добавления путей сканирования ... 94
10.1.3 Редактирование задачи и хоста ... 105
10.1.4 Копирование задачи ... 105
10.1.5 Удаление задачи, хостов и путей сканирования ... 105
10.1.6 Статистика выполнения задач ... 107
Статистика по задачам сканирования ... 107
Статистика по хостам ... 110
Статистика по путям сканирования ... 112
10.2 Вкладка "Информация о файлах" ... 114
10.2.1 Таблица с информацией о файлах ... 115
Режимы отображения информации о файлах ... 115
Добавление записей в таблицу с детальной информацией ... 115
Структура таблицы с детальной информацией о файлах ... 117
Идентификация персон и групп ... 120
Действия с файлами ... 120
10.2.2 Настройка отображения таблицы ... 122
10.2.3 Поиск по примеру ... 123
Сброс результатов поиска ... 124
10.2.4 Выгрузка отчетов с информацией о файлах ... 124
10.2.5 Пример работы с таблицей ... 127
Настройка фильтров ... 127
Результаты фильтрации ... 129
10.3 Вкладка "Дубликаты файлов" ... 130
10.4 Вкладка "Похожие файлы" ... 133
11 Раздел "Персоны" ... 136

11.1 Просмотр досье персоны 136
11.2 Вкладка "Персональная информация" 138
11.3 Вкладка "Хранение файлов" 139
11.4 Создание, редактирование и удаление досье 139
11.5 Добавление контактов персоны 140
11.6 Изменение руководителя персоны 141
11.7 Объединение досье 141
12 Раздел "Расследования" 142
12.1 Создание и архивирование расследования 143
12.2 Добавление материалов в расследование 143
12.3 Совместное проведение расследования 144
12.4 Редактирование, выгрузка, печать и удаление расследования 144
13 Раздел "Отчеты" 146
13.1 Конструктор отчетов 146
13.1.1 Отчет в формате XLSX 148
13.1.2 Отчет по нарушениям в формате PDF или HTML 149
13.1.3 Отчет по персоне в формате PDF или HTML 152
13.2 Архив отчетов 153
13.3 Задачи на регулярную отправку 153
14 ПОЛЬЗОВАТЕЛЬСКОЕ ЛИЦЕНЗИОННОЕ СОГЛАШЕНИЕ 155
15 Глоссарий 160

В настоящем руководстве вы сможете найти сведения по работе и решению задач в InfoWatch Data Discovery (далее Система или Data Discovery) – программном обеспечении, предназначенном для поиска конфиденциальной информации на общих сетевых ресурсах, рабочих станциях, серверах и в хранилищах документов.

1 Аудитория

Данное руководство предназначено для инженеров внедрения и офицеров безопасности, которые будут работать с Системой и заниматься ее администрированием. Руководство рассчитано на пользователей, знакомых с основами работы в среде операционных систем Linux и MS Windows.

2 Комплект документов

В документацию входят:

- «InfoWatch Data Discovery. Руководство по установке». Документ содержит описание процесса установки, обновления и удаления Системы.
- «InfoWatch Data Discovery. Руководство администратора». Документ содержит информацию по настройке Системы.
- «InfoWatch Data Discovery. Руководство пользователя». Содержит описание работы с задачами сканирования.

Сопутствующая документация по системе InfoWatch Traffic Monitor (далее Traffic Monitor) включает в себя:

- «InfoWatch Traffic Monitor. Руководство по установке». Содержит описание установки, настройки, обновления и удаления системы InfoWatch Traffic Monitor.
- «InfoWatch Traffic Monitor. Руководство администратора». Содержит информацию по администрированию системы InfoWatch Traffic Monitor (база данных, серверная часть).
- «InfoWatch Traffic Monitor. Руководство пользователя». Содержит описание работы с InfoWatch Traffic Monitor (настройка конфигурации, экспорт/импорт данных, подготовка политик для обработки объектов).
- «InfoWatch Traffic Monitor. Справочник по конфигурационным файлам». Содержит пояснения к часто используемым конфигурационным файлам.

Кроме основных руководств для пользователей и администраторов Системы предусмотрена База знаний lk.infowatch.ru, которая включает материалы по устранению неисправностей и ответы на вопросы. В случае отсутствия доступа к Базе знаний, просьба обращаться в службу технической поддержки компании InfoWatch по адресу support@infowatch.com.

3 Техническая поддержка пользователей

При возникновении проблем и вопросов, связанных с работой Системы, вы можете обратиться в службу технической поддержки:

- если вы приобрели продукт у партнера компании InfoWatch, то обратитесь в службу технической поддержки партнера;
- если продукт приобретен у компании InfoWatch напрямую, то обратитесь в службу технической поддержки компании InfoWatch по адресу support@infowatch.com.

Вы также можете посетить раздел технической поддержки на нашем сайте: https://www.infowatch.ru/services/support.

4 Обзор

Центр расследований – веб-консоль для работы с продуктами InfoWatch. Вы можете установить в кластере как один, так и несколько продуктов. Установка всех продуктов позволяет получать наиболее полную информацию о нарушениях, персонах, рисках и т.д.

Data Discovery предназначен для поиска конфиденциальной информации на общих сетевых ресурсах, рабочих станциях, серверах и в хранилищах документов. Data Discovery работает следующим образом:

- подключается к различным ресурсам.
- сканирует файлы на ресурсах и собирает метаданные файлов.
- отправляет копии файлов или текстовые данные для анализа в:
    - InfoWatch Traffic Monitor;
    - InfoWatch Data Analysis Service (поиск похожих);
    - Модуль контентного анализа (обнаружение объектов защиты).
- отображает метаданные, результаты анализа и группировки просканированных файлов.

Используя Data Discovery, вы можете:

- подключаться к общим сетевым ресурсам и к распределенным файловым системам (DFS) по протоколу SMB:
    - подключаться к SMB-ресурсам, созданным на хостах под управлением MS Windows;
    - подключаться к SMB-ресурсам, созданным на хостах под управлением Linux;
- подключаться к рабочим станциям и серверам под управлением Linux с аутентификацией по протоколу SSH, в том числе используя безопасный метод с помощью SSH-ключей;
- подключаться к библиотекам документов в SharePoint Server 2016 по протоколу HTTPS;
- подключаться к файловым ресурсам по протоколам FTP/FTPS;
- подключаться к хранилищу данных Nextcloud по протоколу WebDAV;
- подключаться к системе управления контентом Alfresco Community Edition по протоколам HTTP/HTTPS, в том числе с подключением через прокси-сервер;
- подключаться к хранилищу Яндекс Диск, используя REST API Диска, в том числе с подключением через прокси-сервер;
- сканировать файлы на ресурсах:
    - обнаруживать новые и измененные файлы;
    - собирать метаданные файлов;
    - создавать копии файлов для анализа;
    - извлекать текстовые контенты из копий файлов для анализа их содержимого с помощью InfoWatch Data Analysis Service и Модуля контентного анализа;
    - извлекать из копий файлов изображения и распознавать на них текстовые данные с помощью технологии OCR для последующего анализа с помощью InfoWatch Data Analysis Service и Модуля контентного анализа;
- отправлять данные на сервер Traffic Monitor для анализа;
- просматривать статистику о выполнении задач сканирования;
- просматривать информацию о файлах, включая данные о прошлых сканированиях, что позволяет также просмотреть историю изменений файлов;
- искать дубликаты среди просканированных файлов;
- определять с помощью InfoWatch Data Analysis Service документы с одинаковыми терминами (похожие) и группировать их;
- искать дубликаты или похожие файлы по загружаемому образцу;

- обнаруживать объекты защиты в просканированных текстовых файлах с помощью Модуля контентного анализа;
- просматривать извлеченные текстовые контенты с подсветкой обнаруженных объектов защиты;
- выгружать отчеты с информацией о файлах, дубликатах и похожих.

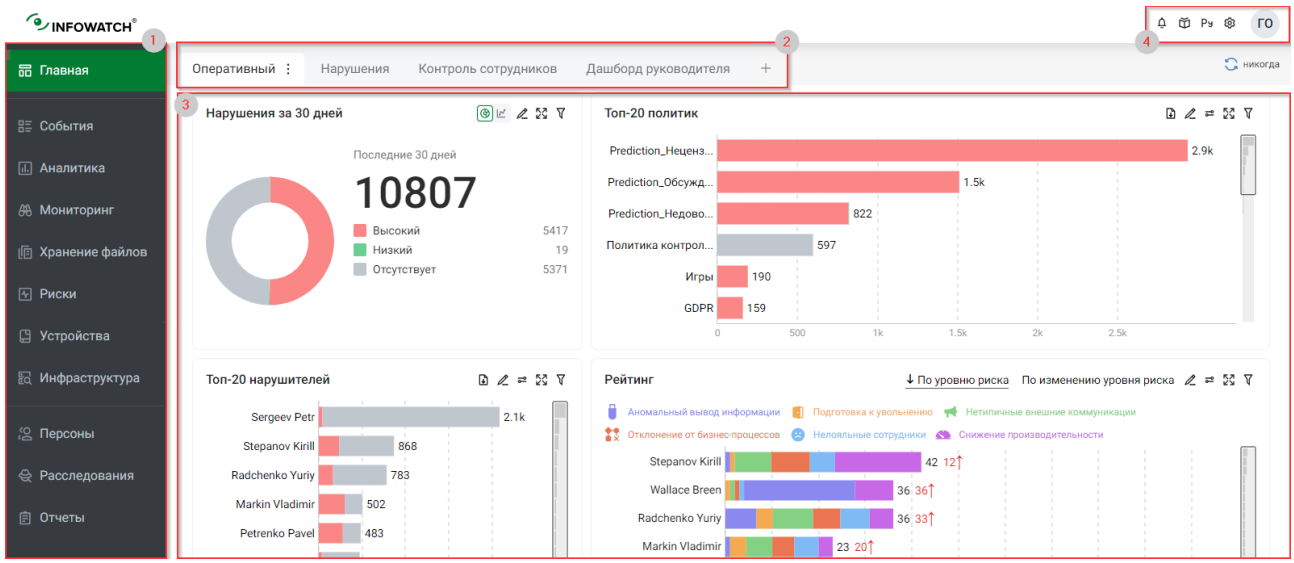
5 Интерфейс Центра расследований

Для работы с Data Discovery используйте веб-консоль Центра расследований InfoWatch.

Центр расследований представляет собой единую графическую среду, в которой могут быть представлены и проанализированы данные установленных продуктов InfoWatch:

- Activity Monitor;
- Data Access Tracker;
- Data Discovery;
- Prediction;
- Vision.

В интерфейсе Центра расследований можно выделить несколько областей. На изображении ниже представлена **Главная** страница.



Элементы интерфейса:

Номер на схеме	Элемент окна	Описание элемента
1	Панель навигации	Рабочие разделы Центра расследований
2	Панель дашбордов	Список предустановленных дашбордов
3	Рабочая область	Отображение виджетов продуктов InfoWatch и работа с ними
4	Меню пользователя	Пользовательское меню, в котором вы можете: • отредактировать свою учетную запись; • сменить пароль; • включить использование горячих клавиш;

Номер на схеме	Элемент окна	Описание элемента
		• узнать контактные данные компании InfoWatch и службы поддержки; • выйти из Системы
	Настройки	Раздел настроек (подробнее см. *InfoWatch Data Discovery. Руководство администратора*)
	Уведомления	Папка уведомлений, в которую попадают сообщения Системы пользователю (подробнее см. InfoWatch Data Discovery. *Руководство администратора,* Системные уведомления)
	Смена языка	Выбор языка интерфейса: *Русский* или *Английский*
	Справочник	Справочник Центра расследований, в котором доступны: • *Сценарии работы с Системой* для наиболее эффективного использования продукта; • *Руководство администратора* с описанием настроек продукта; • *Руководство пользователя* с описанием работы в продукте; • Информация о новых возможностях продукта (*Что нового)*.

Работа в Центре расследований ведется в тематических разделах:

Раздел	Описание
Главная	Основной раздел, в котором размещены предустановленные дашборды с наборами виджетов, позволяющие быстро получить и отобразить информацию о нарушениях и нарушителях на каждом из них с учетом фильтров. Вы также можете создавать собственные дашборды с желаемым набором виджетов.
События	Список событий, согласно установленным фильтрам
Аналитика	Статистическая информация о нарушениях/нарушителях, согласно установленным фильтрам, отображаемая на: • виджетах, содержащих графики, диаграммы и таблицы; • графе связей с визуализацией событий.

Раздел	Описание
	Вкладка *Статистика нарушений* доступна всем продуктам. На вкладке *Граф связей* представлена функциональность продукта InfoWatch Vision (см. InfoWatch Vision. *Руководство пользователя*). На вкладке *Статистика активности* представлена функциональность продукта InfoWatch Activity Monitor (см. InfoWatch Activity Monitor. *Руководство пользователя*)
Мониторинг	Мониторинг активности сотрудников, а также наблюдение за рабочими станциями сотрудников в режиме онлайн. Раздел продукта InfoWatch Activity Monitor (см. InfoWatch Activity Monitor. *Руководство пользователя*)
Хранение файлов	Создание и запуск задач сканирования сетевых ресурсов, просмотр и анализ детальной информации о файлах, включая владельцев, поиск дубликатов и похожих файлов (подробнее см. "Раздел "Хранение файлов")
Риски	Расчет рейтинга сотрудников для прогнозирования рисков информационной безопасности. Раздел продукта InfoWatch Prediction (см. InfoWatch Prediction. *Руководство пользователя*)
Устройства	Контроль внешних устройств хранения Раздел продукта InfoWatch Device Control (см. *InfoWatch Device Control. Руководство пользователя*)
Инфраструктура	Автоматический аудит службы каталогов и доступа к почтовым ящикам; отслеживание изменений прав доступа к папкам и файлам Раздел продукта InfoWatch Data Access Tracker (см. InfoWatch Data Access Tracker. *Руководство пользователя*)
Персоны	Справочник персон компании, а также справочник внешних контактов
Расследования	Проведение расследований с использованием накопленных данных по нарушениям
Отчеты	Работа с шаблонами отчетов и задачами на регулярную отправку почтовых отчетов, выгрузка созданных отчетов

6 Главная страница

Главная страница Центра Расследований содержит инструменты для получения актуальной информации о нарушениях – дашборды. Дашборды состоят из набора виджетов одного или нескольких продуктов InfoWatch. Вы можете создать свои дашборды и добавить на них любые из доступных виджетов или выбрать предустановленные дашборды с наиболее используемыми виджетами. Набор доступных для добавления виджетов зависит от установленных продуктов InfoWatch.

Предустановленные дашборды:

- **Оперативный** - представляет ключевые показатели для мониторинга нарушений:
  - Графики с динамикой нарушений за последние 3 дня или за последние сутки, диаграммы с распределением по уровням нарушений;
  - Лидеры по нарушениям среди сотрудников-отправителей;
  - Топ наиболее нарушаемых политик безопасности;
- **Контроль сотрудников** - представляет статистику нарушений с участием сотрудников:
  - Лидеры по нарушениям среди сотрудников-отправителей;
  - Рейтинг сотрудников (при установленном продукте InfoWatch Prediction);
  - Сотрудник под наблюдением;
- **Файлы** - представляет наиболее часто встречаемые в нарушениях файлы и аудит действий с ними:
  - Топ сработавших объектов защиты;
  - Топ использованных в операциях файлов;
  - Топ использованных в операциях типов файлов;
  - Статистика действий с файлами (при установленном продукте InfoWatch Activity Monitor)
- **Нарушения** - показывает аномалии в действиях сотрудников, изменения поведения, сравнительную статистику по нарушениям:
  - Графики с динамикой нарушений за выбранную дату или период времени, диаграммы с распределением по уровням нарушений;
  - Топ наиболее нарушаемых политик безопасности;
  - Лидеры по нарушениям среди сотрудников-отправителей;
- **Дашборд руководителя** - представляет статистику для руководителя подразделения по действиям его сотрудников:
  - Графики с динамикой нарушений за выбранную дату или период времени, диаграммы с распределением по уровням нарушений;
  - Лидеры по нарушениям среди сотрудников-отправителей;
  - Топ типов событий с нарушениями;
  - Решения пользователей по событиям.

6.1 Дашборды и виджеты

6.1.1 Дашборды

Дашборды с виджетами доступны на **Главной** странице. Офицер безопасности может разместить на рабочей области предустановленные дашборды, создать пользовательские и далее добавить на них виджеты любого из установленных продуктов InfoWatch.
Для этого:

1. Нажмите  на панели дашбордов.

2. Выберите один из предустановленных дашбордов или **Новый дашборд**, введите новое имя, если необходимо, и нажмите **Сохранить.** Новый дашборд будет добавлен на панель.
3. Нажмите  и выберите **+ Добавить виджет**.
4. В открывшемся окне слева укажите, какие виджеты и какого из установленных продуктов InfoWatch вы хотите добавить, или выберите **Все**. Для быстрого поиска нужного виджета по названию воспользуйтесь поисковой строкой.
5. Напротив интересующих виджетов нажмите **+ Добавить** и укажите количество. Можно добавить на дашборд не более 10 виджетов одного типа.
6. Выберите другие виджеты, которые нужно добавить.
7. Нажмите **Применить**. Выбранные виджеты будут расположены на рабочей области.

Условия Единого фильтра, установленные в других разделах Центра расследований, не распространяются на виджеты дашбордов на **Главной** странице. Любой виджет настраивается независимо от других. Чтобы настроить виджет:

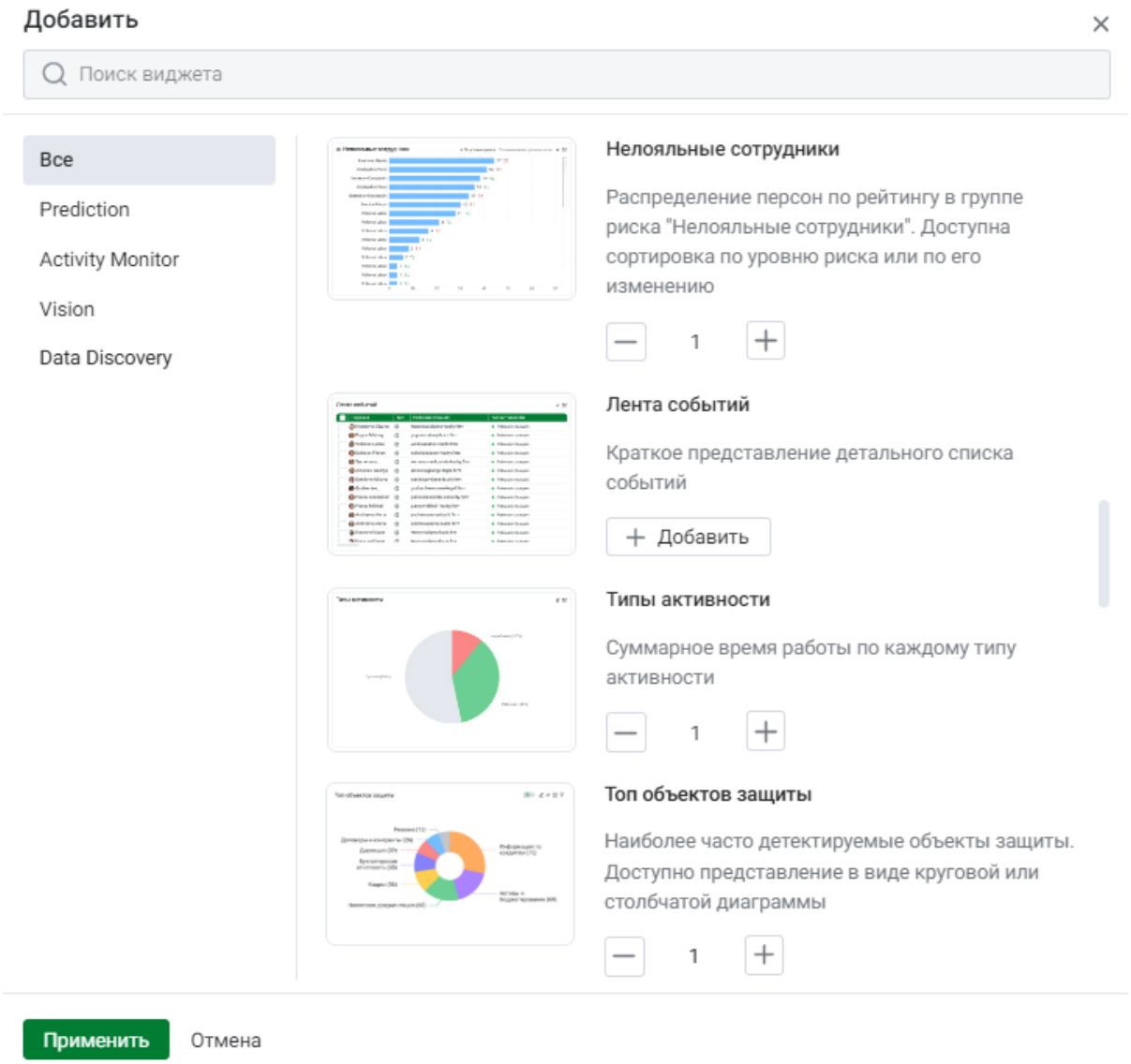
1. Нажмите  в углу нужного виджета.
2. Установите нужные значения фильтра. Если необходимо, добавьте дополнительные атрибуты фильтра (см. "Единый фильтр").
3. Нажмите **Сохранить**.

Данные на виджетах обновляются согласно установленному интервалу. Чтобы установить интервал обновления, в меню правом верхнем углу выберите одно из значений, которое может изменяться от 1 минуты до 24 часов. Если данные обновлять не требуется, выберите **Не обновлять**. Также вы можете обновить данные вручную. Для этого нажмите . Необходимо иметь в виду, что обновляются только виджеты, расположенные в видимой области на дашборде. Чтобы обновить все данные на всех виджетах, пролистайте отображение на экране всех виджетов и нажмите ,

6.1.2 Виджеты

На дашбордах можно расположить виджеты разделов **Аналитика**, **Риски** и **Инфраструктура** (о виджетах Data Discovery см. Вкладка "Файлы").

Набор виджетов, доступных для добавления на дашборды, зависит от продуктов, которые установлены в кластере Центра расследований.

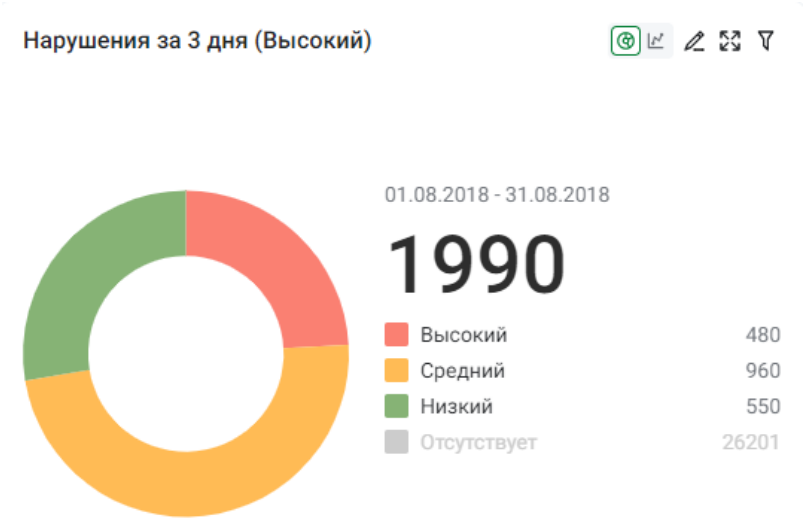


Часть виджетов доступны только на **Главной** странице: **Нарушения** и **Сотрудник под наблюдением**.

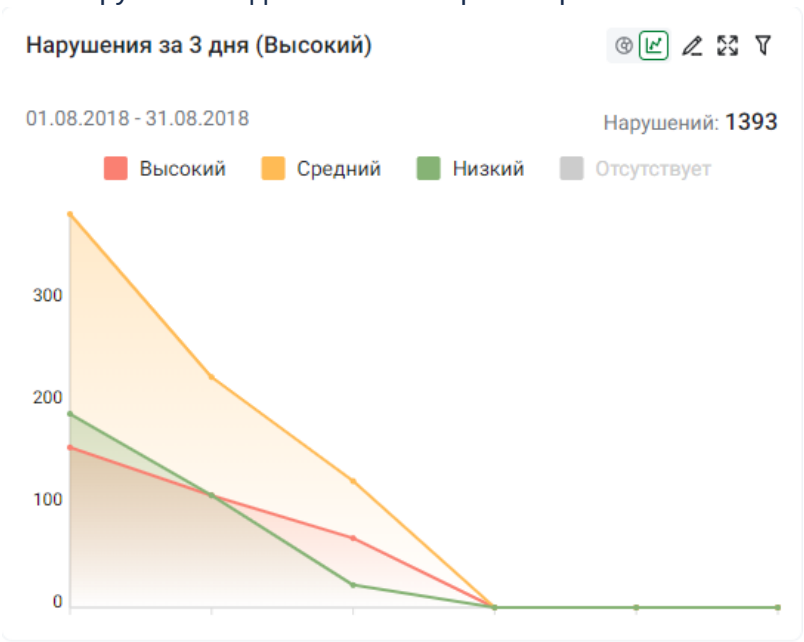
Виджет Нарушения

Этот виджет обрабатывает данные, полученные из Traffic Monitor, и может отображать:

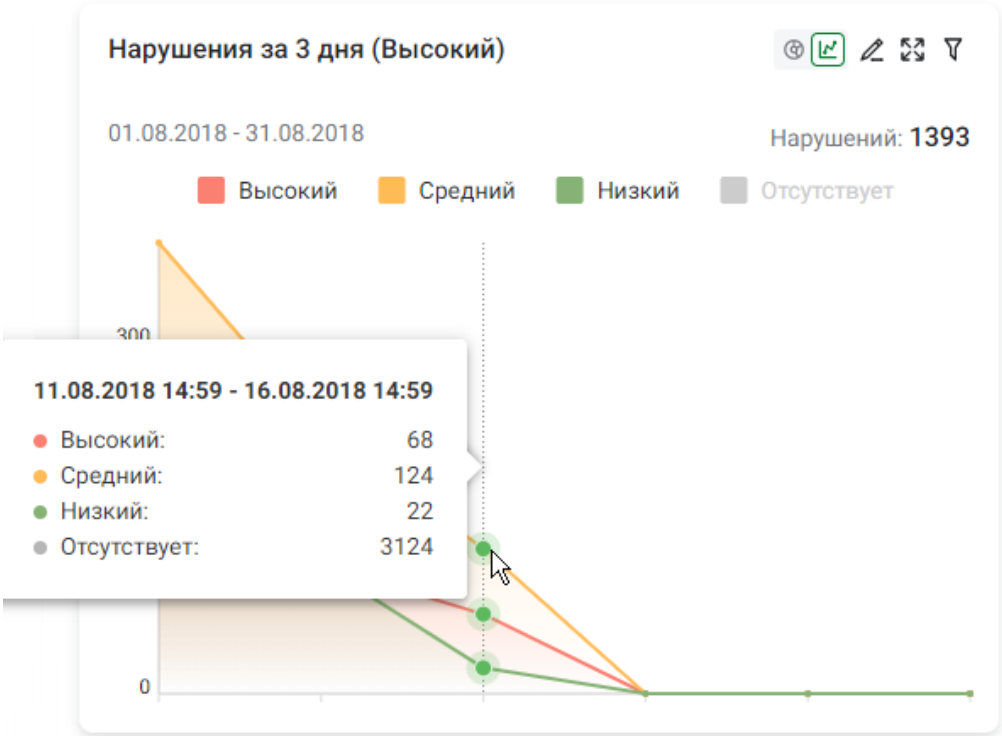
- - общее количество нарушений с распределением по их уровням



- - нарушения в динамике на отрезке времени



В обоих вариантах вы можете управлять отображением, используя легенду, на которой можно включить или отключить графики уровней нарушений:



Виджет Сотрудник под наблюдением

Виджет собирает о сотруднике информацию, на основании которой в дальнейшем Офицер безопасности может начать расследование или взять на особый контроль.

Для настройки виджета:

1. Нажмите  и задайте атрибуты фильтров, применимых к этому виджету.
2. В фильтре **Персона** укажите одну персону из числа сотрудников. При необходимости добавьте другие фильтры и укажите их значения.
3. Нажмите **Сохранить**. Виджет будет отображать:
   - персональную информацию о сотруднике и политики безопасности, в которых сотрудник является участником. Одновременно на виджете может быть не больше 8 политик с наибольшим количеством событий с высоким уровнем нарушений.
   - последнее событие, связанное с персоной, с указанием даты , действия и политики
   - другие персоны из событий, связанных с сотрудником
4. Нажмите , чтобы изменить вид отображения виджета:
   - размер фото сотрудника;
   - диаграммы политик с высоким уровнем нарушений.

7 Единый фильтр

Единый фильтр – универсальный инструмент, позволяющий управлять поиском и отображением данных в Центре расследований InfoWatch.

В разных разделах Центра расследований представлены различные наборы фильтров.

Для Data Discovery Единый фильтр доступен в разделах **Хранение файлов, Аналитика** (вкладка **Статистика нарушений**), **События** и **Персоны**.

Также на доступность фильтра влияют настройки Роли пользователя.

С помощью Единого фильтра вы можете:

- управлять отображением записей о просканированных файлах:
  - выбирать конкретные записи для отображения;
  - исключать из таблицы с информацией о файлах лишние записи;
- сохранять настройки фильтрации, чтобы быстро получать необходимую информацию.

Важно!

При работе с единым фильтром учитывайте:

- все настройки Единого фильтра распространяются только в рамках одной вкладки в браузере;
- при перемещении на шаг назад или вперед (кнопка **Назад**/**Вперед** браузера) фильтр возвращается в то состояние, в каком он был на этом шаге;
- при выходе из Системы все атрибуты фильтра сбрасываются в исходное состояние.

7.1 Фильтры

Единый фильтр – панель, которая содержит группу фильтров.

7.1.1 Фильтры в разделах **Хранение файлов** и **Аналитика → Файлы**

Примечание:

Фильтры Data Discovery уникальны и работают только в разделе **Хранение файлов** на вкладках **Информация о файлах**, **Дубликаты файлов** и **Похожие файлы**, в разделе **Аналитика → Файлы**, в досье в подразделах **Доступные файлы** и **Владелец файлов** (вкладка **Хранение файлов**). Их значения не переносятся при переходе в другие разделы и влияют только на фильтрацию данных Data Discovery.

Пользователь с доступом к разделу **Хранение файлов** может использовать фильтры Data Discovery без дополнительных настроек.

Набор фильтров соответствуют полям таблицы с детальной информацией о файлах:

Название фильтра	Описание	Установка значений
Обнаружен	Период, в который файл был просканирован	Выберите период времени: • Последние сутки - последние 24 часа; • Текущий день - от 0:00 до текущего времени; • Последние 3 дня; • Последние 7 дней; • Последние 30 дней; • Текущий месяц; • Прошлый месяц; • Произвольный период: • Период (укажите даты начала и конца периода); • Выбрать дату (укажите одну или несколько дат). Значение по умолчанию: **Любая дата**
Создан	Период времени, в который файл был создан	
Изменен	Период времени, в который файл был изменен	
Задача	Название задачи сканирования	Чтобы выбрать значение, выполните одно из действий: • Выберите значение из списка. Для поиска в списке введите значение в строку поиска. Отображение элементов в списке не зависит от условий, установленных в других фильтрах. • Введите значение в строку поиска и выберите значение типа "поиск по шаблону". Значение по умолчанию: **Все**
Хост	Имя или IP-адрес хоста	
Расположение	Полный путь к директории	
Имя файла	Имя Файла	
Владелец	Персона или группа, которая является владельцем файла	В списке отображаются имена персон и групп, идентифицированных по SID. Вы можете фильтровать записи одновременно и по группам, и по персонам. Персоны и группы выбираются отдельно, для этого в фильтре используйте переключатель **Персона**/**Группа**. Все выбранные значения объединяются с помощью логического "ИЛИ". Чтобы выбрать значение, выполните одно из действий: • Выберите персону из списка. Для поиска в списке введите имя персоны или SID в строку поиска. • Выберите группу из дерева Active Directory. Для поиска по дереву введите имя группы в строку поиска. Подробнее об идентификации персон и групп в таблице см. Таблица с информацией о файлах. Значение по умолчанию: **Все** Опция **Искать по правам, назначенным персоне напрямую** определяет, должен ли фильтр учитывать вхождение персоны в группы, а также вложенность групп. Если доступ к файлу имеет группа, а не конкретная персона, с включенной опцией

Название фильтра	Описание	Установка значений
Доступен для	Персоны и группы, которые имеют доступ к файлу	поиск по персоне, входящей в группу, не даст результатов. Аналогично действие и для вложенных групп: с включенной опцией поиск по дочерней группе не выведет в результатах файлы, если владение/доступ к ним назначен родительской группе. Если опция выключена, в результатах будут выведены и файлы, владение/доступ к которым назначен персоне или группе напрямую, и файлы, владение/доступ к которым имеет родительская группа. **Примечание:** В фильтре отображаются все группы, добавленные при интеграции с Active Directory. Если у группы нет SID, ее невозможно выбрать в фильтре
Размер	Ограничения на размер файла	В поле **от** укажите минимальный размер файла. В поле **до** укажите максимальный размер файла. Выберите единицу измерения. Значение по умолчанию: **Все** **Примечание**: Если указан только минимальный размер файла, то будут отображаться файлы, чей размер больше или равен минимальному размеру. Если указан только максимальный размер файла, то будут отображаться файлы, чей размер не превышает максимальный размер.
Тип подключения	Тип подключения хоста	Выберите значение из списка. Значение по умолчанию: **Все**
Объект защиты	Наличие в просканированных файлах объектов защиты из конфигурации Traffic Monitor, загруженных в результате интеграции	Доступны следующие значения: • *События без ОЗ* • *События с ОЗ* Для значения *События с ОЗ* **доступен выбор конкретных объектов защиты**. Доступен поиск по объектам защиты. Значение по умолчанию: **Все** - в таблице отображаются все файлы, с объекты защиты и без них.
Отсутствует	Наличие или отсутствие файла при повторном сканировании	Доступны следующие значения: • *Нет* – только обнаруженные файлы. В столбце **Отсутствует** нет значения. • *Да* – только файлы, которые были **не обнаружены** при повторном сканировании по тому же пути. В столбце **Отсутствует** указаны дата и время повторного сканирования, при котором файл не был

Название фильтра	Описание	Установка значений
		обнаружен. Если выбрать это значение фильтра в режиме **Актуально**, будет отображена страница без данных. Значение по умолчанию: **Все** Фильтр доступен только в разделе **Хранение файлов**.

Важно!

В разделе **Хранение файлов** фильтры **Владелец** и **Доступен для** могут работать некорректно после обновления Системы. Отредактируйте или пересоздайте все наборы фильтров (**Сохраненные фильтры**), которые включают эти фильтры.

7.1.2 Фильтры в разделах **События** и **Аналитика → Статистика нарушений**

В таблице указаны фильтры и виджеты, доступные в **События** и **Аналитика → Статистика нарушений**.

Примечание:

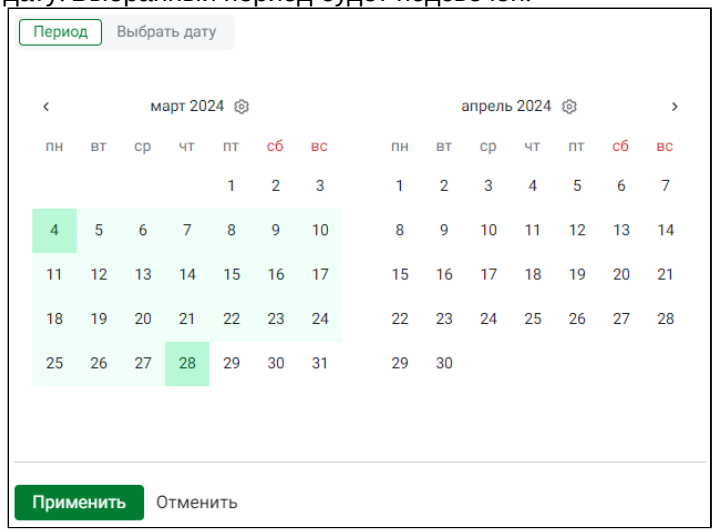
По умолчанию из таблицы ниже доступны виджеты:

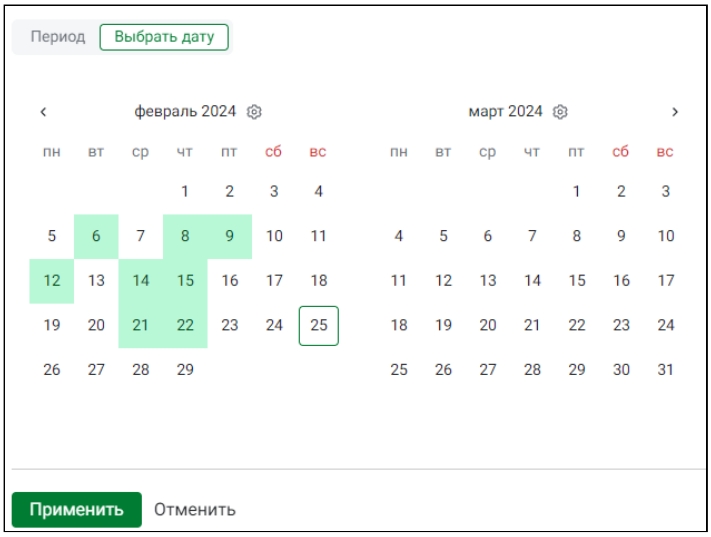
- Дата;
- Отправители;
- Типы событий;
- Уровень нарушений;
- Получатели;
- Теги DLP (о включении см. Получение тегов событий из InfoWatch Traffic Monitor).

Остальные перечисленные в таблице виджеты доступны только при наличии в кластере продукта Vision.

Важно!

Чтобы фильтрация данных была доступна, убедитесь, что в роли пользователя открыт доступ к разделам **События** , **Аналитика → Статистика нарушений** . Подробнее о настройках ролей см. *InfoWatch Data Discovery. Руководство администратора. Роли.*

Название фильтра	Описание и возможные значения
Дата	Атрибуту соответствует виджет **Дата.** По умолчанию задан период *Последние 30 дней*. Выберите период времени, данные за который вас интересуют: • Последние сутки - последние 24 часа; • Текущий день - от 0:00 до текущего времени; • *Последние 3 дня;* • *Последние 7 дней;* • *Последние 30 дней;* • *Текущий месяц;* • *Прошлый месяц;* • *Произвольный период* (можно выбрать даты начала и окончания периода либо список произвольных дат): **Чтобы указать интересующий период времени** 1. Выберите указателем мыши начальную дату в календаре, затем конечную дату. Выбранный период будет подсвечен:  2. Нажмите **Применить**. **Чтобы указать одну или несколько произвольных дат** 1. Выберите указателем мыши одну или несколько нужных дат в календаре. Выбранные даты будут подсвечены:

Название фильтра	Описание и возможные значения
	 2. Нажмите **Применить**.
Часы	Атрибуту соответствует виджет **Время**. Установите время наступления события.
Уровень нарушения	Атрибуту соответствует виджет **Уровень нарушения**. Выберите уровень нарушения (соответствует аналогичному в Traffic Monitor): • *Высокий;* • *Средний;* • *Низкий;* • *Отсутствует.*
Тип события	Атрибуту соответствует виджет **Типы событий**. Создайте выборку по типу события. Для этого выберите тип события Traffic Monitor: • *Работа в приложениях:* • *Буфер обмена;* • *Ввод с клавиатуры;* • *Почта:* • *Электронная почта;* • *Веб-почта;* • *Принтер и МФУ:* • *Печать;* • *Хранение файлов:* • *Data Discovery;* • *Интернет-активность:* • *Web-сообщение;* • *Мессенджер:* • *Facebook;*

Название фильтра	Описание и возможные значения
	• *ICQ;* • *MS Lync;* • *Mail.Ru Агент;* • *MS Teams;* • *Skype;* • *Telegram;* • *Viber;* • *WhatsApp;* • *ВКонтакте;* • *XMPP;* • *Yahoo!;* • *Обмен файлами:* • *Облачные хранилища;* • *Съемное устройство;* • *FTP;* • *Сетевой ресурс;* • *Терминальная сессия;* • *Файловые операции:* • *Чтение файла;* • *Запись файла;* • *Другие:* • *Фотосъемка;* • *SMS*
Объект защиты	Атрибуту соответствует виджет **Объекты защиты**. Наличие в событиях объектов защиты из конфигурации Traffic Monitor, загруженных в результате интеграции. Доступны следующие значения: • *События без ОЗ* • *События с ОЗ* Для значения *События с ОЗ* доступен выбор конкретных объектов защиты. Доступен поиск по объектам защиты
Политика	Атрибуту соответствует виджет **Политики**. Выберите из списка политики, созданные в Traffic Monitor
Вердикт	Атрибуту соответствует виджет **Вердикт**. Выберите из списка интересующие вердикты событий Traffic Monitor: • *Разрешено;* • *Заблокировано;* • *Запросить объяснение;* • *Карантин*

Название фильтра	Описание и возможные значения
Решение пользователя	Атрибуту соответствует виджет **Статистика** с установленным значением **Решение пользователя** для столбца. Выберите из списка нужные решения пользователя по событиям в Traffic Monitor: • *Нарушение;* • *Требуется дополнительная обработка;* • *Решение не принято;* • *Нет нарушения*
Имя вложения	Атрибуту соответствует виджет **Топ вложений**. Введите имя файла в строке поиска или выберите файлы из загруженных файлов-вложений в событиях Traffic Monitor
Тип вложения	Атрибуту соответствует виджет **Типы вложений**. Выберите категории и/или типы файлов-вложений в событиях Traffic Monitor
Филиал	Атрибуту соответствует виджет **Источник**. Выберите филиалы и/или интеграции с ними
Интеграция	Выберите одну или несколько интеграций с системами-источниками
Теги DLP	Позволяет отобразить события, содержащие или не содержащие теги, присвоенные Офицером безопасности в Traffic Monitor. В строке поиска введите первые символы тега. Дополнительно фильтр можно настроить с помощью модификаторов: • *Содержит значение* – будут отображены события с хотя бы одним из тегов, указанных далее; • *Точное совпадение* – будут отображены события, в каждом из которых присутствуют все указанные далее теги. Не будет работать при выборе поиска по шаблону
Тип контакта	Позволяет отобразить события по различным типам контактов, используемых в Системе. Выберите одно или несколько значений: • *Уволенные* – фильтр по событиям, в которых одним из участников является уволившийся ранее сотрудник. Если при интеграции с другими системами не были автоматически загружены личные контакты сотрудника, с ледует добавить их вручную (подробнее см. Добавление контактов персоны) • *Внешние* – фильтр по событиям, в которых хотя бы один участник не принадлежит к списку внутренних доменов • *Внутренние* - фильтр по событиям, в которых хотя бы один участник принадлежит к списку внутренних доменов • *Новые* – фильтр по событиям, в которых данный идентифицированный или неидентифицированный контакт, полученный из Traffic Monitor,

Название фильтра	Описание и возможные значения
	встречается впервые в указанную дату или временной интервал в фильтре **Дата** • *Личные* – фильтр по событиям с сообщениями, отправленными на личные адреса, в том числе при отправке письма одним сотрудником на личный адрес другого сотрудника • *Самому себе* – фильтр по событиям, в которых отправителем и получателем является одна и та же персона. При одновременном выборе значений *Личные* и *Самому себе* будут отображены события с письмами, отправленными сотрудником на свой личный почтовый адрес
Расследования	Позволяет отобразить события, участвующие в одном или нескольких расследованиях. Выберите одно из значений: • *Вне расследования* – фильтр по событиям, которые не участвуют в расследованиях; • *В расследовании* – фильтр по событиям, участвующих в расследованиях. При выборе этого значения укажите в строке поиска одно или несколько наименований расследований. Если оставить поле пустым, поиск будет производиться по всем расследованиям.
Периметр, Вошло периметр, Покинуло периметр	Выберите периметр из списка. Если требуется, задайте модификатор: • *Периметр* – поиск по всем событиям, у которых в атрибутах указаны выбранные периметры; • *Вошло в периметр* – поиск по событиям, у которых в атрибуте **Вошло в периметр** указаны выбранные периметры; • *Покинуло периметр* – поиск по событиям, у которых в атрибуте **Покинуло периметр** указаны выбранные периметры. Если выбрать значение *Неизвестный периметр*, то будут найдены события, в которых периметр входа и/или периметр выхода не был определен в Traffic Monitor
Количество получателей	Позволяет отобразить события с определенным количеством получателей. Указав значения *От* и *До,* можно скрыть массовые рассылки (например, количество получателей до 7) или показать только персональные коммуникации (количество получателей до 1)
Наличие вложения	Позволяет отобразить события, содержащие или не содержащие вложения, а также уточнить их совокупный размер и количество. Например: • указать количество вложений, включающее в себя и вложения внутри архивов • указать размер вложения от 3 КБ, чтобы не учитывать в качестве вложений логотипы компаний из письма или • указать размер вложения от 5 МБ, чтобы найти все события с крупными вложениями

Название фильтра	Описание и возможные значения
Тема письма	Позволяет осуществить полнотекстовый поиск по темам писем в событиях. Поиск ведется от трех введенных символом и может осуществляться: • с учетом морфологии – одновременно по всем введенным словам, включая однокоренные, в любом порядке • по точному совпадению. Если будет найдено более 1000 событий, то введенный запрос необходимо будет дополнить или уточнить. Фильтр работает только в разделе **Статистика**
Контакт*, Контакт отправителя*, Контакт получателя*	Позволяет отобразить события, участником которых были указанные адреса электронной почты или идентификаторы устройств. В строке поиска введите первые символы имя сотрудника, его контактных данных или устройства. Дополнительно фильтр можно настроить с помощью модификаторов: • **Содержит значение** - будут отображены события с участием хотя бы одного из адресов или идентификаторов устройств; • **Точное совпадение** - будут отображены события, в каждом из которых участвуют все выбранные адреса или идентификаторы устройств. Не будет работать при выборе поиска по шаблону. Для поиска сотрудников по имени используйте фильтр **Персона**.
Персона*, Отправитель*, Получатель*	Позволяет отобразить события, участником которых являлись указанные персоны. В строке поиска можно указывать как имя сотрудника, так и его контактные данные. Дополнительно фильтр можно настроить с помощью модификаторов: • **Содержит значение** - будут отображены события с участием хотя бы одной из выбранных персон; • **Точное совпадение** - будут отображены события, в каждом из которых участвуют все выбранные персоны
Группа*, Группа получателя*, Группа отправителя*	Атрибуту соответствует виджет **Аналитика по группам/персонам**. Создает выборку по группам AD и Traffic Monitor. Укажите группы в иерархии в компании, синхронизированные с Traffic Monitor или AD
Должность*, Должность отправителя*, Должность	Укажите все интересующие вас должности в компании, добавленные при синхронизации с Traffic Monitor или AD

Название фильтра	Описание и возможные значения
получателя*	
Отдел*, Отдел отправителя*, Отдел получателя*	Укажите все интересующие вас отделы в компании, добавленные при синхронизации с Traffic Monitor или AD
Статус*, Статус отправителя*, Статус получателя*	Укажите все интересующие вас статусы сотрудников в компании, добавленные при синхронизации с Traffic Monitor
* Для атрибутов, связанных с персонами, можно установить модификаторы: • *Любой;* • *Отправитель;* • *Получатель.* Если установить модификатор *Отправитель/Получатель*, то выбранные персоны, группы персон или контакты будут являться отправителями или получателями в событиях. Если установить модификатор *Любой*, то будут выбраны все участники события: все персоны, группы персон или контакты	

Подробнее про настройку панели фильтров см. Работа с Единым фильтром.

7.2 Работа с Единым фильтром

7.2.1 Добавление и удаление фильтров

В панели фильтров можно использовать один или несколько фильтров.

Чтобы добавить фильтр на панель фильтров:

1. Нажмите ;
2. В открывшемся окне выберите нужный фильтр из списка.
3. Щелкните вне окна со списком фильтров.

Фильтр будет добавлен на панель фильтров.

Чтобы удалить фильтр с панели фильтров:

1. Нажмите ;
2. Снимите флажок с фильтра в списке.
3. Щелкните вне окна со списком фильтров.

Фильтр будет удален с панели фильтров.

Чтобы переместить фильтр на другое место на панели, зажмите  у выбранного фильтра перетащите на любое другое место.

Чтобы скрыть панель Единого фильтра, нажмите .

7.2.2 Значения фильтров

В фильтре необходимо выбрать минимум одно значение. Если не выбрано ни одно значение, то используется значение по умолчанию.

Если изменить значение фильтра, то автоматически обновится таблица в разделе **Хранение файлов > Информация о файлах**. В таблице отображаются только те записи, которые соответствуют условиям фильтров.

Важно!

При выходе из Системы сбрасываются значения фильтров. Рекомендуется сохранить фильтры перед выходом из Системы (см. "Сохранение и загрузка наборов фильтров").

Условие с несколькими фильтрами

Если заданы значения для нескольких фильтров, то условия нескольких фильтров объединяются с помощью логического "И".

Пример:

Укажите значение для нескольких фильтров:

- в фильтре **Имя файла** укажите значение: **отчет .doc**
- в фильтре **Задача** укажите значение: **Сканирование сетевых ресурсов**
- в фильтре **Размер** укажите в поле **до** значение **2** и выберите единицу измерения **МБ**

В результате настройки в таблице с информацией о файлах отображаются только записи, которые одновременно соответствуют всем вышеуказанным условиям.

В таблице отображаются записи о файлах с именем **отчет.doc** из задачи **Сканирование сетевых ресурсов**, если размер файла не превышает 2 мегабайта.

Тип значения "поиск по шаблону"

Тип значения "поиск по шаблону" доступен для следующих фильтров:

- Задача;
- Хост;
- IP-адрес хоста;
- Расположение;

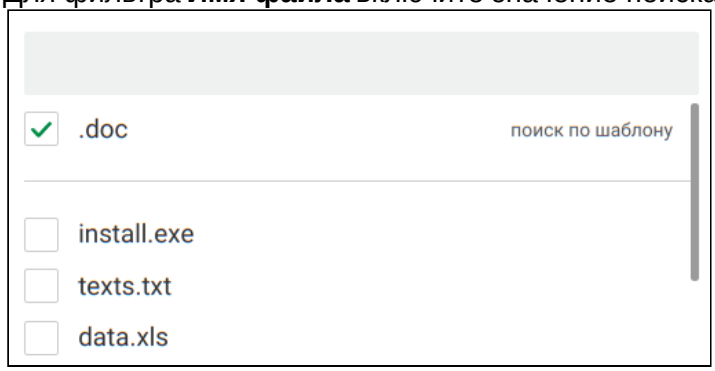
- Имя файла;
- Владелец;
- Доступен для.

Чтобы использовать значение типа "поиск по шаблону", введите значение в строку поиска и выберите значение типа "поиск по шаблону".

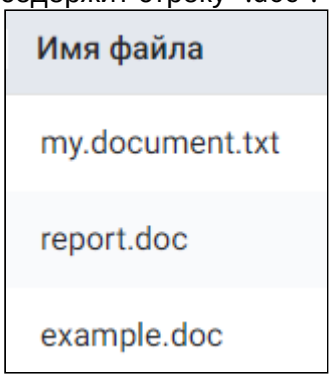
При фильтрации Система проверяет, входит ли значение из строки поиска в полное значение поля записи о файле. Для значений типа "поиск по шаблону" не учитывается регистр.

Пример:

1. Для фильтра **Имя файла** включите значение поиска по шаблону ".doc":



2. В результате фильтрации в таблице отображаются только те файлы, имя которых содержит строку ".doc":



Включение и исключение значений фильтра

В разделе **Хранение файлов** > **Информация о файлах** несколько состояний доступны для значений следующих фильтров:

- Задача;
- Хост;
- IP-адрес хоста;
- Расположение;
- Имя файла;
- Владелец;
- Доступен для;
- Тип подключения.

Доступные состояния для значений вышеуказанных фильтров:

-  – значение не выбрано (состояние по умолчанию);

-  – значение включено;
-  – значение исключено.

Чтобы изменить состояние значения, щелкните по флажку состояния.

Вы можете включить и/или исключить несколько значений, при этом:

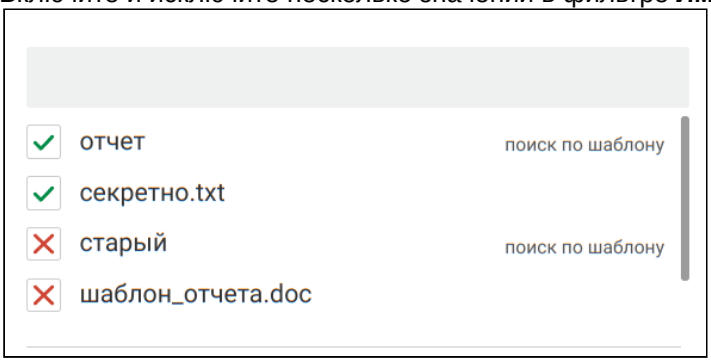
- Несколько включенных значений объединяются с помощью логического "ИЛИ";
- Несколько исключенных значений объединяются с помощью логического "ИЛИ";
- Условие из нескольких включенных значений и условие из нескольких исключенных значений объединяются с помощью логического "И".

Пример:

В таблице находятся 5 записей со следующими значениями в поле **Имя файла**:

- отчет_2021.docx
- старый_отчет.doc
- новый_отчет.pdf
- секретно.txt
- шаблон_отчета.doc

Включите и исключите несколько значений в фильтре **Имя файла**:



В таблице будут отображаться записи о файлах, у которых значение поля **Имя файла** соответствует одновременно двум условиям:

- **Условие по включенным значениям**:
  Условие выполняется, если: имя файла содержит в себе подстроку **отчет** ИЛИ именем файла является **секретно.txt**
- **Условие по исключенным значениям**:
  Условие НЕ выполняется, если: имя файла содержит подстроку **старый** ИЛИ именем файла является **шаблон_отчета.doc**

В результате в таблице отображаются записи со следующими значениями в поле **Имя файла**:

- отчет_2021.docx
- новый_отчет.pdf
- секретно.txt

Сброс значений фильтра

Чтобы сбросить значения фильтра, нажмите на панели фильтра . Для фильтра будет установлено значение по умолчанию.

Чтобы сбросить значения для всех фильтров, нажмите .

7.2.3 Сохранение и загрузка наборов фильтров

Чтобы использовать заданные значения фильтров в дальнейшем, вы можете сохранить их в каталоги одного из типов:

- *Личные* – каталоги и их содержимое доступны только создавшему их пользователю;
- *Общие* – каталоги и их содержимое доступны всем пользователям Системы.

Чтобы сохранить фильтр:

1. Установите требуемые значения фильтров и нажмите .
2. В окне **Сохранить фильтр** задайте имя нового фильтра.
3. Выберите тип каталога, в который Система сохранит фильтр: *Личный* или *Общий*.
4. Выберите каталог, в котором следует сохранить фильтр, или создайте новый. Если каталогов много, используйте поиск по названию.

> **Примечание:**
>
> Если не выбрать каталог, Система сохранит фильтр в корневой директории указанного типа каталогов.

5. Нажмите **Сохранить**.

Чтобы воспользоваться ранее созданным фильтром:

1. В области Единого фильтра нажмите **Сохраненные фильтры**.
2. Откройте каталог, который содержит интересующий фильтр.
3. Нажмите на фильтр, чтобы его активировать.

Чтобы отредактировать каталоги и фильтры:

1. В области Единого фильтра нажмите **Сохраненные фильтры** или на имя примененного фильтра.
2. Нажмите  над списком с фильтрами.
3. В окне **Настройка фильтров и каталогов** вы можете:
    - изменить иерархию каталогов путем перетаскивания;
    - в строке типа каталога:
        -  создать каталог;
    - при наведении курсора на каталог:
        -  переименовать каталог;
        -  создать дочерний каталог;
        -  удалить каталог;
    - перемещать фильтры внутри каталогов и между каталогами;
    - при наведении курсора на имя фильтра:
        -  переименовать фильтр;

-  записать текущий фильтр, т.е. сохранить с этим названием примененные на данный момент значения фильтров;
-  удалить фильтр.

При этом Главный офицер может редактировать и удалять личные и общие каталоги и/или их содержимое. Другие офицеры безопасности могут редактировать и удалять только личные каталоги и/или фильтры.

8 Раздел "События"

> **Примечание:**
>
> Data Discovery не использует в работе события из Traffic Monitor.

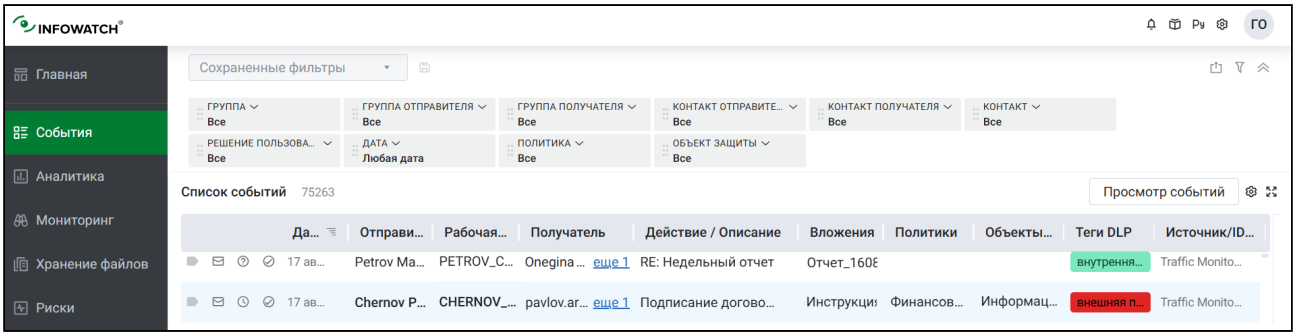
В этом разделе для просмотра доступны данные событий, загруженных в Центр расследований после синхронизации с Traffic Monitor:

- Уровень нарушения в событии;
- Тип события;
- Решение Офицера безопасности по событию;
- Вердикт Системы по событию;
- *Дата* создания события в Traffic Monitor;
- *Теги DLP* – теги, назначенные в Traffic Monitor;
- *Рабочая станция* – компьютеры, связанные с событием;
- *Отправитель* – лицо, совершившее нарушение;
- *Получатель* – лица, которым субъект передал данные;

> **Примечание:**
>
> Отметка  рядом с контактом *Отправителя* или *Получателя* означает, что это новый контакт, который появился в Системе в дату или период времени, указанные в фильтре **Дата**.

- *Вложения* – наличие файлов-вложений в событии, их названия и размер;
- *Действие/Описание* переданных данных;
- *Политики*, сработавшие на данном виде нарушения;
- *Объекты защиты*, которые были переданы;
- *Источник/ID в источнике* – идентификатор события в Traffic Monitor.

Чтобы изменить набор столбцов таблицы с событиями, нажмите  и в меню установите/снимите флажки рядом с названиями столбцов. Чтобы вернуть первоначальные настройки таблицы, нажмите **Вернуть системные настройки**.



Получите интересующую выборку по событиям, настроив фильтры на панели Единого фильтра (подробнее см. "Единый фильтр"). События представлены в разделе в виде таблицы.

Чтобы получить расширенную информацию или анализ выбранного события, нажмите:

-  внизу списка событий. При этом вместе с основными данными события будут отображены и такие, как реплики в чате, текст письма, вложенные изображения и т.д. В детализации событий цветом выделяются слова, которые относятся к обнаруженным объектам защиты;
- , чтобы перейти в режим просмотра диалогов в мессенджерах (доступно, если у вас установлен продукт Vision);
- , чтобы добавить событие в новое или ранее созданное расследование;
- , чтобы посмотреть событие на графе связей (доступно, если у вас установлен продукт Vision);
- , чтобы скопировать ID события для быстрого поиска его в Traffic Monitor;
- , чтобы просмотреть снимки экрана в событии (доступно, если у вас установлен продукт Activity Monitor);
- , чтобы просмотреть событие на таймлайне персоны (доступно, если у вас установлен продукт Activity Monitor).

Чтобы оперативно обработать сразу несколько событий:

1. В списке выберите события:
   - зажав клавишу ***Ctrl*** и выделив нужные события;
   - зажав клавишу ***Shift*** и выбрав диапазон событий (от и до).
2. Нажмите на панели сверху:
   -  для построения частного графа связей по такой выборке (доступно, если у вас установлен продукт Vision);
   -  для начала расследования с указанной выборкой событий (подробнее см. "Добавление материалов в расследование");
   - **Просмотр событий**, чтобы перейти в раздел **События** Traffic Monitor для более подробного анализа. При этом будут отображены только 300 событий из всех отфильтрованных в Центре расследований. Если события принадлежат нескольким источникам, то они будут открыты в разных вкладках браузера.

     Чтобы снять выделение с выбранных событий, нажмите  **Сбросить выбранные события**.

Совет:

Если вы часто переключаетесь между Центром расследований и Traffic Monitor, рекомендуем настроить сквозную аутентификацию, которая позволяет не вводить повторно имя пользователя и пароль. Подробнее о настройке см. документ InfoWatch Traffic Monitor. *Руководство администратора*, Настройки автоматического удаления событий Data Discovery в Traffic Monitor.

Примечание:

Столбцы таблицы с событиями отображаются с учетом настроек роли, выданной пользователю. Поэтому необходимо, чтобы пользователю были выданы нужные права доступа:

- для столбцов *Дата, Тип события, Отправитель, Получатель* включите пункты **События, Статистика нарушений** на вкладке **Функции** в настройке роли;
- для столбцов *Описание, Вложения, Источник, ID события* – пункт **Детали событий;**
- для столбцов *Уровень нарушения, Решение пользователя, Вердикт* – пункт **Детали нарушений;**
- для столбцов *Политики, Объекты защиты, Периметры* – пункт **Политики, Объекты защиты и Периметры**

Действия пользователя:

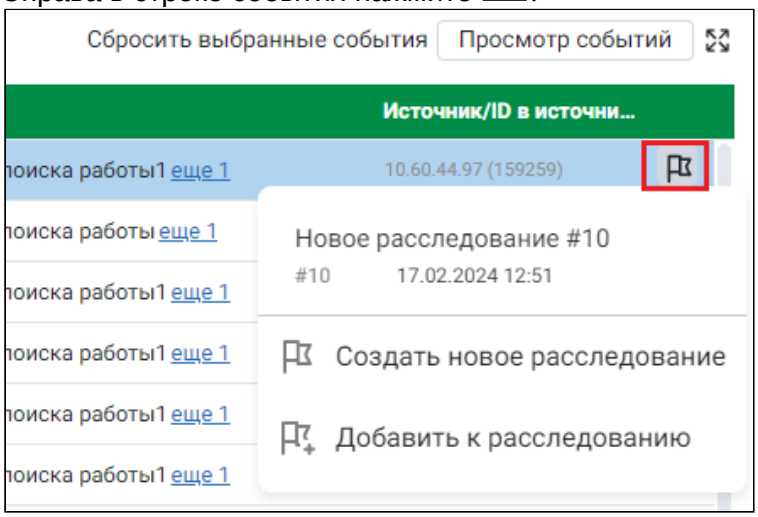
- Просмотр и выгрузка данных событий
- Просмотр досье персоны
- Фильтрация событий по персонам и контактам
- Добавление событий в расследование

8.1 Добавление событий в расследование

Вы можете добавить события из списка в расследование или создать новое расследование на их основе.

Чтобы добавить событие в расследование:

1. Выделите нужное событие в списке щелчком мыши.
2. Справа в строке события нажмите .



3. Выберите:
   - **Добавить к расследованию** – далее начните вводить название расследования, в которое нужно добавить событие, выберите его в выпадающем списке и нажмите **Добавить**;
     или

- **Создать новое расследование** – в этом случае новое расследование будет создано автоматически.

Чтобы добавить одно или несколько событий в расследование:

1. Выделите нужные события, используй клавиши *Ctrl* или *Shift*.
2. В правой части экрана над таблицей с событиями нажмите:
   - **Добавить к расследованию** – далее начните вводить название расследования, в которое нужно добавить событие, выберите его в выпадающем списке и нажмите **Добавить**;
     или
   - **Создать новое расследование** – в этом случае новое расследование будет создано автоматически.

События, участвующие в расследованиях, помечены знаком . Такие события также можно добавлять в новые или уже существующие расследования.

Чтобы перейти в расследование из события:

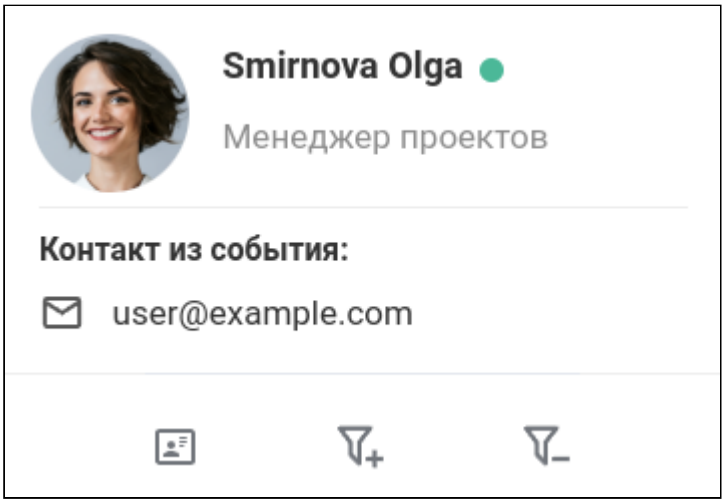
1. Нажмите .
2. Выберите расследование в списке. Оно откроется на новой вкладке браузера.

Чтобы просмотреть события, участвующие в расследованиях, воспользуйтесь фильтром **Расследования** (подробнее см. "Единый фильтр").

8.2 Просмотр и выгрузка данных событий

8.2.1 Просмотр используемых контактов

Карточка персоны в разделе **События** содержит контакты сотрудников, которые участвовали в выбранном событии:



8.2.2 Просмотр данных о вложениях

Чтобы просмотреть информацию о вложениях в событии, нажмите **еще**. В появившемся списке будут представлены все вложения этого события.

Вы можете добавить одно или несколько вложений в фильтр, и в списке будут только содержащие это вложение события.

> **Примечание:**
>
> Для корректной работы необходимо разрешить открывать всплывающие окна с указанных источников в браузере.

8.2.3 Выгрузка данных события

При подготовке отчетов по инцидентам, проведении расследования или составлении подборки интересующих вас событий можно выгрузить таблицу с событиями в виде документа Excel или сохранить её в досье. Для этого:

1. В разделе **События** установите нужные значения атрибутов фильтра. Список событий будет обновлен.
2. В правой верхней части экрана нажмите .
3. Выберите **Выгрузить отчет.**
4. Выберите состав отчета и нажмите **Сохранить**.
5. Дождитесь, пока сформируется отчет.
6. Перейдите в раздел **Отчеты** на вкладку **Архив отчетов** и скачайте XLSX-файл с отчетом, нажав на ссылку.
7. Откройте XLSX-файл. На листе **Дополнительная информация** указаны дата и время создания отчета, примененные к выборке фильтры и их значения, а также ссылка на раздел формирования отчета. На листе **Список событий** приведена таблица с событиями. Если требуется, вы можете отредактировать ее вручную.

Чтобы настроить регулярную отправку отчетов:

1. Настройте фильтр (см. Единый фильтр) и сортировку записей.

   > **Важно!**
   >
   > Если после фильтрации таблица содержит более 10 000 записей, то в отчет добавляются только первые 10 000 записей.

   В отчет будут выгружены данные в выбранном представлении.
2. В правой верхней части экрана нажмите .
3. Выберите **Настроить регулярную отправку**.
4. Определите состав отчета.
5. Задайте параметры отправки:

Параметр	Описание
Период формирования отчета	Укажите, за какой период необходимо отправлять отчет: • **Последние сутки** • **Текущий день** • **Последние 3 дня**

Параметр	Описание
	• **Последние 7 дней** • **Последние 30 дней**
Получатели	Укажите получателей отчета
Частота отправки	Укажите, с какой периодичностью пользователю необходимо получать отчет: • **Каждый день** • **Каждую неделю** Вы можете выбрать дни недели, по которым Центр расследований будет отправлять отчеты. • **Каждый месяц** Отчет будет отправляться 1 числа месяца, следующего за отчетным.
Время	Укажите время отправки отчета

6. Нажмите **Сохранить**. Будет создана задача на регулярную отправку. Список задач вы можете просмотреть в разделе "**Отчеты**" на вкладке "**Задачи на регулярную отправку**". Подробнее см. "Задачи на регулярную отправку".

Важно!

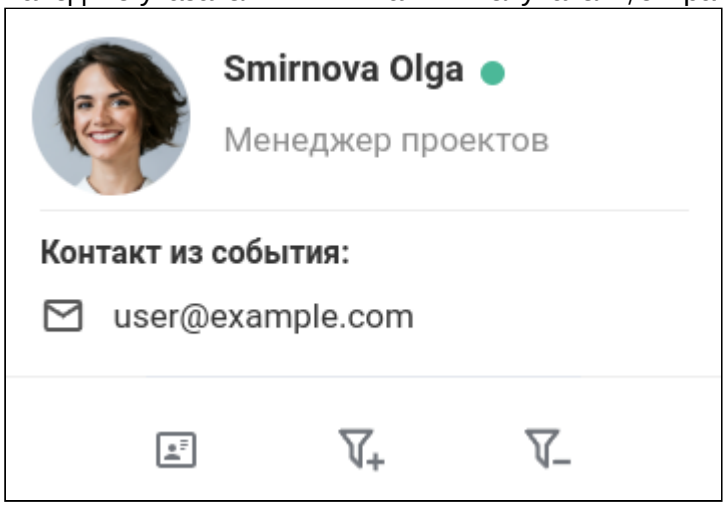
Для регулярной отправки отчетов в Центре расследований должно быть настроено соединение с почтовым сервером (подробнее см. *InfoWatch Data Discovery. Руководство администратора*. *Сервисы для уведомлений*).

Вы также можете сформировать шаблоны отчетов по персонам и нарушениям и выгрузить их в формате XLSX, PDF или HTML. Подробнее см. "Конструктор отчетов".

8.3 Фильтрация событий по персонам и контактам

Чтобы отфильтровать события по нужным получателям или отправителям:

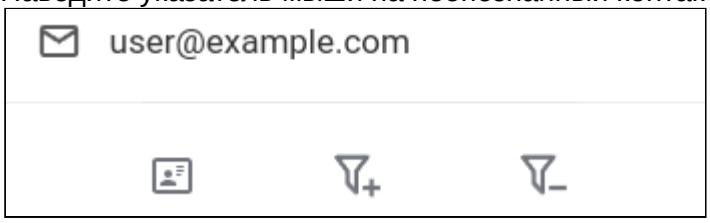
1. Наведите указатель мыши на имя получателя/отправителя/контакт:



2. Нажмите на выбор:
   -  - данный получатель/отправитель будет добавлен как включенный атрибут фильтра **Персона**, и список событий обновится;
   -  - данный получатель/отправитель будет добавлен как исключенный атрибут фильтра **Персона**, и список событий обновится.

Чтобы отфильтровать события по неопознанным контактам:

1. Наведите указатель мыши на неопознанный контакт в событии:



2. Нажмите на выбор:
   -  - данный контакт будет добавлен как включенный атрибут фильтра **Контакты**, и список событий обновится;
   -  - данный контакт будет добавлен как исключенный атрибут фильтра **Контакты**, и список событий обновится.

Чтобы присвоить неопознанный контакт персоне:

1. Наведите указатель мыши на неопознанный контакт в событии.
2. Нажмите .
3. Если требуется завести новое досье с этим контактом:
   a. Нажмите **Создать досье**.
   b. Введите все необходимые данные, как указано в статье "Создание, редактирование и удаление досье".
4. Если требуется добавить контакт в уже существующее досье:
   a. Нажмите **Добавить в досье**.
   b. Укажите досье сотрудника и нажмите **Добавить**.

9 Раздел "Аналитика"

Раздел **Аналитика** предоставляет Офицеру безопасности наглядную статистику в виде таблиц, диаграмм и графиков, построенных по заданным условиям Единого фильтра, по всем полученным из Traffic Monitor событиям. В разделе представлены вкладки:

- **Статистика нарушений** – виджеты статистики (см. "Виджеты статистики");
- **Граф связей** – граф связей продукта Vision (см. *InfoWatch Vision. Руководство пользователя*);
- **Статистика активности** – виджеты Activity Monitor (см. *InfoWatch Activity Monitor. Руководство пользователя*);
- **Файлы** – виджеты Data Discovery (см. Вкладка "Файлы").

9.1 Виджеты статистики

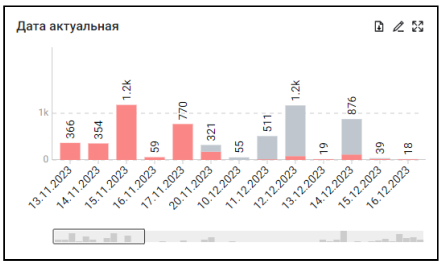
Виджеты на вкладке **Статистика нарушений** наглядно отображают данные от Traffic Monitor, хранящиеся в Системе.

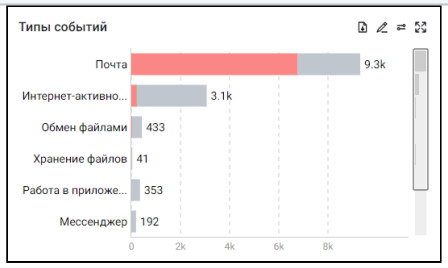
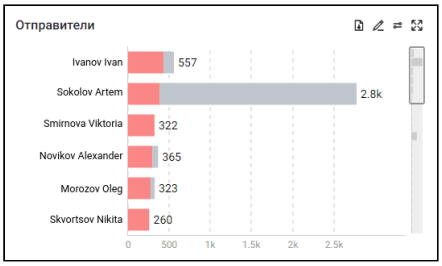
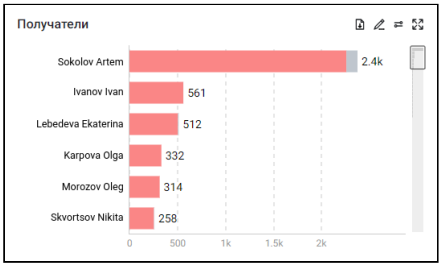
Примечание:

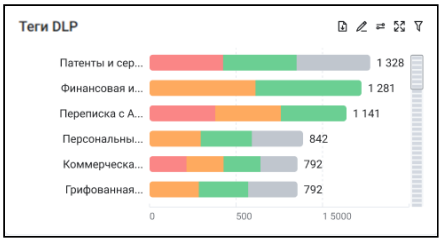
В таблице ниже представлены общедоступные виджеты статистики:

- Дата;
- Типы событий;
- Уровень нарушений;
- Отправители;
- Получатели;
- Теги DLP (о включении см. Получение тегов событий из InfoWatch Traffic Monitor).

Остальные виджеты статистики доступны только при наличии в кластере продукта Vision.

Название виджета	Описание
Дата 	Отображает распределение событий по указанным датам с учетом уровней нарушений. Можно выбрать одну или несколько дат и отфильтровать события за выбранный день.
Типы событий	Отображает распределение событий Traffic Monitor с разными уровнями нарушений по типам. **Чтобы получить расширенную информацию:**

Название виджета	Описание
	Если навести курсор на имя источника, выводится расширенная информация о количестве событий с разными уровнями нарушения в них: 
Уровень нарушений 	Отображает распределение событий по уровням нарушений.
Отправители 	Отображает список персон-отправителей с наибольшим количеством нарушений. Выводятся первые 50 нарушителей, отсортированных по событиям с максимальным уровнем нарушений. **Чтобы настроить отображение на виджете:** 1. Нажмите ⚙ в углу виджета. 2. В открывшемся окне установите: a. Количество отображаемых отправителей: от *50* до *200*; b. Принцип их сортировки: *по уровню нарушения* или *по общему количеству событий.* 3. Нажмите **Сохранить**.
Получатели 	Отображает список получателей, участвующих в событиях. Получатели могут быть идентифицированными или неидентифицированными. Выводятся первые 50 наиболее популярных контактов. **Чтобы добавить контакт из события в фильтр:** 1. Наведите указатель мыши на контакт в событии. 2. Нажмите ⚙. Выбранный контакт добавится в Единый фильтр, и данные событий будут

Название виджета	Описание
	перестроены с его учетом. Чтобы исключить контакт из фильтрации, нажмите . **Чтобы добавить неизвестный контакт в досье:** 1. Наведите указатель мыши на неизвестный контакт на диаграмме. 2. Нажмите . 3. Введите все необходимые данные, как указано в "Создание, редактирование и удаление досье".
Теги DLP 	Отображает распределение по уровням нарушений для событий с тегами, назначенными в Traffic Monitor.

9.2 Выгрузка отчетов с данными статистики

Вы можете выгрузить данные с виджетов в виде отчета, а также настроить регулярную отправку отчета.

Чтобы сформировать и выгрузить отчет:

1. В разделе **События** установите нужные значения атрибутов фильтра. Список событий будет обновлен.
2. В правой верхней части экрана нажмите .
3. Выберите **Выгрузить отчет.**
4. В открывшемся окне **Выгрузка отчета** вы можете определить состав **Детализированного** отчета, который содержит данные из виджетов в виде таблиц.
5. Вы можете включить в отчет сразу все виджеты, установив  напротив **Состав выгрузки**.

   Если необходимо включить в отчет только определенные виджеты, установите  только напротив их названий.
6. Дождитесь, пока сформируется отчет.
7. Перейдите в раздел **Отчеты** на вкладку **Архив отчетов** и скачайте XLSX-файл с отчетом, нажав на ссылку.
8. Откройте XLSX-файл. На листе **Дополнительная информация** указаны дата и время создания отчета, примененные к выборке фильтры и их значения, а также ссылка на раздел формирования отчета. На листе **Список событий** приведена таблица с событиями. Если требуется, вы можете отредактировать ее вручную.

Чтобы настроить регулярную отправку отчетов:

1. Настройте фильтр (см. Единый фильтр) и сортировку записей.

   > **Важно!**
   >
   > Если после фильтрации таблица содержит более 10 000 записей, то в отчет добавляются только первые 10 000 записей.

   В отчет будут выгружены данные в выбранном представлении.
2. В правой верхней части экрана нажмите .
3. Выберите **Настроить регулярную отправку**.
4. Определите состав отчета.
5. Задайте параметры отправки:

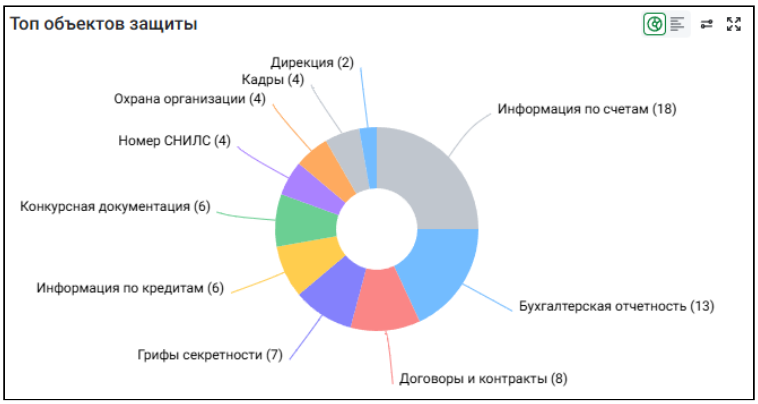
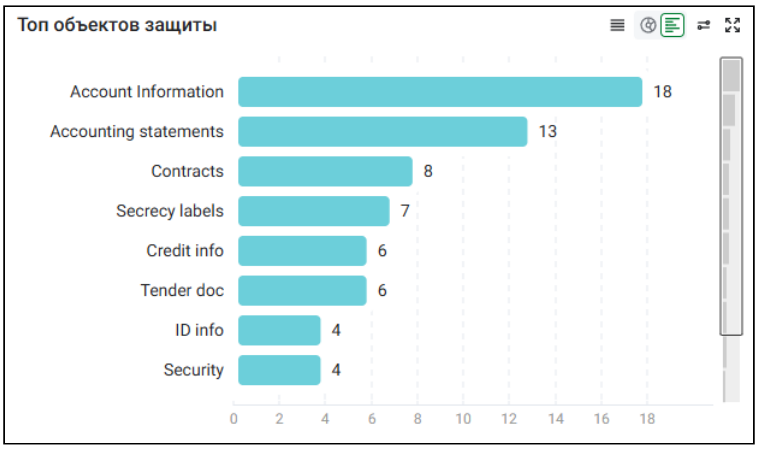
Параметр	Описание
Период формирования отчета	Укажите, за какой период необходимо отправлять отчет: • **Последние сутки** • **Текущий день** • **Последние 3 дня** • **Последние 7 дней** • **Последние 30 дней**
Получатели	Укажите получателей отчета
Частота отправки	Укажите, с какой периодичностью пользователю необходимо получать отчет: • **Каждый день** • **Каждую неделю** Вы можете выбрать дни недели, по которым Центр расследований будет отправлять отчеты. • **Каждый месяц** Отчет будет отправляться 1 числа месяца, следующего за отчетным.
Время	Укажите время отправки отчета

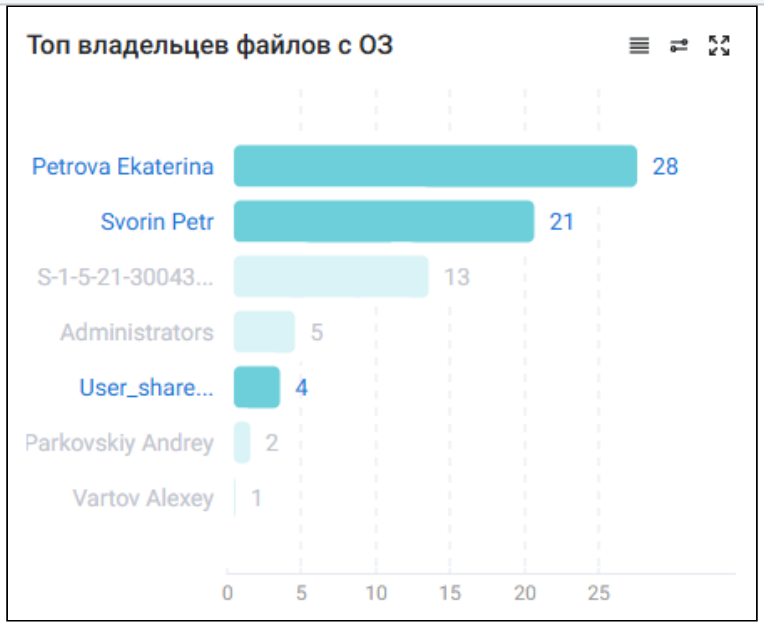
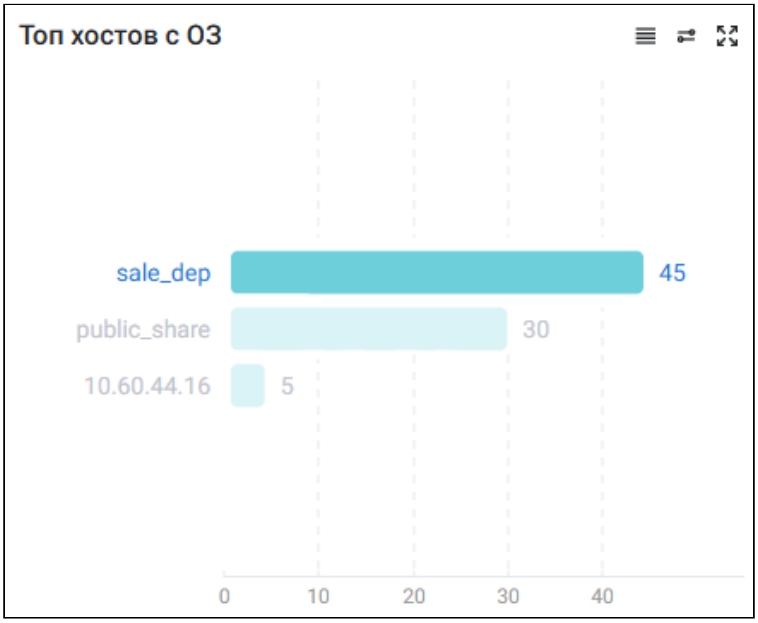
6. Нажмите **Сохранить**. Будет создана задача на регулярную отправку. Список задач вы можете просмотреть в разделе "**Отчеты**" на вкладке "**Задачи на регулярную отправку**". Подробнее см. "Задачи на регулярную отправку".

> **Важно!**
>
> Для регулярной отправки отчетов в Центре расследований должно быть настроено соединение с почтовым серверов (подробнее см. *InfoWatch Data Discovery. Руководство администратора. Сервисы для уведомлений*).

9.3 Вкладка "Файлы"

На вкладке представлены виджеты с информацией, которая получена в результате сканирования файлов продуктом Data Discovery.

Виджет	Описание
Топ объектов защиты  	Объекты защиты, наиболее часто детектируемые в просканированных файлах. Данные представлены в виде круговой или столбчатой диаграммы. **Примечание**: В одном файле может быть обнаружено несколько объектов защиты. Виджет отображает все объекты защиты файлов, удовлетворяющих условиям фильтра, даже те, которые не указаны в фильтре **Объект защиты**.
Топ владельцев файлов с ОЗ	Список владельцев файлов с наибольшим количеством детектированных объектов защиты

Виджет	Описание
	
Топ хостов с ОЗ 	Список хостов с наибольшим количеством детектированных объектов защиты

Вы можете:

- **Просмотреть виджет в полноэкранном режиме**:
  1. Нажмите **Открыть на весь экран**.
  2. Чтобы вернуться в обычный режим, нажмите **Свернуть**.
- **Редактировать виджет**. Вы можете задать количество отображаемых значений:
  1. Нажмите **Редактировать виджет**.

2. В окне **Редактировать виджет** выберите количество из выдающего списка.
3. Нажмите **Сохранить**.

- **Перейти из виджета в раздел Хранение файлов → Информация о файлах**, сохранив выбранные значения фильтров:
  1. Выберите элементы виджета. Для выбора нескольких элементов, нажмите на них, удерживая клавишу Ctrl. Для сброса выделения нажмите на пустой области виджета.
  2. Нажмите  для перехода к таблице с информацией о файлах.
     В открывшейся вкладке **Информация о файлах** будут переданы значения фильтров с вкладки **Файлы**, а также они будут дополнены значениями виджета, например, выбранными объектами защиты.

10 Раздел "Хранение файлов"

Раздел **Хранение файлов** содержит функциональность InfoWatch Data Discovery. В нем пользователь создает и запускает задачи сканирования, просматривает результаты: информацию о файлах, дубликатах и похожих, имена владельцев файлов, обнаруженные объекты защиты.

В разделе представлены вкладки:

- **Задачи** – создание и запуск задач сканирования, просмотр статистики их выполнения;
- **Информация о файлах** – таблицы с данными о файлах, полученными в результате сканирования ресурсов. Информация о файлах представлена в двух режимах отображения: Актуально и История;
- **Дубликаты файлов** – результаты автоматического поиска дубликатов файлов, по хешу или по размеру;
- **Похожие файлы** – результаты автоматического поиска документов с одинаковыми терминами (похожих файлов) с помощью модуля Data Analysis Service.

10.1 Вкладка "Задачи"

Вкладка содержит список задач сканирования и подробную информацию о каждой из них. О работе задач сканирования см. "Задача сканирования".

Основные действия при работе с разделом:

- Создание, настройка и запуск задачи (см. "Создание, запуск и настройка задачи"):
  - Добавление и настройка хостов (см. "Добавление и настройка хостов")
  - Добавление путей сканирования (см. "Добавление путей сканирования")
- Редактирование задачи и хоста (см. "Редактирование задачи и хоста")
- Копирование задачи (см. "Копирование задачи")
- Удаление задачи, хостов и путей сканирования (см. "Удаление задачи, хостов и путей сканирования")
- Просмотр статистики и информации о задачах (см. "Статистика выполнения задач")

10.1.1 Задача сканирования

Задача сканирования – операция проверки ресурсов (путей сканирования) на наличие конфиденциальных данных. Доступны два режима выполнения задачи (см. "Создание, запуск и настройка задачи", параметр **Выполнение задачи**):

- Непрерывное выполнение – Система циклично сканирует ресурсы, пока пользователь не остановит задачу вручную или пока она не будет остановлена автоматически по расписанию.
- Разовое выполнение – задача останавливается автоматически, если был выполнен один проход по каждому ресурсу.

При повторном запуске задачи не осуществляется полное повторное сканирование – обработка ресурсов продолжается с того места, где она закончилась в предыдущем сканировании.

Хост – рабочая станция или сервер, к которому осуществляется подключение для сканирования файлов. Хост содержит список ресурсов, которые называются *Пути сканирования*. При создании задачи сканирования в задачу необходимо добавить хосты, для каждого хоста необходимо добавить пути сканирования (подробнее про создание и настройку задачи см. "Создание, запуск и настройка задачи").

Проход – полное сканирование одного ресурса. Проход может состоять из нескольких сессий.

Сессия – подключение к ресурсу в рамках прохода.

Подробнее об обработке ресурсов см. "Сканирование ресурсов". О событиях, которые Система отправляет в Traffic Monitor см. "Событие Traffic Monitor".

Сканирование ресурсов

Типы ресурсов

Создав задачу, вы можете сканировать:

- общие сетевые ресурсы и распределенные файловые системы (DFS) по протоколу SMB:
  - SMB-ресурсы на хостах под управлением MS Windows;
  - SMB-ресурсы на хостах под управлением Linux;
- файловые системы серверов и рабочих станций под управлением Linux по протоколу SFTP, с аутентификацией по SSH;
- библиотеки документов в SharePoint Server 2016 по протоколу HTTPS;
- файловые ресурсы по протоколу FTP;
- хранилища Nextcloud по протоколу WebDAV;
- серверы с системой управления контентом Alfresco Community Edition по протоколам HTTP/HTTPS;
- хранилища Яндекс Диск по REST API Диска.

> **Примечание:**
>
> Подключение по протоколам **FTPS** (Explicit FTP over TLS) и **HTTPS** (SharePoint, Nextcloud и Alfresco) проходит без проверки подлинности TLS/SSL-сертификатов.
>
> Поддерживается сканирование:
>
> - Nextcloud версии 25 и более новых;
> - Alfresco Community Edition 5.2 или 7.1.1.
>
> В текущей версии не гарантируется сканирование рабочих станций под управлением ОС Astra Linux с включенным мандатным доступом.

Учетная запись для сканирования

Учетная запись, которую Data Discovery использует для сканирования ресурсов, должна иметь следующие права:

- Право на чтение общего сетевого ресурса (при SMB-подключении);
- Право на чтение директорий;
- Право на вывод списка содержимого директорий;
- Право на чтение файлов.

Значения параметров для учетной записи указываются в настройках задачи и в настройках хоста (подробнее см. Создание, запуск и настройка задачи).

Обработка файлов в рамках задачи

В рамках первого прохода по ресурсу Data Discovery выполняет следующие действия со всеми файлами на ресурсе:

- Собирает метаданные файла и добавляет их в таблицу с информацией о файлах (см. Вкладка "Информация о файлах"). Записи о новых файлах будут добавлены в таблицу в обоих режимах, **Актуально** и **История**.
- Если в блоке **Режим обработки** выбран хотя один бы один режим, создает копию файла в хранилище задачи.
- Если выбран режим **Отправлять файлы на анализ в Traffic Monitor**, отправляет событие на сервер Traffic Monitor для анализа. Событие состоит из копии файла и информации о файле.
- Если выбран режим **Проводить контентный анализ файлов**, извлекает текстовые контенты и отправляет в Модуль контентного анализа для обнаружения в них объектов защиты.
- Если выбран режим **Определять похожие файлы**, извлекает текстовые контенты и отправляет для анализа в Data Analysis Service.

> **Примечание:**
>
> Если размер файла превышает 4 ГБ, Data Discovery соберет только его метаданные.

При последующих сессиях Data Discovery обрабатывает только новые и измененные файлы. Чтобы определить измененные файлы, Data Discovery проверяет метаданные файлов на изменения:

- Путь к файлу;
- Имя файла;
- Размер файла;
- Дата и время изменения файла;
- Владелец файла;
- Пользователи/группы, которые имеют доступ к файлу.

> **Важно!**
>
> Проверка следующих метаданных на изменения поддерживается только при сканировании SMB-ресурсов на хостах под управлением ОС семейства Windows:
>
> - Владелец файла
> - Пользователи/группы, которые имеют доступ к файлу

Если файл является новым или измененным, то Data Discovery:

- Добавляет новую запись в таблицу в обоих режимах, **Актуально** и **История** (см. Вкладка "Информация о файлах"):
- Если файл был изменен, данные о его **предыдущем сканировании** будут отображены **только** в таблице режима **История**.
- Если в блоке **Режим обработки** выбран хотя бы один режим, создает копию файла в хранилище задачи.
- Если выбран режим **Отправлять файлы на анализ в Traffic Monitor**, отправляет событие на сервер Traffic Monitor для анализа. Событие состоит из копии файла и информации о файле.
- Если выбран режим **Проводить контентный анализ файлов**, извлекает текстовые контенты и отправляет в модуль контентного анализа для обнаружения в них объектов защиты.

- Если выбран режим **Определять похожие файлы**, извлекает текстовые контенты и отправляет для анализа в Data Analysis Service.

> **Примечание:**
>
> Если в расширенных настройках задачи включен параметр **Дополнительно проверять неизмененные файлы**, то при каждом проходе по ресурсу Data Discovery скачивает файл и проверяет содержимое файла на наличие изменений (подробнее про параметр см. "Создание, запуск и настройка задачи"). Если содержимое файла изменилось, то файл отправляется в составе события в Traffic Monitor.
>
> Если размер файла превышает 4 ГБ, Data Discovery соберет только его метаданные.

Ошибки при сканировании

При сканировании могут возникать ошибки, например:

- Data Discovery не может подключиться к ресурсу;
- у используемой учетной записи недостаточно прав для обработки директории/файла.

Если Data Discovery не может подключиться к ресурсу, он перейдет к обработке другого ресурса. В рамках задачи Data Discovery периодически будет пытаться повторно подключиться к этому ресурсу.

Data Discovery пропустит директорию или файл, если не может обработать их. При следующем подключении к ресурсу Data Discovery попытается повторно обработать пропущенную директорию или файл.

Событие Traffic Monitor

Data Discovery отправляет в Traffic Monitor события, которые содержат информацию о файлах. В каждом событии копия файла является вложением.

> **Примечание:**
>
> Тип события и перехватчик, указанные в событии в Traffic Monitor, зависят от версии Traffic Monitor:
>
> - Для Traffic Monitor версии 7.3 и ниже: тип события – *Краулер*, перехватчик – *Crawler;*
> - Для Traffic Monitor версии выше 7.3: тип события – *Data Discovery*, перехватчик – *Data Discovery*.

Traffic Monitor применяет к событиям Data Discovery специальный antigarbage-фильтр, вы можете отключить его работы или выбрать режим. Подробнее см. *InfoWatch Data Discovery. Руководство администратора. Настройки автоматического удаления событий Data Discovery в Traffic Monitor*.

Атрибуты события, которое формирует Data Discovery для отправки в Traffic Monitor:

Атрибут события	Описание
Отправители	Data Discovery отправляет SID владельца файла, например:

Атрибут события	Описание
	`S-1-5-21-1922741149-3613679056-2300624754-1000` Система получает значение из файловой системы ресурса. **Примечание:** Traffic Monitor может идентифицировать персону по полученному SID.
Получатели	Data Discovery отправляет список SID пользователей и групп, которые имеют доступ к файлам. Система получает значение из файловой системы ресурса. **Примечание: Traffic Monitor** может идентифицировать персону или группу по полученному SID.
Компьютер	IP-адрес хоста или доменное имя хоста, на котором был найден файл. **Примечание: Traffic Monitor** может идентифицировать рабочую станцию по полученной информации о хосте.
Название задачи	Название задачи сканирования в Data Discovery, в рамках которой файл был скачан с ресурса.
Дата запуска задачи	Дата и время начала сессии, в рамках которой файл был скачан с ресурса
Дата создания файла	Дата и время создания файла на хосте. Система получает значение из файловой системы ресурса. **Примечание:** Определение даты создания файла не поддерживается в следующих случаях: • при сканировании SMB-ресурсов на хостах под управлением Linux; • при сканировании хостов под управлением Linux по протоколу SSH. В этих случаях атрибут заполняется значением атрибута события *Дата изменения файла*.
Дата изменения файла	Дата и время изменения файла на хосте. Система получает значение из файловой системы ресурса.
Путь к файлу или адрес приемника	Полный путь к файлу на хосте

Атрибут события	Описание
	Примечание: Полный путь включает в себя путь сканирования ресурса. Для файлов с SharePoint-хостов полный путь также включает в себя путь к сайту.
Дата перехвата	Дата и время, когда файл был скачан Системой
Сервер перехвата	Сервер Data Discovery, который обрабатывал ресурс

Важно!

Получение SID владельца файла и SID пользователей/групп, которые имеют доступ к файлу, поддерживается для:

- SMB-ресурсов на хостах под управлением ОС семейства Windows;
- библиотек документов в SharePoint Server 2016.

Возможно некорректное заполнение атрибутов *Отправители* и *Получатели*, если:

- Производится сканирование SMB-ресурса на хосте под управлением Linux;
- Производится сканирование SMB-ресурса, созданного на сервере SMB, который не поддерживает файловые атрибуты для определения владельца файла и пользователей/групп, которые имеют доступ к файлу;
- Производится сканирование SMB-ресурса, созданного в файловой системе, которая не поддерживает файловые атрибуты для определения владельца файла и пользователей/групп, которые имеют доступ к файлу.

Получение SID владельца файла и SID пользователей/групп, которые имеют доступ к файлу, не поддерживается при сканировании хостов:

- под управлением Linux по протоколу SSH;
- по протоколу FTP;
- по протоколу WebDAV.

В этом случае атрибуты *Отправители* и *Получатели* имеют значение "iw data discovery".

10.1.2 Создание, запуск и настройка задачи

Создание задачи

Чтобы создать новую задачу сканирования:

1. В разделе **Хранение файлов** перейдите на вкладку **Задачи**, чтобы увидеть список задач сканирования.
2. В левом верхнем углу нажмите 

.
3. В открывшемся окне настройте задачу и нажмите **Сохранить** (см. *Настройки задачи*).
4. Добавьте в задачу хосты (см. *Добавление и настройка хостов*).
5. Для хостов добавьте пути сканирования (см. *Добавление путей сканирования*).

Важно!

Вы не можете запустить задачу, если в задаче нет путей сканирования.

См. также *Примеры настроек задачи, хоста и добавления путей сканирования*.

Запуск задачи сканирования

Новая задача имеет статус:

- **Новая**, если для задачи не был настроен автоматический запуск по расписанию;
- **Ожидает запуск по расписанию**, если для задачи был настроен автоматический запуск по расписанию.

Для запуска задачи нажмите  в столбце **Действия** в списке задач сканирования. При этом статус задачи изменится на **Выполняется**, а кнопка запуска задачи изменится на . Нажмите , чтобы остановить задачу.

Важно!

Вы можете запустить задачу сканирования, только если в Системе есть активная лицензия (см. *InfoWatch Data Discovery. Руководство администратора. Лицензирование*).

Подробнее о запущенной задаче можно узнать, просмотрев статистику (см. *Статистика выполнения задач*).

Настройки задачи

Настройки задачи сканирования расположены на нескольких вкладках:

- Вкладка Общие:
- Вкладка Фильтры файлов
- Вкладка Фильтры папок
- Вкладка Расписание
- Вкладка Расширенные настройки
- Вкладка Логирование

Примечание:

Поля, обязательные к заполнению, отмечены в интерфейсе значком .

Вкладка **Общие**:

Основные настройки задачи

Параметр	Описание
Название	Название задачи.
Описание	Краткое описание задачи.
Выполнение задачи	Режим выполнения задачи. **Непрерывное**: • Система циклично сканирует ресурсы, пока пользователь не остановит задачу вручную или пока она не будет остановлена автоматически по расписанию. • При повторном запуске задачи не осуществляется полное повторное сканирование – обработка ресурсов продолжается с того места, где она закончилась в предыдущем сканировании. **Разовое**: • Система выполняет один проход по каждому ресурсу, после чего задача считается завершенной. Задача автоматически останавливается и получает статус **Завершена**. • Если запустить задачу, которая имеет статус **Завершена**, то Система: a. Обработает новые и измененные файлы на каждом ресурсе. b. Остановит задачу и установит статус **Завершена**. • Задача получает статус **Завершена**, если был выполнен проход по каждому доступному ресурсу. Задача со статусом **Завершена** может содержать ресурсы со статусом **Ошибка**, которые Система не смогла обработать. • Если включены режимы обработки, задача получит статус **Завершена**: • для режима **Отправлять файлы на анализ в Traffic Monitor** – когда все данные будут отправлены в Traffic Monitor; • для режима **Проводить контентный анализ файлов** – когда текстовые контенты извлечены и готовы к отправке в Модуль контентного анализа для обработки;

Параметр	Описание
	• для режима **Определять похожие файлы** – когда в Data Analysis Service будет отправлено уведомление о наличии данных для обработки. • Задача может запускаться и останавливаться по расписанию. Задача остановится, если Система выполнила один проход по каждому ресурсу или если наступило время автоматической остановки по расписанию. **Важно**: в обоих случаях задача получит статус **Ожидает запуск по расписанию**.
Режим обработки	Режим обработки файлов. В рамках задачи сканирования Data Discovery собирает метаданные файла и добавляет их в таблицу с информацией о файлах. Вся детальная информация о файлах доступна для просмотра на вкладке Информация о файлах. Также Data Discovery создает копии файлов для анализа. Для этого должен быть выбран хотя бы один из режимов обработки: • **Отправлять файлы на анализ в Traffic Monitor** – для отправки копий файлов на анализ в разделе **Настройки** → **Настройки продуктов** → **Зоны и серверы** должен быть указан доступный для соединения сервер Traffic Monitor. Во избежание дублирования нагрузки рекомендуем не включать отправку файлов в Traffic Monitor, если вы будете использовать OCR в Data Discovery. • **Проводить контентный анализ файлов** – Data Discovery извлекает текстовые контенты и для обнаружения в них объектов защиты использует отдельный Модуль контентного анализа. Модуль контентного анализа поддерживает следующие технологии из конфигурации Traffic Monitor: • Автолингвист • Бланки • Выгрузки из БД • Категории и термины • Текстовые объекты • Эталонные документы Про установку модуля см. *InfoWatch Data Discovery. Руководство по установке. Установка Data Discovery*. **Чтобы режим был доступен:** • установите Модуль контентного анализа в одном кластере с Data Discovery; • добавьте в Центр расследований объекты защиты из конфигурации Traffic Monitor (подробнее см. *InfoWatch Data Discovery. Руководство администратора. Интеграция с InfoWatch Traffic Monitor*). Рекомендуем добавлять объекты защиты

Параметр	Описание
	из той же конфигурации Traffic Monitor, что была использована при установке Модуля контентного анализа; • добавьте в зону выполнения задачи сканирования сервер с ролью **Контентный анализ** (подробнее см. *InfoWatch Data Discovery. Зоны и серверы).* • **Определять похожие файлы** – Data Discovery извлекает текстовые контенты и для поиска похожих файлов использует модуль InfoWatch Data Analysis Service. Про установку и настройку Data Analysis Service см. *InfoWatch Data Analysis Service. Руководство по установке, конфигурированию и администрированию*. **Чтобы режим был доступен:** • установите **Data Analysis Service 1.5.0** или **1.6.0** одном кластере с **Data Discovery**; • включите в **Data Analysis Service** функцию протокластеризации; • установите в **Data Analysis Service** модель `proto_clusterization_resources`; • настройте **Data Analysis Service** на обработку ноды, на которой установлен **Data Discovery**. Если не выбрать ни один режим обработки, **Data Discovery** будет собирать только метаданные файлов, не создавая их копии. **Примечание**: чтобы файл был обработан на определение похожих, извлеченный **текстовый контент** должен быть **не менее 200 байт**. Контент меньше 200 байт не будет отправлен на анализ в Data Analysis Service, но будет доступен для просмотра в веб-интерфейсе Центра расследований в разделе **Хранение файлов**.

Секция "Авторизация для доступа к ресурсам"

Параметр	Описание
По логину и паролю	
Группа SMB	Домен или группа. Параметр используется для хостов с SMB-подключением.
Логин	Имя пользователя.

	Важно: для подключения к SharePoint нужно указать значение в формате: domain\user
Пароль	Пароль.
По SSH-ключу (только для типа подключения SSH)	
Логин	Имя пользователя, которому был передан публичный ключ.
Ключевая фраза	Пароль для зашифрованного приватного ключа.
Приватный ключ авторизации	Приватный ключ. Ключ может быть указан в поле вручную или загружен из файла с помощью кнопки  . Пример использования SSH-ключей см. в статье *Использование SSH-ключей для доступа к ресурсам*.

Примечание:

Если значения параметров авторизации не указаны, то по умолчанию для подключения к SMB-хостам в рамках задачи будет использоваться анонимный доступ.

Чтобы ускорить сканирование хранилища Nextcloud, для подключения используйте пароль приложения (подробнее см. примечание в пункте "Сторонние клиенты WebDAV" из официальной документации Nextcloud).

Вкладка **Фильтры файлов**

Фильтр по размеру

Фильтр **Размер** позволяет установить ограничения на размер файла. Если размер файла подходит под условие, установленное в фильтре, то файл будет просканирован.

- В поле **От** укажите минимальный размер файла. Вы не можете указать значение больше 4 ГБ.
- В поле **До** укажите максимальный размер файла. Вы не можете указать значение больше 4 ГБ.
- Справа от каждого из указанных полей фильтра отображается текущая единица измерения. Чтобы изменить единицу измерения для поля, нажмите на единицу измерения и выберите нужное значение в раскрывающемся списке.

Примечание:

Если указан только минимальный размер файла, то будут просканированы файлы, чей размер больше или равен минимальному размеру.
Если указан только максимальный размер файла, то будут просканированы файлы, чей размер не превышает максимальный размер.

Примечание:

Если размер файла превышает 4 ГБ, то Система не скачивает файл и не отправляет событие в **Traffic Monitor**. Запись о файле добавляется в таблицу с информацией о файлах (см. *Таблица с информацией о файлах*).

Маски имен файлов

Система сканирует файл, если имя файла подходит под условие маски. **Список добавленных масок** содержит маски имен файлов.

Важно!

Если не указать ни одну маску, Система будет сканировать файлы с любыми именами, включая системные файлы. Отправка большого количества файлов на сервер **Traffic Monitor** может значительно повысить нагрузку на **Traffic Monitor**.

Действия с масками:

Чтобы добавить маску:

1. Нажмите на поле **Добавьте маску файла**. В открывшемся списке отображаются предустановленные категории файлов и маски для форматов.
2. Выполните одно из действий:
   - выберите маски/категории в списке. Для поиска начните вводить маску или имя категории в поле;
   - введите маску в поле и выберите эту маску в списке.
3. Кликните вне списка.

Чтобы удалить маску/категорию, нажмите  справа от маски/категории.

Чтобы экспортировать список масок в файл, нажмите .

Вы можете импортировать список масок из файла. Нужен файл с расширением `.txt`, каждая маска должна быть указана на отдельной строке.
Чтобы импортировать список масок:

1. Нажмите .
2. В открывшемся окне выберите файл.

Примечание:

В результате импорта в задаче останутся только маски, указанные в файле, а остальные будут удалены.

Для составления маски имени используйте следующие **символы**:

- символ "*****" – означает любую последовательность символов, в том числе отсутствие символов;
- символ "**?**" – означает один любой символ.
  Примеры масок

  ***.doc** – файлы, имена которых оканчиваются на ".doc" (example.doc, отчет.doc).

  отчет – файлы, имена которых включают "отчет" (новый_отчет.txt, отчетность.pdf).

  отчет????.pdf – файлы с именами вида: "отчет" + 4 символа + ".pdf" (отчет2012.pdf, отчеты_22.pdf).

  file.txt – файлы, у которых имя "file.txt".

- фильтр по именам файлов не учитывает регистр.

Вкладка **Фильтры папок**

Чтобы исключить директорию и все ее поддиректории из сканирования, введите полный путь к этой директории в поле **Исключить** и нажмите . Путь к директории будет добавлен в **Список исключенных папок**.

Примечание:

По умолчанию **Список исключенных папок** содержит системные директории для административных ресурсов Windows.

Чтобы включить директорию и все ее поддиректории в сканирование, введите полный путь к этой директории в поле **Включить** и нажмите . Путь к директории будет добавлен в **Список включенных папок**. Если в **Список включенных папок** не добавлена ни одна директория, то будут просканированы все директории, кроме исключенных из сканирования.

Чтобы удалить директорию из списка включенных/исключенных папок, нажмите  справа от директории в списке.

Важно!

Исключение директорий из сканирования имеет более высокий приоритет, чем включение директорий в сканирование. Например, если исключить директорию `share/documents/folder` и включить директорию `share/documents/folder/example`, то директория `share/documents/folder/example` обработана не будет.

Формат пути к директории:

- Полный путь к директории – последовательность из имен директорий, разделенных символом "/", которая указывает на расположение директории в файловой системе.
- Для **SMB** полный путь к директории должен включать в себя имя ресурса. Полный путь к директории не должен начинаться с символа "/".
- Для **FTP**, **SFTP** (**SSH**), **Nextcloud**, **Alfresco** и **Яндекс Диск** полный путь к директории должен включать в себя путь сканирования. Полный путь к директории должен начинаться с символа "/".
- Для **SharePoint**-ресурсов полный путь к директории должен включать в себя имя библиотеки документов. Полный путь к директории не должен начинаться с символа "/".
- Фильтры по директориям не учитывают регистр.

Примеры путей к директориям

Для **SMB**:

- Если нужно просканировать директорию `/folder/example` и все ее поддиректории на общем сетевом ресурсе `share`, то добавьте в **Список включенных папок** путь `share/folder/example`
- Если нужно исключить из сканирования директорию `/system/files` и все ее поддиректории на общем сетевом ресурсе `resource`, то добавьте в **Список исключенных папок** путь `resource/system/files`

Для **SFTP (SSH)**:

- Если в качестве пути сканирования на SSH-хосте указан путь `/home/folder` и нужно просканировать директорию `/home/folder/example` и все ее поддиректории, то добавьте в **Список включенных папок** путь `/home/folder/example`
- Если в качестве пути сканирования на SSH-хосте указан путь `/home/documents` и нужно исключить из сканирования директорию `/home/documents/old_files` и все ее поддиректории, то добавьте в **Список исключенных папок** путь `/home/documents/old_files`

Для **Alfresco**, **FTP** и **Nextcloud** формат указания путей к директориям аналогичен **SFTP (SSH)**.

Для **Яндекс Диск**:

- Для пользовательского диска укажите путь в формате `/users/<e-mail_пользователя>/<имя_папки>`.
  Пример пути: `/users/ivanov@organization.ru/secret_folder`
- Для общего диска организации укажите путь в формате `/shared/<имя_общего_диска>/<имя_папки>`.
  Пример пути: `/shared/head_office_disk/secret_folder`

Для **SharePoint**:

- Если нужно просканировать директорию `/folder/example` и все ее поддиректории в библиотеке `library`, то добавьте в **Список включенных папок** путь `library /folder/example`
- Если нужно исключить из сканирования директорию `/outdated/reports` и все ее поддиректории в библиотеке `documents`, то добавьте в **Список исключенных папок** путь **`documents/outdated/reports`**

- **Маски для имен директорий:**
  Имя каждой директории в полном пути к директории можно указать в формате маски. Для составления маски имени используйте следующие символы:
  - символ "*****" – означает любую последовательность символов, в том числе отсутствие символов;
  - символ "**?**" – означает один любой символ.

Важно!

Символы "*****" и "**?**" используются только в рамках имени одной директории. Например, путь `share/folder/subfolder/example` не подходит под условие маски `share/*/example`, но подходит под условие маски `share/*/*/example`

Примеры путей с использованием масок

Тип хоста	Пример пути	Описание	Директории, подходящие под условие
SMB	`*/Example`	Директория **Example** и все ее поддиректории, если **Example** находится в корневой директории ресурса.	`C$/Example` `share/example`
	`*/*/*/Example`	Директория **Example** и все ее поддиректории, если **Example** находится на три уровня ниже корневой директории ресурса.	`myshare/path/to/Example`
	`share/Doc*`	Директория и все ее поддиректории, если директория находится в корневой директории	`share/Doc` `share/documents`

Тип хоста	Пример пути	Описание	Директории, подходящие под условие
		ресурса **share** и ее имя начинается с "**Doc**".	
	`?$/Users/*/Documents`	Директория **Documents** и все ее поддиректории у каждого пользователя на скрытом ресурсе.	`C$/Users/Ivanov/Documents` `D$/Users/Petrov/Documents`
	`share/documents/*/*`	Все поддиректории для директории **documents** на ресурсе **share**, но не сама директория **documents**. Пути такого вида поддерживаются только для включающего фильтра (**Список включенных папок**).	`share/documents/reports` `share/documents/templates`
SharePoint	`*/Example`	Директория **Example** и все ее поддиректории, если **Example** находится в корневой директории библиотеки документов.	`my_library/example`
	`*/*/*/Example`	Директория **Example** и все ее поддиректории, если **Example** находится на три уровня ниже корневой директории библиотеки документов.	`documents/2021/folder/Example`
	`Library/Doc*`	Директория и все ее поддиректории, если директория находится в корневой директории библиотеки **Library** и имя директории начинается с "**Doc**".	`Library/Documents`
	`storage/documents/*/*`	Все поддиректории для директории **documents** в библиотеке **storage**, но не сама директория **documents**. Пути такого вида поддерживаются только для включающего фильтра (**Список включенных папок**).	`storage/documents/reports` `storage/documents/templates`

Тип хоста	Пример пути	Описание	Директории, подходящие под условие
SSH, FTP, WebDAV, Alfresco	**`/home*`**	Директория и все ее поддиректории, если директория находится в корневой директории и ее имя начинается с "**home**".	`/home` `/HomeFolder`
	`/home/*/*/Example`	Директория **Example** и все ее поддиректории, если **Example** находится на 3 уровня ниже папки **home**.	`/home/path/to/Example`
	`/home/*/Documents`	Директория **Documents** и все ее поддиректории у каждого пользователя на рабочей станции под управлением Linux (для SSH-хоста).	`/home/ivanov/Documents`
	`/home/*/*`	Все поддиректории для директории **home**, но не сама директория **home**. Пути такого вида поддерживаются только для включающего фильтра (**Список включенных папок**).	`/home/username` `/home/folder`
Яндекс Диск	**`/users/*/Documents`**	Директория **Documents** всех пользовательских дисков.	`/users/ivanov@corp.ru/Documents` `/users/ripley@corp.ru/Documents`
	`/shared/*`	Все директории и поддиректории на общем диске.	`/shared/PR/active` `/shared/HR/welcome`

Вкладка **Расписание**

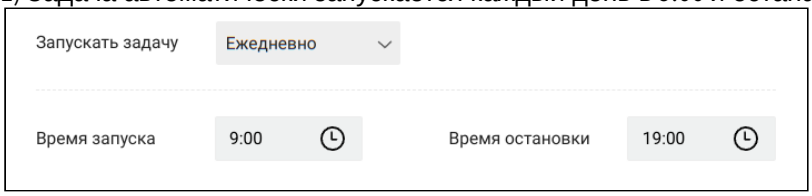
Настройки для автоматического запуска и остановки задачи:

Важно!

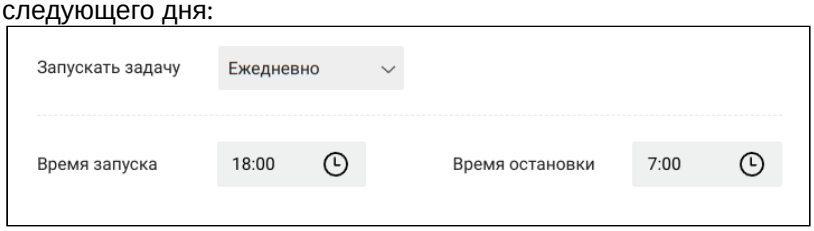
Задача не будет запущена автоматически, если в Системе отсутствует активная лицензия (см. InfoWatch Data Discovery. *Руководство администратора. Лицензирование*).

- Если задачу не нужно запускать и останавливать автоматически, в поле **Запускать задачу** выберите **Вручную**.
- Если задачу нужно запускать и останавливать автоматически каждый день в одно и то же время, в поле **Запускать задачу** выберите **Ежедневно**. В поле **Время запуска** и в поле **Время остановки** укажите время в формате H:mm или HH:mm.
  Примеры ежедневного расписания

  1) Задача автоматически запускается каждый день в 9:00 и останавливается в 19:00:

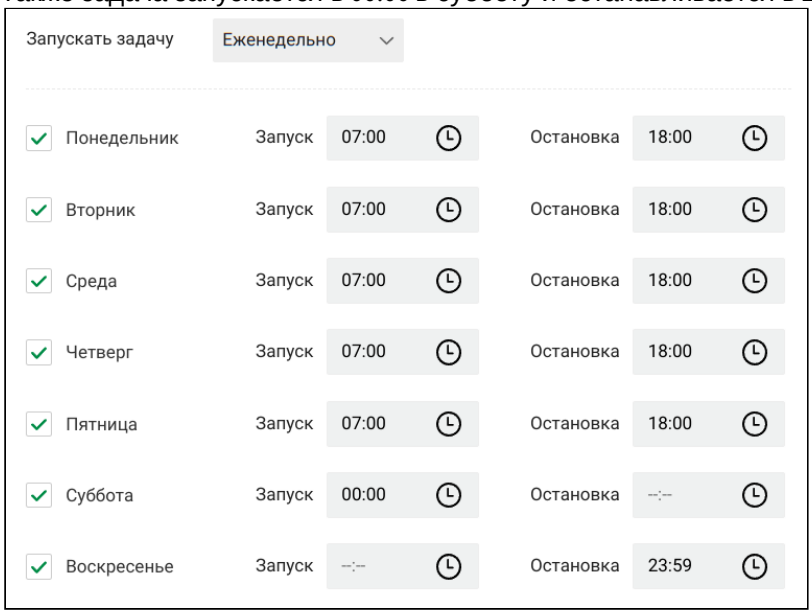


  2) Задача автоматически запускается каждый день в 18:00 и останавливается в 7:00 следующего дня:

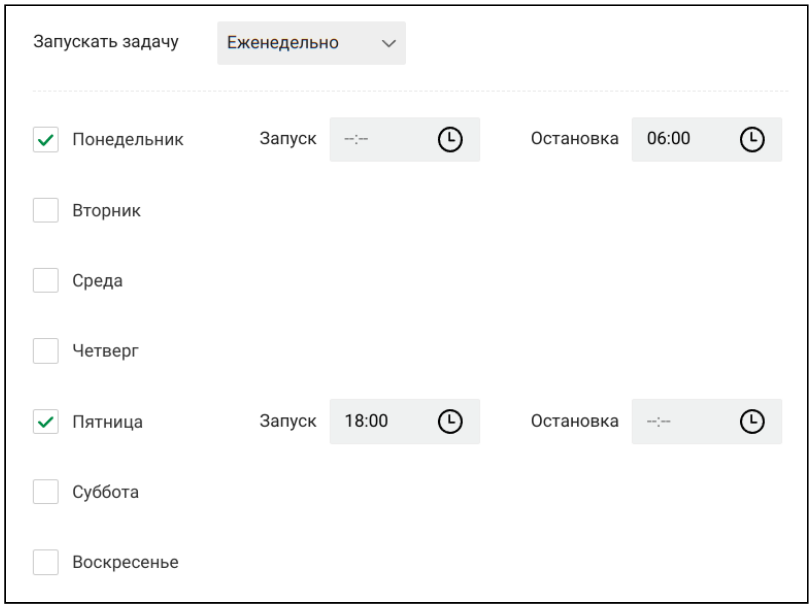


- Если задачу нужно запускать и останавливать автоматически в различные дни недели, в поле **Запускать задачу** выберите **Еженедельно**. Установите флажок слева от того дня недели, в который нужно запускать или останавливать задачу. В поле **Запуск** и/или в поле **Остановка** справа от выбранного дня укажите время в формате H:mm или HH:mm.
  Примеры еженедельного расписания

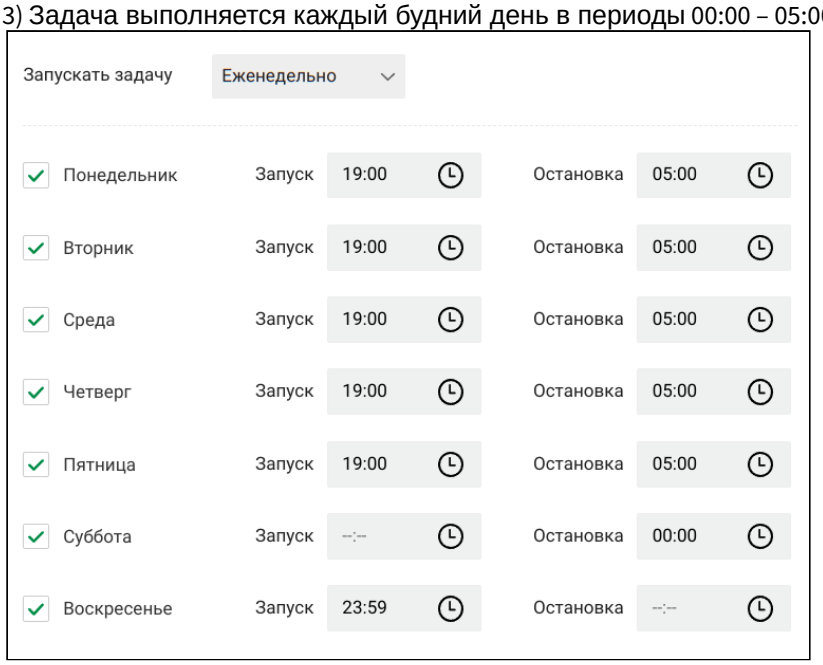
  1) Задача автоматически запускается каждый будний день в 07:00 и останавливается в 18:00. Также задача запускается в 00:00 в субботу и останавливается в 23:59 в воскресенье:



  2) Задача автоматически запускается в пятницу в 18:00 и останавливается в понедельник в 06:00:



3) Задача выполняется каждый будний день в периоды 00:00 – 05:00 и 19:00 – 23:59:



Примечание:

При сохранении расписания указанное время добавляется в Систему в часовом поясе рабочей станции, с которой осуществляется подключение к веб-интерфейсу.

При просмотре расписания время отображается в часовом поясе рабочей станции, с которой осуществляется подключение к веб-интерфейсу.

Примечание:

Если для задачи настроено расписание, ее можно также запускать и останавливать вручную. При этом автоматический запуск и остановка задачи будут выполняться по расписанию.

Вкладка **Расширенные настройки**

Настройка задачи с учетом особенностей инфраструктуры:

Параметр	Описание
Число сканеров	Количество сканеров (экземпляров сервиса Puller). **Data Discovery** использует сканеры для сканирования ресурсов задачи и загрузки файлов с ресурсов. Если указано значение **0**, то Система: • не сканирует файлы и не скачивает новые файлы в хранилища; • отправляет файлы из хранилищ в **Traffic Monitor**. **Примечание**: Для сканирования одного хоста в рамках задачи используется только один сканер. Не рекомендуем указывать число сканеров, которое превышает число хостов в задаче. Если у хоста большое количество путей сканирования, то рекомендуем распределить их по разным задачам. Для этого: 1. Создайте несколько задач. 2. В каждую задачу добавьте хост. 3. В каждой задаче для хоста укажите несколько путей сканирования. Каждый путь сканирования хоста должен быть указан только в одной из задач. Если хост указан в нескольких задачах, то он будет обрабатываться несколькими сканерами. В этом случае может значительно увеличиться нагрузка на хост во время сканирования.
Число передатчиков	Количество передатчиков (экземпляров сервиса Sender). **Data Discovery** использует передатчики для отправки данных в **Traffic Monitor**. Рекомендуем указать значение, которое больше или равно количеству серверов **Data Discovery** в зоне задачи. Параметр доступен, если включен режим обработки **Отправлять файлы на анализ в Traffic Monitor**.
Число экстракторов	Количество экстракторов. Каждый экстрактор представлен парой сервисов Ripper и Contents Processor. **Data Discovery** использует экстракторы для извлечения из файлов текстовых контентов, сохранения их в файловое хранилище и уведомления **Data Analysis Service** о наличии данных для обработки. Параметр доступен, если включен хотя бы один из режимов обработки: **Определять похожие файлы** или **Проводить контентный анализ файлов**.

Параметр	Описание
Интервал обработки (в секундах)	Пауза после завершения прохода по ресурсу. Система повторно просканирует ресурс, только если указанный интервал закончился и очередь вновь дошла до ресурса. Параметр позволяет уменьшить нагрузку на ресурсы.
Приоритет обработки файлов в сессии (в днях)	Размер окна сканирования ресурса. Если дата изменения файла не попадает в окно сканирования, то файл не будет обработан в текущую сессию. Параметр используется для равномерного сканирования ресурсов. Чем ниже значение, тем чаще сканер будет переключаться между ресурсами. Параметр используется, если включен хотя бы один из режимов обработки. В первой сессии сканирования в рамках прохода по ресурсу в качестве окна сканирования устанавливается период от *Newest* до *Newest − Depth*, где: • *Newest* – самая поздняя дата изменения у файлов на ресурсе; • *Depth* – количество дней, указанное в параметре. В каждой следующей сессии в качестве окна сканирования устанавливается период от *Last* до *Last − Depth*, где: • *Last* – самая ранняя дата изменения у файлов, обработанных в предыдущую сессию; • *Depth* – количество дней, указанное в параметре. Новые и измененные файлы обрабатываются в **каждой** сессии. **Пример использования параметра** На ресурсе находятся 4 файла: • файл A, дата изменения которого – 28.05.2025; • файл B, дата изменения которого – 03.05.2025; • файл C, дата изменения которого – 15.04.2025; • файл D, дата изменения которого – 19.03.2025. Если значение параметра равно 30 дням, то: 1. В первой сессии в рамках прохода в качестве окна сканирования устанавливается период 28.05.2025 – 28.04.2025. Обрабатывается файл A и файл B. Сканер переключается на другой ресурс. 2. Во второй сессии в качестве окна сканирования устанавливается период 02.05.2025 – 02.04.2025. Обрабатывается файл C. Сканер переключается на другой ресурс. 3. В третьей сессии в качестве окна сканирования устанавливается период 14.04.2025 – 15.03.2025. Обрабатывается файл D. Проход по ресурсу завершается. **Примечание:** Если вы укажете значение 0, **Data Discovery** будет сканировать все файлы независимо от даты изменения.

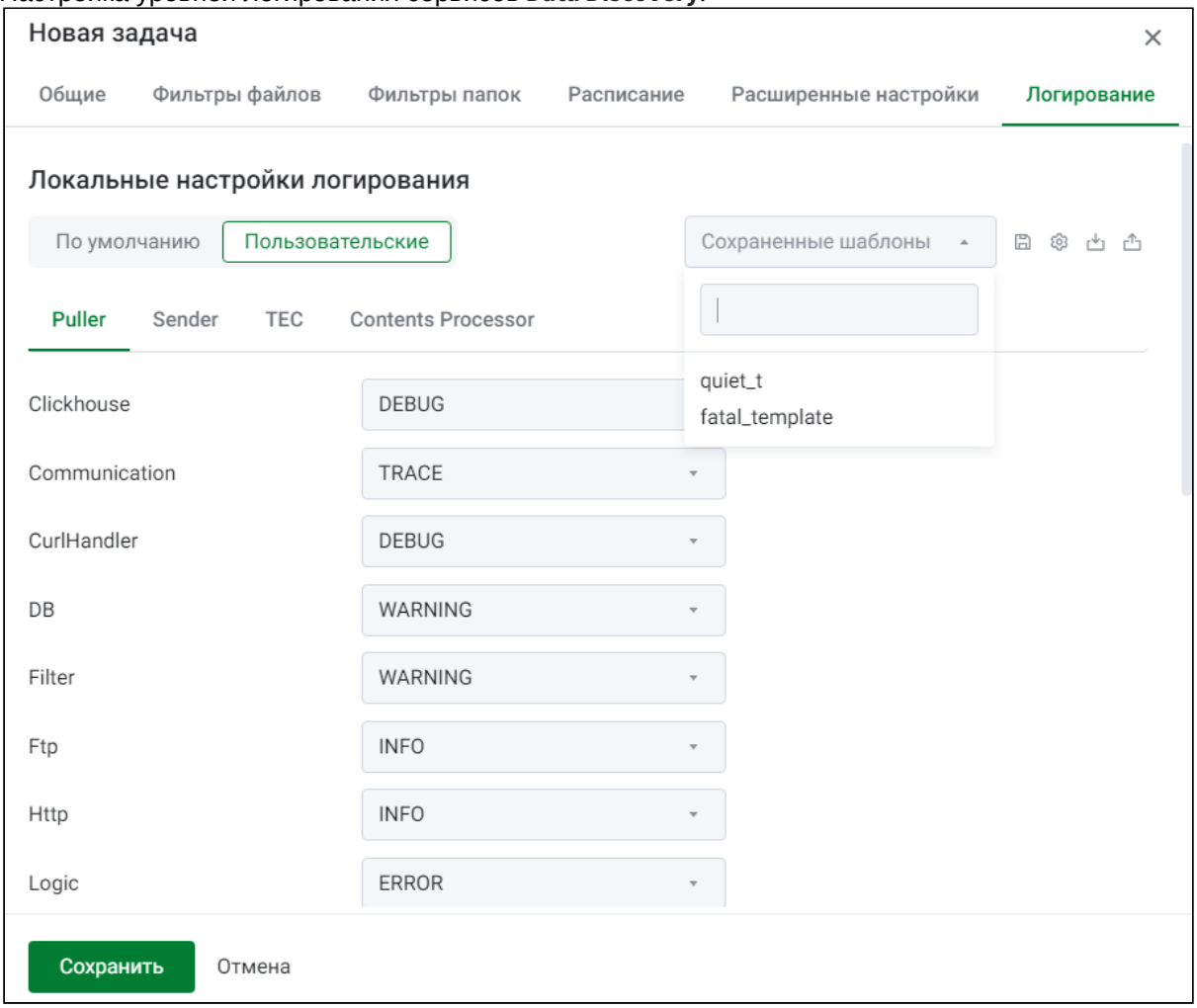
Параметр	Описание
Не обрабатывать файлы старше (в днях)	Ограничение файлов для сканирования по дате изменения. Значение параметра – количество дней до запуска задачи. Параметр определяет интервал дат изменения файлов для сканирования. Файлы с датой изменения **вне интервала просканированы не будут**. Дата обновляется при старте задачи. Если задача сканирует ресурс за несколько сессий, в эти сессии интервал не будет обновлен. Обновление интервала произойдет при следующем старте задачи. **Примечание:** Если вы укажете значение 0, **Data Discovery** будет сканировать все файлы независимо от даты изменения. **Пример использования параметра** На ресурсе находятся 4 файла: • файл A, дата изменения которого – 28.05.2025; • файл B, дата изменения которого – 11.05.2025; • файл C, дата изменения которого – 29.04.2025; • файл D, дата изменения которого – 12.04.2025. Если вы запустите задачу сканирования *30.05.2025* в *14:00* со значением параметра *30*, то будут просканированы только файлы A и B, так как их даты изменения попадают в интервал *30.04.2025, 14:00 – 30.05.2025, 14:00*. Если вы запустите задачу со значением 0, параметр никак не ограничит файлы для сканирования, и **Data Discovery** обработает все 4 файла.
Длительность обработки (в минутах)	Максимальная длительность сессии. Сессия включает в себя получение списка файлов, хранящихся на ресурсе, и обработку файлов. По истечении времени, указанного в параметре, обработка ресурса будет приостановлена, сканер начнет обработку следующего ресурса. Параметр используется для равномерного сканирования ресурсов. Чем ниже значение, тем чаще сканер будет переключаться между ресурсами. **Примечание**: Когда сканер повторно дойдет до ресурса, обработка которого была приостановлена по истечении максимального времени, обработка продолжится с того места, где она была приостановлена в предыдущую сессию. Если вы укажете значение 0, **Data Discovery** будет сканировать ресурс до полного завершения, без ограничения по времени. **Важно**: При каждом подключении к ресурсу сканер получает список файлов, хранящихся на ресурсе. Во время получения списка файлов ресурс имеет статус **Сбор информации**. Обработка ресурса начинается после получения списка файлов. Если на ресурсе находится большое количество файлов, то получение списка файлов может занять некоторое время. В некоторых случаях получение списка файлов может занимать все время, отведенное на одну сессию обработки ресурса. В этих случаях файлы на ресурсе не будут обрабатываться. Рекомендуем увеличить значение параметра **Длительность обработки** при сканировании ресурсов, которые имеют большое количество файлов.

Параметр	Описание
Настройка DNS	Вы можете выбрать: • **Использовать системные настройки** – значение по умолчанию; • **Указать адреса DNS вручную** – например, если сервер **Data Discovery** не в одном домене с хостами, которые нужно просканировать, укажите DNS-серверы, с помощью которых адреса хостов будут определены. Чтобы добавить дополнительное поле ввода адреса DNS-сервера, нажмите кнопку . Вы можете указать до трех адресов DNS-серверов.
Тип аутентификации SMB	Для SMB-подключения доступен выбор протокола аутентификации: • **NTLM** – значение по умолчанию; • **Kerberos**. Для протокола **Kerberos** нужно указать: • **Имя Realm** – имя области (Realm) Kerberos, в которой находятся хосты. Имя области чувствительно к регистру; • **Адрес KDC** – адрес центра распространения ключей (KDC, Key Distribution Center) для указанной области.
Зона	Зона задачи. • Сервисы задачи запускаются на серверах, которые добавлены в зону. • Система скачивает файлы в хранилища серверов зоны. • Задача использует настройки зоны для экспорта данных в **Traffic Monitor**. Вы можете изменить зону задачи, только если задача имеет статус **Новая**. Чтобы изменить зону в других случаях: a. В настройках задачи на вкладке **Расписание** в поле **Запускать задачу** выберите **Вручную** и сохраните настройки. b. Скопируйте задачу. c. Измените зону у копии задачи. Подробнее о зонах и серверах см. *InfoWatch Data Discovery*. *Руководство администратора. Зоны и серверы*.
Дополнительно проверять неизмененные файлы	По умолчанию для файлов, которые были скачаны во время предыдущих проходов по ресурсу, сканер проверяет только метаданные. Если метаданные файла не изменились, то файл не будет обработан повторно. Если параметр включен (), то при каждом проходе по ресурсу сканер будет скачивать файл и проверять содержимое файла на наличие изменений. Это позволяет выявить файлы, у которых изменилось содержимое, но не изменились метаданные. Если содержимое файла изменилось, то файл отправляется в составе события в **Traffic Monitor**. Параметр используется, если включен хотя бы один из режимов обработки.

Параметр	Описание
	Важно: если параметр включен ( ), то увеличится объем трафика и время, требуемое для повторной обработки ресурса.
Применять распознавание текста на изображениях (OCR)	Позволяет распознавать текст на изображениях и отправлять эти контенты на анализ. О выборе OCR-экстрактора см. *InfoWatch Data Discovery. Руководство администратора*. *Выбор OCR-экстрактора и параметры работы по умолчанию*. Чтобы применять распознавание текста на изображениях (OCR), на вкладке **Общие** включите хотя бы один из режимов обработки с экстракцией данных из файлов: **Проводить контентный анализ** и/или **Определять похожие файлы**. Во избежание дублирования нагрузки рекомендуем не включать отправку файлов в **Traffic Monitor**, если вы будете использовать OCR в **Data Discovery**. При включении параметра можно указать в секундах **Максимальное время обработки изображения (OCR)**. Если за указанное время OCR-экстрактор не успеет обработать изображение, он пропустит его и перейдет к следующему. Если к следующему сканированию этой же задачей файл останется без изменений, OCR-экстрактор не будет повторно обрабатывать пропущенное изображение.

Вкладка **Логирование**

Настройка уровней логирования сервисов **Data Discovery**.



Доступны настройки **По умолчанию** или **Пользовательские**.
Если выбраны **Пользовательские** настройки, вы можете задать уровень логирования для каждого компонента сервисов:

- **Puller**
- **Sender**
- **TEC** (Task Execution Controller)
- **Contents Processor**

> **Примечание:**
>
> **Data Discovery** сохраняет лог-файлы в директории `/var/log/infowatch/disc`

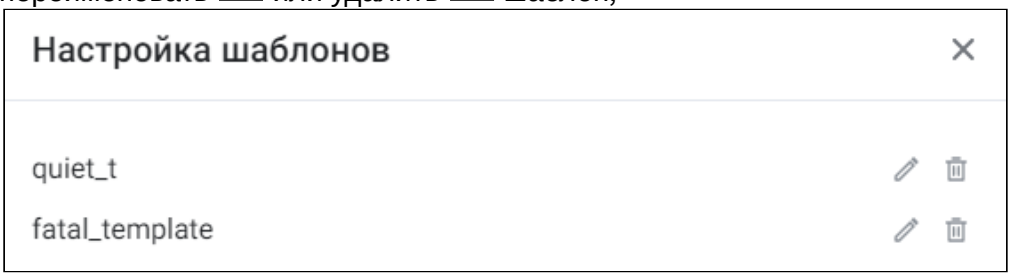
Значения уровней логирования:

- `trace` – наиболее подробное логирование, включая все сообщения с уровнями `debug`, `info` и выше;
- `debug` – вывод отладочной информации, обычно для диагностики ошибок и общения с разработкой/поддержкой;

- `info` – информационные сообщения, могут быть отключены на стабильной системе;
- `warning` – предупреждения;
- `error` – ошибки;
- `fatal` – наиболее важные сообщения, приводящие к завершению работы или отказу.

Вы можете сохранять пользовательские настройки логирования в виде шаблонов, чтобы при необходимости быстро переключаться между ними.
Справа от поля **Сохраненные шаблоны** нажмите:

- – чтобы **сохранить** изменения настроек логирования в новом шаблоне. В открывшемся окне укажите имя нового шаблона;
  Чтобы использовать шаблон, нажмите на поле **Сохраненные шаблоны** и выберите его из списка. После выбора шаблона значения уровней логирования поменяются.
- – чтобы открыть **настройки шаблонов**. В открывшемся окне вы можете переименовать или удалить шаблон;



- – чтобы **экспортировать** все сохраненные шаблоны. Они будут сохранены в файле `presets.json` на вашем компьютере;
- – чтобы **импортировать** шаблоны. В открывшемся окне выберите экспортированный с другого сервера JSON-файл.

Добавление и настройка хостов

Добавление хостов

> **Важно!**
>
> **Добавление хостов из LDAP-каталогов и импорт списка** хостов доступны **только для** хостов с **SMB**-подключением. При добавлении для всех хостов в параметре **Тип подключения** по умолчанию указывается **SMB**. Изменение значения данного параметра недоступно при редактировании добавленных хостов.
>
> Для всех хостов по умолчанию параметр **Наследовать из задачи** включен.

Для добавления хостов нажмите на задачу и используйте один из способов:

- Автоматическая синхронизация с LDAP
- Добавление вручную

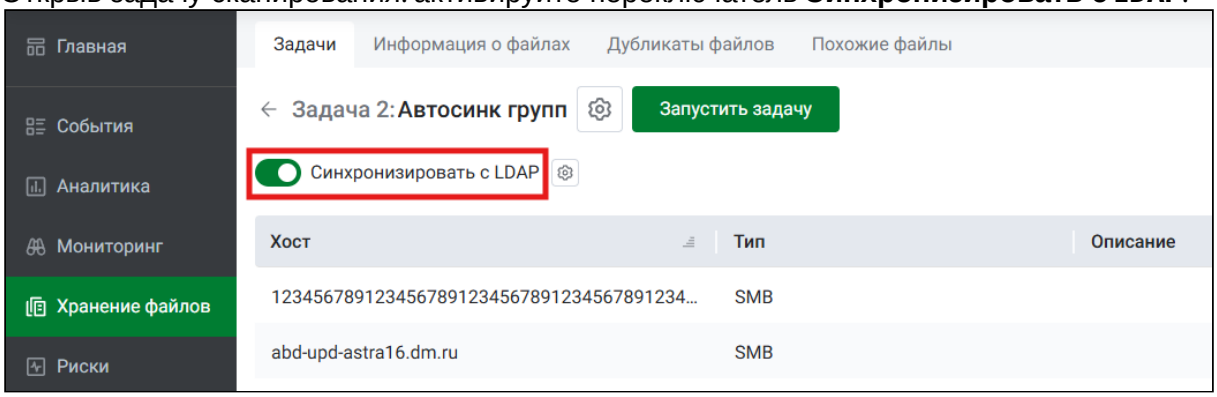
- Добавление из LDAP-каталога
- Импорт и экспорт с использованием JSON-файла

Автоматическая синхронизация с LDAP

В этом режиме пользователю достаточно указать группы из LDAP-каталога, чтобы добавить все хосты из этих групп, а также автоматически обновлять их состав в случае изменений в LDAP-каталоге. Для синхронизации с LDAP в кластере должна быть настроена и запущена интеграция с Active Directory (см. *InfoWatch Data Discovery. Руководство администратора. Интеграция с Active Directory*).

Чтобы включить автоматическую синхронизацию с LDAP:

1. Открыв задачу сканирования. активируйте переключатель **Синхронизировать с LDAP**.


Главная
События
Аналитика
Мониторинг
Хранение файлов
Риски
Задачи
Информация о файлах
Дубликаты файлов
Похожие файлы
Задача 2: Автосинк групп
Запустить задачу
Синхронизировать с LDAP
Хост
Тип
Описание
123456789123456789123456789

Active Directory (см. *InfoWatch Data Discovery. Руководство администратора. Интеграция с Active Directory*).

- При выключении синхронизации с LDAP в задаче останутся добавленные хосты, но их состав не будет обновлен в случае изменений в LDAP-каталоге. Также для задачи будут сброшены настройки выбора групп. При повторном включении будет необходимо заново выбрать группы из LDAP-каталога

Добавление вручную

Чтобы добавить хост вручную:

1. Открыв задачу сканирования, нажмите , чтобы **добавить один хост и указать его адрес вручную**.
2. В открывшемся окне настройте хост (см. Настройки хоста) и нажмите **Сохранить**.

Добавление из LDAP-каталога

1. Открыв задачу сканирования, нажмите , чтобы **добавить хосты из LDAP-каталога**.
2. В открывшемся окне отображаются данные из LDAP-каталогов, полученные Системой в результате синхронизаций с Active Directory. Если удалить интеграцию, то данные этой интеграции не будут отображаться в окне. Подробнее о добавлении данных в Систему см. *InfoWatch Data Discovery. Руководство администратора. Интеграции*.

   > **Примечание:**
   >
   > Если Система установлена совместно с другими продуктами Платформы, в окне также могут отображаться данные, полученные в результате синхронизаций с **Traffic Monitor**.

   Данные в окне представлены в виде дерева. Для элементов LDAP-каталога используются следующие обозначения:

   -  – контейнер;
   -  – группа;
   -  – компьютер.

   Чтобы раскрыть элемент дерева, нажмите слева от элемента. Чтобы искать элементы по имени, введите текст в поисковую строку.

3. С помощью переключателей под строкой поиска вы можете дополнительно **включить отображение групп**:
   - **из неактивных интеграций** - отображает группы, добавленные в результате интеграции, которая уже удалена;
   - **без компьютеров** - отображает группы, в которых не вложены компьютеры напрямую;
   - **без дочерних групп**.
4. Чтобы добавить компьютеры из каталога в задачу:
   a. Выберите в дереве компьютеры и/или элементы, которые включают в себя компьютеры.
   b. Нажмите **Добавить**.

При добавлении компьютера в качестве хоста в задачу, в параметре **Адрес** у хоста автоматически указывается полное доменное имя компьютера.

Импорт и экспорт с использованием JSON-файла

1. Нажмите , чтобы **добавить хосты из JSON-файла**.
2. В открывшемся окне выберите JSON-файл, который содержит описание хостов.
   Для корректного добавления хостов:
   - максимально допустимый размер JSON-файла - **1 МБ**;
   - содержимое файла должно соответствовать формату (подробнее см. Формат JSON-файла для импорта хостов).

Если при импорте **Data Discovery** обнаружит дублирование хостов, они не будут добавлены в систему, а пользователю будет выведено уведомление со списком хостов-дубликатов. Новые хосты будут отображены в списке.

В верхней части экрана со списком хостов доступно управление задачей. Нажмите , чтобы вернуться к списку задач.

> Для хостов с типом подключения Яндекс Диск экспорт и импорт не поддерживаются.

Вы можете **экспортировать хосты в JSON-файл**, чтобы импортировать их на другой сервер **Data Discovery**.

Для экспорта хостов нажмите . В результате браузер сохранит на ваш диск JSON-файл с описанием хостов.

Особенности экспорта:

- Фильтры не влияют на экспорт хостов.
- Если отметить хосты в списке, Система экспортирует только их.
  Если не отметить ни один, Система экспортирует все хосты в задаче.
- Экспортированные описания хостов **не содержат пароли** учетных записей, от имени которых **Data Discovery** подключался к ресурсам для сканирования.
  Для сканирования хостов, импортированных из JSON-файла, укажите пароли одним из способов:
  - Перед импортом добавьте в файл для описания каждого хоста поле `password` или `ssh_key` и укажите значение (подробнее см. Формат JSON-файла для импорта хостов).
  - После импорта в веб-интерфейсе **Data Discovery** добавьте пароли в описания хостов в секции **Авторизация для доступа к ресурсам**.

Настройки хоста

> **Примечание:**
>
> Поля, обязательные к заполнению, отмечены в интерфейсе значком .

Все параметры подключения к хранилищу Яндекс Диск выбираются в поле **Параметры авторизации**.

1. Вкладка **Общие**:
   a. Основные настройки хоста:

Параметр	Описание
Адрес	Доменное имя или IP-адрес. Система получает IP-адреса для всех доменных имен: • при запуске задачи; • регулярно во время выполнения задачи: через 30 секунд после того, как все адреса были получены. **Примечание**: Для подключения к SharePoint необходимо указать адрес, по которому доступно веб-приложение SharePoint. В некоторых случаях веб-приложение SharePoint может быть доступно, только если адрес указан в формате доменного имени.
Порт	Порт подключения. Если порт подключения не указан, то используется значение по умолчанию: • для SMB-подключения – 445; • для SFTP (SSH)-подключения – 22; • FTP-подключения - 21; • Nextcloud - 443 (HTTPS), 80 (HTTP); • Alfresco - 443 (HTTPS), 80 (HTTP). Для SharePoint-подключения поле обязательно для заполнения.
Тип подключения	Доступные значения: • SMB – для сканирования общих сетевых ресурсов; • SFTP (SSH) – для сканирования серверов и рабочих станций под управлением ОС Linux по протоколу SFTP, с аутентификацией по SSH; • SharePoint – для сканирования библиотек документов в SharePoint Server 2016 по протоколу HTTPS; • FTP – для сканирования файловых ресурсов по протоколу FTP; • Nextcloud – для сканирования хранилищ данных Nextcloud по протоколу WebDAV; • Alfresco – для сканирования серверов с системой Alfresco по протоколам HTTP/HTTPS; • Яндекс Диск – для сканирования хранилищ Яндекс Диск с помощью REST API Диска. Изменение значения данного параметра недоступно при редактировании хоста.
Путь к	Путь внутри SharePoint к сайту. Сайт должен содержать библиотеки документов, которые требуется просканировать. Параметр доступен, если для хоста выбрано SharePoint-подключение.

Параметр	Описание
сайту	**Пример**: Адрес библиотеки **documents** – `sp.example.com/my_path/site/documents`. В этом случае путь к сайту – `my_path/site`
Протокол	Параметр доступен для типов подключения FTP и WebDAV (Nextcloud). Доступные значения: • **Для FTP:** • FTP • FTPS (Explicit FTP over TLS) • **Для Nextcloud и Alfresco:** • HTTPS • HTTP
Путь подключения	Для подключения к Nextcloud по протоколу WebDAV в поле **Путь подключения** необходимо вставить часть ссылки из хранилища (подробнее см. Получение ссылки для подключения к Nextcloud). Параметр доступен, если для хоста выбран тип подключения Nextcloud.
Адрес прокси-сервера	Если подключение к серверу Alfresco происходит через прокси-сервер, укажите его адрес. Подключение к прокси-серверу поддерживается без авторизации. Параметр доступен, если для хоста выбран тип подключения Alfresco.
Порт для прокси-сервера	Порт для подключения к прокси-серверу. Укажите номер порта, если для подключения к Alfresco используется прокси-сервер. Параметр доступен, если для хоста выбран тип подключения Alfresco.
Параметры авторизации	Набор параметров подключения и авторизации к хранилищу Яндекс Диск. Добавление подключений доступно в разделе **Настройки** → **Настройки продуктов** → **Параметры авторизации**, подробнее см. *InfoWatch Data Discovery. Руководство администратора. Параметры авторизации*). Параметр доступен, если для хоста выбран тип подключения Яндекс Диск.
Описание	Краткое описание хоста

> **Примечание:**
>
> Подключение по протоколам **FTPS** (Explicit FTP over TLS) и **HTTPS** (SharePoint, Nextcloud и Alfresco) проходит без проверки подлинности TLS/SSL-сертификатов.

b. Секция **Параметры авторизации**:

Параметр	Описание
Наследовать из задачи	Если параметр включен ( ), то для подключения к хосту используются значения параметров авторизации, указанные в настройках задачи. Чтобы ввести значения, отличные от указанных в настройках задачи, отключите параметр. При добавлении хоста параметр включен по умолчанию.
По логину и паролю	
Группа SMB	Домен или группа. Параметр доступен, если для хоста выбрано SMB-подключение.
Логин	Имя пользователя **Важно**: Для подключения к SharePoint необходимо указать значение в формате: domain\user
Пароль	Пароль
По SSH-ключу (только для типа подключения SFTP (SSH))	
Логин	Имя пользователя, которому был передан публичный ключ
Ключевая фраза	Пароль для зашифрованного приватного ключа
Приватный ключ авторизации	Приватный ключ. Ключ может быть указан в поле вручную или загружен из файла с помощью кнопки . Пример использования SSH-ключей см. в статье Использование SSH-ключей для доступа к ресурсам

> **Примечание:**

При подключении к SMB-хосту используется анонимный доступ в следующих случаях:

- Если параметр **Наследовать из задачи** выключен и значения параметров авторизации для хоста не указаны;
- Если параметр **Наследовать из задачи** включен и в настройках задачи значения параметров авторизации не указаны.

2. Вкладка **Расширенные настройки**:

Параметр	Описание
Таймаут (мс)	Время ожидания при подключении к хосту. Для SharePoint-хостов не рекомендуется указывать значение меньше **60000** (60 секунд). Если при сканировании ресурса возникают ошибки вида "Connection timed out / Connection timeout", рекомендуется увеличить значение параметра. Подробнее о статусах ресурсов см. Статистика выполнения задач.
Попыток повторения	Максимальное количество попыток соединения
Уровень отладки	Параметр контролирует представление отладочной информации. **Для SMB-подключения:** Для изучения возникающих проблем рекомендуется использовать значение **2**. Подробнее см. документацию smbclient (параметр debuglevel). **Для SFTP (SSH)-подключения:** Для изучения возникающих проблем рекомендуется использовать следующие значения: • **8** – информация об аутентификации; • **64** – отладочная информация протокола SFTP; • **128** – информация об ошибках; • **512** – информация о работе с сокетами. Для получения отладочной информации одновременно по нескольким значениям, введите сумму значений. Например, чтобы одновременно получить отладочную информацию протокола SFTP (значение параметра: **64**) и информацию об ошибках (значение параметра: **128**), укажите в поле **Уровень отладки** значение **192** (64+128). Подробнее см. документацию libssh2. **Для SharePoint-подключения и Alfresco:** Для изучения возникающих проблем рекомендуется использовать следующие значения: • **1** – информация сокетах; • **2** – информация о запросах и ответах по HTTP; • **8** – информация об HTTP-аутентификации; • **16** – информация об HTTP-аутентификации (plaintext);

Параметр	Описание
	• **128** – информация о теле HTTP-ответов; • **256** – информация об SSL/TLS. Для получения отладочной информации одновременно по нескольким значениям, введите сумму значений. Например, чтобы одновременно получить информацию об аутентификации (значения параметра: **8** и **16**) и информацию об SSL/TLS (значение параметра: **256**), укажите в поле **Уровень отладки** значение **280** (8+16+256). **Для FTP-подключения**, **Nextcloud** и **Яндекс Диск**: Доступны следующие значения: **0** - отключить логирование сторонних библиотек; **1** - включить логирование сторонних библиотек.
Пауза до повтора (сек)	Временной интервал между попытками соединения
Размер чанка (байт)	Размер фрагмента передаваемых данных

Использование SSH-ключей для доступа к ресурсам

Data Discovery может подключаться к ресурсам сканирования, используя безопасный метод соединения с помощью SSH-ключей.

Важно!

Учитывайте, что особенности генерации ключей зависят от настроек используемой ОС.

Ниже приведен пример генерации ключей на ОС Red Hat Enterprise Linux 7.9, в которой по умолчанию выбрано использование алгоритма шифрования RSA.

Чтобы при сканировании использовать авторизацию с помощью SSH-ключей, выполните действия:

1. Сгенерируйте пару SSH-ключей с помощью команды:

```
ssh-keygen
```

Вы можете сгенерировать их непосредственно на сервере **Data Discovery** или на другом компьютере.

При необходимости вы можете использовать другую утилиту или указать предпочитаемый алгоритм шифрования.

2. Вы можете задать имя файлов с ключами или нажать **Enter** без ввода для создания файлов с именем по умолчанию.

3. Далее вам будет предложено задать ключевую фразу (passphrase), чтобы использовать зашифрованный ключ.
   Если вы не хотите использовать зашифрованный ключ, нажмите **Enter** без ввода ключевой фразы.
4. Если вы использовали имя файлов по умолчанию, будут созданы два файла с ключами:
   - `/<имя_пользователя>/.ssh/id_rsa` - приватный ключ авторизации;
   - `/<имя_пользователя>/.ssh/id_rsa.pub` - публичный ключ.
5. Далее необходимо передать публичный ключ (`id_rsa.pub`) **на все хосты**, на которых **Data Discovery** будет авторизоваться с помощью SSH-ключей.
   Для этого используйте утилиту `ssh-copy-id` с помощью команды вида:

```
ssh-copy-id <имя_пользователя_на_хосте>@<адрес_хоста>
```

Пример команды передачи публичного ключа:

```
ssh-copy-id ERipley@10.70.12.34
```

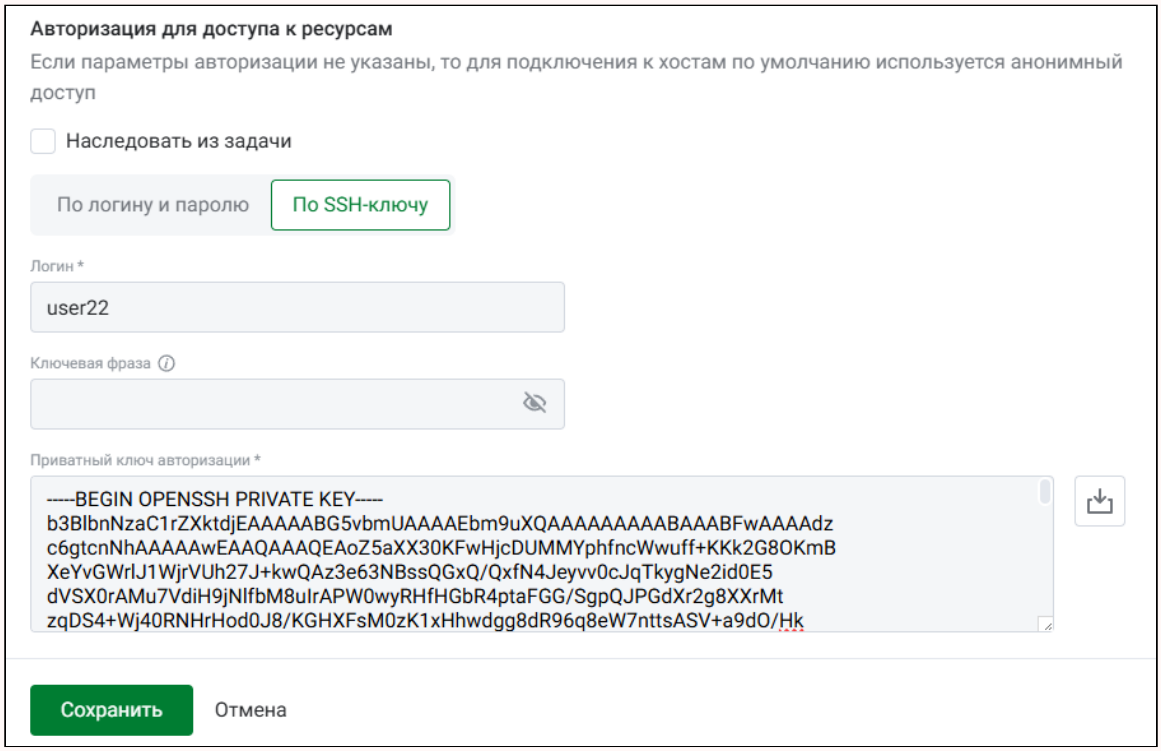
> **Примечание:**
>
> После передачи публичного ключа на хост, крайне рекомендуется проверить возможность авторизации с его помощью.

6. В окне создания задачи сканирования или добавлении хоста укажите в поле:
   a. **Логин** - имя пользователя на хосте, которому вы скопировали публичный ключ.
   b. **Ключевая фраза** - укажите ключевую фразу, если задали ее при генерации ключей.
   c. **Приватный ключ авторизации** - скопируйте содержимое файла `id_rsa` или загрузите его содержимое с помощью кнопки .

> **Важно!**
>
> Необходимо использовать приватный ключ в формате PEM.
>
> Если вы сгенерировали ключ в другом формате, конвертируйте его в формат PEM.
>
> **Пример корректного ключа OpenSSH:**

```
-----BEGIN OPENSSH PRIVATE KEY-----
b3BlbnNzaC1rZXktdjEAAAAABG5vbmUAAAAEbm9uXQAAAAAAAAABAAABFwAAAAdz
c6gtcnNhAAAAAwEAAQAAAQEAoZ5aXX30KFwHjcDUMMYphfncWwuff+KKk2G8OKmB
XeYvGWrlJ1WjrVUh27J+kwQAz3e63NBssQGxQ/QxfN4Jeyvv0cJqTkygNe2id0E5
dVSX0rAMu7VdiH9jNlfbM8uIrAPW0wyRHfHGbR4ptaFGG/SgpQJPGdXr2g8XXrMt
zqDS4+Wj40RNHrHod0J8/KGHXFsM0zK1xHhwdgg8dR96q8eW7nttsASV+a9dO/Hk
Zbu0SYZa01gJ53OFw1K5+raQ+yILM76QKOZPmZ/LvefCJtbjZdtH/bIc/dNN+M3I
h8HBTZ8zI5A99gdDPJIhldhyZSjYqdtJrVhscRmpgQwmTQAAA9CHFZ2chxWdnAAA
```

```
AAdzc2gtcnNhAAABAQChnlpdffQoXAeNwNQwximF+dxbC59/4oqTYbw4qYFd5i8Z
auQnVaOtVSHbsn6TBADPd7rc0GyxAbFD9DF83gl7K+/RwmpOTKA17aJ3QTl1VJfS
sAy7tV2If2M2V9szy4isA9bTDJEd8cZtHim1oUYb9KClAk8Z1evaDxdesy3OoNLj
5aPjRE0eseh3Qnz8oYdcWwzTMrXEeHB2CDx1H3qrx5bue22wBJX5r1078eRlu7RJ
hlrTWAnnc4XDUrn6tpD7IgszvpAo5k+Zn8u958Im1uNl20f9shz90034zciHwcFN
nzMjkD32B0M8kiGV2HJlKNip20mtWGxxGamBDCZNAAAAAwEAAQAAAQA+RUgmC5ie
UE8sw+Y94KwxDcNkZQcQfBXOkXK3FkBdEIqHLZM4wJtu9oU0JAL23Cz/wu8Bk2C9
lpxZRWYn6VkX1ybw2AqjnoyAGs5lSoE7/tr8d46AaHHgXMy6LIGt5DcFp/kGminz
sqs0cdG6RKhBrzCSQ2B84NqdywDDISB+IwQjYmL3329v9jk3wP8ZljcrxPyD9nh0
qjHb9sxXpQXId1DGMEcXDsawbO0GIFP40SzqpCtlnlOTf7+5WGqEweKNM6rQnpY0
vItLdecKm2nztI9LWunLQkB2JM5WKdcbpAF7JgI8dlyUnsiTFhdKPhLmBBY0sKBv
/Zj6LOk9nqXhAAAAgQDZodI48pRKX3NQH4m04hH2T2zF9XZCqQxqW/sOjpuWkWRw
ly8yOGa47D9vLO5HkCjQZQiqX/sItI9z6dXBo882Q0egcubS0xFPnpNIoEzcGeCB
9ive6OW1zcRQNTAziEQmfvQqLpuuhb3oP2de+NGhTAVzy13G4nyRTtV/FQ9hKgAA
AIEA76Oh699EgVcIKmijyBFvgWpklKTQ/UNFOUuOBMp3gujMQep/cXtchvWBBtBd
9W6tzlwLh45A4wWz0ncQx9E5Y8Jh/CVbc1cJwHpOUABUwbqq0UqWyzIwBiGsbYvt
oJIIZoByZvRuVGBqSAPFiYywlLihIs84x5FmL2vCgME3XykAAACBAKynF1Hr01tq
Anr5CJTgRfpJT+AeX6X15Pk9eEH2alAhqVohWBPvk3hzomAAFC5MCAKzpop1msY5
04nh4Dw4TpQK5JOUezFOSlXXZVQyQqvB0yDRY2W6kRYoCzb5u/CMthbFwk1KGqxe
vdPBUEkTc9ZCEkmPVwnAj17ve835/saFAAAAEHJzYS1rZXktMjAyNTAxMjgBAgME
BQYHCAkK
-----END OPENSSH PRIVATE KEY-----
```

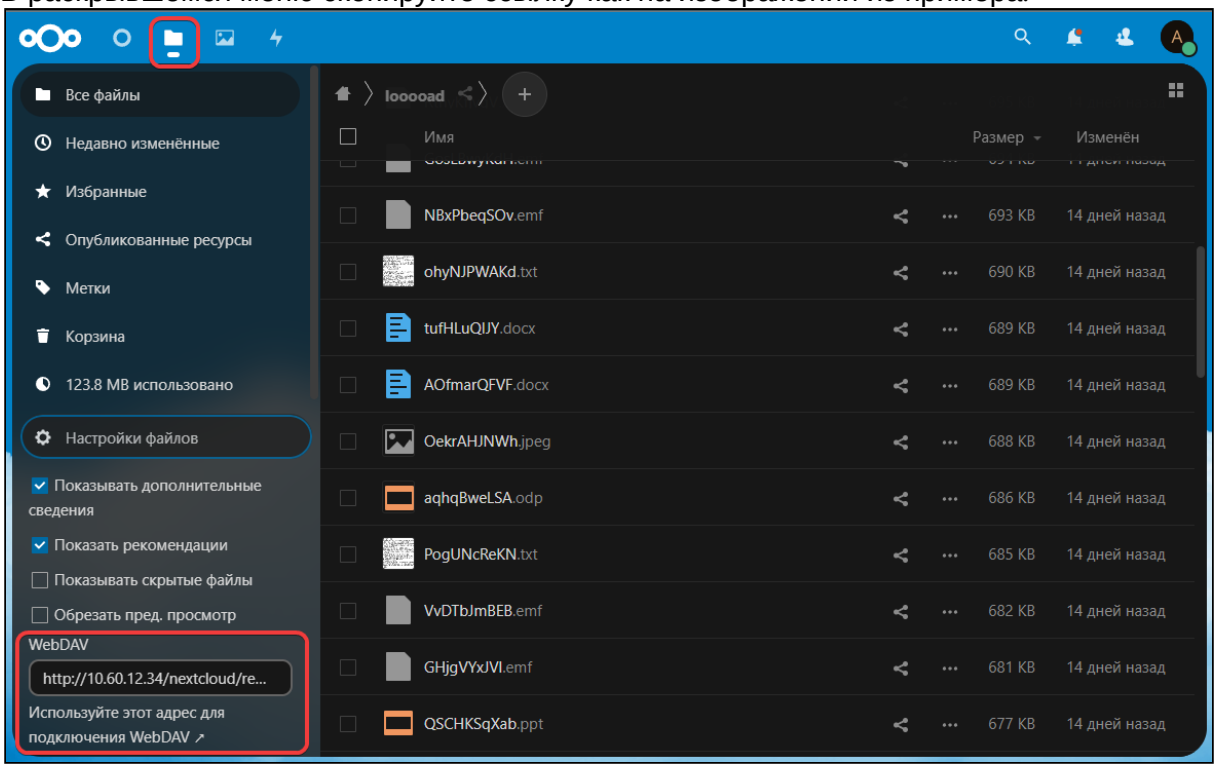



Получение ссылки для подключения к Nextcloud

Чтобы Data Discovery подключился к хранилищу Nextcloud по протоколу WebDAV:

1. Откройте веб-интерфейс хранилища Nextcloud.
2. Авторизуйтесь под учетной записью, от имени которой будет подключаться Data Discovery.

3. На верхней панели выберите вкладку **Файлы**.
4. В левом нижнем углу нажмите на **Настройки файлов**.
5. В раскрывшемся меню скопируйте ссылку как на изображении из примера:



6. Из скопированной ссылки в окне добавления хоста:
   a. В поле **Адрес** укажите ip-адрес хранилища.
      В нашем примере: `10.60.12.34`
   b. В поле **Протокол** укажите протокол передачи.
      В нашем примере: `http`
   c. В поле **Путь подключения** укажите оставшуюся часть ссылки.

Пример ссылки для подключения к Nextcloud:
`/nextcloud/remote.php/dav/files/admin/`

   Ссылка ведет в корневой каталог хранилища, доступного этому пользователю. Для этого хоста вы будете указывать пути сканирования относительно корневого каталога.

Формат JSON-файла для импорта хостов

Data Discovery позволяет добавить в задачу хосты из JSON-файла. Это позволит быстрее настроить систему для выполнения задач.

> **Примечание:**
>
> Максимально допустимый размер импортируемого JSON-файла - **1 МБ**.
>
> Для хостов с типом подключения Яндекс Диск экспорт и импорт не поддерживаются.

Для импорта в файле обязательно должен быть указан IP-адрес или доменное имя хоста. Если в файле не указаны часть параметров, для них **Data Discovery** использует значения по умолчанию.

В таблице ниже представлен состав JSON-файла:

Код (пример содержимого файла)	Формат / Возможные значения	Описание
`{`		{
`"ver": "1.0",`		Версия JSON-схемы
`"hosts": [`		Общая секция, включающая описания всех хостов
`{`		
`"host": "10.60.44.194",`	Строка, до 255 символов	IP-адрес или доменное имя хоста
`"type": "WEBDAV",`	• `SMB` • `SSH` • `SHAREPOINT` • `FTP` • `WEBDAV` • `ALFRESCO`	Тип подключения. Значение по умолчанию: SMB
`"protocol": "HTTP",`	• `HTTPS` • `HTTP` • `FTPS` • `FTP`	Параметр доступен для типов подключения FTP, Nextcloud и Alfresco Доступные значения: • **Для FTP:** • FTPS • FTP • **Для Nextcloud и Alfresco:** • HTTPS • HTTP
`"site": "nextcloud/remote.php/dav/files/admin/",`	Строка, до 255 символов	Параметр используется для типов подключения: • SHAREPOINT - путь к сайту; • Nexcloud - путь подключения.

Код (пример содержимого файла)	Формат / Возможные значения	Описание
		Важно! Для подключения к хранилищу Nextcloud по протоколу WebDAV необходимо вставить ссылку из хранилища, но **без IP-адреса** (подробнее см. "Получение ссылки для подключения к Nextcloud").
`"resources": [ "archive/cases", "demo", "lv-426" ],`		Пути сканирования
`"port": 2121,`	Целое число от 0 до 65535	Порт для подключения к хосту
`"proxy_host": 10.12.34.56,`	Строка, до 255 символов	Адрес прокси-сервера для подключения к Alfresco
`"proxy_port": 2121,`	Целое число от 0 до 65535	Порт для подключения к прокси-серверу
`"note": "Архив с данными",`	Строка, до 255 символов	Описание хоста
`"group_name": {},`	Строка, до 255 символов	Группа пользователя для типа подключения SMB
`"user_name": "admin",`	Строка, до 255 символов	Логин для соединения
`"auth_type": "SSH_KEY",`	• `PASSWORD` • `SSH_KEY`	Тип авторизации. Доступные значения: • PASSWORD - авторизация по логину и паролю; • SSH_KEY - SSH-ключ.
`"ssh_key": "-----BEGIN RSA PRIVATE KEY-----\nProc-Type: 4,ENCRYPTED\n DEK-Info: AES-128-CBC,9EDACF3CCDBC0F30ECEA3D4CCFCF8F34\n`	Строка	Приватный ключ авторизации. **Важно!**

Код (пример содержимого файла)	Формат / Возможные значения	Описание
`MIIEogIBAAKCAQEApAT92vPvEUIuHMImewQHIWyJbwLNE9r2lW5Zyz7sQpg+p08w\nPpg/5DBVCyOFKBjJVRb3CRVK7l8qdo5cxVFQqhjx2DvGcdtn+JKGh9+oSVS6r9rh\nGSF4e32z2Ozc5zQdv5FQ+WrUNaElNI7aAG6GC/R6Imich7Igy0Jdw/iuiP1KxG4F\njFVmvcZwKzMGVuyAKtM1YKP+sJN8TCk99dFM/hGKGh4GSx+9ELyVGMyJBhps7eHh\nEaz+2364RYYyCUzwCv5eXs8dSszDtlmJTjt86w4fNfAF919P1WokkwU/u5FrY7mU\n1nhG0wZFi+SzcZmCcGpTOr+Y464XYKIN8WRaIQIDAQABAoIBAGNZhq7mjQvGK4pN\n6+MduHf9d23fo3sAadZv5kIfVhqsf172r4Fq1WCSCiFaHaEBCorZAz/kPuQOjvkL\nGJSIwbSWsX9KDdUPknXVw1lwF0573S0CBtEjHginRpYCO3hscAyGqCa8F9h4SMIC\nOXiDefyp1IUdGptzQTv9cjBvZMEiu7A02Tp5RU7HTkavsLbP4sB9wngJuhGhsBMd\noMqrbSDKzIhgWuFFLeyGBEqUxfUDCKfWIP6y3VDb5ZBgRP59qxrFx0TAeaXHoiK8\ndqywSIrGhGyE87Otb6pOQXXwxT8MdT5F3vmVRMUc2eegz694zm2mMaJDiSOLDowA\ndjETpsECgYEAz73B3N3ibYWgU+tm5mgbdEO5PcoEeWKKOeyrueniw+kP+FMdkqEa\niGHyWbJHkNHgvt7fg79zpAhalYkJr4t+AsZ4Fvaa6GzJ8mvbuUccYn3A1tc+eNlF\nLKAzcLntGJkC6Vbvt/0JwmkTNNDprDpBidtF6qp7XbtQRjOklzoR81kCgYEAyh8h\nKFgR63NiUQOFKluAa1mIcVFzlYIrlPpDvbxxHPPpPgZmsLLPSsqVu67/ZBuIYJvY\n2RG8h98sX+U5ji2lE47fw04tkxOrihLuRX4drtjMIVqndrxrpHAjKFa6ttrwCd1O\nDlqvN13zWv4EqSbEVOoOgZ7uHEynI9LmTFf7rAkCgYBKqyX3QTmVrCYIg+AkiN71\nVzJJ50bcanLMhyjiVhLP4AGBfD5w6b8jrjDx1IasowNpF6q3LCVgnDAS94Knuje8\nv6GCf3kGebPMksATRGALdq/NiOhHf7DBbYOPzQRPhw/o1vqq4ySS4DmrDjwvT3En\nH81tn4FgG2erE/zg7lptYQKBgCnXfea7LKKAVWcOAoA9Rv1yIHPM98+IjfseP+3E\nV+o1eCB19MY5jf4eqBadaJNRu868FI0DRfMeuTH04xBS1KyS1DG7OmYLGWTQ/HB4\nZi6vBLsW6DoFUYdOC4krQyXddvmjzafdEA63LUy170MRGm21smEVBJfpKmSZIUHG\nqPv5AoGAeDLKkqUg4rlUM9khzaa1i04g8g8mmh2zTfQMiImr/G1K8UCXGLsIIg/M\nipvF+Bsuxb5X6YxeqGZCvfaCQXgmfjzOzrISW`		• Ключ должен быть указан одной строкой, для этого после копирования замените в нем все переносы строк на "`\n`". • Если в файле указан приватный ключ авторизации, будет выбран тип авторизации `SSH_KEY`, даже если в `auth_type` было указано значение `PASSWORD`. Параметр не экспортируется и указывается только вручную для импорта.

Код (пример содержимого файла)	Формат / Возможные значения	Описание
`gh/X262zzLTNYzdGpyuIsB3b2M+\nXc0cBaFm8hNc6SKc/hXLzoae33ITtUhTcAhZGTjxujpuRPzkVUg=\n-----END RSA PRIVATE KEY-----",`		
`"ssh_key_pass": "city17",`	Строка, до 255 символов	Ключевая фраза. Ключевую фразу необходимо указать, если вы используете зашифрованный приватный ключ авторизации. **Важно!** Параметр не экспортируется и указывается только вручную для импорта.
`"password": "12pass34",`	Строка, до 255 символов	Пароль для соединения. **Важно!** Параметр не экспортируется и указывается только вручную для импорта.
`"inherit_from_task": false,`	• `true` • `false`	Наследование из задачи логина и пароля для соединения. Если в импортируемом файле указан логин (`user_name`), значение `inherit_from_task` будет проигнорировано, а параметры авторизации не будут наследоваться из задачи даже при значении `true` . Значение по умолчанию: `true`
`"auto_detect_resources": false,`	• `true` • `false`	Использование режима *Сканировать все ресурсы*
`"timeout_msec": 5000,`	Целое неотрицательное число	Таймаут соединения или операций с хостом, в миллисекундах.

Код (пример содержимого файла)	Формат / Возможные значения	Описание
		Значение по умолчанию: 5000
`"retry_count": 5,`	Целое неотрицательное число	Максимальное количество попыток соединения. Значение по умолчанию: 5
`"pause_sec": 3,`	Целое неотрицательное число	Временной интервал между попытками соединения, в секундах. Значение по умолчанию: 3
`"io_chunk_size": 32768,`	Целое неотрицательное число	Размер фрагмента передаваемых данных, в байтах. Значение по умолчанию: 32768
`"client_debug_level": 0`	Целое неотрицательное число	Уровень отладки. Значение по умолчанию: 0
`}`		
`]`		
`}`		

Пример JSON-файла для импорта нескольких хостов:

```
{
  "ver": "1.0",
  "hosts": [
    {
      "host": "10.60.3.195",
      "type": "SSH",
      "protocol": "HTTPS",
      "resources": [
        "share"
      ],
      "user_name": "root",
      "auth_type": "PASSWORD",
      "password": "SSHpass54321",
      "inherit_from_task": false,
      "auto_detect_resources": false,
      "timeout_msec": 5000,
```

```
    "retry_count": 5,
    "pause_sec": 3,
    "io_chunk_size": 32768,
    "client_debug_level": 0
  },
  {
    "host": "10.60.12.34",
    "type": "ALFRESCO",
    "resources": [
      "/folder/new"
    ],
    "port": 443,
    "protocol": "HTTPS",
    "proxy_host": "10.21.33.56",
    "proxy_port": 445,
    "user_name": "Alpharius",
    "auth_type": "PASSWORD",
    "password": "alfpass42",
    "inherit_from_task": false,
    "auto_detect_resources": false,
    "timeout_msec": 5000,
    "retry_count": 5,
    "pause_sec": 3,
    "io_chunk_size": 32768,
    "client_debug_level": 0
   },
   {
    "host": "10.60.44.194",
    "type": "WEBDAV",
    "protocol": "HTTP",
    "resources": [
      "Photos"
    ],
    "site": "nextcloud/remote.php/dav/files/admin/",
    "user_name": "NCadmin",
    "auth_type": "PASSWORD",
    "password": "NCpass42",
    "inherit_from_task": false,
    "auto_detect_resources": false,
    "timeout_msec": 5000,
    "retry_count": 5,
    "pause_sec": 3,
    "io_chunk_size": 32768,
    "client_debug_level": 0
  },
  {
    "host": "10.60.44.233",
    "type": "SMB",
    "protocol": "HTTPS",
    "resources": [
      "demo",
      "shares",
      "smoke"
    ],
    "inherit_from_task": true,
    "auto_detect_resources": false,
    "timeout_msec": 5000,
    "retry_count": 5,
```

```
      "pause_sec": 3,
      "io_chunk_size": 32768,
      "client_debug_level": 0
    },
    {
      "host": "10.70.229.170",
      "type": "FTP",
      "protocol": "HTTPS",
      "resources": [
        "/dd"
      ],
      "port": 2121,
      "user_name": "admin",
      "auth_type": "PASSWORD",
      "password": "SDpass",
      "inherit_from_task": false,
      "auto_detect_resources": false,
      "timeout_msec": 5000,
      "retry_count": 5,
      "pause_sec": 3,
      "io_chunk_size": 32768,
      "client_debug_level": 0
    }
  ]
}
```


Добавление путей сканирования

Для сканирования хостов пользователю доступно:

- Добавление путей сканирования вручную
- Выбор из доступных путей сканирования
- Сканирование всех ресурсов

Добавление путей сканирования вручную

Чтобы добавить путь сканирования:

1. Перейдите в задачу сканирования.
2. Нажмите на хост, чтобы перейти к списку путей сканирования.
3. В поле слева над таблицей укажите путь сканирования:
   - Для хоста с **SMB**-подключением укажите имя общего сетевого ресурса, например:

     ```
     share
     ```

     Примечание:

     Для хостов с SMB-подключением не поддерживается возможность указать поддиректорию сетевого ресурса (например: share/folder) в качестве пути сканирования.

Чтобы сканировать поддиректорию SMB-ресурса, настройте фильтры для директорий на вкладке **Фильтры папок** в настройках задачи.

- Для хоста с **Alfresco**, **Nextcloud**, **SFTP (SSH)**- или **FTP**-подключением укажите полный путь к директории, например: `/home/username/Documents`

  Примечание:

  Для типа подключения Nextcloud указывайте пути сканирования относительно корневого каталога, указанного в поле **Путь подключения** при описании хоста.

- Для хоста с **SharePoint**-подключением вы можете указать значение в следующих форматах:
  - имя библиотеки документов, например: `Shared Documents`
  - путь к директории внутри библиотеки, например: `Shared Documents/folder/example`
- Для хоста с типом подключения **Яндекс Диск** – полный путь к папке на Яндекс Диске.
  Для пользовательского диска укажите путь в формате `/users/<e-mail_пользователя>/<имя_папки>`.
  Пример пути: `/users/ivanov@organization.ru/secret_folder`
  Для общего диска организации укажите путь в формате `/shared/<имя_общего_диска>/<имя_папки>`.
  Пример пути: `/shared/head_office_disk/secret_folder`

4. Нажмите .

В результате добавленный путь сканирования отображается в списке.

В верхней части экрана со списком путей сканирования доступно управление задачей и хостом.

Нажмите , чтобы вернуться к списку хостов.

Выбор из доступных путей сканирования

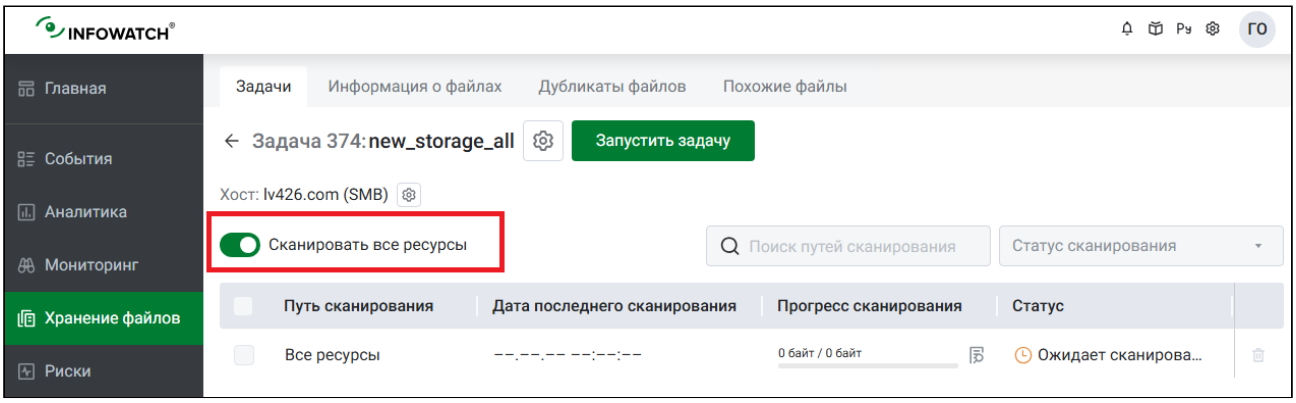
По запросу пользователя **Data Discovery** определяет все доступные пути сканирования на хосте и представляет возможность выбрать их из дерева папок.

Для выбора путей сканирования:

1. Перейдите в задачу сканирования.
2. Нажмите на хост, чтобы перейти к списку путей сканирования.
3. Нажмите .
4. В открывшемся окне выберите пути сканирования из дерева папок.
   Поиск осуществляется только по данным, которые отображаются в окне. Перед поиском рекомендуется раскрыть элементы дерева и убедиться, что элементы загрузились и отображаются в окне.
5. Нажмите **Добавить**.

Сканирование всех ресурсов

Для хостов с SMB-подключением доступен режим **Сканировать все ресурсы**. Для включения/выключения режима воспользуйтесь переключателем над таблицей слева. Режим включен по умолчанию.



С включенным режимом **Data Discovery** будет сканировать все доступные ресурсы этого хоста. При появлении новых доступных путей сканирования **Data Discovery** будет автоматически обрабатывать их без дополнительных действий пользователя. При включенном режиме в таблице будет отображаться только одна запись "*Все ресурсы*", а кнопки добавления, удаления и автоматического определения путей сканирования станут недоступны.

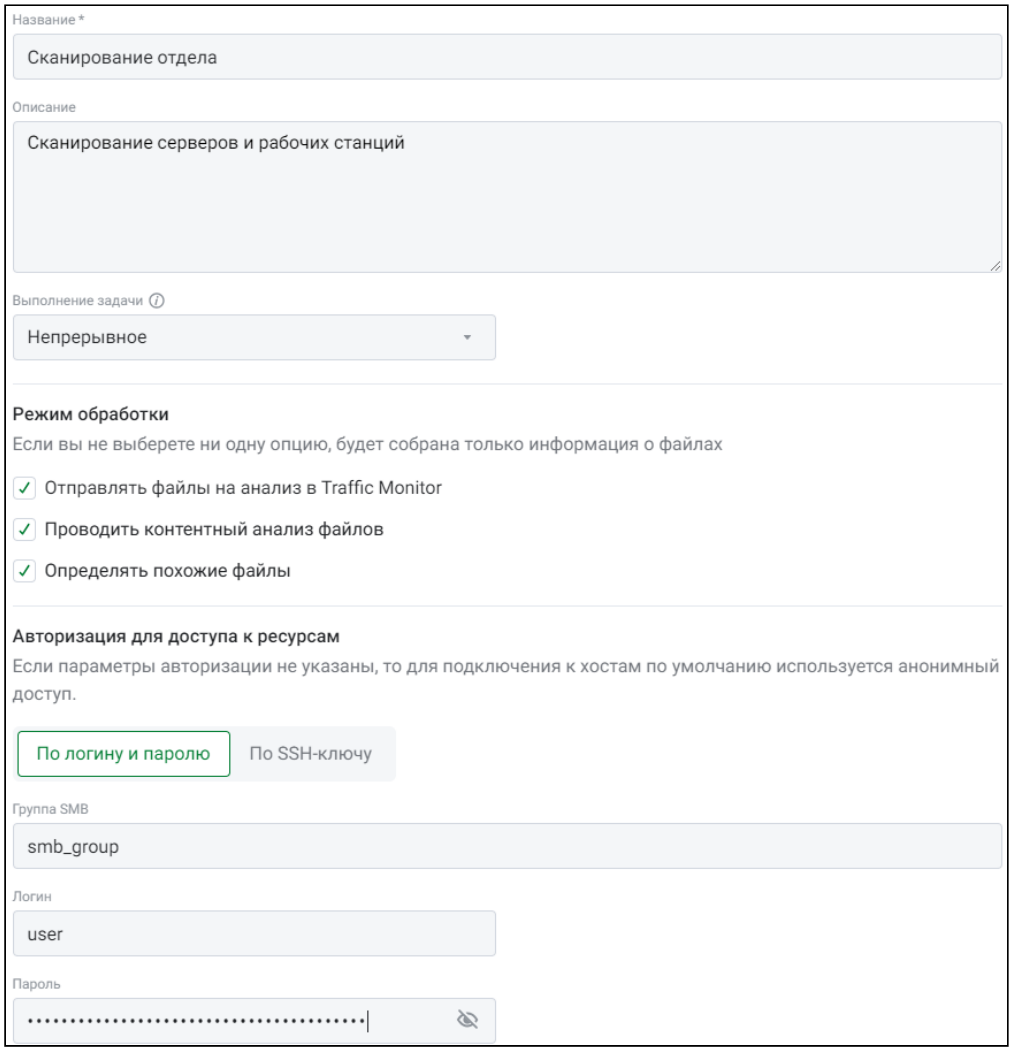
При включении режима Сканировать все ресурсы все ранее добавленные для хоста пути сканирования будут удалены. При выполнении задачи ранее просканированные файлы будут повторно отправлены на анализ.

После отключения режима Сканировать все ресурсы необходимо заново добавить для хоста пути сканирования. После добавления путей ранее просканированные файлы будут повторно отправлены на анализ.

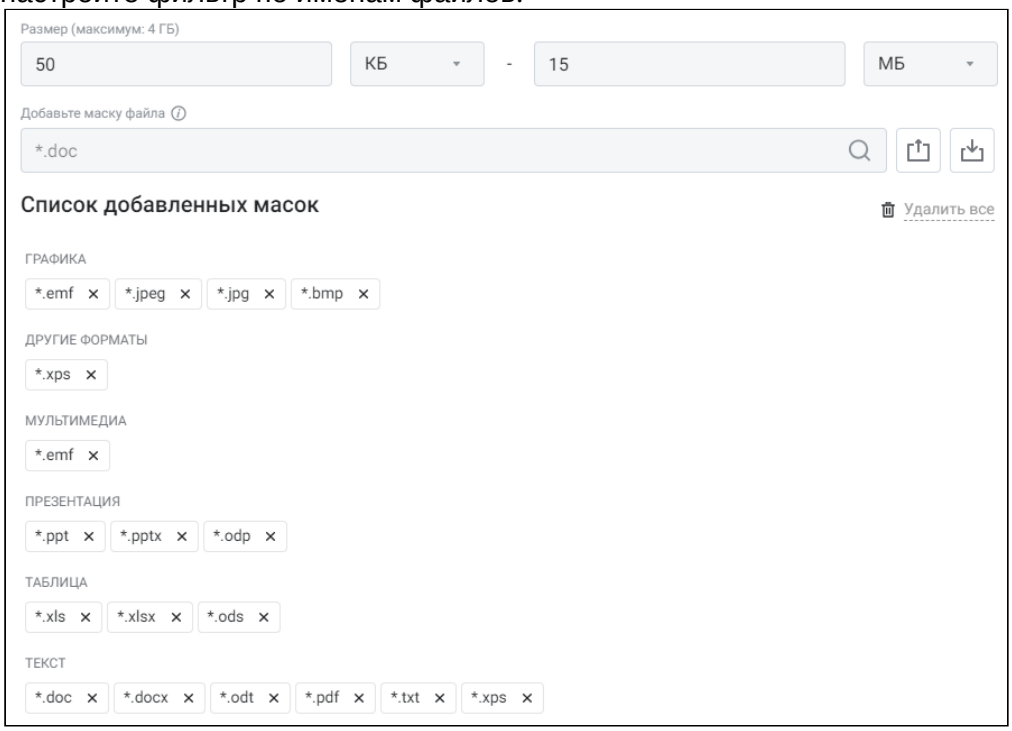
Примеры настроек задачи, хоста и добавления путей сканирования

Пример настроек задачи

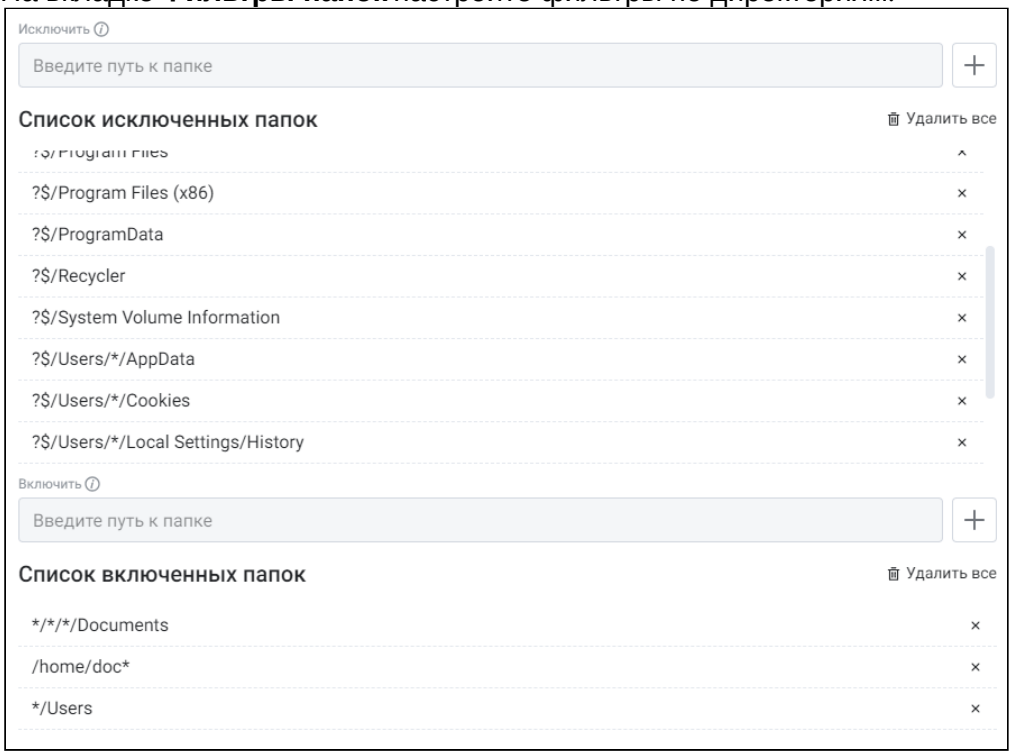
1. На вкладке **Общие** укажите значения параметров задачи:


Название *
Сканирование отдела
Описание
Сканирование серверов и рабочих станций
Выполнение задачи
Непрерывное
Режим обработки
Если вы не выберете ни одну опцию, будет собрана только информация о файлах
Отправлять файлы на анализ в Traffic Monitor
Проводить контентный анализ файлов
Определять похожие файлы
Авторизация для доступа к ресурсам
Если параметры авторизации не указаны, то для подключения к хостам по умолчанию используется анонимный доступ.
По логину и паролю
По SSH-ключу
Группа SMB
smb_group
Логин
user
Пароль

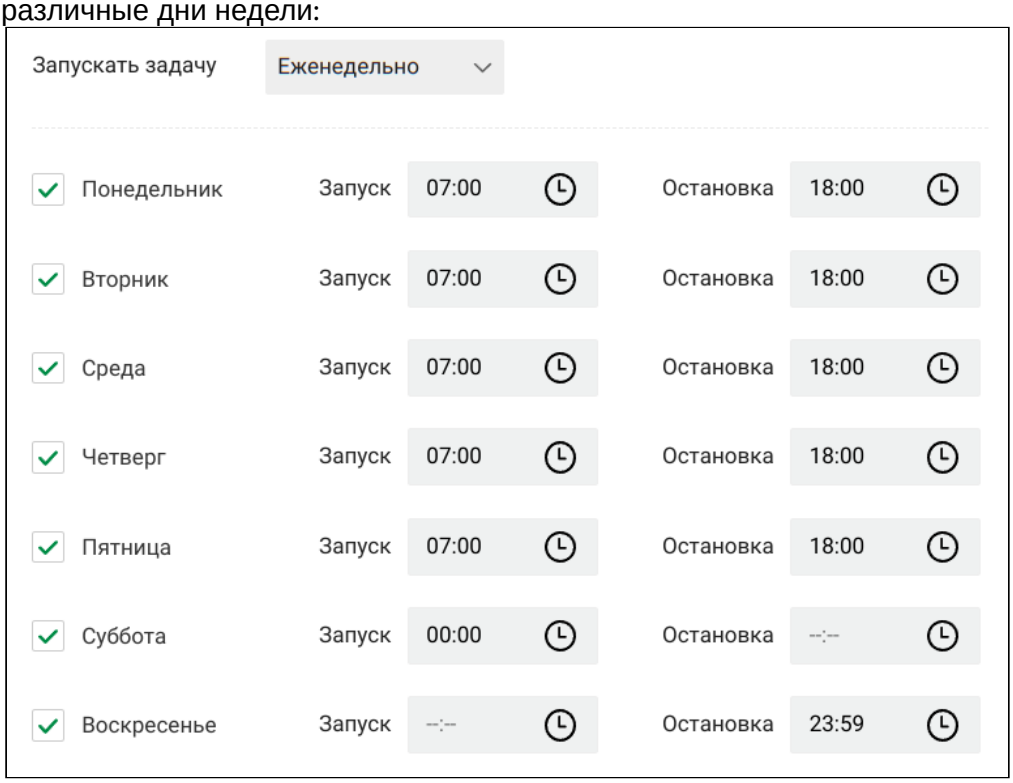
2. На вкладке **Фильтры файлов** установите ограничение на размер для каждого файла и настройте фильтр по именам файлов:



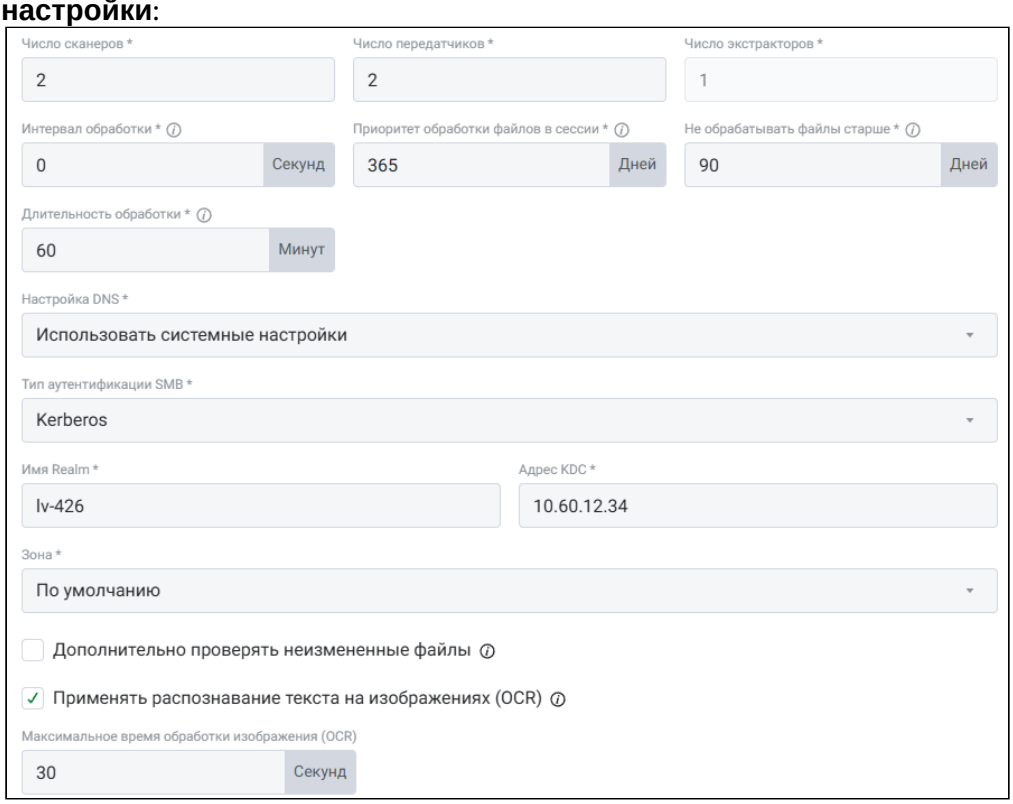
3. На вкладке **Фильтры папок** настройте фильтры по директориям:



4. На вкладке **Расписание** настройте автоматический запуск и остановку задачи в различные дни недели:



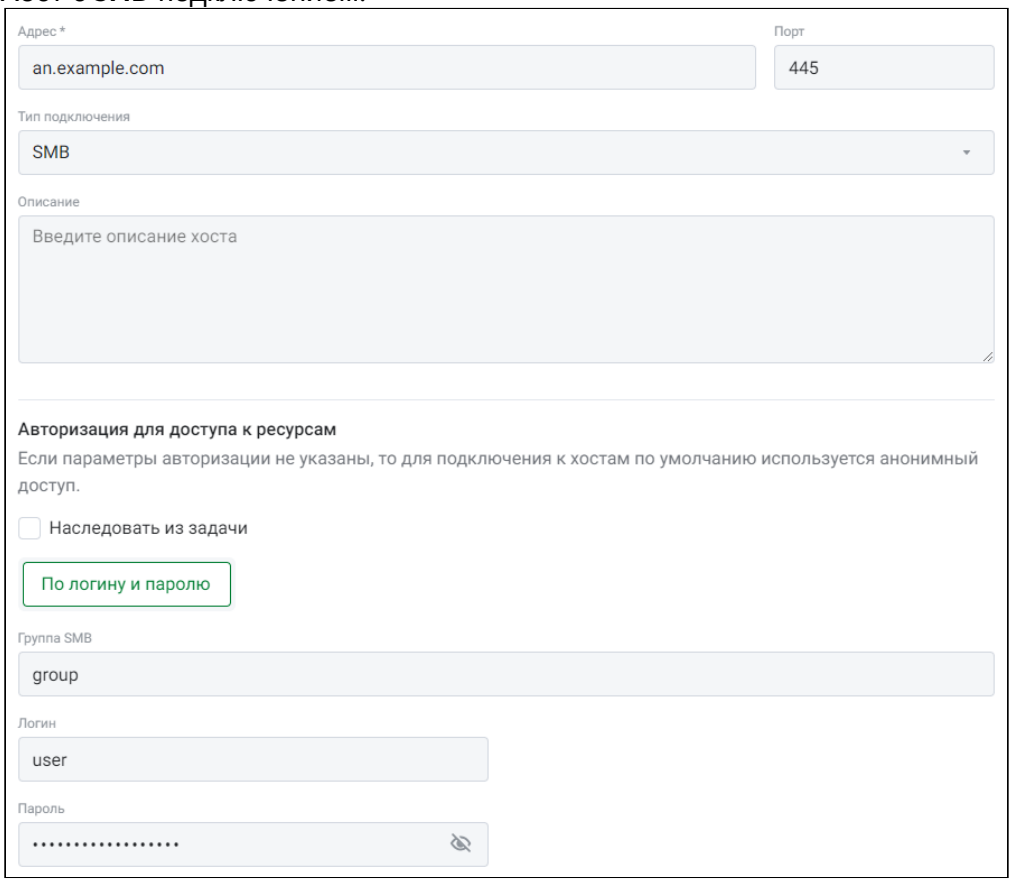
5. При необходимости измените значения параметров на вкладке **Расширенные настройки**:



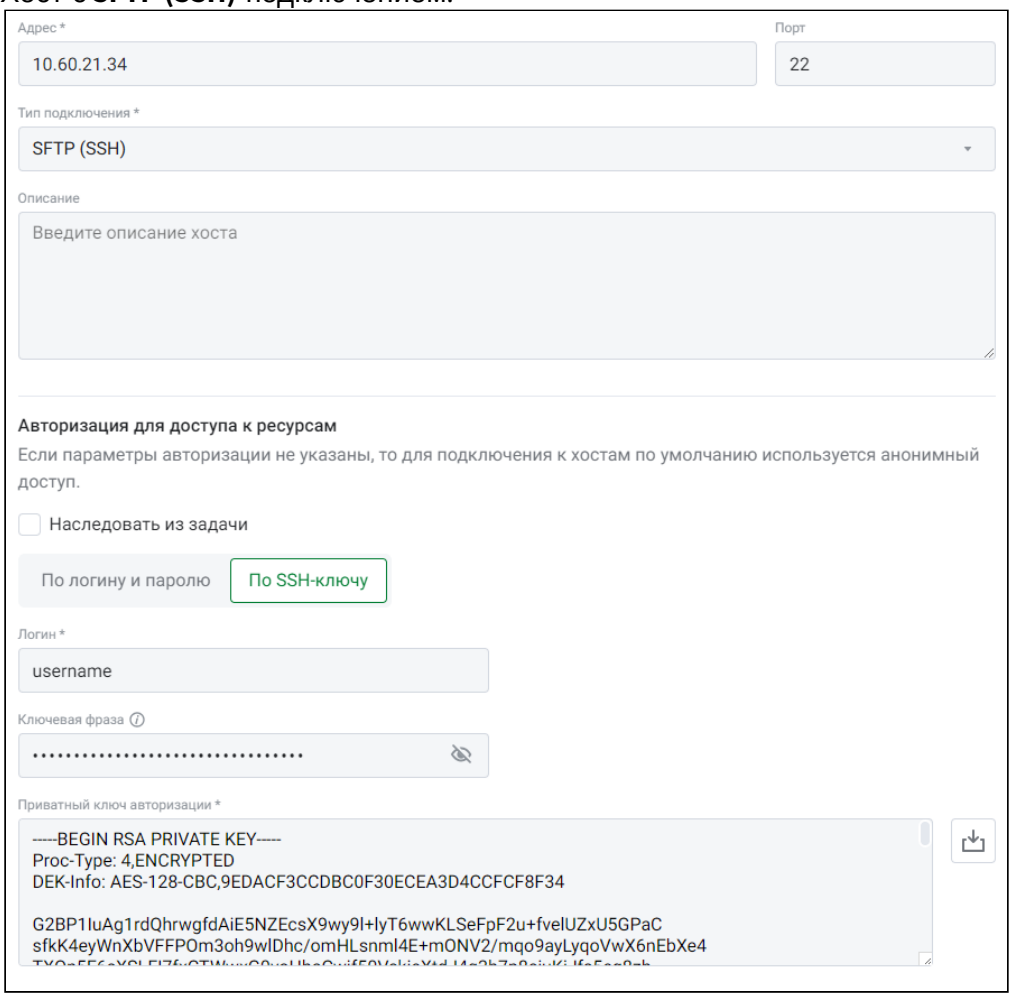
На этой вкладке вы можете включить распознавание текста на изображениях (OCR).

Пример настроек хостов

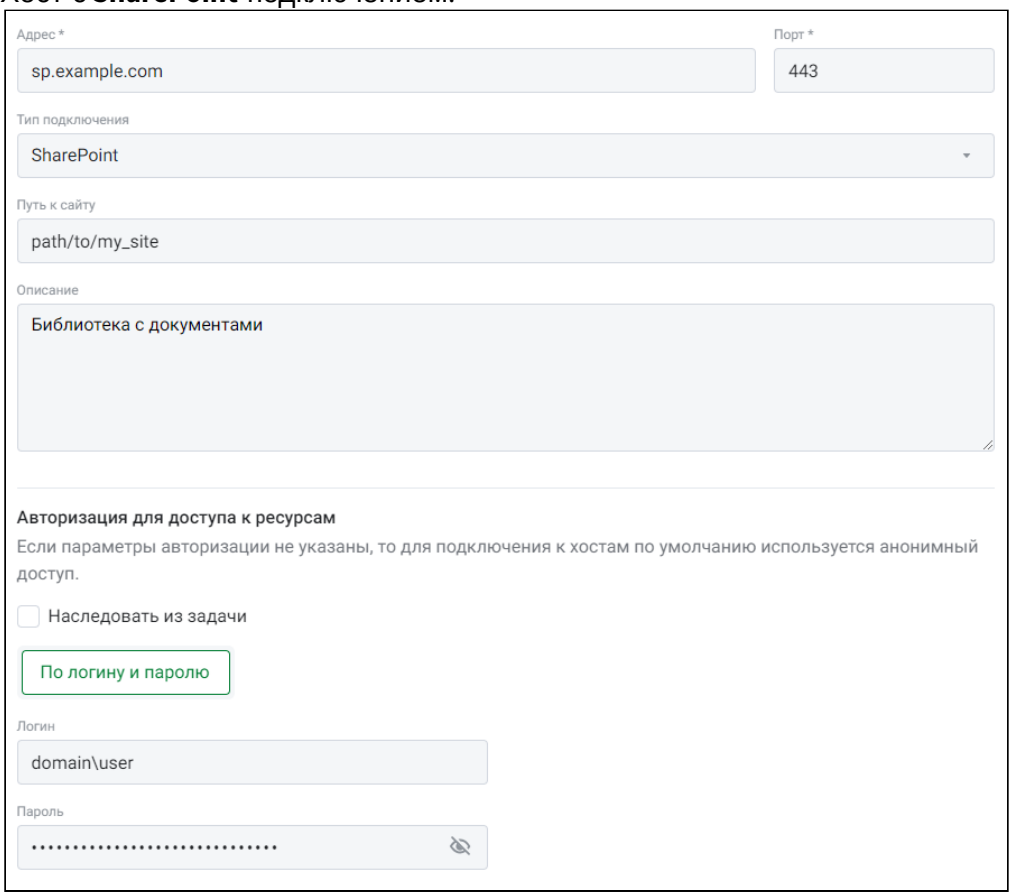
1. На вкладке **Общие** укажите значения параметров для подключения к хосту:
   - Хост с **SMB**-подключением:



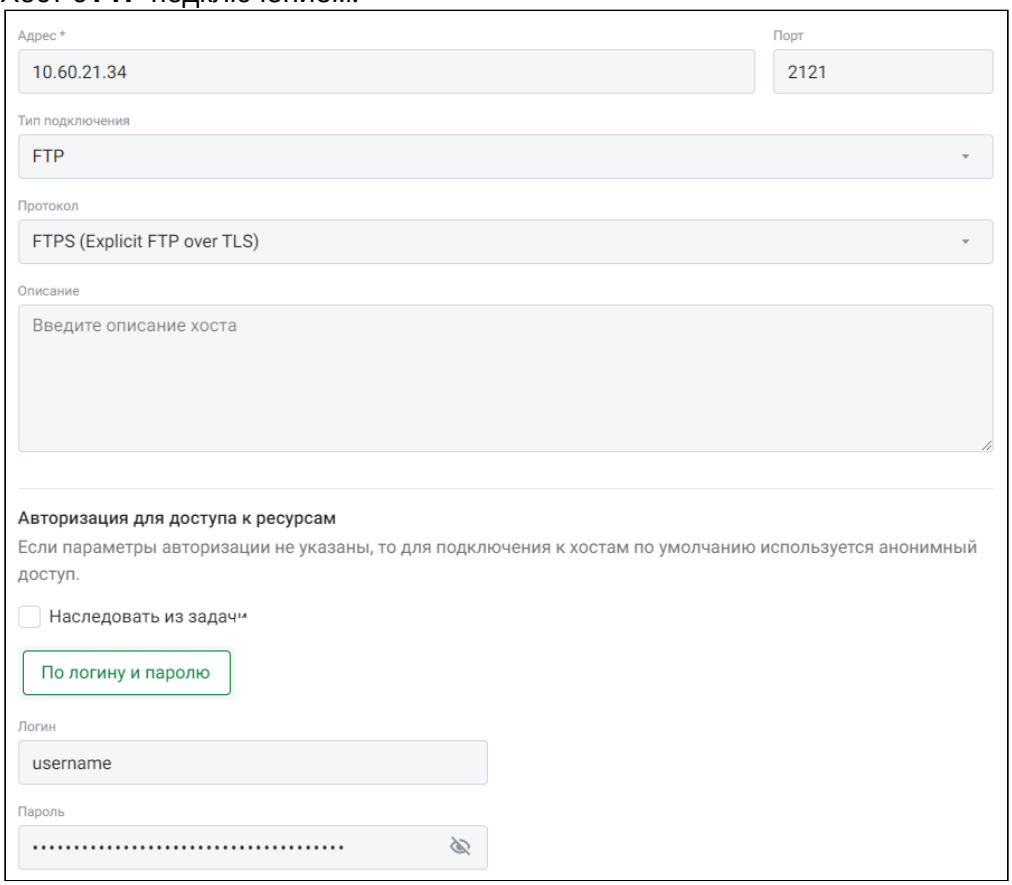
- Хост с **SFTP (SSH)**-подключением:



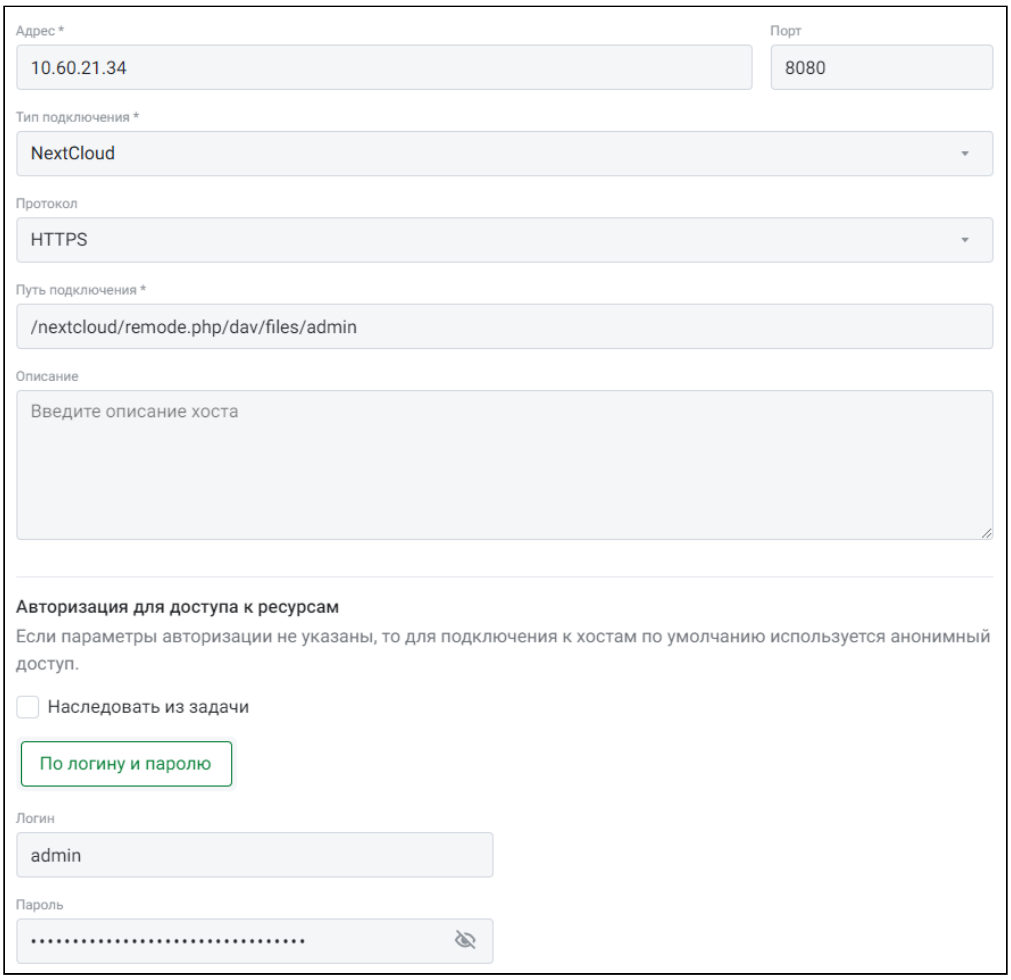
- Хост с **SharePoint**-подключением:



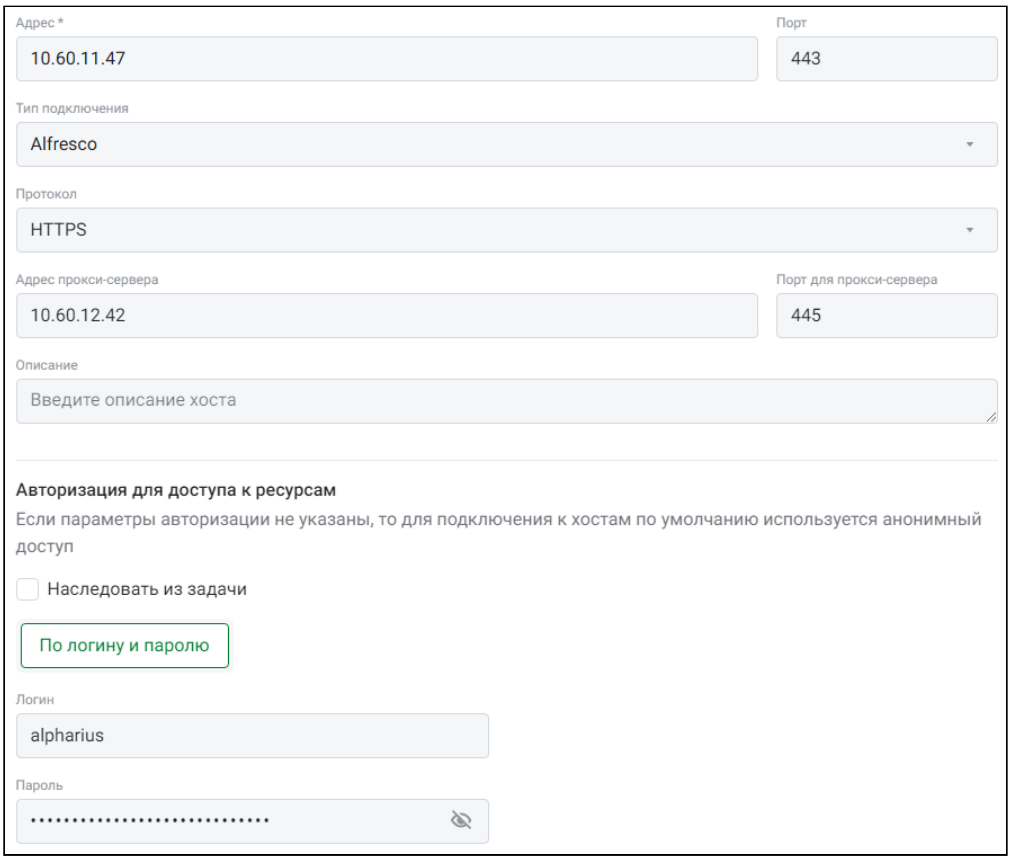
- Хост с **FTP**-подключением:



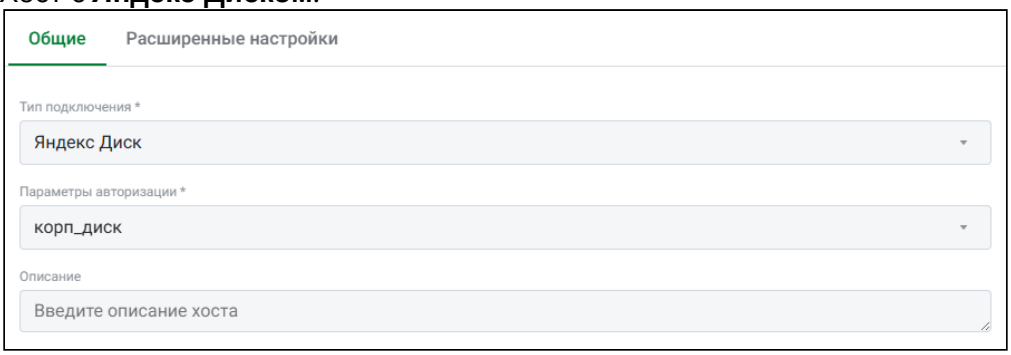
- Хост с **Nextcloud**:



- Хост с **Alfresco**:



- Хост с **Яндекс Диском**:

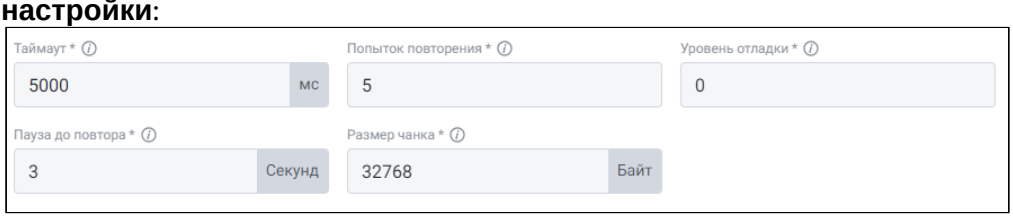


Примечание:

Чтобы ускорить сканирование хранилища Nextcloud, для подключения используйте пароль приложения (подробнее см. примечание в пункте Сторонние клиенты WebDAV из официальной документации Nextcloud).

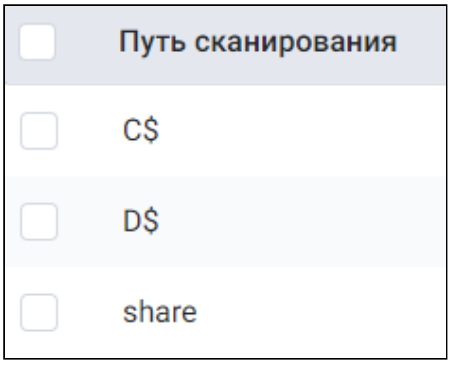
Подключение по протоколам **FTPS** (Explicit FTP over TLS) и **HTTPS** (SharePoint, Nextcloud и Alfresco) проходит без проверки подлинности TLS/SSL-сертификатов.

2. При необходимости измените значения параметров на вкладке **Расширенные настройки**:



Примеры добавления путей сканирования

Для хоста с **SMB**-подключением в качестве путей сканирования укажите имена сетевых ресурсов:

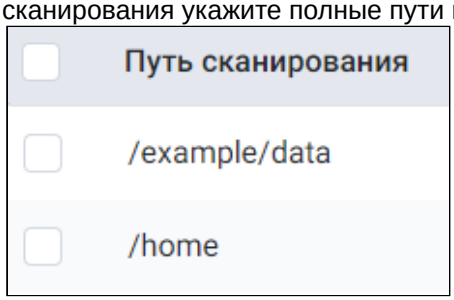


Примечание:

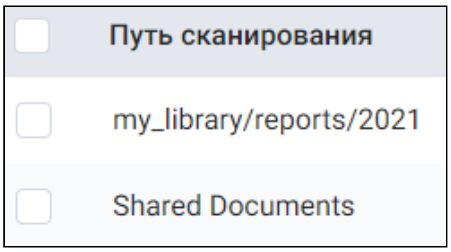
Не поддерживается возможность указать поддиректорию сетевого ресурса (например: share/folder) в качестве пути сканирования.

Чтобы сканировать поддиректорию SMB-ресурса, настройте фильтры для директорий на вкладке **Фильтры папок** в настройках задачи.

Для хоста с **Яндекс Диск**, **Alfresco, Nextcloud, SFTP (SSH)**-, **FTP**-подключением в качестве путей сканирования укажите полные пути к директориям:



Для хоста с **SharePoint**-подключением укажите имена библиотек и/или пути к директориям внутри библиотек:



10.1.3 Редактирование задачи и хоста

Редактирование настроек задачи и хостов доступно только для задач со статусами **Новая**, **Ожидает запуск по расписанию**, **Остановлена**, **Завершена**. Вы можете нажать , чтобы просмотреть текущие настройки задачи/хоста, даже если редактирование недоступно.

Чтобы отредактировать задачу:

1. Нажмите на раздел **Задачи**, чтобы перейти к списку задач сканирования.
2. Справа в строке задачи, которую требуется отредактировать, нажмите .
3. Измените настройки задачи (о настройках задачи см. "Создание, запуск и настройка задачи").
4. Нажмите **Сохранить**.

Чтобы отредактировать хост в задаче:

1. Нажмите на задачу, чтобы перейти к списку хостов.
2. Справа в строке хоста, который требуется отредактировать, нажмите .
3. Измените настройки хоста (о настройках хоста см. "Добавление и настройка хостов").
4. Нажмите **Сохранить**.

10.1.4 Копирование задачи

Чтобы скопировать задачу:

1. Откройте вкладку **Задачи**, чтобы перейти к списку задач сканирования.
2. Нажмите  справа в строке задачи, копия которой вам нужна.

В списке задач отобразится новая задача. Название новой задачи состоит из названия оригинальной задачи и слова "Копия". Новая задача имеет те же настройки, хосты и пути сканирования.

Важно!

Файлы задачи и метаданные файлов не будут скопированы.

В рамках новой задачи Система полностью повторно просканирует ресурсы.

10.1.5 Удаление задачи, хостов и путей сканирования

Удаление задачи, хостов и путей сканирования доступно только для задач со статусами **Новая**, **Ожидает запуск по расписанию**, **Остановлена**, **Завершена**.

Важно!

При удалении задачи сканирования **будут удалены**:

- копии просканированных файлов, которые не были отправлены в Traffic Monitor;

- метаданные просканированных файлов из обоих режимов, и **Актуально**, и **История**;
- извлеченные текстовые контенты.

Примечание:

Если вы удалите задачу, файлы задачи будут удалены из хранилищ зоны и не будут отправлены в Traffic Monitor.

Рекомендуется отправить файлы в Traffic Monitor перед удалением задачи:

1. В настройках задачи:
   - Включите параметр **Отправлять файлы на анализ в Traffic Monitor**.
   - Укажите следующие значения:
     - **0** для параметра **Число сканеров**.
     - **N** для параметра **Число передатчиков**, где **N** больше или равно количеству серверов Data Discovery в зоне задачи.
2. Запустите задачу. Система отправляет файлы в Traffic Monitor и удаляет их из хранилищ, но не скачивает новые файлы в хранилища зоны.
3. Дождитесь, когда в хранилищах зоны не останется файлов задачи. Чтобы получить количество файлов задачи на сервере Data Discovery:
   a. Подключитесь к ssh-консоли сервера.
   b. Выполните команду `ls PATH/NUMBER -1 | wc -l`, где:
      - `PATH` – путь к хранилищу сервера, по умолчанию: `/opt/disctps`;
      - `NUMBER` – номер задачи сканирования.

      Пример: `ls /opt/disctps/7 -1 | wc -l`
4. Остановите задачу.

Чтобы удалить задачу:

1. Откройте вкладку **Задачи**, чтобы перейти к списку задач сканирования.
2. Нажмите  справа в строке задачи, которую необходимо удалить.
3. Подтвердите удаление, нажав **Удалить**.

Чтобы удалить хосты:

1. Нажмите на задачу, чтобы перейти к списку хостов.
2. Если требуется удалить один хост, нажмите  в столбце **Действия** у хоста.

   Если требуется удалить несколько хостов, отметьте слева от каждого из них  и нажмите  в меню сверху.

   Если требуется удалить все хосты на текущей странице, установите  напротив заголовка столбца **Хост** и нажмите  в меню сверху.
3. Подтвердите удаление, нажав **Удалить**.

> **Важно!**
>
> Если вы добавите хост повторно, Система заново просканирует его ресурсы.

Чтобы удалить пути сканирования:

1. Нажмите на хост, чтобы перейти к списку путей сканирования.
2. Если требуется удалить один путь, нажмите  справа от записи в таблице.

   Если требуется удалить несколько путей, отметьте слева от каждого из них  и нажмите  в меню сверху.

   Если требуется удалить все пути на текущей странице, установите  напротив заголовка столбца **Путь сканирования** и нажмите  в меню сверху.
3. Подтвердите удаление, нажав **Удалить**.

> **Важно!**
>
> Если вы добавите путь повторно, Система заново его просканирует.

10.1.6 Статистика выполнения задач

В Data Discovery представлена статистика – информация о выполнении каждой задачи сканирования. Также доступна статистика по хостам и путям сканирования.

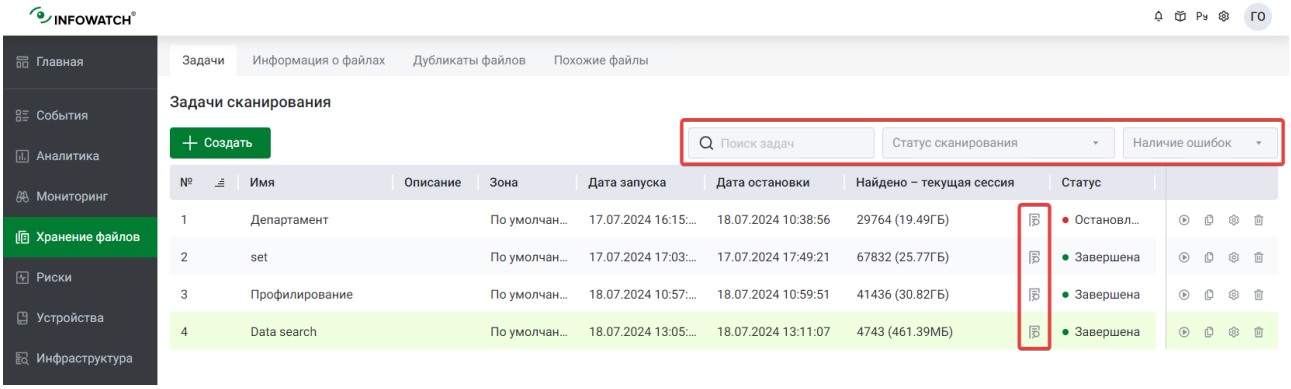
Статистика по задачам сканирования

Статистика доступна в разделе **Хранение файлов** на вкладке **Задачи**. Навигация по страницам списка задач осуществляется с помощью панели, расположенной внизу экрана. Статистика по задачам включает в себя следующие данные:

Параметр	Описание
№	Номер задачи. Номер является уникальным
Имя	Название задачи
Описание	Краткое описание задачи сканирования
Зона	Зона задачи
Дата запуска	Дата и время запуска задачи, значение отображается в формате dd.MM.yyyy HH:mm:ss

Параметр	Описание
Дата остановки	Дата и время остановки задачи, значение отображается в формате dd.MM.yyyy HH:mm:ss. Для задачи в статусе *Выполняется* данные отсутствуют
Найдено – текущая сессия	Число и общий объем файлов, найденных на всех ресурсах задачи с учетом фильтров. Система получает число файлов на ресурсе при каждом подключении (в каждой сессии). Число найденных файлов может уменьшиться, если с ресурса удалили файлы или если изменились настройки фильтров в задаче
Статус	Текущий статус задачи: • Новая • Ожидает запуск по расписанию • Запускается • Выполняется • Останавливается • Остановлена • Определение путей сканирования • Завершена Если возникла ошибка при работе с ресурсом, рядом со статусом отображается ⚠. Чтобы просмотреть информацию об ошибке, наведите курсор на ⚠.

Работа со списком задач



Сортировка задач возможна по следующим столбцам: **№**, **Имя**, **Дата запуска**, **Дата остановки**, **Статус**. Чтобы изменить порядок сортировки строк в списке задач, нажмите на заголовок того столбца, по которому нужно выполнить сортировку.

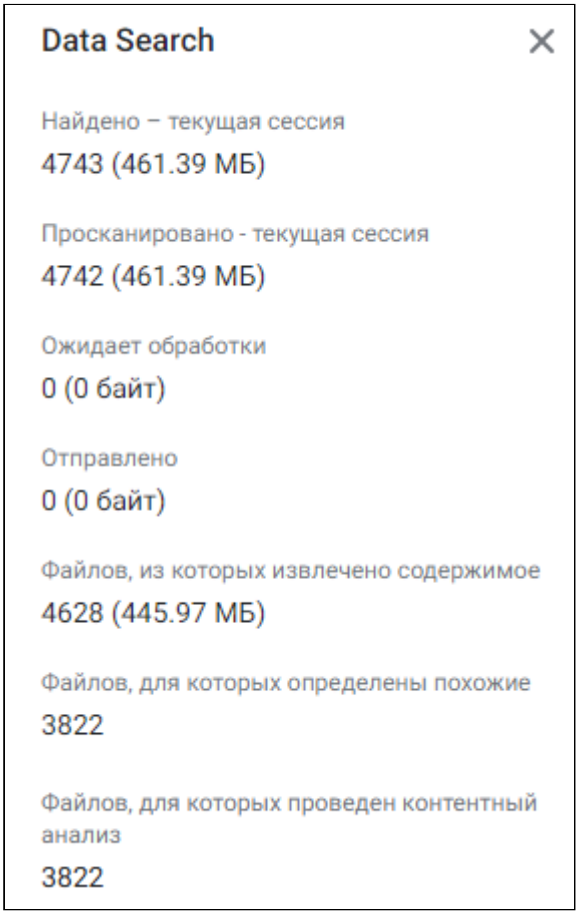
Чтобы искать задачи по имени, введите первые символы имени в поле **Поиск задач**, расположенное над таблицей справа. В списке отображаются задачи, имя которых начинается с символов, введенных в поле **Поиск задач**. Поиск осуществляется без учета регистра.

Чтобы отфильтровать список задач, выберите значения в выпадающих списках над таблицей справа. Вы можете фильтровать список задач:

- **по Статусу сканирования** – вы можете выбрать один из статусов;
- **по Наличию ошибок** – вы можете выбрать одно из значений: **С ошибками** или **Без ошибок**.

Чтобы сбросить значение фильтра, в поле с выбранным значением нажмите .

Чтобы открыть подробную Статистику выполнения задачи, в строке задачи в столбце **Найдено - текущая сессия** нажмите . Справа откроется окно с именем задачи и статистикой ее выполнения:



Если в момент просмотра задача запущена, значения полей будут меняться по ходу выполнения. Статистика содержит следующие данные:

- **Найдено - текущая сессия** – количество и общий объем файлов, найденных в текущей сессии с учетом фильтра. Data Discovery получает число файлов на ресурсе при каждом подключении (в каждой сессии). Число найденных файлов может уменьшиться, если с ресурса удалили файлы или если изменились настройки фильтров в задаче;
- **Просканировано - текущая сессия** – количество и общий объем файлов, просканированные файлы в текущей сессии. В большинстве случаев число просканированных файлов возрастает с каждой сессией. В конце прохода по ресурсу число файлов сбрасывается, в следующем проходе сканер проверяет файлы и обрабатывает новые;
- **Ожидает обработки** – количество и общий объем файлов, ожидающих обработки в текущий момент;
- **Отправлено** – количество и общий объем файлов, отправленных в Traffic Monitor за все время выполнения задачи;

- **Файлов, из которых извлечено содержимое** – количество и общий объем файлов, из которых извлечены текстовые данные для поиска документов с одинаковыми терминами (похожих), а также для обнаружения в них объектов защиты (контентный анализ);
- **Файлов, для которых определены похожие** – количество файлов, которые **прошли обработку** на определение похожих. Значение в поле отобразится только после обработки всех данных модулем Data Analysis Service, при этом задача сканирования к этому моменту уже будет в статусе **Завершена**. До завершения обработки в поле будет отображено значение 0;
- **Файлов, для которых проведен контентный анализ** - количество файлов, которые проверены на наличие в них объектов защиты. Значение в поле будет изменяться по мере обработки данных Модулем контентного анализа.

Чтобы скопировать в буфер обмена номер или имя задачи наведите курсор мыши на ячейку таблицы и нажмите .

Чтобы переместить столбец, нажмите на заголовок столбца и, не отпуская, перемещайте заголовок вдоль строки заголовков.

Чтобы изменить ширину столбца, нажмите на правую границу заголовка столбца и, не отпуская, перемещайте границу столбца.

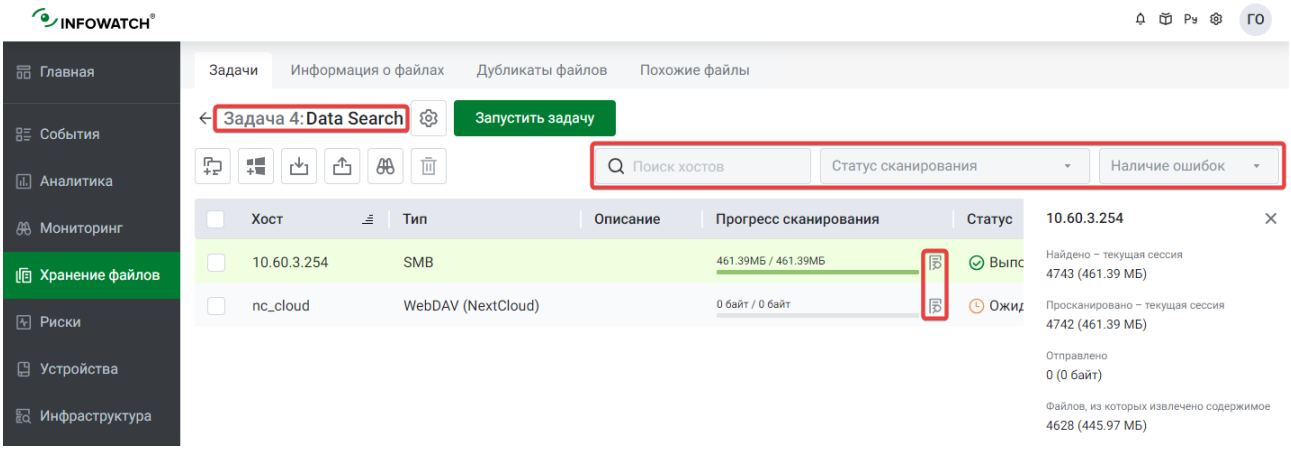
Статистика по хостам

Нажмите на задачу, чтобы перейти к списку хостов. Слева вверху будет указан номер и имя задачи сканирования. Навигация по страницам списка хостов осуществляется с помощью панели навигации, расположенной внизу экрана. Статистика содержит следующую информацию:

Параметр	Описание
Хост	Адрес хоста
Тип	Тип подключения хоста
Описание	Описание хоста
Прогресс сканирования	Общий объем просканированных файлов / общий объем найденных файлов
Статус	**Примечание:** Статус хоста зависит от статусов ресурсов хоста. Текущий статус хоста: •  *Идет сканирование* – хотя бы один ресурс хоста имеет статус *Подключение*, *Сбор информации* или *Обработка файлов* •  *Выполнено* – все ресурсы имеют статус *Выполнено*, за исключением тех, на которых возникла ошибка

Параметр	Описание
	• *Ожидает сканирования* – статус отображается, если не выполняются условия для статусов, указанных выше Если возникла ошибка при работе с ресурсом, рядом со статусом отображается ⚠. Чтобы просмотреть информацию об ошибке, наведите курсор на ⚠.

Работа со списком хостов



Сортировка хостов возможна по столбцам: **Хост**, **Тип**. Чтобы изменить порядок сортировки строк в списке хостов, щелкните левой клавишей мыши по заголовку того столбца, по которому нужно выполнить сортировку.

Чтобы искать хосты по адресу, введите первые символы IP-адреса/доменного имени в поле **Поиск хостов**, расположенное над таблицей справа. В списке отображаются хосты, адрес которых начинается с символов, введенных в поле **Поиск хоста**. Поиск осуществляется без учета регистра.

> **Примечание:**
>
> Поиск хостов по адресу осуществляется по значениям, указанным при создании/редактировании хоста. Например, если при создании/редактировании хоста адрес был указан в виде доменного имени, то хост не будет найден по IP-адресу.

Чтобы отфильтровать список хостов, выберите значения в выпадающих списках над таблицей справа. Вы можете фильтровать список хостов:

- **по Статусу сканирования** – вы можете выбрать один из статусов;
- **по Наличию ошибок** – вы можете выбрать одно из значений: **С ошибками** или **Без ошибок**.

Чтобы сбросить значение фильтра, в поле с выбранным значением нажмите ✕.

Чтобы открыть подробную Статистику сканирования хоста, в строке хоста в столбце **Прогресс сканирования** нажмите . Справа откроется окно с адресом хоста и статистикой сканирования:

Если в момент просмотра задача запущена, значения полей будут меняться по ходу выполнения. Статистика содержит следующие данные:

- **Найдено - текущая сессия** – количество и общий объем файлов, найденных на этом хосте в текущей сессии с учетом фильтра. Data Discovery получает число файлов на ресурсе при каждом подключении (в каждой сессии). Число найденных файлов может уменьшиться, если с ресурса удалили файлы или если изменились настройки фильтров в задаче;
- **Просканировано - текущая сессия** – количество и общий объем файлов, просканированных на этом хосте в текущей сессии. В большинстве случаев число просканированных файлов возрастает с каждой сессией. В конце прохода по ресурсу число файлов сбрасывается, в следующем проходе сканер проверяет файлы и обрабатывает новые;
- **Отправлено** – количество и общий объем файлов, отправленных с этого хоста в Traffic Monitor за все время выполнения задачи;
- **Файлов, из которых извлечено содержимое** – количество и общий объем файлов, из которых извлечены текстовые данные для поиска документов с одинаковыми терминами (похожих).

Чтобы переместить столбец, нажмите на заголовок столбца и, не отпуская, перемещайте заголовок вдоль строки заголовков.

Чтобы изменить ширину столбца, нажмите на правую границу заголовка столбца и, не отпуская, перемещайте границу столбца.

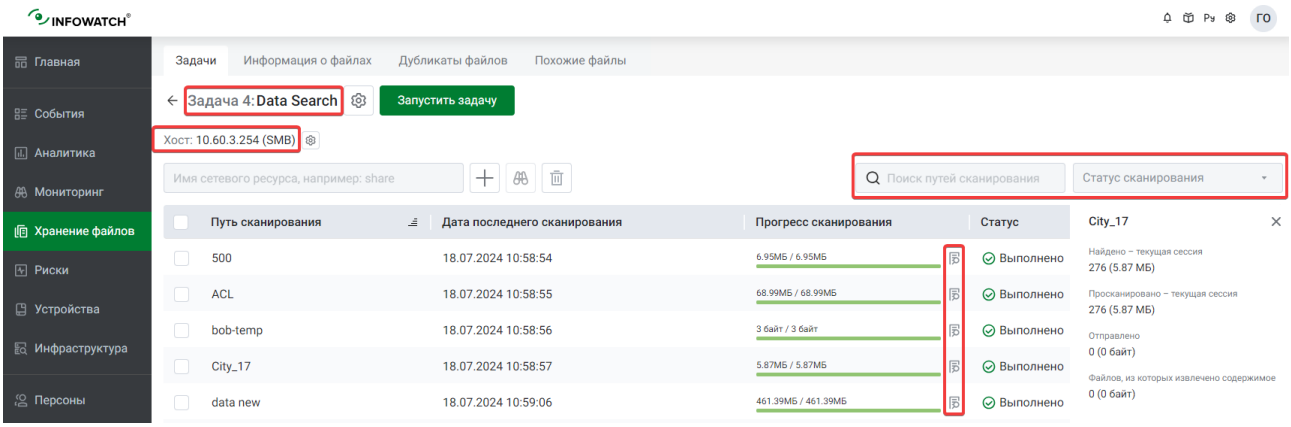
Статистика по путям сканирования

Нажмите на хост, чтобы перейти к списку путей сканирования. Слева вверху будет указан номер и имя задачи сканирования, а также адрес хоста. Навигация по страницам списка путей сканирования осуществляется с помощью панели навигации, расположенной внизу экрана. Для пути сканирования статистика содержит следующую информацию:

Параметр	Описание
Путь сканирования	Путь сканирования хоста (ресурс)
Дата последнего сканирования	Дата и время последнего подключения к ресурсу, значение отображается в формате dd.MM.yyyy HH:mm:ss
Прогресс сканирования	Общий объем просканированных файлов / общий объем найденных файлов
Статус	Текущий статус пути сканирования: •  *Ожидает сканирования* – подключение к ресурсу еще не осуществлялось в рамках задачи •  *Подключение* – осуществляется подключение к ресурсу

Параметр	Описание
	• *Сбор информации* – Система получает список файлов и информацию о файлах • *Обработка файлов* – Система сохраняет информацию о файлах и загружает копии файлов в хранилище • *Остановлено* – завершилась сессия сканирования ресурса **или** была остановлена задача сканирования • *Выполнено* – завершился проход по ресурсу, были обработаны все файлы, находящиеся на ресурсе • *Ошибка* – произошла ошибка при работе с ресурсом (чтобы просмотреть информацию об ошибке, наведите курсор на статус ресурса) **Примеры ошибок** В некоторых случаях сообщения об ошибке содержат следующий текст: • Operation not permitted • Permission denied • Authentication failed Ошибки такого вида могут означать, что используются неверные параметры авторизации или у учетной записи недостаточно прав доступа для сканирования ресурса.

Работа со списком путей сканирования



Сортировка путей сканирования возможна по столбцу **Путь сканирования**. Чтобы изменить порядок сортировки строк в списке путей сканирования, щелкните левой клавишей мыши по заголовку столбца **Путь сканирования**.

Чтобы искать пути сканирования, введите первые символы пути в поле **Поиск путей сканирования**, расположенное над таблицей справа. В списке отображаются пути, которые

начинаются с символов, введенных в поле **Поиск путей сканирования**. Поиск осуществляется без учета регистра.

Чтобы отфильтровать список путей сканирования, выберите значение в выпадающем списке над таблицей справа. Вы можете фильтровать список путей по **Статусу сканирования** – вы можете выбрать один из статусов.

Чтобы сбросить значение фильтра, в поле с выбранным значением нажмите .

Чтобы открыть подробную Статистику сканирования ресурса, в строке ресурса в столбце **Прогресс сканирования** нажмите . Справа откроется окно с именем пути и статистикой сканирования:

Если в момент просмотра задача запущена, значения полей будут меняться по ходу выполнения. Статистика содержит следующие данные:

- **Найдено - текущая сессия** – количество и общий объем файлов, найденных по этому пути сканирования в текущей сессии с учетом фильтра. Data Discovery получает число файлов на ресурсе при каждом подключении (в каждой сессии). Число найденных файлов может уменьшиться, если с ресурса удалили файлы или если изменились настройки фильтров в задаче;
- **Просканировано - текущая сессия** – количество и общий объем файлов, просканированных по этому пути в текущей сессии. В большинстве случаев число просканированных файлов возрастает с каждой сессией. В конце прохода по ресурсу число файлов сбрасывается, в следующем проходе сканер проверяет файлы и обрабатывает новые;
- **Отправлено** – количество и общий объем файлов, отправленных с этого пути сканирования в Traffic Monitor за все время выполнения задачи;
- **Файлов, из которых извлечено содержимое** – количество и общий объем файлов, из которых извлечены текстовые данные для поиска документов с одинаковыми терминами (похожих).

Чтобы переместить столбец, нажмите на заголовок столбца и, не отпуская, перемещайте заголовок вдоль строки заголовков.

Чтобы изменить ширину столбца, нажмите на правую границу заголовка столбца и, не отпуская, перемещайте границу столбца.

10.2 Вкладка "Информация о файлах"

На вкладке отображается детальная информация о файлах, которые Data Discovery просканировал при выполнении задач, а также результаты контентного анализа.

Детальная информация о файлах представлена в виде таблицы, а результаты группировки - в виде списка групп с раскрываемыми записями:

- о составе таблицы с детальной информацией, добавлении данных в таблицу и действиях с файлами см. "Таблица с информацией о файлах";
- о настройке отображения записей в таблице и в группах см. "Настройка отображения таблицы";
- о фильтрации записей в таблице см. "Единый фильтр";
- о поиске дубликатов или похожих по загружаемому примеру см. "Поиск по примеру";
- о выгрузке данных из таблицы см. "Выгрузка отчетов с информацией о файлах".

Также см. "Пример работы с таблицей".

10.2.1 Таблица с информацией о файлах

На вкладке **Информация о файлах** Data Discovery отображает результаты сканирования ресурсов. Информация о файлах представлена в виде таблицы. В конце каждой сессии обработки ресурса в таблицу добавляется информация о просканированных за эту сессию файлах.

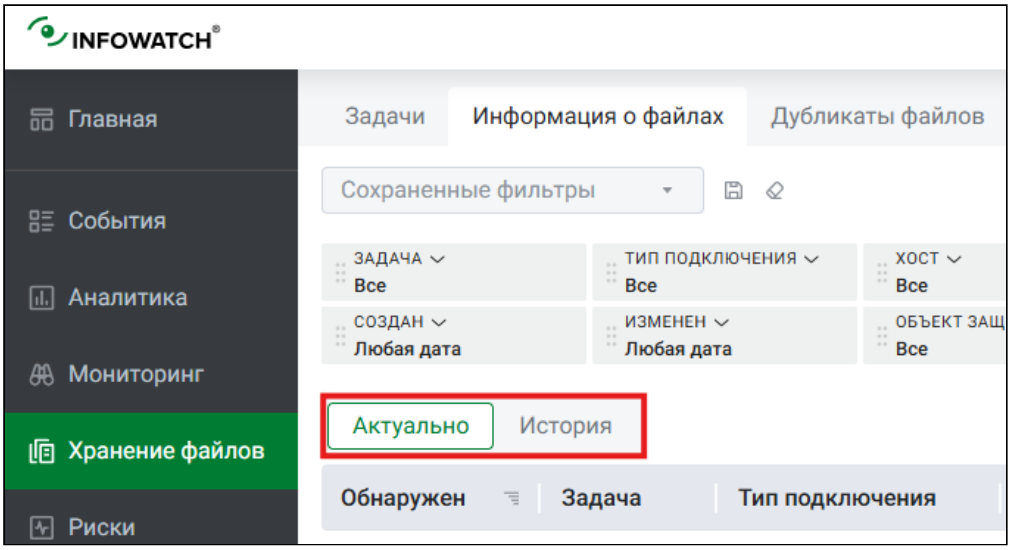
Режимы отображения информации о файлах

Data Discovery позволяет просматривать информацию о файлах в двух режимах: **Актуально** и **История**.

В режиме **Актуально** в таблице представлены **данные последних выполненных сканирований**. В этом режиме в таблице для каждого файла отображается только одна запись.

В режиме **История** – **данные всех сканирований, и последних, и предыдущих**. В этом режиме для каждого файла отображаются результаты всех его прошлых сканирований, что позволяет пользователю использовать историю изменений метаданных и содержимого файла.
Просмотр записей об отсутствующих файлах – у которых заполнено поле **Отсутствует**, доступен только в режиме **История**.

Для выбора режима отображения воспользуйтесь переключателем над таблицей.



Структура таблицы с данными одинакова в обоих режимах.

Добавление записей в таблицу с детальной информацией

Проход – полное сканирование одного ресурса. Проход может состоять из нескольких сессий.

Сессия – подключение к ресурсу в рамках прохода.

В рамках первого прохода по ресурсу Data Discovery собирает метаданные всех файлов на ресурсе и добавляет для каждого файла запись в таблицу обоих режимов, **Актуально** и **История**.

При последующих сессиях Система обрабатывает только новые и измененные файлы. Чтобы определить измененные файлы, Система проверяет метаданные файлов на изменения:

- Путь к файлу;
- Имя файла;
- Размер файла;
- Дата и время изменения файла;
- Владелец файла;

- Пользователи/группы, которые имеют доступ к файлу.

> **Примечание:**
>
> Проверка следующих метаданных на изменения поддерживается только при сканировании SMB-ресурсов на хостах под управлением ОС семейства Windows:
>
> - Владелец файла
> - Пользователи/группы, которые имеют доступ к файлу

Если просканирован новый или измененный файл, то Система в конце сессии добавляет новую запись в таблицу обоих режимов, **Актуально** и **История**. Запись о прошлом сканировании измененного файла Система отобразит только в таблице режима **История**.

> **Важно!**
>
> В режиме **Актуально** извлеченное содержимое файла будет отображаться, только если **последнее завершенное сканирование** было выполнено задачей с включенным режимом **Проводить контентный анализ файлов** и/или **Определять похожие файлы**. Если в задаче, в рамках которой выполнено последнее сканирование файла, не был выбран ни один из режимов с экстракцией содержимого, в таблице в режиме **Актуально** будут отображаться только метаданные файла.
>
> Если файл сканируют несколько задач, в таблице в режиме **Актуально** будет отображаться результат только последнего сканирования, независимо от результатов предыдущих. **Не рекомендуется** сканировать один ресурс несколькими задачами с отличающимися режимами обработки.

> **Пример:**
>
> Файл последовательно просканирован двумя задачами с разными режимами обработки:
>
> 1. Просканирован первой задачей, в которой включен режим **Проводить контентный анализ файлов** – в таблице в обоих режимах добавлена запись о файле с возможностью просмотра содержимого, в столбце **Задача** указано название первой задачи.
> 2. Выполнено повторное сканирование файла второй задачей, в которой выключены все режимы обработки – в таблице в режиме **Актуально** запись заменена новой, без содержимого и только с метаданными, в столбце **Задача** указано название второй задачи.
>   Прошлая запись с возможностью просмотра содержимого отображается в таблице в режиме **История**, в столбце **Задача** указано название первой задачи.
>
> Если вы удалите вторую задачу сканирования, вторая запись будет удалена, а первая с возможностью просмотра содержимого снова будет отображена в таблице в режиме **Актуально**.

Структура таблицы с детальной информацией о файлах

Каждая запись таблицы содержит информацию об одном просканированном файле. Каждая запись содержит следующие поля:

Поле	Описание
Обнаружен	Дата и время, когда файл был обнаружен на ресурсе. Дата и время указываются в формате `dd.MM.yyyy HH:mm:ss` **Примечание**: Если неизмененный файл обнаружен повторно в следующих сессиях в рамках задачи, то: • в таблицу не добавляются новые записи; • значение поля **Обнаружен** в записи о файле не обновляется.
Задача	Название задачи сканирования, в рамках которой файл был обнаружен на ресурсе **Примечание**: Если файл на ресурсе обнаружен в рамках нескольких задач сканирования, то в таблицу добавляется запись о файле для каждой такой задачи
Хост	Доменное имя или IP-адрес хоста, на котором был обнаружен файл. **Примечание**: для файлов, обнаруженных при сканировании Яндекс Диска, запись будет иметь вид `<ID_организации.disk.yandex.ru>`. Например, `<8120422.disk.yandex.ru>`
Расположение	Расположение файла – полный путь к директории, в которой был обнаружен файл. **Примечание**: Полный путь включает в себя путь сканирования ресурса. Для файлов с SharePoint-хостов полный путь также включает в себя путь к сайту.
Имя файла	Имя файла. Система получает значение из файловой системы ресурса.
Размер	Размер файла. Система получает значение из файловой системы ресурса.
Создан	Дата и время создания файла. Дата и время указываются в формате `dd.MM.yyyy HH:mm:ss` Система получает значение из файловой системы ресурса. **Примечание**: Определение даты создания файла не поддерживается в следующих случаях:

Поле	Описание
	• при сканировании SMB-ресурсов на хостах под управлением Linux; • при сканировании хостов под управлением Linux по протоколу SSH. В этих случаях поле **Создан** заполняется значением поля **Изменен**.
Изменен	Дата и время изменения файла. Дата и время указываются в формате `dd.MM.yyyy HH:mm:ss` Система получает значение из файловой системы ресурса. **Примечание**: Дата и время изменения файла в поле указывают на следующие изменения: • При сканировании SMB-ресурсов на хостах под управлением ОС семейства Windows: • последнее изменение содержимого файла; • последнее изменение имени файла; • последнее изменение пути к файлу; • последнее изменение владельца файла или списка пользователей и групп, которые имеют доступ к файлу; • При сканировании SMB-ресурсов на хостах под управлением Linux / хостов под управлением Linux по протоколу SSH: • последнее изменение содержимого файла; • При сканировании библиотек документов в SharePoint Server 2016: • последнее изменение содержимого файла; • последнее изменение имени файла.
Отсутствует	Дата и время окончания прохода в рамках задачи, если файл не был обнаружен повторно на ресурсе по результатам прохода. Дата и время указываются для первого прохода, в который файл не был обнаружен. Дата и время указываются в формате `dd.MM.yyyy HH:mm:ss` Система не обнаружит файл повторно в рамках задачи, если: • Файл был удален с ресурса. • У файла изменились метаданные. В этом случае в таблицу добавляется новая запись. • В настройках задачи были изменены фильтры файлов/директорий и при повторном сканировании файл не соответствовал условиям фильтров. Записи об отсутствующих файлах отображаются только в режиме История.

Поле	Описание
Владелец	В поле **Владелец** Data Discovery указывает имя персоны или группы, которая является владельцем файла.
Доступен для	В поле **Доступен для** Data Discovery указывает список имен персон и групп, которые имеют доступ к файлу. Data Discovery идентифицирует персон и группы по SID, значения которых он получает из файловой системы ресурса. Data Discovery определяет имена по SID на основе данных из Active Directory (см. "Идентификация персон и групп"). Если данные для идентификации персоны или группы отсутствуют, Data Discovery вместо имени укажет SID, например: `S-1-5-21-1922741149-3613679056-2300624754-1000`
Объекты защит	Объекты защиты, детектированные в этом файле с помощью Модуля контентного анализа

Важно!

Получение SID владельца файла и SID пользователей/групп, которые имеют доступ к файлу, поддерживается для:

- SMB-ресурсов на хостах под управлением ОС семейства Windows;
- библиотек документов в SharePoint Server 2016.

Возможно некорректное заполнение полей **Владелец** и **Доступен для**, если:

- Производится сканирование SMB-ресурса на хосте под управлением Linux;
- Производится сканирование SMB-ресурса, созданного на сервере SMB, который не поддерживает файловые атрибуты для определения владельца файла и пользователей/групп, которые имеют доступ к файлу;
- Производится сканирование SMB-ресурса, созданного в файловой системе, которая не поддерживает файловые атрибуты для определения владельца файла и пользователей/групп, которые имеют доступ к файлу.

Получение SID владельца файла и SID пользователей/групп, которые имеют доступ к файлу, не поддерживается при сканировании хостов:

- под управлением Linux по протоколу SSH;
- по протоколу FTP;
- по протоколу WebDAV;
- с системой Alfresco.

В этом случае атрибуты *Отправители* и *Получатели* имеют значение "iw data discovery".

Примечание:

Дата и время в полях **Обнаружен**, **Изменен**, **Создан**, **Не обнаружен** отображается в часовом поясе рабочей станции, с которой осуществляется подключение к веб-интерфейсу Data Discovery.

Идентификация персон и групп

Data Discovery может идентифицировать персон и группы по SID.

В случае успешной идентификации:

- вместо SID **персоны** в полях **Владелец** и **Доступен для** Data Discovery отобразит имя, указанное в досье;
- вместо SID **группы** в полях **Владелец** и **Доступен для** Data Discovery отобразит название группы, полученное из Active Directory.

Идентификация учетных записей ОС Windows по известным (well-known) SID

Система идентифицирует следующие учетные записи ОС Windows (подробнее об известных SID см. документацию Microsoft Windows):

SID	Учетная запись	Описание
S-1-5-18	System	Учетная запись, которую использует операционная система
S-1-1-0	Everyone	Все пользователи

Данные об указанных SID по умолчанию добавлены в Систему. Система производит идентификацию во всех записях таблицы.

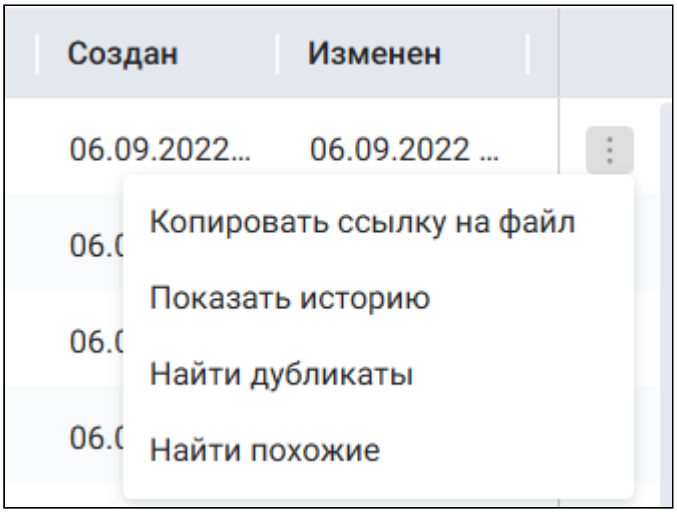
Идентификация персон и групп из Active Directory

Все персоны, добавленные в Систему результате синхронизации с Active Directory имеют тип контакта SID, что позволяет идентифицировать их. Если группы из Active Directory имеют SID, они будут идентифицированы. Подробнее о добавлении данных о персонах и группах в Систему см. *InfoWatch Data Discovery. Руководство администратора. Интеграции*.

Система идентифицирует персон и группы во всех записях таблицы по актуальным данным каждой интеграции. Идентификация по данным последней синхронизации может занять некоторое время – до 15 минут с момента завершения синхронизации.

Действия с файлами

Справа от каждой записи вы можете нажать кнопку ⋮, чтобы выбрать одно из доступных действий с файлом:



- **Копировать ссылку на файл**;
- **Показать историю** – вы можете переключить таблицу в режим **История** и просмотреть результаты всех сканирований выбранного файла;
- **Найти дубликаты** – если в задаче сканирования был выбран хотя бы один из режимов обработки, Data Discovery получит информацию о содержимом файлов и сможет искать дубликаты файла, выбранного в таблице;
- **Найти похожие** – если выбранный файл прошел обработку на определение похожих, вы можете найти файлы похожие с ним смыслу. Похожие будут найдены в числе файлов, также прошедших обработку на определение похожих. В результате Data Discovery отобразит файлы, которые по результатам анализа Data Analysis Service входят в одну тематическую группу.

> **Примечание:**
>
> В режиме **История** в результатах поиска похожих могут быть записи прошлых сканирований одного файла.

Оба действия, и **Найти дубликаты**, и **Найти похожие**, аналогичны Поиску по примеру, но в качестве примера для поиска будет использован файл из таблицы.

Поиск дубликатов или похожих по файлу из таблицы будет недоступен при выполнении любого из условий:

- размер файла превышает 4 ГБ;
- при сканировании не были созданы копии файлов - в настройках задачи сканирования не выбран ни один из режимов обработки.

Чтобы файл был обработан на определение похожих, извлеченный **текстовый контент** должен быть **не менее 200 Байт**. Контент меньше 200 Байт не будет отправлен на анализ в Data Analysis Service, но будет доступен для просмотра в веб-интерфейсе Центра расследований в разделе **Хранение файлов**.

Если файлы были просканированы Data Discovery версии **ниже 1.8**, информация об их содержимом отсутствует. Поиск по таким файлам не осуществляется.
В этом случае:

1. Скопируйте задачу, в которой были просканированы файлы.
2. Повторно просканируйте файлы в новой задаче.

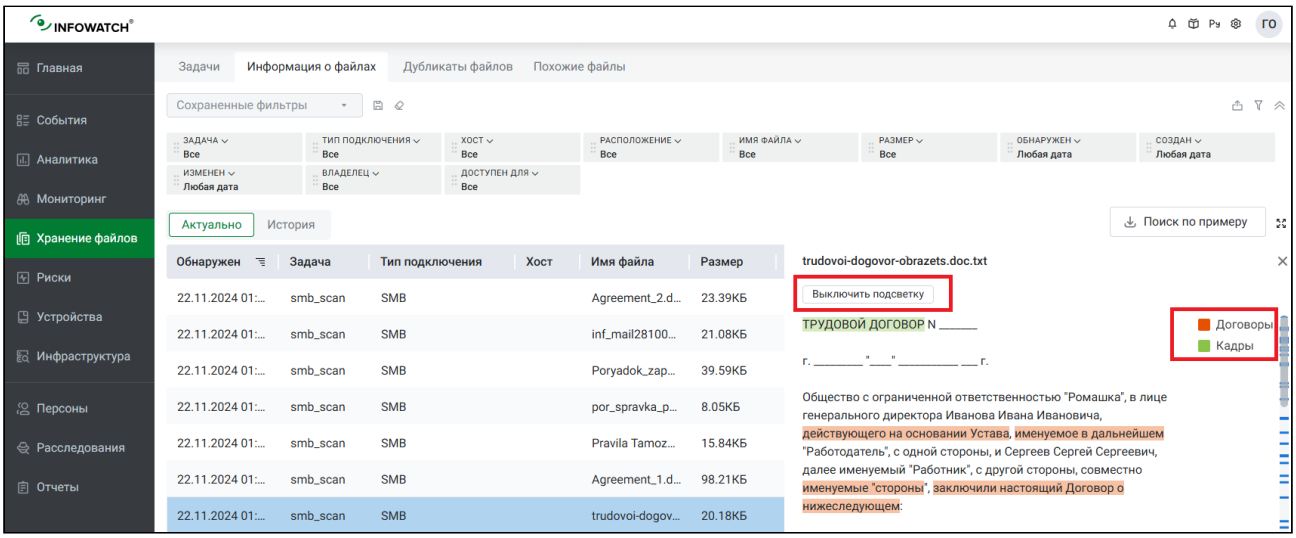
Просмотр извлеченных текстовых контентов и результатов контентного анализа

Чтобы просмотреть извлеченный из файла текстовый контент, выберите строку в таблице. Справа откроется окно с именем файла и извлеченными текстовыми данными. Если из файла не извлечены текстовые данные, окно будет содержать только имя файла.

Из одного файла может быть извлечено несколько текстовых контентов, но для анализа и отображения они будут объединены в один.

Пример: если экстрактор обрабатывает файл-архив, содержащий несколько документов, из каждого может быть извлечен текстовый контент, но в итоге они будут объединены в один общий, относящийся к файлу-архиву.

Если Модуль контентного анализа обнаружил в текстовом контенте **объекты защиты**, относящиеся к ним фрагменты будут выделены цветом.



Модуль контентного анализа поддерживает следующие технологии из конфигурации Traffic Monitor:

- Автолингвист;
- Бланки;
- Выгрузки из БД;
- Категории и термины;
- Текстовые объекты;
- Эталонные документы.

На легенде справа от текстовых данных указаны названия обнаруженных объектов защиты. Легенда интерактивна: чтобы отключить/включить подсветку фрагментов, относящихся к конкретному объекту защиты, нажмите на его название в легенде. Если навести курсор на выделенный цветом текстовый фрагмент, отобразится подсказка с именем объекта защиты.

Вы можете убрать все выделение цветом с помощью кнопки **Выключить подсветку** над содержимым.

10.2.2 Настройка отображения таблицы

Чтобы изменить порядок сортировки записей о файлах в таблице, нажмите на заголовок столбца, по которому нужно выполнить сортировку. Все записи таблицы будут отсортированы по возрастанию/убыванию значений выбранного столбца.

Чтобы переместить столбец, нажмите на заголовок столбца и, не отпуская, перемещайте заголовок вдоль строки заголовков.

Чтобы изменить ширину столбца, нажмите на правую границу заголовка столбца и, не отпуская, перемещайте границу столбца.

Навигация по страницам таблицы осуществляется с помощью панели, расположенной внизу экрана.

10.2.3 Поиск по примеру

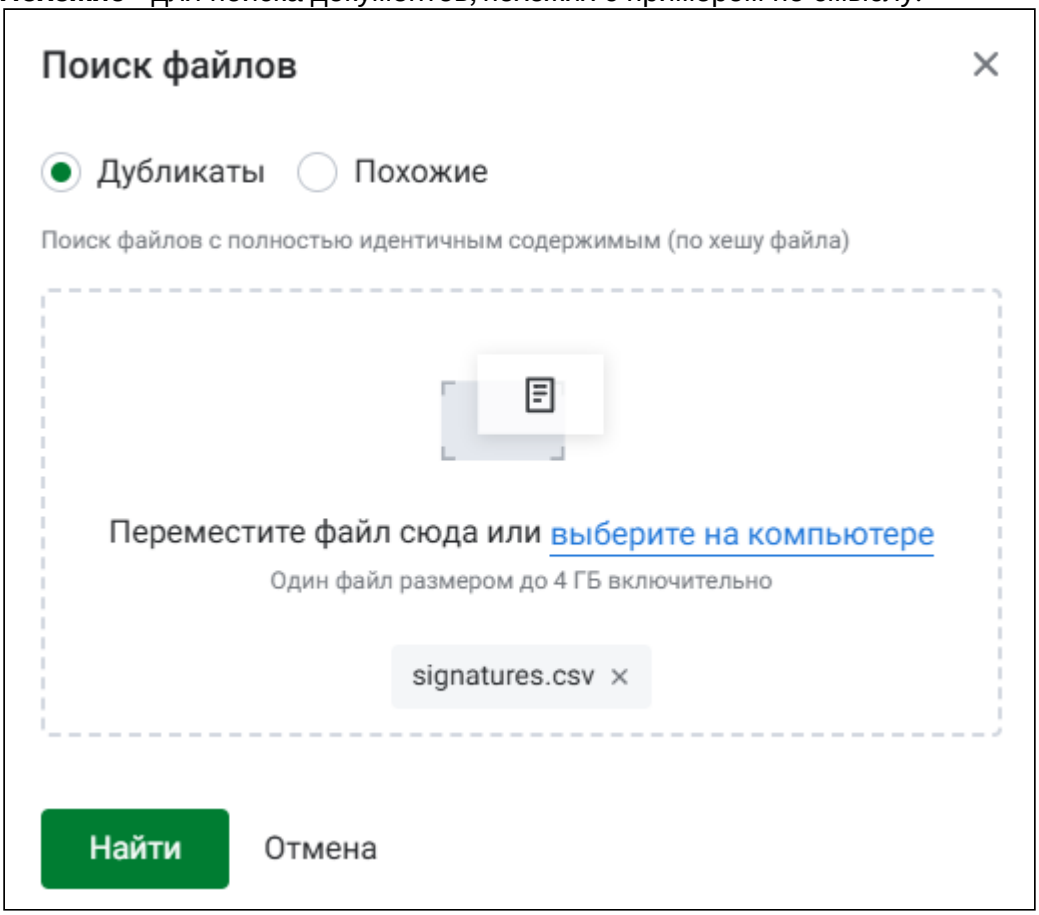
Data Discovery позволяет искать дубликаты или похожие файлы по загружаемому примеру.

Поиск по примеру имеет следующие ограничения:

- **Для дубликатов**: максимальный размер загружаемого примера **не должен превышать 4 ГБ**;
- **Для похожих**:
  - **максимальный** размер загружаемого примера должен быть **менее 300 МБ**;
  - **максимальный** размер загружаемого в качестве примера **текстового файла** - **не более 25 МБ**.

Для поиска файлов по примеру:

1. Справа над таблицей с информацией о файлах нажмите **Поиск по примеру**.
2. В открывшемся окне выберите тип поиска:
   - **Дубликаты** – для поиска файлов с полностью идентичным содержимым;
   - **Похожие** – для поиска документов, похожих с примером по смыслу.



> **Примечание:**
>
> В режиме **История** в результатах поиска похожих могут быть записи прошлых сканирований одного файла.

3. Загрузите пример для поиска одним из предложенных способов.
4. Нажмите **Найти**.
5. Дождитесь окончания поиска.

В зависимости от выбранного типа поиска таблице будут отображены записи о файлах:

- дубликатах загруженного примера
  или
- входящих в одну тематическую группу с примером.

К записям будут применены условия Единого фильтра (см. "Единый фильтр").

> **Важно!**
>
> Если файлы были просканированы:
>
> - Data Discovery 1.7 или ниже - информация об их содержимом отсутствует. Поиск дубликатов по таким файлам не осуществляется;
> - Data Discovery 1.9 или ниже - из них не были извлечены текстовые контенты. Поиск похожих по таким файлам не осуществляется.
>
> Также поиск не осуществляется по файлам, которые ранее были просканированы, но в задаче сканирования не были выбраны режимы обработки.
>
> Для этих случаев:
>
> 1. Скопируйте задачу, в которой были просканированы файлы.
> 2. Повторно просканируйте файлы в новой задаче с нужным режимом обработки.

Сброс результатов поиска

Чтобы сбросить результаты поиска, нажмите  справа от имени загруженного примера.

10.2.4 Выгрузка отчетов с информацией о файлах

В разделе **Хранение файлов** на вкладках **Информация о файлах**, **Дубликаты файлов** и **Похожие файлы** вы можете выгрузить представленную информацию в отчет. Формат файла отчета – XLSX.

Чтобы выгрузить данные:

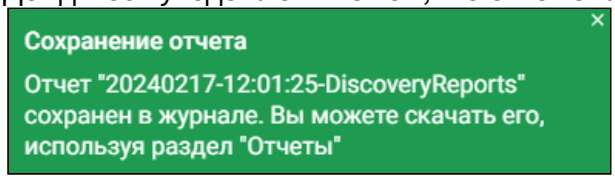
1. Выберите данные и настройте их отображение:
   a. Выберите режим отображения: **Актуально** или **История**.
   b. Настройте фильтр (см. Единый фильтр).

> **Важно!**

Если после фильтрации таблица содержит более 10 000 записей, то в отчет добавляются только первые 10 000 записей.

c. Выберите сортировку записей (см. Настройка отображения таблицы).
d. Если необходимо, выполните поиск дубликатов или похожих по загружаемому примеру (см. Поиск по примеру).
В отчет будут выгружены данные в выбранном представлении.

2. В правой верхней части экрана нажмите .
3. Выберите **Выгрузить отчет**.
4. В открывшемся окне определите состав отчета:
   a. Чтобы добавить лист с дополнительной информацией в отчет, выберите в списке **Дополнительная информация**. Лист содержит:
      - дату и время формирования отчета;
      - настройки фильтров;
      - ссылку для перехода в Консоль управления.
   b. В зависимости от выбранной вкладки вы можете выбрать:
      - **Информация о файлах**;
      - **Дубликаты файлов**;
      - **Похожие файлы**.
   c. Нажмите **Сохранить**.
5. Дождитесь уведомления о том, что отчет был сохранен.



6. Для загрузки файла отчета перейдите в раздел **Отчеты → Архив отчетов**.
7. Нажмите на название отчета, чтобы сохранить файл на ваш компьютер.

На вкладке **Информация о файлах** вы можете **настроить регулярную отправку отчетов.** Для этого:

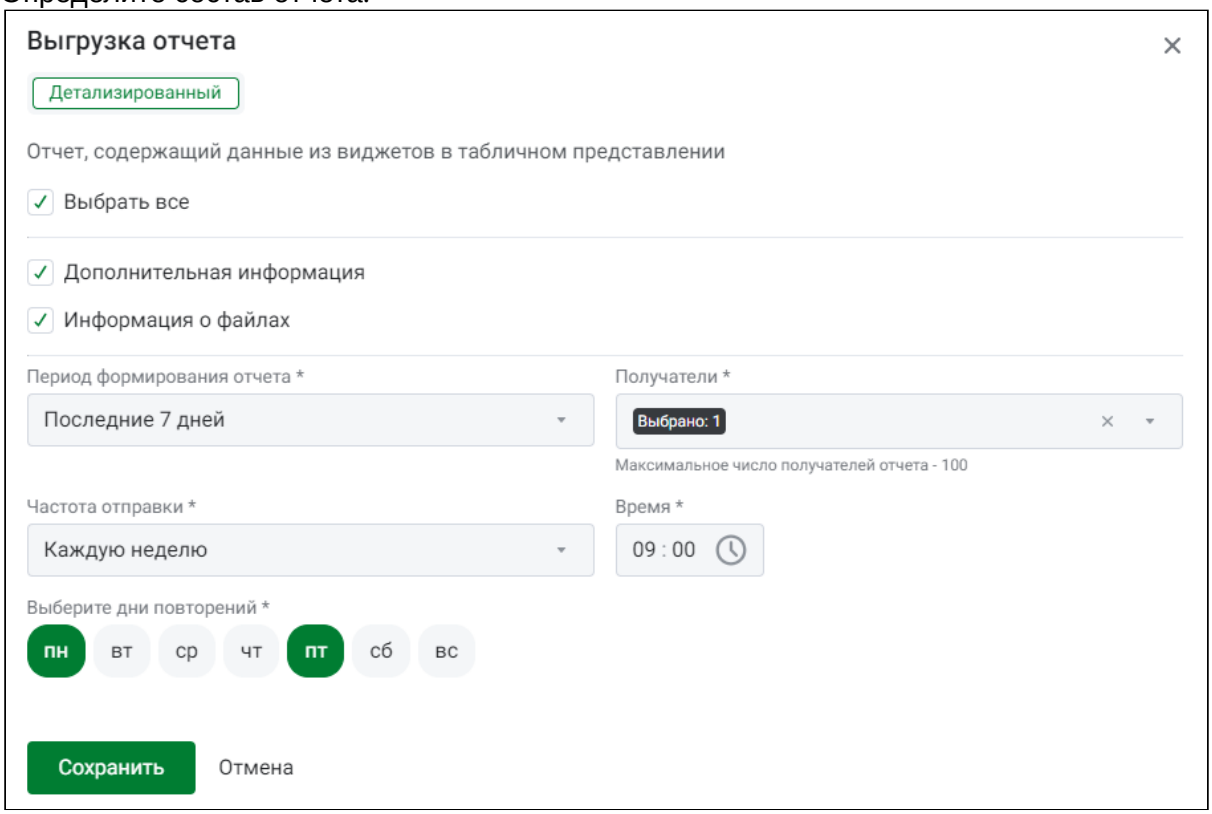
1. Настройте фильтр (см. Единый фильтр), сортировку записей (см. Настройка отображения таблицы) и/или поиск дубликатов по образцу (см. Поиск по примеру).

   Важно!

   Если после фильтрации таблица содержит более 10 000 записей, то в отчет добавляются только первые 10 000 записей.

   В отчет будут выгружены данные в выбранном представлении.
2. В правой верхней части экрана нажмите .
3. Выберите **Настроить регулярную отправку**.

4. Определите состав отчета.



5. Также задайте параметры отправки:

Параметр	Описание
Период формирования отчета	Укажите, за какой период необходимо отправлять отчет: • **Последние сутки** • **Текущий день** • **Последние 3 дня** • **Последние 7 дней** • **Последние 30 дней**
Получатели	Укажите получателей отчета
Частота отправки	Укажите, с какой периодичностью Офицеру необходимо получать отчет: • **Каждый день** • **Каждую неделю** Вы можете выбрать дни недели, по которым Data Discovery будет отправлять отчеты. • **Каждый месяц** Отчет будет отправляться 1 числа месяца, следующего за отчетным.
Время	Укажите время отправки отчета

6. Нажмите **Сохранить**. Будет создана задача на регулярную отправку. Список задач вы можете просмотреть в разделе **Отчеты** на вкладке **Задачи на регулярную отправку**. Подробнее см. Задачи на регулярную отправку.

> **Важно!**
>
> Для регулярной отправки отчетов в Центре расследований должно быть настроено соединение с почтовым серверов (подробнее см. *InfoWatch Data Discovery. Руководство администратора. Сервисы для уведомлений*).

10.2.5 Пример работы с таблицей

В таблице содержится информация о файлах, просканированных на различных ресурсах в рамках нескольких задач сканирования.

Необходимо определить:

- на каких рабочих станциях располагался файл с информацией о доходах;
- в какой временной период файл располагался на рабочих станциях.

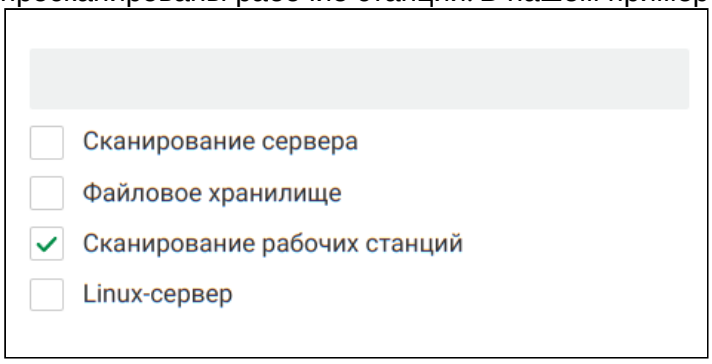
О файле с информацией о доходах известно, что:

- имя файла включает в себя строку "доход";
- файл был создан в текущем месяце;
- размер файла не превышает 5 МБ.

Настройка фильтров

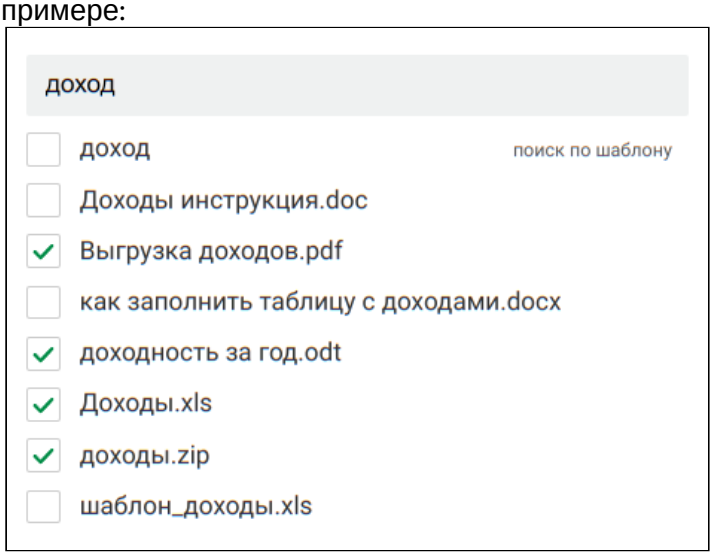
Настройте фильтрацию для записей в таблице:

1. В фильтре **Задача** выберите из списка имя задачи, в рамках которой были просканированы рабочие станции. В нашем примере:



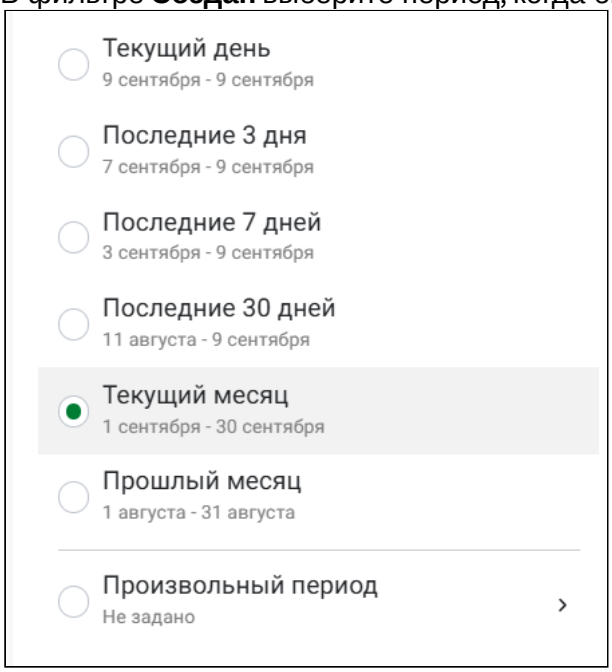
2. Выберите значения для фильтра **Имя файла**:
   a. В фильтре введите в строку поиска "доход";

b. Выберите интересующие имена файлов из предложенных в списке. В нашем примере:

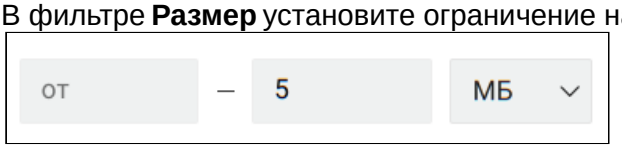


> **Примечание:**
>
> Отображение элементов в списке не зависит от условий других фильтров. Например, в списке могут присутствовать файлы из той задачи сканирования, которая не подходит под условие фильтра **Задача**.

3. В фильтре **Создан** выберите период, когда был создан файл. В нашем примере:



4. В фильтре **Размер** установите ограничение на размер файла. В нашем примере:



Результаты фильтрации

В таблице с информацией о файлах отображаются только записи, которые соответствуют одновременно всем указанным условиям:

- имя задачи сканирования, в которой был обнаружен файл – "Сканирование рабочих станций";
- имя файла – "Выгрузка доходов.pdf", "доходность за год.odt", "Доходы.xls" или "доходы.zip";
- файл был создан в текущем месяце;
- размер файла не превышает 5 МБ.

Выполните сортировку в таблице по столбцу **Хост**:

Обнаружен	Задача	Хост	Имя файла	Размер	Создан	Изменен	Отсутствует
02.09.2022 15:22:17	Сканирование рабочих станций	user1.example.com	доходы.xls	4.21 МБ	02.09.2022 15:15:17	02.09.2022 15:15:17	--.--.-- --:--:--
04.09.2022 11:22:17	Сканирование рабочих станций	user2.example.com	доходы.xls	4.21 МБ	04.09.2022 11:07:38	04.09.2022 11:07:38	04.09.2022 14:59:10
04.09.2022 13:52:11	Сканирование рабочих станций	user2.example.com	Выгрузка доходов.pdf	2.34 МБ	04.09.2022 13:42:36	04.09.2022 13:42:36	--.--.-- --:--:--
05.09.2022 18:01:29	Сканирование рабочих станций	user3.example.com	доходы.xls	4.21 МБ	05.09.2022 17:59:20	05.09.2022 17:59:20	06.09.2022 13:01:29
06.09.2022 13:01:15	Сканирование рабочих станций	user3.example.com	доходы.xls	4.17 МБ	05.09.2022 17:59:20	06.09.2022 13:00:01	--.--.-- --:--:--

Файлы, соответствующие условиям, были найдены на 3 рабочих станциях:

Имя рабочей станции	Описание записей о файлах
user1.example.com	Система обнаружила на хосте файл **доходы.xls**. Файл был впервые обнаружен 02.09.2022 в 15:22:17. Поле **Отсутствует** не заполнено, что возможно в следующих случаях: • Неизмененный файл **доходы.xls** еще находится на ресурсе. • Файл был изменен/удален, но повторное сканирование ресурса после изменения/удаления файла не проводилось.
user2.example.com	1. Система обнаружила на хосте файл **доходы.xls**. Файл был впервые обнаружен 04.09.2022 в 11:22:17. Поле **Отсутствует** заполнено – Система не обнаружила файл при повторном сканировании, что возможно в следующих случаях: • Файл был удален с ресурса. • Файл был изменен. Если после изменения файл соответствует условиям фильтров в настройках задачи, то при повторном сканировании в таблицу добавляется новая запись о измененном файле. **Примечание**: Запись о измененном файле может не отображаться в таблице, если файл не подходит под текущие условия фильтров таблицы. • Условия фильтров в настройках задачи были изменены и при повторном сканировании файл не соответствовал новым условиям.

Имя рабочей станции	Описание записей о файлах
	2. Система обнаружила на хосте файл **Выгрузка доходов.pdf**. Файл был впервые обнаружен 04.09.2022 в 13:52:11. Поле **Отсутствует** не заполнено.
user3.example.com	1. Система обнаружила на хосте файл **доходы.xls**. Файл был впервые обнаружен 05.09.2022 в 18:01:29. Поле **Отсутствует** заполнено. 2. Система обнаружила на хосте файл **доходы.xls**. Файл был впервые обнаружен 06.09.2022 в 13:01:15. Поле **Отсутствует** не заполнено. Дата создания и имя файла в обеих записях одинаковые. В записях различаются размер файла и дата изменения файла. Файл **доходы.xls** мог быть отредактирован на хосте, после чего при повторном сканировании: a. Был обнаружен измененный файл **доходы.xls**, в таблицу была добавлена новая запись. b. Не был обнаружен файл **доходы.xls** с неизмененными метаданными, у существующей записи Система заполнила дату и время, когда файл не был обнаружен.

10.3 Вкладка "Дубликаты файлов"

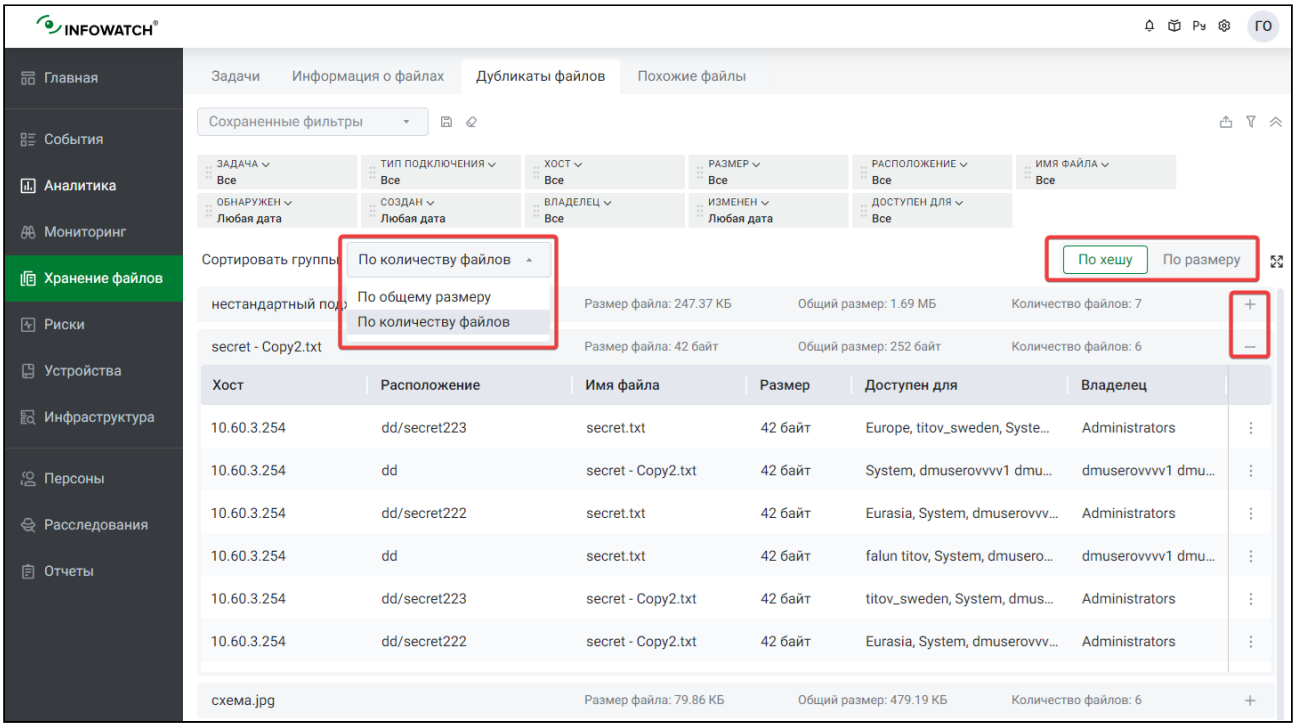
Помимо поиска по образцу Data Discovery автоматически сравнивает просканированные файлы:

- ***по хешу*** – если при сканировании Data Discovery создавал копии файлов. Для этого в задаче сканирования **должен быть выбран хотя бы один из режимов обработки**. В противном случае Data Discovery сможет искать дубликаты только по размеру, так как не будут созданы копии файлов для анализа;
- ***по размеру*** (до байта) – если при сканировании Data Discovery не создавал копии файлов и собирал только метаданные.

Файлы, которые совпадают по хешу или по размеру, Data Discovery считает дубликатами и объединяет в группы.

> **Примечание:**
>
> Точность поиска по хешу значительно выше.

Результаты поиска дубликатов представлены на вкладке **Дубликаты файлов** в виде списка групп, где каждая запись представляет собой таблицу с информацией о файлах, которые в нее входят.



Имя группы - имя первого файла, который в нее входит.
Для каждой группы указаны:

- размер файла;
- общий размер группы;
- количество файлов в группе.

Доступна сортировка групп:

- **по общему размеру**;
- **по количеству файлов**.

Чтобы открыть группу, нажмите справа .

Для каждого файла группы указана следующая информация:

Поле	Описание
Хост	Доменное имя или IP-адрес хоста, на котором был обнаружен файл
Расположение	Расположение файла – полный путь к директории, в которой был обнаружен файл. **Примечание**: Полный путь включает в себя путь сканирования ресурса. Для файлов с SharePoint-хостов полный путь также включает в себя путь к сайту.
Имя файла	Имя файла. Система получает значение из файловой системы ресурса.
Размер	Размер файла.

Поле	Описание
	Система получает значение из файловой системы ресурса.
Владелец	В поле **Владелец** Data Discovery указывает имя персоны или группы, которая является владельцем файла.
Доступен для	В поле **Доступен для** Data Discovery указывает список имен персон и групп, которые имеют доступ к файлу. Data Discovery идентифицирует персон и группы по SID, значения которых он получает из файловой системы ресурса. Data Discovery определяет имена по SID на основе данных из Active Directory (см. Идентификация персон и групп в статье Таблица с информацией о файлах). Если данные для идентификации персоны или группы отсутствуют, Data Discovery вместо имени укажет SID, например: `S-1-5-21-1922741149-3613679056-2300624754-1000`
Объекты защиты	Объекты защиты, детектированные в этом файле с помощью Модуля контентного анализа

Справа от каждой записи вы можете нажать ⋮ и выбрать **Копировать ссылку на файл**.

Чтобы изменить порядок сортировки записей о файлах в таблице, нажмите на заголовок столбца, по которому нужно выполнить сортировку. Все записи таблицы будут отсортированы по возрастанию/убыванию значений выбранного столбца.

Чтобы переместить столбец, нажмите на заголовок столбца и, не отпуская, перемещайте заголовок вдоль строки заголовков.

Чтобы изменить ширину столбца, нажмите на правую границу заголовка столбца и, не отпуская, перемещайте границу столбца.

Навигация по страницам таблицы осуществляется с помощью панели, расположенной внизу экрана.

Указанные поля и логика их заполнения аналогичны полям из таблицы с детальной информацией.

На странице с дубликатами вам так же доступен **просмотр извлеченных из файла текстовый контентов**. Для этого выберите строку в таблице. Справа откроется окно с именем файла и извлеченными текстовыми данными. Если из файла не извлечены текстовые данные, окно будет содержать только имя файла.

> **Примечание:**
>
> Из одного файла может быть извлечено несколько текстовых контентов, но для анализа и отображения они будут объединены в один.
>
> **Пример**: если экстрактор обрабатывает файл-архив, содержащий несколько документов, из каждого может быть извлечен текстовые контент, но в итоге они будут объединены в один общий, относящийся к файлу-архиву.

Если Модуль контентного анализа обнаружил в текстовом контенте **объекты защиты**, относящиеся к ним фрагменты будут выделены цветом. Подробнее см. Таблица с информацией о файлах.

Вы можете отфильтровать данные и **выгрузить отчет**, подробнее см. Выгрузка отчетов с информацией о файлах.

10.4 Вкладка "Похожие файлы"

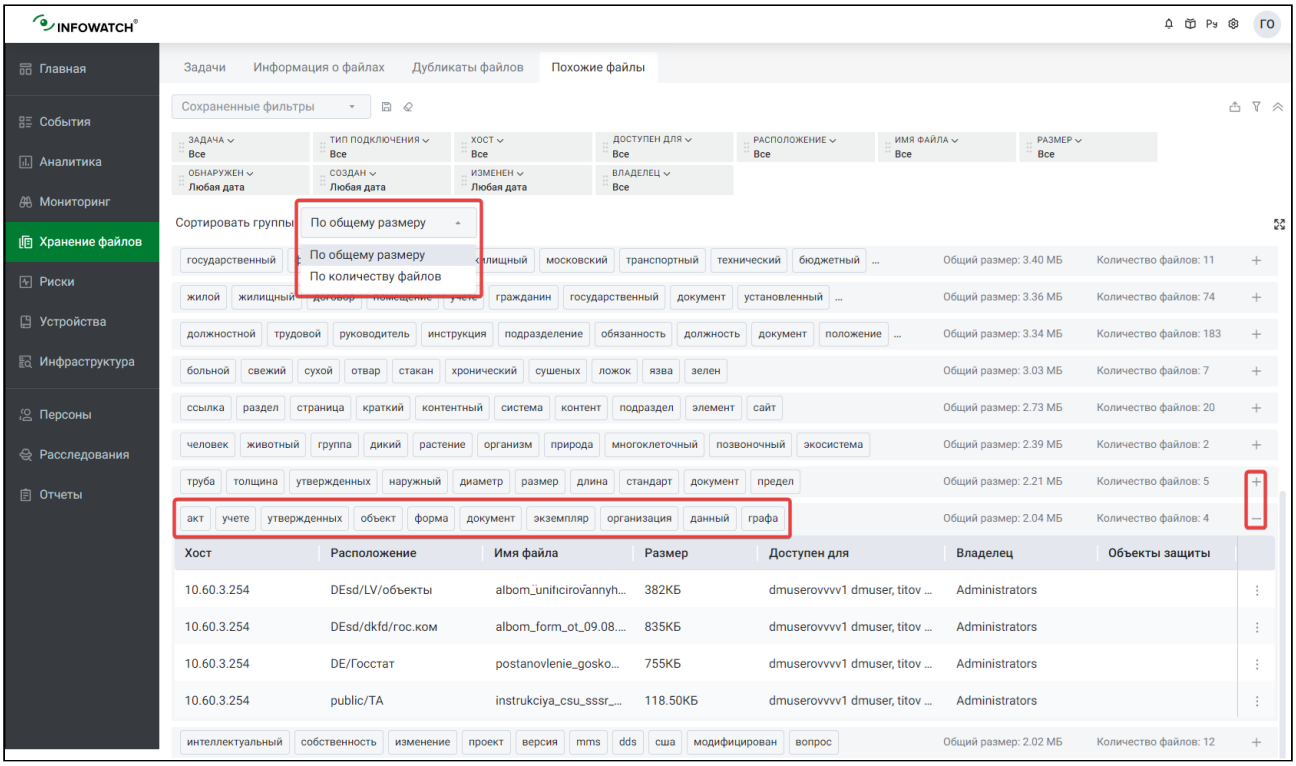
Data Discovery позволяет обнаруживать документы с одинаковыми терминами, похожие по смыслу. Функциональность предназначена для обнаружения мест хранения и идентификации прав доступа к документам одной тематики. Для поиска похожих файлов Data Discovery взаимодействует с модулем InfoWatch Data Analysis Service. При сканировании ресурсов Data Discovery извлекает из копий файлов текстовые контенты, в том числе текст с изображений с помощью OCR, сохраняет их в хранилище и уведомляет модуль Data Analysis Service о наличии данных для обработки. Про установку и настройку Data Analysis Service см. *InfoWatch Data Analysis Service. Руководство по установке, конфигурированию и администрированию*.

Для поиска похожих:

- Data Analysis Service 1.5.0 или 1.6.0 должен быть установлен в одном кластере с Data Discovery;
- в Data Analysis Service должна быть включена функция протокластеризации;
- Data Analysis Service должен быть настроен на обработку ноды, на которой установлен Data Discovery;
- в задаче сканирования должен быть выбран режим обработки **Определять похожие файлы**.

> **Примечание:**
>
> Чтобы файл был обработан на определение похожих, извлеченный **текстовый контент** должен быть **не менее 200 байт**. Контент меньше 200 байт не будет отправлен на анализ в Data Analysis Service, но будет доступен для просмотра в веб-интерфейсе Центра расследований в разделе **Хранение файлов**.
>
> Из одного файла может быть извлечено несколько текстовых контентов, но для анализа и отображения они будут объединены в один.
>
> **Пример**: если экстрактор обрабатывает файл-архив, содержащий несколько документов, из каждого может быть извлечен текстовый контент, но в итоге они будут объединены в один общий, относящийся к файлу-архиву.

Результаты поиска похожих представлены на вкладке **Похожие файлы** в виде списка тематических групп, где каждая запись представляет собой таблицу с информацией о файлах, которые в нее входят. Файлы, объединенные в группу, близки по смыслу.



В результате анализа контентов могут быть выделены группы, состоящие из одного файла. Это означает, что среди просканированных не обнаружено файлов, похожих по смыслу. Подобная уникальность также может иметь значение для Офицера безопасности при анализе тематики файлов. Искать похожие вы также можете на вкладке Информация о файлах, выбрав файл из таблицы для сравнения.

Для каждой группы указаны:

- **теги**, наиболее точно **определяющие общий смысл** документов;
- общий размер группы;
- количество файлов в группе.

Доступна сортировка групп:

- **по общему размеру**;
- **по количеству файлов**.

Чтобы открыть группу, нажмите  справа.

Для каждого файла группы указана следующая информация:

Поле	Описание
Хост	Доменное имя или IP-адрес хоста, на котором был обнаружен файл.
Расположение	Расположение файла – полный путь к директории, в которой был обнаружен файл. **Примечание**: полный путь включает в себя путь сканирования ресурса. Для файлов с SharePoint-хостов полный путь также включает в себя путь к сайту.

Поле	Описание
Имя файла	Имя файла. Система получает значение из файловой системы ресурса.
Размер	Размер файла. Система получает значение из файловой системы ресурса.
Объекты защиты	Объекты защиты, детектированные в этом файле с помощью Модуля контентного анализа.
Владелец	В поле **Владелец** Data Discovery указывает имя персоны или группы, которая является владельцем файла.
Доступен для	В поле **Доступен для** Data Discovery указывает список имен персон и групп, которые имеют доступ к файлу. Data Discovery идентифицирует персон и группы по SID, значения которых он получает из файловой системы ресурса. Data Discovery определяет имена по SID на основе данных из Active Directory (см. "Идентификация персон и групп" в статье *Таблица с информацией о файлах*). Если данные для идентификации персоны или группы отсутствуют, Data Discovery вместо имени укажет SID, например: `S-1-5-21-1922741149-3613679056-2300624754-1000`

Справа от каждой записи вы можете нажать  и выбрать **Копировать ссылку на файл**.

Чтобы просмотреть извлеченный из файла текстовый контент, выберите строку в таблице. Справа откроется окно с именем файла и извлеченными текстовыми данными. Если из файла не извлечены текстовые данные, окно будет содержать только имя файла.

Если Модуль контентного анализа обнаружил в текстовом контенте **объекты защиты**, относящиеся к ним фрагменты будут выделены цветом. Подробнее см. *Таблица с информацией о файлах*.

Чтобы изменить порядок сортировки записей о файлах в таблице, нажмите на заголовок столбца, по которому нужно выполнить сортировку. Все записи таблицы будут отсортированы по возрастанию/убыванию значений выбранного столбца.

Чтобы переместить столбец, нажмите на заголовок столбца и, не отпуская, перемещайте заголовок вдоль строки заголовков.

Чтобы изменить ширину столбца, нажмите на правую границу заголовка столбца и, не отпуская, перемещайте границу столбца.

Навигация по страницам таблицы осуществляется с помощью панели, расположенной внизу экрана.

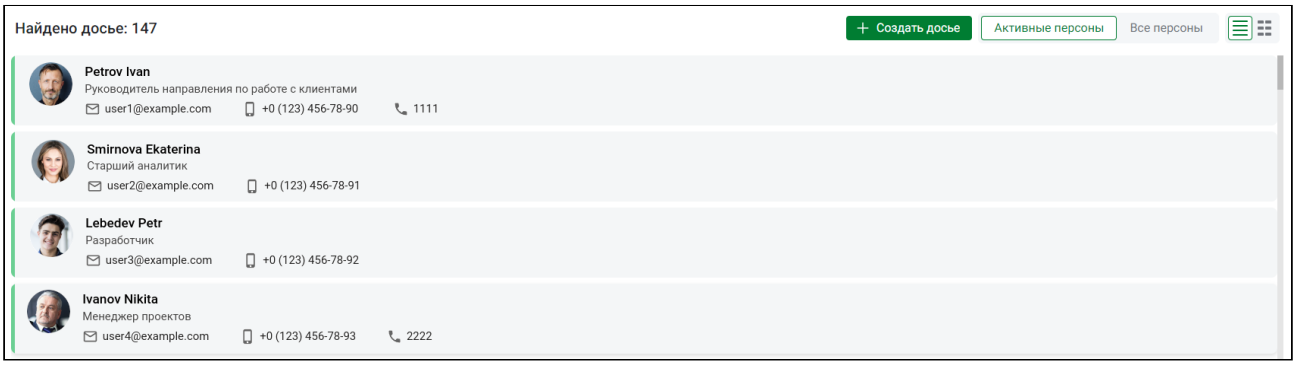
Указанные поля и логика их заполнения аналогичны полям из таблицы с детальной информацией.

Вы можете отфильтровать данные и **выгрузить отчет**, подробнее см. *Выгрузка отчетов с информацией о файлах*.

11 Раздел "Персоны"

Раздел содержит справочник персон, данные которых загружены в Систему при синхронизации с Active Directory или Traffic Monitor, и персон, данные которых внесены вручную.

Для просмотра досье персоны, выберите ее левой кнопкой мыши. Для возврата к списку персон нажмите  в левом верхнем углу рабочей области.



В списке персон отображаются только персоны, удовлетворяющие условиям единого фильтра. Для настройки отображения досье используйте кнопки в правой части рабочей области:

- **Активные персоны** – отобразить досье только сотрудников с активной учетной записью в Active Directory и хотя бы одним событием;

  > **Примечание:**
  >
  > Если у персоны нет событий, либо она заблокирована/удалена в AD, то такая персона считается неактивной и отображается в списке **Все персоны**.
  >
  > Для просмотра активных персон используйте фильтры. Если ни в одном фильтре не выбрано ни одно значение, Система не выведет список активных персон.

- **Все персоны** – отобразить досье всех персон, включая когда-либо работавших в компании;
-  – отображать в виде таблицы;
-  – отображать в виде плитки.

В этом разделе также доступны следующие действия:

- Просмотр досье персоны
- Просмотр и настройка виджетов персональной информации
- Просмотр информации о файлах, которыми владеет персона и к которыми имеет доступ
- Создание, редактирование и удаление досье
- Добавление контактов персоны
- Изменение руководителя персоны
- Объединение досье

11.1 Просмотр досье персоны

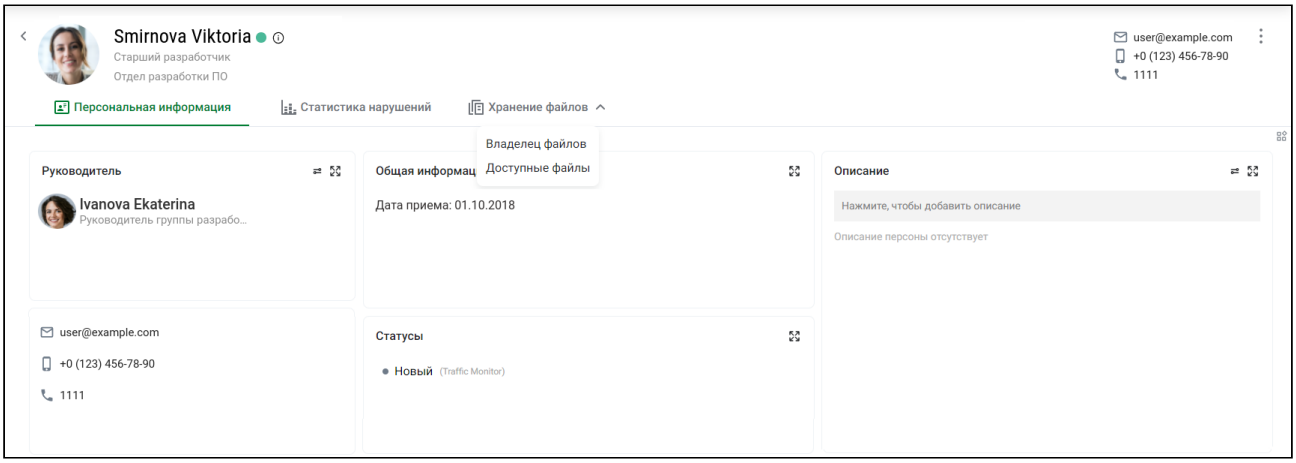
Досье содержит информацию о конкретном сотруднике в организации.

В Досье персоны отображается:

- ФИО сотрудника;
- Должность;
- Отдел;
- Контакты: учетная запись из Active Directory, личные или рабочие контакты персоны, включая адрес электронной почты, номер телефона, номер мобильного телефона.

> **Примечание:**
>
> Все персоны, добавленные при интеграции с Active Directory, имеют тип контакта SID, используемый для их идентификации.
>
> В зависимости от заполненности атрибутов персоны в каталоге домена, Система может добавить в досье учетные записи вида:
>
> - Полное доменное имя (FQDN): *<sAMAccountName>@<UPN-суффикс>*
>   Пример: *ivanov@company.ru*, *pavel.petrov@exampledomain.com*
> - Краткое доменное имя (NetBIOS или DNS): *<sAMAccountName>@<NetBIOS-домен>*
>   Пример: *ivanov@company*, *pavel.petrov@ed*

> **Важно!**
>
> Не создавайте персон с одинаковым значением контакта SID. Наличие в Системе двух и более персон с одинаковыми SID приведет к некорректному определению владельца файла



Справа от имени персоны отображается цветовой индикатор статуса:

-  - активные сотрудники, в том числе персоны, добавленные вручную, а также персоны, у которых истек срок действия учетной записи в AD, но статус в AD остается Активным;
-  - сотрудники, отключенные в AD или имеющие неопределенный статус.

Нажмите на иконку  справа от имени персоны, чтобы увидеть:

- дату создания досье;

- дату обновления досье;
- владельца досье – Система или Офицер безопасности (если досье было добавлено вручную).

Чтобы изменить данные в досье, используйте кнопку  в правой части рабочей области (подробнее см. "Создание, редактирование и удаление досье").

Информация в досье персоны представлена на вкладках:

- **Персональная информация** - содержит виджеты досье, специфичные для персоны: **Общая информация**, **Руководитель**, **Контакты**, **Группы**, **Описание** (см. "Вкладка "Персональная информация"");
- **Статистика нарушений** - виджеты статистики, относящиеся к персоне (см. "Виджеты статистики");
- **Хранение файлов** - при нажатии доступен выбор одного из подразделов:
  - **Владелец файлов** - содержит таблицу с файлами, для которых персона определена владельцем (см. "Вкладка "Хранение файлов"");
  - **Доступные файлы** - содержит таблицу с файлами, к которым персона имеет доступ (см. "Вкладка "Хранение файлов"").

> **Примечание:**
>
> Выше указаны вкладки, доступные при установке Data Discovery.
>
> Набор вкладок в досье персоны зависит от продуктов, установленных в кластере Центра расследований.

Чтобы вернуться к списку персон, нажмите  в верхнем левом углу рабочей области.

11.2 Вкладка "Персональная информация"

Виджеты на вкладке **Персональная информация** позволяют Офицеру безопасности просмотреть все данные по персоне, хранящиеся в Системе.

Название виджета	Описание
Руководитель	Отображает список руководителей персоны, а также позволяет задать руководителя вручную
Контакты	Позволяет управлять контактами персоны, а также отображает контакты, загруженные из подключенных источников
Общая информация	Отображает состояние персоны в различных источниках: дата приема, стаж работы, рабочее место, дата увольнения

Название виджета	Описание
Группы	Отображает список групп из ActiveDirectory или InfoWatch Traffic Monitor, в которые входит персона
Описание	Отображает описание персоны, позволяет задавать и редактировать его вручную

Чтобы развернуть виджет на весь экран, нажмите .

Чтобы отредактировать данные на виджете, нажмите .

11.3 Вкладка "Хранение файлов"

Вкладка **Хранение файлов** включает в себя два подраздела:

- **Владелец файлов**;
- **Доступные файлы**.

В обоих подразделах досье структура таблиц с файлами аналогична разделу **Хранение файлов** (подробнее см. "Таблица с информацией о файлах"), но все записи относятся к выбранной персоне. О фильтрации записей в таблицах см. "Единый фильтр".

Вы также можете выгрузить отчет с представленными данными в формате XLSX (подробнее см. "Выгрузка отчетов с информацией о файлах").

Важно!

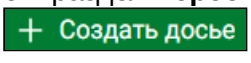
Владельца файлов и возможность доступа персоны к файлам Data Discovery определяет с помощью идентификатора безопасности SID. Если в контактах персоны отсутствует SID, в подразделах **Владелец файлов** и **Доступные файлы** информация будет недоступна.

Персоны, добавленные в результате интеграции с Active Directory (см. *InfoWatch Data Discovery. Руководство администратора. Интеграция с Active Directory*) имеют контакт типа SID.

Если персона добавлена вручную, чтобы Data Discovery мог определять ее как владельца файлов или определять наличие у нее доступа к файлам, добавьте персоне контакт типа SID (см. "Добавление контактов персоны").

11.4 Создание, редактирование и удаление досье

Чтобы создать досье новой персоны или дополнительное досье существующей в Системе персоны:

1. Перейдите в раздел **Персоны**.
2. Нажмите 

.
3. Заполните персональные данные:

Поле	Описание
Имя персоны	Имя персоны

Поле	Описание
Должность	Должность в организации
Отдел	Отдел, где работает персона
Описание	Дополнительное описание

4. Добавьте фотографию.
5. Добавьте контактные данные персоны: личный или рабочий номер телефона, адрес электронной почты, аккаунты в социальных сетях.
6. Нажмите + Создать досье. Новое досье будет доступно в списке досье сотрудников.

Чтобы внести изменения в досье:

1. Откройте нужное досье.
2. Нажмите ⋮ и выберите **Редактировать**.
3. Внесите изменения и нажмите **Сохранить**.

Чтобы удалить досье:

1. Откройте нужное досье.
2. Нажмите ⋮ и выберите **Удалить**.
3. Нажмите **ОК**. Досье будет удалено без возможности восстановления.

> **Важно!**
>
> Если вы удалите досье с контактом типа SID, при идентификации персоны по этому SID в полях **Владелец** и **Доступен для** будет отображаться имя персоны из удаленного досье. Чтобы это исправить, создайте новое досье и добавьте ему этот SID.
>
> Чтобы обновить значения полей **Владелец** и **Доступен для** выполните повторное сканирование.

11.5 Добавление контактов персоны

При синхронизации с Active Directory и InfoWatch Traffic Monitor становятся доступны контакты персон: номера телефона, адреса электронной почты, аккаунты в социальных сетях (при синхронизации с Traffic Monitor) и прочее. При этом пользователь может добавить дополнительные контакты вручную. Для этого:

1. Перейдите в досье персоны.
2. На вкладке **Персональная информация** на виджете **Контакты** нажмите ⚙.
3. В открывшемся окне будут отображены все автоматически добавленные при интеграции из других систем контакты персоны. При этом по умолчанию показаны по одному контакту типа: логин в домене, телефон, электронная почта, мобильный телефон.
   Нажмите **Добавить контакт**.

4. Укажите тип контакта (например, личный номер телефона) и введите его значение. Возможно добавить несколько позиций и оставить комментарий к контакту, для этого нажмите .

> **Примечание:**
>
> Для идентификации персоны, владеющей файлом или имеющей доступ к файлу, добавьте ей контакт типа SID.

> **Важно!**
>
> Не создавайте персон с одинаковым значением контакта типа SID. Наличие в Системе двух и более персон с одинаковыми SID приведет к некорректному определению владельца файла.

5. Нажмите **Сохранить**. Новый контакт доступен в досье. Наличие комментария к контакту отмечается значком . Чтобы прочитать комментарий, нажмите на него.

11.6 Изменение руководителя персоны

На вкладке **Персональная информация** на виджете **Руководитель** указан руководитель персоны. Данная информация добавляется при синхронизации с Active Directory и InfoWatch Traffic Monitor.

Офицер безопасности может изменить руководителя персоны, который был добавлен вручную. Руководитель персоны, загруженный из системы-источника, не редактируется.
Для изменения руководителя:

1. На виджете нажмите .
2. В открывшемся окне **Изменение руководителя** в поле *Добавлено вручную* введите имя нового руководителя из числа персон, зарегистрированных в Системе. Допускается добавить несколько имен.
3. Нажмите **Сохранить**. На виджете **Руководитель** отображаются указанные руководители.

Для перехода в досье руководителя щелкните левой клавишей мыши на его имени на виджете **Руководитель**.

11.7 Объединение досье

Вы можете объединить два или несколько досье в одно. При этом новое досье будет содержать всю информацию из досье, существовавших до объединения. Для этого:

1. Откройте досье, которое требуется объединить с другими.
2. Нажмите  и выберите **Объединить досье**.
3. В открывшемся окне в поле **Объединить с** выберите для объединения одно или несколько досье из списка.
4. Нажмите **Объединить**.

12 Раздел "Расследования"

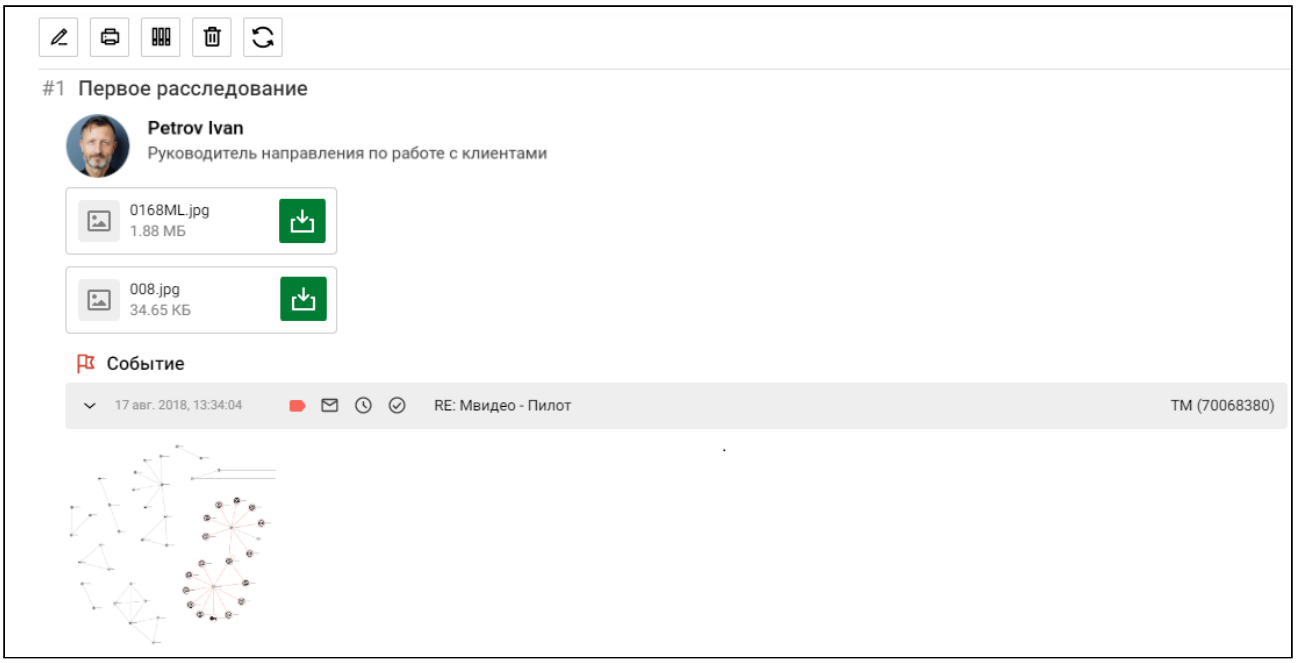
В этом разделе вы можете накапливать собранную из разных источников информацию и обобщать в расследования для последующего принятия решений или формирования отчетов. Сбор материалов и проведение расследования происходит следующим образом:

1. Система представляет факты нарушения политик в виде событий.
2. Офицер безопасности вручную добавляет такие события в расследование, при необходимости давая описания, прикрепляя файлы, снимки экрана, а также досье персон, подозреваемых в нарушении.
3. К расследованию подключаются другие Офицеры безопасности, вносят свои комментарии, прикладывают другие материалы, таким образом оперативно обмениваясь информацией.
4. По результатам расследования накопленная информация оформляется ими в виде отчета.

На панели навигации находится список всех открытых расследований и архив – база расследований, работа с которыми в данный момент не ведется.

На рабочей области отображается карточка выбранного расследования, которая может включать:

- ссылки на досье персон, имеющих отношение к расследованию;
- карточки события Traffic Monitor из раздела **События**;
- файлы-вложения различных форматов;
- графические файлы;
- текстовый комментарий в свободной форме.



Чтобы подробнее ознакомиться с приложенными материалами, можно:

- нажать на ссылку досье и перейти в раздел **Персоны**;
- нажать  и скачать на жесткий диск приложенный файл-вложение;
- раскрыть карточку события и ознакомиться с расширенной информацией по нему;
- просмотреть графические файлы.

Для удобства порядок следования материалов в карточке расследования можно менять. Для этого зажмите  и перетащите материал в нужное место. Чтобы удалить неактуальный материал из расследования, нажмите  напротив него и выберите **Удалить**.

Чтобы обновить состояние карточки расследования, нажмите .

Действия пользователя:

- Создание и архивирование расследования
- Добавление событий в расследование
- Добавление материалов в расследование
- Совместное проведение расследования
- Редактирование, выгрузка, печать и удаление расследования

12.1 Создание и архивирование расследования

Чтобы создать расследование:

1. На панели навигации слева нажмите .
2. В открывшемся окне введите наименование для нового расследования.
3. Нажмите **Сохранить**. Карточка нового расследования будет открыта для добавления материалов.

Чтобы в процессе работы над расследованием другой пользователь случайно не отредактировал данные, рекомендуется заархивировать расследование. Также следует архивировать расследование после его окончания.

Чтобы поместить расследование в архив:

1. Перейдите в нужное расследование на панели навигации.
2. Нажмите .
3. В открывшемся окне нажмите **Поместить в архив**. При этом расследование в архиве можно просмотреть без возможности внесения изменений.

Чтобы вернуться к работе над архивированным расследованием:

1. Выберите нужный архив.
2. Нажмите .
3. В открывшемся окне нажмите **Разархивировать**. После этого расследование станет доступным для изменений.

12.2 Добавление материалов в расследование

Для проведения расследования необходимо собрать материалы в как можно более полном объеме. Для этого:

1. Перейдите в карточку расследования на панели навигации.
2. В рабочей области можно добавить различные виды материалов в расследование:
   - **Текстовый комментарий** – для этого:
     a. Введите в строке комментарий в свободной форме.

b. Если необходимо, отформатируйте текст комментария с помощью встроенных стандартных инструментов. Они станут доступны по двойному клику указателем мыши на тексте комментария.
c. Нажмите **Enter**, чтобы сохранить комментарий.

- **Ссылку на досье персоны** – для этого:
  a. Нажмите .
  b. В открывшемся окне выберите нужное досье и нажмите **Применить.**
- **Файлы-вложения** – для этого:
  a. Нажмите .
  b. Выберите на диске компьютера один или несколько файлов любого формата и нажмите **Открыть**. Процесс может занять некоторое время и зависит от размера файлов. За один раз можно загрузить до 10 файлов общим размером до 100 МБ. Выбранные файлы будут прикреплены к расследованию. Просмотреть их можно после скачивания на жесткий диск любым из допущенных к расследованию Офицеров безопасности.
- **Графические файлы** – для этого:
  a. Нажмите .
  b. Выберите на жестком диске компьютера один или несколько графических файлов с расширениями: jpg, jpeg, png, webp, gif и нажмите **Открыть**. Файлы будут отображены в карточке в режиме предварительного просмотра.
- **Ссылку на событие Traffic Monitor** (подробнее см. "Добавление событий в расследование").

12.3 Совместное проведение расследования

Одной из основных возможностей для достижения высокой эффективности в расследованиях является коллективная работа над ними. Для этого реализовано совместное редактирование любого расследования. При этом:

- если первый ОБ открыл на просмотр расследование, и в этот же момент другой ОБ добавил новый материал, то первому ОБ придет сообщение, что расследование было изменено. При этом первый ОБ будет иметь возможность сохранить свои изменения поверх изменений второго ОБ, и все внесенные ими изменения будут объединены.
- если второй ОБ отредактировал и сохранил объект, который первый ОБ ранее удалил, то второй ОБ получит уведомление о невозможности сохранить изменения, так как удаленный материал уже недоступен.

12.4 Редактирование, выгрузка, печать и удаление расследования

Чтобы поменять название расследования:

1. Перейдите в карточку расследования.
2. Нажмите .
3. Введите новое название и нажмите **Сохранить**.

Чтобы выгрузить или распечатать карточку расследования:

1. Перейдите в карточку расследования.

2. Нажмите .
3. Выберите действие:
    - Распечатать PDF-документ. Информация, содержащаяся в карточке, будет распечатана с помощью стандартных средств браузера.
      или
    - Сохранить архив на жестком диске. Архив содержит PDF-документ с текстовым содержимым расследования и файлы вложений.

Чтобы удалить расследование:

1. Перейдите в карточку расследования.
2. Нажмите .
3. В открывшемся окне нажмите **Удалить**. Удалить расследование может только его автор.

13 Раздел "Отчеты"

В разделе доступны вкладки:

- Конструктор отчетов - где вы можете работать с шаблонами отчетов;
- Архив отчетов - где вы можете просмотреть список отчетов и скачать нужный;
- Задачи на регулярную отправку - где вы можете просмотреть список задач на регулярную отправку, приостановить отправку или удалить задачу.

13.1 Конструктор отчетов

Вам доступно создание шаблонов отчетов о персонах и нарушениях на основе данных, полученных из Traffic Monitor. Шаблоны содержат несколько настраиваемых разделов на выбор. Вы можете выгрузить отчет сразу после создания шаблона, настроить регулярную отправку отчета, а также сохранить шаблон, чтобы выгрузить отчет позже.

На вкладке не отображаются шаблоны, созданные другими пользователями.

> **Примечание:**
>
> В текущей версии данные Data Discovery не используются в конструкторе отчетов.

Чтобы создать новый шаблон отчета:

1. На вкладке **Конструктор шаблонов** нажмите  **Новый шаблон отчета**.
2. Укажите **Параметры отчета**:
   a. **Название**
   b. **Формат**: *XLSX, PDF* или *HTML*
   c. **Тип**:
      - для формата *XLSX*:
        - *Детализированный*
      - для форматов *PDF* и *HTML*:
        - *По нарушениям*
        - *По персонам*
3. При необходимости настройте **Фильтры**:
   - *Персона*
   - *Уровень нарушения*
   - *Дата*

   > **Примечание:**
   >
   > По умолчанию отображаются лишь три перечисленных фильтра. Вы можете настроить дополнительные фильтры, нажав  в правом верхнем углу секции **Фильтры**.
   >
   > Чтобы скрыть панель фильтров, нажмите .

4. Выберите разделы, которые будут добавлены в отчет, и настройте их отображение. См. статьи по типам шаблонов:

- "Отчет в формате XLSX"
- "Отчет по нарушениям в формате PDF или HTML"
- "Отчет по персоне в формате PDF или HTML"

5. Нажмите **Сохранить**. Карточка шаблона будет отображена на вкладке **Конструктор отчетов**.

Чтобы выгрузить отчет:

- на карточке шаблона нажмите **Выгрузить отчет**
или
- откройте шаблон, нажав на его карточку, и нажмите **Выгрузить** → **Выгрузить отчет**.

Чтобы настроить регулярную отправку отчета:

1. Откройте шаблон, нажав на его карточку.
2. Нажмите **Выгрузить** → **Настроить регулярную отправку**.
3. Укажите параметры регулярной отправки и нажмите **Сохранить**. Будет создана задача на регулярную отправку. Список задач вы можете просмотреть в разделе **Отчеты** на вкладке **Задачи на регулярную отправку**. Подробнее см. "Задачи на регулярную отправку".

> **Примечание:**
>
> Настройка регулярной отправки отчетов, созданных по настраиваемым шаблонам, аналогична настройке регулярной отправки отчета раздела **События**. См. "Просмотр и выгрузка данных событий".

Чтобы отредактировать шаблон:

1. Откройте шаблон, нажав на его карточку.
2. Внесите изменения.
3. Нажмите **Сохранить**.

Вы можете **дублировать шаблон**, чтобы создать на его основе новый. Для этого:

- на карточке шаблона нажмите **Дублировать шаблон отчета**
или
- откройте шаблон, нажав на его карточку, и нажмите **Дублировать**.

Для быстрого **поиска** требуемого **шаблона** введите часть его названия в поле **Поиск по названию** . Отобразятся все шаблоны, чьи названия содержат указанную последовательность символов.

Чтобы удалить шаблон:

- на карточке шаблона нажмите Удалить.
или
- откройте шаблон, нажав на его карточку и нажмите **Удалить**.

13.1.1 Отчет в формате XLSX

Вы можете добавить в шаблон отчета следующие разделы:

Раздел отчета	Описание
Настройки раздела	
Дополнительная информация	Дополнительная информация: дата и время формирования отчета, примененные фильтры, временной период, за который формируется отчет
Дата	Количество событий для каждой даты за выбранный период
Отправители/ получатели	Количество событий по уровням нарушений для каждого отправителя/ получателя. Для персон внутри периметра указаны отдел и должность
Отображаемых значений не более	Количество персон для отображения: *3*, *5*, *10* (по умолчанию), *20*
Сортировать по	Признак, по которому данные будут отсортированы: • *Общее количество событий* (по умолчанию) • *Уровень нарушения*
Типы событий	Количество событий для каждой категории и для каждого типа
Отображаемых значений не более	Количество персон для отображения: *3*, *5*, *10* (по умолчанию), *20*
Сортировать по	Признак, по которому данные будут отсортированы: • *Общее количество событий* (по умолчанию) • *Уровень нарушения*
Уровень нарушений	Количество событий для каждого уровня нарушения
Круг общения	Количество событий по уровням нарушений для каждого отправителя/ получателя на одной странице. Для персон внутри периметра указаны отдел и должность
Отображаемых значений не более	Количество персон для отображения: *3*, *5*, *10* (по умолчанию), *20*
Сортировать по	Признак, по которому данные будут отсортированы:

Раздел отчета	Описание
Настройки раздела	
	• *Общее количество событий* (по умолчанию) • *Уровень нарушения*

13.1.2 Отчет по нарушениям в формате PDF или HTML

Вы можете добавить в шаблон отчета следующие разделы:

Раздел отчета	Описание
Настройки раздела	
Титульный лист	Титульный лист, на котором указаны название, автор отчета, дата событий в отчете, общее количество страниц
Указать автора отчета	Указывать ли имя Офицера безопасности, сформировавшего отчет
Оглавление	Список разделов в отчете и их страницы
Динамика нарушений	Вы можете добавить в отчет график и/или таблицу, представляющие динамику нарушений за указанный период **Примечание:** Элемент можно добавить только при наличии доступа к функции *Детали нарушений.*
График	Количество событий каждого уровня нарушения по датам/часам согласно фильтрам
Детализация	Таблица, в которой для каждого временного периода указано количество событий каждого уровня нарушения
Отправители/ получатели	Вы можете добавить в отчет гистограмму и/или таблицу, представляющие число событий для наиболее активных отправителей/получателей **Примечание:** Если у Офицера безопасности нет доступа к функции *Детали нарушений*, то уровни нарушения не отображаются.
Отображаемых значений не более	Количество персон для отображения: *3*, *5*, *10* (по умолчанию), *20*

Раздел отчета		Описание
	Настройки раздела	
	Сортировать по	Признак, по которому данные будут отсортированы: • *Общее количество событий* (по умолчанию) • *Уровень нарушения*
	Топ отправителей/ получателей	Гистограмма, отображающая для каждой персоны суммарное количество событий, в которых персона была отправителем/получателем, и уровни нарушения
	Детализация	Таблица, отображающая для каждой персоны и для каждого уровня нарушения суммарное количество событий, в которых персона была отправителем/получателем
События по каналам		Вы можете добавить в отчет таблицы, представляющие суммарное количество событий для разных типов/категорий **Примечание:** Если у Офицера безопасности нет доступа к функции *Детали нарушений*, то уровни нарушения не отображаются.
	Отображаемых значений не более	Количество событий для отображения: *3*, *5*, *10* (по умолчанию), *20*
	Сортировать по	Признак, по которому данные будут отсортированы: • *Общее количество событий* (по умолчанию) • *Уровень нарушения*
	Категории событий	Таблица, отображающая суммарное количество событий для каждой категории событий
	Типы событий	Таблица, отображающая суммарное количество событий для каждого типа событий
Детализация категории событий		График и таблица, представляющие число событий для каждой выбранной категории **Примечание:** Если у Офицера безопасности нет доступа: • к функции *Политики, Объекты защиты и Периметры*, то нет возможности выгрузить подраздел *Нарушения по политикам*; • к функции *Детали нарушений:* • уровни нарушения не отображаются;

Раздел отчета	Описание
Настройки раздела	
	• нет возможности выгрузить график *Динамика нарушений* и таблицу с детализацией динамики.
Отображаемых значений не более	Количество событий для отображения: *3*, *5*, *10* (по умолчанию), *20*
Сортировать по	Признак, по которому данные будут отсортированы: • *Общее количество событий* (по умолчанию) • *Уровень нарушения*
Категории	Категории событий, для которых будет выгружена детализация: • *Почта* • *Мессенджер* • *Обмен файлами* • *Хранение файлов* • *Принтер и МФУ* • *Интернет-активность* • *Работа в приложениях* • *Файловые операции* • *Другие*
Состав	Наполнение раздела: • *График Динамика нарушений* • *Детализация динамики нарушений* • *Топ отправителей* • *Детализация топ отправителей*
События с нарушениями	Вы можете добавить в шаблон таблицу со всеми событиями с учетом фильтра в отчете, для которых решение пользователя – *Нарушение*. Отображаются первые 100 событий Для каждого события доступны следующе атрибуты: • *ID события* • *Источник события* (адрес сервера Traffic Monitor) • *Дата и время события* • *Уровень нарушения* • *Тип события* • *Решение пользователя* • *Вердикт* • *Отправители* • *Получатели* • *Описание*

Раздел отчета Настройки раздела	Описание
	• *Вложения* • *Политики* • *Объекты защиты* • *Детализация события* **Примечание:** Если у Офицера безопасности нет доступа: • к функции *Политики, Объекты защиты и Периметры*, то нет возможности выгрузить подраздел *Нарушения по политикам*; • к функции *Детали нарушений:* • уровни нарушения не отображаются; • нет возможности выгрузить график *Динамика нарушений* и таблицу с детализацией динамики.

13.1.3 Отчет по персоне в формате PDF или HTML

Вы можете добавить в шаблон отчета следующие разделы:

Раздел отчета Настройки раздела	Описание
Титульный лист	Титульный лист, на котором указаны название, автор отчета, дата событий в отчете, общее количество страниц
Указать автора отчета	Указывать ли имя Офицера безопасности, сформировавшего отчет
Справка о персоне	Раздел с личными данными и информацией о персоне: фото, имя, должность, стаж работы, контакты, отдел, имя руководителя, список расследований, связанных с персоной, а также комментарий к досье
Оглавление	Список разделов в отчете и их страницы
Круг общения	В разделе представлены: • Гистограмма и граф *Топ внутренних/внешних коммуникаций*, отображающие персон внутри/вне периметра, которые были участниками в общении с целевой персоной, и их досье • Гистограммы *Топ исходящих*, *Топ входящих*, отображающие персоны и внешние контакты, которые в общении с целевой персоной были получателями/отправителями

Раздел отчета		Описание
	Настройки раздела	
	Отображаемых значений не более	Количество событий для отображения: *3*, *5*, *10* (по умолчанию), *20*
	Сортировать по	Признак, по которому данные будут отсортированы: • Общее количество событий *(по умолчанию)* • Уровень нарушения
	Гистограмма	Добавить ли гистограммы в отчет
Нарушения		Вы можете добавить в отчет график и таблицу с детализацией нарушений. **Примечание:** Если у Офицера безопасности нет доступа к функции *Детали нарушений*, этот раздел не может быть добавлен.
	График	График *Динамика нарушений*, отображающий количество событий для каждого уровня нарушения по датам/часам, в которых данная персона была участником
	Детализация	Таблица *Детализация динамики*, отображающая уровни нарушений и количество событий для каждого уровня

13.2 Архив отчетов

В разделе представлена таблица с файлами всех отчетов, созданных в Центре расследований.

Для каждого отчета указаны:

- Название;
- Размер;
- Дата и время формирования.

Вы можете изменять ширину столбцов таблицы.

Чтобы выгрузить отчет, нажмите на его название. Файл отчета будет сохранен на вашем компьютере.

Таблица содержит последние 50 отчетов, которые хранятся 30 дней. Если количество файлов превысит 50, Система удалит файлы, созданные первыми.

13.3 Задачи на регулярную отправку

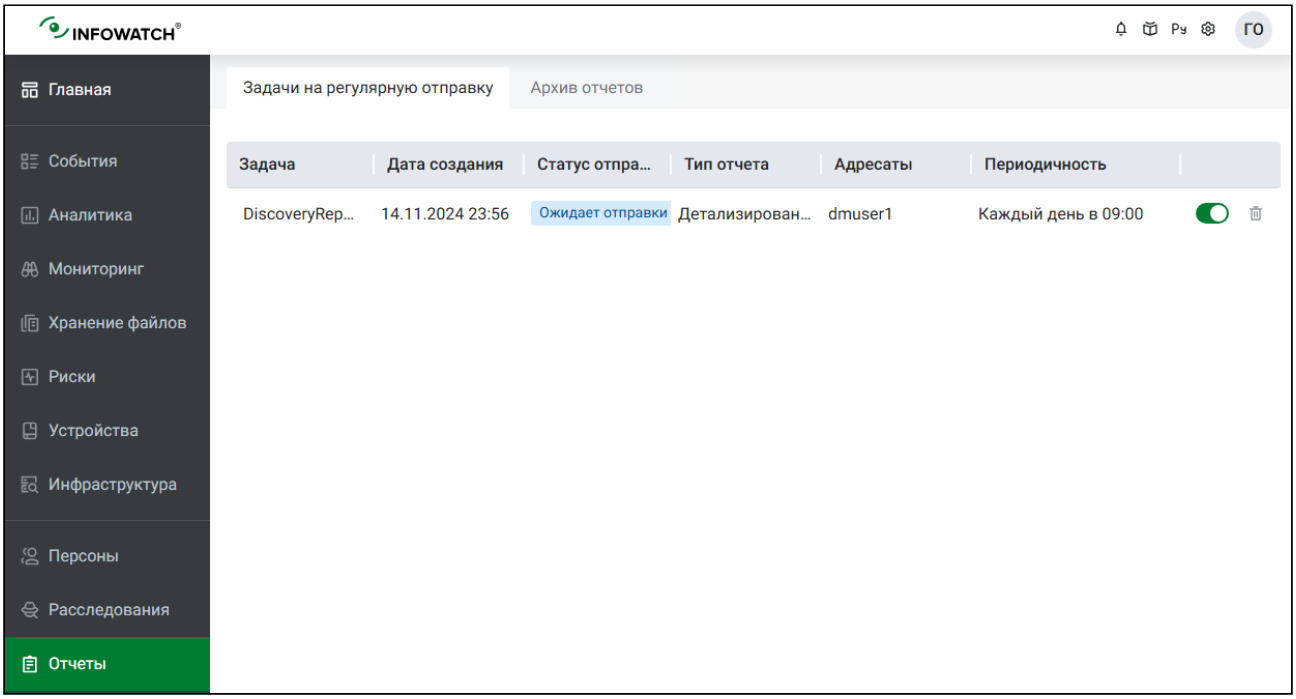
Список задач на регулярную отправку содержит следующие поля:

- **Задача** – имя задачи. Имя по умолчанию зависит от раздела или вкладки, откуда выгружается отчет. Имя задачи вы можете изменить;
- **Дата создания** – дата и время, когда был создана задача на отправку;
- **Статус отправки**;

- **Тип отчета** – детализированный, презентационный;

> Data Discovery формирует только детализированные отчеты.

- **Адресаты** – пользователи, которые получают отчет;
- **Периодичность** – дни недели и время, когда адресаты получают отчет;



В разделе вы можете:

- приостановить отправку отчета;
- удалить задачу из списка;
- изменить ширину столбцов таблицы;
- изменить наименование задачи.

Чтобы изменить имя задачи, нажмите на текущее название задачи и введите новое.

Чтобы отключить отправку отчета на какой-то период, воспользуйтесь переключателем .

Чтобы удалить задачу:

1. В строке задачи нажмите .
2. В открывшемся диалоговом окне нажмите ОК. Задача будет удалена.

14 ПОЛЬЗОВАТЕЛЬСКОЕ ЛИЦЕНЗИОННОЕ СОГЛАШЕНИЕ

ВНИМАНИЕ! Внимательно ознакомьтесь с условиями настоящего Лицензионного соглашения с конечным пользователем (далее – «Соглашение») перед началом работы с программным обеспечением. Установка и/или использование программного обеспечения означает ваше полное и безоговорочное согласие со всеми условиями настоящего Соглашения.

Если вы не согласны с условиями настоящего Соглашения, или если вы заключаете настоящее Соглашение от имени юридического лица, но у вас нет таких полномочий, вы должны прервать установку и/или прекратить использование программного обеспечения и удалить его копии.

1. Предоставление лицензии

1.1. Настоящее Лицензионное соглашение с конечным пользователем является договором присоединения, заключаемым между вами и ООО «Лаборатория ИнфоВотч», ИНН 7734583888 (далее – «Правообладатель»).

1.2. По настоящему Соглашению вам на условиях личной, не подлежащей передаче неисключительной лицензии предоставляются права на использование программного обеспечения Правообладателя (далее — «ПО»), а именно права загрузки, установки, резервного копирования и использования программного обеспечения по функциональному назначению в соответствии с условиями, указанными в Документации к ПО (Руководстве Пользователя, Руководстве Администратора, Руководстве по установке), настоящем Соглашении и Договоре, заключенном между вами и вашим лицензиаром. Документация к ПО признается частью настоящего Соглашения.

1.3. ПО предназначено только для использования в вашей текущей операционной деятельности. В случае если вы получили, загрузили и/или установили ПО, предназначенное для ознакомительных целей, вы имеете право использовать ПО только в целях ознакомления и только в течение ознакомительного периода. Любое использование ПО для других целей или по завершении ознакомительного периода запрещено.

1.4. Если вы используете ПО разных версий или версии ПО для разных языков, если вы получили ПО на нескольких носителях, если вы иным способом получили несколько копий ПО или получили ПО в составе пакета другого программного обеспечения, то общее число используемых вами лицензий не должно превышать их количества, определенного Договором между вами и вашим лицензиаром.

1.5. Вы имеете право изготовить копию ПО при условии, что эта копия предназначена только для архивных целей или для замены правомерно приобретенного экземпляра в случаях, когда оригинал утерян, уничтожен или стал непригоден для использования. Такая копия не может быть использована для иных целей и должна быть уничтожена в случае, если владение экземпляром ПО перестанет быть правомерным.

1.6. Вы самостоятельно несете ответственность и обеспечиваете соблюдение применимого экспортного и импортного законодательства, а также применимых торговых санкций и эмбарго в отношении передачи прав и использования ПО.

2. Условия использования

2.1. Программное обеспечение предназначено исключительно для использования в целях анализа и контроля информационных потоков в корпоративной среде, в том числе, для выявления и предотвращения случаев несанкционированного использования конфиденциальных данных.

2.2. Вы обязаны использовать ПО исключительно в соответствии с применимыми к вам требованиями законодательства Российской Федерации и иных стран, которые, не ограничиваясь, могут включать:

- законодательство о персональных данных;
- законодательство о коммерческой тайне;
- законодательство о противодействии неправомерному использованию инсайдерской информации;

- законодательство о противодействии коррупции;
- законодательство о критической информационной инфраструктуре;
- законодательство о государственных информационных системах;
- законодательство о защите данных в Национальной платежной системе;
- нормы Трудового кодекса РФ;
- а также иные применимые источники права.

2.3. Вы не вправе использовать ПО для любых целей или любым способом, нарушающим применимое законодательство, настоящее Соглашение и/или права третьих лиц. Правообладатель не несет ответственность за нарушение вами условий настоящего пункта.

2.4. Если вы не являетесь резидентом РФ, вы обязаны использовать ПО в соответствии с применимым правом, включая законы и нормативные акты, регламентирующие защиту персональных данных и коммерческой тайны.

2.5. Во избежание нарушения вами статей 138, 138.1 УК РФ, статьи 23 Конституции РФ и иных правовых норм перед началом использования программного обеспечения вам рекомендовано письменно проинформировать сотрудников о том, что:

- на рабочей станции используется данное ПО, о функциональности данного ПО и о данных, которые могут быть получены в результате применения данного ПО;
- рабочая станция и используемые каналы связи (электронная почта, локальная корпоративная сеть, телефония и IP-телефония и прочее) являются вашей собственностью и передаются им во временное пользование для выполнения служебных обязанностей;
- вся информация, созданная на рабочей станции, является вашей собственностью и подлежит контролю и логированию на законных основаниях.

2.6. Обращаем ваше внимание, что вам необходимо получить согласие лиц, данные которых будут обрабатываться при использовании ПО, на соответствующий сбор, хранение, использование, обработку, распространение и иную обработку указанных данных.

2.7. Напоминаем, что установка программы на компьютере другого лица без его ведома, получение личной информации третьих лиц без их ведома, нарушение неприкосновенности частной жизни, нарушение тайны переписки, телефонных переговоров, телеграфных и иных видов сообщений является уголовно наказуемым деянием и преследуется в соответствии с действующим законодательством Российской Федерации.

2.8. Правообладатель не имеет доступа к данным, созданным и/или полученным в рамках использования вами ПО. Вы имеете исключительный доступ к указанным данным. С учетом этого Правообладатель не осуществляет сбор, хранение, использование и распространение указанных данных.

2.9. Вы самостоятельно несете риск того, что данные, созданные и/или полученные в рамках использования вами ПО, могут содержать данные неэтического и/или противоправного характера, компьютерные вирусы или программы (или ссылки на них), способные прервать или нарушить нормальную функциональность компьютерного оборудования, программного обеспечения или средств телекоммуникации любых лиц.

2.10. Вам не разрешается:

- целиком или частично редактировать, изменять, адаптировать, переводить, декомпилировать, дизассемблировать, осуществлять инженерный анализ, создавать или воссоздавать исходный код, модернизировать, исправлять явные ошибки, модифицировать ПО и/или его компоненты, осуществлять попытки предпринять аналогичные действия в отношении ПО и/или его компонентов, а также создавать производные работы, основанные на ПО и/или Документации;
- использовать ПО сверх лицензируемого количества учетных записей, предусмотренных договором с вашим непосредственным лицензиаром;

- осуществлять интеграцию ПО в сторонние сайты и программные продукты, за исключением случаев, когда такая интеграция прямо разрешена Правообладателем и правообладателем сторонних сайтов и программных продуктов. Допускается интеграция ПО в сторонние программные продукты посредством встроенного модуля API в рамках функциональных возможностей указанного модуля;
- удалять или изменять уведомления об авторских правах, правах собственности на ПО и/или его элементы;
- использовать ПО с целью извлечения прибыли, сдавать ПО в аренду, прокат или во временное пользование, предоставлять подписку на ПО, предоставлять услуги аутсорсинга с использованием ПО, сублицензировать, распространять или иным образом предоставлять любому физическому или юридическому лицу любые права в отношении ПО и/или Документации, разглашать результаты стендовых испытаний ПО, за исключением случаев, прямо разрешенных настоящим Соглашением либо письменным согласием Правообладателя;
- передавать и предоставлять ссылку на дистрибутив и/или доступ к лицензионному ключу третьим лицам в нарушение положений настоящего Соглашения и Договора, заключенного между вами и вашим лицензиаром;
- использовать ПО с целью создания данных или кода, предназначенных для использования другими программными продуктами;
- самостоятельно устанавливать дополнительные приложения к ПО, заменять компоненты ПО;
- осуществлять незаконный сбор, систематизацию, хранение и распространение конфиденциальной и/или персональной информации;
- использовать ПО для несанкционированного удаленного доступа к компьютерам третьих лиц, а также изменения конфигурации компьютера так, чтобы такие возможности могли появиться у стороннего ПО;
- собирать, хранить, использовать, обрабатывать и распространять с помощью ПО любые информационные данные, которые нарушают применимое законодательство, права и законные интересы третьих лиц.

2.11. Лицензионный ключ является конфиденциальной информацией, его несанкционированная передача третьим лицам запрещена. Правообладатель оставляет за собой право использовать средства для проверки подлинности установленного у вас лицензионного ключа.

3. Ограничение гарантий и ответственности Правообладателя

3.1. Программное обеспечение предоставляется «КАК ЕСТЬ» («AS IS») в соответствии с общепринятым в международной практике обычаем делового оборота. За проблемы, возникающие в процессе эксплуатации ПО, в том числе: проблемы совместимости с другими программными продуктами (пакетами, драйверами и др.), проблемы, возникающие из-за неоднозначного толкования Документации в отношении ПО, несоответствия результатов использования ПО вашим ожиданиям и т.п.) и иные, Правообладатель ответственности не несет.

3.2. Вы подтверждаете, что вам известны функциональные свойства ПО, в отношении которых предоставляется право использования, и самостоятельно несете риск соответствия ПО своим индивидуальным пожеланиям и потребностям, а также за результаты, полученные с его помощью.

3.3. Правообладатель не предоставляет никаких гарантий, выраженных или подразумеваемых, и настоящим отказывается от всех гарантий.

Правообладатель не гарантирует, что ПО не содержит ошибок, будет работать непрерывно, быстро, без неточностей, и не несет никакой ответственности за прямые или косвенные последствия использования ПО, в том числе возникшие из-за возможных ошибок в ПО или опечаток в ПО и/или прилагаемой или переданной позднее Документации.

Правообладатель не дает никаких гарантий и не несет никакой ответственности перед вами в случае любых изменений в программном обеспечении третьих лиц, произошедшее после установки/ внедрения ПО и повлекшее потерю функциональности ПО (включая, но не ограничиваясь, изменением протокола передачи данных, формата хранения данных, логике работы стороннего программного обеспечения, обновлением программного обеспечения, которое перестает поддерживать работу с ПО). Правообладатель не дает никаких гарантий, условий, представлений или положений (выражаемых в явной или в подразумеваемой форме) на все, включая без ограничений нарушения прав третьих лиц, коммерческое качество, интеграцию или пригодность для определенных целей.

3.4. Данные, которые становятся доступными при использовании ПО, включая файлы, ссылки, изображения и текст, являются исключительной ответственностью физического или юридического лица, от которого они были получены, и являются собственностью соответствующего лица. Настоящее Соглашение не дает никаких прав на такие данные. Вы принимаете на себя ответственность за возможные последствия, риски безопасности и любые убытки, вызванные такими данными, просмотренными или полученными Вами с помощью ПО.

3.5. Правообладатель не несет ответственности за какие-либо убытки, ущерб, независимо от причин его возникновения (включая, но не ограничиваясь прямым или косвенным ущербом, убытками связанными с недополученной прибылью, прерыванием коммерческой или производственной деятельности, утратой деловой информации, небрежностью, или какими-либо иными убытками), возникшие вследствие использования или невозможности использования ПО, даже если Правообладатель был уведомлен о возможных убытках. При любых обстоятельствах ответственность Правообладателя ограничена возмещением документально подтвержденного реального ущерба, ограниченного 3 000 (тремя тысячами) рублей, и возлагается на него при наличии в его действиях вины.

4. Права на объекты интеллектуальной собственности

4.1. Вы соглашаетесь с тем, что исключительные права на любые объекты интеллектуальной собственности, воплощенные в ПО и /или любой предоставленной вам документации, принадлежат Правообладателю. По настоящему Соглашению ПО не передается в собственность, не продается, а предоставляются права его использования на условиях ограниченной неисключительной лицензии. Ничто в данном Соглашении не предоставляет вам никаких прав на указанные объекты интеллектуальной собственности иные, чем предоставленные вам по Договору, заключенному между вами и вашим лицензиаром.

4.2. Вы соглашаетесь с тем, что исходный код, лицензионный ключ для ПО являются собственностью Правообладателя.

5. Права на информацию, доступ к которой получен вами в рамках осуществления настоящего Соглашения

5.1. Вы соглашаетесь с тем, что вам не принадлежат никакие права на любую информацию, не являющуюся объектом интеллектуальной собственности в соответствии с разделом 4, доступ к которой получен вами в рамках осуществления настоящего Соглашения.

5.2. К указанной информации, включая, но не ограничиваясь, относятся системы, методы работы, иная информация.

5.3. Указанная выше информация будет использоваться Вами только в целях осуществления прав на ПО, предоставленных вам по настоящему Соглашению и Договору с лицензиаром, без права использования указанной информации в собственных интересах и за пределами Соглашения и Договора, заключенного между вами и вашим лицензиаром.

6. Открытое программное обеспечение

Вы проинформированы **о том, что ПО содержит открытое программное обеспечение, распространяемое под определенными лицензиями, с которыми вы можете ознакомиться в файле licenses.inf (license.info), распространяемом с ПО в составе дистрибутива.**

Вы соглашаетесь с тем, что входящие в состав ПО программные компоненты внешних поставщиков предоставляются без гарантий доступа к обновленным версиям компонент и к иным программам, оборудованию, сертификатам на обучение или услугам внешних поставщиков.

7. Заключительные положения

7.1. Настоящее Соглашение вступает в силу с начала установки ПО и действует на протяжении всего периода использования.

7.2. В случае нарушения вами какого-либо из условий Соглашения и/или Договора, заключенного между вами и вашим лицензиаром, или в случае, если исполнение Соглашения или Договора, заключенного между вами и вашим лицензиаром будет нарушать или нарушает применимое право, Правообладатель вправе отказаться от исполнения Соглашения и расторгнуть его (включая, но не ограничиваясь, путем блокировки лицензионного ключа, прекращения предоставления прав) в любое время без уведомления вас. В указанном в настоящем пункте случае стоимость ПО/размер вознаграждения за права на ПО, а также возможные убытки по настоящему Соглашению и/или Договору, заключенному между вами и вашим лицензиаром, вам не возмещаются.

7.3. При расторжении или прекращении Соглашения Пользователь обязан уничтожить все копии установленного ПО.

7.4. Вы настоящим уведомляетесь о том, что при включении в ПО функции сбора статистики использования, ПО может осуществлять сбор и запись информации в обезличенном виде:

а) об аппаратном обеспечении, об операционной системе устройства, на котором установлено и эксплуатируется ПО, а также иная техническая информация, содержащая сведения об использовании аппаратных средств, нагрузке на аппаратные средства и быстродействии ПО;

б) данные о действиях конечного пользователя в пользовательском интерфейсе ПО.

Правообладатель вправе использовать указанные данные для принятия решений о развитии функциональности и улучшении эксплуатационных характеристик ПО.

Правообладатель обязуется обеспечить конфиденциальность полученных данных, не разглашать и не передавать их третьим лицам.

7.5. Настоящее Соглашение, а также любые споры, возникающие из или в связи с настоящим Соглашением, регулируются и истолковываются в соответствии с законодательством Российской Федерации.

7.6. Все споры и разногласия, возникающие в процессе исполнения Соглашения, будут по возможности разрешаться путем переговоров. В случае недостижения согласия, споры будут рассматриваться в судебном порядке в Арбитражном суде г. Москвы.

7.7. Настоящее Соглашение может быть изменено Правообладателем в любой момент без какого-либо уведомления. Новая редакция Соглашения вступает в силу с момента ее размещения по указанному в настоящем пункте адресу. С условиями действующей редакции Соглашения вы всегда можете ознакомиться по адресу https://kb.infowatch.com/display/EULA/. При этом продолжение использования ПО после внесения изменений и/или дополнений в настоящее Соглашение означает согласие с такими изменениями и/или дополнениями. В случае несогласия с условиями такого измененного Соглашения, вы обязаны немедленно отказаться от дальнейшего использования ПО.

8. Контактная информация Правообладателя

Тел./факс: +7(495)229-00-22

Коммерческий департамент: sales@infowatch.com

Служба технической поддержки: support@infowatch.com

Веб-сайт: www.infowatch.ru

15 Глоссарий

Термин	Определение
Администратор	Администратор – пользователь Системы, выполняющий установку, настройку и поддерживающий работу Системы. См. также: Офицер безопасности, Пользователь
Группа SMB	Параметр, позволяющий использовать Систему как часть определенной рабочей группы для сетевого управления папками, расположенными на рабочих станциях, терминальных серверах, файловых хранилищах и т.д. (например, домен)
Задача сканирования	Операция проверки ресурсов (путей сканирования) на наличие конфиденциальных данных. Задача содержит описание ресурсов сканирования, загрузки и передачи файлов с них в Traffic Monitor. Включает в себя список хостов, список путей сканирования, фильтр по именам сканируемых файлов, фильтр по размеру сканируемых файлов, фильтр по директориям и расписание для автоматического запуска и остановки задачи
Зона	Группа серверов Data Discovery. В зону могут входить несколько серверов из кластера Системы. Каждая зона имеет отдельные настройки для экспорта данных в Traffic Monitor
Интерфейс пользователя	Совокупность средств и методов, при помощи которых пользователь взаимодействует с Системой
Информация о файле	Информация о файле включает в себя метаданные файла и данные о сканировании файла/ресурса. К данным о сканировании файла/ресурса относятся: • название задачи; • дата обнаружения файла; • IP-адрес хоста и др. См. также: Метаданные файла
Консоль управления	Графический интерфейс пользователя, предназначенный для управления системой Data Discovery (администрирование и настройка Системы, работа с задачами сканирования)
Лицензия	Право на использование Системы. Получается при приобретении Системы и определяет допустимые действия с Системой

Термин	Определение
Маска	Шаблон поиска – метод описания поискового запроса с использованием метасимволов. Маски используются для фильтра по именам сканируемых файлов и для фильтра по директориям
Метаданные файла	Данные о файле, которые Data Discovery получает из файловой системы ресурса: • Полный путь к файлу; • Имя файла; • Размер файла; • Дата и время создания файла; • Дата и время изменения файла; • Владелец файла; • Пользователи/группы, которые имеют доступ к файлу. См. также: Информация о файле
Офицер безопасности	Основной пользователь Консоли управления. Также – предустановленная роль пользователя Консоли управления, имеющая привилегии на все действия в Системе, за исключением административных
План выполнения задачи	Формируемый сервисом Execution Controller порядок действий для сервиса сканера (Puller). Определяет последовательность обработки ресурсов
Пользователь	Пользователь системы Data Discovery – администратор, офицер безопасности и др. См. также: Администратор, Офицер безопасности
Платформа	Принципиальная конструкция продукта – комплекс частей, подсистем, интерфейсов и производственных процессов. Продуктовая Платформа выступает в качестве базы для развертывания семейства продуктов и сборки готовых изделий из стандартизированных компонентов
Проход	Полное сканирование одного ресурса (т.е. все файлы на ресурсе должны быть обработаны сервисом сканера (Puller)). Проход может состоять из нескольких сессий
Рабочая станция/Компьютер	В терминах Системы подразумевается контролируемая рабочая станция или терминальное устройство
Расписание	Расписание для автоматического запуска и остановки задачи. Доступно ежедневное и еженедельное расписание

Термин	Определение
Ресурс (Resource)	Ресурс хоста, который содержит файлы. Система сканирует файлы на ресурсах и загружает файлы во временное хранилище. Для подключения к ресурсу необходимо указать путь сканирования хоста
Сессия	Подключение к ресурсу в рамках прохода. Длительность сессии сканирования (параметр **Длительность обработки**) задается в расширенных настройках задачи
Сервис ClickHouse	Распределенная аналитическая столбцовая СУБД, предназначенная для проведения анализа по хранящимся данным в реальном времени. Компонент продуктовой Платформы
Сервис PostgreSQL	СУБД, которая используется в продукте Data Discovery для хранения настроек и конфигураций задач, а также статистики о задачах. Компонент продуктовой Платформы
Фильтры папок	Набор условий для обработки директорий на ресурсах. Представляет из себя список директорий, включенных в обработку, и список директорий, исключенных из обработки
Фильтры файлов	Набор условий в настройках задачи для сканирования и загрузки файлов с ресурсов. Представляет из себя список масок для фильтрации по именам файлов, а также фильтры по максимальному и минимальному размеру файлов
Хост (Host)	Рабочая станция или сервер, к которому осуществляется подключение для сканирования файлов и загрузки с него файлов для дальнейшего анализа. Содержит список ресурсов для сканирования
Временное хранилище (Storage)	Временное место хранения скачанных файлов. Файл хранится здесь до отправки в Traffic Monitor
Экспорт данных в Traffic Monitor	Описание подключения к Traffic Monitor для отправки скачанных с ресурсов файлов и их дальнейшего анализа. Указывает на конкретную инсталляцию Traffic Monitor, требуется токен авторизации и список адресов XAPI. В меню Консоли управления Системы относится к разделу **Зоны и серверы**
Яндекс Диск	Облачное хранилище файлов, часть экосистемы Яндекс 360
Alfresco Community Edition	Система управления корпоративным контентом (ECM)

Термин	Определение
CASI	Сервис для передачи извлеченных текстовых контентов в Модуль контентного анализа
Contents Processor	Сервис передает данные в Ripper и уведомляет Data Analysis Service. Пара экземпляров сервисов Ripper и Contents Processor - экстрактор. Экстракторы извлекают текстовые данные из файлов и уведомляют Data Analysis Service о наличии данных для обработки
Discoperator	Сервис представляет из себя бэкенд и используется для работы Консоли управления.
FTP	Протокол передачи данных по сети
InfoWatch Data Analysis Service	Data Analysis Service предназначен для потоковой постобработки и анализа данных, предоставляемых продуктами компании InfoWatch. Для Data Discovery задействован модуль протокластеризации текстовых данных
InfoWatch Traffic Monitor	Traffic Monitor - программный комплекс, предназначенный для контроля различных видов трафика (SMTP, IMAP, POP3, HTTP, HTTPS, IMAP, XMPP, ICQ, NRPC) и теневых копий данных, копируемых на съемные носители и отправляемых на печать
Kubernetes	Программное обеспечение для оркестрации контейнеризированных приложений – автоматизации их развертывания, масштабирования и координации в условиях кластера
NATS	Система обмена сообщениями для компонентов
Nextcloud	Набор клиент-серверных программ для создания и использования хранилища данных
Puller	Сервис для сканирования ресурсов и загрузки файлов с них. Экземпляры сервиса Puller – сканеры. Сканеры собирают метаданные файлов и загружают файлы с ресурсов во временное хранилище
Ripper	Сервис для извлечения текстовых контентов из файлов. Пара экземпляров сервисов Ripper и Contents Processor - экстрактор. Экстракторы извлекают текстовые данные из файлов и уведомляют Data Analysis Service о наличии данных для обработки
Sender	Сервис для отправки загруженных файлов в Traffic Monitor. Экземпляры сервиса Sender – передатчики. Передатчики отправляют файлы из временного хранилища в Traffic Monitor

Термин	Определение
	для анализа. После отправки файла в Traffic Monitor передатчик удаляет файл из временного хранилища
SMB (Server Message Block)	Сетевой протокол для удаленного доступа к файлам, принтерам и другим сетевым ресурсам, а также для межпроцессного взаимодействия
SSH (Secure Shell)	Сетевой протокол, позволяющий производить удаленное управление операционной системой и туннелирование TCP-соединений (например, для передачи файлов)
Task Execution Controller	Сервис для управления выполнением задачи. Контролирует работу над задачей всех остальных сервисов
TPS Agent	Сервис для мониторинга состояния временного хранилища. Помогает работе с временным хранилищем сервисам Puller и Sender, предоставляет статистику по хранилищу
WebDAV	Расширение протокола HTTP, позволяющее пользователям управлять файлами, хранящимися на удаленных серверах
XAPI	Приватный API для передачи файлов в Traffic Monitor