INDICA Manuals

INDICA

May 31, 2023

ADMINISTRATOR DOCUMENTATION

1 System Requirements **3**

2 Installation **7**
- 2.1 Prerequisites . . . 7
- 2.2 Installation . . . 7

3 Activation **9**

4 Admin Page Overview **11**
- 4.1 Services . . . 12
- 4.2 Indexer Progress . . . 13
- 4.3 Spider Status . . . 15
- 4.4 Modules . . . 15
- 4.5 System Status . . . 17

5 System Settings **21**
- 5.1 Users . . . 22
  - 5.1.1 Users List . . . 23
  - 5.1.2 Creating a user . . . 24
  - 5.1.3 Editing a user . . . 27
  - 5.1.4 Deleting a user . . . 29
  - 5.1.5 Resetting a user's password . . . 29
  - 5.1.6 Resetting the admin password . . . 30
- 5.2 Roles and Permissions . . . 31
  - 5.2.1 Introduction to Roles and Permissions . . . 31
  - 5.2.2 Configuring Roles . . . 33
  - 5.2.3 Default permissions matrix . . . 37
- 5.3 Cases . . . 38
  - 5.3.1 Cases List . . . 39
  - 5.3.2 Adding a case . . . 40
  - 5.3.3 Editing a case . . . 42
  - 5.3.4 Archiving a case . . . 43
  - 5.3.5 Restoring an archived case . . . 44
  - 5.3.6 Deleting a case . . . 45
- 5.4 Networking . . . 45
  - 5.4.1 Schedule bandwith throttling . . . 46
  - 5.4.2 Network interface . . . 48
  - 5.4.3 Mailer settings . . . 50
  - 5.4.4 Proxy settings . . . 50
  - 5.4.5 Zabbix monitoring . . . 51

5.4.6 SSL Settings . . . 52
5.5 Security . . . 56
5.5.1 Enabling Active Directory integration . . . 56
5.5.2 Enabling Single Sign-on . . . 58
5.6 Software & Support . . . 60
5.6.1 Activation Key . . . 61
5.6.2 Software version . . . 61
5.6.3 Updates . . . 63
5.6.4 Remote Connection . . . 63
5.7 General Settings . . . 64
5.8 Backup Settings . . . 65
5.9 Look and Feel . . . 65

6 Datasources Settings **69**
6.1 Archive Settings . . . 69
6.1.1 Archive Location . . . 70
6.1.2 Archive Settings . . . 71

7 Manager Page Overview **73**
7.1 Services . . . 74
7.2 Progress Status . . . 76
7.3 Checks & Balances . . . 76
7.4 Index Stats . . . 77
7.5 System Status . . . 78
7.6 Workflow . . . 79
7.7 Tag Info . . . 80
7.8 Shares . . . 81
7.9 File Distribution Info . . . 82
7.10 File Processing Info . . . 84

8 Case Manager Exporting **87**
8.1 Create Export . . . 88
8.1.1 Creating a PDF Export . . . 89
8.2 Downloading an export . . . 89

9 Case Manager Inventory Exports **91**

10 Case Manager Metadata Upload **93**

11 Case Manager Path Reindex **95**

12 Case Manager Saved Queries **97**

13 Case Manager Scheduled Queries **101**
13.1 Scheduled Query Table . . . 102
13.2 Adding Scheduled Query . . . 102

14 Case Manager Tags **105**
14.1 Introduction . . . 105
14.2 Tags . . . 106
14.2.1 General Information . . . 106
14.2.2 Adding Tags . . . 107
14.2.3 Re-ordering and Nesting Tags . . . 107
14.2.4 Removing Tags . . . 107
14.2.5 Locking the Tags . . . 108

14.2.6 Untagging a Tag . . . 108
14.2.7 View Tag Audit Log . . . 109
14.3 Export . . . 109
14.4 Tag by CSV . . . 110

15 Case Manager Workflow **115**
15.1 Create workflow . . . 116
15.2 State options . . . 116
15.3 Visual representation workflow . . . 117

16 Reporting **119**
16.1 Report rules . . . 119
16.2 Reports (Report groups) . . . 120

17 Case Settings **123**
17.1 General Settings . . . 124
17.1.1 General Settings . . . 125
17.1.2 OCR Settings . . . 127
17.1.3 Inventory Listings . . . 128
17.2 Display Settings . . . 128
17.2.1 Facet Settings . . . 128
17.2.2 Result list Settings . . . 130
17.2.3 Index Progress Notification Settings . . . 131
17.2.4 Detail View Settings . . . 132
17.2.5 List Style Settings . . . 133
17.2.6 Redaction Feature Settings (Advanced PDF Viewer) . . . 133
17.3 Search Settings . . . 134
17.3.1 Stopword List . . . 134
17.3.2 Synonym List . . . 134
17.3.3 Boosting . . . 134
17.4 API Settings . . . 135
17.4.1 Polling API . . . 135
17.4.2 PowerBI Connection . . . 138
17.4.3 Polling API Examples . . . 147

18 GDPR Data Inventory Manage **155**
18.1 IT Assets . . . 156
18.2 Data Inventory . . . 156
18.3 Business Rules . . . 157

19 GDPR Settings Manage **159**
19.1 GDPR Settings . . . 160
19.2 GDPR Patterns . . . 161
19.3 GDPR Threat Matrix . . . 162
19.4 GDPR regular vs. sensitive personal information . . . 163
19.5 Bag of Words . . . 163

20 Look & Feel **165**
20.1 Logo . . . 166
20.2 Fonts, Interface & Text colors . . . 166
20.3 GDPR fonts and colors . . . 167

21 Data Sources **169**
21.1 Collections . . . 171
21.1.1 Adding a collection . . . 171

21.1.2 Editing a collection . . . 172
21.1.3 Deleting a collection . . . 172
21.2 Datatabs . . . 173
21.2.1 DataTabs List . . . 174
21.2.2 DataTabs Order & Security . . . 180
21.3 E-mail . . . 181
21.3.1 E-mail integration . . . 181
21.3.2 Map email to user . . . 182
21.3.3 EWS Accounts . . . 183
21.3.4 POP3 Settings . . . 183
21.4 Fileshares . . . 184
21.4.1 Adding a fileshare . . . 185
21.4.2 Editing a fileshare . . . 187
21.4.3 Deleting a fileshare . . . 189
21.5 Local . . . 189
21.5.1 Adding a local source . . . 189
21.5.2 Editing a local source . . . 191
21.5.3 Deleting a local source . . . 191
21.6 Queries . . . 192
21.6.1 Adding Queries . . . 192
21.6.2 Editing Queries . . . 194
21.6.3 Deleting Queries . . . 195
21.7 S3 Buckets . . . 195
21.7.1 Adding S3 Buckets . . . 195
21.7.2 Editing S3 Buckets . . . 197
21.7.3 Deleting S3 Buckets . . . 198
21.8 SharePoint . . . 198
21.8.1 General Settings . . . 198
21.8.2 Adding a sharepoint source . . . 199
21.8.3 Editing a sharepoint source . . . 207
21.8.4 Deleting a sharepoint source . . . 207
21.9 Box . . . 207
21.9.1 Adding a box source . . . 208
21.9.2 Editing a box source . . . 209
21.9.3 Deleting a box source . . . 209

22 Manager Setup **211**
22.1 Export . . . 212
22.2 Import . . . 212

23 Login and Credentials **215**
23.1 Two-Factor Authentication . . . 216

24 User Settings **219**

25 Switching Cases **221**

26 Filtering and Daterange **223**
26.1 Sidebar . . . 223
26.1.1 Expanding and collapsing the sidebar . . . 225
26.2 Selecting multiple filters in one category . . . 227
26.3 Daterange . . . 227

27 GDPR Dashboarding **229**
27.1 Overview Page . . . 229

27.1.1 GDPR Items Graph . . . 231
27.1.2 GDPR Sources Graph . . . 231
27.1.3 GDPR Buildup Over Time Graph . . . 232
27.1.4 GDPR Progress Graph . . . 233
27.2 Landscape Page . . . 233
27.2.1 GDPR Risk Radar Graph . . . 234
27.2.2 GDPR User Group Access To Data Sources Graph . . . 235
27.2.3 GDPR Items Per Data Source Graph . . . 236
27.2.4 GDPR User Group Access To GDPR Items Graph . . . 237
27.3 Issues Page . . . 238
27.3.1 Privacy Issues List . . . 239
27.3.2 Total Progress Graph . . . 240
27.3.3 Progress over time Graph . . . 240
27.4 Data Inventory Page . . . 241

28 Previewer **243**
28.1 Breadcrumbs . . . 243
28.2 Switching Views . . . 243
28.3 Data Lineage (Tab) . . . 244
28.4 PDF Controls . . . 245
28.4.1 Advanced PDF Viewer . . . 246
28.5 Document Actions . . . 247
28.5.1 Advanced PDF Viewer . . . 249
28.6 Redacting . . . 249
28.7 Redacting (Advanced PDF Viewer) . . . 250
28.8 Date-changer . . . 252
28.9 Page Deletion (Advanced PDF Viewer) . . . 253
28.10 Searching in documents . . . 253

29 Querying **255**
29.1 Writing queries . . . 255
29.1.1 Search By . . . 256
29.1.2 Exact match . . . 256
29.1.3 Terms and Phrases . . . 256
29.1.4 Boolean Operators . . . 256
29.1.5 Grouping . . . 257
29.1.6 Range Searches . . . 257
29.1.7 Wildcards . . . 258
29.1.8 Fuzzy Searches . . . 258
29.1.9 Word Distance Searches . . . 258
29.1.10 Boosting terms . . . 258
29.2 Advanced query builder . . . 259
29.3 Escaping special characters . . . 261

30 Query to CSV **263**
30.1 Exporting to CSV from the front-end . . . 263
30.2 Exporting to CSV and XLS(X) . . . 266

31 Saved Queries **267**

32 Selecting and de-selecting **269**
32.1 Select documents on this page . . . 270
32.2 Select all query results . . . 271
32.3 Deselect all . . . 272

33 Sorting **273**

34 Table & List View **275**
34.1 List view . . . 275
34.2 Table view . . . 276
34.3 Table view (with flipped table) . . . 276

35 Tagging **277**
35.1 Selecting items for tagging . . . 277
35.2 Tagging selected items . . . 278
35.3 Tagging individual items . . . 279

36 Visuals **281**
36.1 Location . . . 282
36.2 Per Country . . . 282
36.3 Per type . . . 283
36.4 Email . . . 284
36.5 NLP Graph . . . 285

37 Workflow **287**
37.1 States . . . 288
37.2 Buckets . . . 288
37.3 Moving documents . . . 288

38 Contacting INDICA **289**

Welcome to the >INDICA documentation.

This documentation can be used as a reference guide when working with INDICA. Different documentation is available depending on the version of INDICA you're working with. You can select documentation for the version you're using in the bottom left corner.

The manual is divided into 3 different sections:

- User Documentation
- Manager Documentation
- Administrator Documentation

For the best understanding of all the options, settings, and features it is advised to read all sections.

CHAPTER

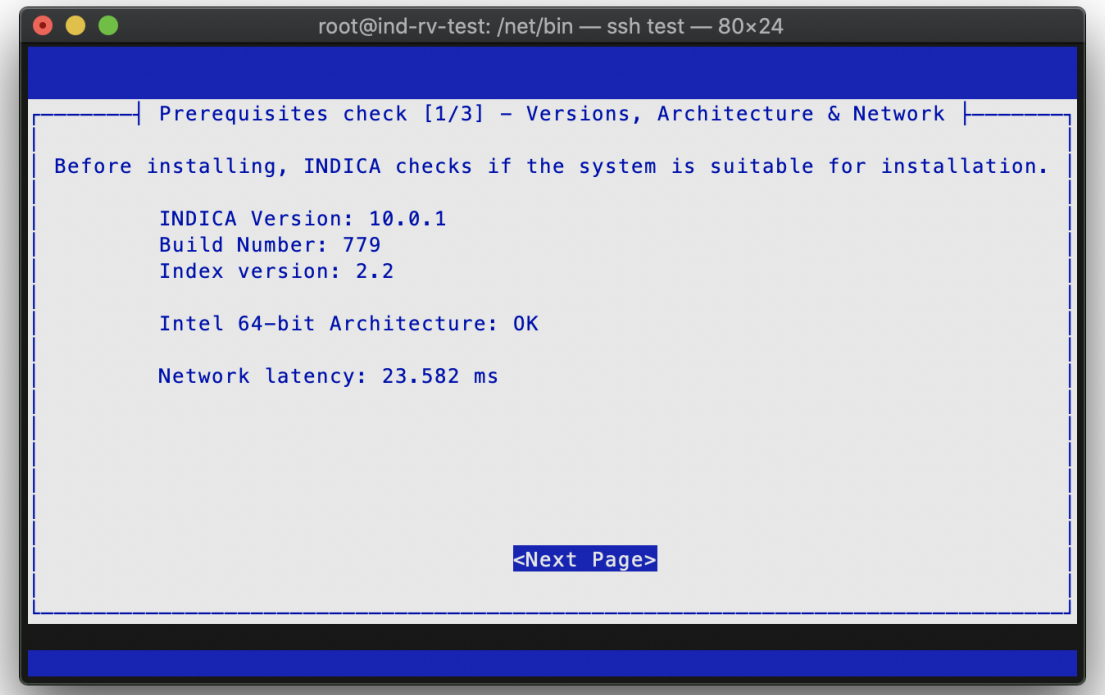
ONE

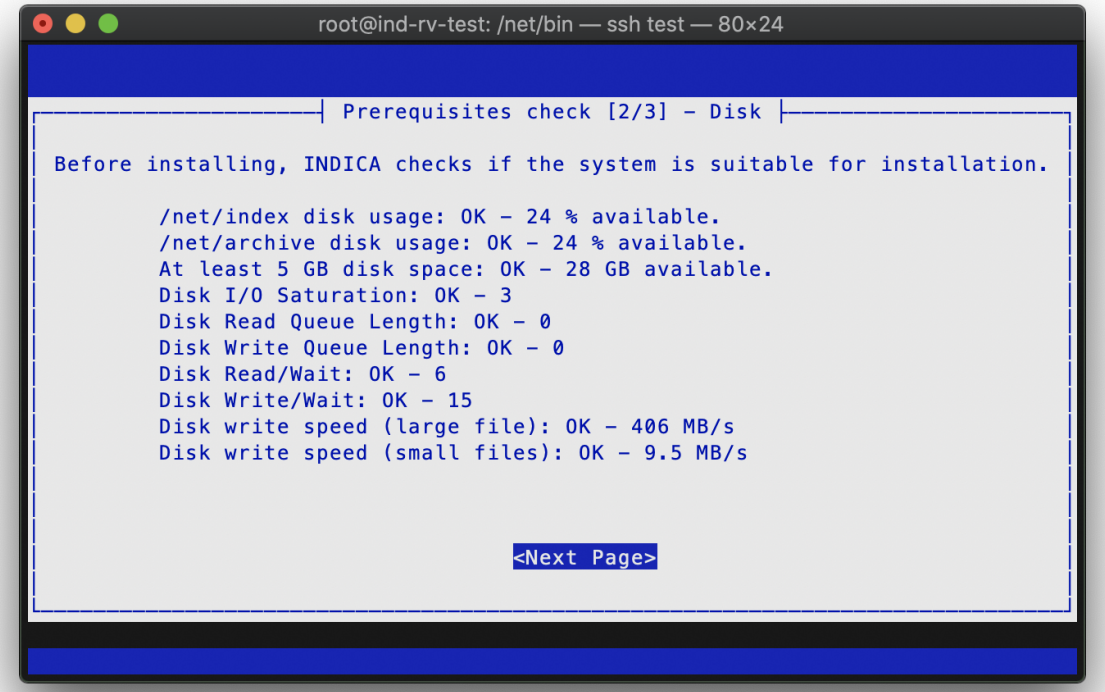
SYSTEM REQUIREMENTS

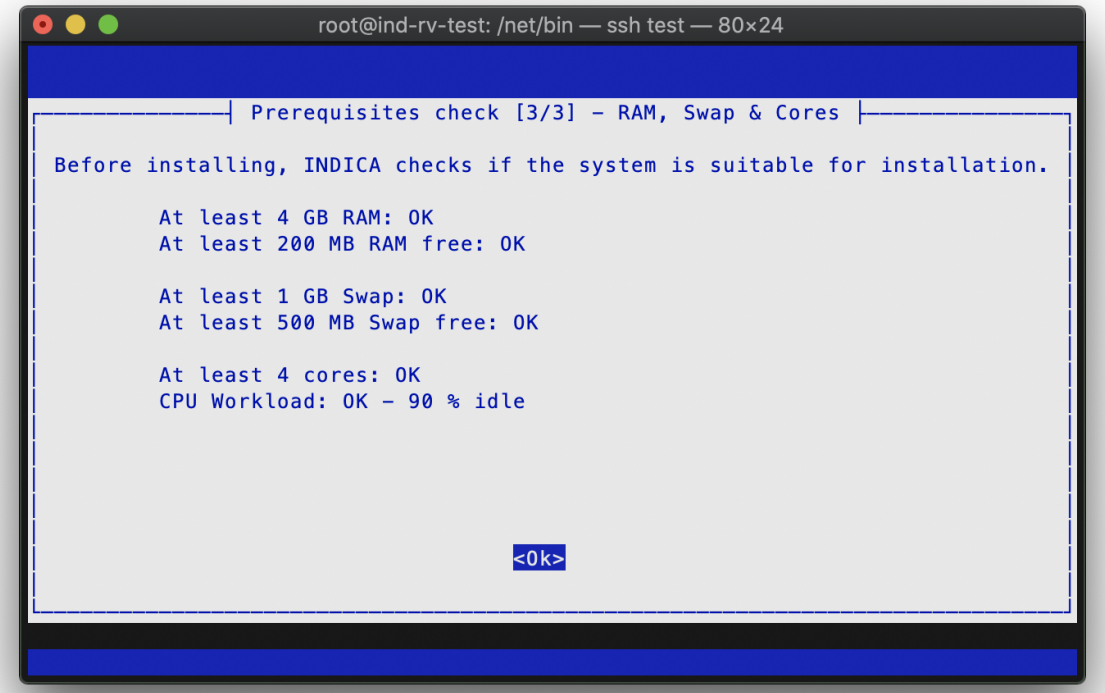
Before installing INDICA, please make sure that the system that will run INDICA meets the minimal specifications. Please refer to our knowledge base to see the requirements. Also keep in mind that the requirements are dependant on the expected workload of the system. INDICA can help you scope and calculate this.

Prior to installation, INDICA will check if the system is suitable for installation.

Note: Please note that the displayed requirements in the screen captures are not the actual system requirements. The displayed requirements are the bare minimum for a tiny INDICA test instance, not an actual production server.




root@ind-rv-test: /net/bin — ssh test — 80×24
Prerequisites check [2/3] - Disk
Before installing, INDICA checks if the system is suitable for installation.
/net/index disk usage: OK - 24 % available.
/net/archive disk usage: OK - 24 % available.
At least 5 GB disk space: OK - 28 GB available.
Disk I/O Saturation: OK - 3
Disk Read Queue Length: OK - 0
Disk Write Queue Length: OK - 0
Disk Read/Wait: OK - 6
Disk Write/Wait: OK - 15
Disk write speed (large file): OK - 406 MB/s
Disk write speed (small files): OK - 9.5 MB/s
<Next Page>


root@ind-rv-test: /net/bin — ssh test — 80×24
Prerequisites check [3/3] - RAM, Swap & Cores
Before installing, INDICA checks if the system is suitable for installation.
At least 4 GB RAM: OK
At least 200 MB RAM free: OK
At least 1 GB Swap: OK
At least 500 MB Swap free: OK
At least 4 cores: OK
CPU Workload: OK - 90 % idle
<Ok>

CHAPTER

TWO

INSTALLATION

2.1 Prerequisites

Before installing can begin, please make sure the following is in place:

1. A Ubuntu server with the requested OS version is in place;
2. The server meets the minimum system requirements;
   Please refer to System Requirements for more information.
3. The machine has an active network connection and is capable of reaching the internet;
4. The INDICA installable archive is available, either locally or by download link.

2.2 Installation

Installing and setting up INDICA is very easy. INDICA comes pre-packaged with almost everything it needs to be installed. For ease of use, an installation script is provided. Some manual steps need to be taken before the installation script can be used.

1. Gain SSH access to the target machine.
   This can be done with command-line SSH, PuTTY, or equivalent.
2. Elevate your rights to root-level access.
   This can be done using the command `sudo su` and typing in the password.
3. Create the appliance directory called "net" in the root of the filesystem "/"
   Use the command `mkdir /net`
4. Navigate to the newly created directory.
   Use the command `cd /net` to do this.
5. Transfer the INDICA .tar.gz file to this directory.
   This can be done by transferring with scp, winSCP, or the download link provided by INDICA.
   Make sure the file is in `/net`.
6. Extract the downloaded package using "tar -zxvf [filename]".
7. Run the installation script: "/net/bin/install.sh".

The install script will ask some questions and guide you through the installation.

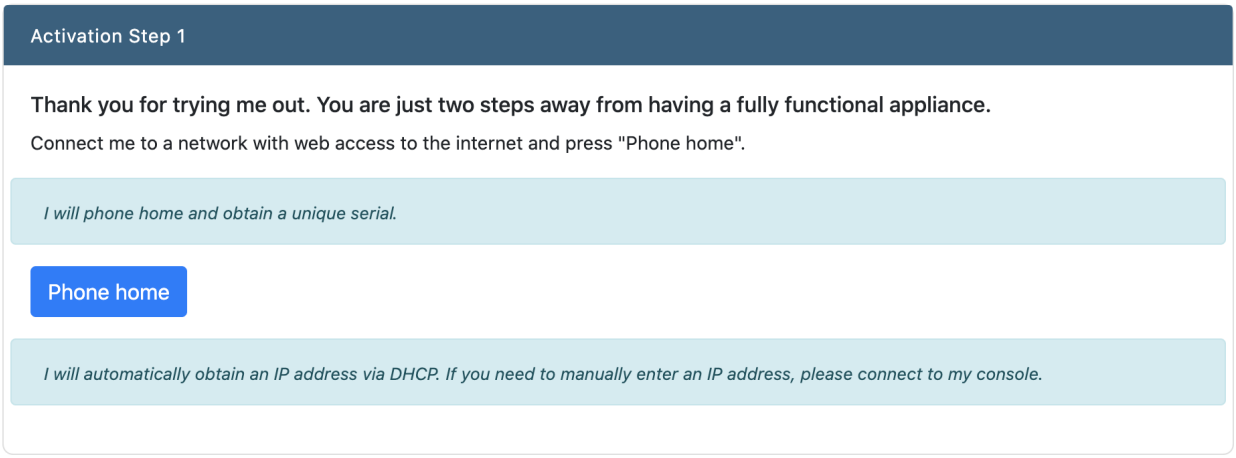
CHAPTER

THREE

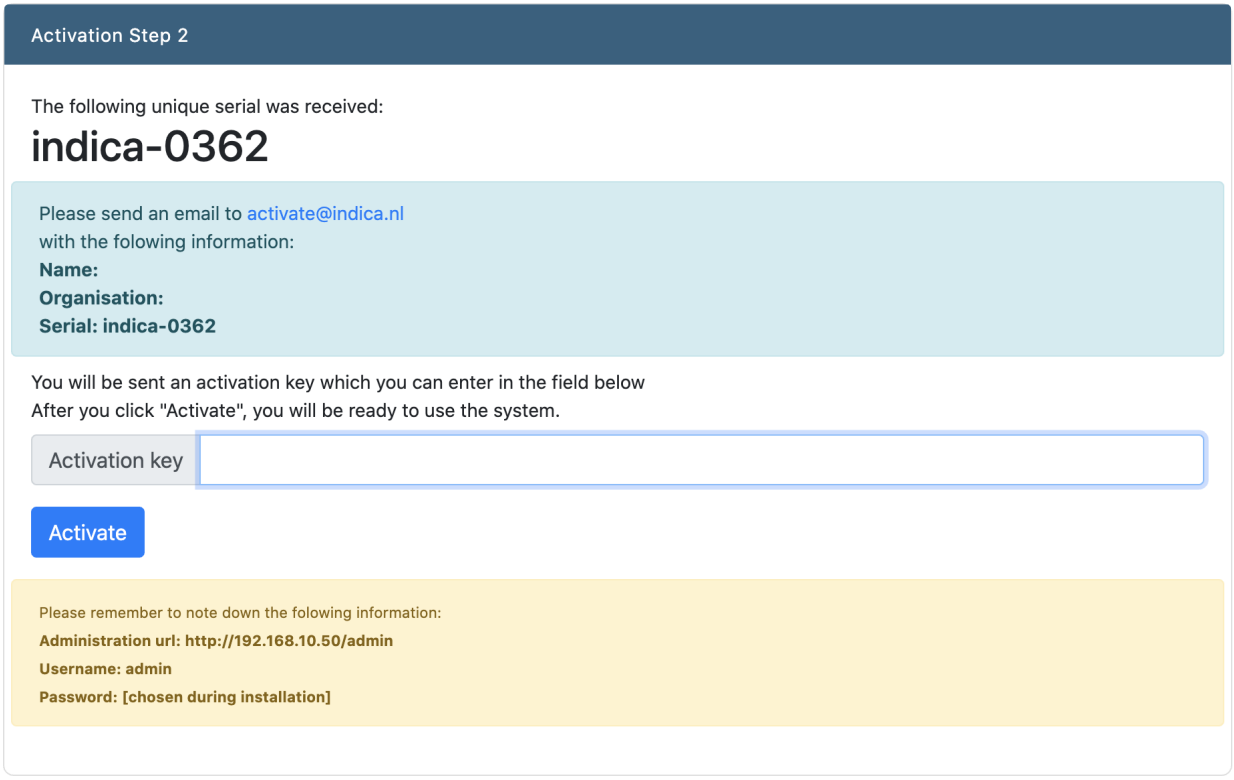
ACTIVATION

When installation is done, please connect to the web interface of your appliance. This can be found on `http://[ip]`. The IP address here is the appliance IP. The IP of the appliance is shown in the console.

When connected, you will be presented with the welcome page:



Simply click the “Phone Home” button. Your appliance will connect to the INDICA API and retrieve an appliance serial. When the serial has been recieved, you will be redirected to the second step in the activation process:



When this step is presented, please send an email (or create a support ticket) to INDICA Support containing the information displayed on the screen. INDICA will then provide you with a unique activation key to activate your appliance.

CHAPTER

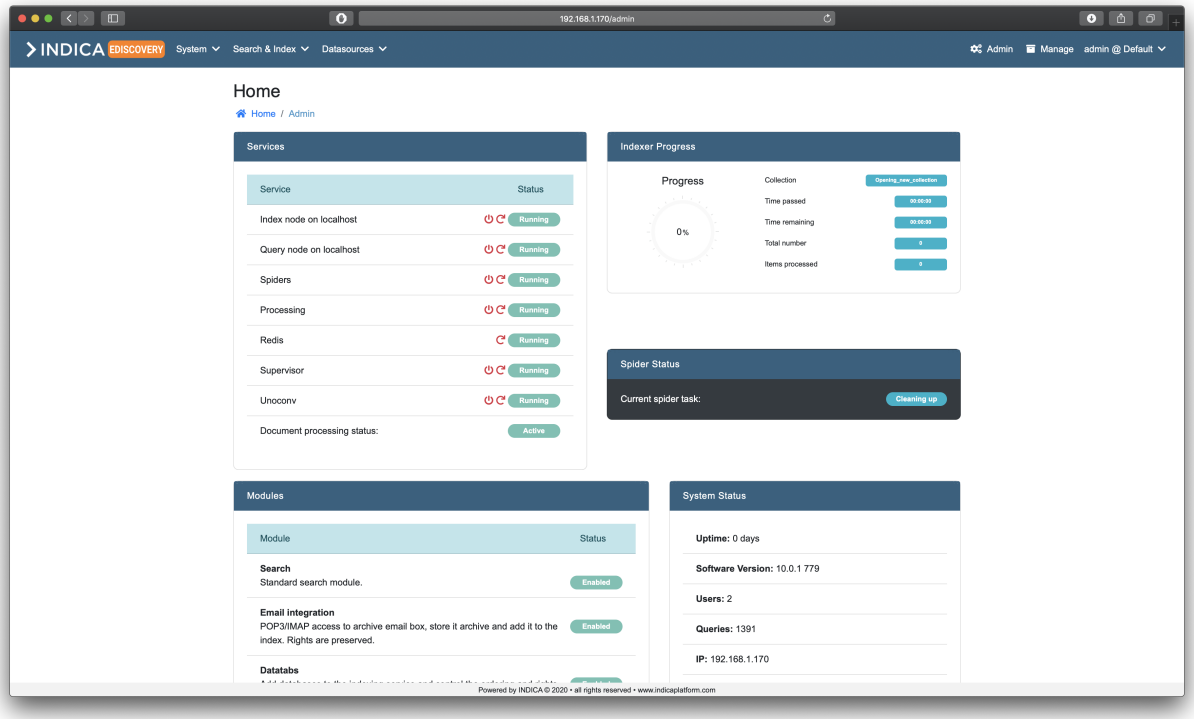
FOUR

ADMIN PAGE OVERVIEW

Table of Contents

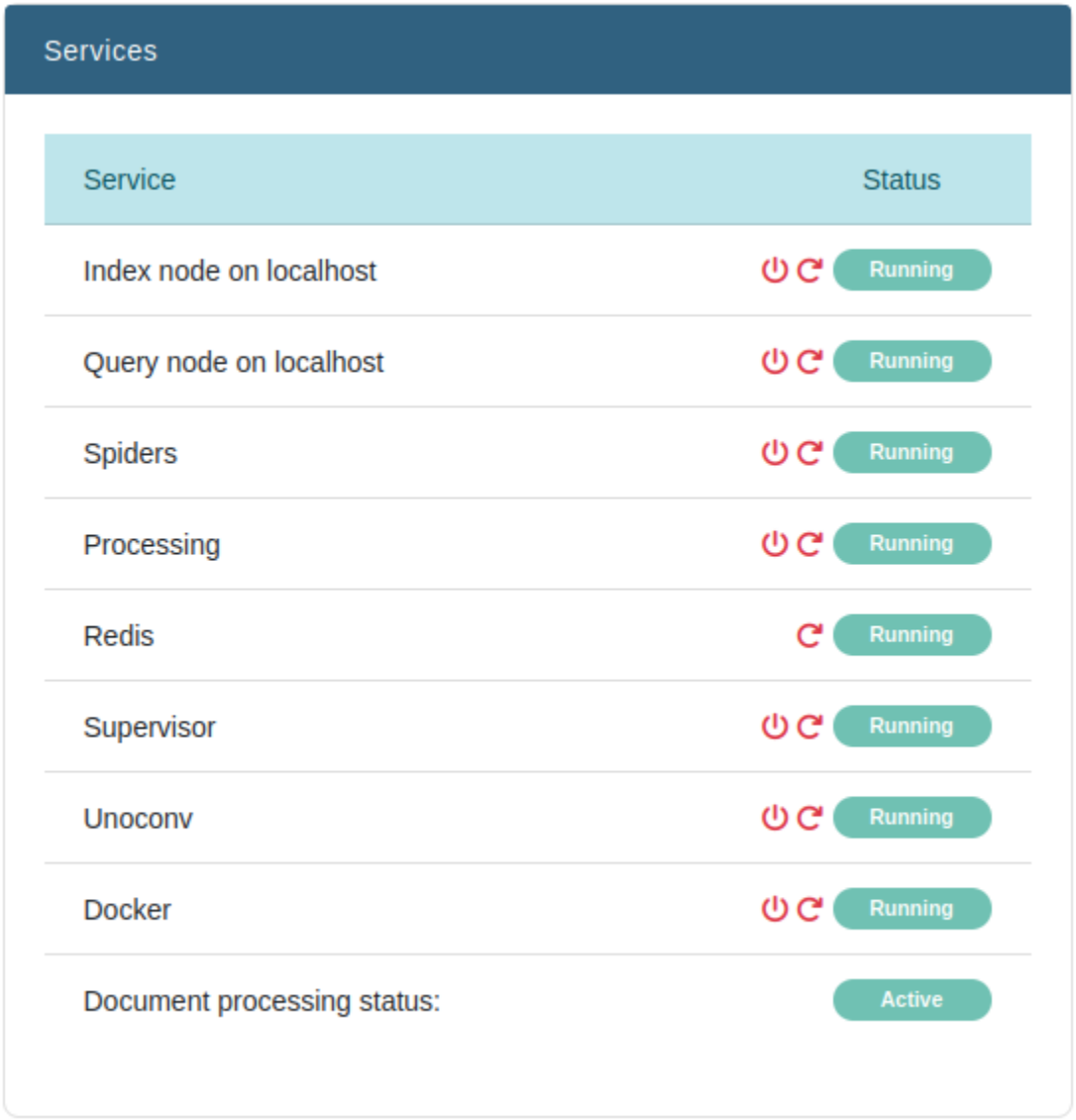
- *Admin Page Overview*
  - *Services*
  - *Indexer Progress*
  - *Spider Status*
  - *Modules*
  - *System Status*

This page shows the status of the system services, the indexer and spider status, system information and the enabled modules.



4.1 Services

In this panel a number of system services and their status is shown. On a single-node install, all services are running on the same machine. With cloud setups, it's possible to host some services on different nodes.



Most services can be stopped, started or restarted from this overview.

- **Index node**

  This displays the status of the Solr service on the Index node. When stopped, the index is not hosted anymore.
  The index and query node can be different depending on the setup.

- **Query node**

  This displays the status of the Solr service on the Query node. When stopped, the index is not hosted anymore.

The index and query node can be different depending on the setup.

- **Spiders**

  This displays the status of the Spider service. The spider is responsible for discovering new and/or changed data in the connected sources. Stopping the spider here will disable the spider for all cases on the appliance. The spider can also be disabled per-case in the “Cases” overview.

- **Processing**

  This displays the status of the Apache Tika service. Apache Tika is responsible for parsing document content to plain text for storing in the index.

- **Redis**

  Redis is needed to make the web front-end work. When this is disabled, most parts of the front-end will not work. Therefore, it can only be restarted.

- **Supervisor**

  Supervisor is needed to enable running of jobs on basis of a queue. Without supervisor, some functions will not work properly.

- **Unoconv**

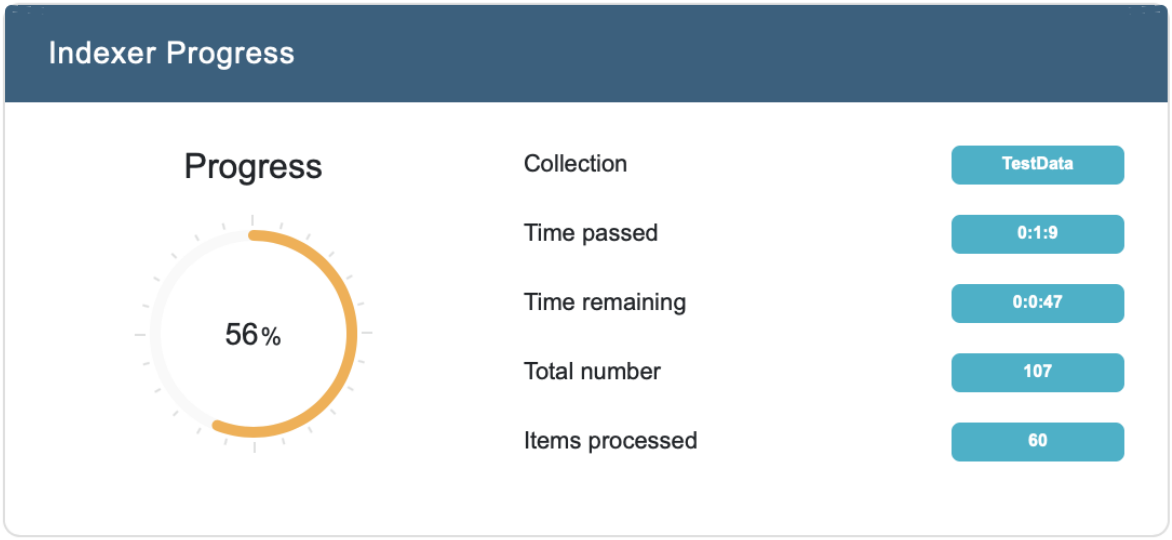
  Unoconv is responsible for converting eligible documents to PDF format so they can be previewed properly and nicely.

- **Document processing status**

  This displays the overall status of the document processing.

4.2 Indexer Progress

The Indexer Progress will display the general status of the Indexer.



It displays the following statistics:

- **Progress**

  Indexing progress of the current collection.

- **Collection**

  Name of the collection that is currently being indexed.

- **Time passed**

  Time passed since starting indexation on this collection.

- **Time remaining**

  Estimate on the time needed to fully index the current collection.

- **Total number**

  Number of items in this collection that need to be indexed.

- **Items processed**

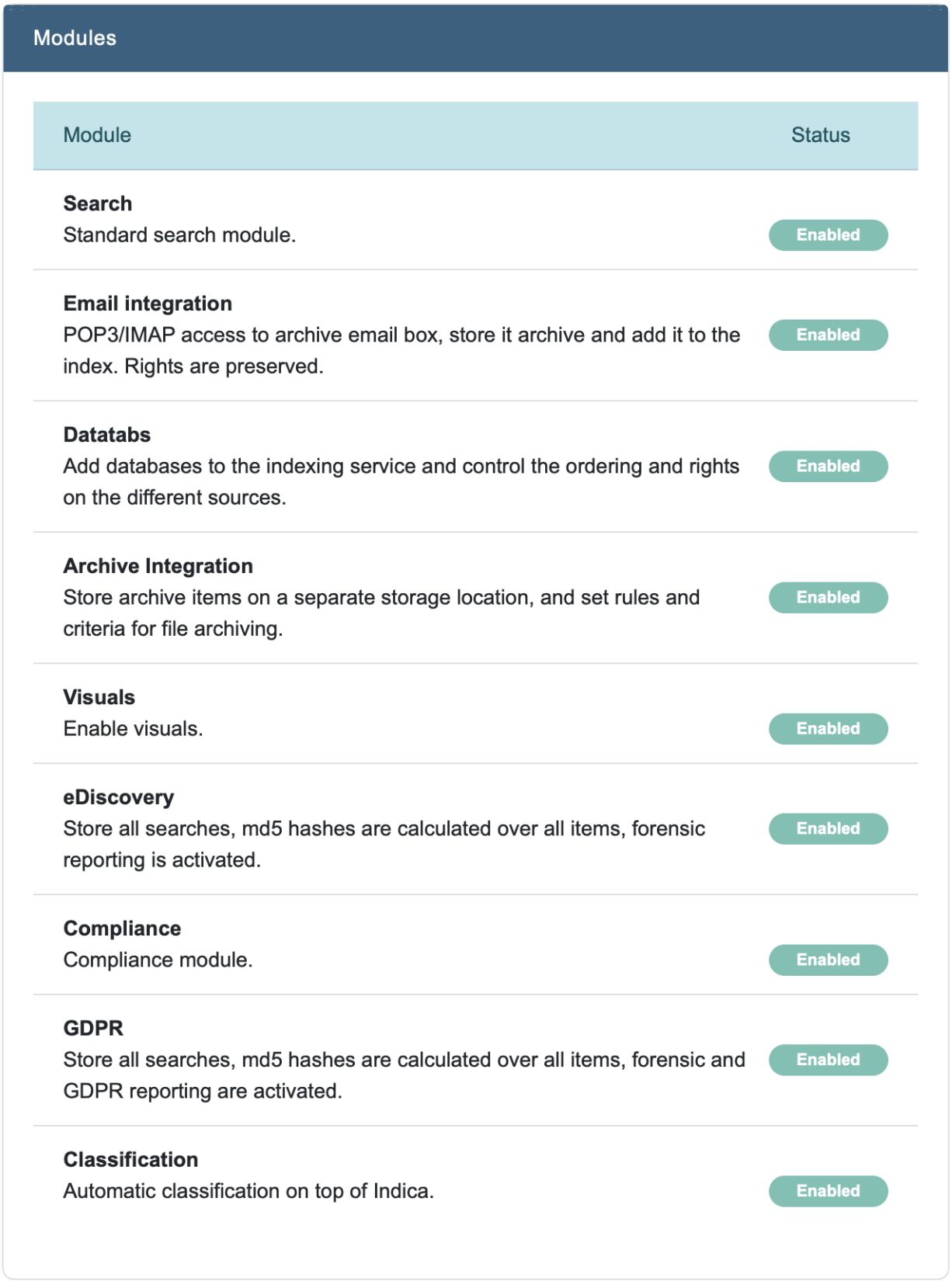
  Number of items in this collection that have been indexed.

4.3 Spider Status

Shows the current status of the spider.

4.4 Modules

Modules show what modules are enabled for your appliance. This is dependent on the license key you received during activation.



The following modules are available:

- **Search**

  Standard Enterprise Search module.

- **Email integration**

  POP3/IMAP access to archive email box, store it archive and add it to the index. Rights are preserved.

- **Datatabs**

  Add databases to the indexing service and control the ordering and rights on the different sources.

- **Archive Integration**

  Enabled store archive items on a separate storage location, and set rules and criteria for file archiving.

- **Visuals**

  Enable visuals on the search page.

- **eDiscovery**

  Store all searches, md5 hashes are calculated over all items, forensic reporting is activated.

- **Compliance**

  Compliance module.

- **GDPR**

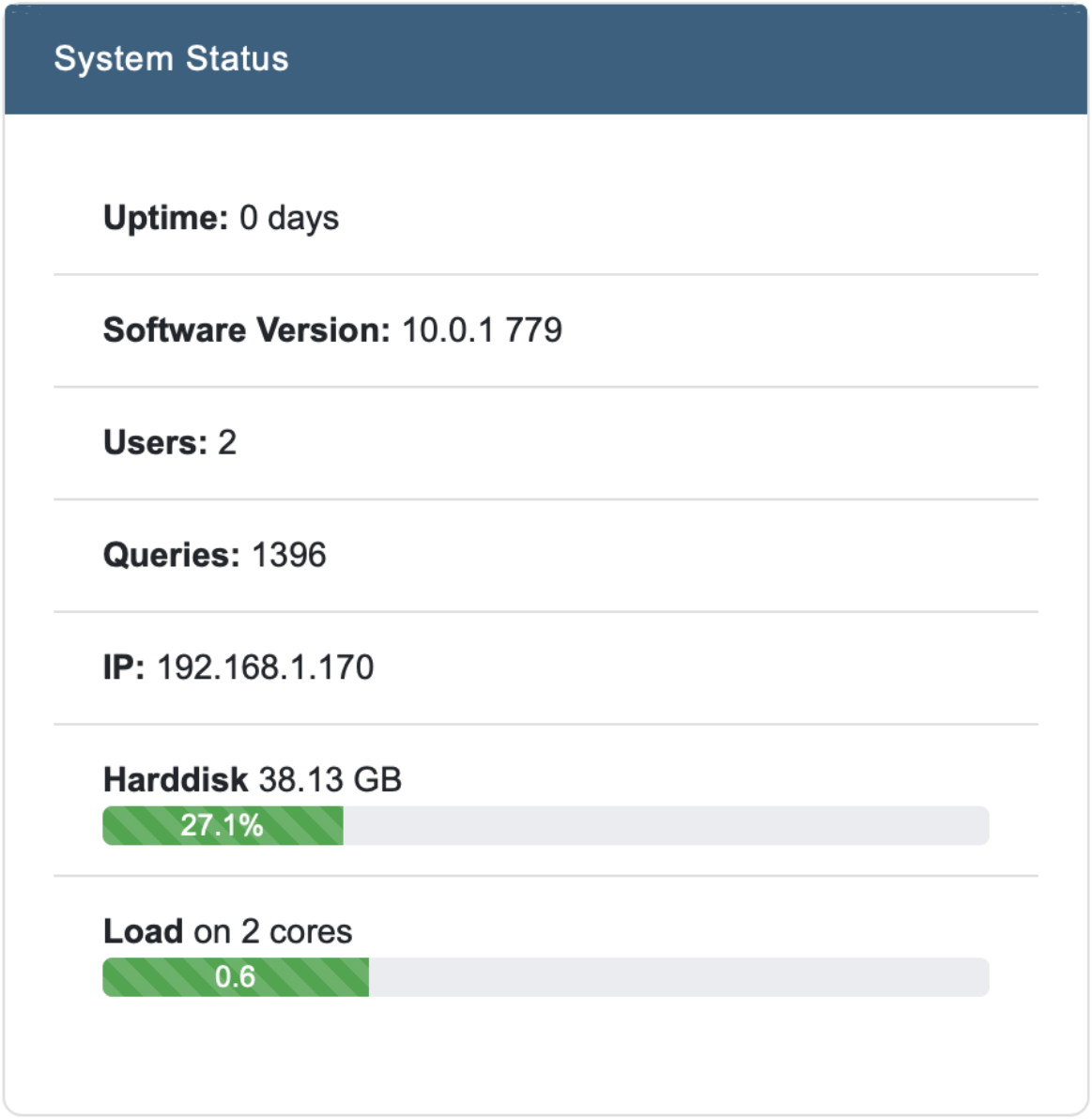
  Store all searches, md5 hashes are calculated over all items, forensic and GDPR reporting are activated.

- **Classification**

  Automatic classification on top of INDICA.

4.5 System Status

System status displays some basic system statistics and the appliances current workload.



It displays the following statistics:

- **Uptime**
  Current uptime of the appliance.
- **Software Version**
  Current software and build number version.
- **Users**
  Number of users in the system.
- **Queries**
  Total number of queries ran on the appliance. This includes system queries as well.
- **IP**
  Current IP address of the appliance.

- **Harddisk**
  Current disk usage of the appliance.

  Note: External mounts are not taken into account.

- **Load**
  Current workload on the appliance.

CHAPTER

FIVE

SYSTEM SETTINGS

Table of Contents

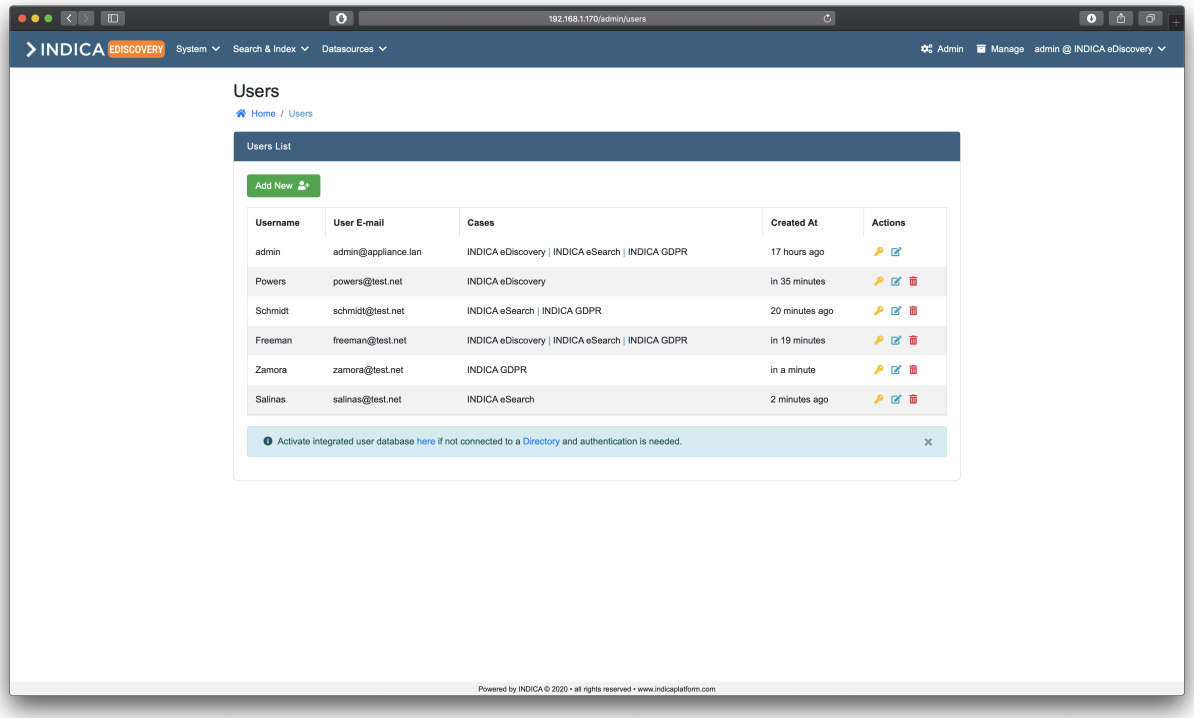
- *System Settings*
  - *Users*
    - *Users List*
    - *Creating a user*
    - *Editing a user*
    - *Deleting a user*
    - *Resetting a user's password*
    - *Resetting the admin password*
  - *Roles and Permissions*
    - *Introduction to Roles and Permissions*
      - *General information*
      - *Roles*
      - *Permissions*
    - *Configuring Roles*
      - *Adding a new role*
      - *Editing an existing role*
      - *Deleting an existing role*
    - *Default permissions matrix*
  - *Cases*
    - *Cases List*
    - *Adding a case*
      - *Mapping AD groups to roles*
    - *Editing a case*
    - *Archiving a case*
    - *Restoring an archived case*

        * *Deleting a case*
    - *Networking*
        * *Schedule bandwith throttling*
            · *Setting a daily schedule*
            · *Setting a custom schedule*
        * *Network interface*
        * *Mailer settings*
        * *Proxy settings*
        * *Zabbix monitoring*
        * *SSL Settings*
            · *Use a self-signed certificate*
            · *Submitting a CSR*
            · *Upload own certificate and keyfiles*
    - *Security*
        * *Enabling Active Directory integration*
        * *Enabling Single Sign-on*
    - *Software & Support*
        * *Activation Key*
        * *Software version*
        * *Updates*
        * *Remote Connection*
    - *General Settings*
    - *Backup Settings*
    - *Look and Feel*

5.1 Users

Overview of the User administration page. Here it's possible to create, edit, and delete users. You can also reset their password from this page.

Note: When Active Directory integration is enabled, the users and managers are administrated through the Active Directory.



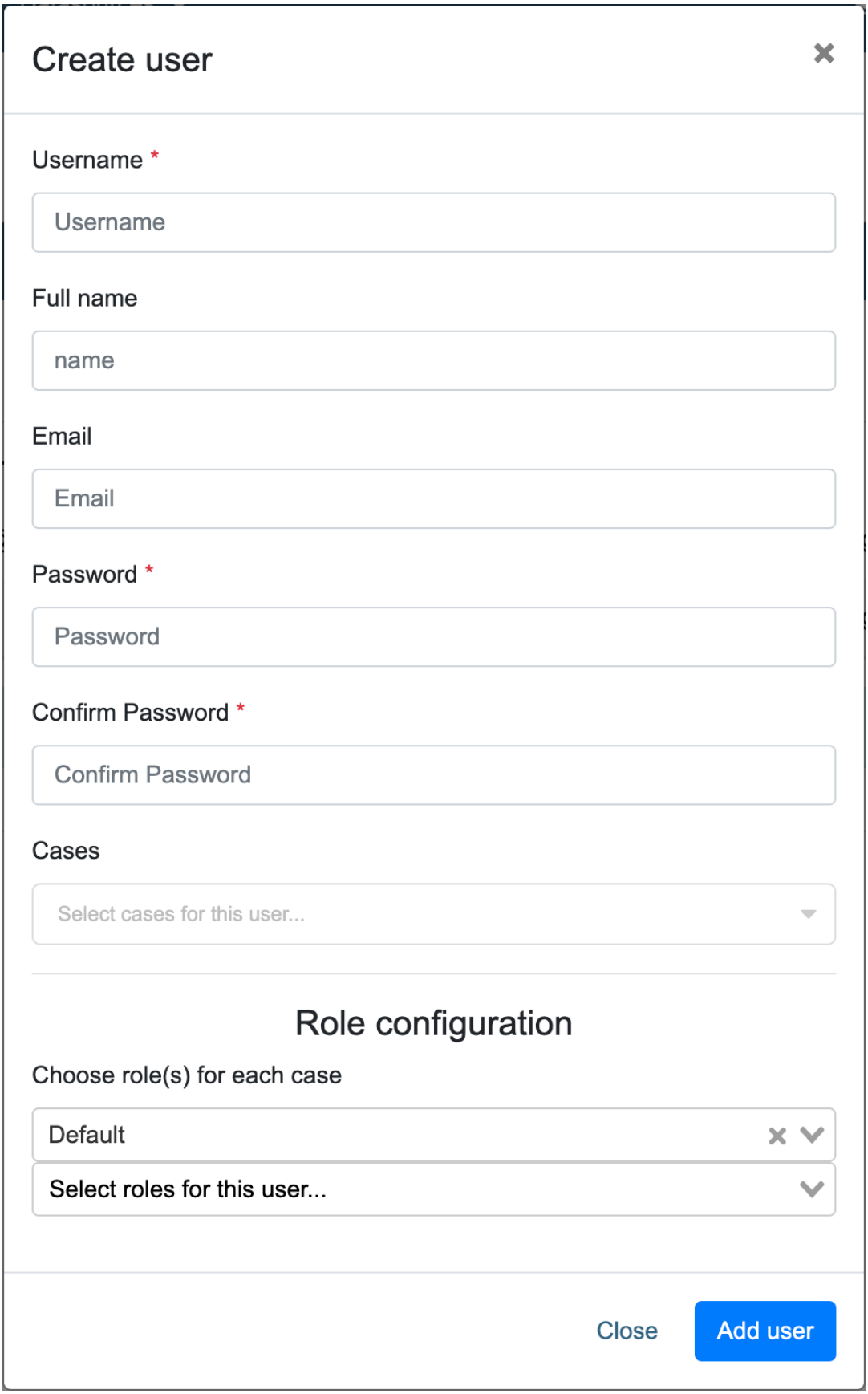
5.1.1 Users List

The listview provides administrators with some useful information and features. In the table the following columns are present:

- **Username**
  Username of the user.
- **User E-mail**
  E-mail address of the user.
- **Cases**
  List of cases that the user has access to, separated by the pipe (“|”) symbol.
- **Created At**
  Creation timestamp of the user.
- **Actions**
  - Yellow key button: Reset a user’s password.
  - Blue paper button: Edit a user.
  - Red trashbin button: Delete a user.

5.1.2 Creating a user

Creating a user is done by clicking the green “Add New” button. A popup will open:


Create user
Username *
Username
Full name
name
Email
Email
Password *
Password
Confirm Password *
Confirm Password
Cases
Select cases for this user...
Role configuration
Choose role(s) for each case
Default
Select roles for this user...
Close
Add user

Simply fill in the following details, and click “Add User” when done.

- **Username**

  Username for the new user. The username must be unique.
  This username is used to sign in into INDICA.

- **Full Name**

  Full name of the user. If given, this name will be displayed instead of the username.

- **E-mail**

  E-mail address of the user.

- **Password (2x)**

  Choose a password for the user.
  Use a strong password to prevent unauthorised access.

- **Case(s)**

  Choose the case(s) that the user should have access to.

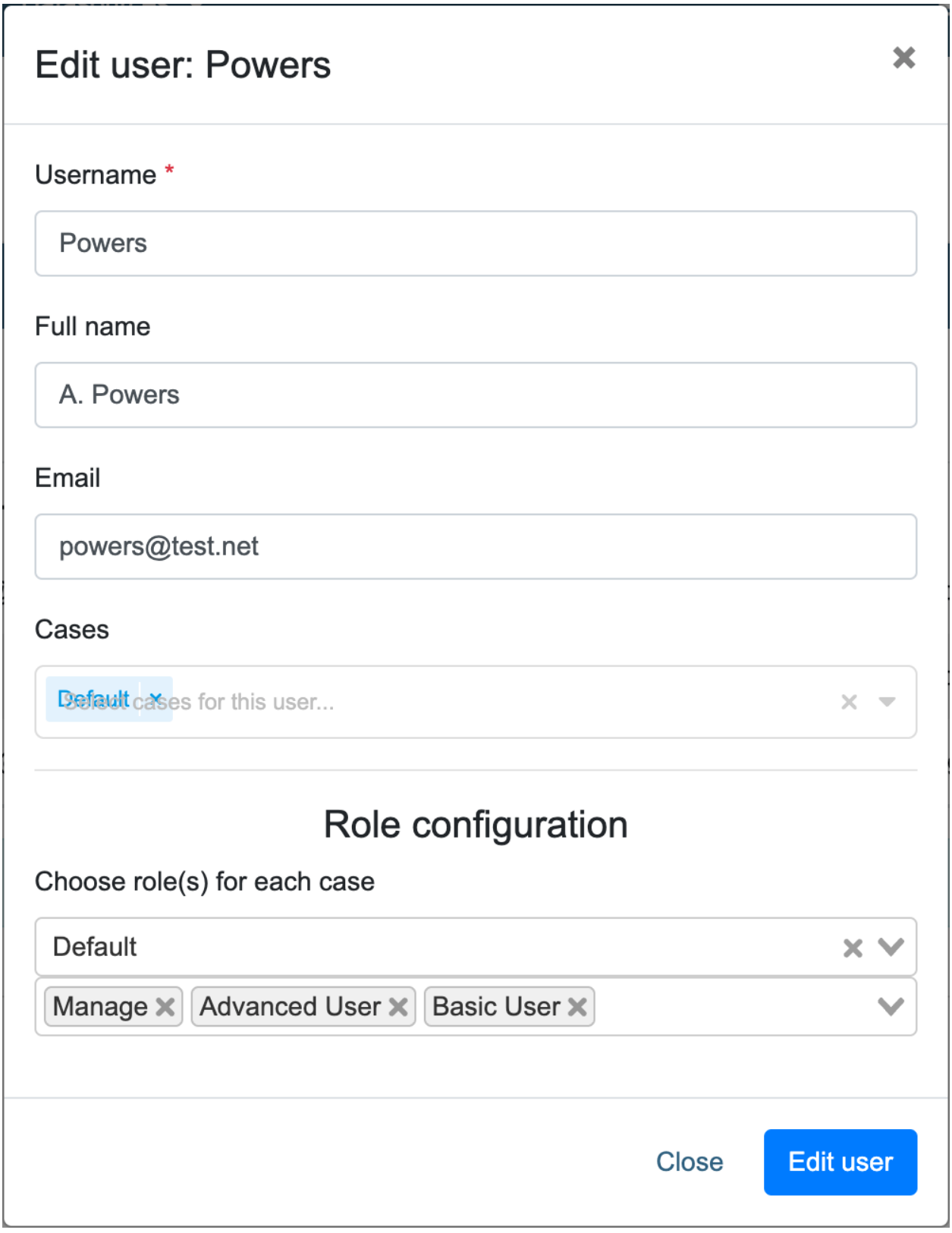
- **Role configuration (*only available when AD integration is not activated*)**

  Choose role(s) for the new user.
  Available roles are dependant on the setup of *Roles and Permissions*.
  In the top drop-down, cases can be selected. Once a case is selected, roles can be selected in the bottom dropdown for each case.

Note: Creating users is not needed when AD integration is enabled and configured. INDICA will import the user’s information from the AD upon signing in.

5.1.3 Editing a user

Editing a user is done by clicking the blue paper button in the “Actions” column of the “Users List”. A popup will open:



Here you can edit the settings of the user. When done, click “Edit user” to save the changes.

5.1.4 Deleting a user

Deleting a user is done by clicking the red trashbin button in the "Actions" column of the "Users List".

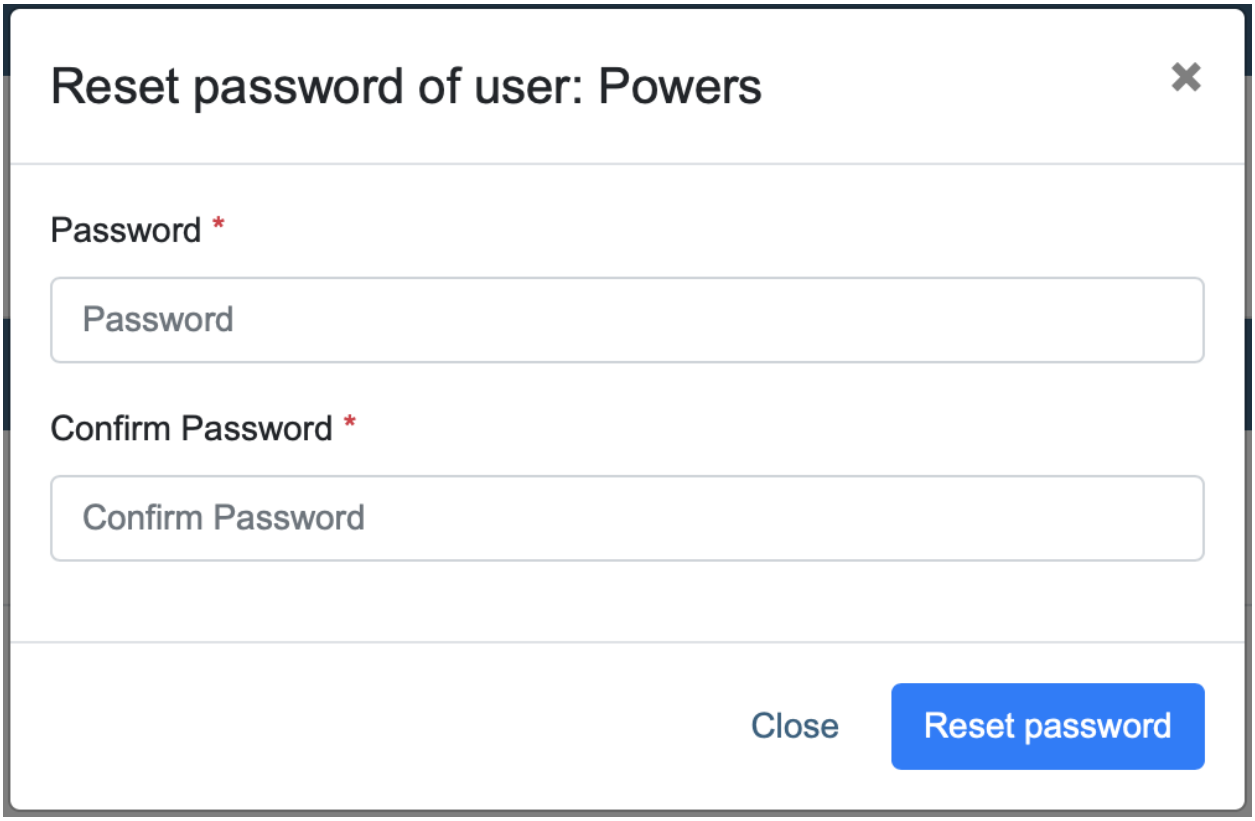
A confirmation popup will open. Click "Delete" to permanently delete the user. This action cannot be undone.

Note: Deleting a user that signed in using AD integration does not prohibit that user from signing in again. The user needs to be removed from the AD groups on the AD side.

5.1.5 Resetting a user's password

Resetting a user's password only works when there is no AD integration active.

Resetting a user's password is done by clicking the yellow key button in the "Actions" column of the "Users List". A popup will open:



Simply pick a new password, confirm it, and click "Reset Password". The user can now authenticate using the new password.

Note: The user will not be notified by INDICA that their password has changed.

5.1.6 Resetting the admin password

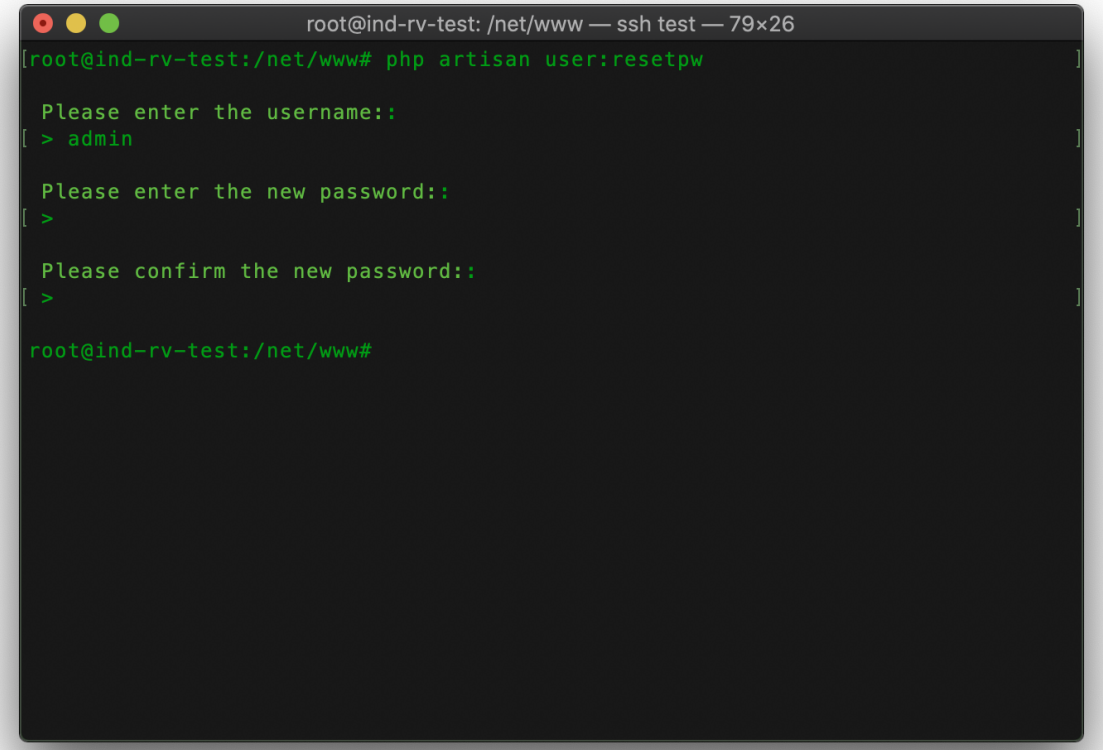
When the password of the only Admin account has been lost, there is a way to reset the password.

To do this, SSH access and root-level permissions on the appliance are needed. When logged in into the appliance, navigate to the following directory:

cd /net/www

Then, run the following command:

php artisan user:resetpw



The command will ask for the username and then for a new password.

Note: When typing the new password, nothing will be displayed.

You should now be able to authenticate in the front-end with the new password.

5.2 Roles and Permissions

When working with sensitive data, protecting that data and making sure users will only see that they need to see is very important. To accommodate this, INDICA features a Roles and Permissions system. With this system, it is possible to define custom roles and assign permissions to those roles. This gives complete control over what a user can and cannot do, while still keeping it maintainable and scalable.

5.2.1 Introduction to Roles and Permissions

General information

Roles and Permissions are a part of the Admin portal. This means that *only* administrators (or anyone with the corresponding role, if assigned by an administrator) can change the setup of roles and permissions. Users with the Admin role are unlimited in their actions. However, every action an administrator takes is logged for audit purposes. Logged actions cannot be changed by any administrator.

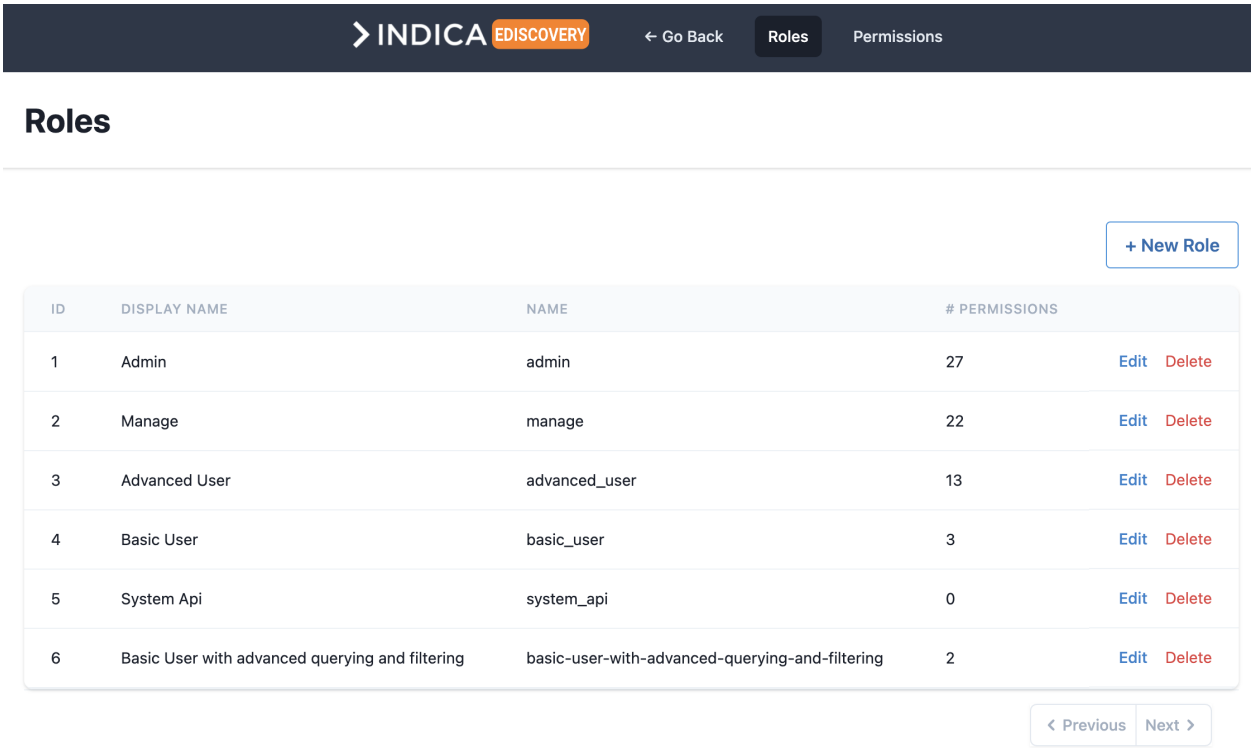
Changing the setup of roles and permissions can (by default) only be done by Administrators. However, changing the roles and permissions is a permission in itself, meaning it can be assigned to other roles if desired.

Roles

A role can be assigned to a user or a group of users. These roles have permissions assigned to them. A user can have multiple roles, and roles can have multiple permissions.

Roles can be mapped to AD groups, more information about mapping can be found in the section *Mapping AD groups to roles*.

The following roles are configured by default:



- **Admin**

  Typically assigned to administrators.
  Administrators can see and do anything on the system.
  They have unlimited access to all cases and all settings.
  Administrators are responsible for setting up and maintaining the INDICA Appliance.

- **Manage**

  This role is intended for Case Managers.
  Case managers can see and change settings to the cases they are assigned on.
  They are responsible for case settings like workflows, tags, data sources, etc.

- **Advanced User**

  The "Advanced User" role is for trained and experienced users.
  This role gives more information and features on the Search page.

- **Basic User**

  The "Basic User" role is for new users with limited rights.
  This role gives access to a limited Search page with only basic functionality.

- **System Api**

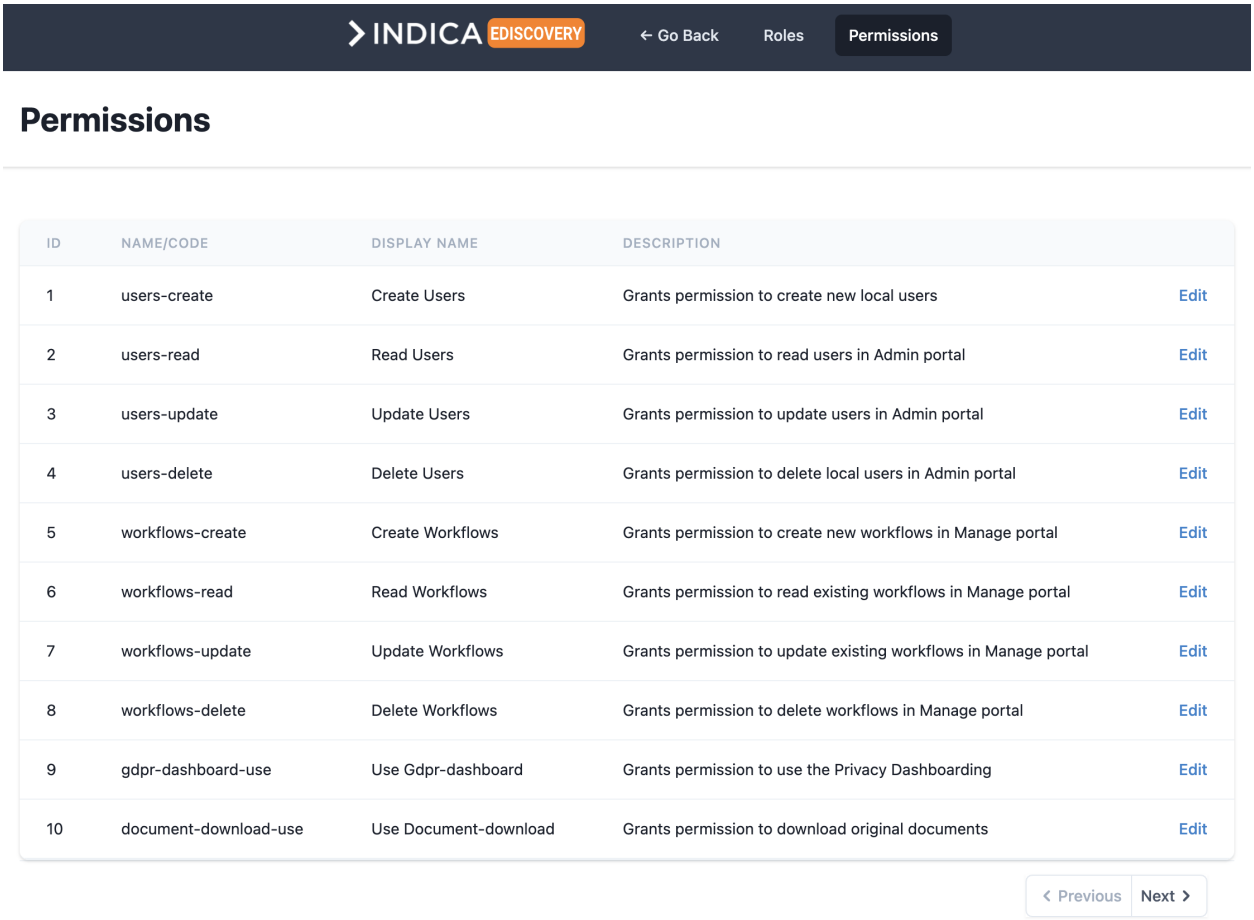
  Internal system user. Is only used for INDICA's inner workings.
  Changing or assigning this role is not needed.

Note: Roles are system-wide. This means that when a role is edited, the permissions of that role will apply to all cases that use this role. If this behavior is not desired, it is possible to create roles for a specific Case and use those roles for that Case. This allows administrators to have Case-specific roles (and permissions).

Permissions

A permission is something a user can see or do in the system. This can be assigning a tag, viewing document content, changing settings, etc. These permissions are attached to a role, and roles are attached to users (or AD groups, see *Mapping AD groups to roles*).

New permissions cannot be created by the administrator or any other user. This is because permissions are defined in the sourcecode of the INDICA Appliance. New permissions will be included in future releases of INDICA. If you're an existing customer, they can also be requested. Contact INDICA for more details.



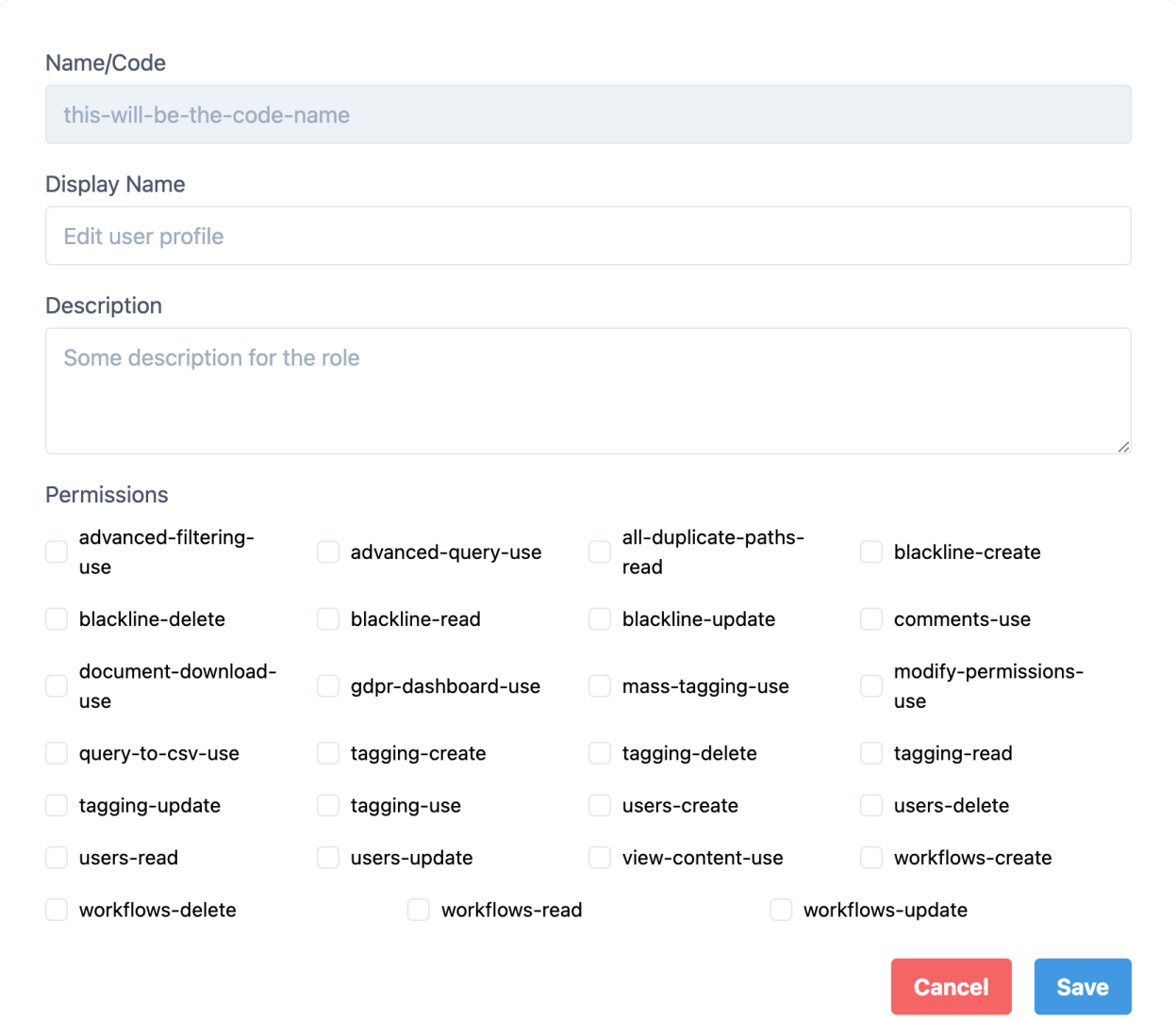
For a full list of all permissions, their description and which roles have which permission, please refer to the *Default permissions matrix*.

Note: Permissions cannot be deleted. They can be edited, but only the display name and the description can be changed.

5.2.2 Configuring Roles

Adding a new role

A new role can be created by clicking the "+ New Role" button on the "Roles" setup page. The following screen will open:



In this form, the following information is required:

- **Name/Code**

  Cannot be filled in by the user, will be generated based on the Display Name.

- **Display Name**

  Name of the role, for example “Basic User with advanced querying and filtering”.

- **Description**

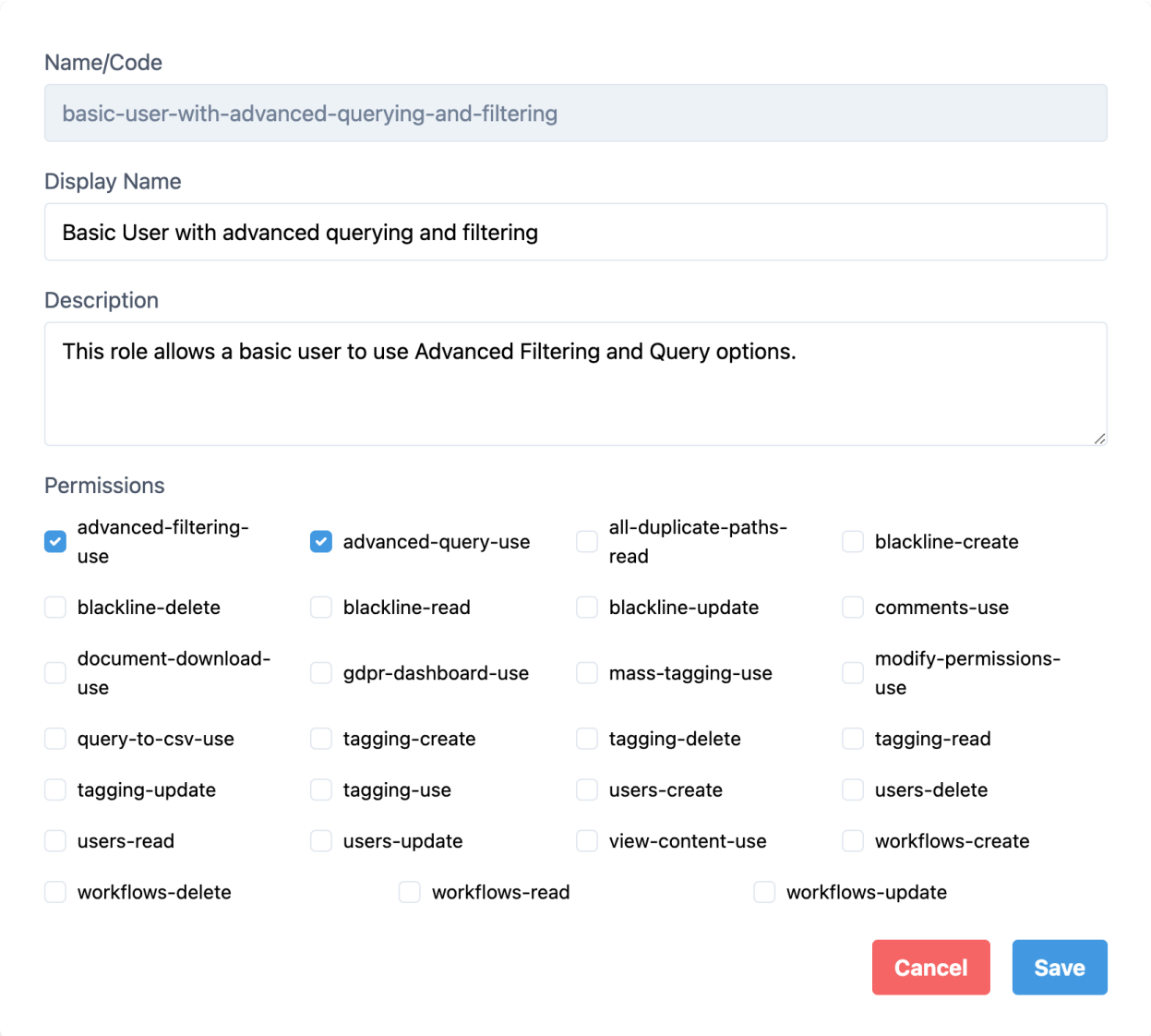
  Description of this role, so they can be distinguished.

- **Permissions**

  Check the boxes for the required permissions.

Note: In this example, the new role is an extention of an already existing role. This means that if this new role is assigned to a user alongside the already existing role, the permissions of both roles will be combined. This makes it unnecessary to select all permissions while setting up this new role. The best course of action will be to select the extra permissions and assign both roles to the user(s). However, it is also possible to add the permissions of the "Basic User" role to this role as well, and only assign this role to the user(s).

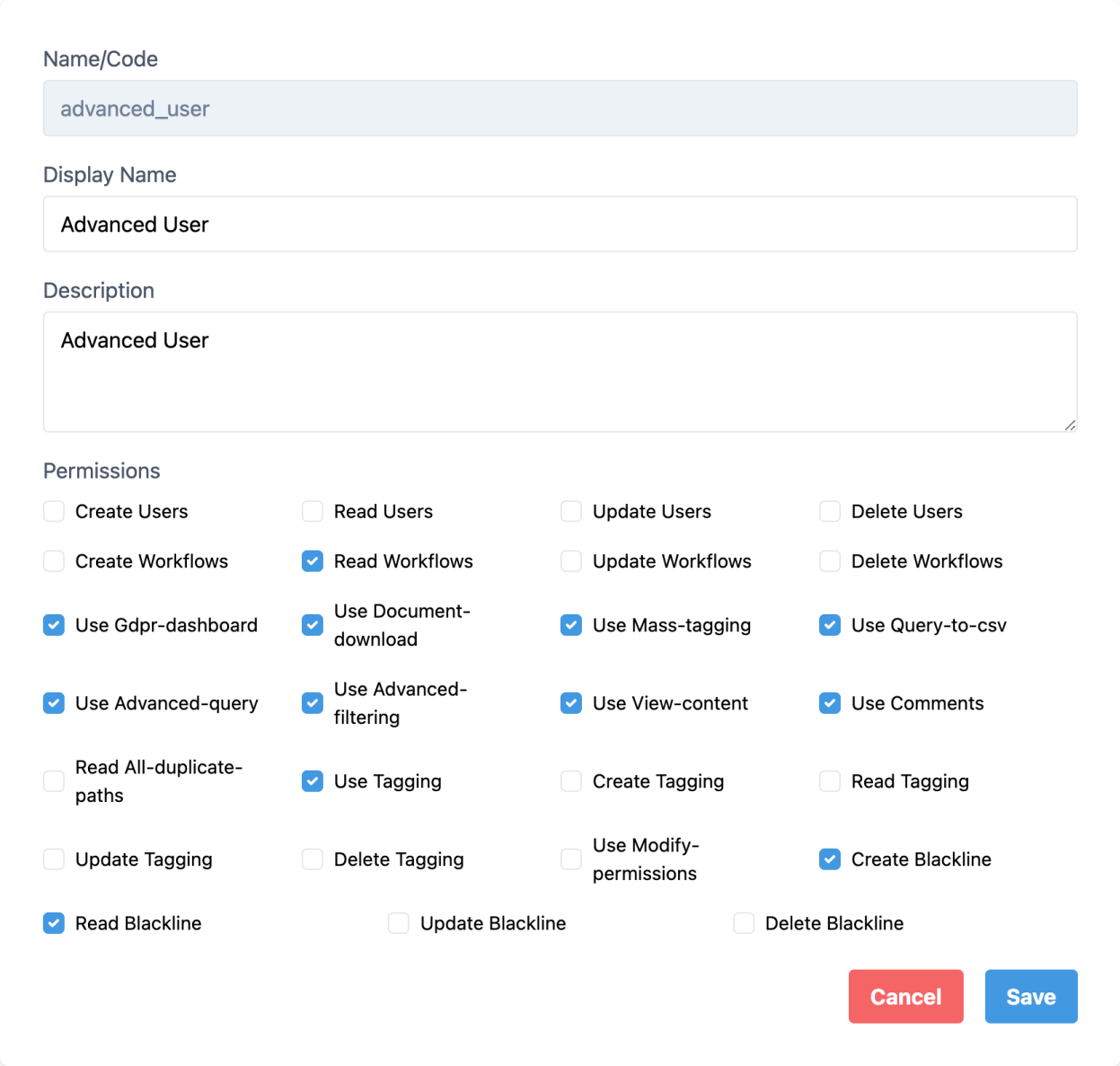
Once all information is entered, it will look like the following image:



Once done, click "Save" to save your new role. It should now be displayed in the overview. This role can now be assigned to user(s).

Editing an existing role

To edit an existing role, click the “Edit” option. The following page will be displayed:



Simply make the edits needed, and click “Save” to save the changes.

Note: The system will immediately recognise the changes and assign the changed permissions to the users that have this role assigned. In some situations it may be necessary to sign out and sign back in for the changes to take effect.

Deleting an existing role

To delete an existing role, click the "Delete" option. A confirmation prompt will open, asking you if you're sure you want to delete this record. Click "OK" to delete this role.

Note: Deleting a role that is assigned to user(s) may have implications on the permissions of the user(s). It is advised to deattach the role before deleting it.

5.2.3 Default permissions matrix

This table shows the Authorisation Matrix of the default roles and permissions.

Table 1: Permission Matrix

Permission name	Description	Role(s)
users-create	Grants permission to create new local users	Admin
users-read	Grants permission to read users in Admin portal	Admin
users-update	Grants permission to update users in Admin portal	Admin
users-delete	Grants permission to delete local users in Admin portal	Admin
workflows-create	Grants permission to create new workflows in Manage portal	Admin Manage
workflows-read	Grants permission to read existing workflows in Manage portal	Admin Manage Advanced User
workflows-update	Grants permission to update existing workflows in Manage portal	Admin Manage
workflows-delete	Grants permission to delete workflows in Manage portal	Admin Manage
gdpr-dashboard-use	Grants permission to use the Privacy Dashboarding	Admin Manage Advanced User
document-download-use	Grants permission to download original documents	Admin Manage Advanced User
mass-tagging-use	Grants permission to use tagging in batches by query and page	Admin Manage Advanced User
query-to-csv-use	Grants permission to export query results to CSV file	Admin Manage Advanced User
advanced-query-use	Grants permission to use the Advanced Query Builder	Admin Manage Advanced User
advanced-filtering-use	Grants permission to use Advanced Filtering	Admin Manage Advanced User

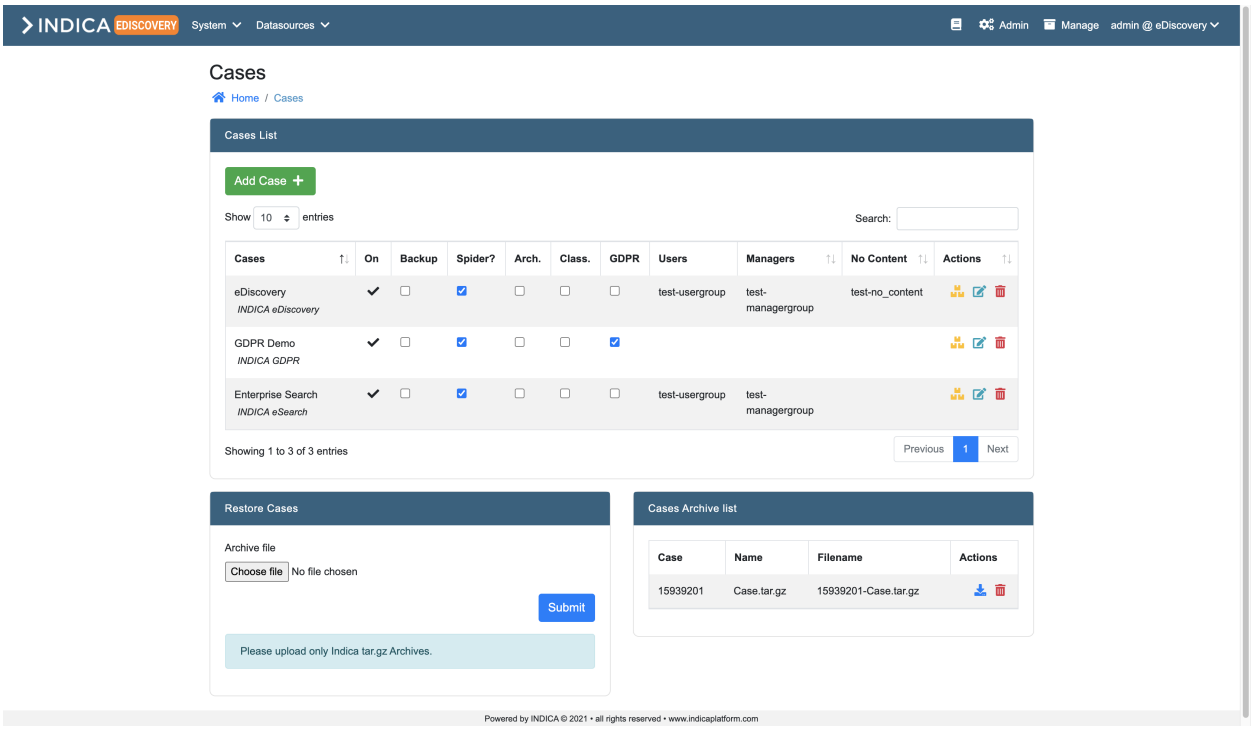
continues on next page

Table 1 – continued from previous page

Permission name	Description	Role(s)
view-content-use	Grants permission to view document content	Admin Manage Advanced User Basic User
comments-use	Grants permission to create read update and delete comments on documents	Admin Manage Advanced User Basic User
all-duplicate-paths-read	Grants permission to view all paths of a document with duplicates	Admin Manage
tagging-use	Grants permission to use the tagging functionality	Admin Manage Advanced User Basic User
tagging-create	Grants permission to create new tags in Manage portal	Admin Manage
tagging-read	Grants permission to view existing tags in Manage portal	Admin Manage
tagging-update	Grants permission to update existing tags in Manage portal	Admin Manage
tagging-delete	Grants permission to delete tags in Manage portal	Admin Manage
modify-permissions-use	Grants permission to modify Roles & Permissions	Admin
blackline-create	Grants permission to redact documents	Admin Manage Advanced User
blackline-read	Grants permission to view redacted documents	Admin Manage Advanced User
blackline-update	Grants permission to update already redacted documents	Admin Manage
blackline-delete	Grants permission to remove redacted documents	Admin Manage

5.3 Cases

Overview of the Case administration page. Here it's possible to create, edit, archive and restore cases.



5.3.1 Cases List

The listview offers some settings and features. The checkboxes can be used to toggle functionality or settings on a case. The columns in the table are:

- **On**
  Whether the case is enabled or not. This setting can be toggled on the “Edit Case” page.
- **Backup**
  Whether the case will be backed up when backup settings are configured.
- **Spider?**
  Enable or disable the spider for this case.

  > **Caution:** Do not disable the spider during indexing.

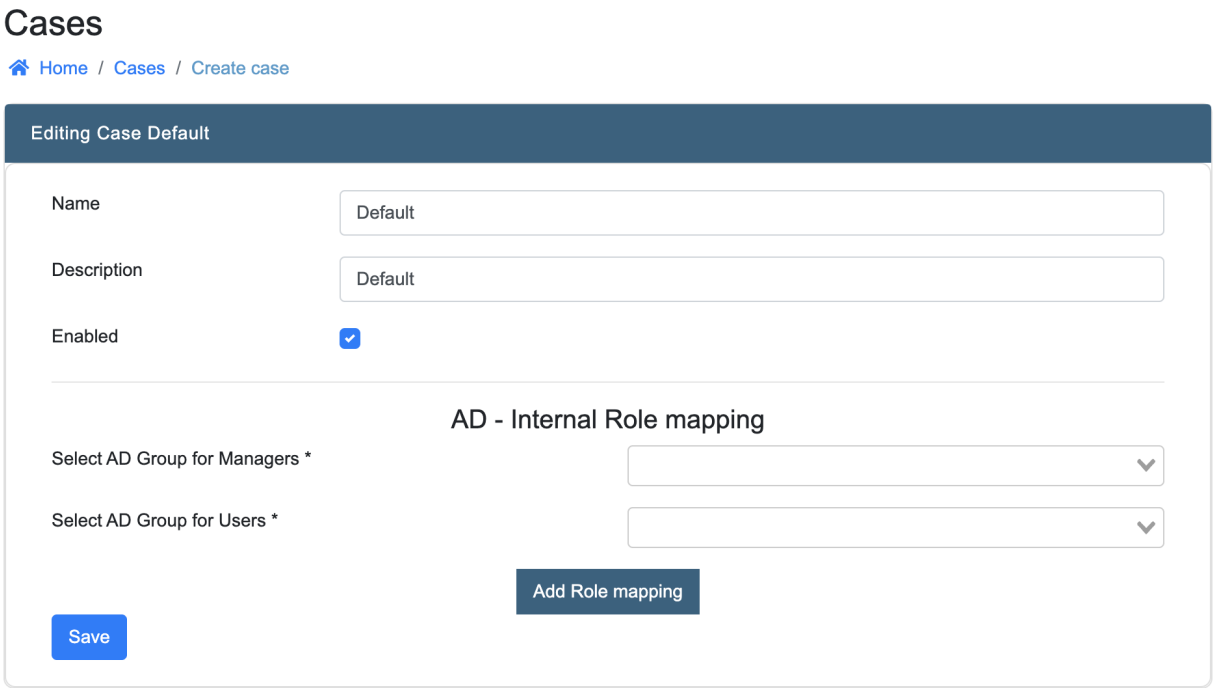
- **Arch.**
  Whether the case is archived or not.
- **Class.**
  Enable or disable the Classification Module.
- **GDPR**
  Enable or disable the GDPR module.
- **Users**
  Name of the Active Directory group to grant “User”-level access to this case.
- **Managers**
  Name of the Active Directory group to grant “Manager”-level access to this case.
- **Actions**

- Yellow boxes button: Click this button to archive the case.
- Blue paper button: Edit this case.
- **Red trashbin button: Delete this case.**

Note: All but one case can be deleted. The system needs at least one running case.

5.3.2 Adding a case

From this page, it's possible to create a new case on the appliance. Simply click the "Add Case" button. You will be redirected to the following page:



Here you can fill in the following details:

- **Name**
  Name of the case.
- **Description**
  Description of the case.
- **Enabled**
  Whether the case is enabled or not. If unchecked, the case will not show up for users and managers.
- **AD Group for managers (*Only available when AD integration is active and configured*)**
  The Active Directory group that is used to grant users access to this case with "Manager"-level.
- **AD Group for users (*Only available when AD integration is active and configured*)**
  The Active Directory group that is used to grant users access to this case with "User"-level.
- **Extra role mappings (*Only available when AD integration is active and configured*)**
  This option allows you to map extra roles to AD groups. See *Mapping AD groups to roles*.

Tip: If AD integration is not enabled, it is still possible to set up *Roles and Permissions* for users.

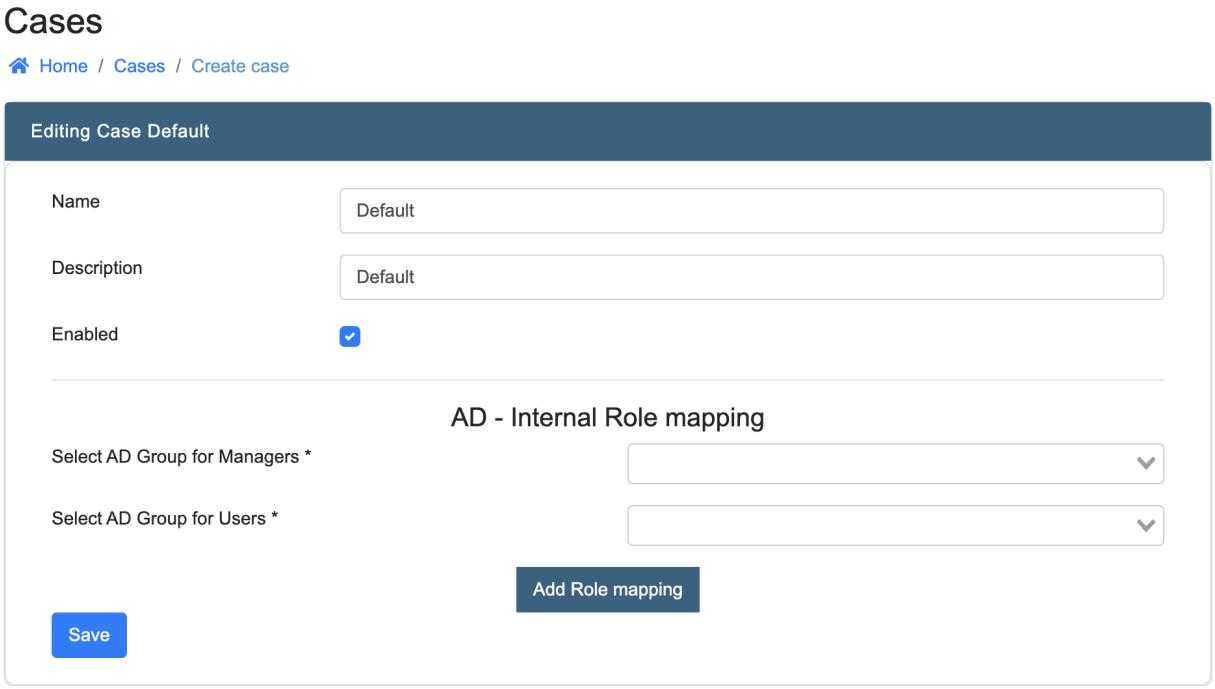
When all details are filled in, click the "Create" button. A new case will be created. This may take several seconds, depending on the size of the system and the amount of indexed data.

Mapping AD groups to roles

If more roles need to be connected to AD groups, it is possible to map them. To do this, navigate to the Case settings page and edit the case in question.

It can also be setup during the creation of a new case.

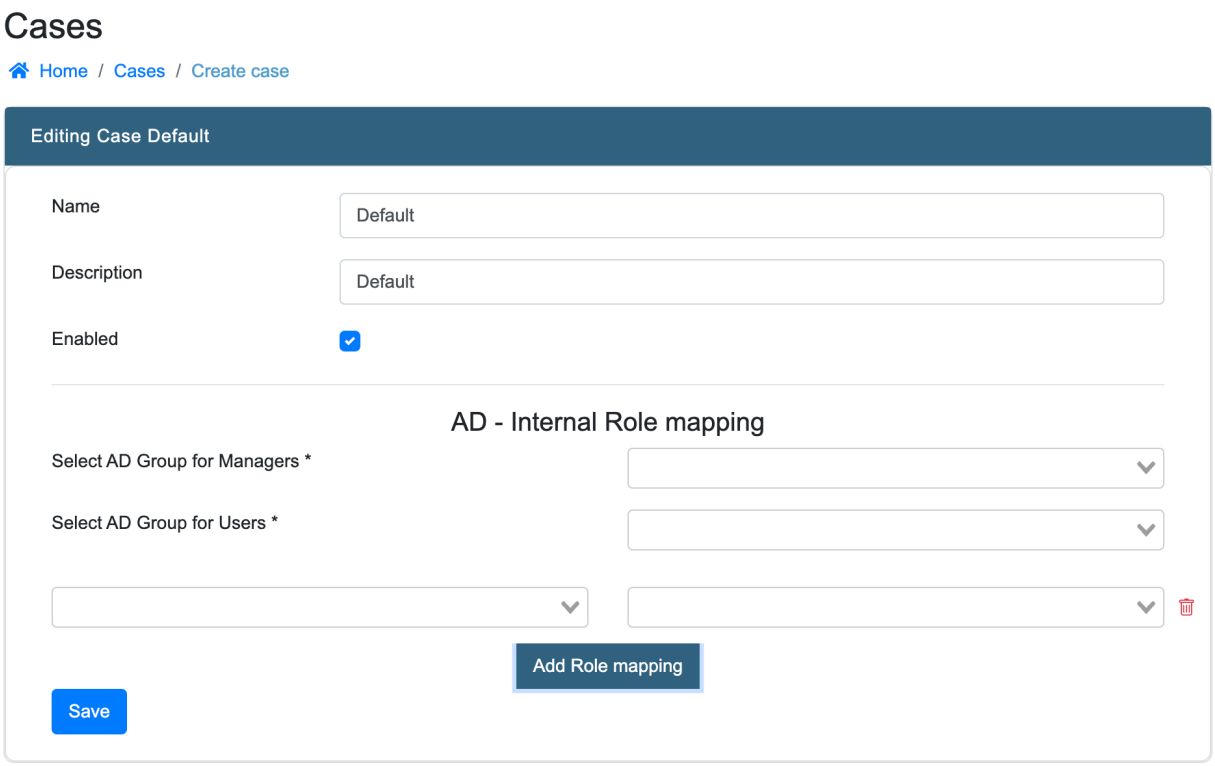
You will see the following:



Two possible mappings may be presented already:

- **Select AD Group for Managers**
  Use this option to select an AD group that will give its users the Manager role in INDICA
- **Select AD Group for Users**
  Use this option to select an AD group that will give its users the User role in INDICA

In addition, the button "Add Role mapping" may be pressed to present a blank line:



In this line, select the desired role on the left side and the corresponding AD group on the right side. This step can be repeated for every role that needs an AD mapping.

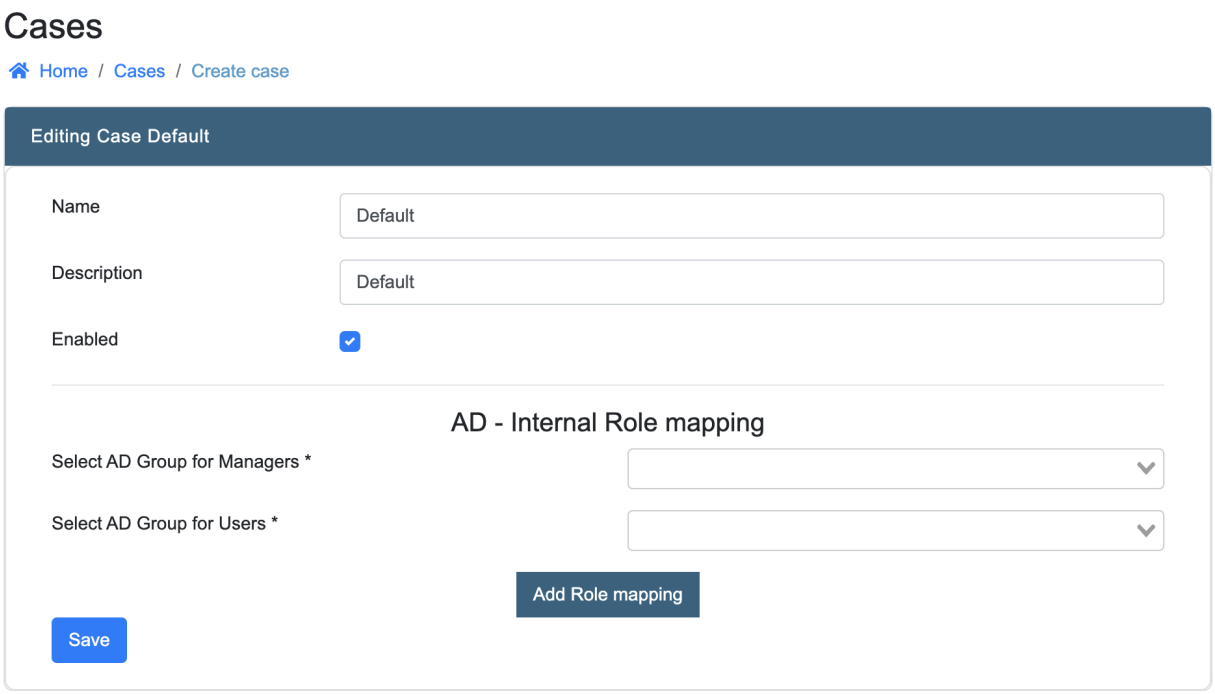
Mappings can also be removed by clicking the red trashbin icon on the right side of the row.

In some situations, a message stating “Modifying AD mappings will logout all users belonging to the group” may be displayed. This means that users will get a forced logout in order for the new settings to take effect.

When done, click “Save” to store the settings.

5.3.3 Editing a case

To edit a case, navigate to the “Cases List” and click the blue paper icon in the “Actions” Column. The following page will open:



From here, it's possible to edit the case. See "Adding a case" for details about the fields. Click "Save Changes" when the desired changes have been made.

This page also shows the users that have access to this case and the created collections in this case.

5.3.4 Archiving a case

To archive a case, navigate to the "Cases List" and click the yellow boxes icon in the "Actions" column. INDICA will now automatically archive this case. The case will be disabled and cannot be accessed anymore.

When the archiving is done, the case will be displayed in the "Cases Archive list".

Note: This list displayes the archived cases by their Case ID as this is unique for every case.

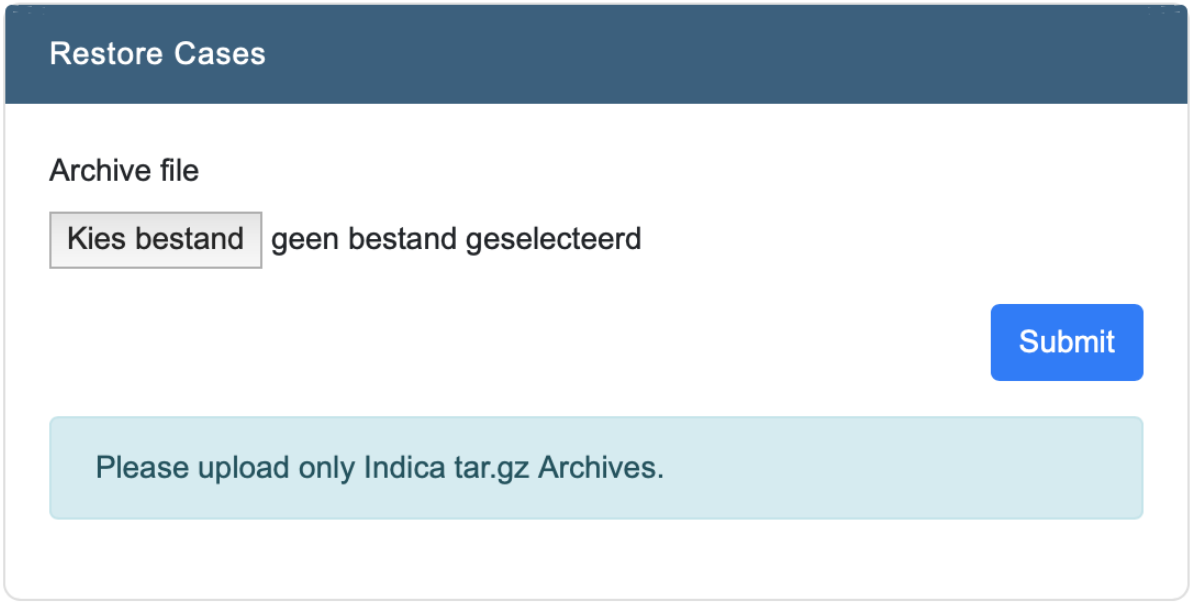
From this list, the archived case can be downloaded in a .tar.gz format. The downloaded file will contain everything that is needed to later restore the case.

If the archived case is no longer needed, it can be deleted by clicking the red trashbin icon.

Caution: This action is irreversible and the data is lost permanently.

5.3.5 Restoring an archived case

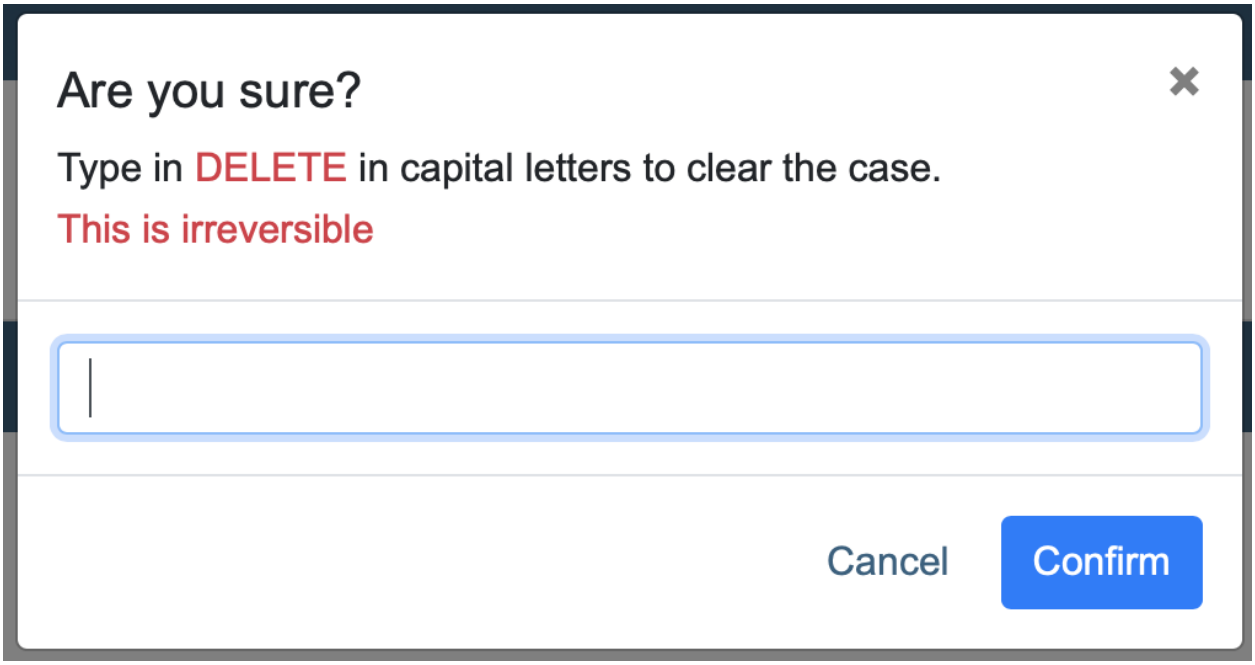
INDICA offers the possibility to restore a previously archived case. To do this, navigate to the "Cases List" page. Here a section is shown where you can upload INDICA .tar.gz files.



Click the "Choose File" button to select the .tar.gz archive. After choosing the correct file, the upload will start. When the upload is done, click the "Submit" button to restore the case from the archive. The case will now be restored and displayed in the "Cases List".

5.3.6 Deleting a case

Deleting a case is done from the "Cases List" page. In the "Actions" column, a red trashbin icon is displayed. When that icon is clicked, a confirmation box will pop up. This is to prohibit accidental deletion.

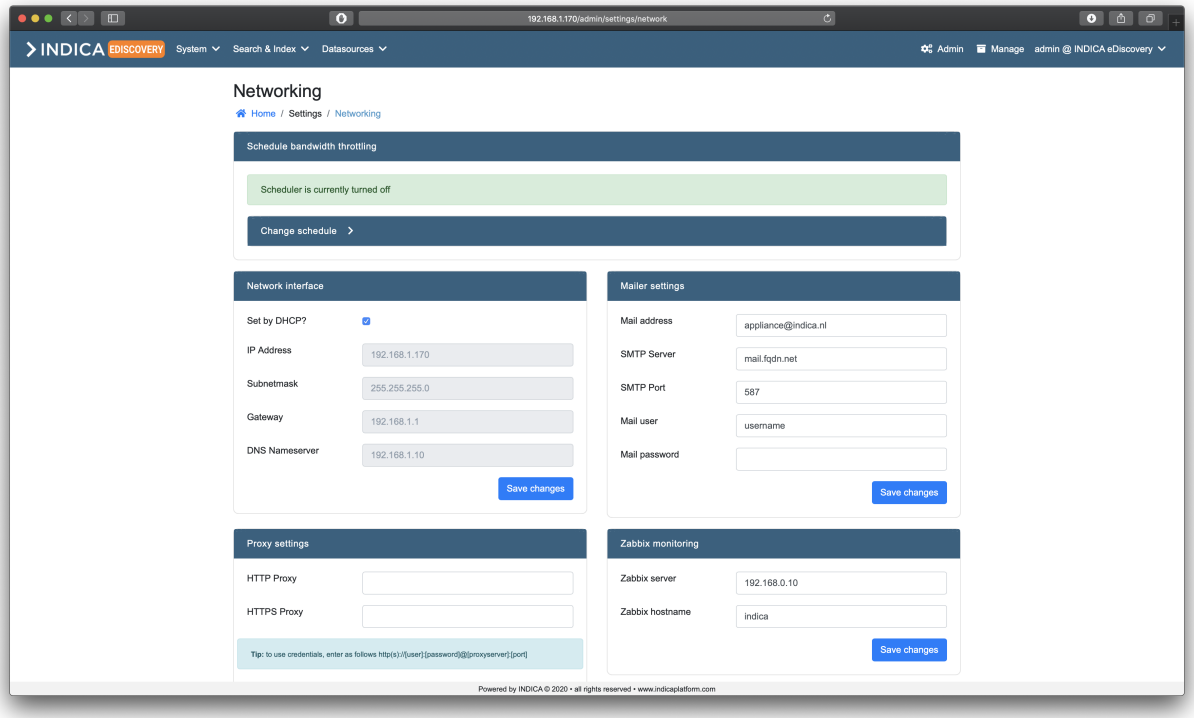


Caution: Deletion of a case is permanent and irreversible. This action cannot be undone.

There must always be one case in the system. The last case cannot be deleted.

5.4 Networking

Overview of the Networking page.

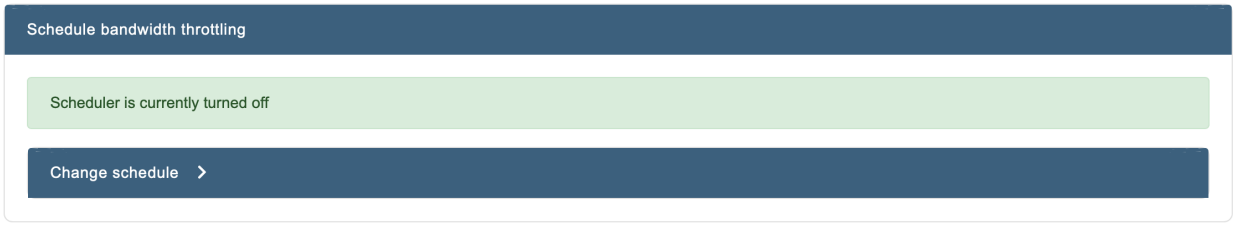


On this page, several settings regarding networking can be configured:

- **Schedule bandwith throttling**
  Schedule how much bandwith the INDICA appliance may use at certain times/days.
- **Network interface**
  Basic settings for configuring the network interface
- **Mailer settings**
  Connect INDICA to a mail server so it can send out notifications
- **Proxy settings**
  Configure a proxy server
- **Zabbix monitoring**
  Configure zabbix monitoring
- **SSL Settings**
  Configure a Secure Socket Layer certificate for the web front-end.

5.4.1 Schedule bandwith throttling

While indexing, the INDICA appliance can take up quite a bit of network bandwith. To make sure that the infrastructure will not be overloaded, you can schedule the maximum allowed bandwith to be used by the INDICA appliance.

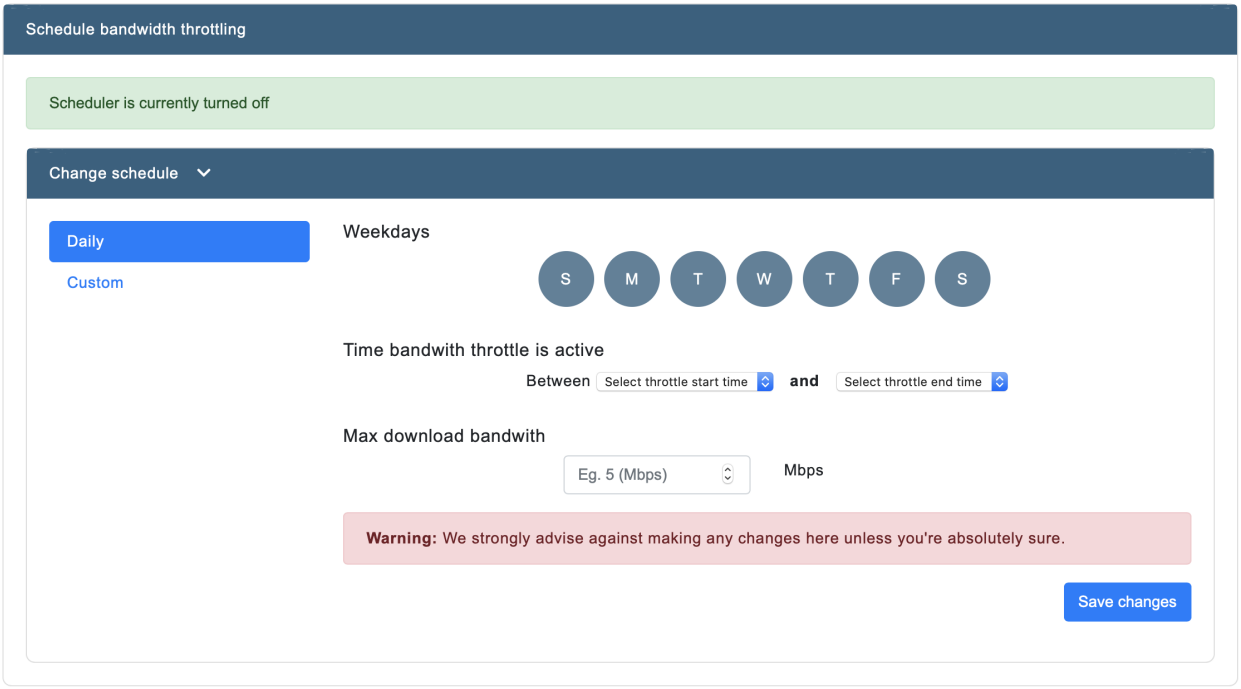


There are two options for scheduling:

- **Daily Schedule**
  This schedule will apply every day. It allows you to set a throttle start time and stop time, and the maximum Mbps it may use.
- **Custom Schedule**
  This schedule allows you to select the days. It allows you to set a throttle start time and stop time, and the maximum Mbps it may use.

Setting a daily schedule

First, click the "Change schedule" button. You will be presented with the following options:

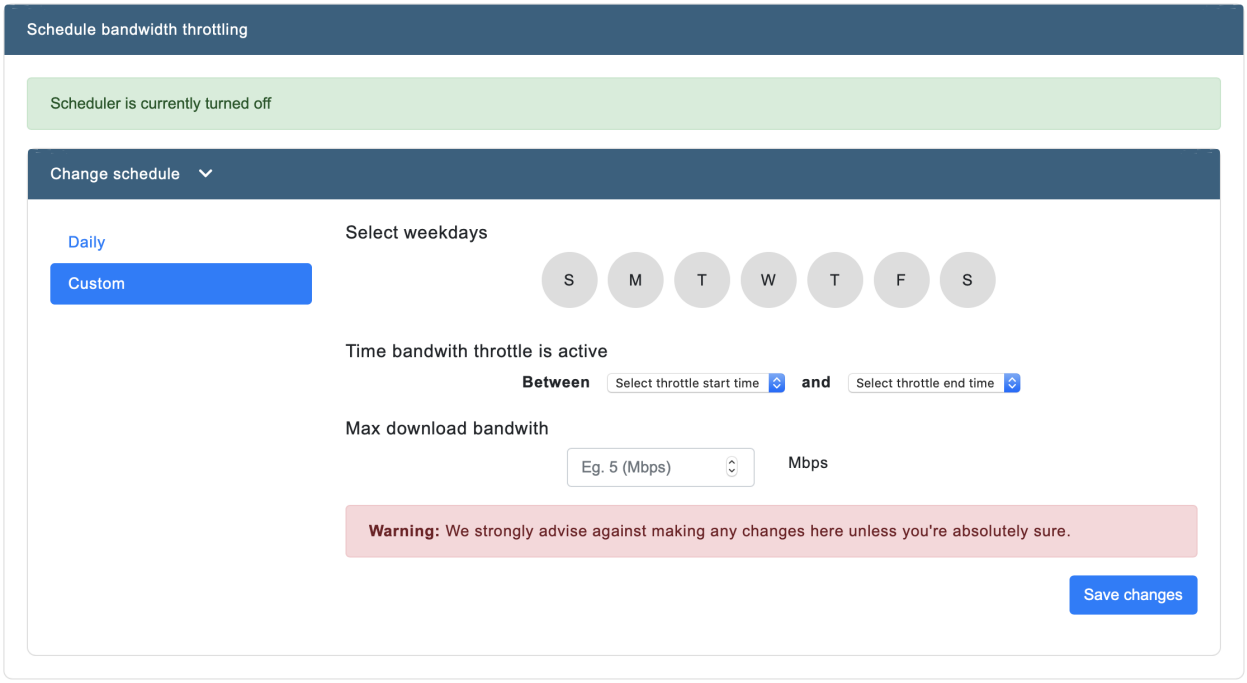


Then, pick a time for the throttle to start and to stop. Lastly, enter a value for the maximum Mbps the appliance may use.

Click "Save changes" to apply the schedule. The schedule should now be active.

Setting a custom schedule

First, click the “Change schedule” button. You will be presented with the following options:



Now click the days you want the schedule to be active. Then, pick a time for the throttle to start and to stop. Lastly, enter a value for the maximum Mbps the appliance may use.

Click “Save changes” to apply the schedule. The schedule should now be active.

5.4.2 Network interface

In most cases, the network will be configured by DHCP. In case that is not true in your setup, this panel allows you to set the correct network settings.

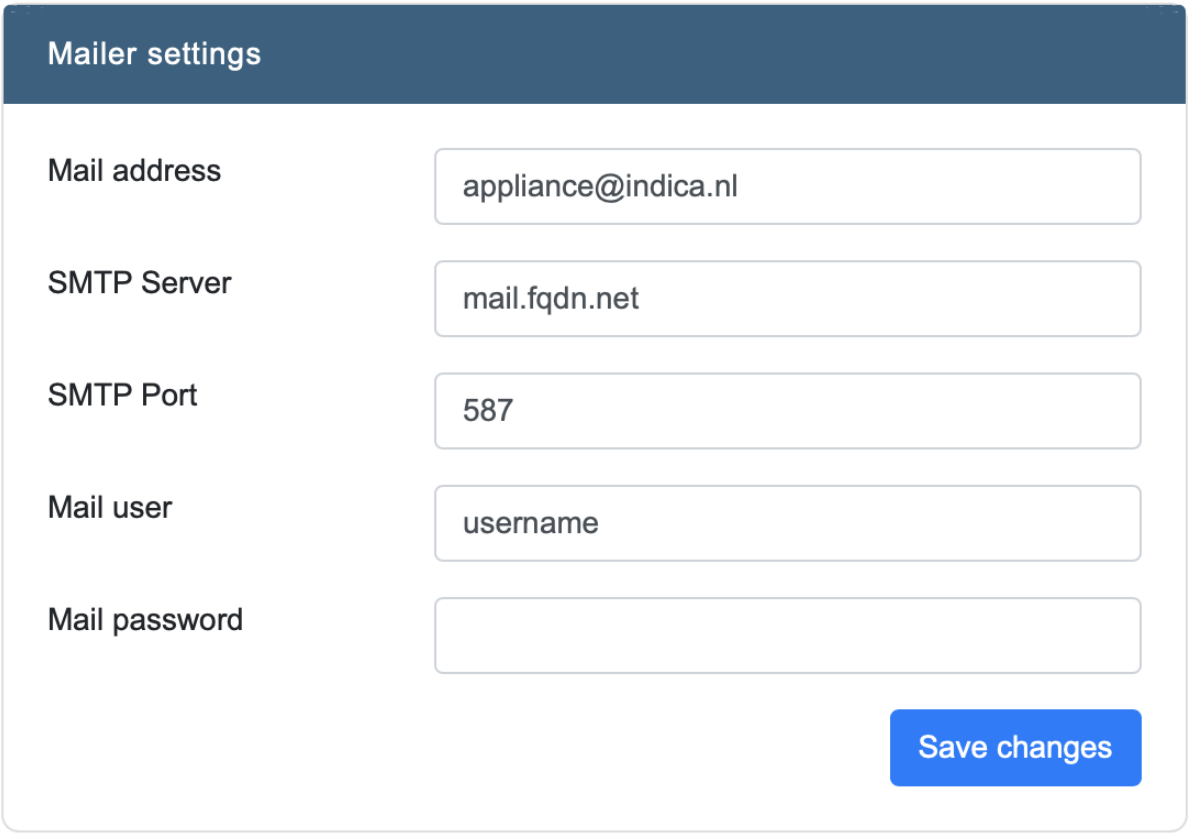
Here you can configure the following settings:

- IP address
- Subnetmask
- Gateway
- DNS Nameserver

Note: When the checkbox "Set by DHCP?" is checked, these settings cannot be changed.

5.4.3 Mailer settings

Here you can configure INDICA to connect to a mail server so it can send out emails.



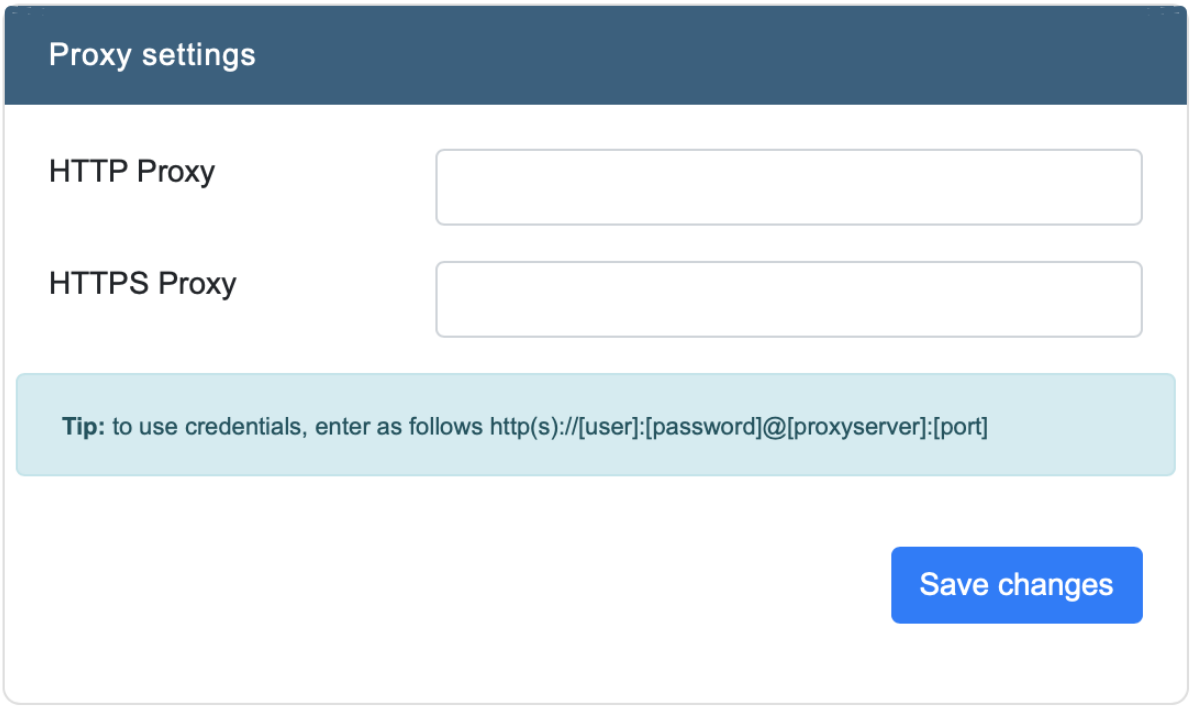
It needs the following information:

- Mail address
- SMTP server
- SMTP port
- Mail user
- Mail password

After entering the correct information, click “Save Changes” to apply the changes.

5.4.4 Proxy settings

This panel contains the Proxy settings.



Note: To use credentials, enter the settings in the following format:

http(s)://[user]:[password]@[proxyserver]:[port]

5.4.5 Zabbix monitoring

Here you can enter the settings for Zabbix monitoring.

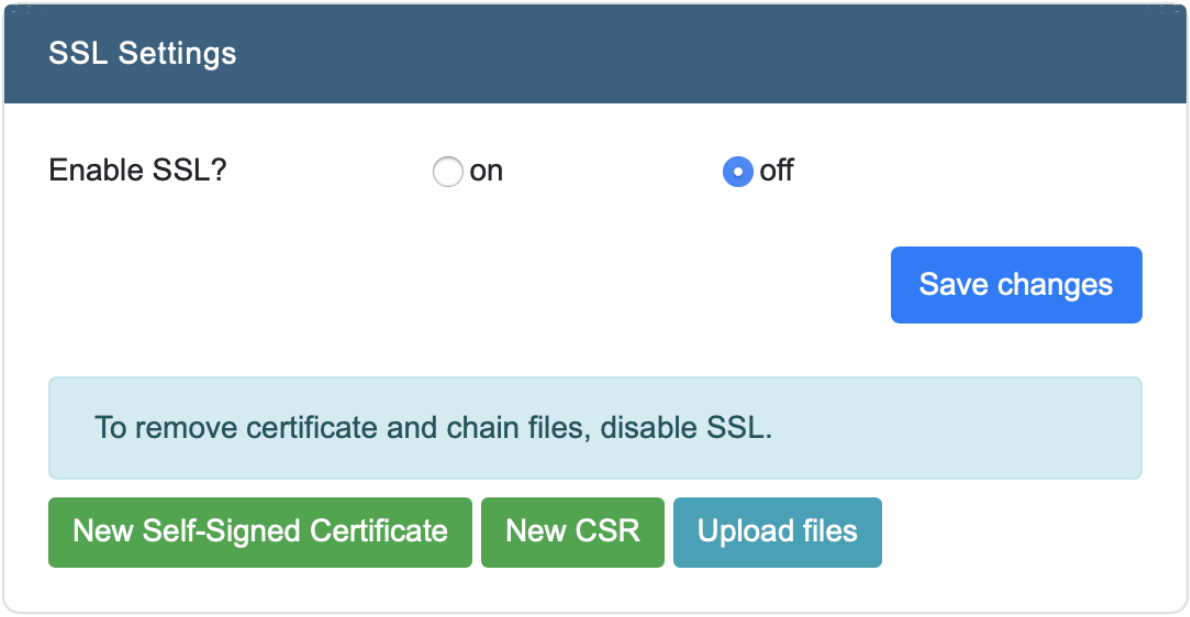
It needs:

- Zabbix Server (IP address)

- Zabbix Hostname (usually ‘indica’)

5.4.6 SSL Settings

SSL Settings allows you to use SSL to secure the client-server traffic of the web-based frontend.



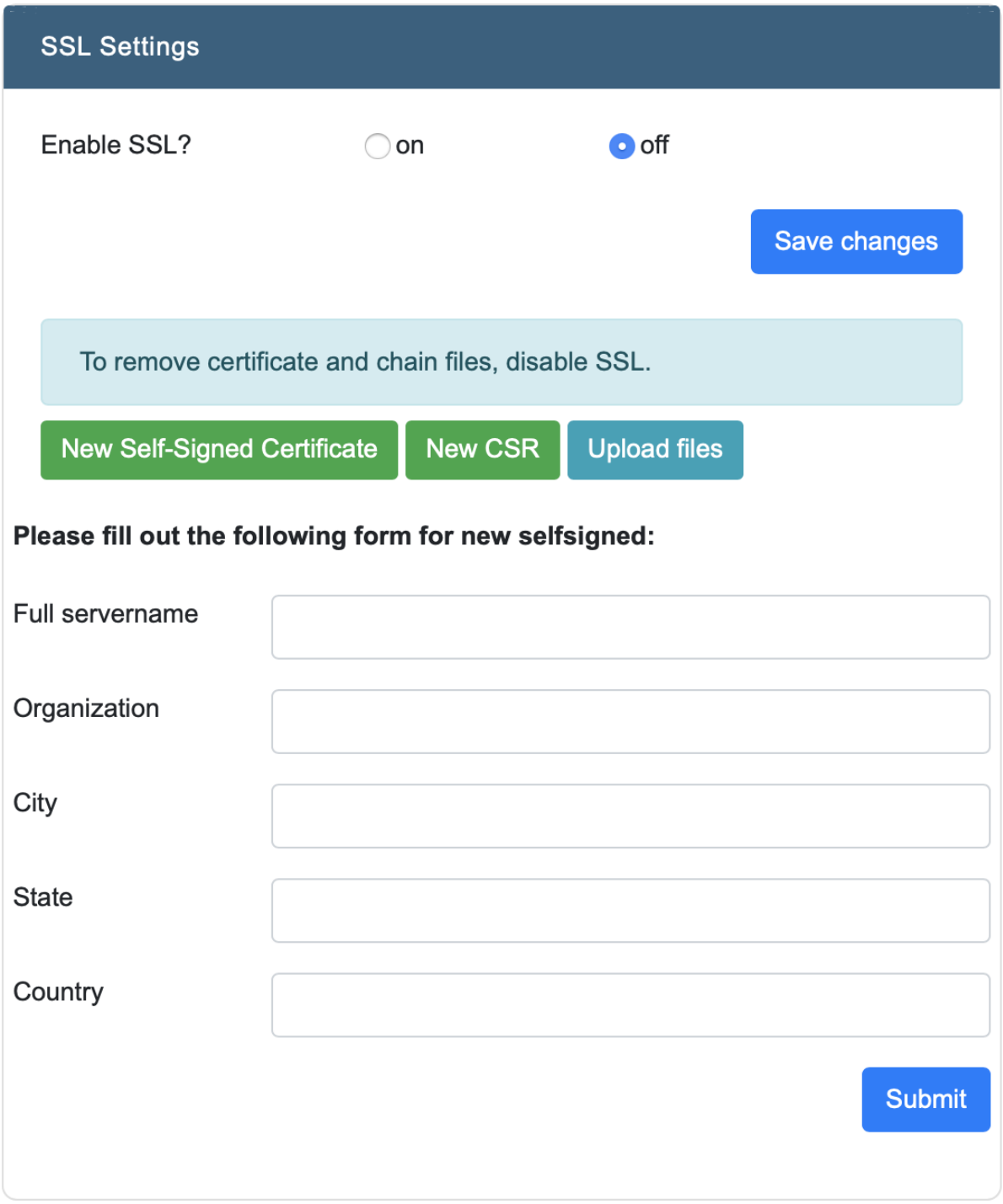
Note: Make sure all configuration is added correctly before enabling SSL. If SSL is enabled before all information is given, it may make the front-end inaccessable.

SSL can be set up in three ways:

- Use a self-signed certificate
- Submit a CSR (Certificate Signing Request)
- Upload own certificate and keyfiles

Use a self-signed certificate

When clicking the button “New Self-Signed Certificate”, you will be prompted to fill in some information:



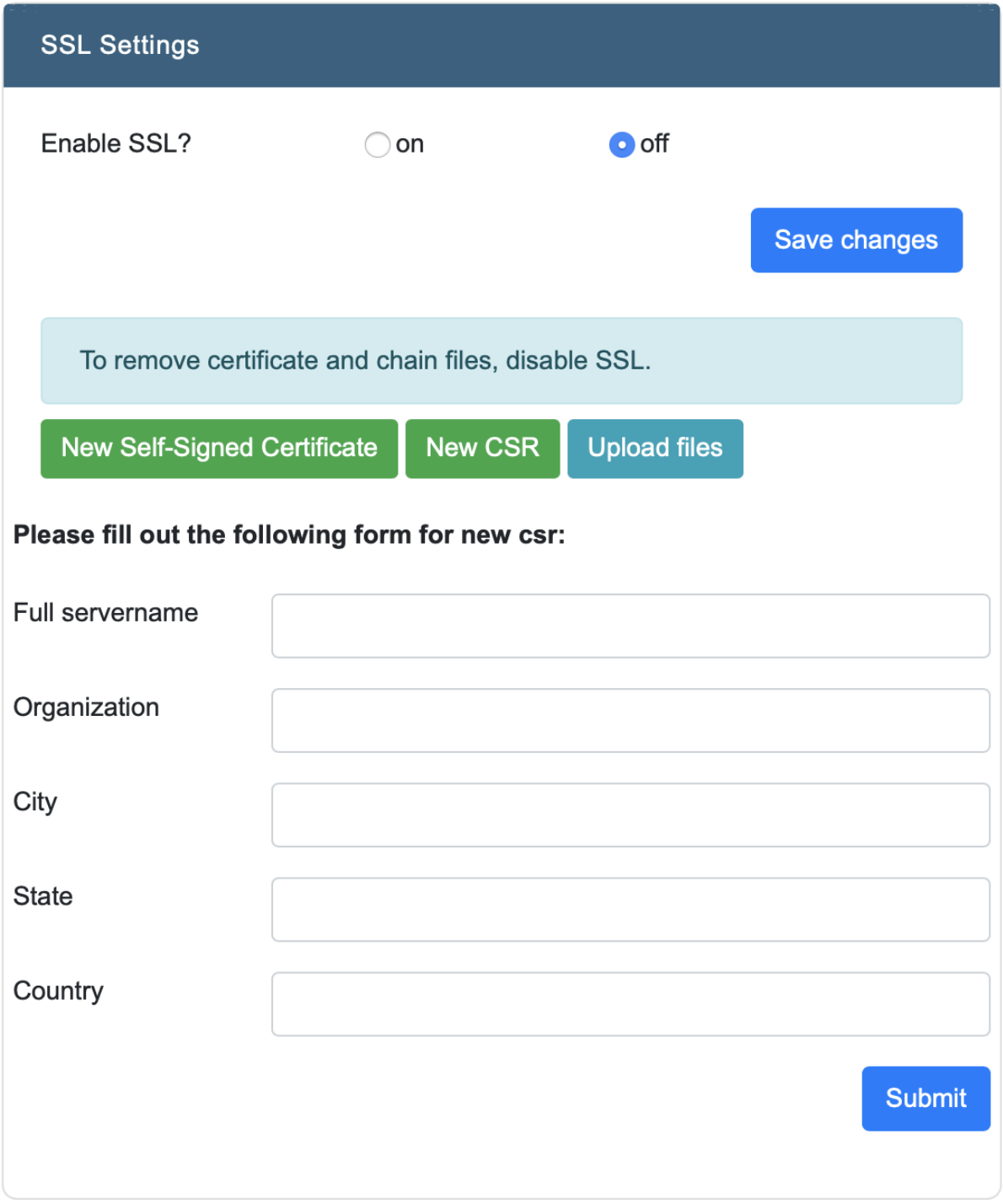
After filling in the fields, click “Submit”. The certificate is now active.

Important: Modern browsers may still notify to the user that the connection may not be secure when a self-signed

certificate is used.

Submitting a CSR

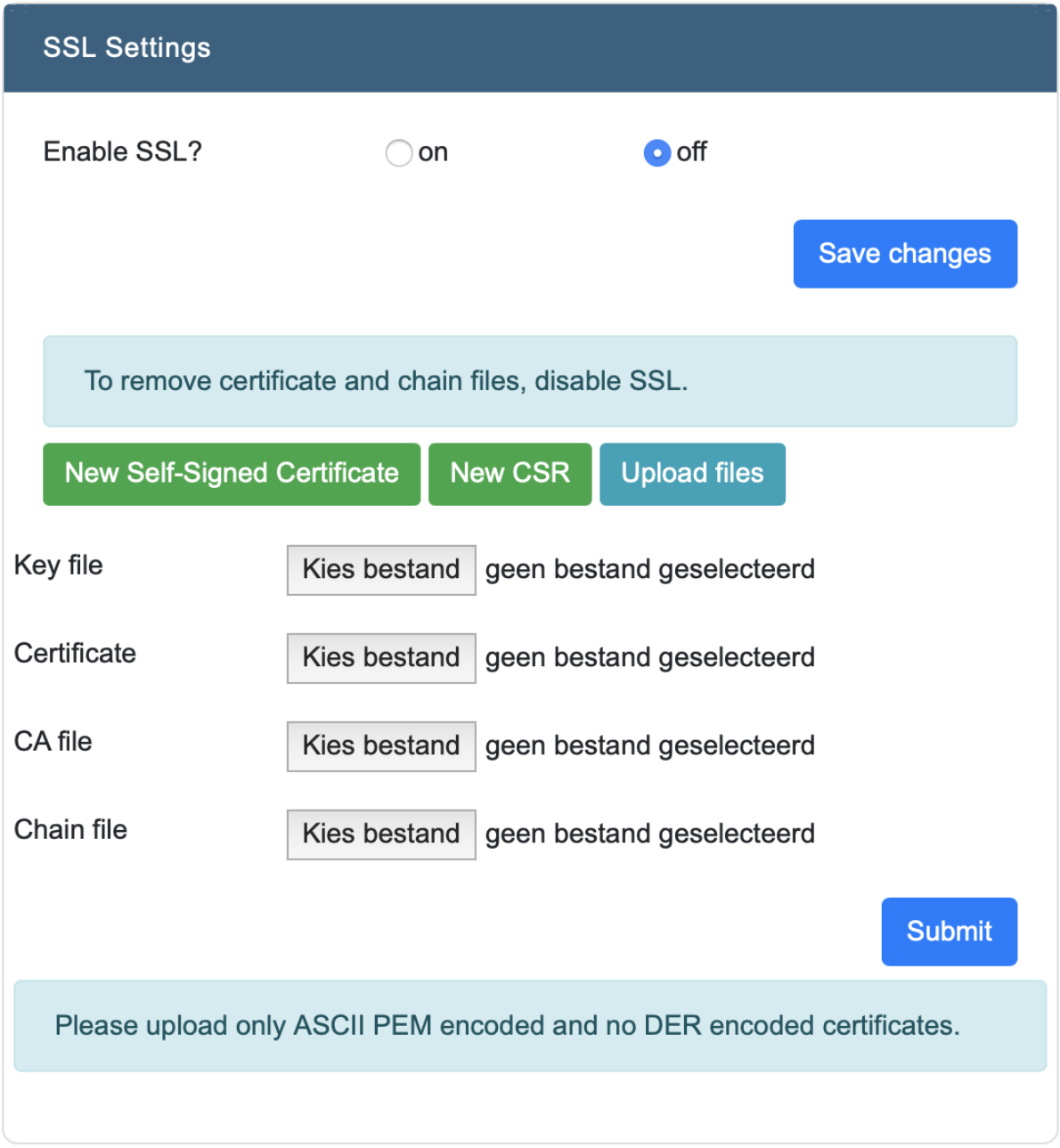
When clicking the button “New CSR”, you will be prompted to fill in some information:



After filling in the fields, click “Submit”.

Upload own certificate and keyfiles

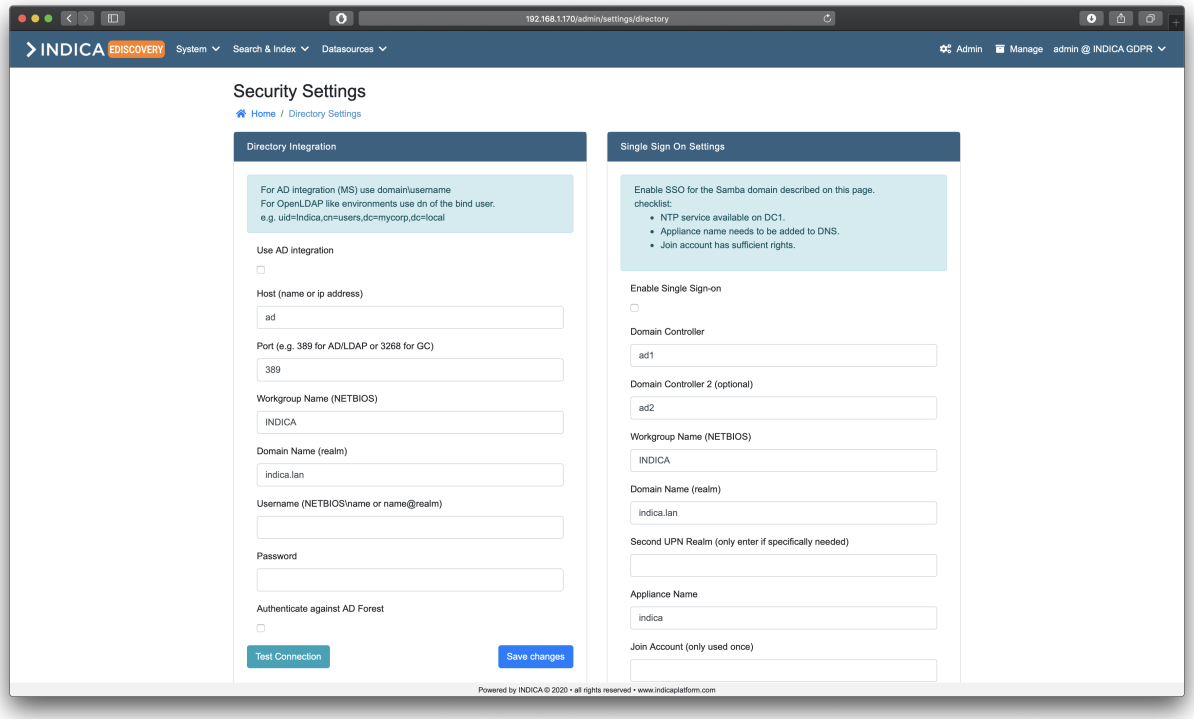
INDICA also provides the option to use your organisation’s own certificate and keyfiles. They can be uploaded after clicking the “Upload files” button. It wil then prompt the following files:



Upload the corresponding files, and then click “Submit” to activate.

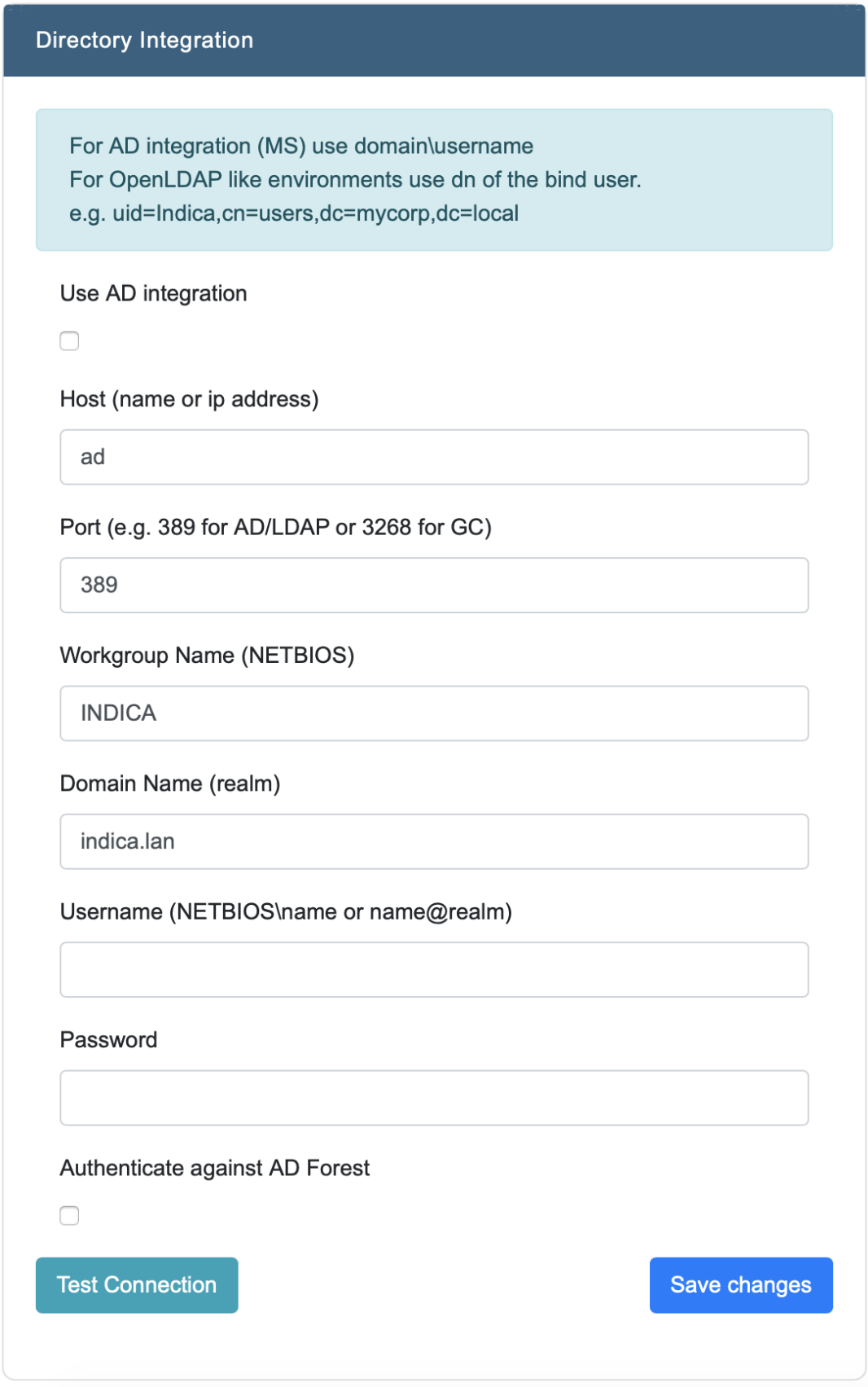
5.5 Security

Security Settings allows different ways to authenticate users. It is possible to setup an integration with an Active Directory (AD) and OpenLDAP, as well as Single Sign On (SSO). There is also a setting to disable reading the ACL from shares. This means that every user in INDICA can see all documents, regardless of Whether or not they are allowed to see the document on the original data source.



5.5.1 Enabling Active Directory integration

Active Directory integration allows for user and group management in the existing AD/LDAP environment. To make use of it, an account which can be used to 'bind' to the AD/LDAP is needed. Then, fill in the following information:


Directory Integration
For AD integration (MS) use domain\username
For OpenLDAP like environments use dn of the bind user.
e.g. uid=Indica,cn=users,dc=mycorp,dc=local
Use AD integration
Host (name or ip address)
ad
Port (e.g. 389 for AD/LDAP or 3268 for GC)
389
Workgroup Name (NETBIOS)
INDICA
Domain Name (realm)
indica.lan
Username (NETBIOS\name or name@realm)
Password
Authenticate against AD Forest
Test Connection
Save changes

- Check the 'Use AD integration' box
- Enter the hostname or IP. Usually, IP works better due to resolving the name.
- Enter the correct port
- Enter the Workgroup Name (NETBIOS)
- Enter the Domain Name (realm)
- Enter the Bind Username (with NETBIOS/realm notation)
- Enter the password of the Bind User
- (optional; if needed) Check the box 'Authenticate against AD Forest'

Then, click on 'Test Connection'. If everything is setup and entered correctly, INDICA should report back the first 10 AD groups it can find. If this is the case, click 'Save Changes' to store and apply the settings.

Now, the existing AD structure can be used to map existing AD groups to INDICA cases and roles.

This is done by setting up two (or more, depending on the *Roles and Permissions* setup) groups per INDICA case:

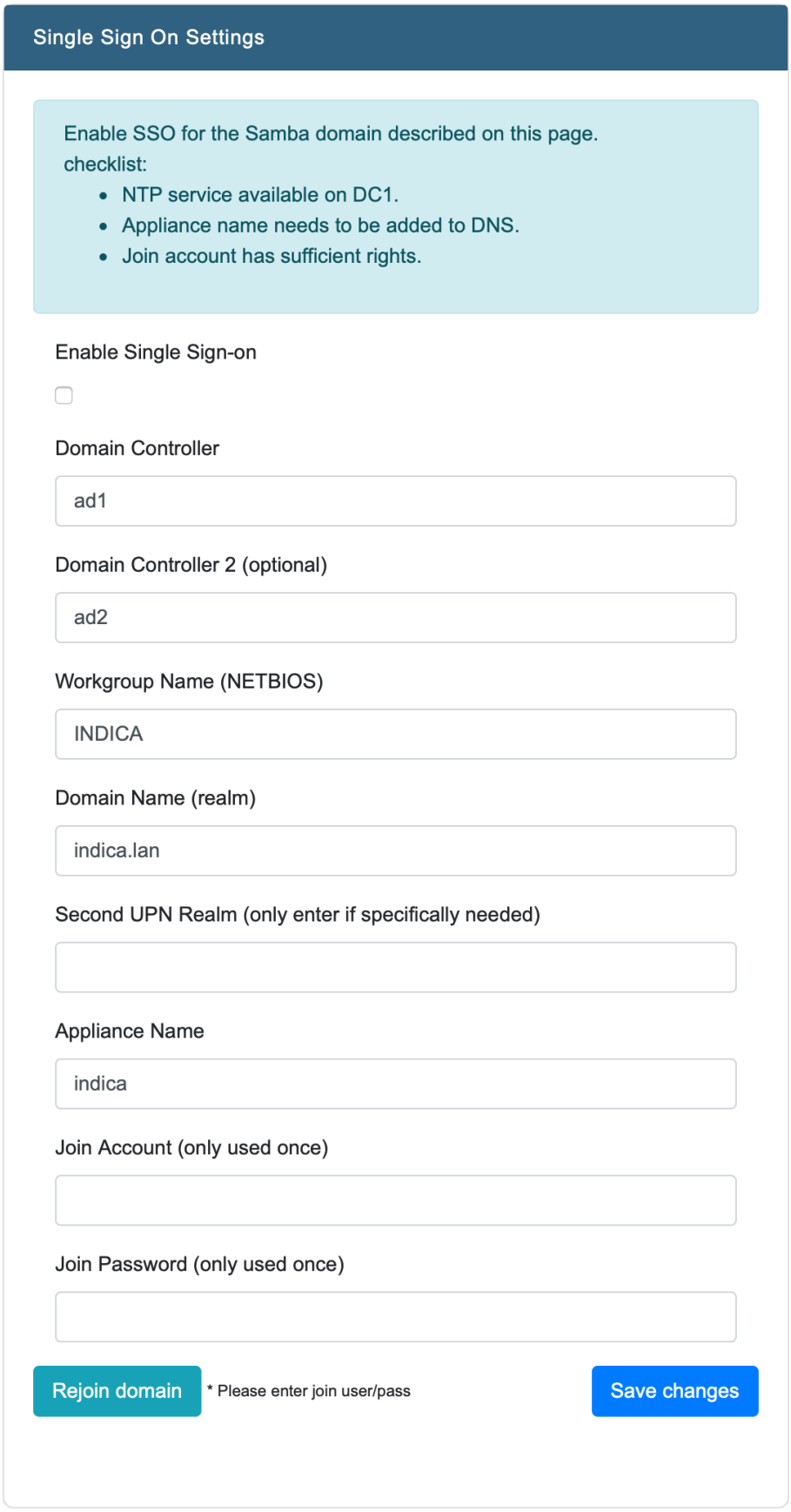
- A User-level access group.
- A Manager-level access group.

These two groups can then be linked to the corresponding INDICA case in the Case Settings. When all groups are set up, users can be added to or removed from the group at will. Edits to the groups in the AD/LDAP environment will be reflected in INDICA immediately.

Tip: More groups can be created and mapped to specific roles. See *Mapping AD groups to roles* for more information.

5.5.2 Enabling Single Sign-on

Single Sign-on(SSO) allows users to access the INDICA interface without having to bother with a username and password. INDICA will automatically sign in users if they are already authenticated.


Single Sign On Settings
Enable SSO for the Samba domain described on this page.
checklist:
NTP service available on DC1.
Appliance name needs to be added to DNS.
Join account has sufficient rights.
Enable Single Sign-on
Domain Controller
ad1
Domain Controller 2 (optional)
ad2
Workgroup Name (NETBIOS)
INDICA
Domain Name (realm)
indica.lan
Second UPN Realm (only enter if specifically needed)
Appliance Name
indica
Join Account (only used once)
Join Password (only used once)
Rejoin domain
* Please enter join user/pass
Save changes

SSO needs the following information:

- Domain Controller
- Domain Controller 2 (optional)
- Workgroup Name (NETBIOS)
- Domain Name (realm)
- Second UPN Realm (only enter if specifically needed)
- Appliance Name
- Join Account (only used once)
- Join Password (only used once)

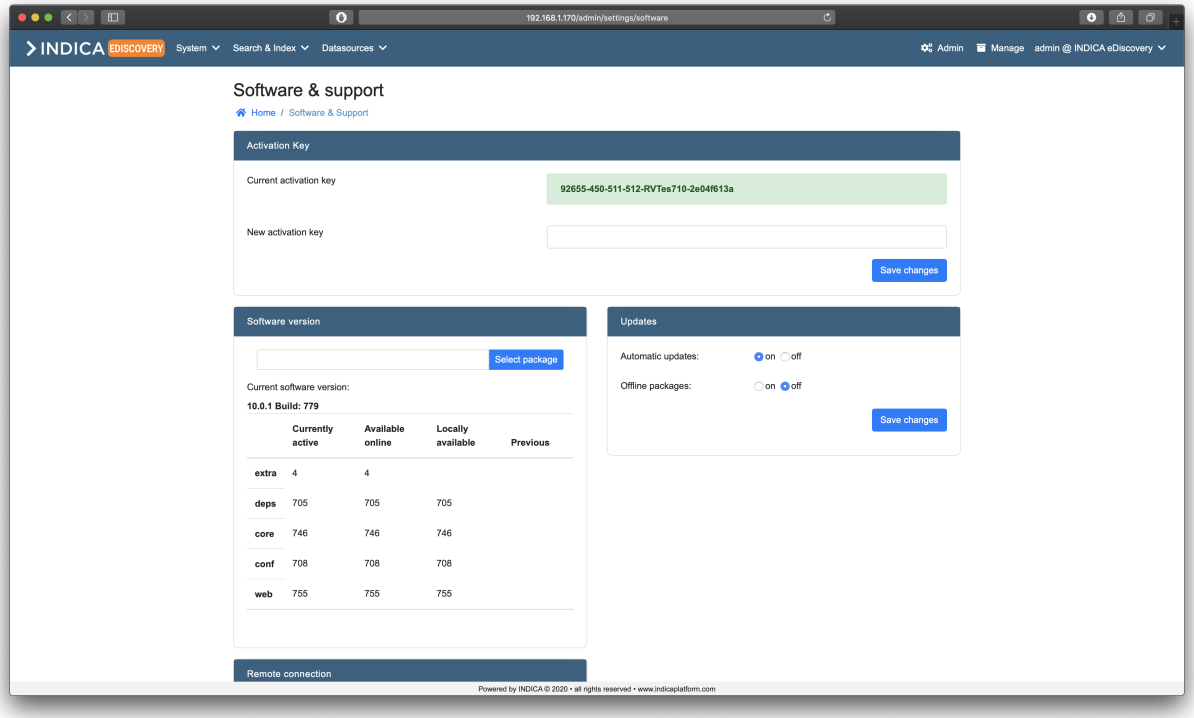
When using SSO, some actions may be required on different parts of the infrastructure (outside the INDICA appliance). Use the following list to make sure everything is set-up properly:

- NTP service available on DC1.
- Appliance name needs to be added to DNS.
- Join account has sufficient rights.

When everything is set-up properly, users should be able to access the INDICA interface without having to authenticate themselves.

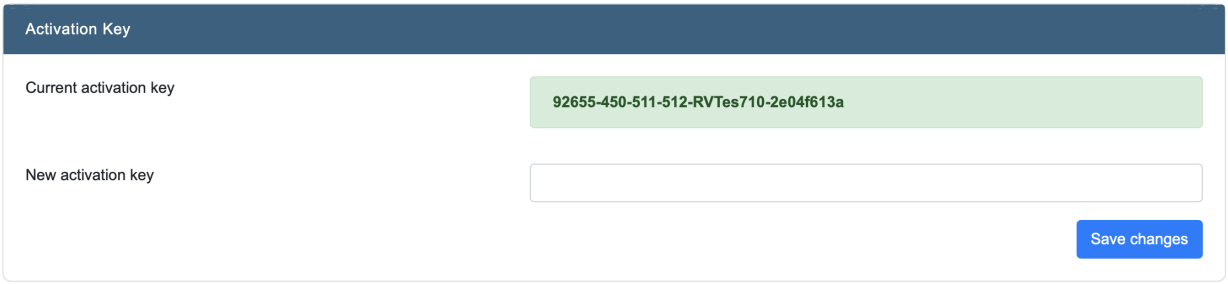
5.6 Software & Support

The Software & Support section helps managing the current license, software versions, and support connection.



5.6.1 Activation Key

Here you can view and edit the software key that is used to activate your INDICA appliance.

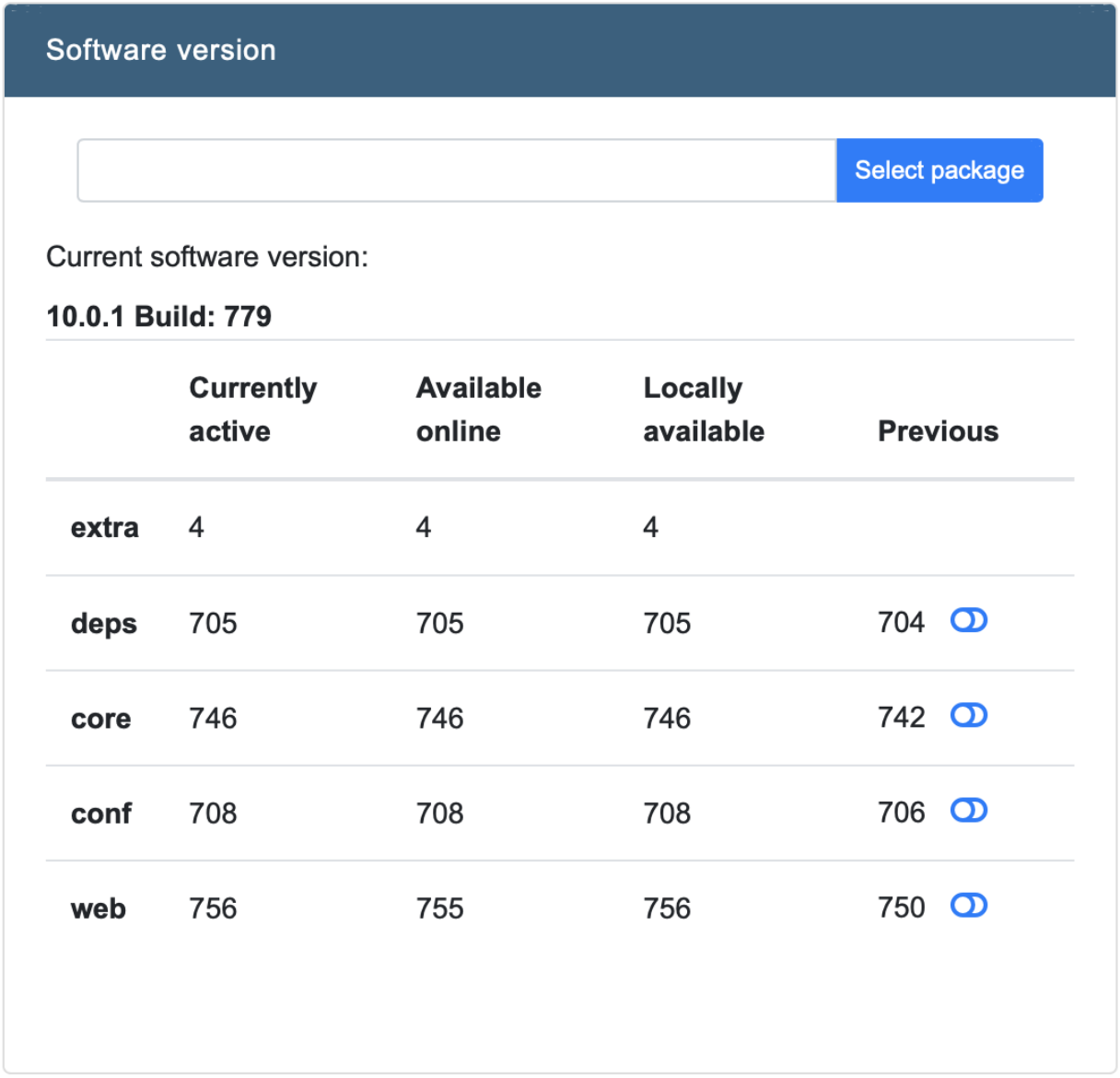


Note: Please keep in mind that the keys are not interchangable between appliances, as they are tied to the hardware as well. In case a new appliance needs to be activated, contact INDICA support to assist you.

5.6.2 Software version

This panel allows you to check the current software version.

It shows you the current version number, as well as the version numbers of all the packages as well. It is possible to update packages by hand, and if needed, roll back to previous versions. This can only be done if there was a previous version installed on the appliance.

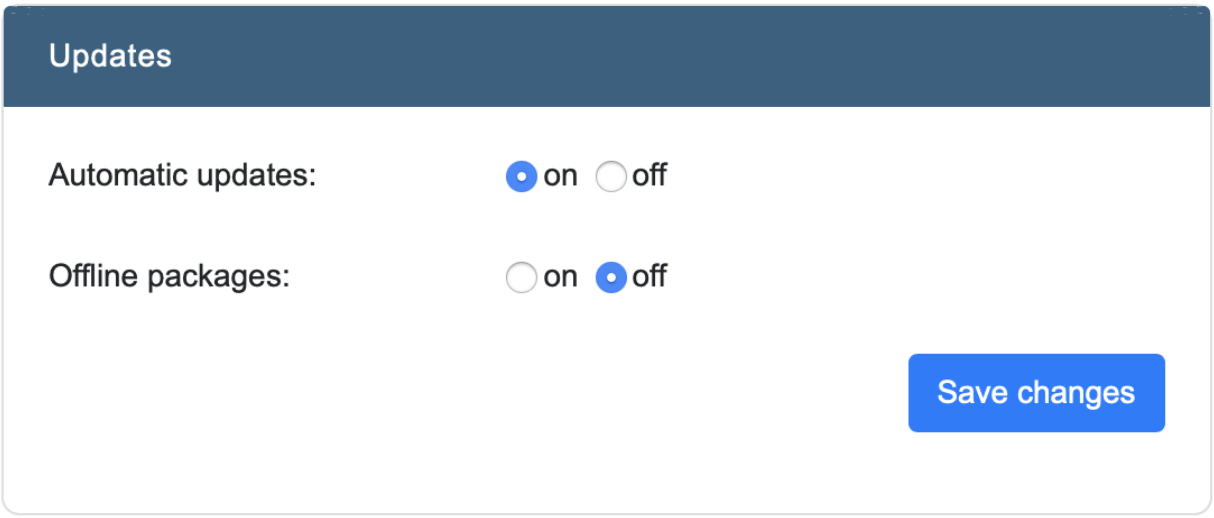


With the ‘Select package’ option, a package can be uploaded to the appliance. INDICA will automatically recognise the type of package uploaded, and it’s version and add the package to the overview if it has a newer version than currently active.

Clicking the switch in the ‘Previous’ column rolls back the version of active package.

5.6.3 Updates

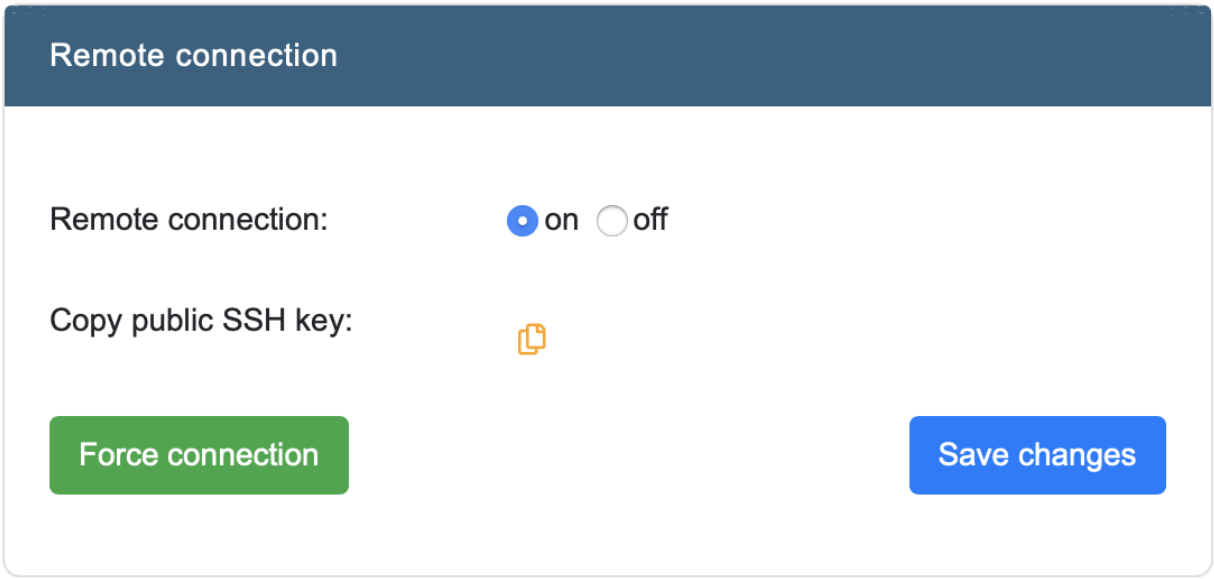
This panel allows you to toggle automatic updates on or off. You can also specify Whether offline packages should be used or not.



5.6.4 Remote Connection

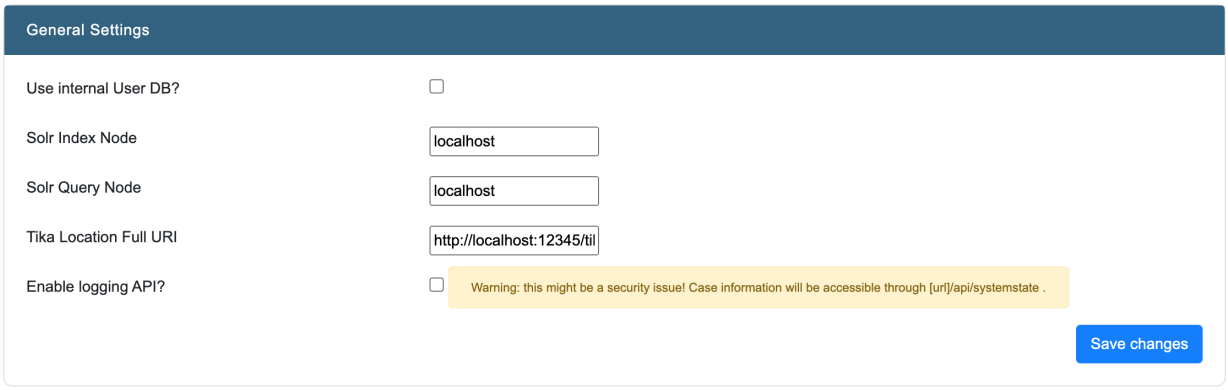
This option allows appliance administrators to toggle remote connections. The remote connections are used by INDICA engineers to remotely troubleshoot and support your appliance. It is advised to leave this setting turned on.

Here you can also copy the public SSH key of the appliance, and force a remote connection to the INDICA server.



5.7 General Settings

This section of the settings allow you to configure a few things:

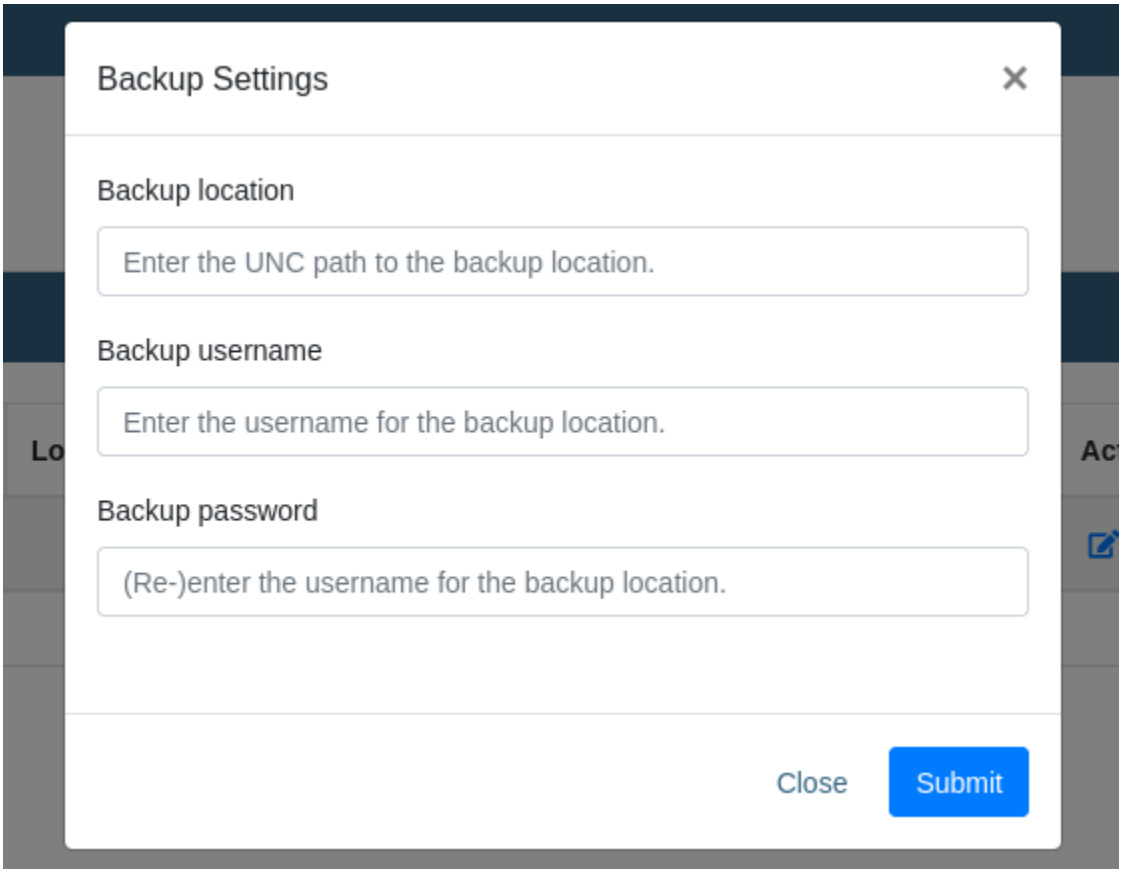


- **Use internal user DB**
  When not connected to an external authentication service, it is possible to use the internal authentication database.
- **Solr Index Node**
  Location of the Solr index node. Only to be edited when running in a cluster Be careful, as changing this setting may lead to an unstable/broken cluster
- **Solr Query Node**
  Location of the Solr query node. Only to be edited when running in a cluster Be careful, as changing this setting may lead to an unstable/broken cluster
- **Tika Location Full URI**
  If Apache Tika is hosted externally, then edit this value to point the indexer to the correct location
- **Enable logging API**
  Enable logging API so logging can be used with external programs/services

> **Warning:** This might be a security issue! Case information will be accessible externally

5.8 Backup Settings

This page allows you to setup the backup location. It needs the following information:

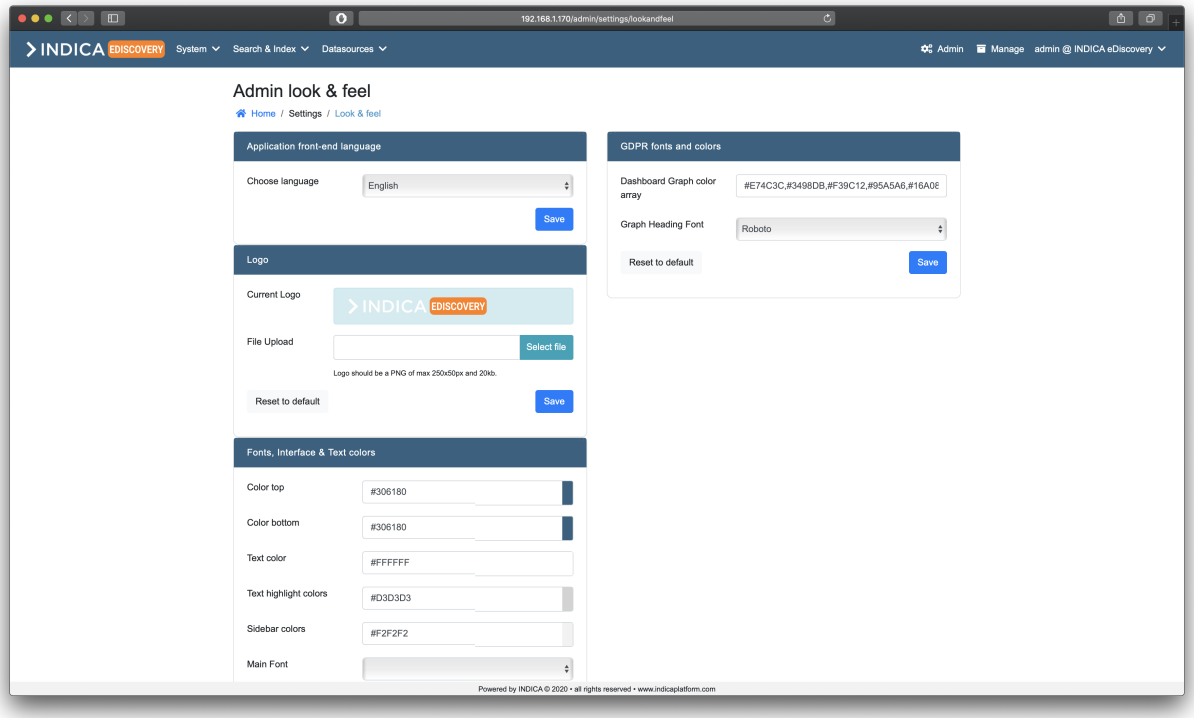


- **UNC path of the backup location**
  Location where INDICA can read and write it's backups to
- **Backup username**
  Username for INDICA to authenticate on the backup location
- **Backup password**
  Password of the corresponding user that INDICA can use to authenticate on the backup location

5.9 Look and Feel

INDICA allows you to apply your custom Look and Feel to the appliance. This can be done on two levels:

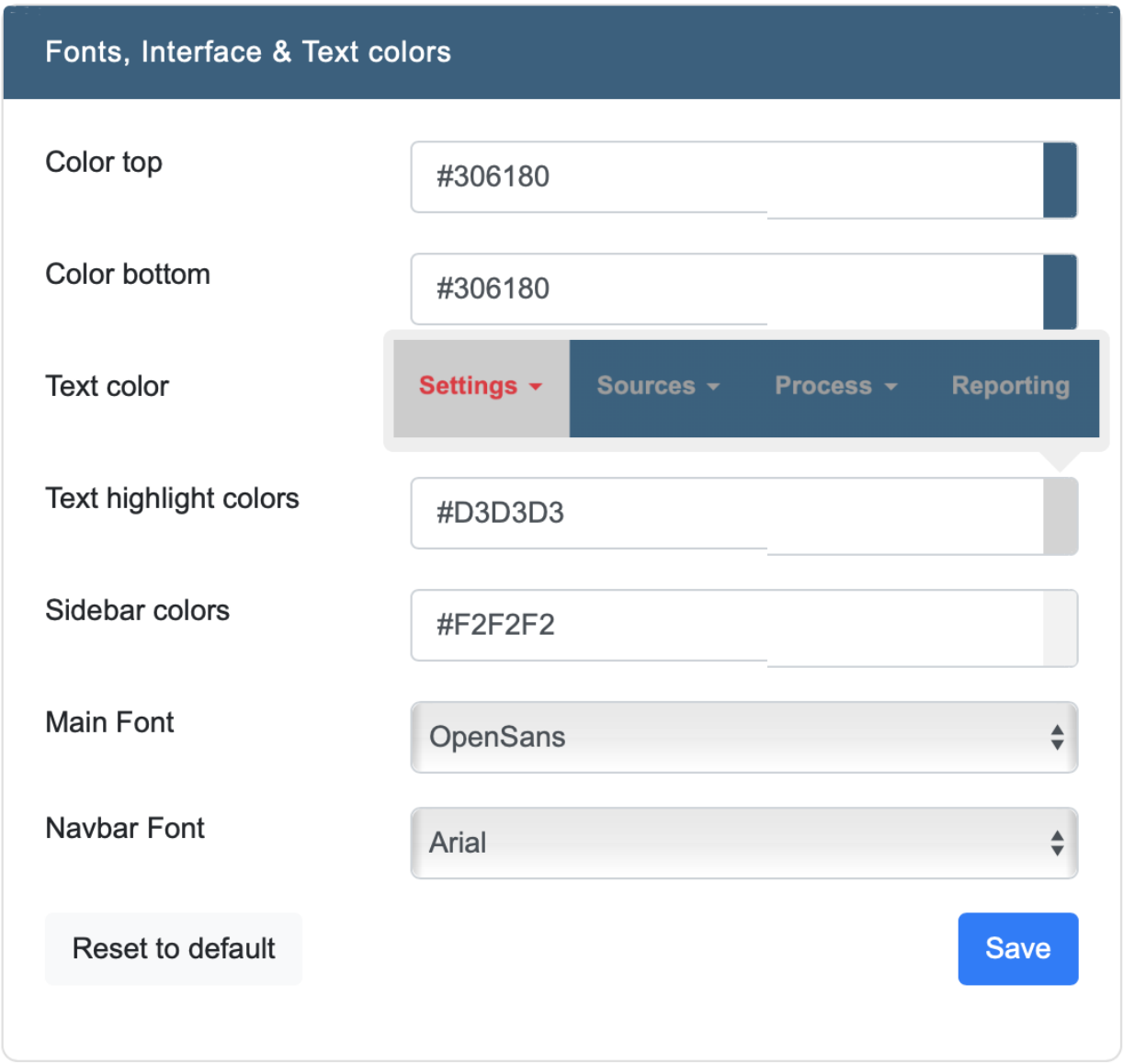
- **Appliance level**
  This is configured in the admin interface and applies to the admin section
- **Case level**
  This is configured in the case management interface and applies to that case only This allowes for customised look and feel in different cases



The settings allow you to change the following:

- **Application front-end language**
  Supported languages: Dutch, English, French, German
- **Logo**
  PNG format, 250x50 pixels, max 20kb
- **Fonts, interface and text colors**
  Allows changing the colors and fonts of the main interface
- **GDPR fonts and colors**
  Allows changing the colors and fonts of the GDPR interface

When hovering the mouse over the color pickers, a preview is shown:



All values can be changed individually, and reset to default if needed.

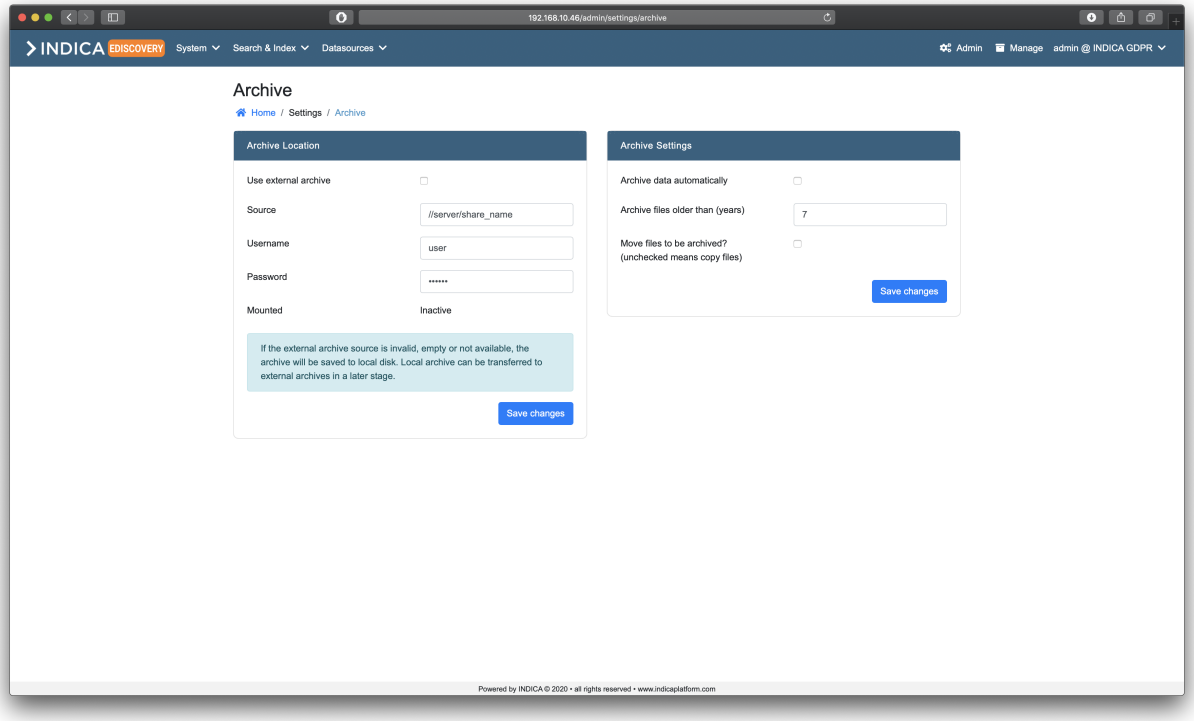
CHAPTER

SIX

DATASOURCES SETTINGS

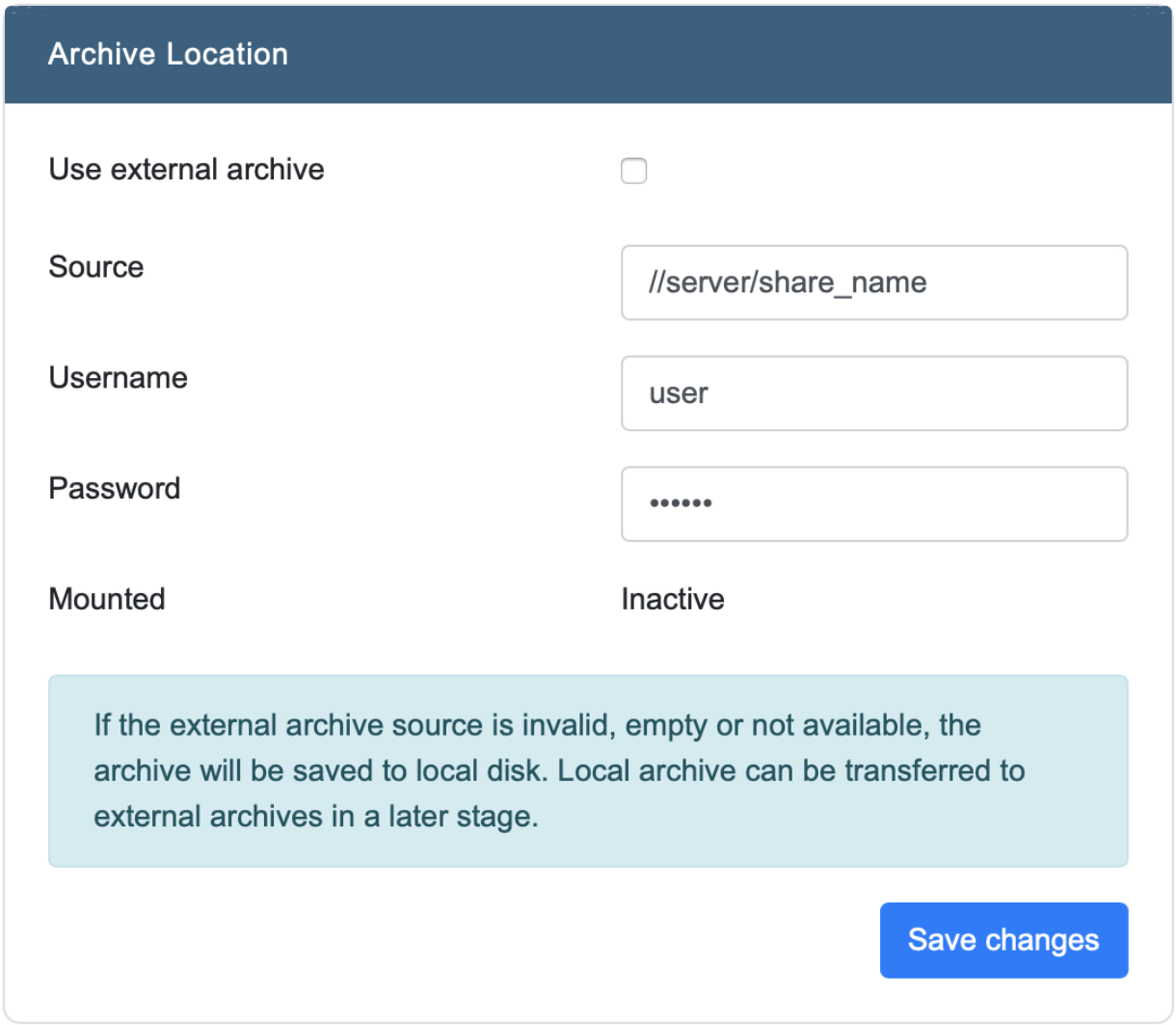
6.1 Archive Settings

In the Datasources page you can enable/disable the archiving function and choose where to locate your archived files. Email archiving is automatically set up when configuring email connectivity. The default archiving location is on the internal storage, this can be changed to an external CIFS or SMB source. Automatic archiving can also be enabled in this page. All archived cases will be list down on Cases page.



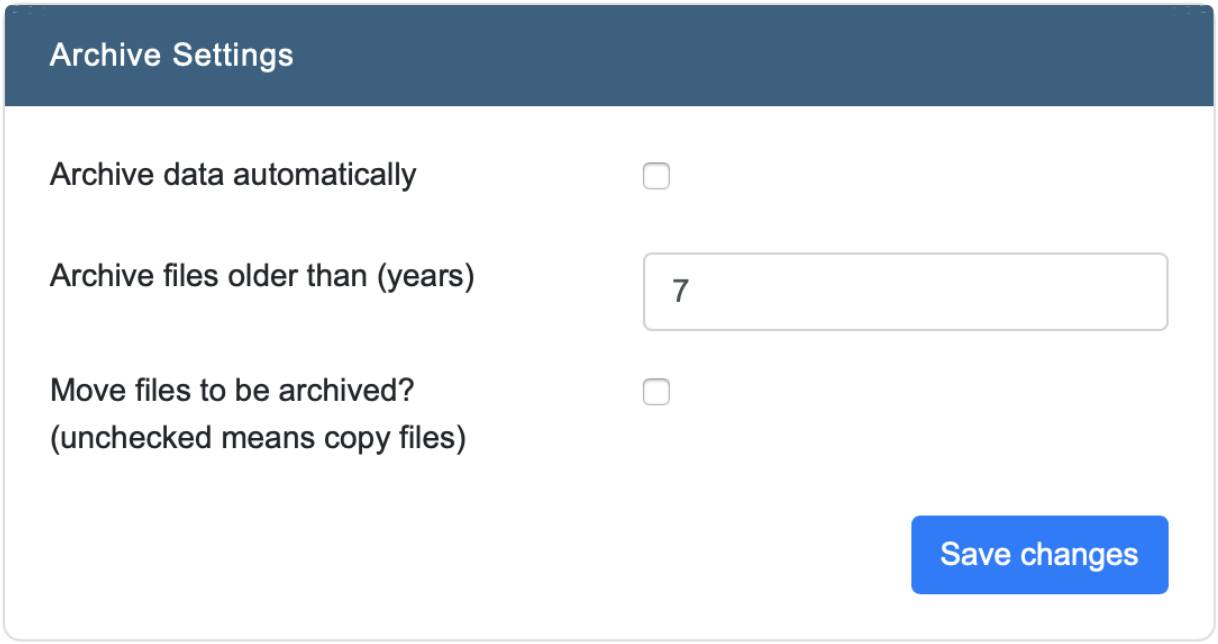
6.1.1 Archive Location

To archive to an external location, please enter the location in UNC format with forward slashes: //server/share_name and its corresponding username and password. Please make sure the archiving credentials contain a user that has "write and create" rights on the share. If you want to archive to other shared storage than CIFS or SMB, please contact your reseller. When setting up an external source with incorrect credentials, the appliance will default the archiving to local storage.



6.1.2 Archive Settings

Choose whether to automatically archive the files by selecting 'yes' or 'no'. INDICA will archive files older than the chosen years (default: 7). If you allow INDICA to move the files, they will be removed from the original location and moved to the archive location.



CHAPTER

SEVEN

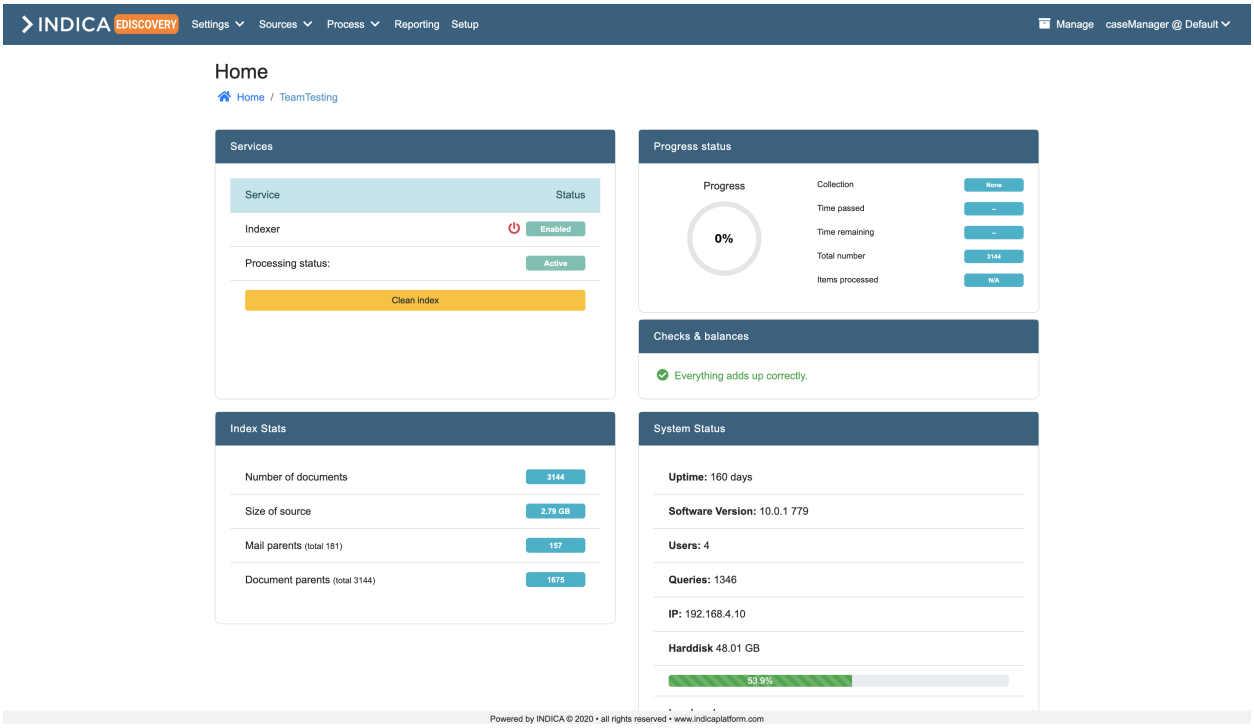
MANAGER PAGE OVERVIEW

Table of Contents

- *Manager Page Overview*
  - *Services*
  - *Progress Status*
  - *Checks & Balances*
  - *Index Stats*
  - *System Status*
  - *Workflow*
  - *Tag Info*
  - *Shares*
  - *File Distribution Info*
  - *File Processing Info*

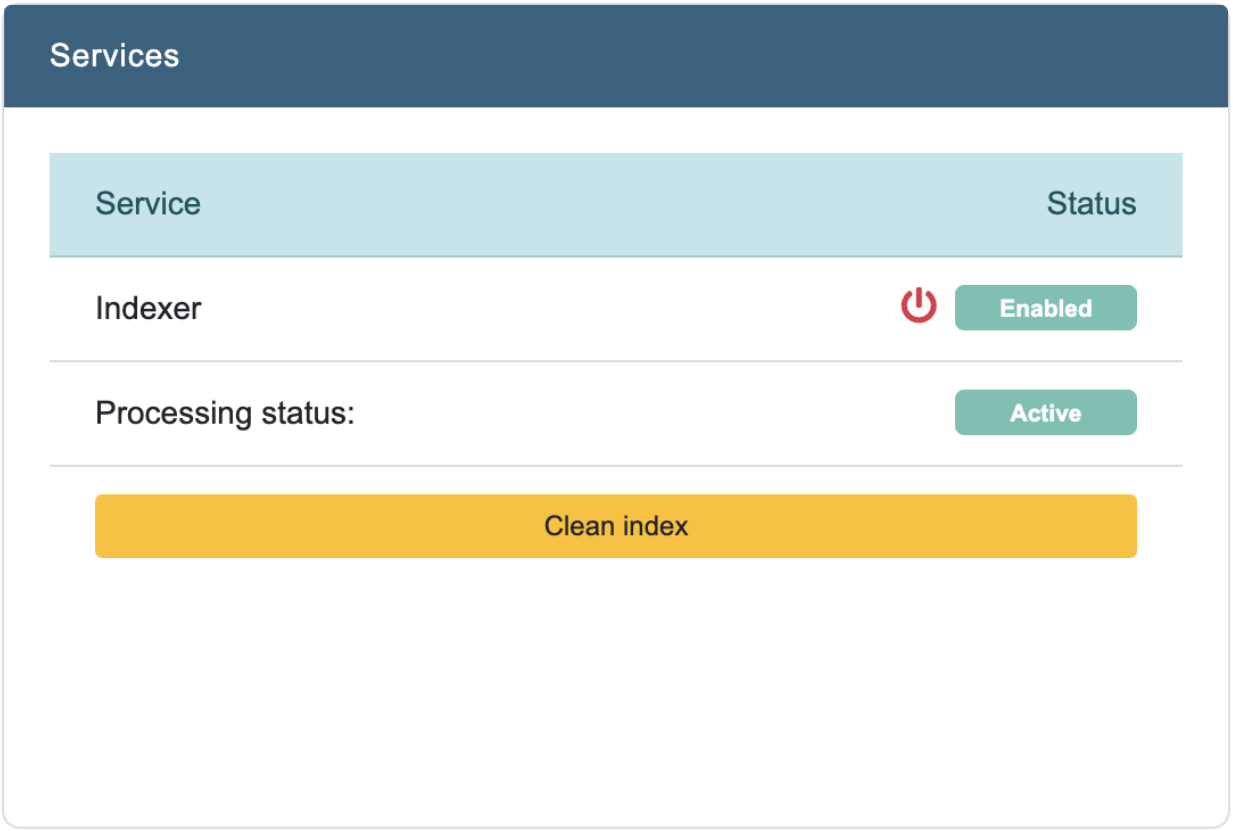
This panel shows the status of the system services, the indexer and spider status, system information and the enabled modules.

The manager page is unique for each case. The results, counts, balances, and content of the page can be different depending on the currently active case.



7.1 Services

This panel shows the services viewable for case managers.

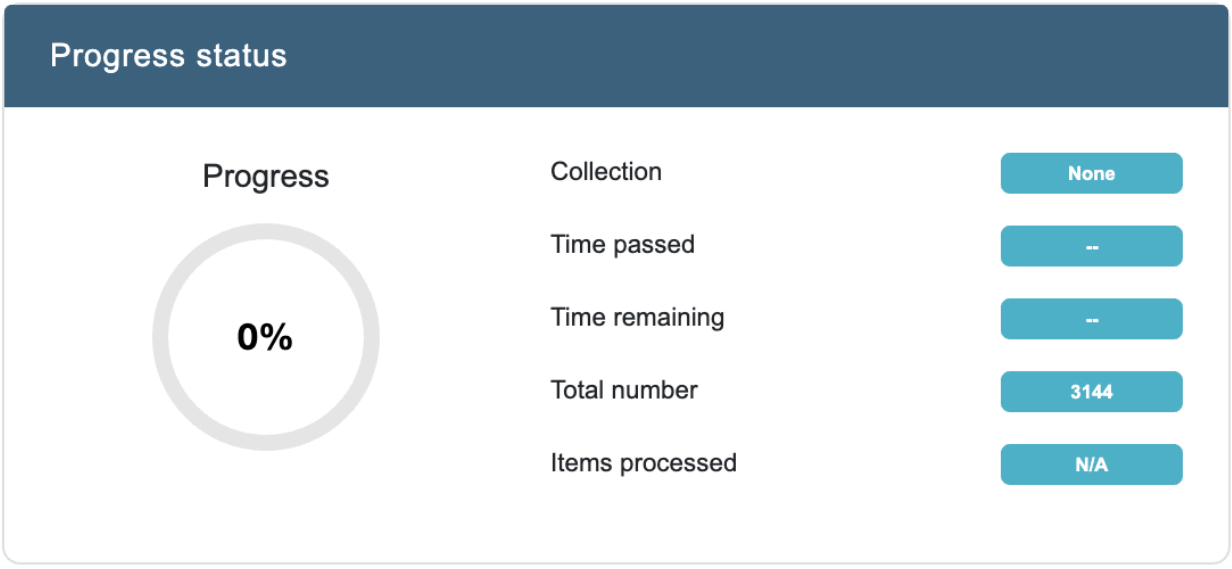


- **Indexer**
  This displays whether the the indexer is enabled or not. Here it's also possible to stop or restart the indexer.
- **Processing status**
  This displays whether the indexer is busy (active) processing or not.
- **Clean index**
  It's possible to clean the index within a case.

> **Caution:** If you decide to clean the index please keep in mind that the whole case will need to be reindexed and that all progress you've done so far will be lost. This action is irreversible, so be cautious when using this feature.

7.2 Progress Status

This panel displays the general status of the indexer within the case you're viewing.



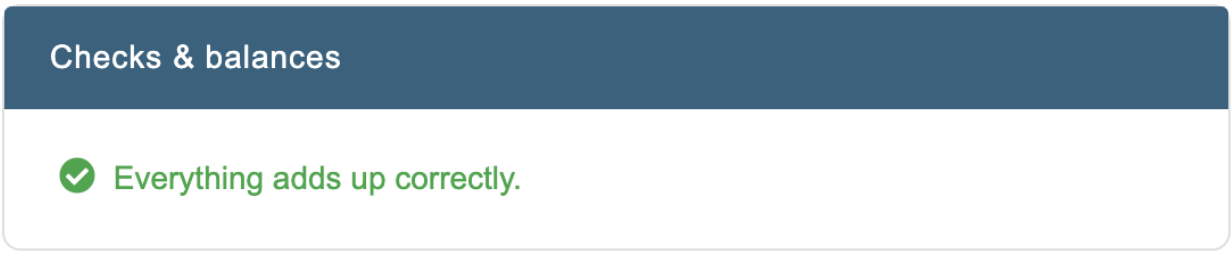
It displays the following statistics:

- **Progress**
  Percentage of collection index progress.
- **Collection**
  Name of the collection that is currently being indexed.
- **Time passed**
  Time passed since starting indexation on this collection.
- **Time remaining**
  Estimate on the time needed to fully index the current collection.
- **Total number**
  Number of items in this collection that need to be indexed.
- **Items processed**
  Number of items in this collection that have been indexed.

For more information about collections, visit the section *Collections*.

7.3 Checks & Balances

Shows whether the indexed documents adds up correctly. There are three different statuses that can be shown:

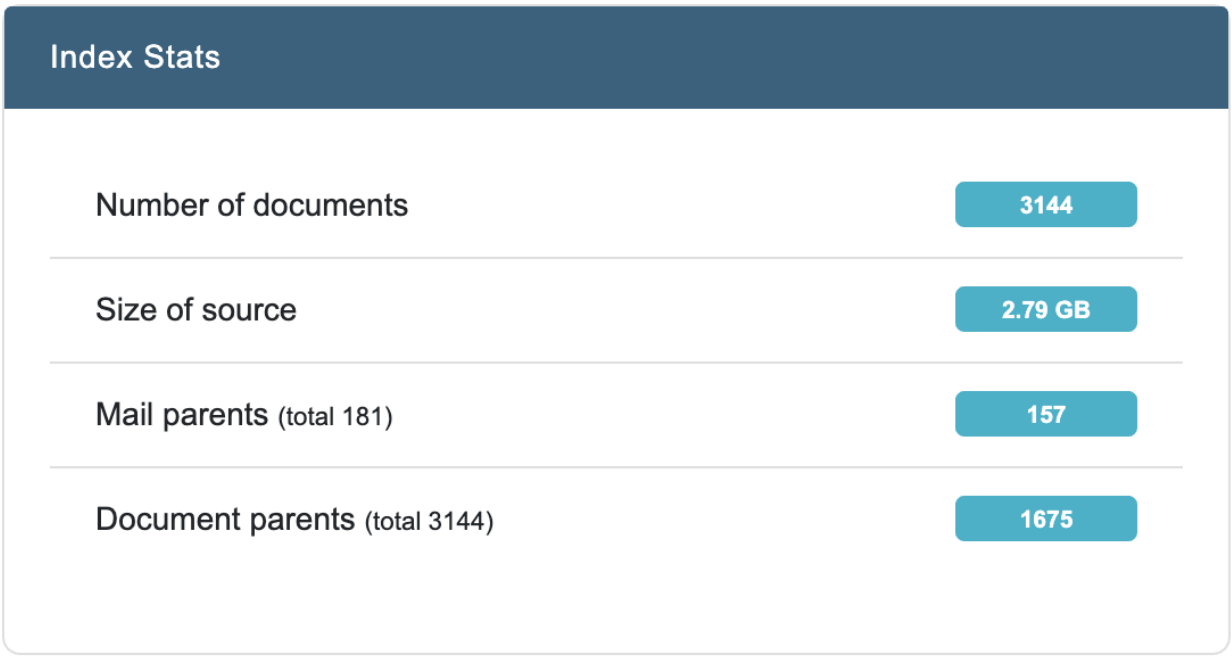


- **Everything adds up correctly**
  This means that the sum of documents in the case match with what the indexer found.
- **There are missing documents**
  This means that the sum of documents in the case do not match with what the indexer found. In this case there are missing documents and you can investigate these by clicking the link provided.
- **Sum of docs in workflow is different from total in index**
  This means that the sum of documents in the case do not match with what the indexer found.

Important: Please keep in mind that while the case is still indexing the checks & balances status will vary. After the indexation process of your case is completely done, everything will usually add up correctly.

7.4 Index Stats

Index stats displays some basic information about the indexed source(s).

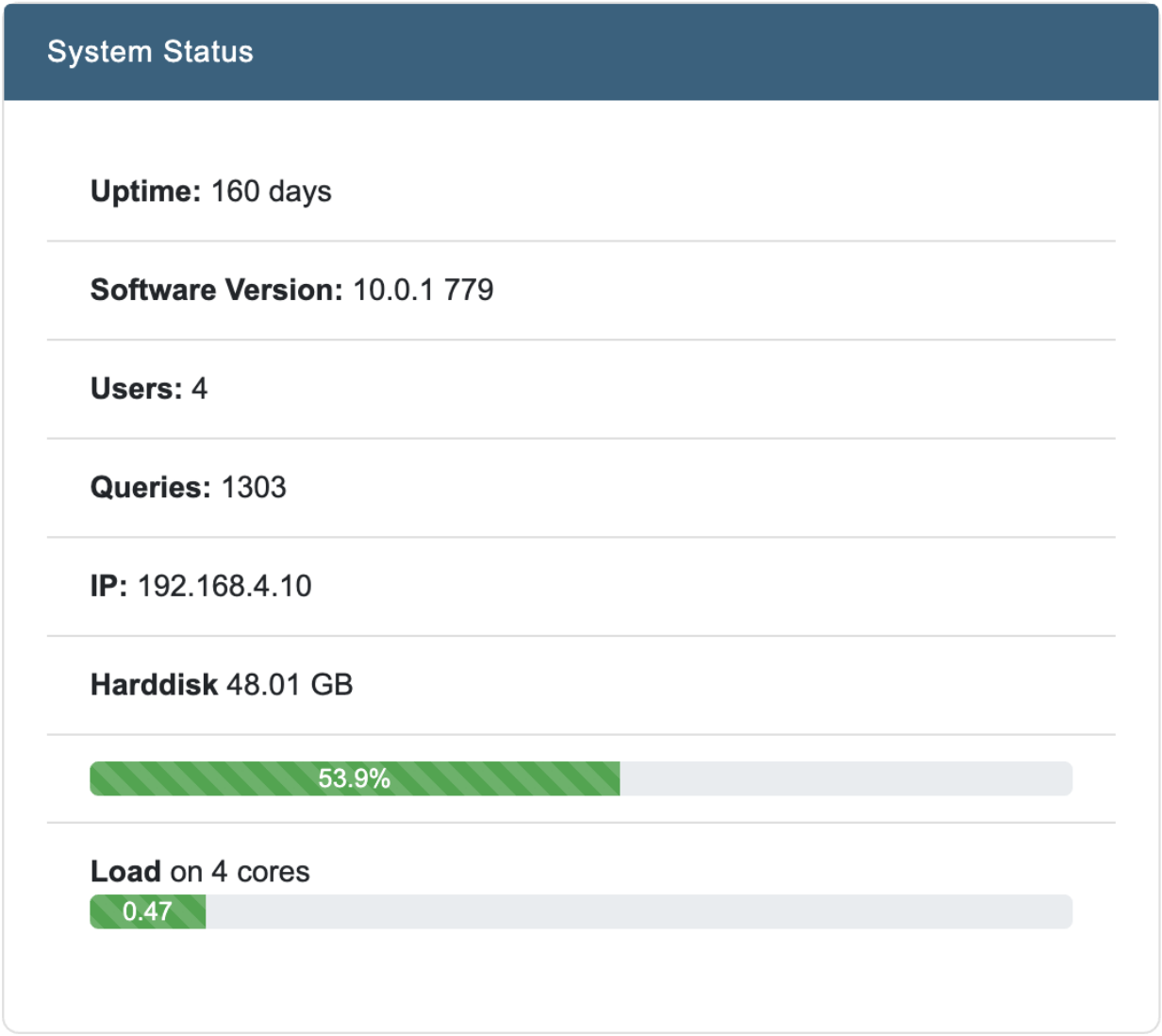


- **Number of documents**
  Shows the number of indexed documents.
- **Counted number of documents**
  This only shows up when there are missing documents and displays the total count of documents.

- **Missing documents**
  This only shows up when there are missing documents and provides a link to see the missing documents.
- **Not indexed documents**
  This only shows up when there are documents that aren't indexed. Shows the number of not indexed documents.
- **Size of source**
  Shows the size of the source(s) connected in MB.
- **Mail parents**
  Shows the total number of indexed mail parents.
- **Document parents**
  Shows the total number of indexed documents.

7.5 System Status

System status displays some basic system statistics and the appliances current workload.



It displays the following statistics:

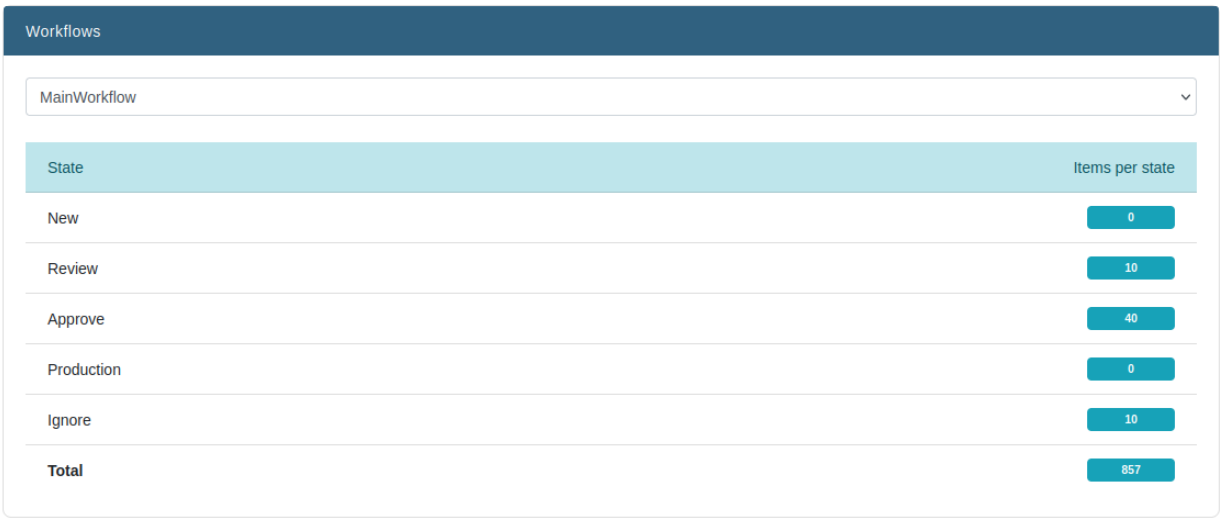
- **Uptime Current**
  Uptime of the appliance.
- **Software Version**
  Current software version.
- **Users**
  Number of users in the system.
- **Queries**
  Total number of queries ran on the appliance. This includes system queries as well.
- **IP Current**
  IP address of the appliance.
- **Harddisk Current**
  Disk usage of the appliance.

  Note: External mounts are not taken into account.

- **Load**
  Current workload on the appliance.

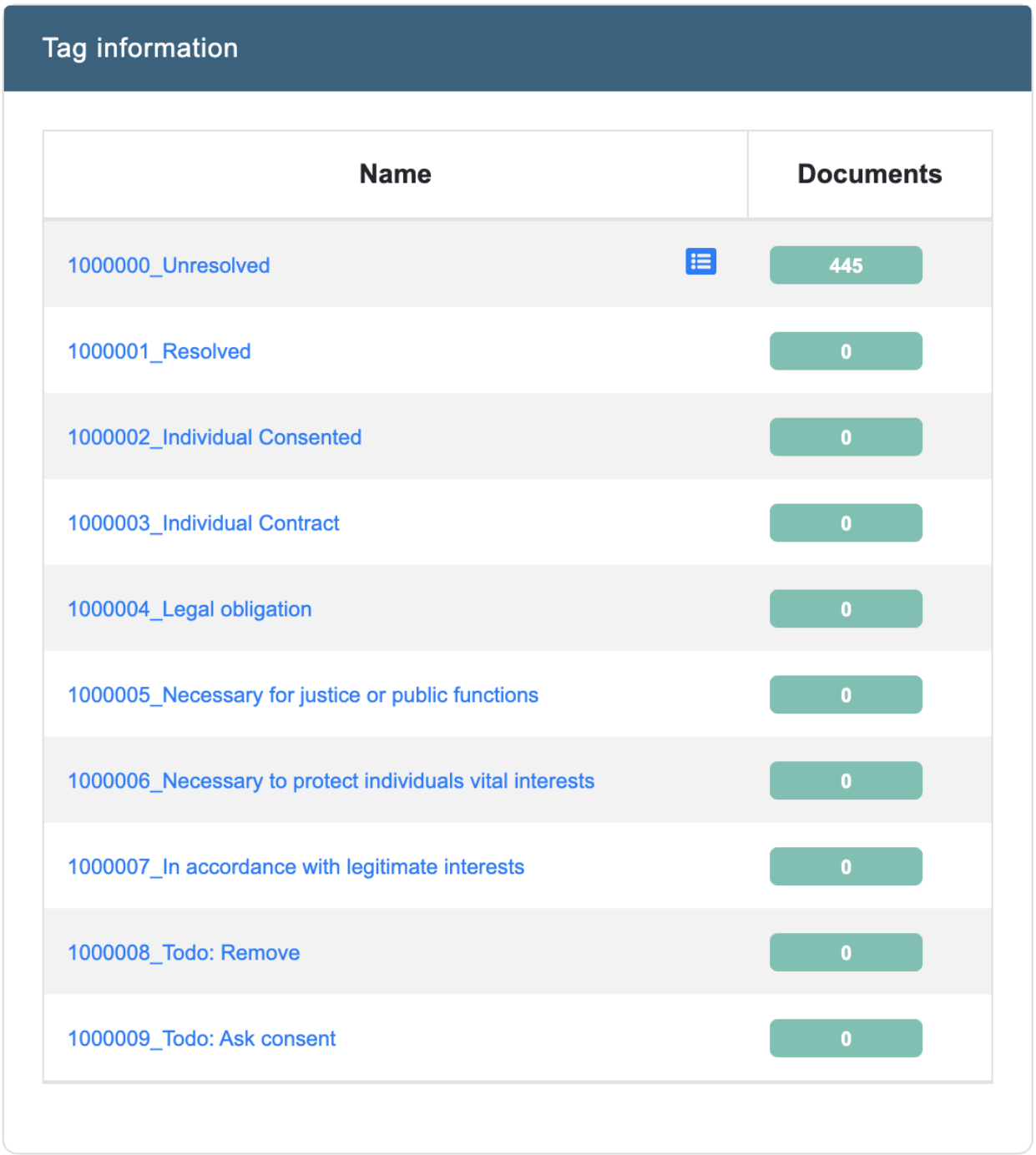
7.6 Workflow

If workflow is enabled, this panel shows the workflow states and how many documents are in the different available states. The image below is an example of our standard workflow, but this might look different depending on how the workflow within your case is set up.



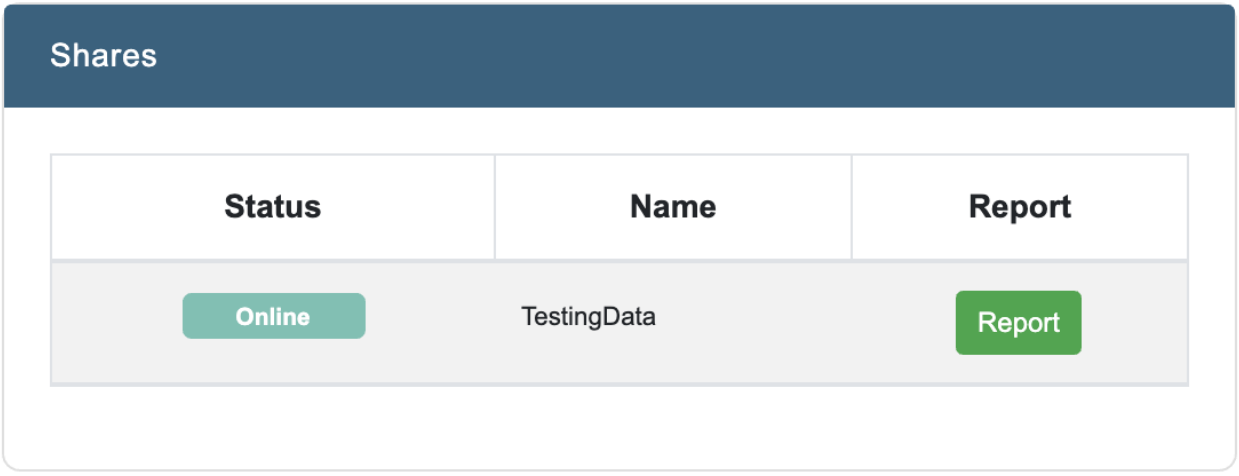
7.7 Tag Info

Within this panel an overview is given of the available tags and how many documents are tagged with the specified tag. If there are tagged documents you can also download a CSV list (up until max 50K documents within that tag).

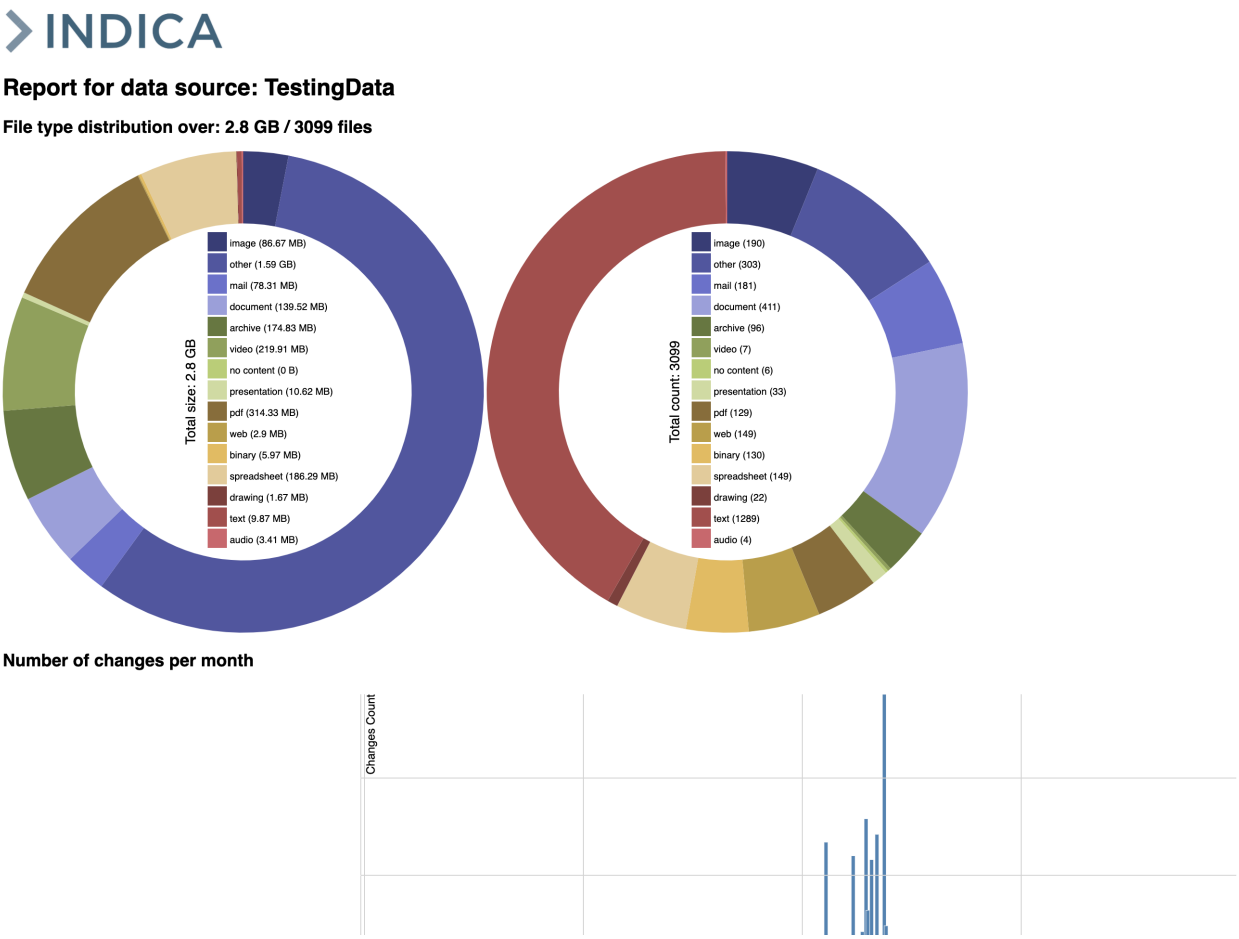


7.8 Shares

This panel shows an overview of the connected shares.



It's also possible to view a report of a share by clicking the button report. This will bring you to a new page with a report about that source. (See part of an example in the image below)



7.9 File Distribution Info

Within the table in this panel an overview is given over the different types of documents. This also shows the count of documents found within that type and how big the size is of all those documents from that type.

File Distribution info

type	count	size
image	190	86.67 MB
other	303	1.58 GB
mail	181	78.30 MB
document	411	139.52 MB
archive	96	174.82 MB
video	7	219.90 MB
ziesodatatab	45	45 B
no content	6	0 B
presentation	33	10.61 MB
pdf	129	314.32 MB
web	149	2.89 MB
binary	130	5.97 MB
spreadsheet	149	186.29 MB
drawing	22	1.66 MB
text	1289	9.87 MB
audio	4	3.41 MB
totals	**3144**	**2.79 GB**

7.10 File Processing Info

Within the table in this panel an overview is given over the different types of processed documents. This also shows the total count and you're able to download a CSV file with a list of the processed documents (up until max 500K).

File Processing Info

desc	count
Processed items	3099
Source: shares	1534
Compressed	1373
RAW text extracted	334
Source: PST/OST	141
No data extracted	137
Attachment	51
Image PDF	50
OCR PDF	50
GDPR generated	45
Structured content	45
OCR	23
Possibly Encrypted	22
Oversized container	3
Unsupported archive	3
PST Warning	1

Note: downloading lists over 500,000 docs is not possible. Long lists are impossible to import in Excel. List download will take a while.

CHAPTER

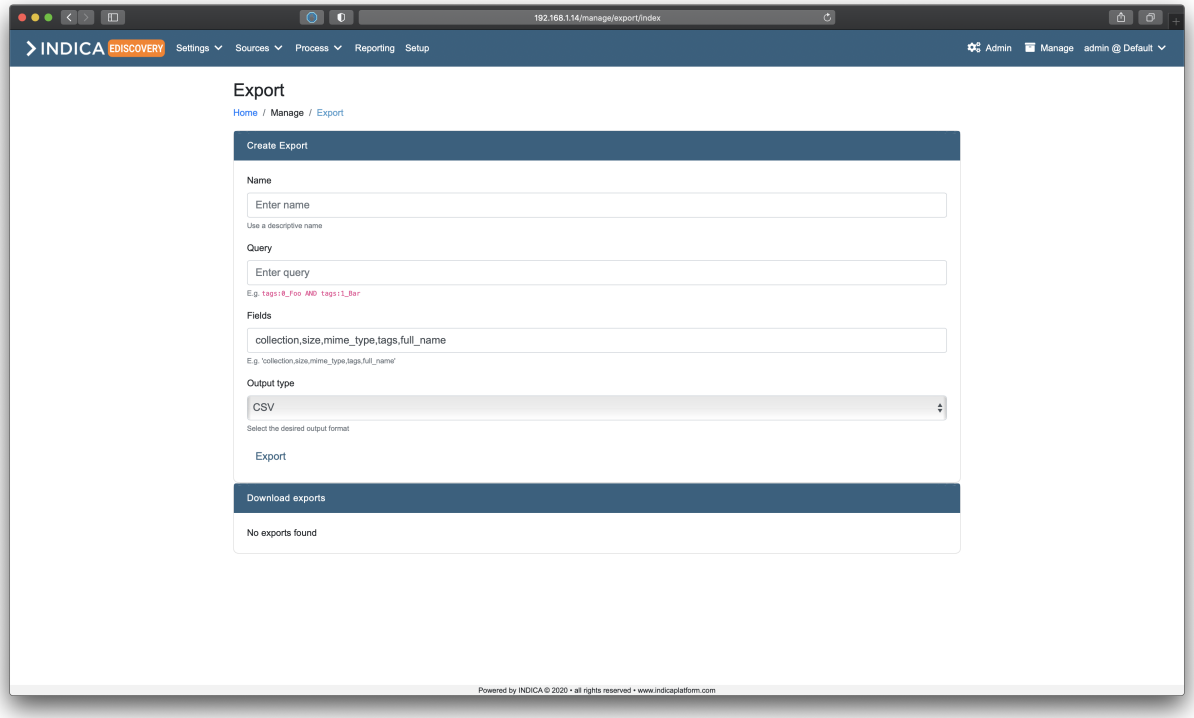
EIGHT

CASE MANAGER EXPORTING

Table of Contents

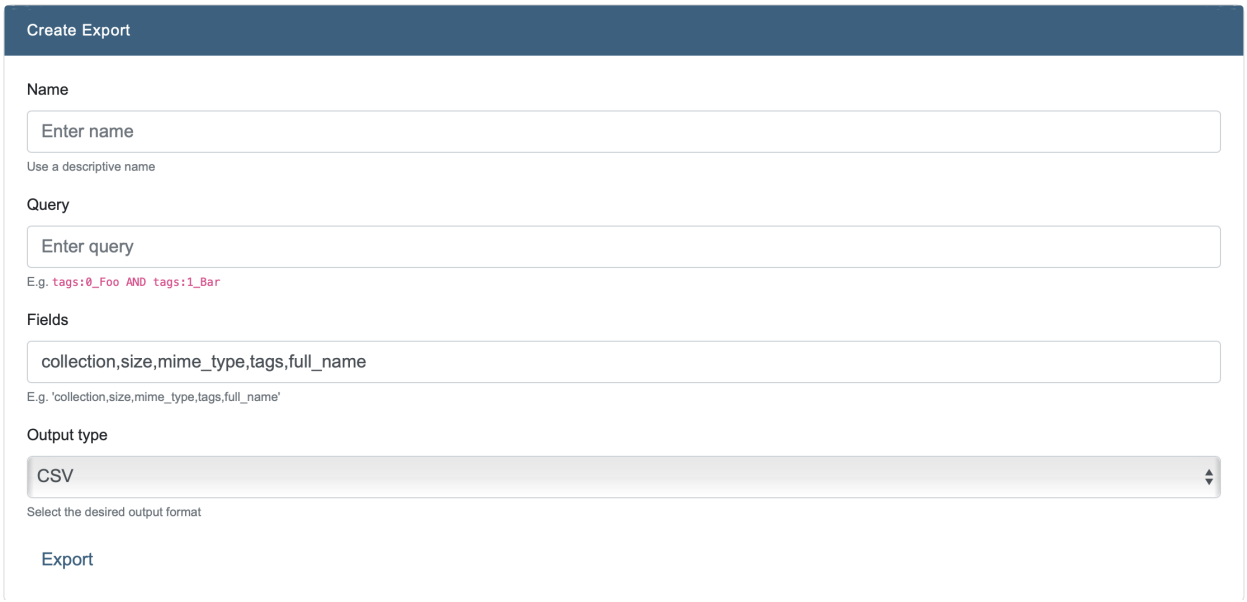
- *Case Manager Exporting*
  - *Create Export*
    - *Creating a PDF Export*
  - *Downloading an export*

Exporting query results to CSV or XLS(X) makes it possible to use INDICA to create lists that can be used for several actions. It allows case managers to export counts, tags, tokens, etc. These can be used to create statistics and dashboarding. This also opens up the possibility to export file lists and to take (automated) actions based on the export.



8.1 Create Export

Creating an export job is done by filling in the following fields:



- **Name**

  Name of the export. This name will be used as filename.

- **Query**

  Query that will be used to create the export.
  Any query that works on the Search page will work here as well.
  It is advised to create the query on the Search page and copy-paste it here.

- **Fields**

  List of all fields that need to be exported.
  The order of the filled in fields will be reflected in the export.

- **Output Type**

  Choose in what format the export will be made.
  A choice can be made between CSV, XLS, XLSX, Original (documents in the export will be in the original format) and PDF (all documents will be converted).

When all fields have been filled in, simply click the "Export" button to start the export. A notification will be displayed that the export is running.

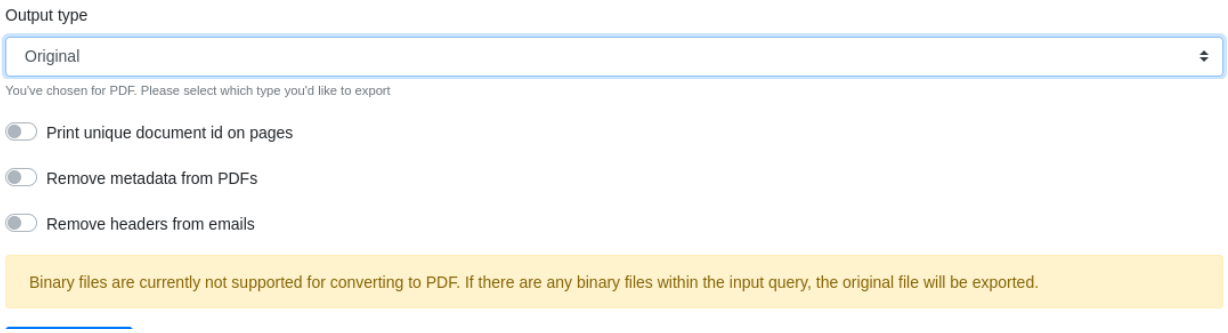
When the export is done, another notification will be displayed. Upon refreshing the page, the export will be shown in "Download Exports"

8.1.1 Creating a PDF Export

As seen in the section above, one of the export options is "converted". This allows you to create an export with all of the files converted to PDF, i.e. all your images, e-mails, spreadsheets, and documents will be in PDF format in the downloaded export.

When creating a converted export, there are two options: "Original" and "Redacted". The names are straightforward - the latter option will export only the PDF files that are redacted (including mails), and the former will redact all files as PDF.

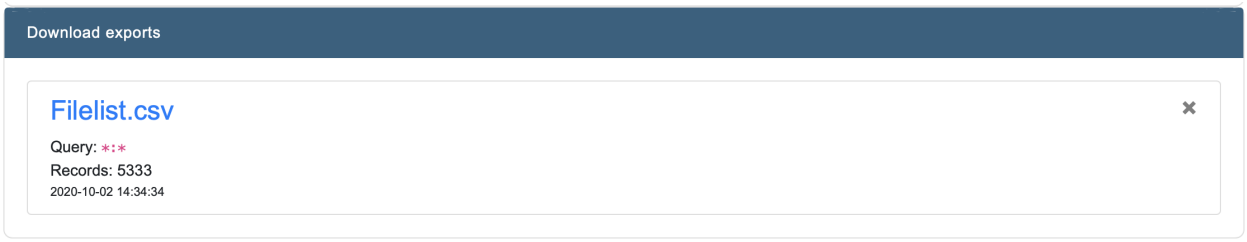
The next three switches are the following:



- This option will print the UUID of the document on each page of the PDF.
- This option will remove any custom metadata on a PDF document.
- This option will remove any headers from mails that are converted to a PDF, rendering the mail to purely the text content of the mail.

8.2 Downloading an export

When an export is made and finished, it will be displayed in the "Download exports" field. All previous exports will be kept here, available for download.



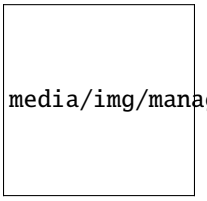
Simply click the name of the export to begin downloading. When an export is no longer needed, click the "X" button to remove it.

This panel also shows you the name of the export, the query that was used to create the export, the number of records that are exported, and the timestamp when it was exported.

CHAPTER

NINE

CASE MANAGER INVENTORY EXPORTS



Here the user can create inventory listing exports for WOO-cases, using the exclusions and judgements tags that are configured in the Display settings of the case. This page has a couple of inputs:

- **Query**
  This is the query by which documents will be fetched. If left blank, then all documents of the case will be fetched.
- **Exclusions**
  Here the user can pick which tags from a pre-configured list will be used as exclusions.
- **Judgements**
  Here the user can pick which tags from a pre-configured list will be used as judgements.
- **Use metadata dat**
  If selected, the date in the document metadata will be used to date the document. Otherwise, the date from the index will be used.

Once the inventory listing is created, it will take some time for it to show up on the page. Once the exporting of the inventory listing is done, the user can downloading the listing from the list of completed listings.

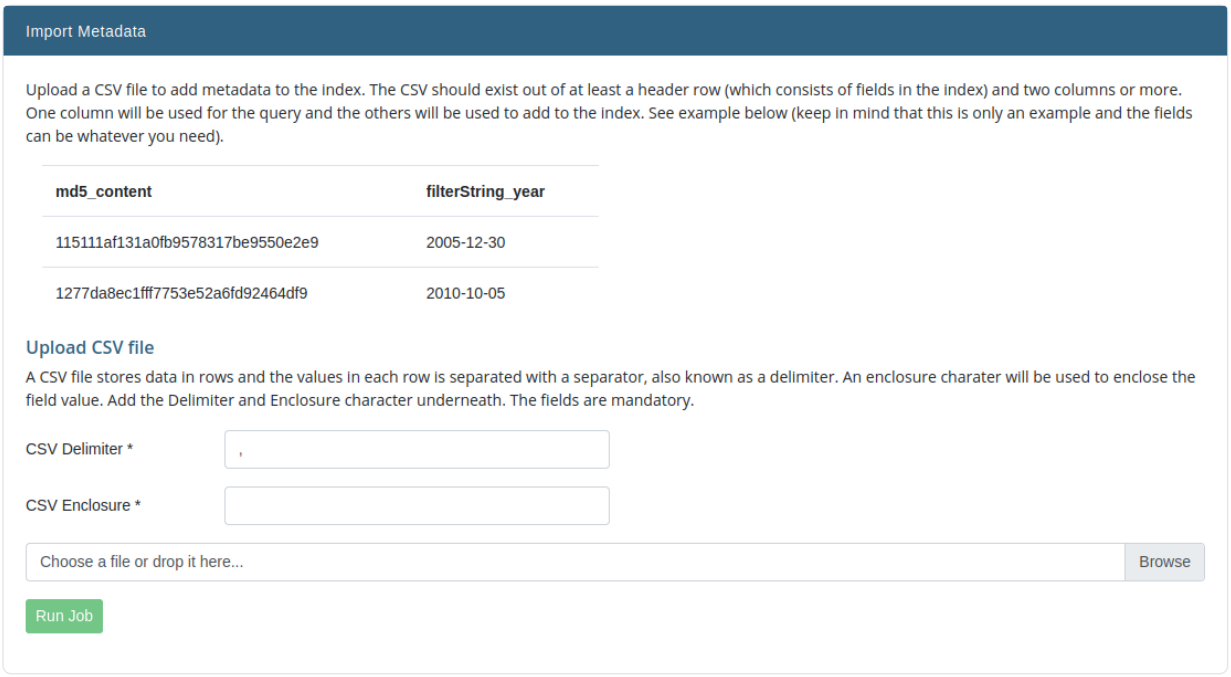
CHAPTER

TEN

CASE MANAGER METADATA UPLOAD

There may be a situation where the metadata of documents needs to be changed en-masse. This can be done with the help of this feature. The user can upload a CSV file with the md5_content of the files that need to be changed and columns with the metadata to change. The CSV must have a header. The user also has to include the delimiter and the enclosure of the CSV files.

Once the files are uploaded and the delimiter and enclosure are given, the job may run. It may take a while depending on the size of the CSV.

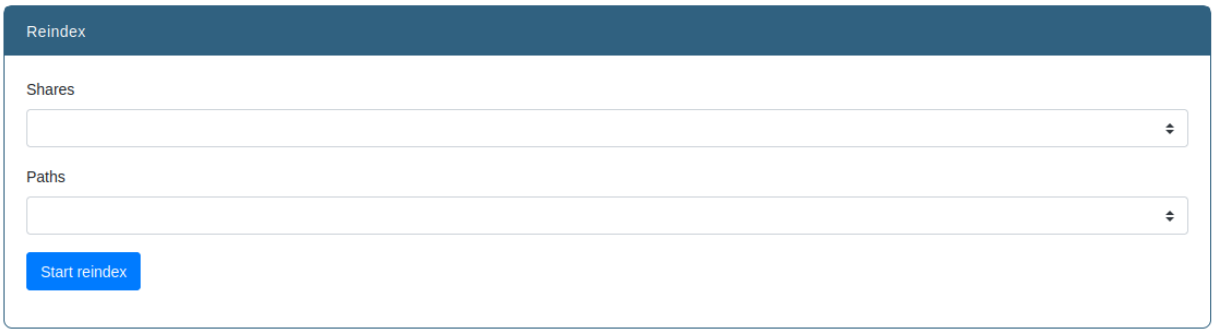


CHAPTER

ELEVEN

CASE MANAGER PATH REINDEX

The case manager can choose to re-index certain documents in the case if necessary. This can be done by selecting the share to re-index, as well as the path to reindex.



CHAPTER

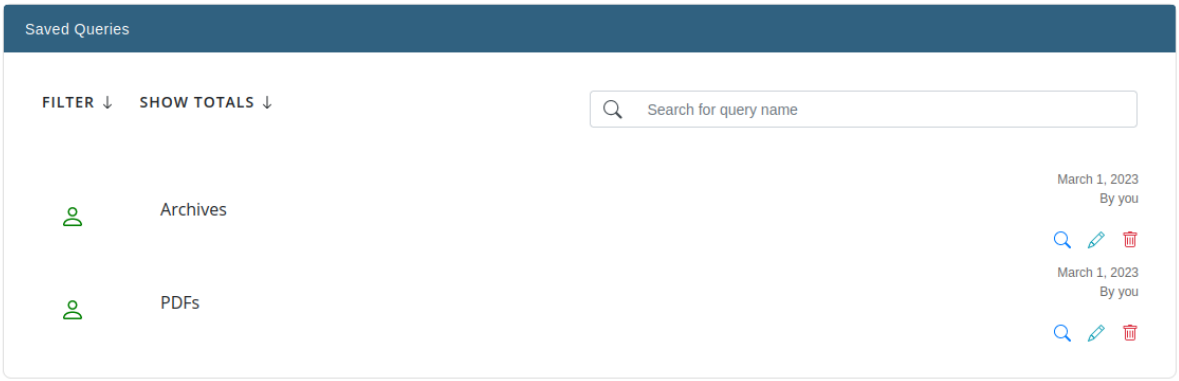
TWELVE

CASE MANAGER SAVED QUERIES

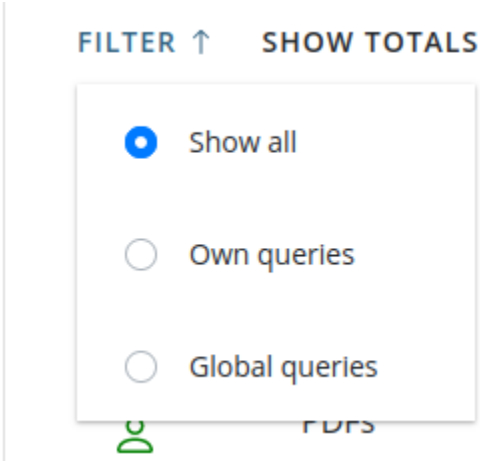
Table of Contents

- *Case Manager Saved Queries*

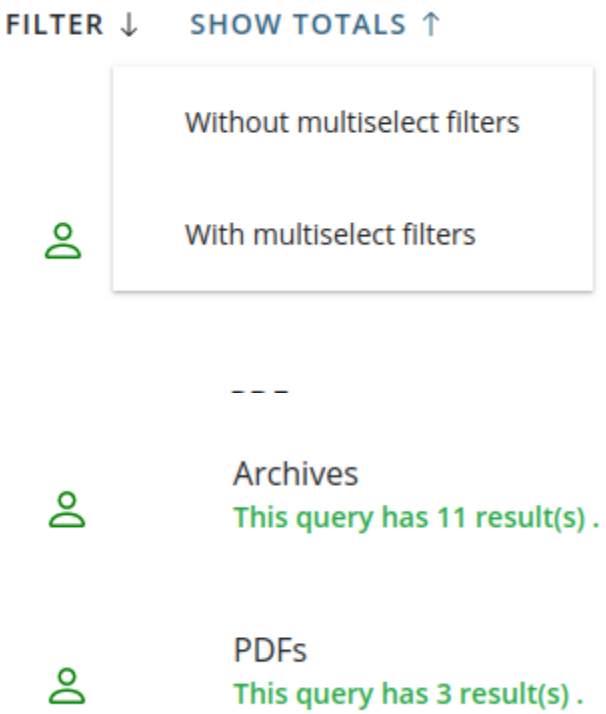
Saved Queries can be found underneath the *Process* menu item within the manage interface. Within this page an overview will be shown of the current saved queries within the case. These are the queries added by hand and the queries added by users. Queries added by users are green colored by default and have a user icon in front of the query name. Queries added here by hand are the color black by default and have a database icon in front of the query name.



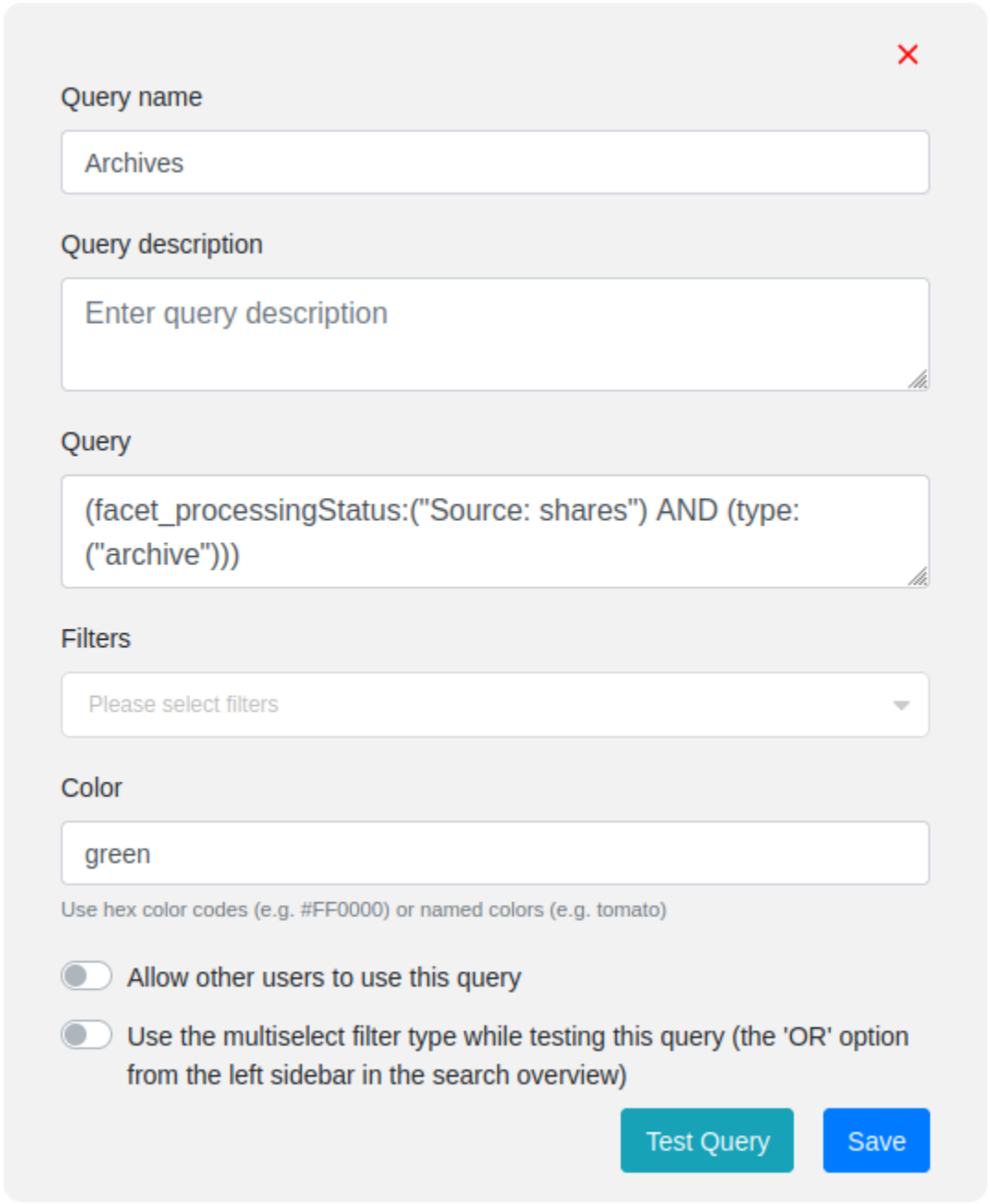
There are a few possible actions within this page. If you have a lot of saved queries, you can search through them. You can also filter through global queries or only the ones that are made by you.



Additionally, you can show the totals of the queries, with the multiselect filters on or off. The result is showing the amount of search results the queries listed give.



Each saved query has three actions associated with it: you can search on the query, edit it or delete it. When editing the saved query, you may test the edited query to check if it is still valid. You may also select if the query is allowed to be used by other users, meaning it will be available to be used by other users on the case.



Read more about creating saved queries here: *Saved Queries*

CHAPTER

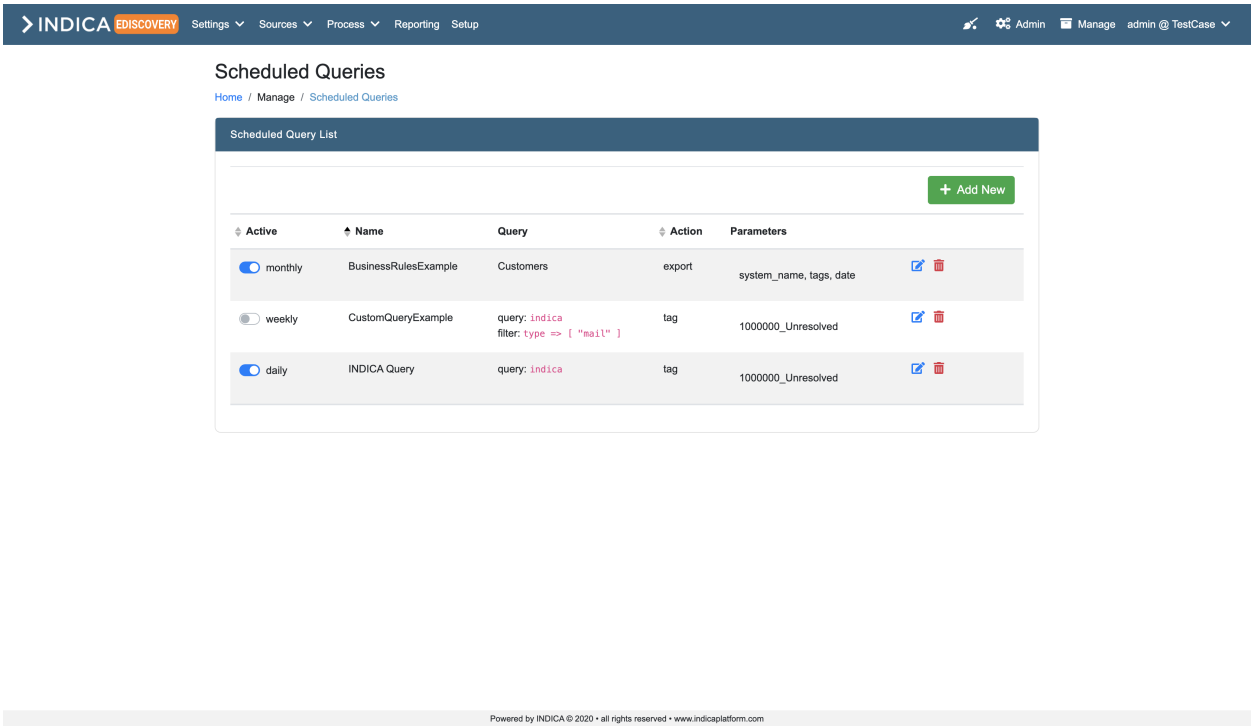
THIRTEEN

CASE MANAGER SCHEDULED QUERIES

Table of Contents

- *Case Manager Scheduled Queries*
  - *Scheduled Query Table*
  - *Adding Scheduled Query*

Scheduled Queries is a **new feature** since INDICA 11.0.0 and can be found underneath the *Process* menu item within the manage interface. Within this page an overview will be shown of the current (in)active scheduled queries within the case. It's possible to have a custom scheduled queries or a scheduled "query" on a business rule.



13.1 Scheduled Query Table

Within the table it's possible to do a few things:

- **Activate / Deactivate queries**

  By clicking the switch at the "Active" column, it's possible to activate or deactivate the scheduled query. How often the scheduled query runs is shown next to switch.

- **Actions**

  Within the table it's possible to edit or delete an existing query by clicking the corresponding icons.

13.2 Adding Scheduled Query

A scheduled query can be added by clicking on the green *Add New* button. This will open up a modal with a form to fill in.

- **Active**

  Whether you want the scheduled query to be active or not.

- **Type**

  There are two possibilities here. You can either choose a business rule or go for a custom query.

- **Name**

  Pick a name for your scheduled query so you can identify it.

- **Query / Business Rule**

  This option depends on what you picked by Type. Type in your custom query or choose a business rule out of the list.

- **Action**

  Pick which action you want the scheduled query to perform. You can choose tag or export.

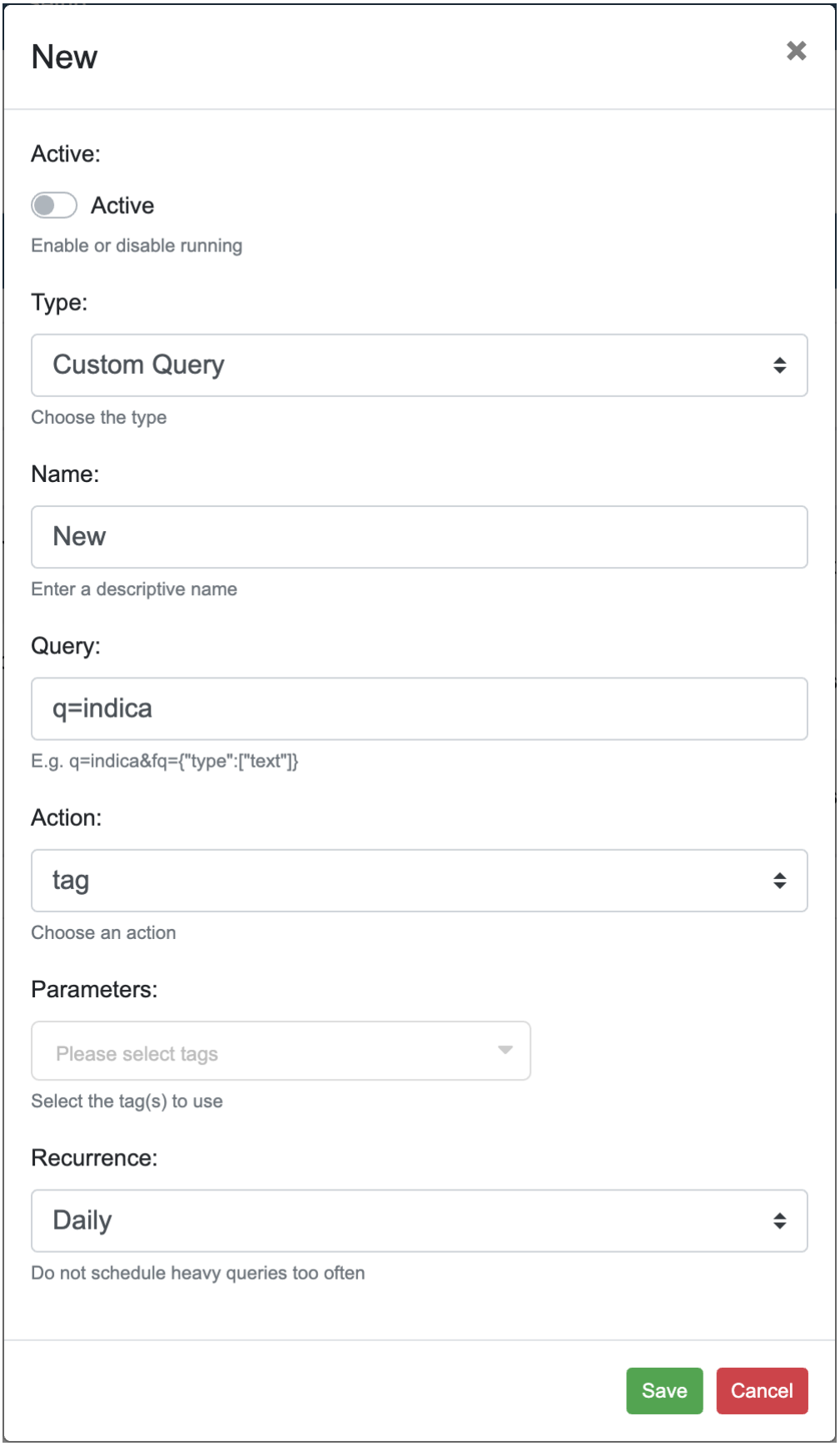
- **Parameters**

  This option depends on what you picked by Action. You can either select the tag(s) you wish the scheduled query to tag, or type in the fields you wish to export.

- **Recurrence**

  How often you want the scheduled query to tag documents or export your fields.

> **Caution:** Please keep in mind that it's *highly advised* not to schedule heavy queries too often.


New
Active:
Active
Enable or disable running
Type:
Custom Query
Choose the type
Name:
New
Enter a descriptive name
Query:
q=indica
E.g. q=indica&fq={"type":["text"]}
Action:
tag
Choose an action
Parameters:
Please select tags
Select the tag(s) to use
Recurrence:
Daily
Do not schedule heavy queries too often
Save
Cancel

CHAPTER

FOURTEEN

CASE MANAGER TAGS

Table of Contents

- *Case Manager Tags*
  - *Introduction*
  - *Tags*
    - *General Information*
    - *Adding Tags*
    - *Re-ordering and Nesting Tags*
    - *Removing Tags*
    - *Locking the Tags*
    - *Untagging a Tag*
    - *View Tag Audit Log*
  - *Export*
  - *Tag by CSV*

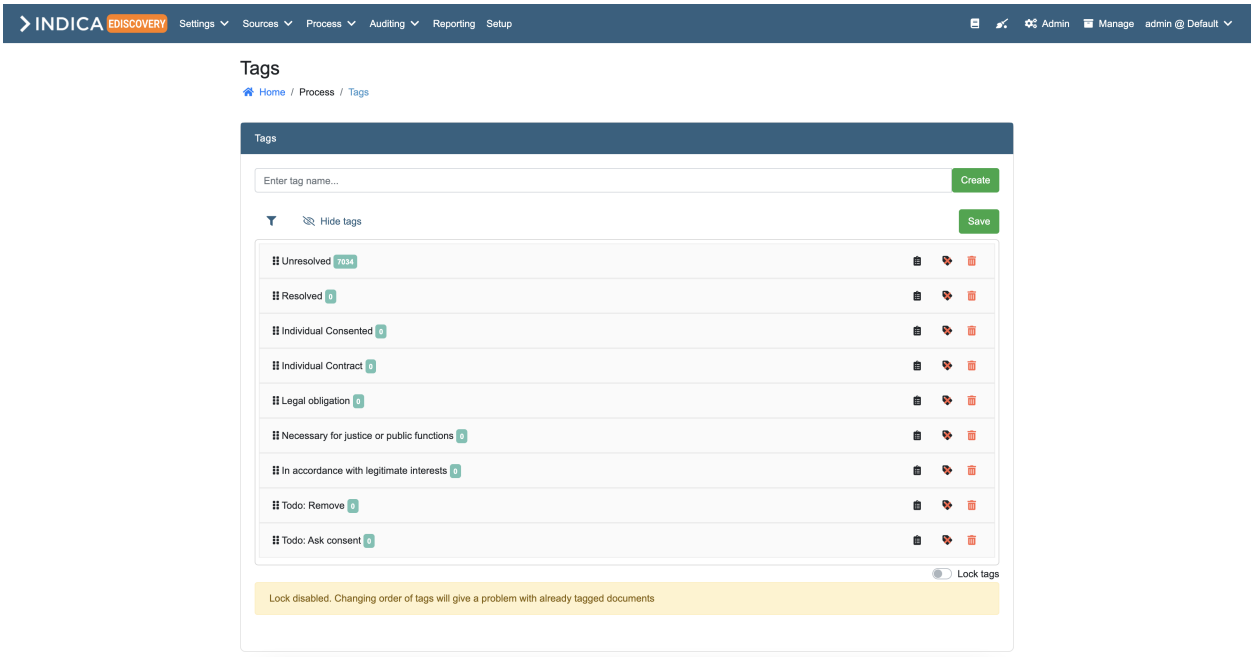
14.1 Introduction

Tags are labels that can be assigned to documents. They contain a small bit of extra information, like (for example) if a document is "Relevant" or "Can be deleted" or "Migrated". Tags are incredibly versatile and can be used for many situations. Users can assign tags to documents once the tag has been created. This can be done for a single document, for a page of documents, or for all query results. It is also possible to add tags automatically, or with help of CSV files.

Case managers have different actions they can do with tags in the manage interface. Tags can be found underneath the Process menu item within the manage interface. In this page you can do the following things:

- Add new tags;
- Re-order or nest tags;
- Delete old or unwanted tags;
- Lock and unlock tags;
- Untag the tag from documents;

- Audit usage of the tag;
- Export (selected) tags;
- Upload a CSV file to tag documents by CSV.

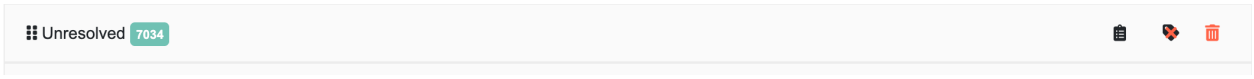


14.2 Tags

14.2.1 General Information

In this overview, every tag is displayed in their own row. This row contains (from left to right):

- Anchor to drag the tag around;
- Name of the tag;
- Number of documents that have this tag assigned to them;
- Button to open the tag's audit log;
- Button to untag this tag from the documents;
- Button to delete this tag.



14.2.2 Adding Tags

Adding tags is easy. Just enter a tag name in the input field and click on create. The tag will show up in the list underneath (see image below for an example).

14.2.3 Re-ordering and Nesting Tags

You can also move tags around by clicking on the tag and dragging it to a different place. This will change the order of tags. It's also possible to have nested tags. To do this, click a tag, drag it on top and a bit to the right of another tag. You will see a preview of the tags placement whenever moving tags around.

Caution: Keep in mind that changing the order of tags when you already have tagged documents in your case, will give some problems with the already tagged documents.

14.2.4 Removing Tags

To delete a tag, click on the red trashbin icon to the right of the tag. A confirmation will open. This confirmation will ask you if you are sure you want to delete this tag. It gives the following options:

- **Cancel**

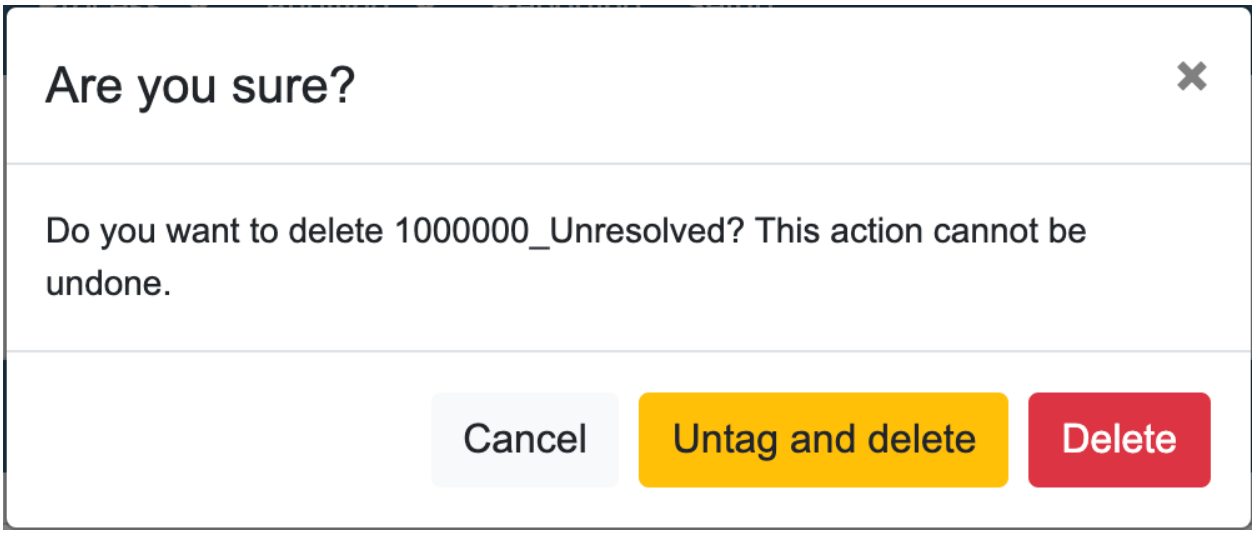
  Closes the confirmation and does not delete the tag.

- **Untag and delete**

  Sends a command to the index to remove this tag from all documents and then deletes the tag.

- **Delete**

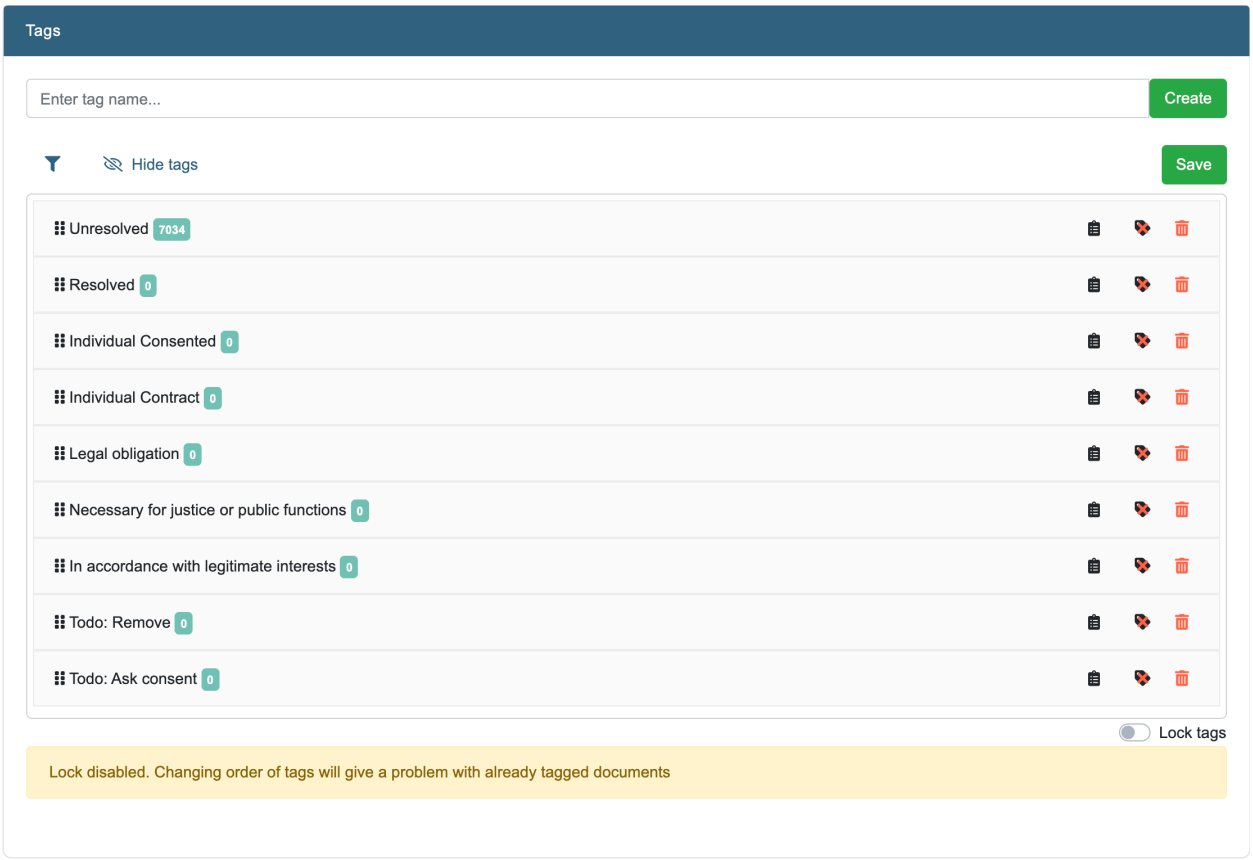
  Leaves the tag assigned to documents and deletes the tag from the overview only.



14.2.5 Locking the Tags

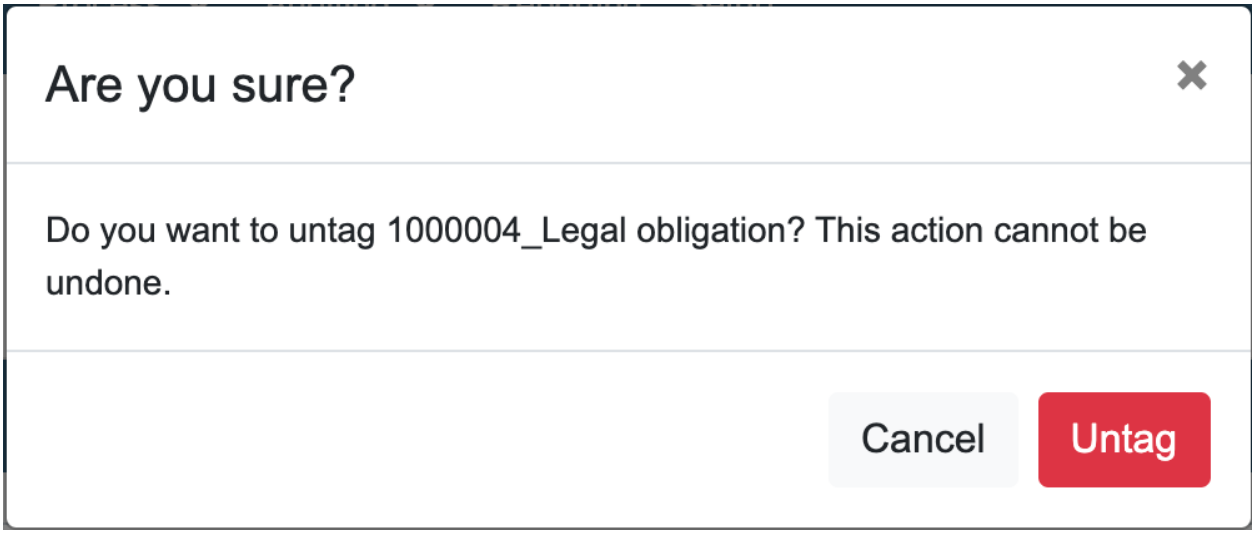
There's also a possibility to lock the tags. When you're done creating tags and moving them around, it's recommended to lock the tags, as this will ensure their placement. This way, when you (or another case manager) comes to the overview, you don't drag any tags by accident.

Important: If GDPR is enabled on your case, some standard tags will show up for you to use. You can delete these if you wish to do so. However, when the indexer encounters a document with privacy-sensitive data, it will still assign the corresponding tags. Keep in mind that deleted tags cannot be re-added as their internal ID will be different.



14.2.6 Untagging a Tag

In some situations it may be needed to remove a tag from all documents. It is possible to do this manually from the Search page, but that is inconvenient. Every tag in the overview as a button (label icon with a red 'x') to remove that tag from all documents. When clicked, a confirmation opens:



When clicking "Untag", a job will be started to remove this tag from all documents.

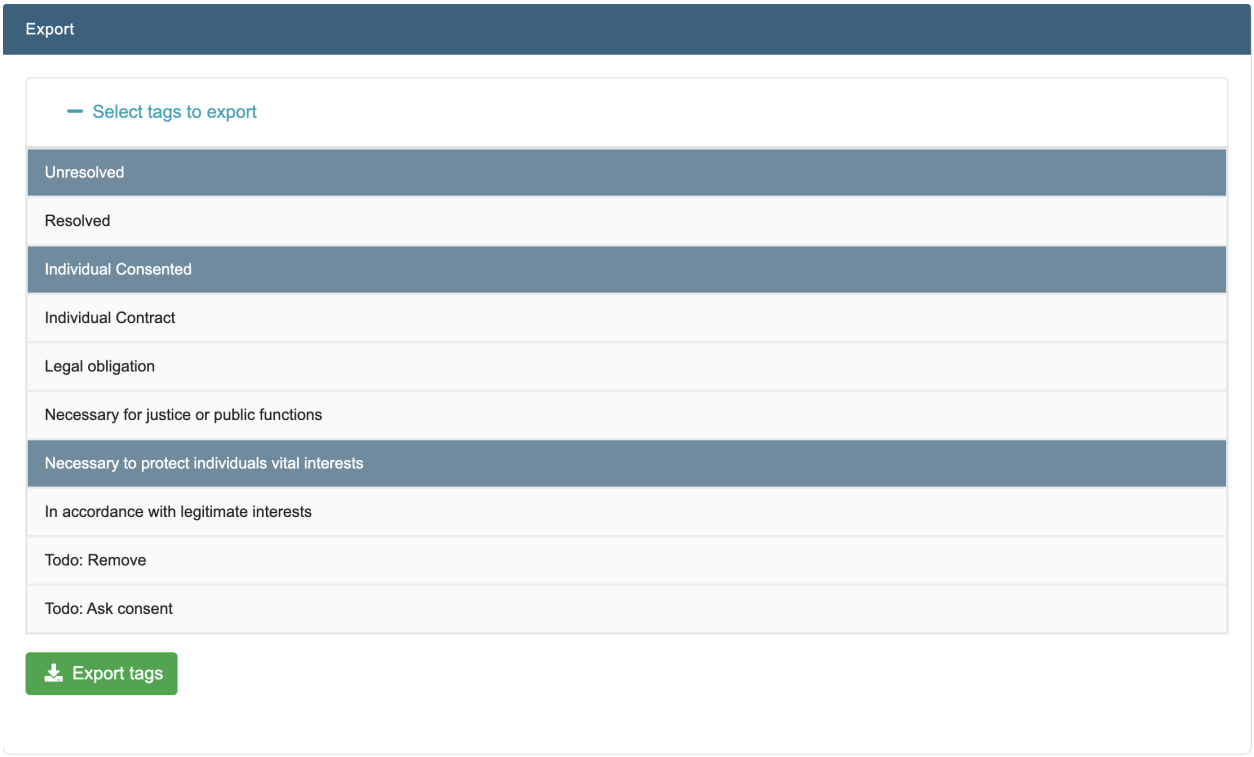
14.2.7 View Tag Audit Log

It is also possible to view the tag audit log. This can be done by clicking the clipboad item in the overview. More information on this functionality is available here: ref-view-tag-audit

14.3 Export

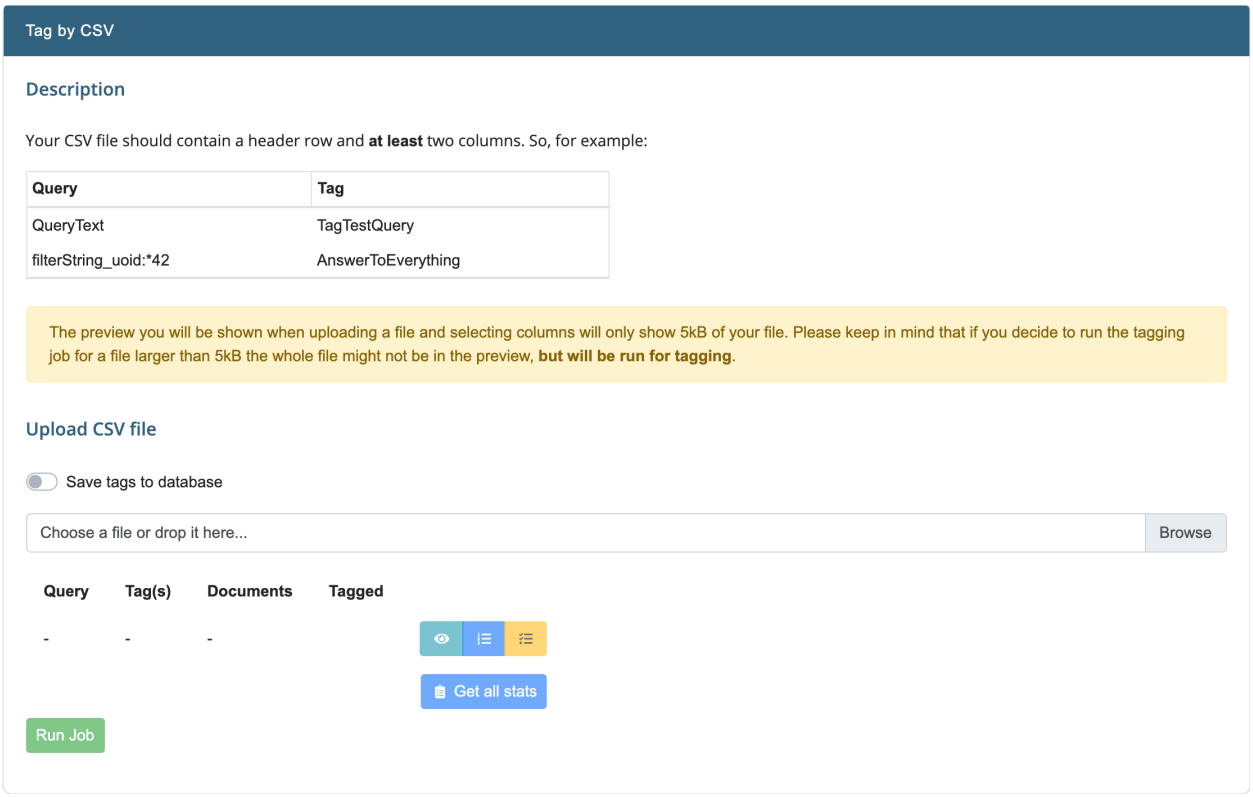
It's also possible to export documents with specific tags. Just click on the tags you wish to export, this will highlight the items in the list. After you've selected your items, you can click on the 'Export tags' button to start the export.

Exporting might take a while, especially when you have a lot of tagged documents. After exporting is done you can click on the download icons to download the file(s).



14.4 Tag by CSV

With INDICA it's also possible to tag documents by CSV. The CSV file should contain a header row and at least two columns. An example has been given in the table.



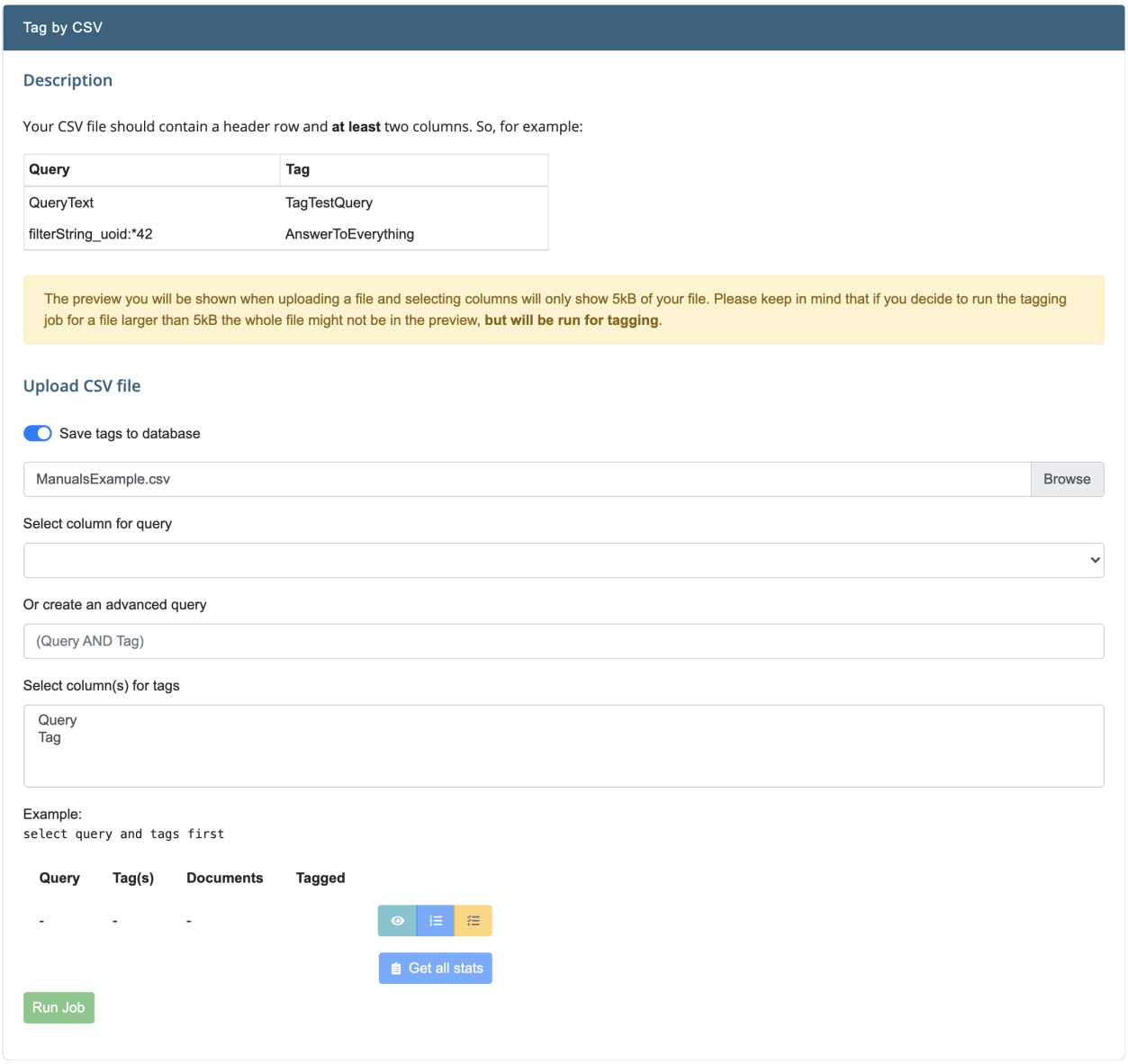
To use this functionality, follow the following steps:

1. Prepare a CSV file that conforms to the given example.
   It should (at least) contain a header row, and two columns.
   The fist column is for the query that will be used, and the second column is for the tag that will be assigned.
   See the example below for the raw contents of the CSV file.

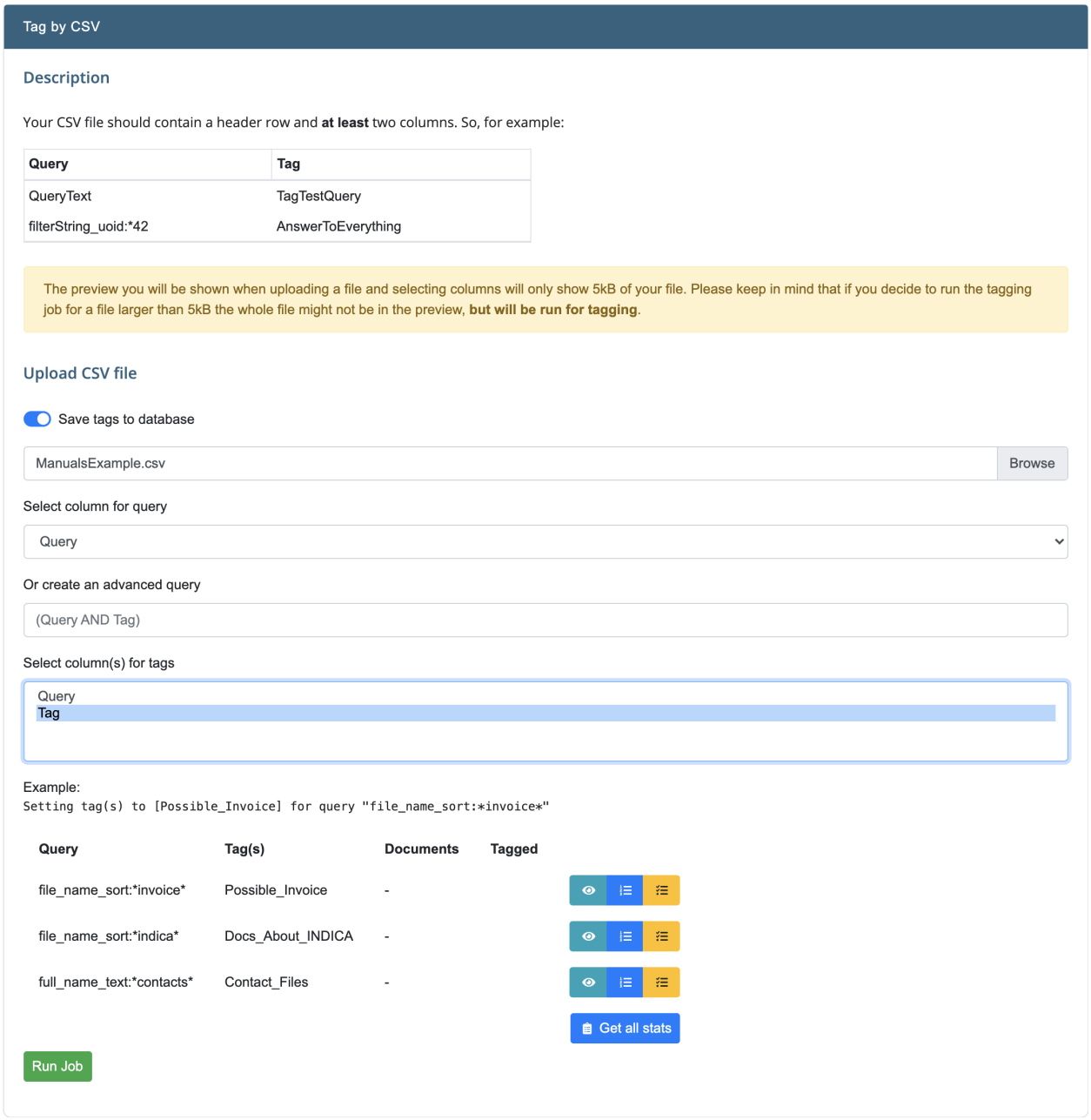
```
Query,Tag
file_name_sort:*invoice*,Possible_Invoice
file_name_sort:*indica*,Docs_About_INDICA
full_name_text:*contacts*,Contact_Files
```

2. Upload the CSV.
   Once uploaded, the overview will change and some questions about the file that was uploaded will be asked:

   1. **Save tags to database** If enabled, the tags will also be saved to the database. This will make them available in all places that work with normal tags and thus makes them dynamic and changeable. Alternatively, you may choose to leave this switch disabled to write static tags to the index that cannot be changed.
   2. **Select column for query** Select the column that contains the query here. In the example, the column is called *Query*.
   3. **Or create an advanced query** This option gives the possibility to use multiple columns as part of the query. The search terms that come from the CSV can be referred to by the column name.
   4. **Select column(s) for tags** Select which column(s) will be used for the tag(s). Do note that selecting multiple options will assign multiple tags to the documents.



When all required information is filled in, the overview will show you the queries and tags that it will assign once the jobs are run:



3. Choose an action to perform.
   The bottom section now contains an example of the job it will run, and a list of all queries and tags that it wil run.
   Each row shows some information and some actions that can be performed. From left to right these are:

   - The query as read from the uploaded CSV file;
   - The tag(s) as read from the uploaded CSV file;
   - The number of documents that are returned for this query *(Only visible after getting the stats)*;
   - The number of documents that already have this tag assigned to them *(Only visible after getting the stats)*;
   - **And three buttons to perform the following actions:**
     - Eye icon (teal): See results of this query. A new tab will open with the query applied, showing

you the documents;

- Numbered list icon (blue): get statistics of this row;
- Checked list icon (yellow): assign the tag to the documents for this row only.

Example:

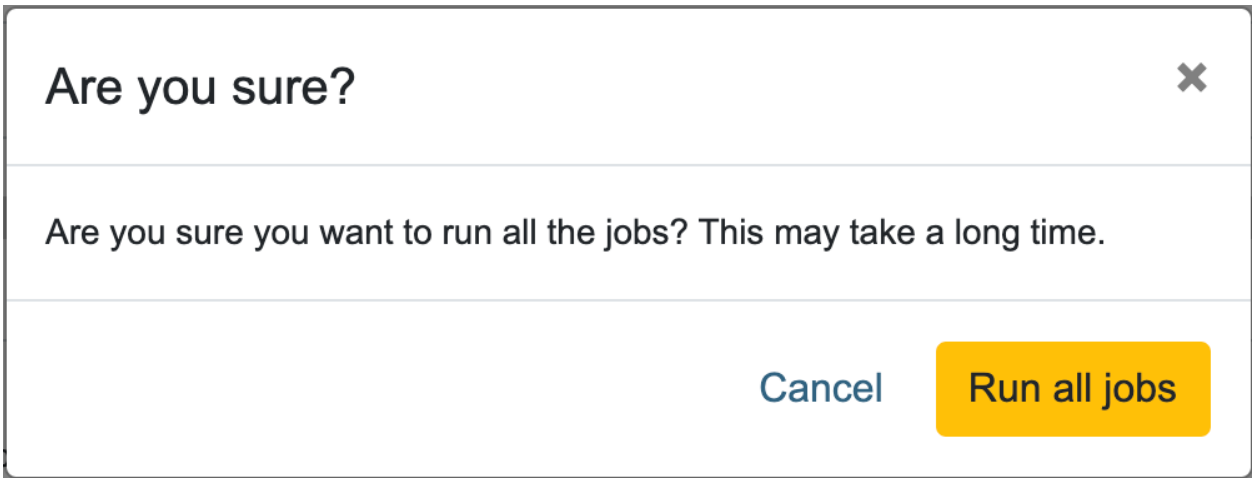
```
Setting tag(s) to [Possible_Invoice] for query "file_name_sort:*invoice*"
```

Query	Tag(s)	Documents	Tagged	
file_name_sort:*invoice*	Possible_Invoice	-		

After getting the statistics, the row contains some more information:

Query	Tag(s)	Documents	Tagged	
file_name_sort:*invoice*	Possible_Invoice	20	Possible_Invoice 0	

1. Run actions in large batches.
   At the bottom of the functionality two additional buttons can be found:
   - Get all stats: gets all the statistics for all rows.
   - Run Job: performs the queries and assigns the tags for all rows. A confirmation will be displayed before running the jobs:



Important: The preview you will be shown when uploading a file and selecting columns will only show 5kB of your file. Please keep in mind that if you decide to run the tagging job for a file larger than 5kB the whole file might not be in the preview, but will be run for tagging. If you wish to preview all the results, it is advised to split up the CSV file into two or multiple batches.

CHAPTER

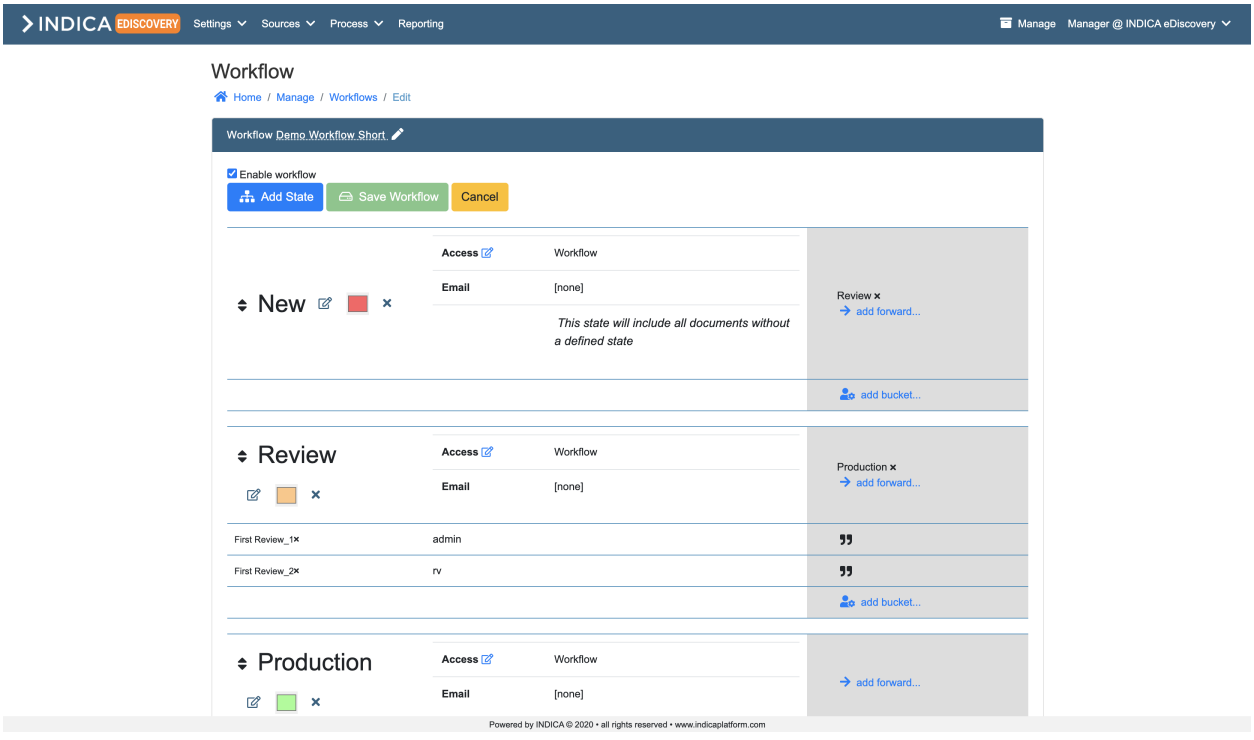
FIFTEEN

CASE MANAGER WORKFLOW

Table of Contents

- *Case Manager Workflow*
  - *Create workflow*
  - *State options*
  - *Visual representation workflow*

The workflow can be found within the manage interface underneath the *Process* menu item. Here you can see an overview of existing workflows and create, edit or delete workflows.



15.1 Create workflow

To create a workflow click the *Create new workflow* button. This will bring you to a new page where you can do the following actions:

- **Edit Workflow Name**

  To edit the workflow name, click the pen icon next to the workflow name. Edit the name and press enter to finalize the changes. Please do not forget to click the *Save Workflow* button to save your changes.

- **Enable Workflow**

  To enable the workflow, check the box next to 'Enable workflow'. If this box is unchecked, the workflow will not show up to users. Please do not forget to click the *Save Workflow* button to save your changes.

- **Save Workflow**

  Make sure you save your changes by clicking the *Save Workflow* button, otherwise your changes will not be saved.

- **Cancel**

  Cancel all changes you made.

- **Add State**

  To add a workflow state, click the *Add State* button. Fill in a name for the state and an email address if necessary and click on the *Save* button next to the input fields. This will add a new state to your workflow. See all options within the chapter 'State options' below.

  The first state in a workflow will always be the state where all documents *without a defined state* will be included.

15.2 State options

After a new state is added there are a few possibilities you can do with the new state.

- **Assign colors**

  Assigning colors to your states will make the overview a little easier to read. These colors are also used in the icons next to the state within the front-end.

- **Access**

  Select a group that has access to the state. Without a group selected the state will not be accessible to users.

- **Add bucket**

  Add bucket(s) to your state. Buckets are assignable to a specific user.

- **Add forward**

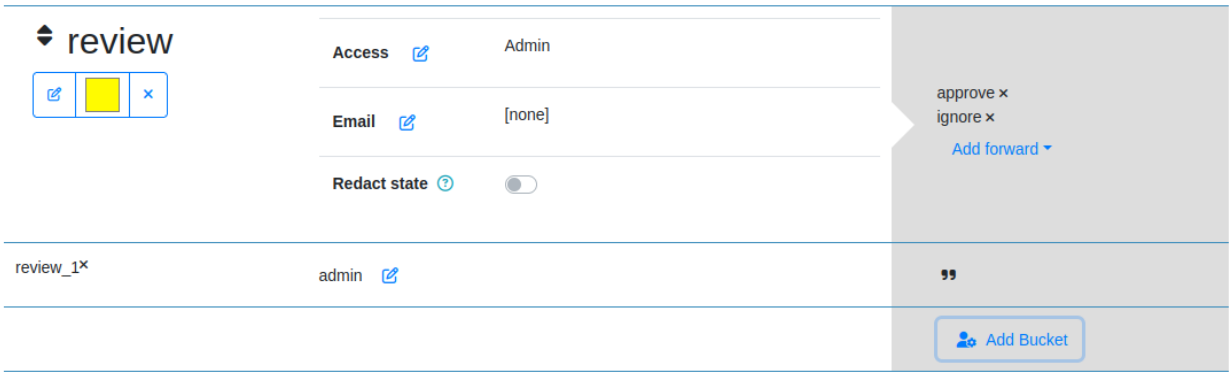
  Add a forward from one state to another. It's possible to add it to the next workflow state in the 'list', or add a forward to a state further down.

- **Move state**

  It's possible to move state in the 'list' of states. Just click the icon with the arrows pointing up and down to drag the state above or below another state.

- **Redact state**

  With PSPDFKit, it is possible to select a workflow state as a "redact" state. When documents with concepts are moved to a redact state, the redactions are automatically applied. A report will be created about all documents that are processed and placed in a workflow redact state, which can be found by clicking the printer icon in the navigation bar.



15.3 Visual representation workflow

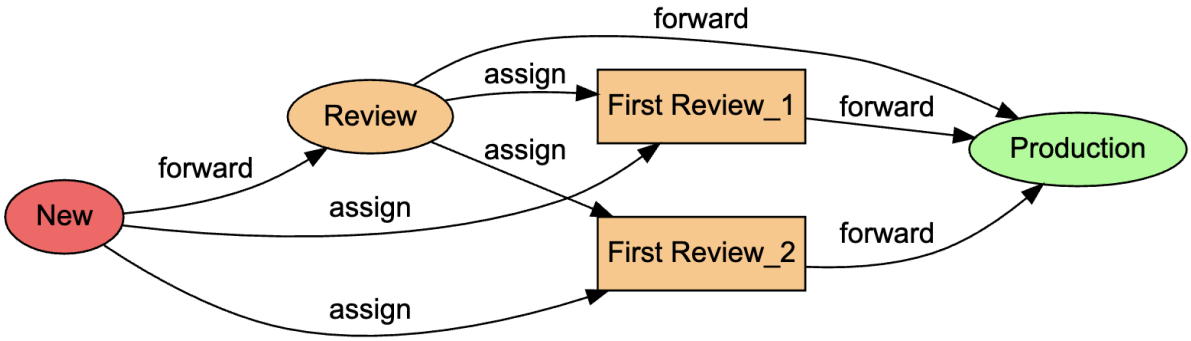
While adding states and buckets, assigning colors and forwards an overview will pop up at the bottom of the page. This is a visual representation of your current workflow.

- **Round shape**

  A round shape is a visual representation of a state. A state can have an arrow from either another state or another bucket. Both of them will have the text 'forward' next to the line of the forward arrow. If the state has a background color, a color has been assigned to that state.

- **Square shape**

  A square shape is a visual representation of a bucket. An arrow will point from a state to a bucket with the text *assign* next to line of the assign arrow. A bucket can also have a background color, but they inherit the color of the state they're placed in.



CHAPTER

SIXTEEN

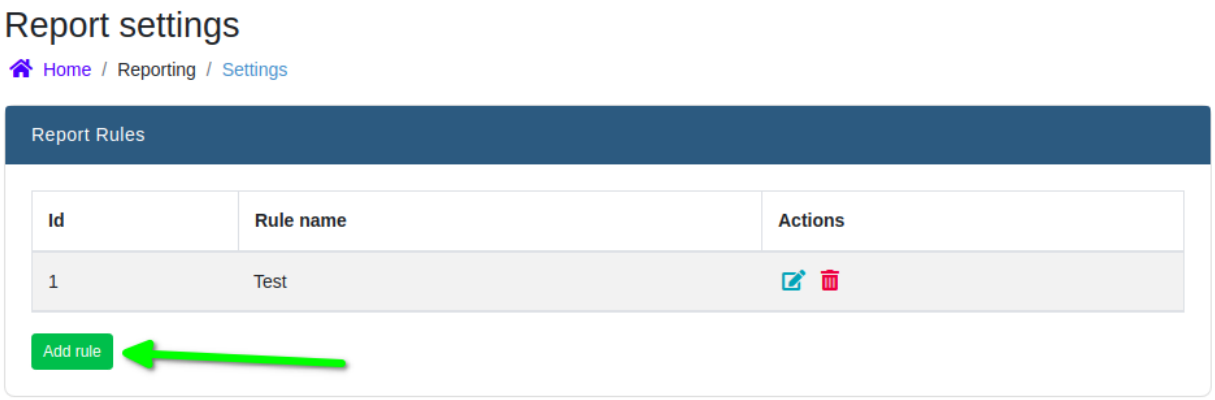
REPORTING

Reporting tool allows you to keep up to date with your data landscape without opening the INDICA appliance. You can setup and receive automated emails from your appliance.

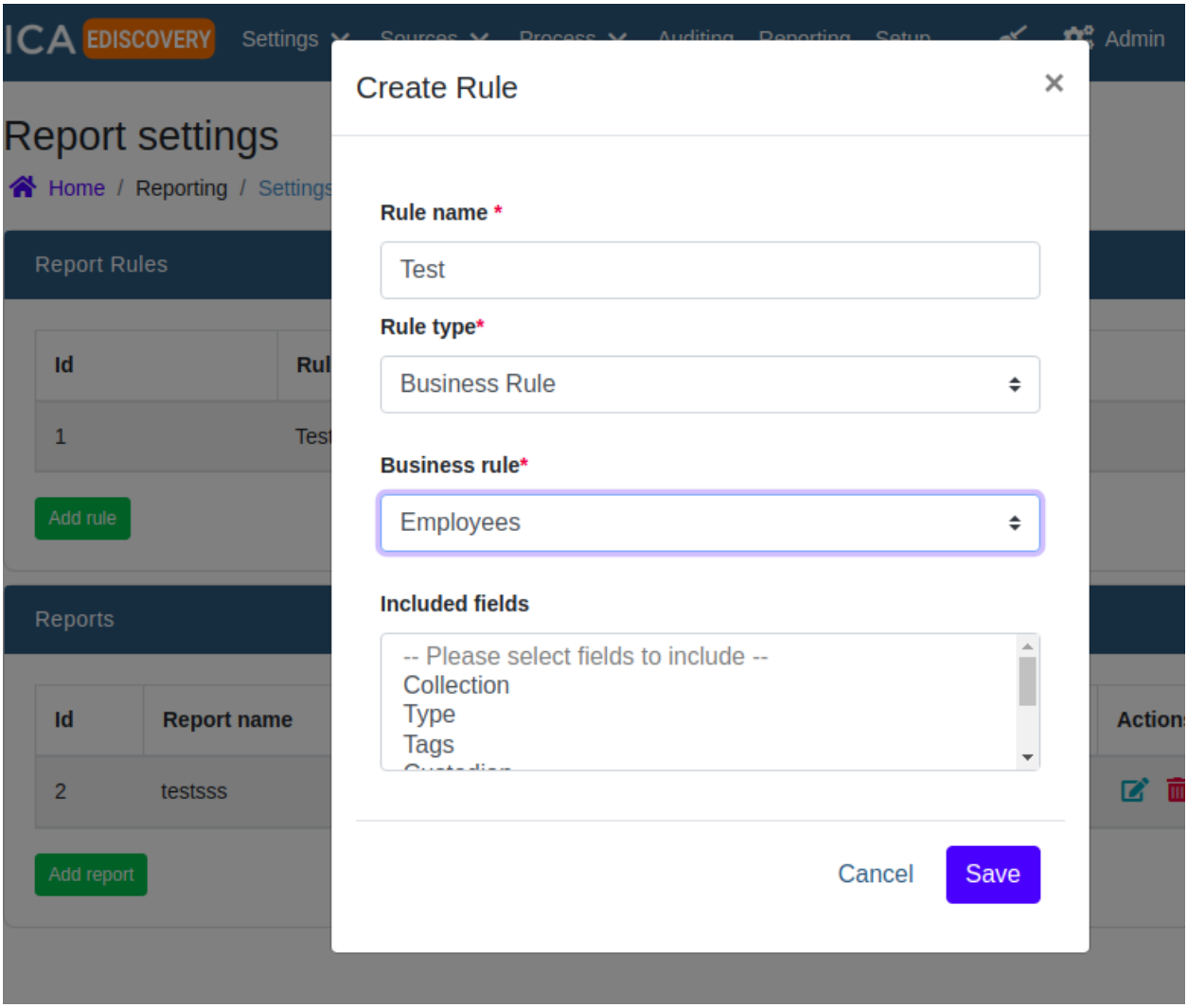
Note: Keep in mind, every user receives reports which are generated based on their rights.

16.1 Report rules

To create an automated report you first have to setup query rules which the report will consist of.



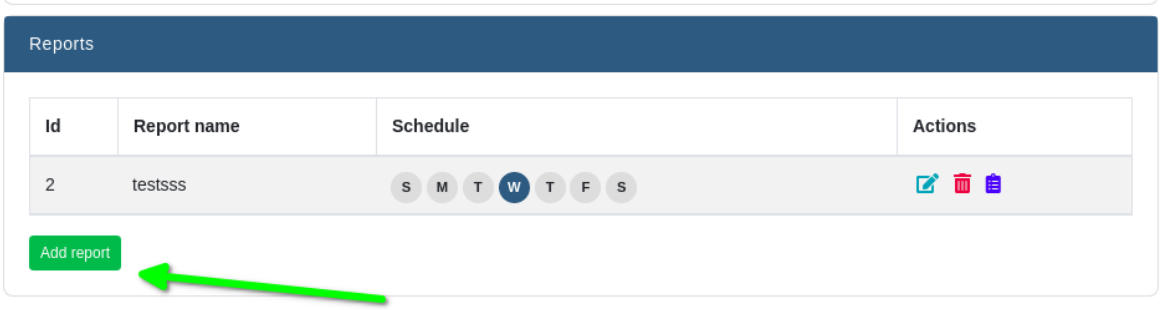
You have an option to create few types of rules based on **Business Rules**, and **Custom Queries**. You also have to choose which fields to include in the report.



Custom Queries follow the same rules as the main search engine of INDICA. You can find what is available in *Querying*

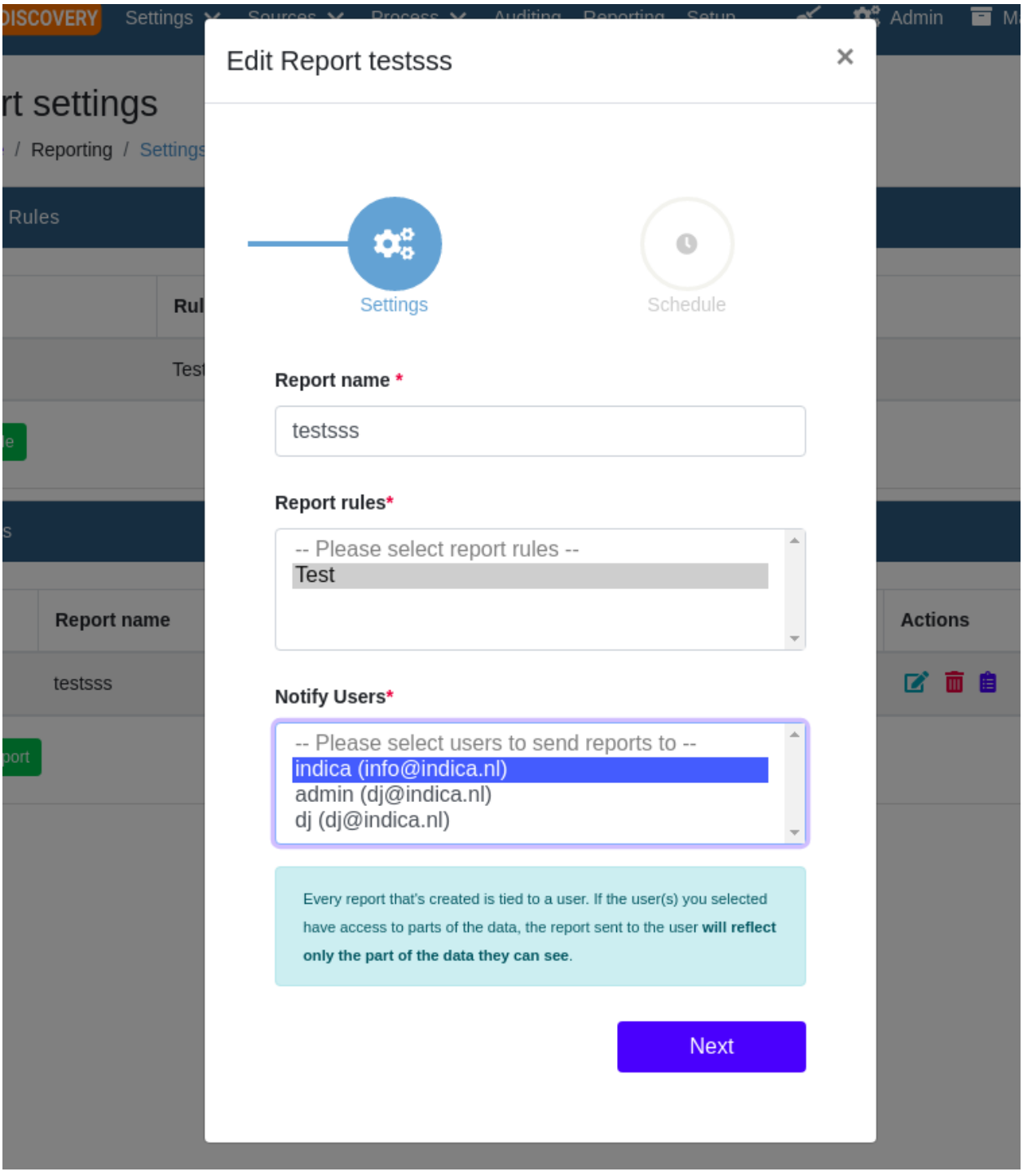
16.2 Reports (Report groups)

Once you are satisfied with the report rules you have set, you can combine different rules to include them in the report email.

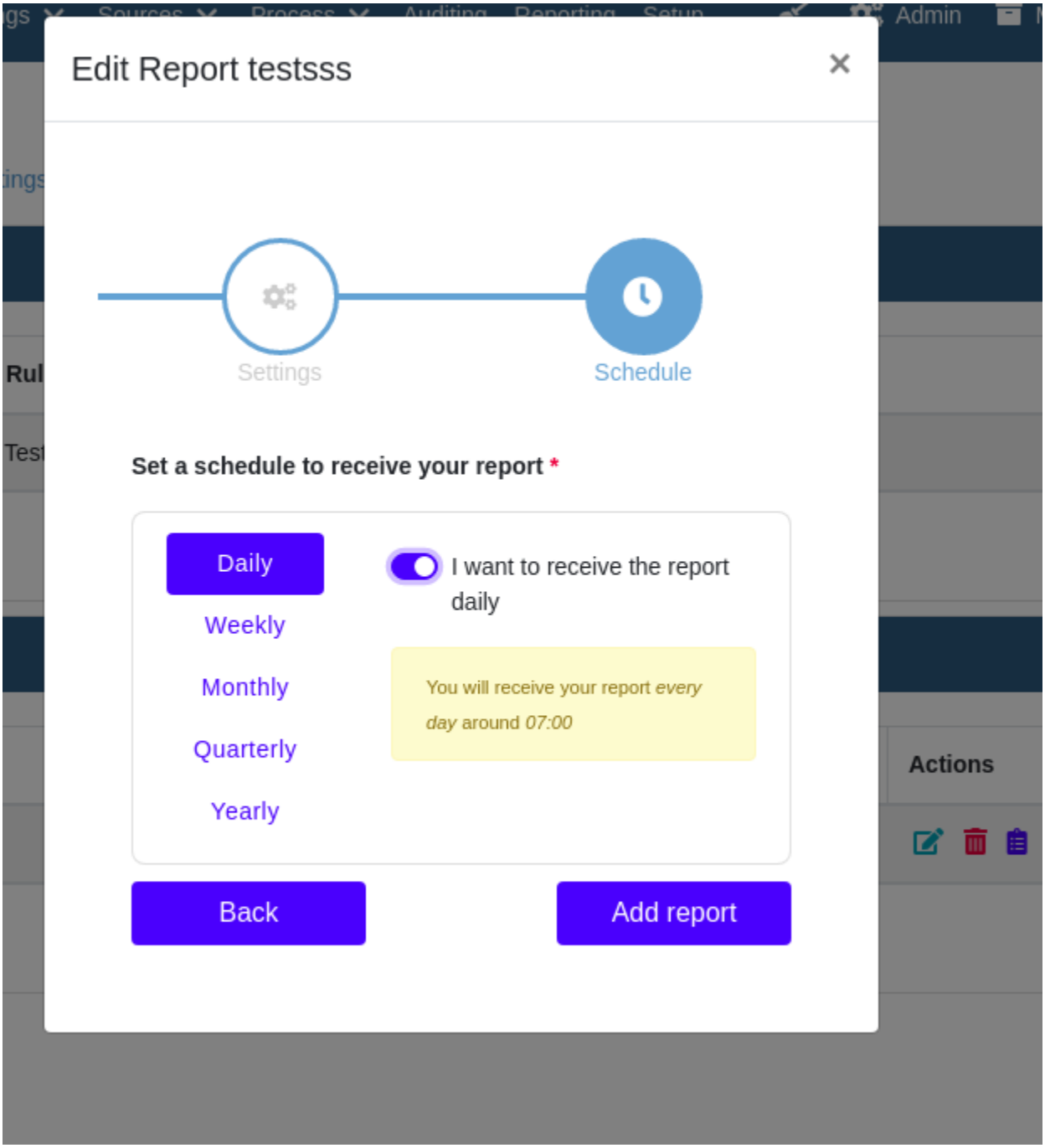


Once you click to add a new report, you have decide what rule to include in the email. To select multiple rules hold

CTRL and click the left mouse button.



By following the reporting wizard you will also get to choose how to schedule the emails.



Once everything is set up, reports will be sent out automatically. It is also possible to see the report instantaniously by clicking the purple clipboard icon in the "Actions" column of the table.

Note: Please keep in mind that delivery times and delays may be influenced by your mail server and hardware

Important: It is also important to know that in case it is not a cloud instance of INDICA, emails will be sent out from your mail server. To see how to configure your mail server go to Networking or ask your administrator

CHAPTER

SEVENTEEN

CASE SETTINGS

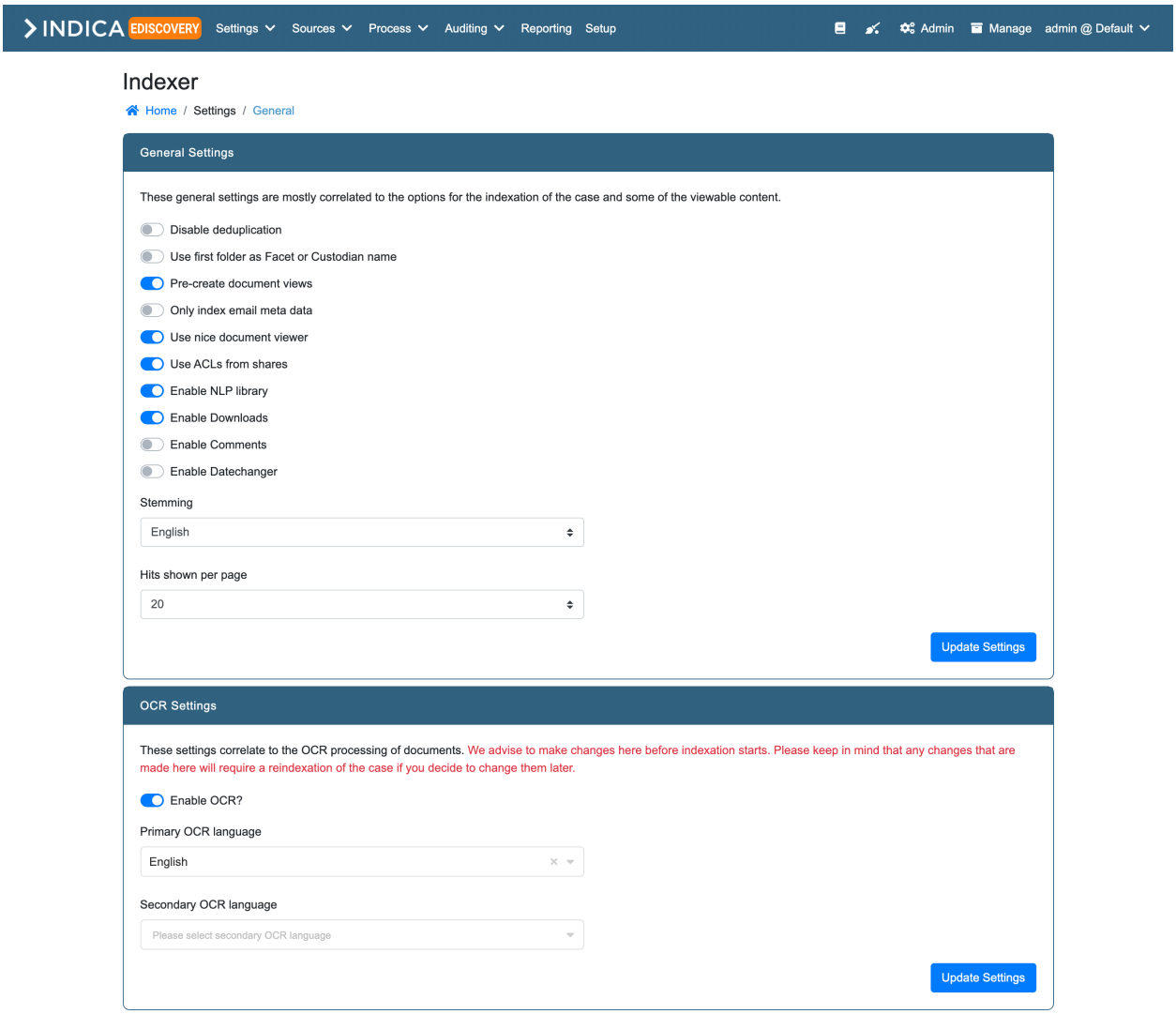
Table of Contents

- *Case Settings*
  - *General Settings*
    - *General Settings*
    - *OCR Settings*
    - *Inventory Listings*
  - *Display Settings*
    - *Facet Settings*
    - *Result list Settings*
    - *Index Progress Notification Settings*
    - *Detail View Settings*
    - *List Style Settings*
    - *Redaction Feature Settings (Advanced PDF Viewer)*
      - *Auto-save*
      - *Redaction Patterns*
      - *Redaction Styling*
  - *Search Settings*
    - *Stopword List*
    - *Synonym List*
    - *Boosting*
      - *Boost Query*
      - *Boost Function*
  - *API Settings*
    - *Polling API*
      - *Setting up your Application*
      - *Batched result responses*

        - *Setting up a polling job on the INDICA appliance*
    * *PowerBI Connection*
        - *Creating an OAuth client*
        - *Creating a Personal Access Token*
        - *Connecting PowerBI to INDICA*
        - *Creating request URLs*
    * *Polling API Examples*
        - *Job GET Endpoint JSON*
        - *Job Result JSON*
        - *Document Query JSON*
        - *Document Query Response JSON*
        - *Subject Details Request JSON*
        - *Export Documents JSON*
        - *Document Export Response JSON*
        - *GDPR Query JSON*
        - *GDPR Query Response JSON*
        - *Batch Response JSON*

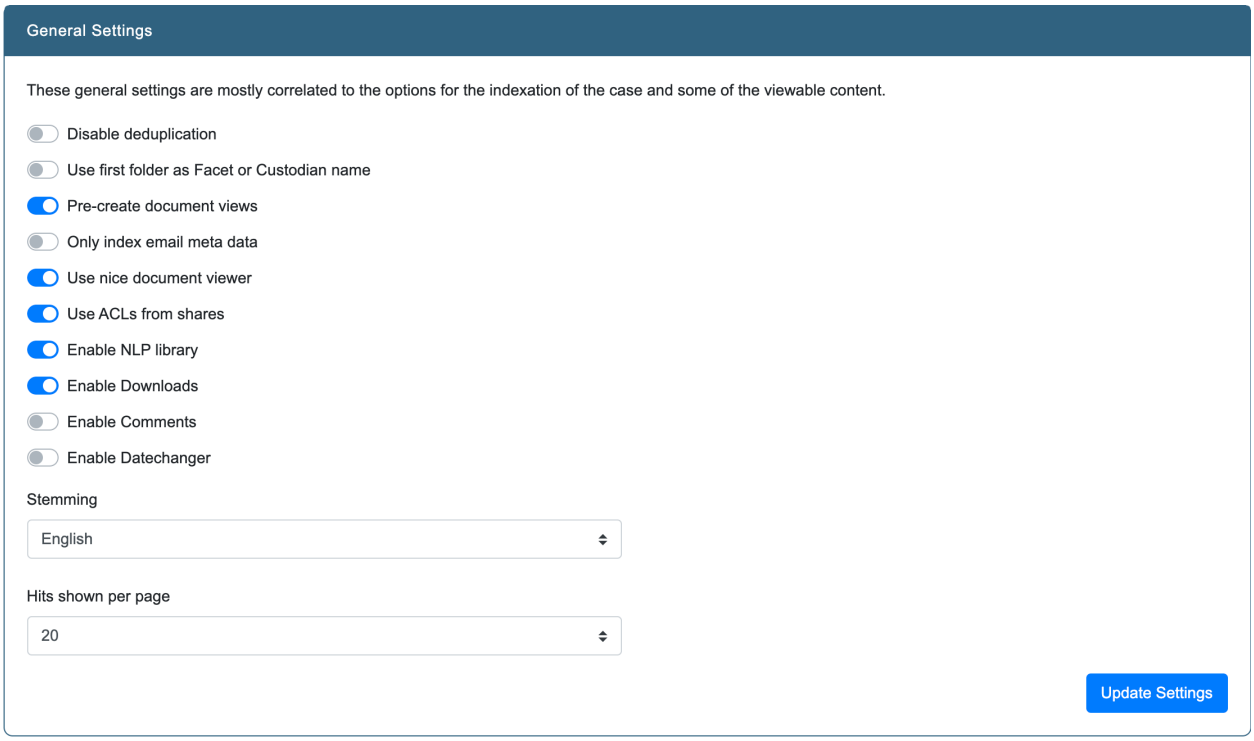
17.1 General Settings

The "General Settings" page allows you to tune some settings like hits per page, enable comments, datechanger, OCR, etc. It has two main sections, General Settings and OCR Settings. Both sections and their settings are explained below.



17.1.1 General Settings

General settings allow you to switch some features on or off. It also allows you to set some parameters that will influence indexer behaviour in order to get the best possible results. It is advised to make sure these settings are configured correctly before starting the indexation of data.



- **Enable deduplication**
  Enables de-duplication of documents based on their content.

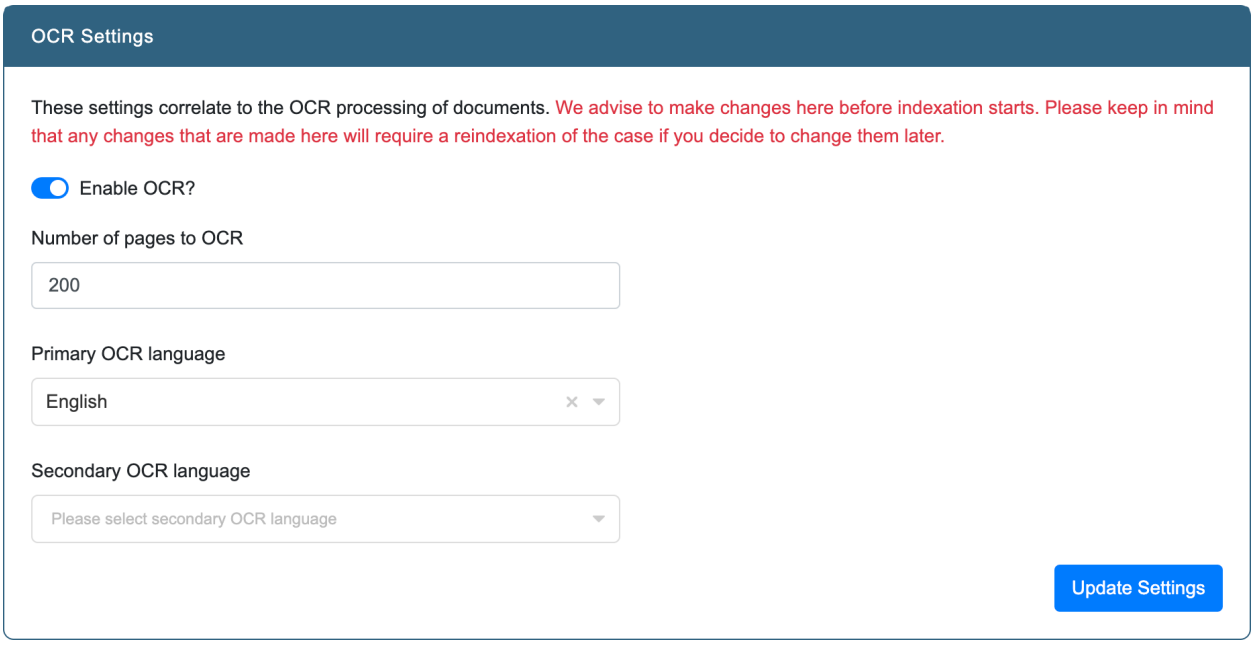
  Note: The case needs to be re-indexed when this setting is changed.

- **Use first folder as Facet or Custodian name**
  This feature is especially useful for eDiscovery investigations. Enabling this option allows you to place data items that belong to a custodian in a separate folder. INDICA will recognise this folder as a custodian and adds filtering capabilities on custodian in the facet bar. The filter options will be displayed in the categry "Custodian". Please do make sure that the category Custodian is visible in the Facet Settings.
- **Pre-create document views**
  This option allows you to pre-create document views. This will result in nicer document previews and faster loading times. More disk space will be used.
- **Only index email meta data**
  When this option is enabled, only the metadata of emails is processed. The email content will not be included in the index.
- **Use ACLs from shares**
  When this option is disabled, the Access Control Lists of shares will not be taken into account when indexing. This means that all INDICA users can see all indexed documents, regardless of whether or not they can see the documents on the source.
- **Enable NLP library**
  When this option is enabled, INDICA will use it's NLP algorithm to extract NLP keywords from documents.
- **Enable Datechanger**
  Datechanger enables the user to alter the date of a document. This is especially useful when a document is scanned in from paper. The file creation date will not necessarily reflect the original (paper) document date. The datechanger can change the document date in the index.

- **Stemming**
  When Stemming is enabled, the indexer will trim down verbs to their stem for indexing.
- **Hits shown per page**
  Number of hits per page. Can be changed to “Show None” to allow metadata investigations.

17.1.2 OCR Settings

These settings allows you to tweak the OCR processing of the indexer. If needed, OCR can be completely disabled as well. Tweaking parameters like the OCR settings will greatly benefit the accuracy of the OCR processing.



The following settings can be changed:

- **Enable OCR?**

  This option specifies whether OCR should be enabled for this case or not.
  If disabled, the indexer will keep track of files that need OCR processing in case OCR will be done at a later stage.

- **Number of pages to OCR**

  This setting changes the number of pages that will be processed by the OCR engine.
  Lowering this number makes indexing faster, but may decrease indexing accuracy for OCR documents.
  This has nothing to do with the OCR quality, but rather the amount of data that is processed.

- **Primary OCR language**

  This should be the language that occurs the most times in the data set.

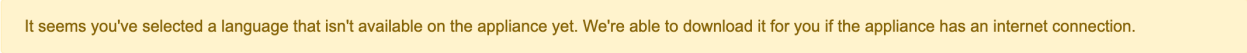
- **Secondary OCR language**

  This should be the language that occurs the most times *after* the primary language in the data set.

- **Extra OCR language(s)** ***(Only visible when Primary and Secondary languages are set)***

  Extra OCR languages can be entered here.

OCR depends on language packages for it's OCR engine. INDICA comes with a few languages pre-installed. If you select a language that is not installed, a message will be displayed:



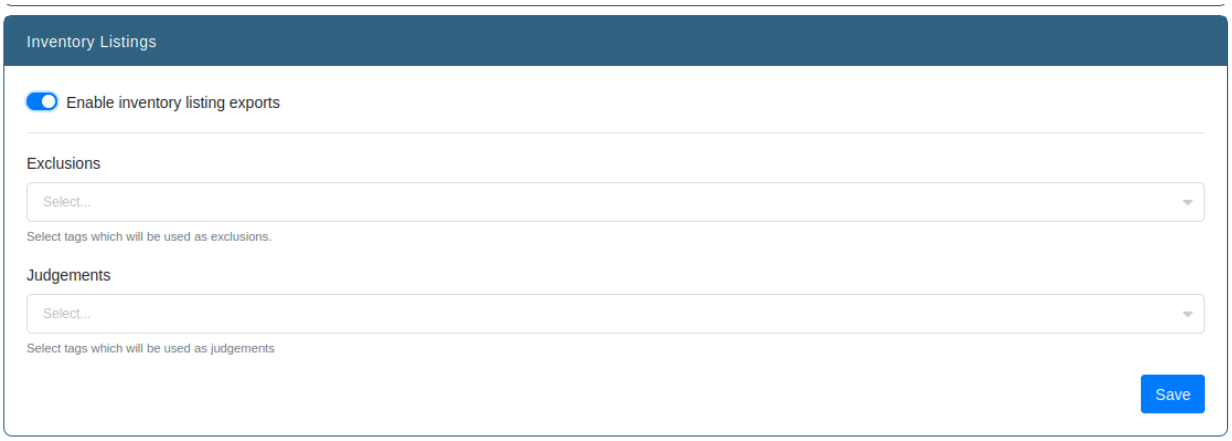
If the machine has an internet connection, the language package will be downloaded automatically.

Tip: It is advised to select (at least) the primary language. The secondary and extra languages are optional. If no languages are selected, the OCR engine will fall back to the default language; English.

Warning: The more languages are enabled, the slower OCR will become because it will try all selected languages on the encountered document. This can drastically slow down the OCR performance.

17.1.3 Inventory Listings

These settings allow to configure settings for creating inventory exports. You can read more about inventory exports here: *Case Manager Inventory Exports*.

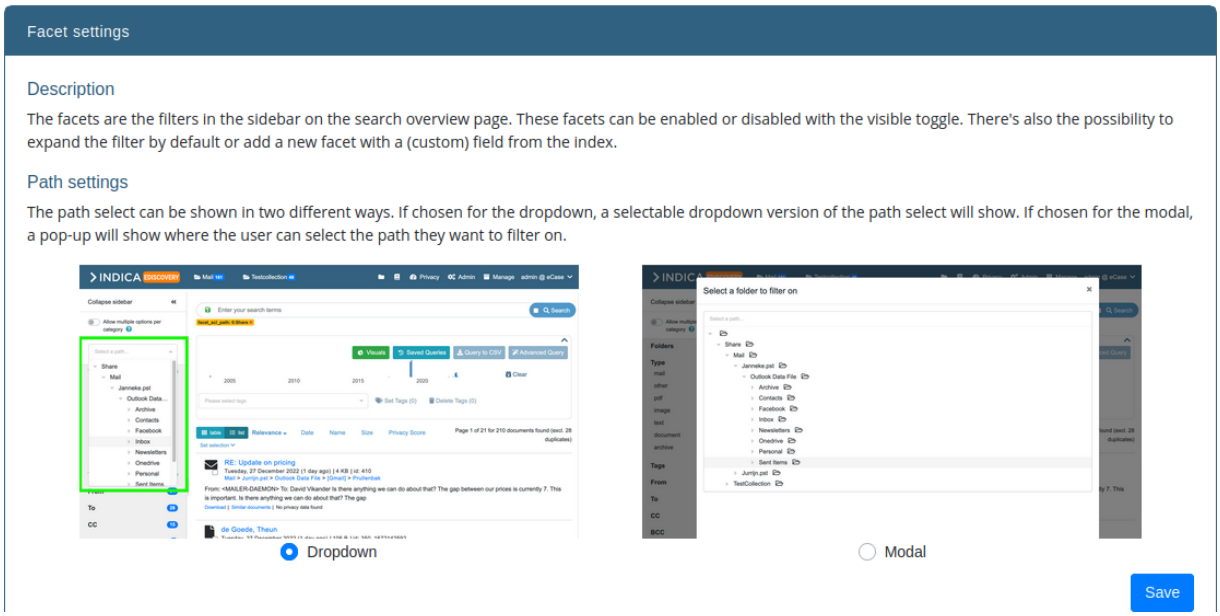


Here you can choose which tags are used as exclusions and which as judgements. These tags will be the ones that are available to choose from when creating inventory exports.

17.2 Display Settings

17.2.1 Facet Settings

Facet Settings allows you to manage and create filtering capabilities in the INDICA front-end.



The first thing available to configure is the path settings. The path select is available to the user in two ways: as a dropdown or as a modal.

The setting shows the list of currently active filters (facets) in the front-end. Don't see the facet list in your front-end? Click on the "Reset to default" button to seed the current facets. From the overview, it's possible to:

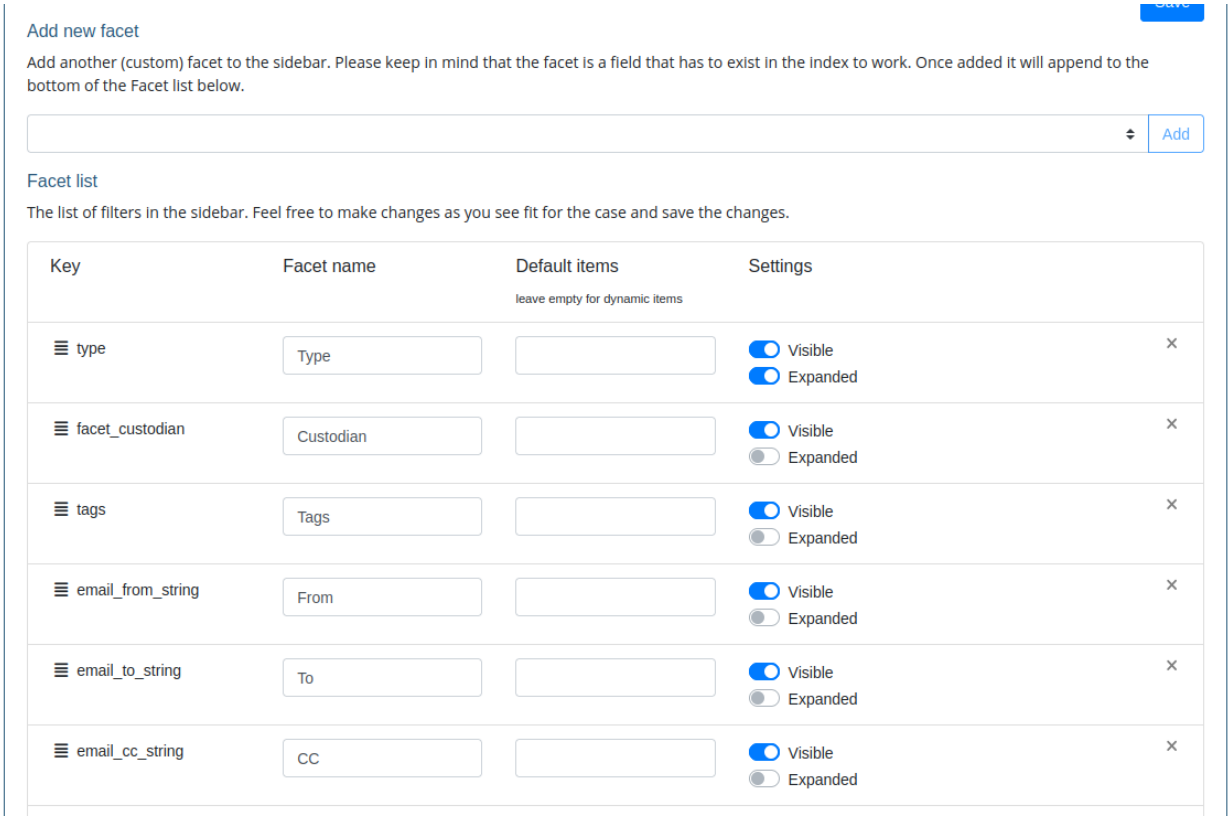
- **Add new facet**
  Select a new facet to add apart from the default ones from the dropdown, then click the "Add button".
- **Re-order facets**
  Re-ordering facets is possible by dragging the icon with the four horizontal lines. Changing the order of the facets will be reflected on the front-end.
- **Delete facets**
  Click the "X" icon on the right to remove a facet. It can always be re-added if needed.
- **Edit facets**
  It is possible to change some characteristics of a facet. You can change the following properties:
  - **Facet name**

    The name of the facet - what it will be called in the facets list on the Search page.
  - **Default items**

    These filter options will always be displayed, even when there are no documents associated with them.
  - **Settings**

    "Visible" either hides or displays this facet. Useful when the facet should be hidden.
    "Expanded" defines whether the facet category should be collapsed or expanded on the Search page.
- **Reset facets to default**
  When needed, the facets can be reset to the default setup by clicking the "Reset to default" button. This

will discard all changes made to the facets and restores the default set of facets. Use only when you want to go back to the default facets.

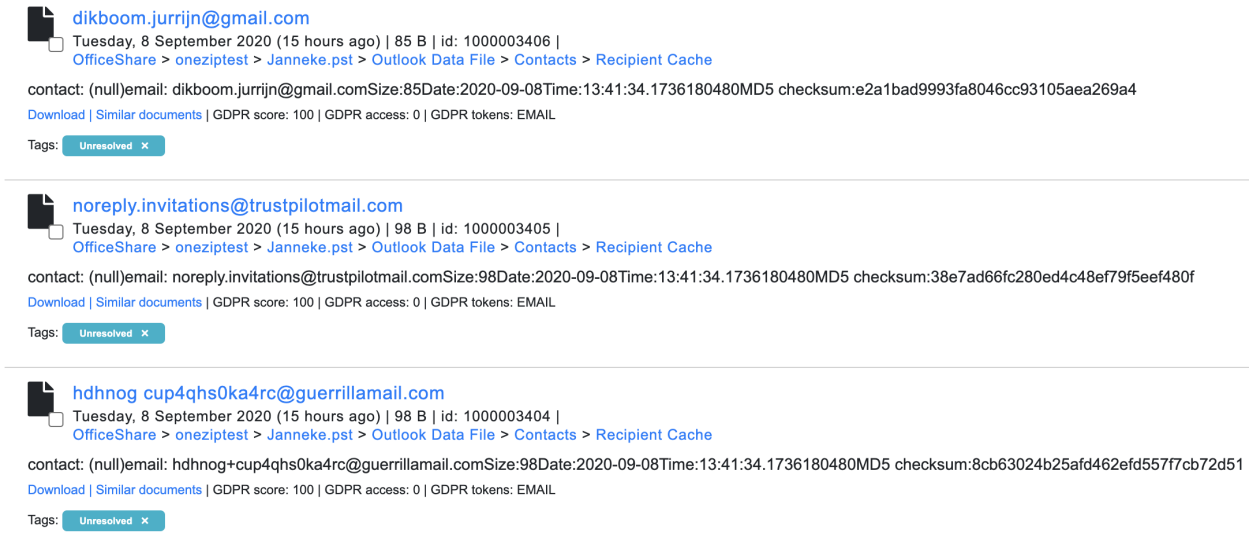


17.2.2 Result list Settings

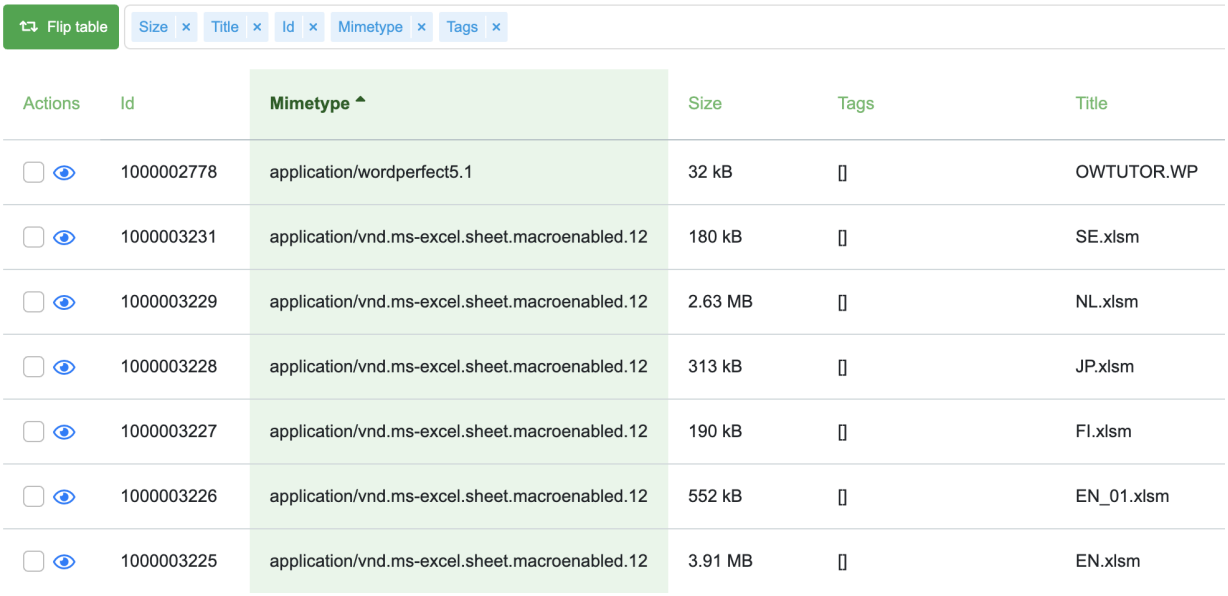
The following option can be changed on this page:

- **Allow user to toggle between list and table style**
  When enabled, a button will appear on the search page that allows the user to switch between list-style results and table-style results. The table-style results can be used for a custom overview of query results, as the displayed information can be chosen by the user. This feature is very powerful, but less easy to use.

Example of list-style preview:



Example of table-style preview:



Actions	Id	Mimetype	Size	Tags	Title
	1000002778	application/wordperfect5.1	32 kB	[]	OWTUTOR.WP
	1000003231	application/vnd.ms-excel.sheet.macroenabled.12	180 kB	[]	SE.xlsm
	1000003229	application/vnd.ms-excel.sheet.macroenabled.12	2.63 MB	[]	NL.xlsm
	1000003228	application/vnd.ms-excel.sheet.macroenabled.12	313 kB	[]	JP.xlsm
	1000003227	application/vnd.ms-excel.sheet.macroenabled.12	190 kB	[]	FI.xlsm
	1000003226	application/vnd.ms-excel.sheet.macroenabled.12	552 kB	[]	EN_01.xlsm
	1000003225	application/vnd.ms-excel.sheet.macroenabled.12	3.91 MB	[]	EN.xlsm

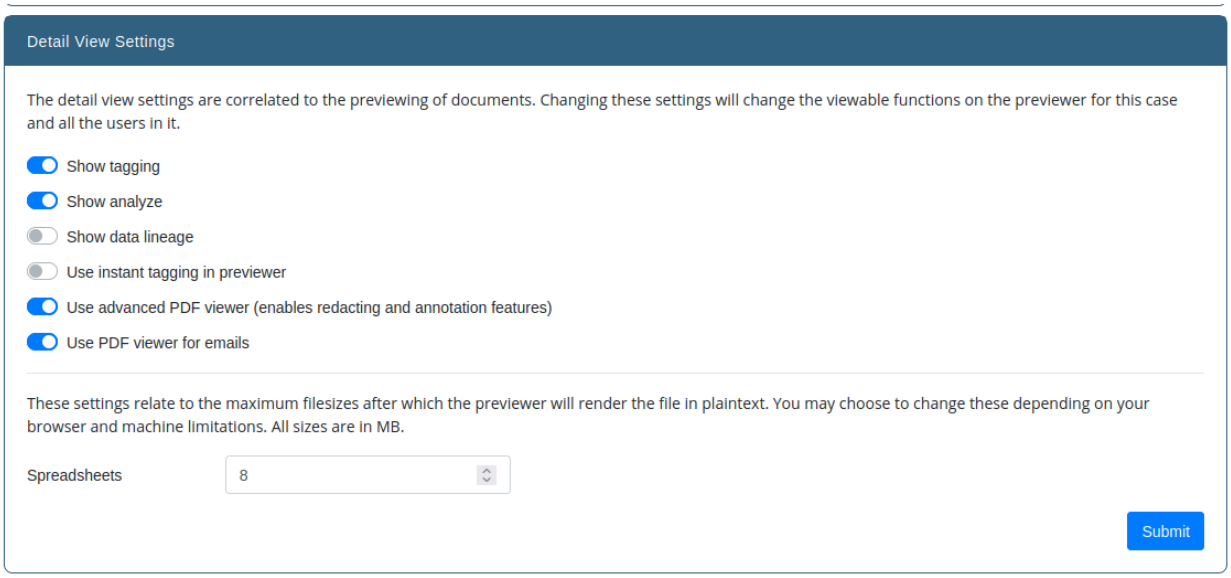
17.2.3 Index Progress Notification Settings

This setting allows to change the type of indexer progress notification. The user can choose to not display the message, have it be smaller or keep it in its normal format.

17.2.4 Detail View Settings

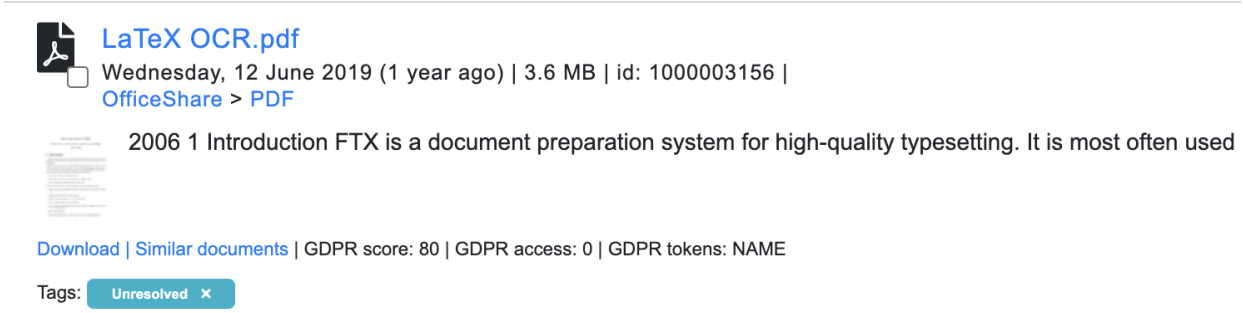
This panel allows you to change the following options:

- **Show Tagging**
  This option enables and disables the options to tag documents in the front-end. It is not possible to alter tags when this option is disabled.
- **Show analyze**
  This option enables and disabled the option to analyze a document in the front-end. It is not possible to alter tags when this option is disabled.
- **Show data lineage**
  This option enables and disables the options to show data lineage button in the front-end. The data lineage feature gives the user the possibility to select text and click on the lineage button. This will fire off a query with the selected text and the results will open automatically in the new "Data Lineage" tab. This feature makes it easier to find the 'single source of truth' when dealing with duplicate inform
- **Use instant tagging in the previewer.**
  When this option is enabled, when selecting a tag in the previewer it is immediately applied to the document, without having to take the extra step to "save" the applied tags.
- **Use advanced PDF viewer (enabled redacting and annotation features)**
  This setting requires PSPDFKit to be configured on your INDICA system. This setting enables the advanced PDF viewer for documents.
- **Use PDF viewer for emails**
  This setting enabls the advanced PDF viewer for viewing e-ma
- **Maximum spreadsheet size configuration**
  This setting sets the maximum spreadsheet size (in MB) that will be rendered in the front-end. Anything above this number will be rendered as plain text.



17.2.5 List Style Settings

These options allow for customizing the information in the list-style results.



The following options can be changed:

- **Show date?**
  Enable or disable the displaying of the date from result item.
- **Show file size?**
  Enable or disable the displaying of the file size from result item.
- **Show ID?**
  Enable or disable the displaying of the document ID from result item.
- **Show path?**
  Enable or disable the displaying of the file path from result item.
- **Show summary?**
  Enable or disable the displaying of the document summary from result item.
- **Show similar?**
  Enable or disable the displaying of the “Similar Documents” button from result item.
- **Show duplicates?**
  Enable or disable the displaying of the “Duplicates” button from result item.
- **show GDPR?**
  Enable or disable the displaying of GDPR tokens from result item.

17.2.6 Redaction Feature Settings (Advanced PDF Viewer)

Note: The Advanced PDF Viewer is not available on all INDICA appliances.

These settings allow you to customise the redaction feature in the advanced PDF viewer.

Auto-save

The first option is to auto-save redaction drafts. After drawing redactions, the draft needs to be saved manually. When this setting is enabled, this is not necessary anymore. The auto-save feature will only save the drafts of the redactions, applying them permanently to the document still needs to be done manually.

Redaction Patterns

These settings set the regex patterns that you can search by in the previewer. You can create new patterns, delete pre-existing ones, and temporarily disable the patterns making them not available to search by.

Redaction Styling

By default, the style of a redaction is a black box with no overlay text. There may be a situation where you want to, for example, redact all phone numbers in a document with a red box with white overlay text saying "PHONE". This is possible with this setting. Here you can set the name of the style, the color of the box and the overlay text. These styles will then be available in the Search & Redact option in the previewer's document actions.

17.3 Search Settings

The search settings allow you to manipulate the search results. This can be done by defining synonyms, stopwords, and editing the boosting settings.

17.3.1 Stopword List

During indexing, it is possible to exclude a list of stopwords from the index. Those words are generally words without informational value, like "a", "and", "this", "the", etc.

Words can be added by typing them in the text field and then pressing the "Add" button.

INDICA comes with a default lists of stopwords, which can be changed here as well.

17.3.2 Synonym List

It is also possible to define synonyms. This allows you to automatically broaden search results that contain words. Synonyms can be added in the text field and then clicking the "Add" button.

Synonyms need to be added as a comma-separated list, for example: "hello,hi,hey".

17.3.3 Boosting

Boosting can be done by adding a boost query, or by adding a boost function.

Boost Query

The Boost Query specifies an additional query clause that will be added to the user's main query to influence the score.

INDICA provides the relevance level of matching documents based on the results found. To boost a query, use the caret, "^", symbol with a boost factor (a number) at the end of the query you are searching. The higher the boost factor, the more relevant the query will be. Boosting allows you to control the relevance of a document by boosting its query. For example, if you are searching for

jakarta apache

And you want the term "jakarta" to be more relevant, boost it using the ^ symbol along with the boost factor next to the query. You would type:

jakarta^4 apache

This will make documents with the term jakarta appear more relevant. You can also boost Phrase Terms as in the example:

"jakarta apache"^4 "Indica search"

By default, the boost factor for each term or phrase is 1. Although the boost factor must be positive, it can be less than 1 (e.g. 0.2).

Boost Function

A Boost Function can also be added here. This feature is for advanced users. Please refer to the official documentation for help.

17.4 API Settings

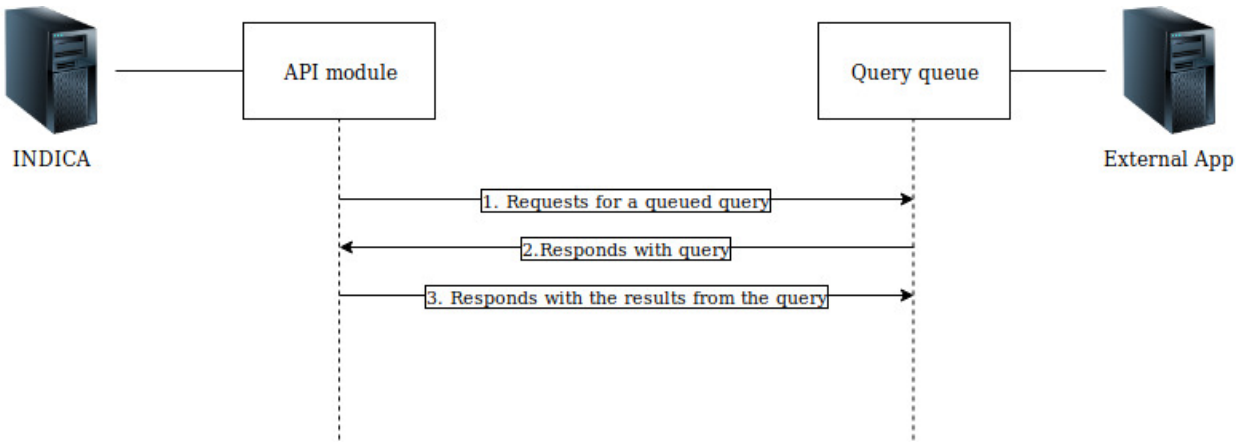
INDICA supports two types of API's. The polling API and the high level API.

17.4.1 Polling API

The INDICA Polling API service for external systems is a secure API that enables external systems to query the INDICA index in a limited way. It is designed for asynchronous communication initiated by the INDICA appliance (polling for instructions). The API is only able to transfer meta information, no actual content of document object can ever be transmitted to the external system asking for information.

The external system creates a queue of queries according to its needs in its own external API end point (the command set) to be run against the INDICA index. The INDICA pull API then returns the results back to the external systems designated API endpoint. Query results can return metadata from the index with addition to GDPR issues and for authorized users links to preview in the INDICA system itself.

This API is designed for on-premise INDICA systems and Internet-facing applications.



Setting up your Application

For the API to work, your application requires a very basic job system. A single Basic Auth or preferably OAuth protected endpoint(GET) that returns a query of commands/queries for INDICA to process. And a single protected endpoint(POST) which receives the results from INDICA.

Endpoint paths are configurable and you can create it according to your system limits/needs, but it should reside on two of the endpoints (not more) and explicitly POST endpoint has to match the specified pattern.

Table 1: Required Endpoints

Type	Path	Ext. App. JSON structure	Description
GET	/api/indica/jobs	*Job GET Endpoint JSON*	Response body has to match one of the job type structures
POST	/api/indica/jobs/{job_id}	*Job Result JSON*	These will be result that you will receive from INDICA once specific job type is done.

INDICA also accepts more job types than the ones shown in the examples above. API endpoints stay the same on your system, but required request end response JSON structures change based on every type.

Table 2: Job Types

Type	Job Queue JSON	Response JSON	Description
document_query	*Document Query JSON*	*Document Query Response JSON*	Document query returns hit counts and meta fields you requested
subject_details	*Subject Details Request JSON*	*Document Query Response JSON*	Returns the same values as Document Query. The only difference is that the query is generated by INDICA
document_export	*Export Documents JSON*	*Document Export Response JSON*	Tag export prepares a package on specified tags and uploads it to a specified vendor with specified security measures
gdpr_query	*GDPR Query JSON*	*GDPR Query Response JSON*	Returns GDPR issues on all or specific assets.

Batched result responses

All responses can be batched if the job request includes "batch" : 1 and batch_size fields in the JSON. batch_size has to be an integer of what is the maximum character count in the resulting JSON will be. Batched results have one extra JSON key ("batch_id") and an extra status - "inprogress". In which "batch_id" represents the order in a batch and status tells if the batch is still "inprogress" or is "processed".

Batch Response JSON

Setting up a polling job on the INDICA appliance

For job polling system to work correctly it is required to follow strict guidelines.

Settings configuration currently resides in Settings > API Settings inside the case management.

To create a new polling job, press the "Add new API polling job" button.

Polling jobs will run according to your settings:

- **Result data type**
  Can be changed based on what data you want to be returned to the external source. Currently, there is "Document query" - which returns document fields that you select further down in "Return fields", and there is also a choice to return "GDPR issues". Structure of returned data is described below;
- **Run query in user scope**
  Option specifies if the query should include only the data which adheres to a specific user's rights (could be API user with your setup in your AD);
- **User**
  Setting of a user who's scope will be used;
- **Time between polling**
  Is used to specify how often the system will check if there are new jobs in the GET endpoint, and run them. Running jobs too often might degrade the performance of your system;
- **Endpoint Authentication type**
  Specifies what type of authentication external endpoints use;
- **Job Queue GET endpoint**
  Specifies endpoint from which jobs list in a JSON format will be retrieved;
- **Result POST endpoint**
  Specifies endpoint to which the queried Document field(s) or GDPR issues will be sent;
- **Require acknowledge**
  Specifies if acknowledgment of received jobs should be given. It is strongly recommended to use acknowledgment since it negates duplicate jobs when polling time is shorter, or queries are more difficult. It guarantees that the same job doesn't get picked up twice or more;
- **Job acknowledgment endpoint**
  Specifies an endpoint to which acknowledgment will be done.

17.4.2 PowerBI Connection

With INDICA, it is possible to connect PowerBI and use the modeling capabilities of PowerBI to create your own dashboarding.

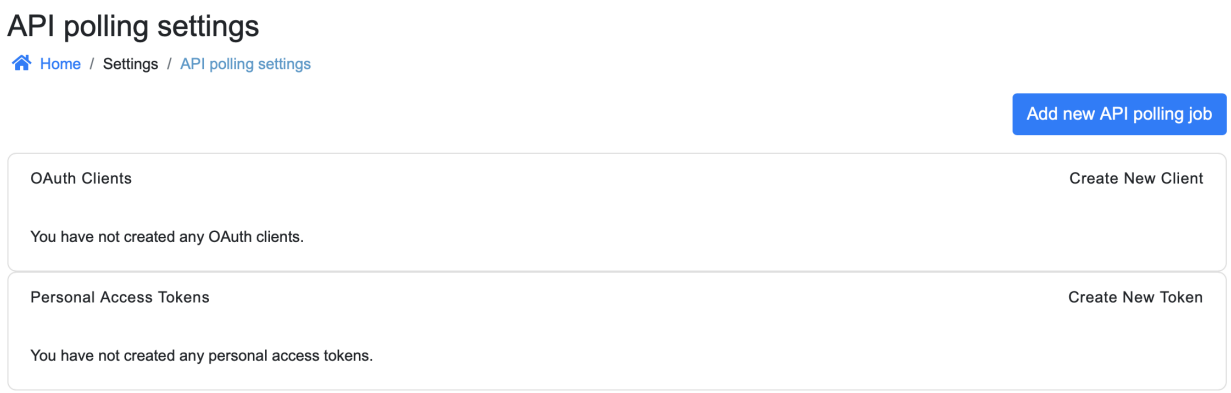
Setting up an API connection between INDICA and PowerBI requires some setup. The steps needed to take are:

1. Creating an OAuth client
2. Creating a personal access token
3. Connecting PowerBI to INDICA
4. Creating request URLs

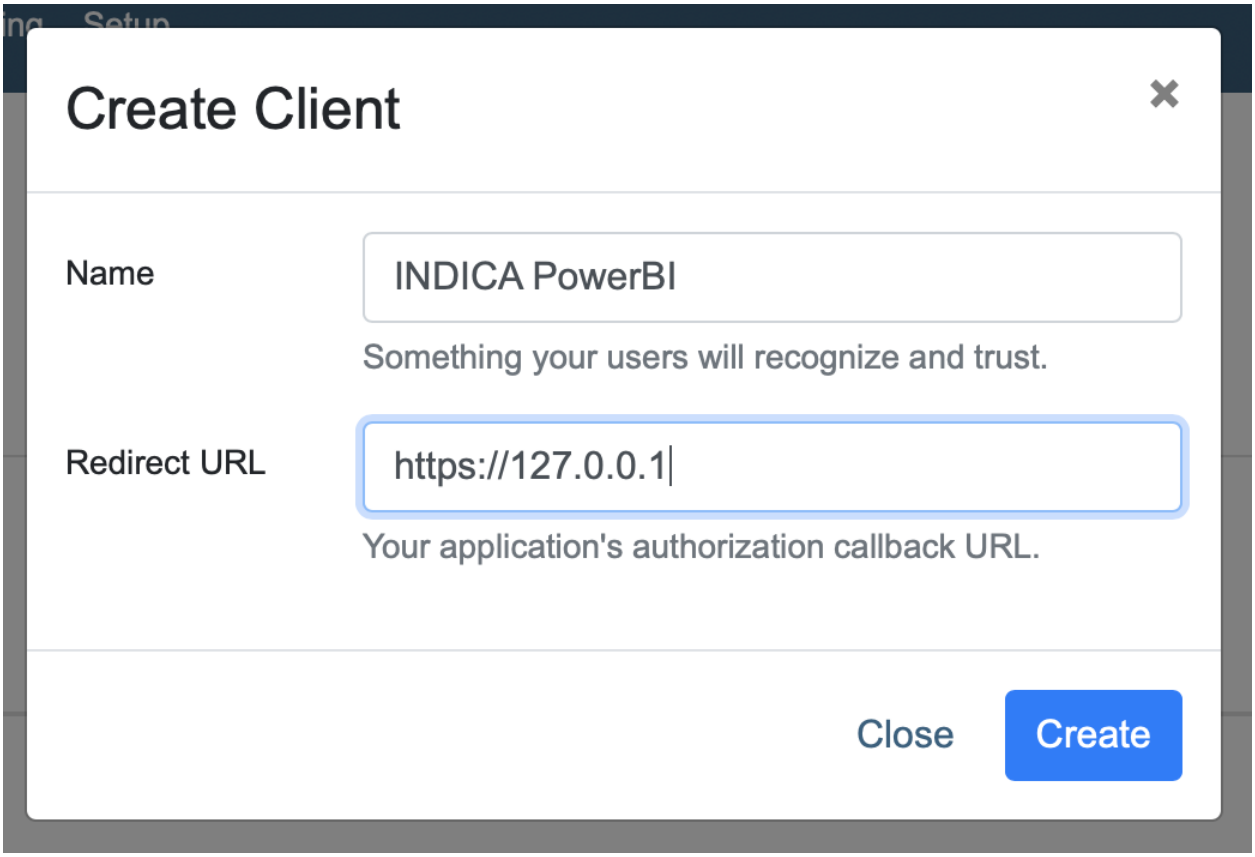
In the following sections, the above steps will be described in detail.

Creating an OAuth client

First, navigate to the API Settings page under 'Manage' > 'Settings' > 'API Settings'. You will be presented with the following page:



Here, we can create a new OAuth client by clicking the "Create New Client" button. A popup will open:

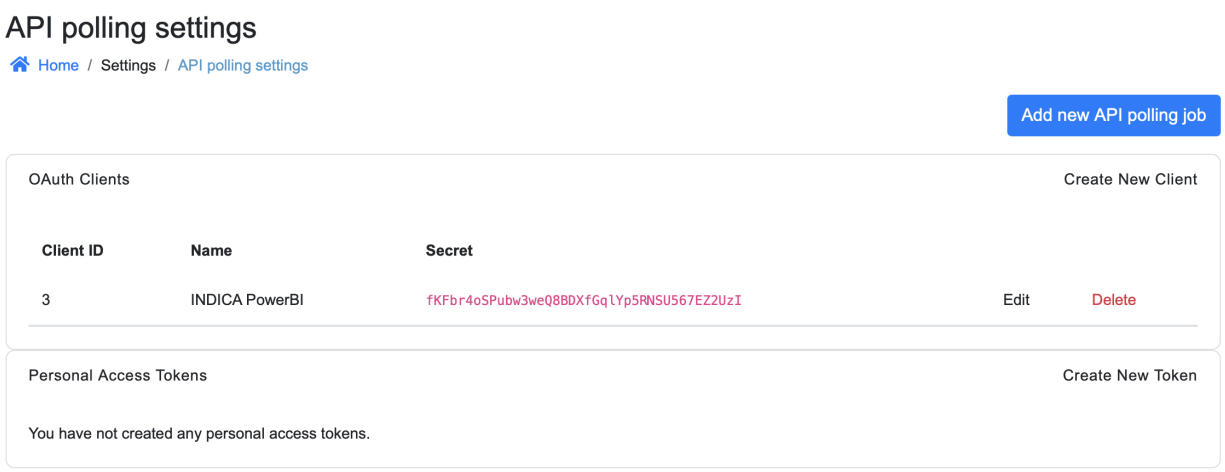


The following information is required:

- **Name**
  Choose a descriptive name so you will recognise it later on.
- **Redirect URL**
  If you don't know what to fill in as redirect URL, you can use *https://127.0.0.1*, this will be a placeholder.

After filling in all fields, click "Create" to create the OAuth client.

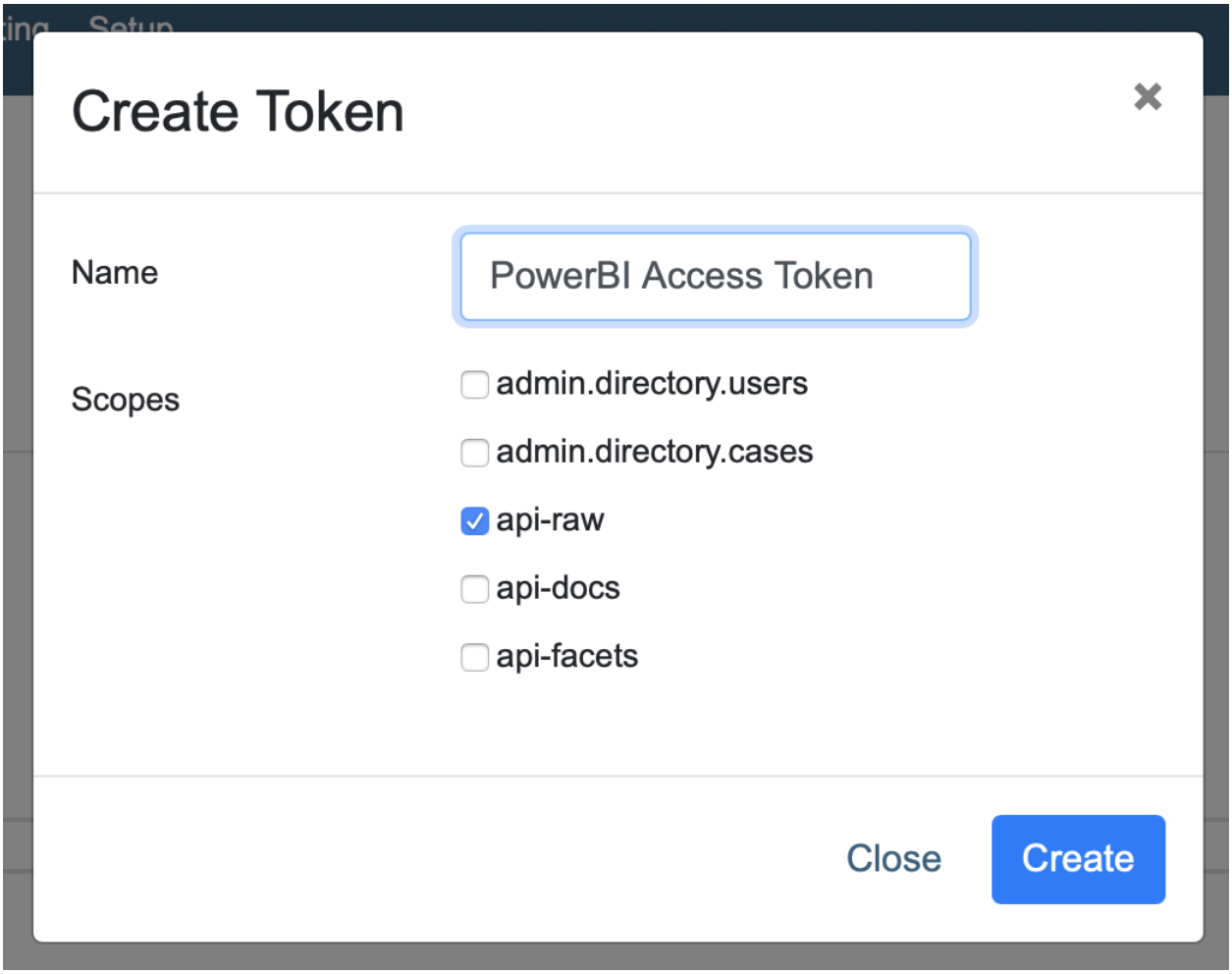
The settings page should look like this:



Creating a Personal Access Token

The next step is creating a Personal Access Token. This token will be used by PowerBI to let INDICA know that it has access to the data. The access token will give unlimited access to the data within INDICA; be careful with it. Do not share it with untrusted parties.

A new token can be created by clicking the "Create New Token" button. A popup will open:



The following information is required:

- **Name**
  Choose a descriptive name so you will recognise it later on.
- **Scopes**
  Scopes define the kind of information is retrievable by the API connection.

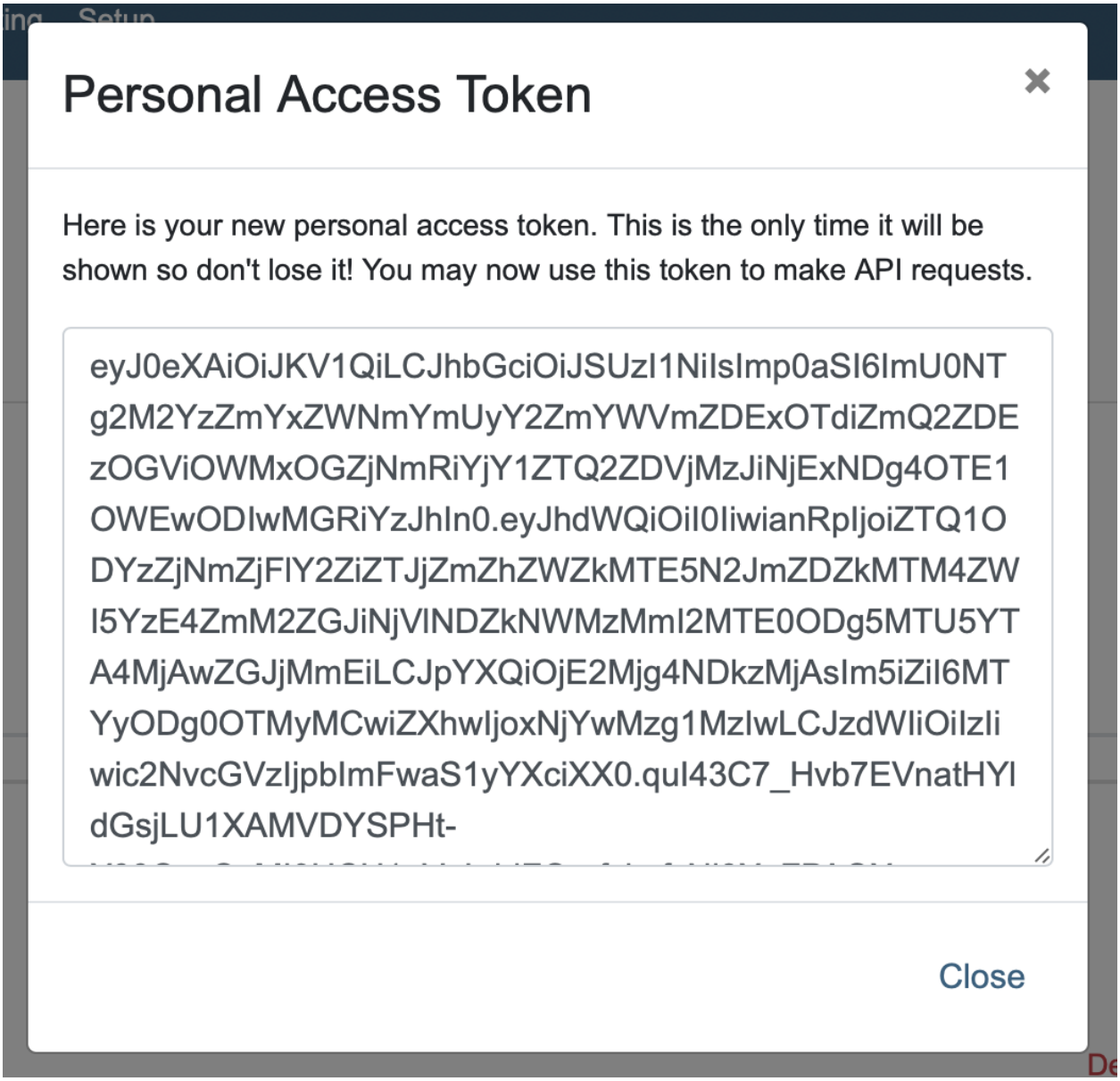
The following 5 scopes are available to the user:

- **admin.directory.users**
  This scope gives the user the ability to: Read and modify users.
- **admin.directory.cases**
  This scope gives the user the ability to: Read and modify cases.
- **api-raw**
  This is the most powerful scope, this scope allows the user to send an receive raw SOLR data.

- **api-docs**
  This scope gives the user access to the documents in the index.
- **api-facets**
  This scope gives the user access to the facets.

For the PowerBI connection, we will be using the scope "api-raw". After checking the box for `api-raw` click "Create" to generate your Personal Access Token.

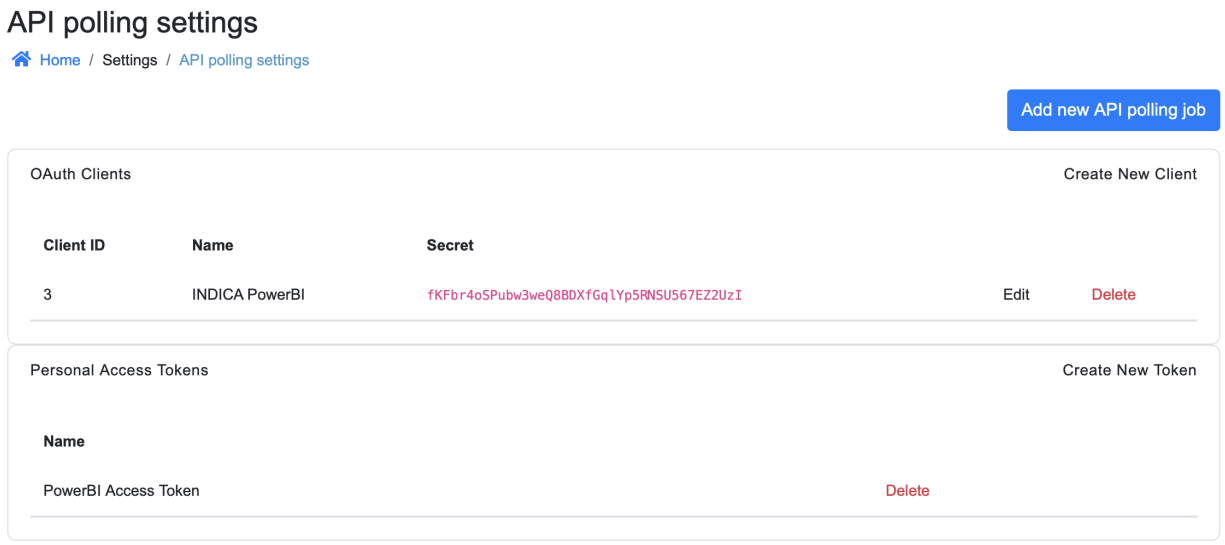
You will now be presented with your personal access token. Make sure to save this token somewhere safe. It will be needed later on.



Warning: This is the only time you will be presented with this token. When the popup is closed, there is no way to retrieve it. Make sure to store this somewhere safe.

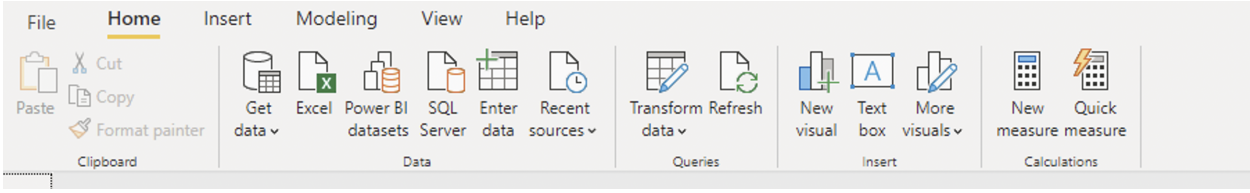
When done, click the "Close" button to close the popup.

The settings page should look like this:

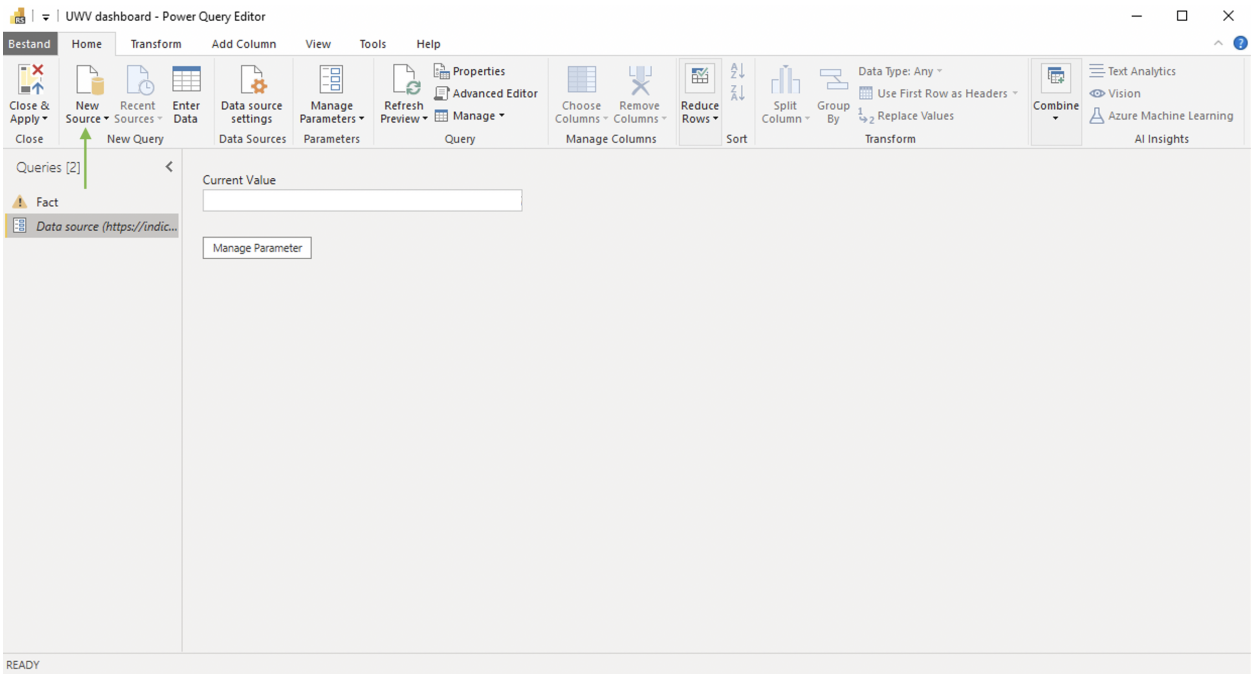


Connecting PowerBI to INDICA

When above steps are completed, the connection between INDICA and PowerBI can be made. Open PowerBI and click the button “Transform data”.

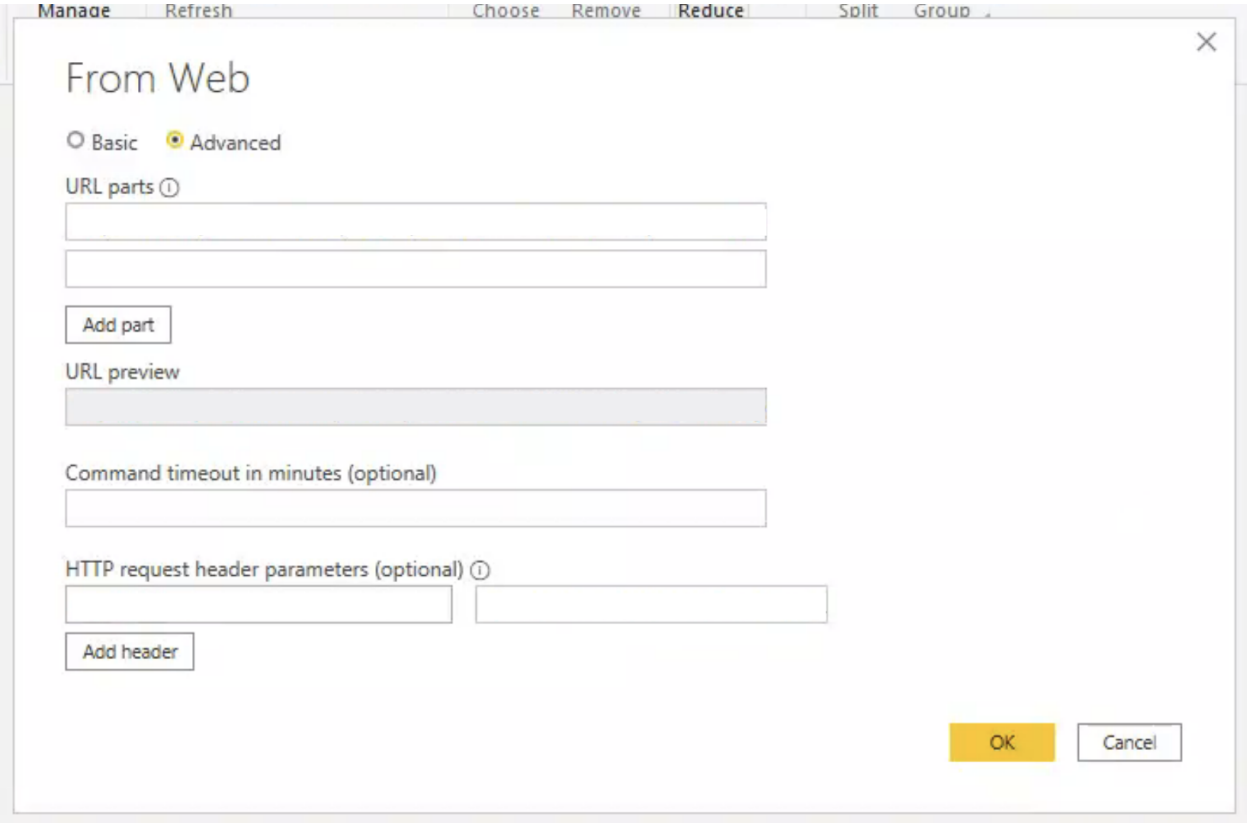


When clicked, the Power Query Editor will open.



Click “New Source” in the menubar on the top (marked by the green arrow). Choose “Web” from the list of available options. A popup will open:

In this popup, choose the “Advanced” option. The popup will now look like this:



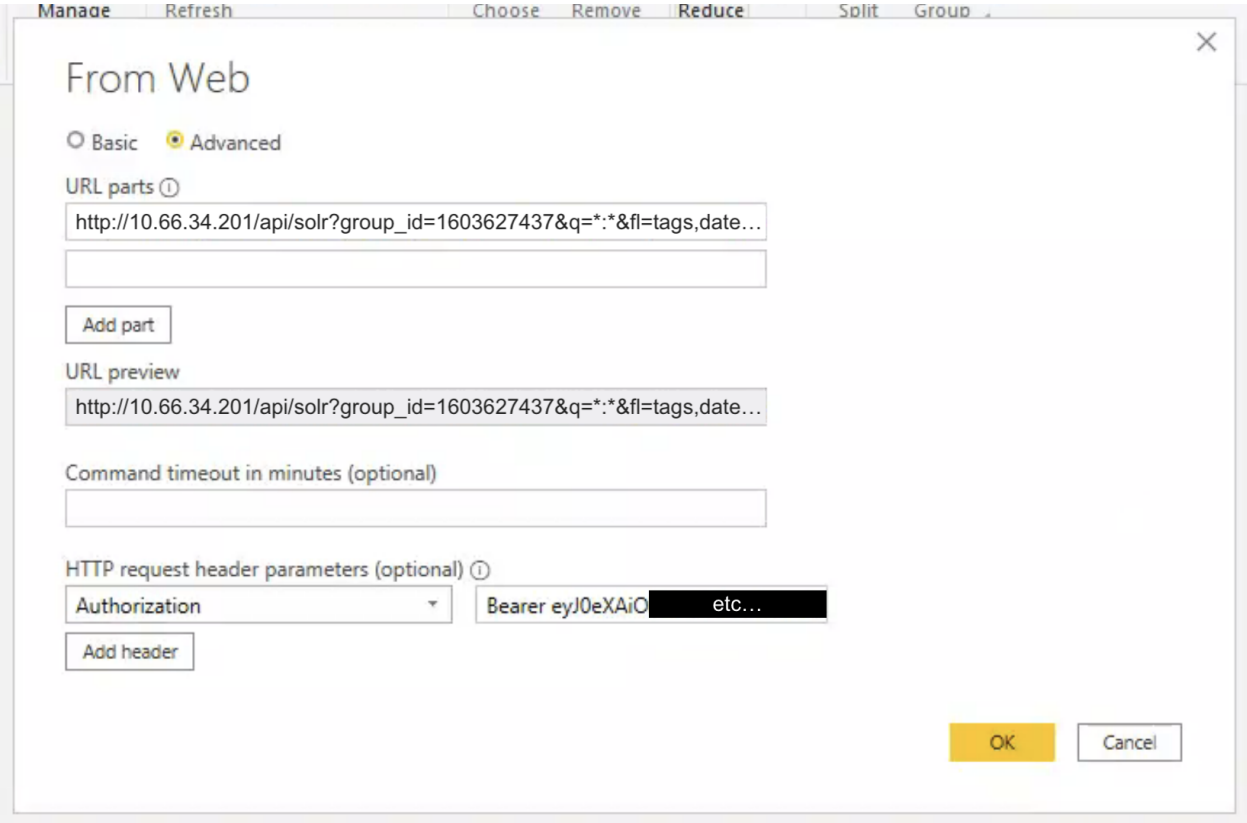
The following fields are required:

- URL parts
- HTTP request header parameter

Adding the access token to the request

INDICA will not return any data when the access token is not set correctly. To configure PowerBI to include the access token, type the word `Authorization` in the field on the left under "HTTP request header parameters". Then, in the field on the right, type in the word `Bearer` followed by a space, and then the complete Personal Access Token.

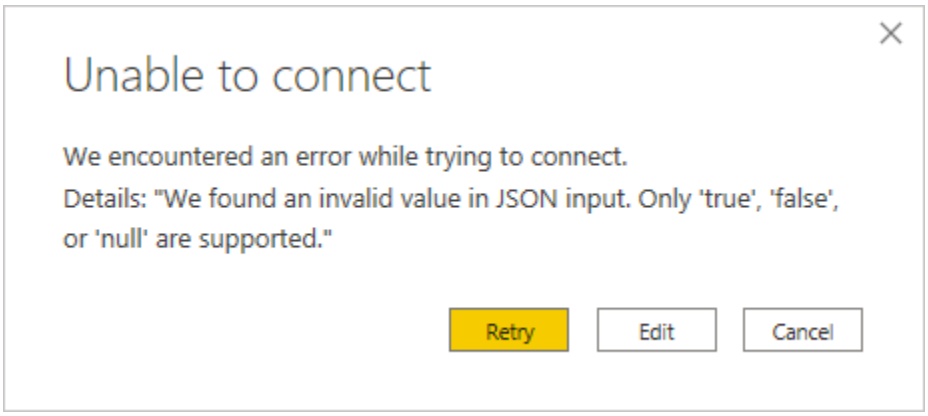
The popup should now look like this:



When everything is filled in, click "OK".

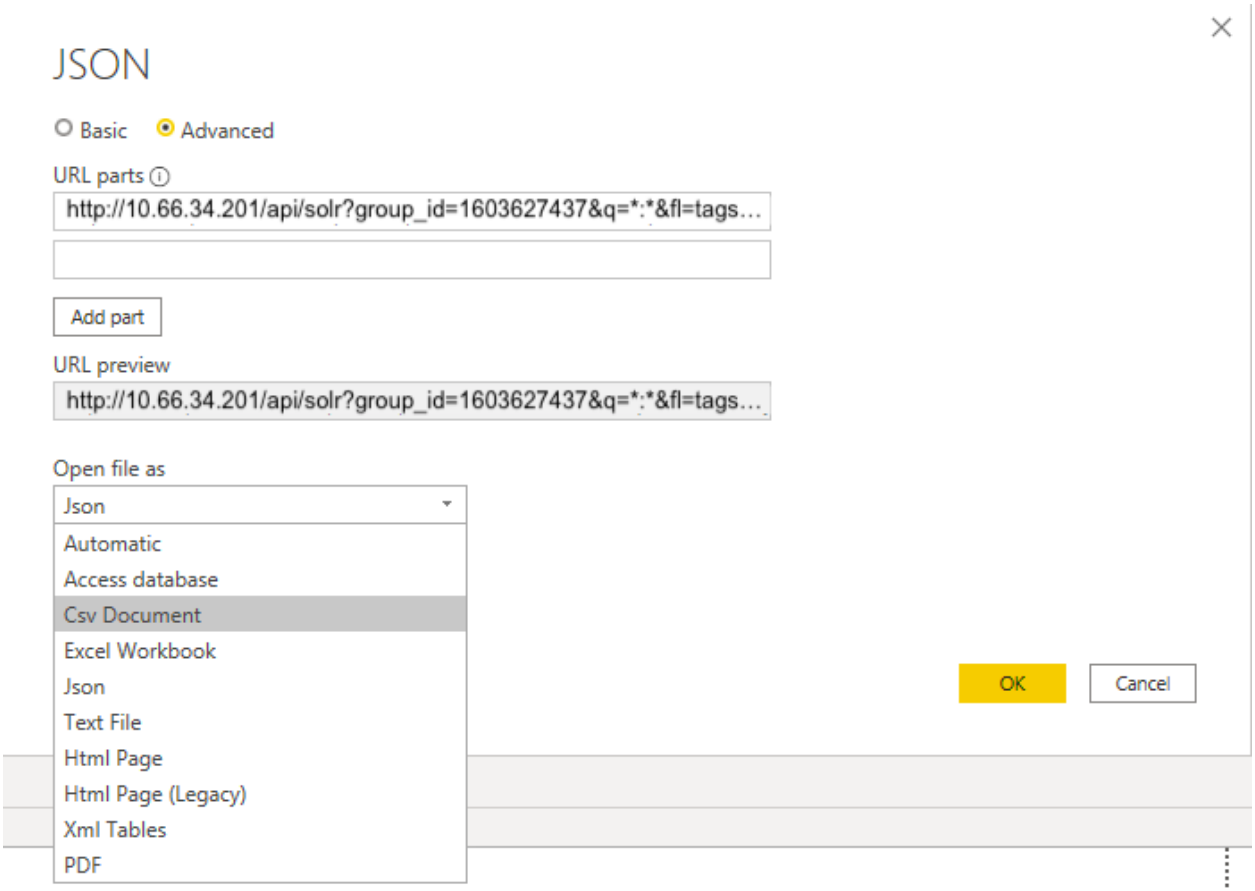
If a question pops up asking "Select which level to apply these settings to", pick the option that ends with `/api/solr`. Then, click "Connect". Your connection with PowerBI should now be made.

In some occasions, PowerBI will return an error stating that it is unable to connect:



This happens because INDICA will return data in CSV format, while PowerBI expects JSON. Click "Edit" when the above error is displayed.

Then the following popup will open:



Under "Open file as", select the option for "Csv Document". PowerBI will now recognise the output that INDICA generates. Finally, click "OK" to confirm. You should now be presented with a preview of the selected data.

The INDICA and PowerBI connection is now ready for use.

Creating request URLs

The URL can be created by combining serveral parameters:

- The URL or IP of the INDICA appliance (called `URL` in the example);
- The ID of the case on the INDICA appliance (called `GROUP_ID` in the example);
- The query you want to use (called `QUERY` in the example);
- A list of the fields to retrieve (called `FIELD_LIST` in the example);
- The number of rows that will be retrieved (called `NUM_ROWS` in the example).

Above parameters may be placed in the following format:

```
http(s)://URL/api/solr?group_id=GROUP_ID&q=QUERY&fl=FIELD_LIST&wt=csv&rows=NUM_ROWS
```

Note that depending on the appliance, either `http` or `https` needs to be used. A working example of this would be:

```
http://10.66.34.201/api/solr?group_id=1603627437&q=*:*&fl=tags,date,
type&wt=csv&rows=100000
```

Note: For no query, use *:* as query

When the URL has been prepared, you can copy and paste it in the "URL parts" field in PowerBI.

17.4.3 Polling API Examples

Job GET Endpoint JSON

```
[
    {
      "job_id": "47ec1fe7db822300e5c4e1bb4b961972",
      "job_type": "document_query",
      "q": "\"David Vikander \"~0",
      "fq": "",
      "return_fields": [
        "file_name",
        "mime_type",
        "size"
      ],
      "status": "processed"
    },
    {
      "job_id": "47ec1fe7db822300e5c4e1bb4b961972",
      "job_type": "document_query",
      "q": "\"Dead Kennedys\"~0 OR \"test\"~0",
      "fq": "",
      "return_fields": [
        "file_name",
        "mime_type"
      ],
      "status": "processed"
    }
]
```

Job Result JSON

```
{
    "job_id": "47ec1fe7db822300e5c4e1bb4b961972",
    "job_type": "document_query",
    "result_url": "http://192.168.2.128/search?q=%22David%20Vikander%22~0&page=1&fq={
↪%22GDPR_Combined%22:[%22NAME%22]}&sort=score%20desc&mlt=0",
    "result_collection": {
      "count": 2,
      "entries": [
        {
          "file_name": "John_Doe.doc",
          "mime_type": "doc"
        },
        {
```

(continues on next page)

(continued from previous page)

```
          "file_name": "John Doe.pst",
          "mime_type": "pst"
        }
      ]
    },
    "status": "processed",
    "error_message": "error"
}
```

Document Query JSON

```
[
    {
      "job_id": "47ec1fe7db822300e5c4e1bb4b961972",
      "job_type": "document_query",
      "q": "\"David Vikander \"~0",
      "fq": [
        {
          "tags": [
            "2_todelete"
          ]
        }
      ],
      "return_fields": [
        "file_name",
        "mime_type",
        "size"
      ],
      "exporting": true,
      "status": "processed"
    },
    {
      "job_id": "47ec1fe7db822300e5c4e1bb4b961972",
      "job_type": "document_query",
      "q": "\"Dead Kennedys\"~0 OR \"test\"~0",
      "fq": "",
      "return_fields": [
        "file_name",
        "mime_type"
      ],
      "exporting": true,
      "status": "pending"
    }
]
```

Document Query Response JSON


```
{
    "job_id": "47ec1fe7db822300e5c4e1bb4b961972",
    "job_type": "document_query",
    "result_url": "http://192.168.2.128/search/SCA0001052?q=a",
    "result_collection_urls": [
      {
        "name": "TestCollections",
        "url": "http://192.168.2.128/search/SCA0001052?module=TestCollections&q=a"
      },
      {
        "name": "Sites",
        "url": "http://192.168.2.128/search/SCA0001052?module=Sites&q=a"
      }
    ],
    "result_collection": {
      "count": 2,
      "entries": [
        {
          "file_name": "John_Doe.doc",
          "mime_type": "doc"
        },
        {
          "file_name": "John Doe.pst",
          "mime_type": "pst"
        }
      ]
    },
    "tags": [
      {
        "title": "To Delete",
        "value": "2_todelete"
      },
      {
        "title": "Unresolved",
        "value": "1_unresolved"
      }
    ],
    "status": "processed",
    "error_message": "error"
}
```

Subject Details Request JSON

```
{
    "job_id": "SCA0001017",
    "job_type": "subject_details",
    "subject_details": {
      "name": "John Doe",
      "tel": "88123456789",
      "email": "jonh@example.com",
      "address": "Elm Street 13"
    },
    "status": "pending"
}
```

Export Documents JSON

```
[
    {
      "job_id": "47ec1fe7db822300e5c4e1bb4b961972",
      "job_type": "document_export",
      "packaging": 1,
      "export_vendor": "box",
      "query_job_id": "56rsbs0e5css51972",
      "protection": "pass",
      "subject_email": "test@example.com",
      "status": "processed"
    },
    {
      "job_id": "8e6d5fe7db822300e5c4e1bb4b961947",
      "job_type": "document_export",
      "packaging": 1,
      "export_vendor": "box",
      "query_job_id": "56rsbs0e5css51972",
      "protection": "pass",
      "subject_email": "test@example.com",
      "status": "processed"
    }
]
```

Document Export Response JSON

```
{
    "job_id": "47ec1fe7db822300e5c4e1bb4b961972",
    "job_type": "document_export",
    "package_url": "https://example.com",
    "status": "processed",
    "error_message": "error"
}
```

GDPR Query JSON

```
[
    {
      "job_id": "e6e31e4ddb032300e5c4e1bb4b9619fa",
      "job_type": "gdpr_query",
      "q": "",
      "fq": "",
      "status": "processed"
    }
]
```

GDPR Query Response JSON

```
{
    "job_id": 64,
    "job_type": "gdpr_query",
    "result_collection": {
      "count": 1,
      "entries": [
        {
          "asset": {
            "type": "collection",
            "identifiers": {
              "name": "Unstructured Data",
              "ip": "198.168.1.0",
              "path": "/dropbox/JohnDoe/things"
            }
          },
          "issues": [
            {
              "type": "NAME",
              "count": 227
            },
            {
              "type": "EMAIL",
              "count": 53
            },
            {
              "type": "TEL",
              "count": 11
            },
            {
              "type": "CC",
              "count": 4
            }
          ]
        }
      ]
    },
    "status": "processed",
```

(continues on next page)

(continued from previous page)

```
    "error_message": "error"
}
```

Batch Response JSON

Response 1

```
[
    {
      "job_id": "47ec1fe7db822300e5c4e1bb4b961972",
      "job_type": "document_query",
      "batch_id": 0,
      "result_url": "indica.lan/query=JonhDoe",
      "result_collection": {
        "count": 2,
        "entries": [
          {
            "file_name": "John_Doe.doc",
            "mime_type": "doc"
          },
          {
            "file_name": "John Doe.pst",
            "mime_type": "pst"
          }
        ]
      },
      "status": "inprogress"
    }
]
```

Response 2

```
[
    {
      "job_id": "47ec1fe7db822300e5c4e1bb4b961972",
      "job_type": "document_query",
      "batch_id": 1,
      "result_url": "indica.lan/query=JonhDoe",
      "result_collection": {
        "count": 2,
        "entries": [
          {
            "file_name": "test.pst",
            "mime_type": "email"
          },
          {
            "file_name": "test.txt",
            "mime_type": "text"
          }
        ]
      },
```

(continues on next page)

(continued from previous page)

```
        "status": "processed"
    }
]
```

CHAPTER

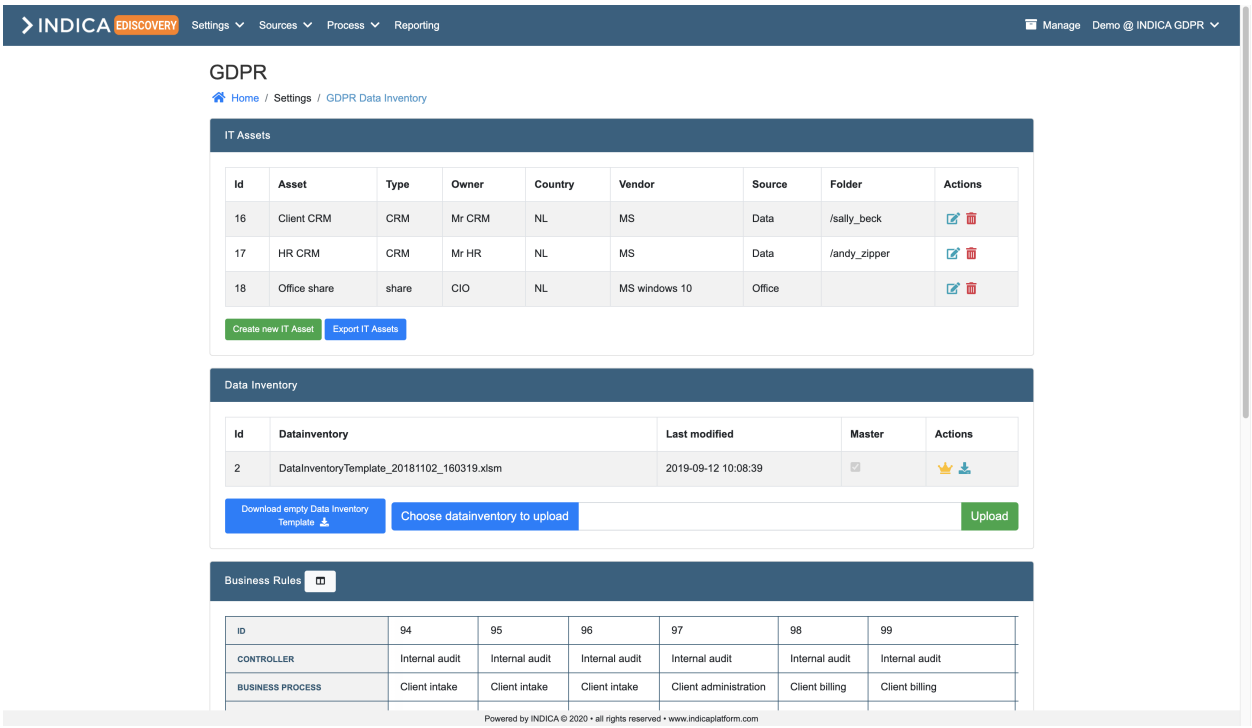
EIGHTEEN

GDPR DATA INVENTORY MANAGE

Table of Contents

- *GDPR Data Inventory Manage*
  - *IT Assets*
  - *Data Inventory*
  - *Business Rules*

GDPR Data Inventory can be found in the manage interface, underneath the *Settings* menu item. This item only shows up if GDPR is enabled within your case.



18.1 IT Assets

Within this panel it's possible to add and export IT Assets. To add an asset, click the 'Create new IT Asset' button. An overlay will open in which you can fill in the following fields:

- *Asset*

  Name of your asset. This field is required.

- *Type*

  The type of asset your new asset will be. This field is required.

- *Owner*

  Who the owner(s) are of the asset. This field is required.

- *Country*

  Which country the asset applies to.

- *Vendor*

  From which vendor the IT asset is. This field is required.

- *Collection*

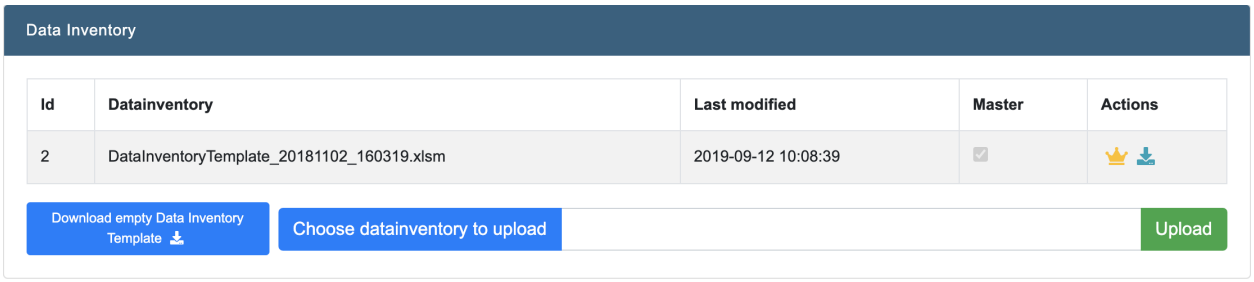
  Which collection the asset applies to. This field is required.

- *Folder*

  Which folder the asset data can be found.

18.2 Data Inventory

Within this panel it's possible to add data inventory items. You can either download an empty template for you to fill in, or upload a file you have ready. To upload your data inventory file, click on the 'Choose data inventory to upload' button and click on your file to add. Then click the upload button and the file will upload. Afterwards, you should see your data inventory item(s) in the table.



18.3 Business Rules

A business rule can be added in two different ways. You can import a file with business rules if you have one, or add them manually. To add them manually, click the 'Create new rule' button. An overlay with a form will open, which you can fill in. Fill in the fields and click on the 'Save' button to add the business rule to your set.

The existing business rules will show the fields you filled in. If you're missing fields, or want some fields 'out of view', you can click the button next to the 'Business Rules' title in the header. Here, you can (un)check boxes whether you want to see them or not. With the six-dotted icons most left to the items you can drag them around to sort them in the way you want to view them.

The business rules will also show how many records are within that rule. You can view them by clicking on the eye-icon. This will open a new tab in your browser with the filters active on the search page.

Business Rules

ID	94	95	96	97	98	99
CONTROLLER	Internal audit	Internal audit	Internal audit	Internal audit	Internal audit	Internal audit
BUSINESS PROCESS	Client intake	Client intake	Client intake	Client administration	Client billing	Client billing
PERSONAL DATA TYPES	NAME	EMAIL	TEL	NINO	NI	CC
PROCESSING ACTIVITY	Contacting	Contacting	Contacting	Billing	Billing	Billing
CATEGORIES OF DATA SUBJECTS	Customers	UNKNOWN	Customers	Customers	Customers	Customers
CATEGORIES OF PERSONAL DATA						
IT ASSET	Client CRM	Client CRM	Client CRM	Client CRM	Client CRM	Client CRM
PURPOSES OF PROCESSING						
LEGAL GROUND		1.e				
PROCESSOR						
INTERNATIONAL TRANSFER	0	0	0	0	0	0
RETENTION TERM	0	0	0	0	0	0
DATA TYPE CLASSIFICATION						
TECHNICAL MEASURES	Authentication	Authentication	Authentication	Authentication	Authentication	Authentication / encryption
ORGANISATIONAL MEASURES						
RECORDS	199	0	11	5	4	1
OPEN						
ACTIONS						

Create new rule + Import Data Inventory

CHAPTER

NINETEEN

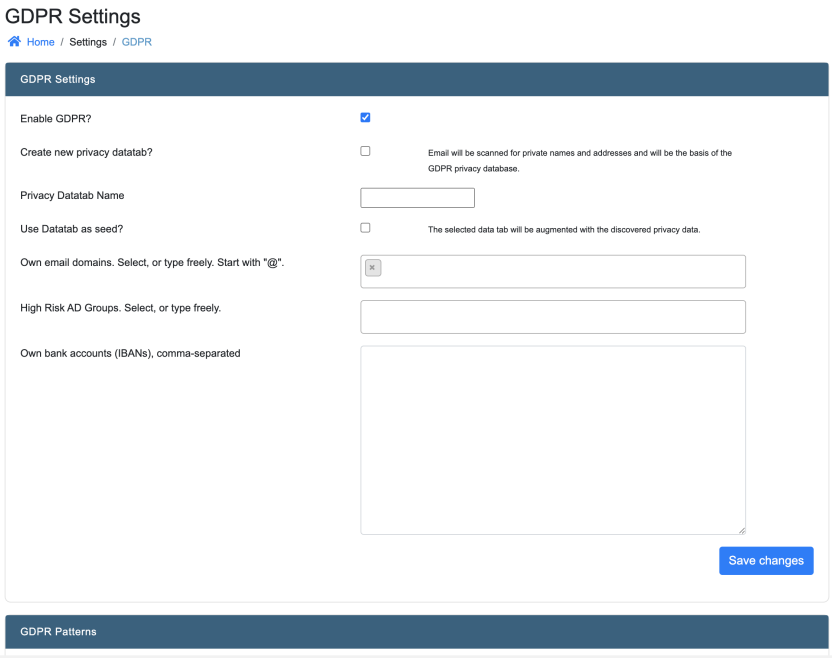
GDPR SETTINGS MANAGE

Table of Contents

- *GDPR Settings Manage*
  - *GDPR Settings*
  - *GDPR Patterns*
  - *GDPR Threat Matrix*
  - *GDPR regular vs. sensitive personal information*
  - *Bag of Words*

GDPR Settings can be found in the manage interface, underneath the *Settings* menu item. This item only shows up if GDPR is enabled within your case.

Note: To fully enable GDPR within your case please make sure you enable GDPR within the GDPR settings page *before indexing*! GDPR can also be enabled afterwards, but your case will need to be fully reindexed, so it's better to enable it beforehand.



19.1 GDPR Settings

Within this panel you can do the following:

- *Enable / Disable GDPR*

  Enabling GDPR is needed if you want to fully enable GDPR within your case.

- *Create new privacy datatab*

  This will create a seperate datatab that can function as a filter.

- *Privacy datatab name*

  The name for your privacy datatab so you can find it easily in the menu.

- *Enable / Disable use datatab as seed*

  With this option you're able to enable the datatab as a seed.

- *Own email domains*

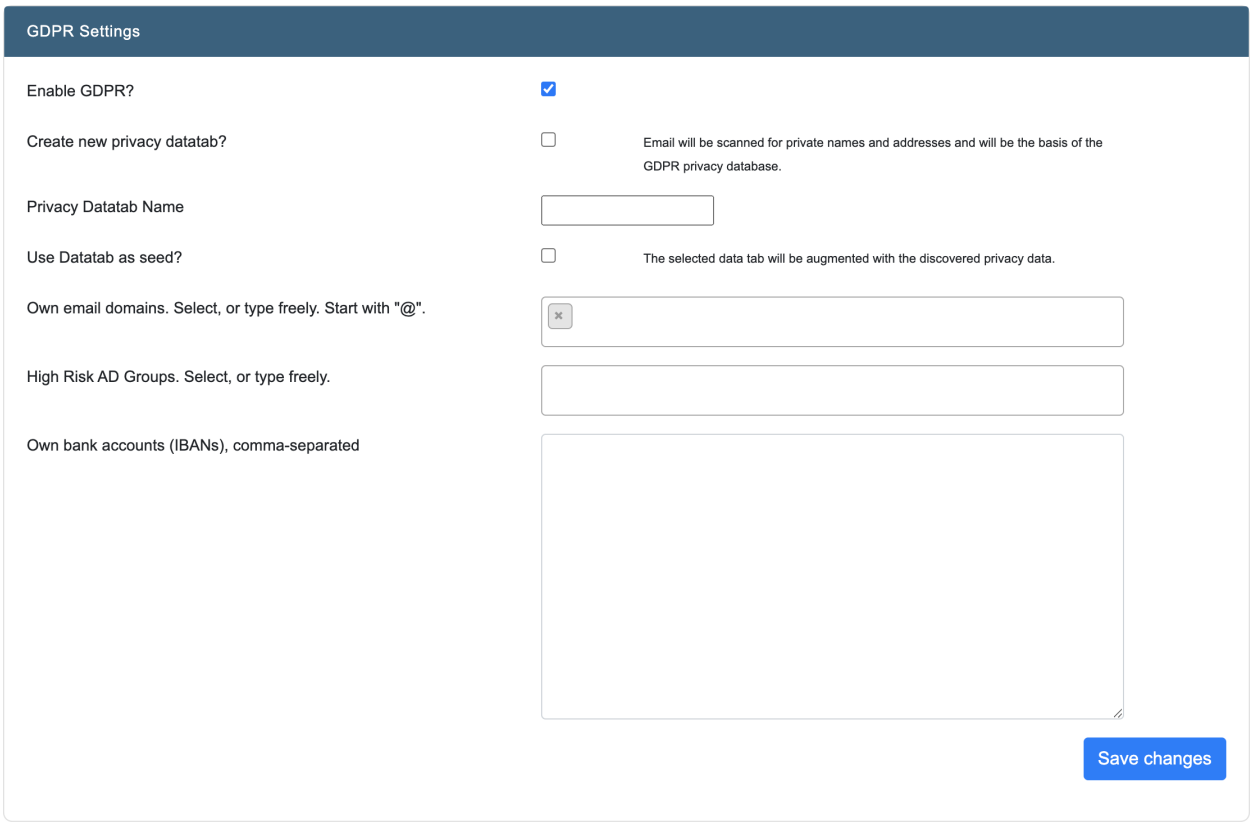
  Adding your own email domains will make sure they are not added as a privacy record.

- *High Risk AD Groups (if AD is connected)*

  If you have an AD connected, add AD groups here that are possible high risks for your case.

- *Own bank accounts (IBANs)*

  Add own bank accounts (IBANs) so these are filtered out of the sensitive data.



19.2 GDPR Patterns

Within this panel it's possible to add GDPR patterns. These 'patterns' are regular expressions which will capture the specified values in your data. INDICA comes with a few standard GDPR patterns already in there, but it's also possible to add your own.

To add your own GDPR pattern click the 'Add pattern' button at the bottom of the panel. This will add a new, empty row with the following fields to fill in:

- *Name*

  The name you want to give your pattern. If documents fall into the pattern, this name will show up under the GDPR filter in the sidebar.

- *Pattern*

  The pattern you want to add. This is a standard regular expression. The actual value that is captured must be placed within parentheses. e.g.: selector(theValueToCapture)

- *Confidence*

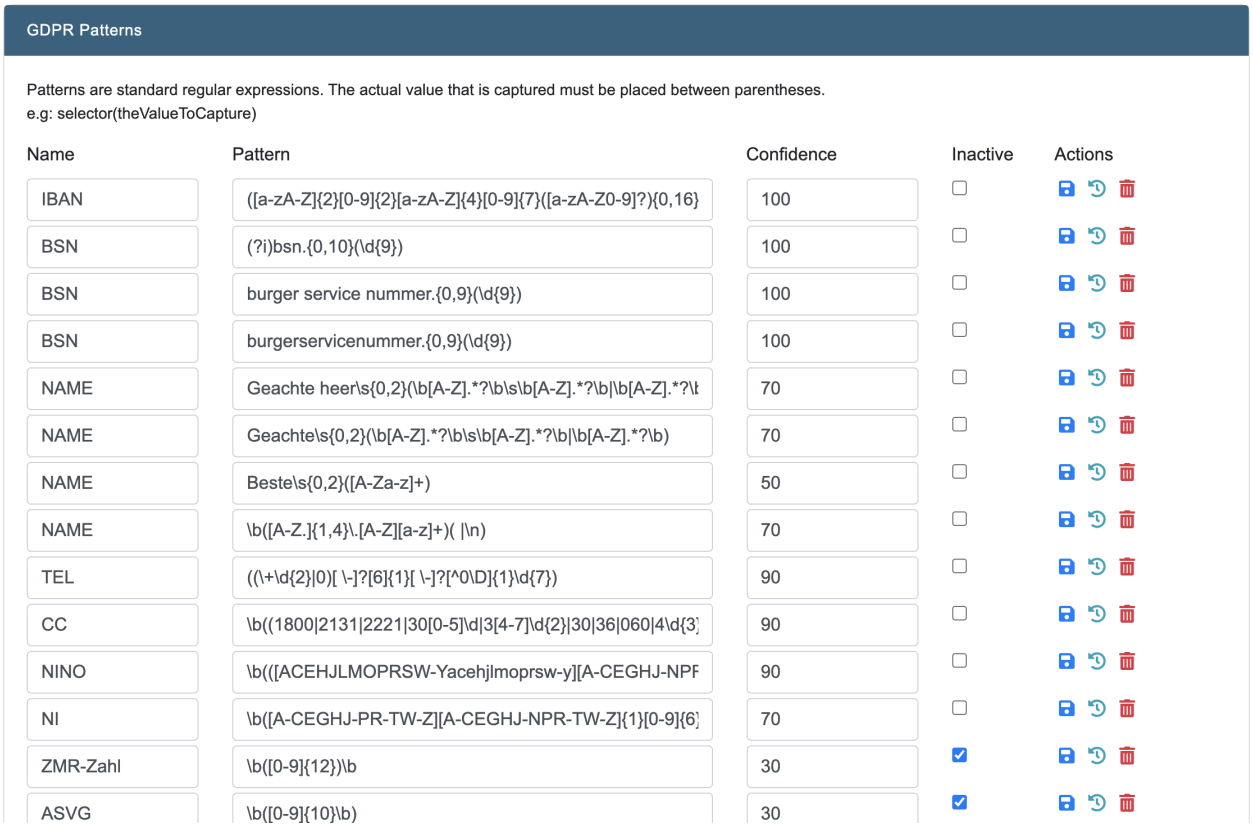
  The confidence score you wish to add to pattern. This is a score between 0 (low) and 100 (high).

- *Inactive*

  Select this if you wish to have the GDPR pattern **not active** within your case.

- *Actions*

  Here you can save, restore and delete your GDPR patterns. Please **do not forget** to save your new pattern(s)



19.3 GDPR Threat Matrix

Within this panel it's possible to add a GDPR threat matrix.

- *Name*

  Add a name for your threat matrix

- *Collection*

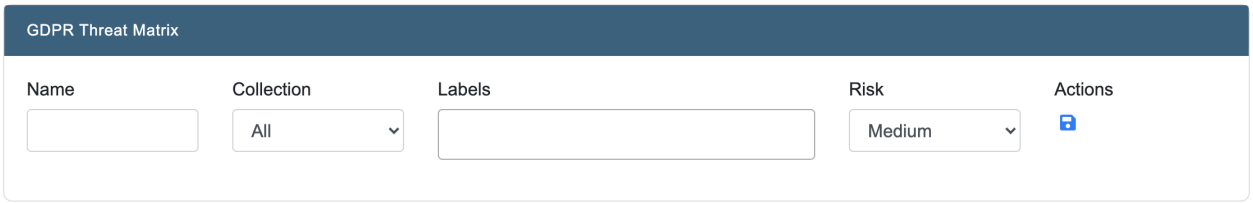
  Select which collection you wish to apply the threat matrix to, or if you want to apply it to all of your collections.

- *Labels*

  Select which labels (GDPR patterns) you wish to add to this matrix.

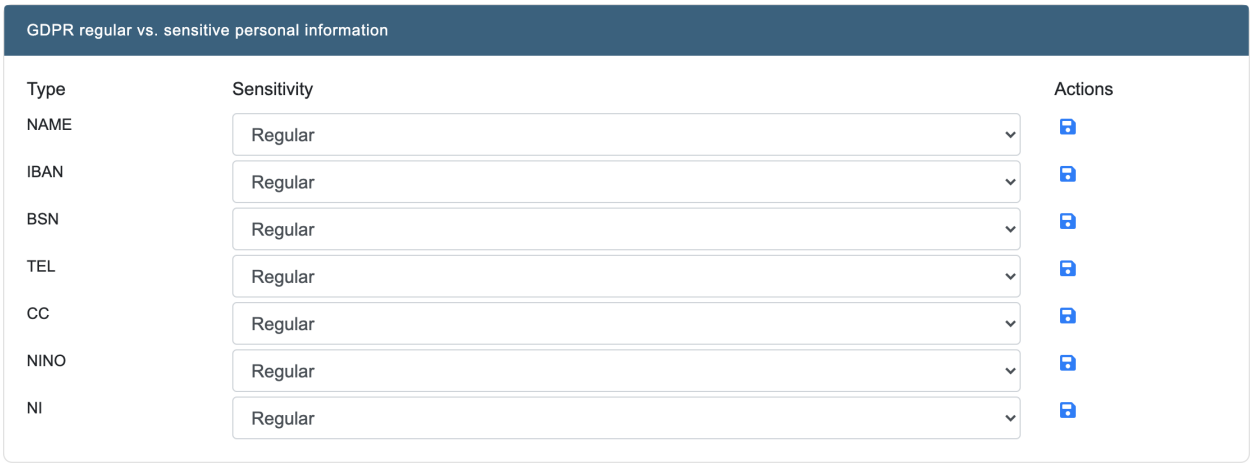
- *Risk*

  Select which risk level applies to your threat matrix: low, medium or high.



19.4 GDPR regular vs. sensitive personal information

Here you can select which types are either regular or sensitive personal information.



19.5 Bag of Words

Within this panel it's possible to add a "Bag of Words" as we call it.

- *Add new*

  Add a name to identify your bag of words. Every word within the bag of words should be surrounded by quotes. eg. "Word1" "Word2" "Word3"

- *Import List*

  The name for the bag of words and actual bag of words should be separated by a comma. eg. Name, "Word1" "Word2" "Word3"

- *Export List*

  It's also possible to export your bag of words list. When you decide to export your bag of words, a .csv file will be downloaded.

CHAPTER

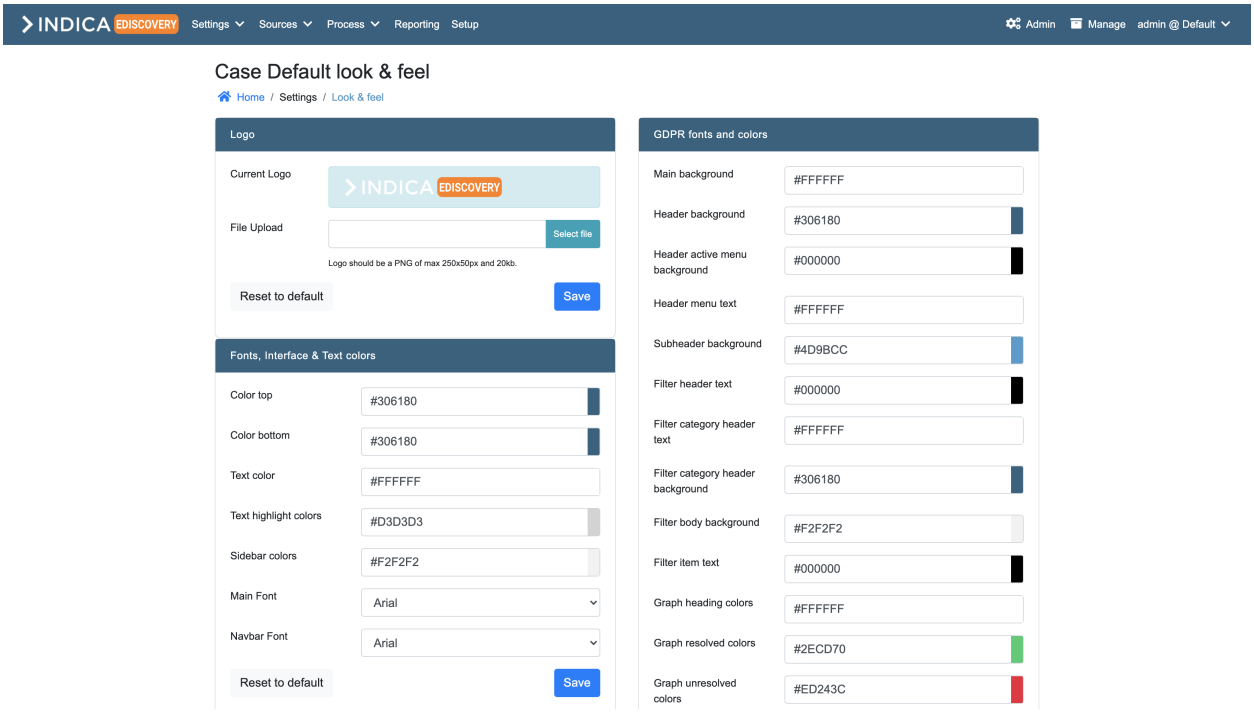
TWENTY

LOOK & FEEL

Table of Contents

- *Look & Feel*
  - *Logo*
  - *Fonts, Interface & Text colors*
  - *GDPR fonts and colors*

Within INDICA it's possible to change the interface to the needs of your company. Within the manage interface, go to "Settings" within the menu and click on "Look & Feel". You will get to a page that looks like the following picture:



20.1 Logo

It is possible to change the default INDICA logo to the logo of your company. This will change the logo in the menu at the top left. Just upload an image which meets te following requirements:

- Image should be a PNG
- Dimensions of max 250px x 50px
- Max 20kb

20.2 Fonts, Interface & Text colors

Within this panel it's possible to change different colors and fonts within the interface. Just hover over one of the colors to see which part of the interface it will change. The red color within the screenshot indicates the part that will change.

See image below for an example:

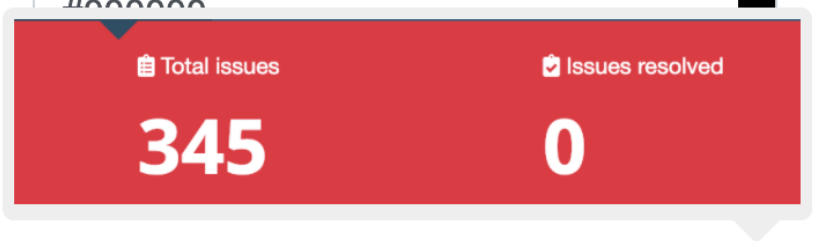
20.3 GDPR fonts and colors

If GDPR is enabled within your case, you can also change the look and feel of the GDPR interface. Just like the panel "Fonts, Interface * Text colors", hover over one of the colors to see which part of the interface it will change. The red color within the screenshot indicates the part that will change.

See image below for an example:

GDPR fonts and colors

Main background	#FFFFFF
Header background	#306180
Header active menu background	
Header menu text	
Subheader background	#4D9BCC
Filter header text	#000000
Filter category header text	#FFFFFF
Filter category header background	#306180
Filter body background	#F2F2F2
Filter item text	#000000
Graph heading colors	#FFFFFF
Graph resolved colors	#2ECD70



CHAPTER

TWENTYONE

DATA SOURCES

Table of Contents

- *Data Sources*
  - *Collections*
    - *Adding a collection*
    - *Editing a collection*
    - *Deleting a collection*
  - *Datatabs*
    - *DataTabs List*
      - *CSV File*
      - *Database Connection*
      - *Unit4 Financial Data Suite*
      - *Alure (Innolan)*
    - *DataTabs Order & Security*
  - *E-mail*
    - *E-mail integration*
    - *Map email to user*
    - *EWS Accounts*
    - *POP3 Settings*
  - *Fileshares*
    - *Adding a fileshare*
    - *Editing a fileshare*
    - *Deleting a fileshare*
  - *Local*
    - *Adding a local source*
    - *Editing a local source*
    - *Deleting a local source*

- *Queries*
  - *Adding Queries*
  - *Editing Queries*
  - *Deleting Queries*
- *S3 Buckets*
  - *Adding S3 Buckets*
  - *Editing S3 Buckets*
  - *Deleting S3 Buckets*
- *SharePoint*
  - *General Settings*
  - *Adding a sharepoint source*
    - *Preparation*
    - *Adding a source*
    - *Online sites*
    - *On-premise sites*
    - *Online personal drives*
  - *Editing a sharepoint source*
  - *Deleting a sharepoint source*
- *Box*
  - *Adding a box source*
    - *Box Preparation*
    - *Connecting Box to INDICA*
  - *Editing a box source*
  - *Deleting a box source*

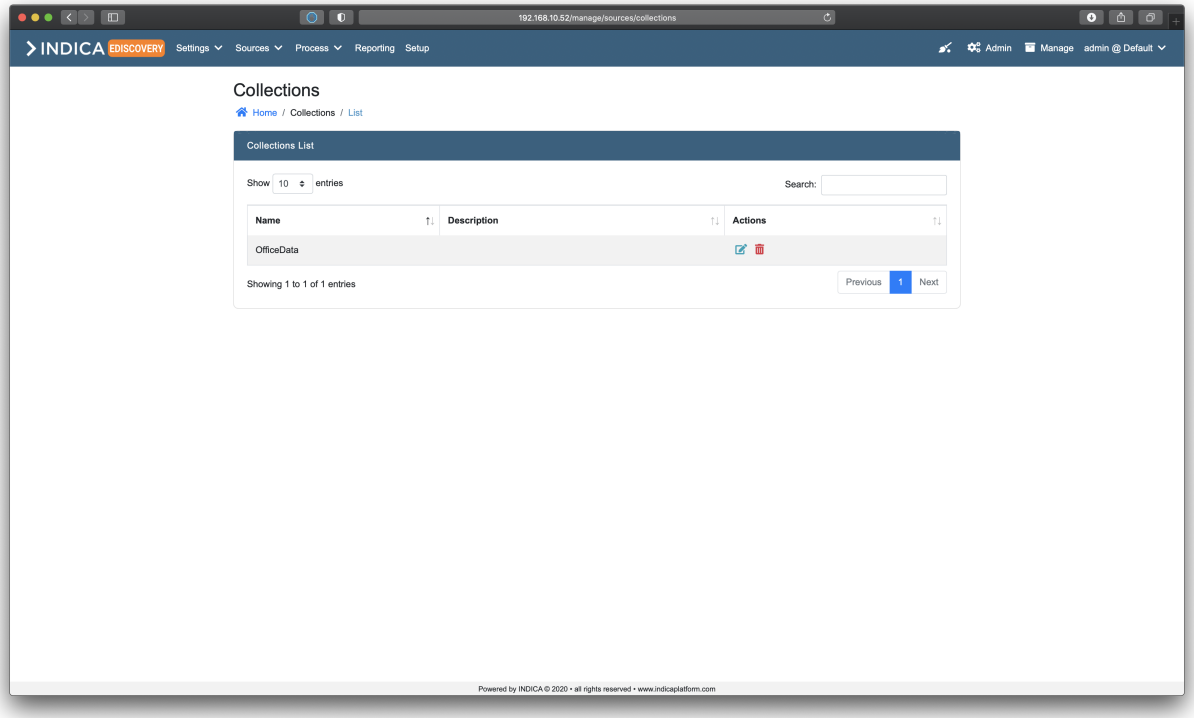
INDICA supports several data sources out of the box. These sources are:

- Collections
- Datatabs
- E-mail
- Fileshares
- Local
- Queries
- S3 Buckets
- SharePoint
- Box

Connecting a source is easy; for some sources no additional rights are required. In some cases it is needed to assign additional rights to a special 'System user' that INDICA uses to read the data. Typically, the needed rights are limited

to read-only rights. By design, INDICA does not alter, move, or delete the source data.

21.1 Collections



A collection allows you to group multiple sources in one collection. This makes it possible to have a collection that contains multiple data sources belonging to a department or person, or a type of data source. It allows you to group the data sources in such a way that it benefits the end user.

There is no limit on the number of collections in a case.

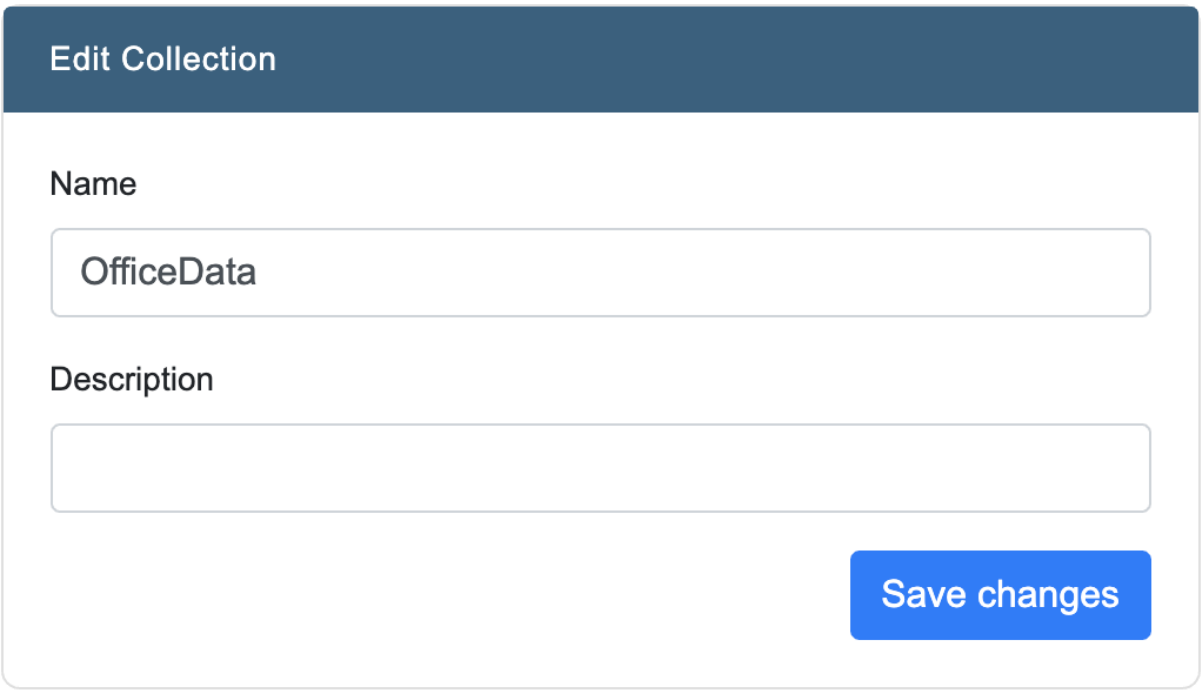
Caution: The name of the collection is stored in the index. Changing the name of a collection or deleting the collection may break your index. Please do not edit or delete collections after data has been indexed.

21.1.1 Adding a collection

When adding a data source, INDICA will ask you to enter a collection name. An existing collection name can be entered to add the new source to the existing collection. If needed, a new collection name can also be picked. INDICA will then create a new collection and add the data source to it.

21.1.2 Editing a collection

After creating a collection, it is possible to edit the collection as well. This can be done by clicking the paper icon (Edit) in the actions column. A new window will open:

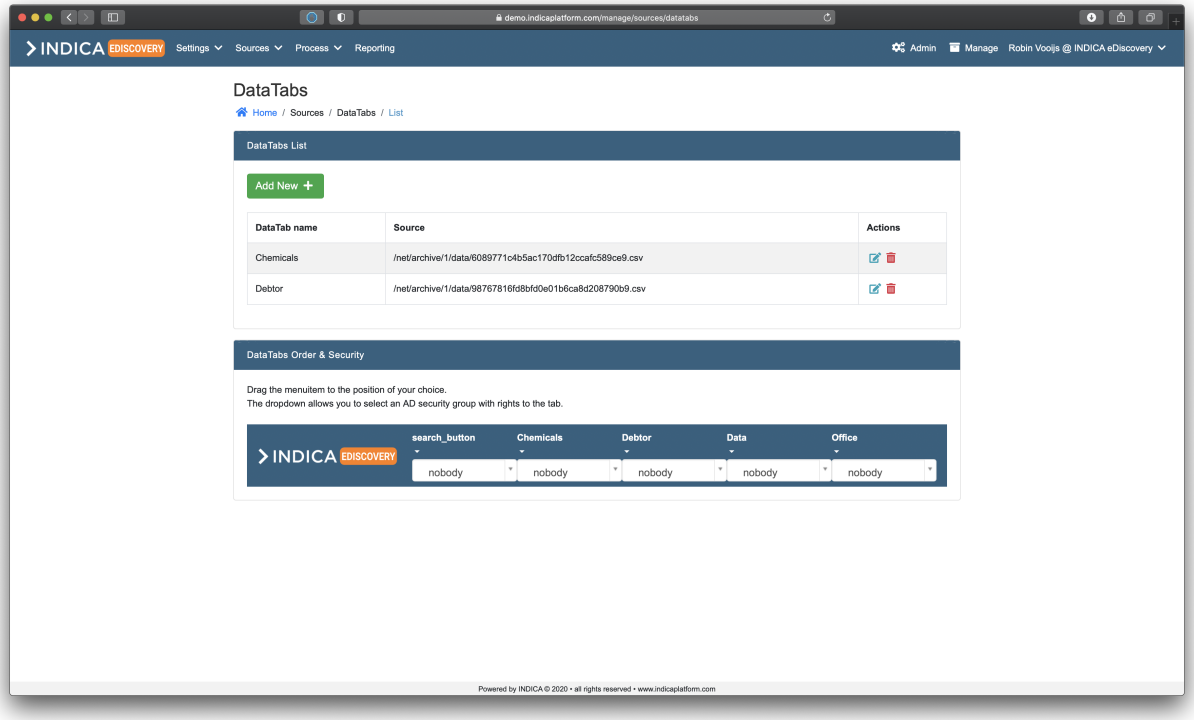


On this page it's possible to edit the name and description of the collection. Click "Save changes" when done.

21.1.3 Deleting a collection

Deleting a collection can be done from the collections overview by clicking the red trash-bin icon (delete). Deleting a collection is irreversible, so use with caution.

21.2 Datatabs

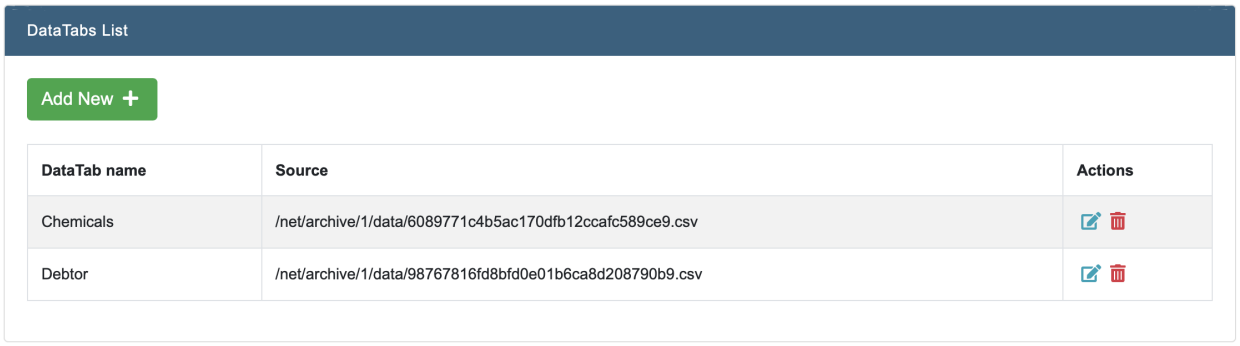


Datatabs are a really powerful tool in INDICA to give insight into the unstructured data. Datatabs make it possible to add a structured source to the index and combine the structured information to the unstructured data. This means that it is possible to load, for example, the database of customers into the index so links can be made from unstructured documents to the corresponding customers. INDICA will create the links between unstructured and structured datasources with the click of a button.

By default, it is possible to add the following sources:

- CSV Files
- Database Connection
- Unit4 Financial Data Suite
- Alure (Innolan)

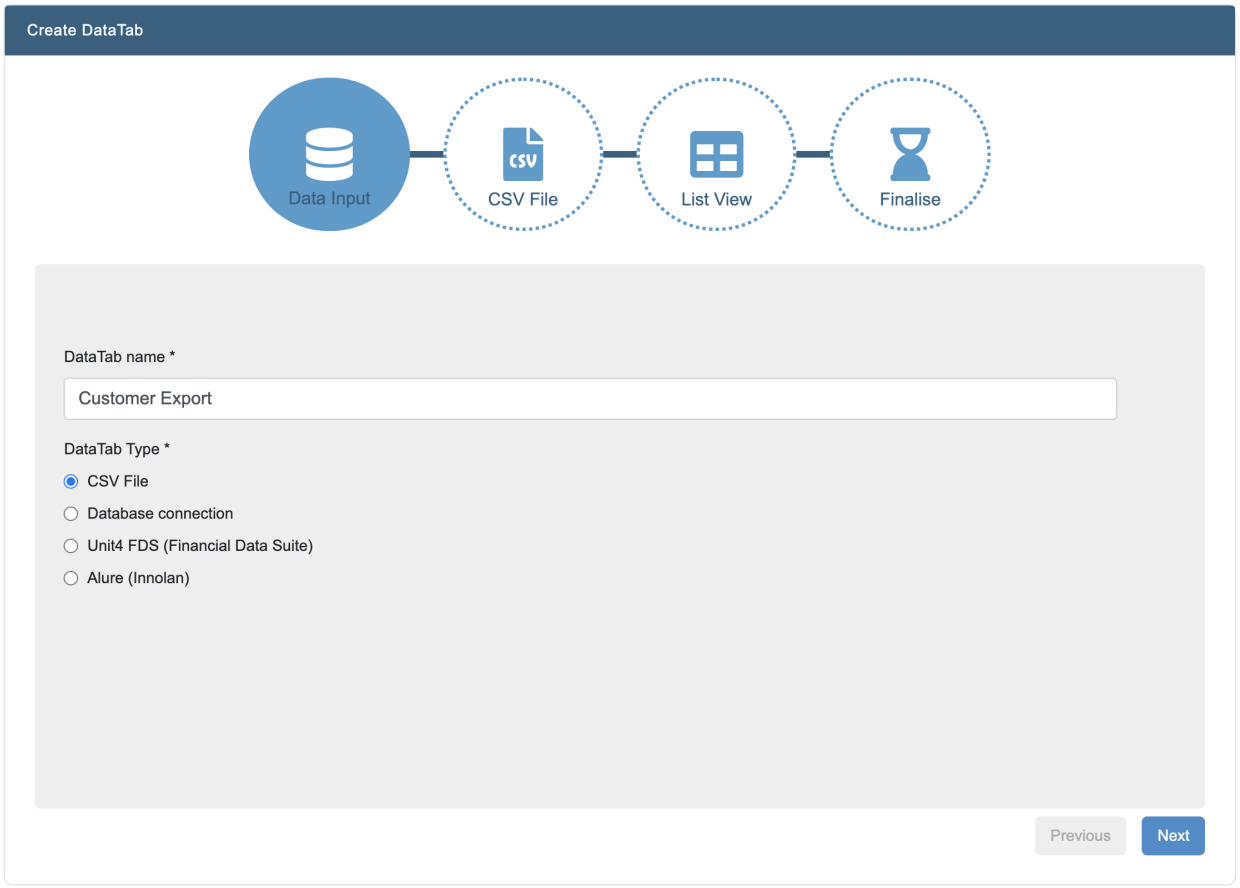
21.2.1 DataTabs List



The datatab list displays the existing datatabs, and allows for creating, updating and deleting datatabs.

CSV File

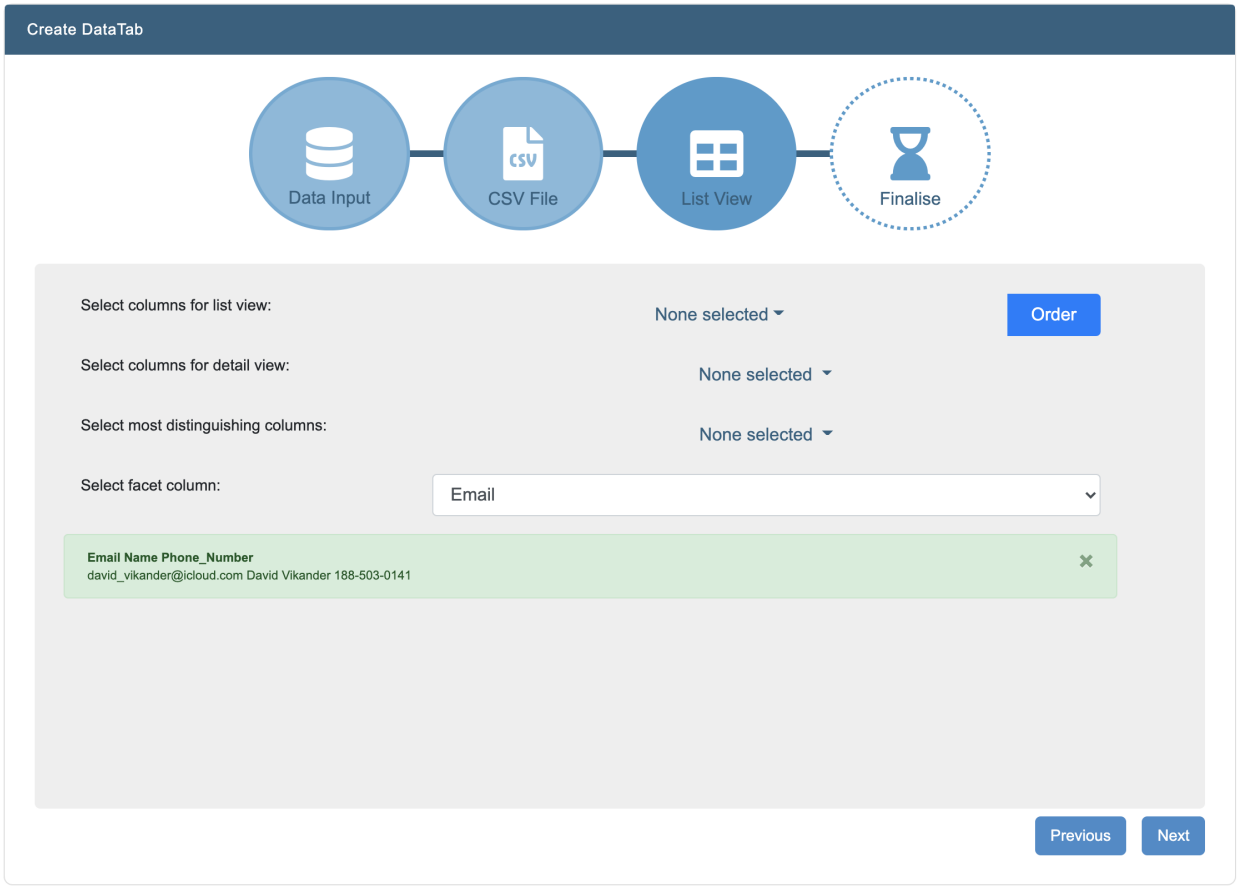
To add a CSV file to the index, click the green "Add New" button in the data tabs list. A new window will open:



Simply enter a name for the datatab, and click the correct option for datatab type. By default, "Database Connection" is selected. Click "CSV File", and then click "Next".

The next step is for uploading a CSV file. Click the "Choose File" button and pick the file you want to upload. After that, click "Next".

You will be presented with a number of options:



- **Select columns for list view**
  These columns will be presented in the previewer when a hit is found when clicking the "Analyze" button;
- **Select columns for detail view**
  These columns will be displayed in the search results list;
- **Select most distinguishing columns**
  Select the columns that are the most distinguishing;
- **Select facet columns**
  Select the column that will be used for creating a facet in the index.

Click "Next" when all settings have been filled in.

The last step provides you with the following options:

- **Select indexing recurrence**
  Choose how often this datatab will be indexed. It is possible to repeat the indexing step multiple times when needed. This allows to automatically update the index contents when the uploaded file is changed on disk.
- **Collection name**
  Choose an existing name to add this datasource to an existing collection. Choose a new name if a new collection should be created.

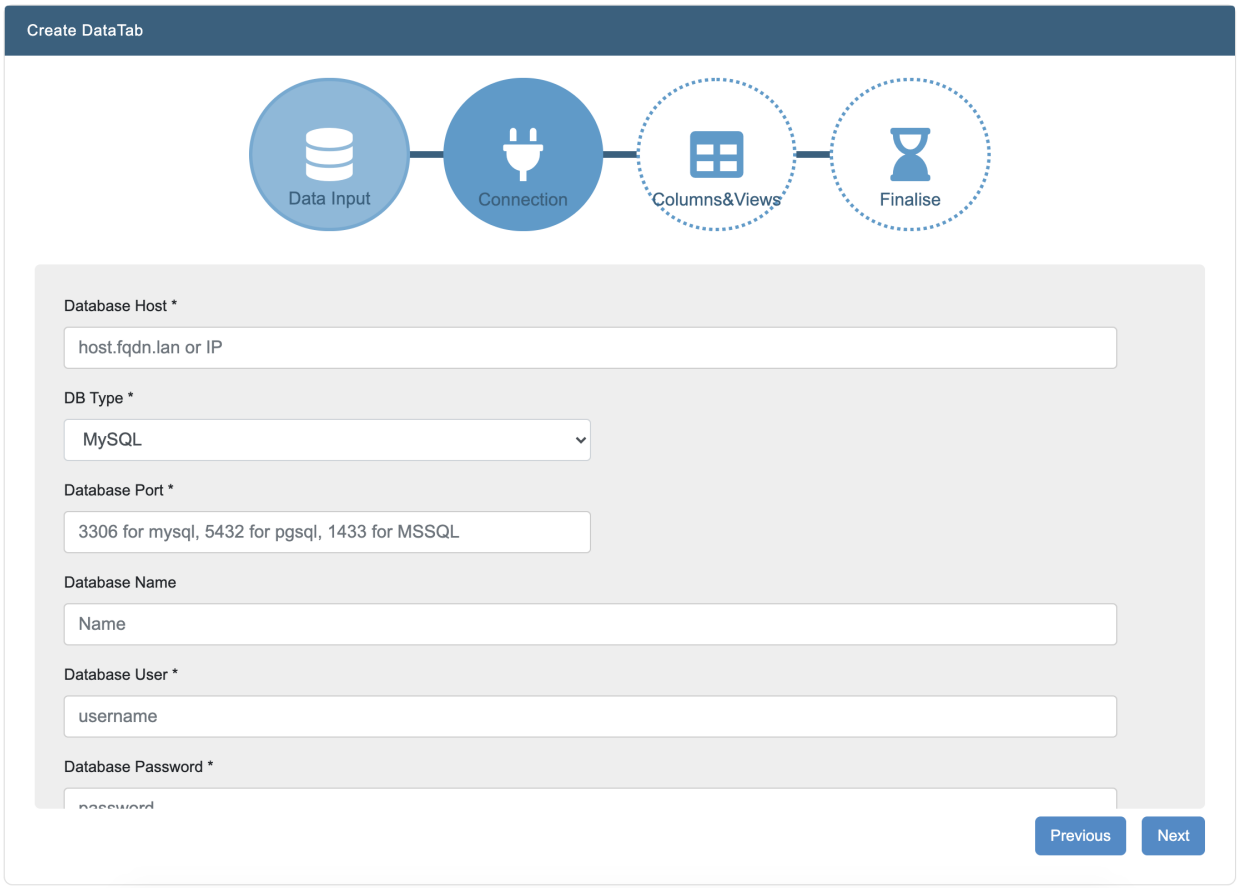
Click "Finish" when done. The CSV file will be added to the index if the spider is enabled.

Database Connection

INDICA can connect directly with most common database types. This makes it possible to connect directly to a lot of services and appliances, as most appliances and applications rely on a database in the background.

To connect to a database, click the "Add New" button in the data tabs list.

A new window will open. Choose a name for the datatab, and select "Database Connection" in the list of options. When done, click "Next".



A number of questions will be asked in step 2. These are:

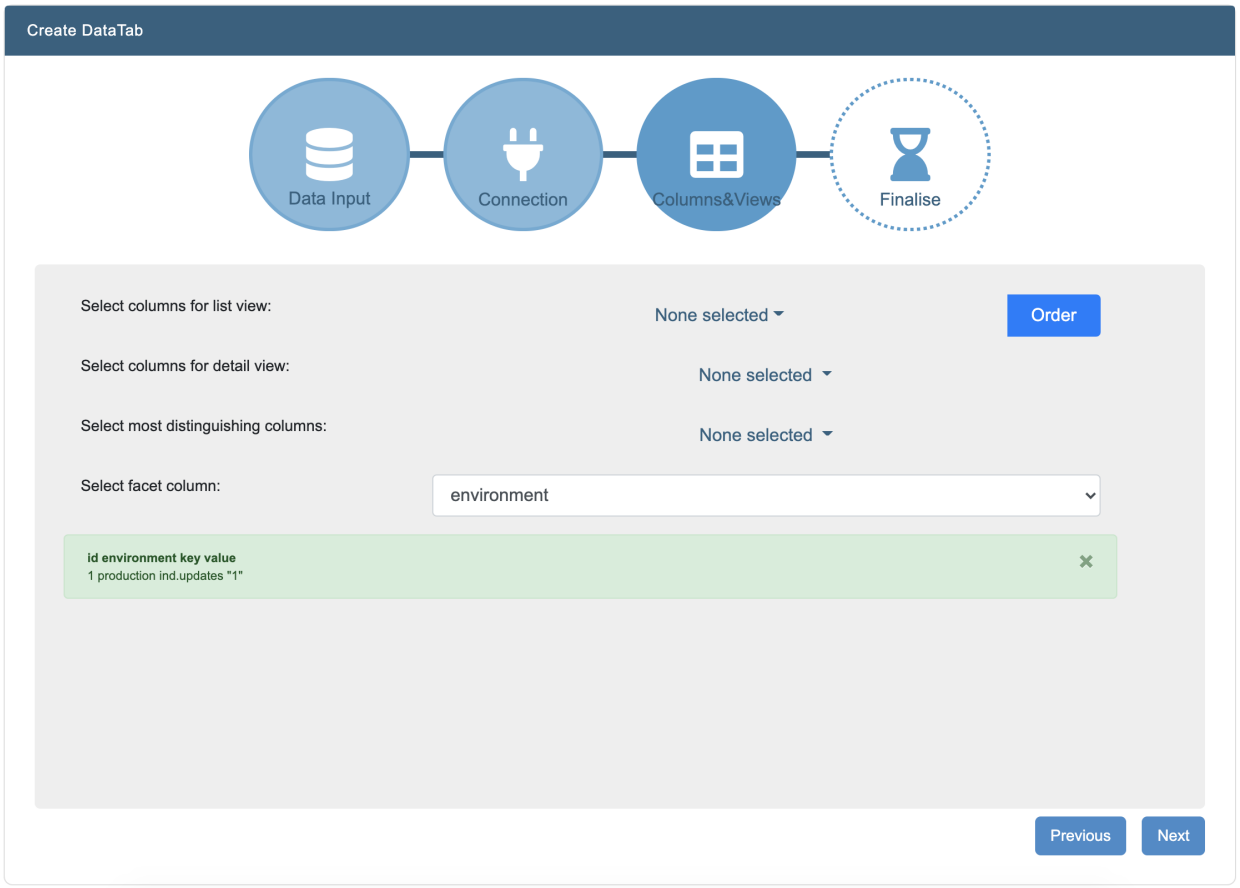
- **Database Host**
  Typically the IP address of the database host;
- **DB Type**
  Choose between MySQL, Postgres, or MS SQL;
- **Database Port**
  The port on which the database is hosted. Default ports will be filled in when selecting database type, and can be changed;
- **Database Name**
  Name of the database;
- **Database User**
  User that can be used to authenticate;
- **Database Password**
  Password of the user in question.

To set up a connection, the user needs to be able to read the database and execute queries on the database.

When everything is filled in, click the "Test Connection" button. If everything is correctly set up and filled in, the test should be succesful.

Next, enter the query that will be executed on the database to get the necessary information. When filled in, click the "Save SQL" button. A message will be displayed at the top of the page containing the selected columns and the first row of values. If this is correct, go to the next step with the "Next" button.

You will be presented with a number of options:



- **Select columns for list view**
  These columns will be presented in the previewer when a hit is found when clicking the "Analyze" button;
- **Select columns for detail view**
  These columns will be displayed in the search results list;
- **Select most distinguishing columns**
  Select the columns that are the most distinguishing;
- **Select facet columns**
  Select the column that will be used for creating a facet in the index.

Click "Next" when all settings have been filled in.

The last step provides you with the following options:

- **Select indexing recurrence**
  Choose how often this datatab will be indexed. It is possible to repeat the indexing step multiple times when needed. This allows to automatically update the index contents when the database is changed.

- **Collection name**
  Choose an existing name to add this datasource to an existing collection. Choose a new name if a new collection should be created.

Click "Finish" when done. The retrieved information will be added to the index if the spider is enabled.

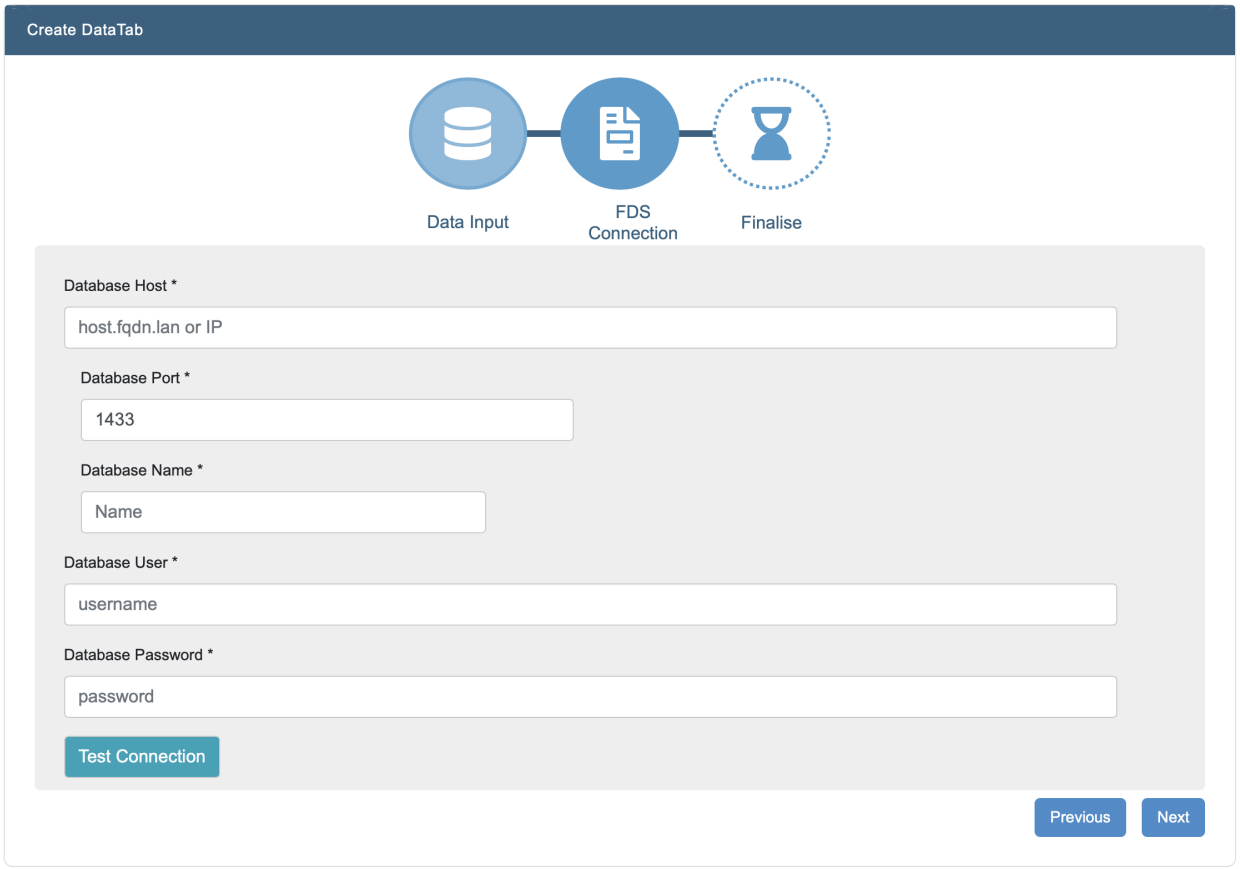
Unit4 Financial Data Suite

It is also possible to connect INDICA directly to Unt4 Financial Data Suite.

To connect to Unit4 FDS, click the "Add New" button in the data tabs list.

A new window will open.

Choose a name for the datatab, and select "Unit4 FDS (Financial Data Suite)" in the list of options. When done, click "Next".



A number of questions will be asked in step 2. These are:

- **Database Host**
  Typically the IP address of the database host;
- **Database Port**
  The port on which the database is hosted. Default port will be filled in, and can be changed;
- **Database Name**
  Name of the database;
- **Database User**
  User that can be used to authenticate;

- **Database Password**
  Password of the user in question.

When everything is filled in, click the "Test Connection" button. If everything is correctly set up and filled in, the test should be succesful.

Click "Next" when all settings have been filled in.

The last step provides you with the following options:

- **Select indexing recurrence**
  Choose how often this datatab will be indexed. It is possible to repeat the indexing step multiple times when needed. This allows to automatically update the index contents when the database is changed.
- **Collection name**
  Choose an existing name to add this datasource to an existing collection. Choose a new name if a new collection should be created.

Click "Finish" when done. The retrieved information will be added to the index if the spider is enabled.

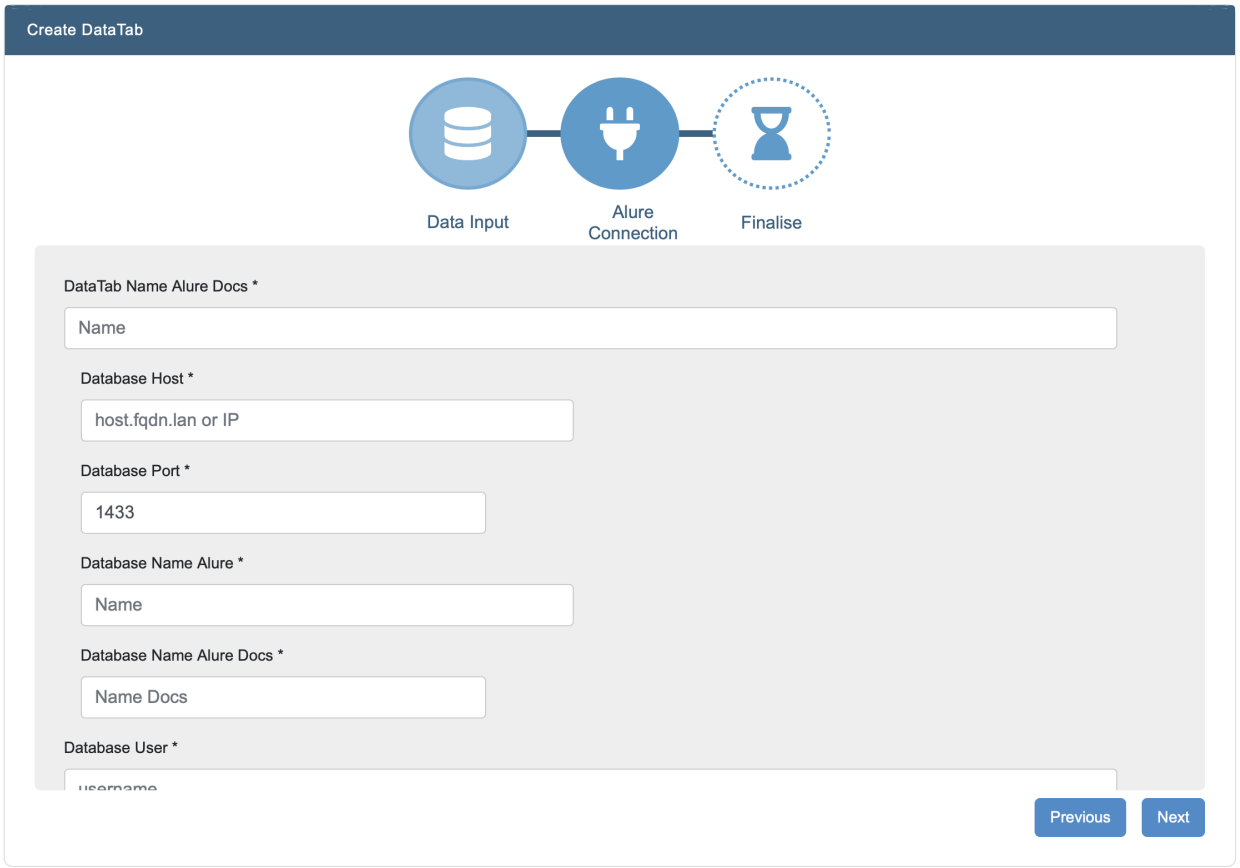
Alure (Innolan)

INDICA can connect with Alure instances as well.

To connect to Alure, click the "Add New" button in the data tabs list.

A new window will open.

Choose a name for the datatab, and select "Alure (Innolan)" in the list of options. When done, click "Next".



A number of questions will be asked in step 2. These are:

- DataTab Name Alure Docs
- Database Host
- Database Port
- Database Name Alure
- Database Name Alure Docs
- Database User
- Database Password

When everything is filled in, click the "Test Connection" button. If everything is correctly set up and filled in, the test should be succesful.

Click "Next" when all settings have been filled in.

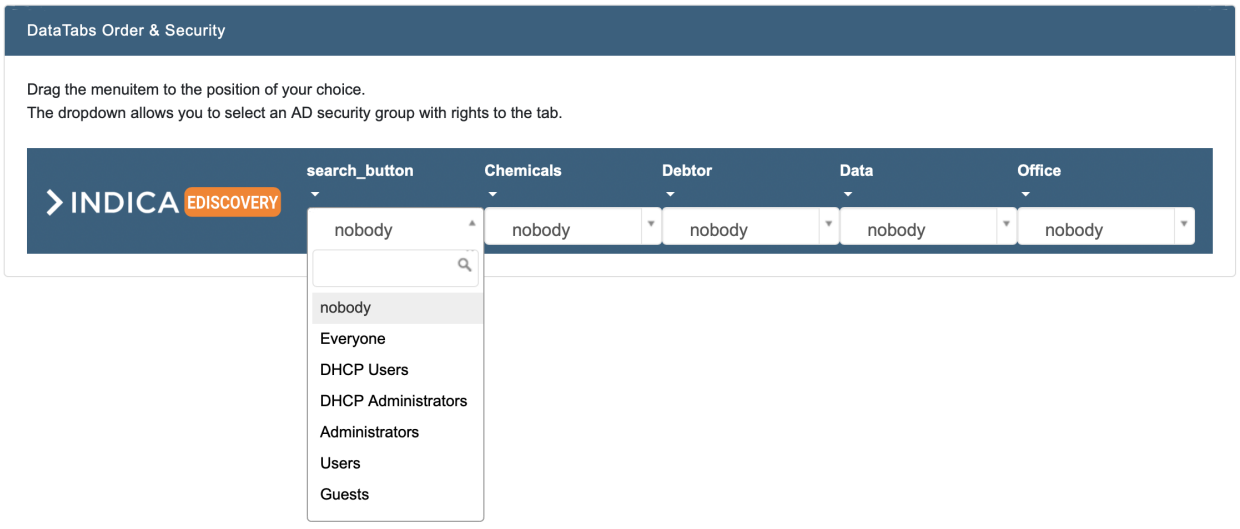
The last step provides you with the following options:

- **Select indexing recurrence**
  Choose how often this datatab will be indexed. It is possible to repeat the indexing step multiple times when needed. This allows to automatically update the index contents when the database is changed.
- **Collection name**
  Choose an existing name to add this datasource to an existing collection. Choose a new name if a new collection should be created.

Click "Finish" when done. The retrieved information will be added to the index if the spider is enabled.

21.2.2 DataTabs Order & Security

Case managers can change the order and access rights of the datatabs.

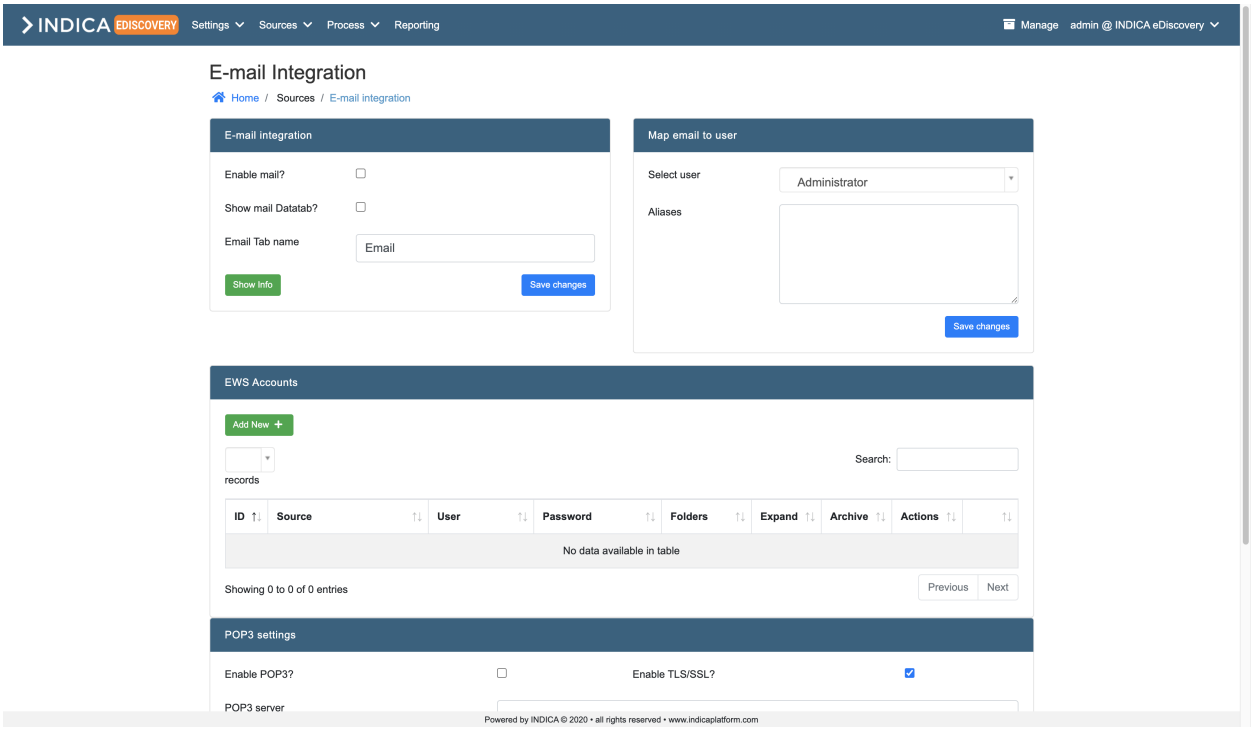


Changing the order of datatabs in the navigation bar is as easy as clicking and dragging a datatab.

To set a security filter on a datatab, open the dropdown menu and select the AD group that should have access to the datatab.

21.3 E-mail

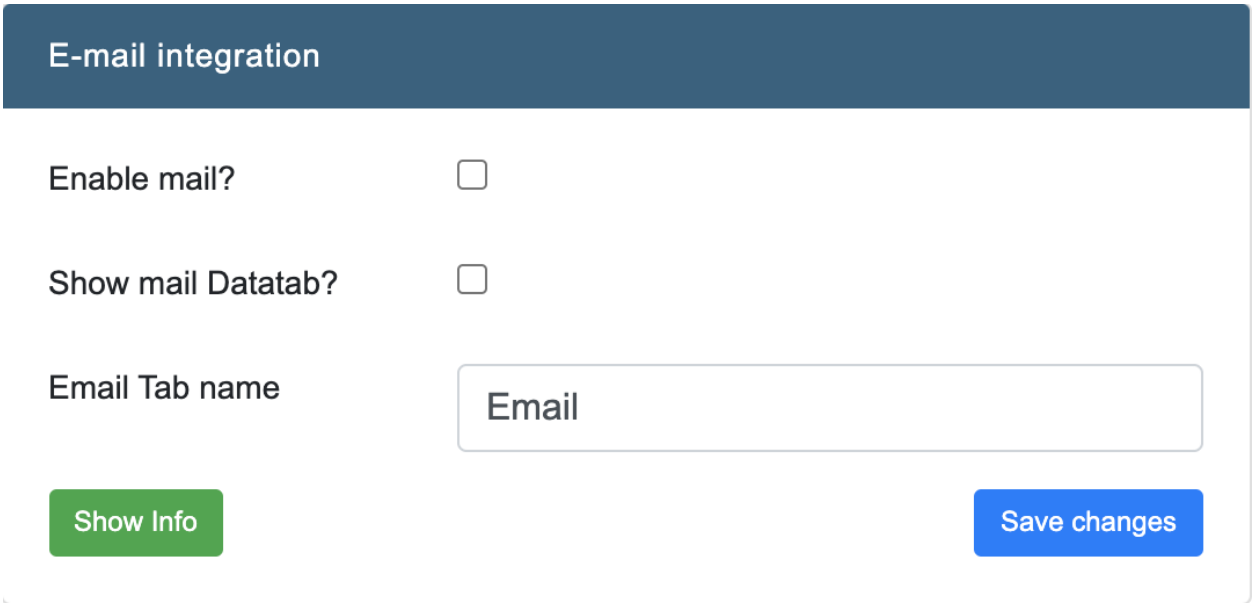
Email can be connected in different ways, of which the most common way is a direct connection to the Webserver through the Exchange Web Services API. Other possibilities are direct reading of PST or OST files, or a connection with POP3 or IMAP boxes. To connect email sources, within the manage interface click on *Sources* and navigate to *E-mail*. Within this overview you'll find a few different panels.



21.3.1 E-mail integration

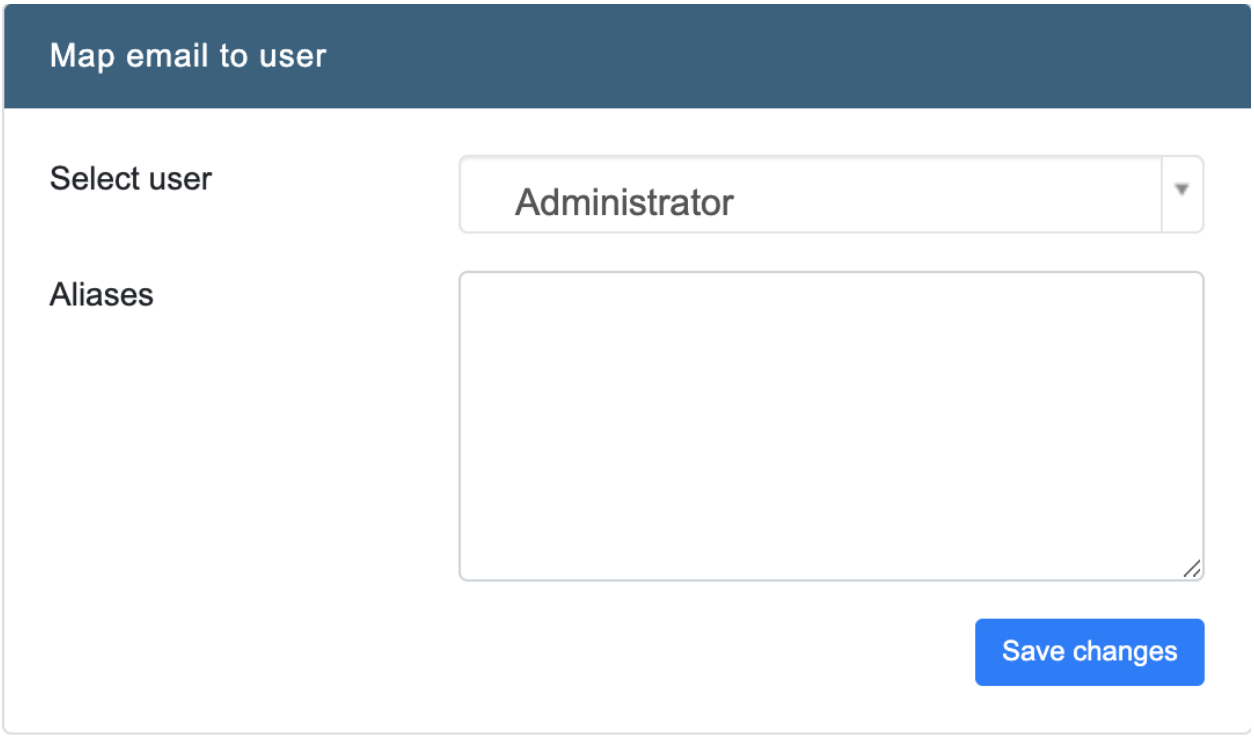
After enabling the e-mail functionality, the e-mail connectivity can be chosen, either a POP archiving/catch all mailbox or EWS connections (see corresponding chapters). In order to store the email archive on an external share, please set the share information in the Archiving section. Within this panel you have a few options:

- **Enable mail?**
  Whether you want to enable mail or not
- **Show mail datatab?**
  Whether you want the mail to show up in a datatab
- **Email tab name**
  The name for your datatab, if you've enabled the datatab option
- **Mail domain name**
  If mail is enabled and there is no (Active) Directory integration enabled, a custom email domain can be set here too. Default it's set to indica.lan.
- **Show info**
  Gives a report overview. This only works when mail is enabled and everything is indexed.



21.3.2 Map email to user

This panel gets the users from the connected AD. If you do not have an AD connected, this will not show any users. Select the user which you wish to add email aliases to.

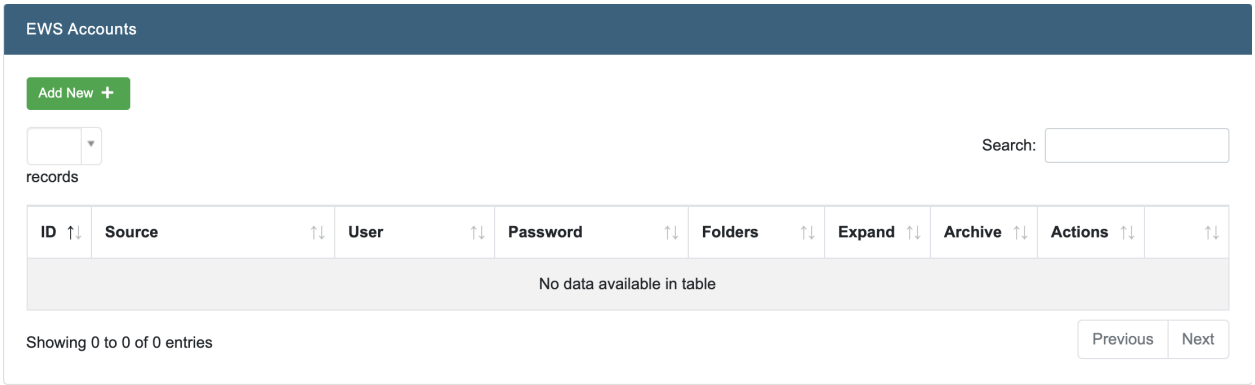


21.3.3 EWS Accounts

Most modern mailservers have EWS API connectivity. Choose expand if you want all shared mailboxes too. A special user can be created to spider all the emails.

Note: Keep in mind that INDICA needs to create an Email archive, so enough storage should be added to the system. To add a new EWS account click on the green *Add new* button within the EWS Accounts panel. A new row within the table will appear, which needs to be filled in:

- **Source**
  The location of the EWS source
- **User**
  The user that's able to connect to the EWS source
- **Password**
  The password corresponding with the user
- **Folders**
  Which folders you wish to index (if unspecified, it will index all folders)
- **Expand**
  Whether you wish to expand (or not) upon the mail boxes being read (behind the @) than the one that has been read.
- **Archive**
  Whether you wish to enable or disable archiving. Archive is an option for mail servers that work according to the exchange guidelines to read mail on a separate archive server.

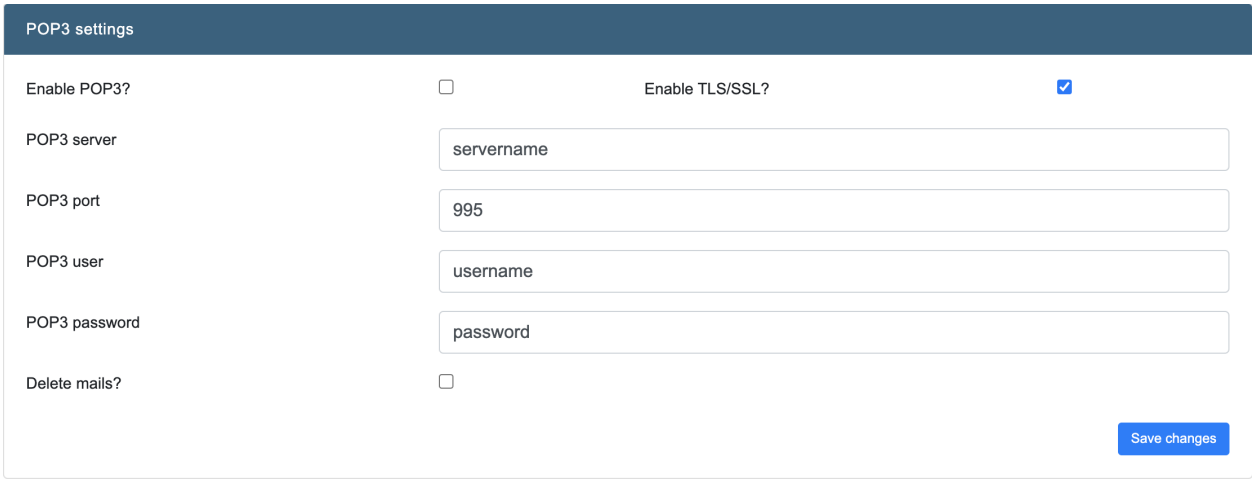


21.3.4 POP3 Settings

The system will connect using the Post Office Protocol 3 (POP3), and store the email on the default archive location. You can choose to access it through a separate tab. All emails from the archive POP mailbox will be downloaded and removed from the mailbox (if chosen). If needed, enter extra email aliases to the user in the list. There are a few options you'll have to fill in before you can connect using the POP protocol:

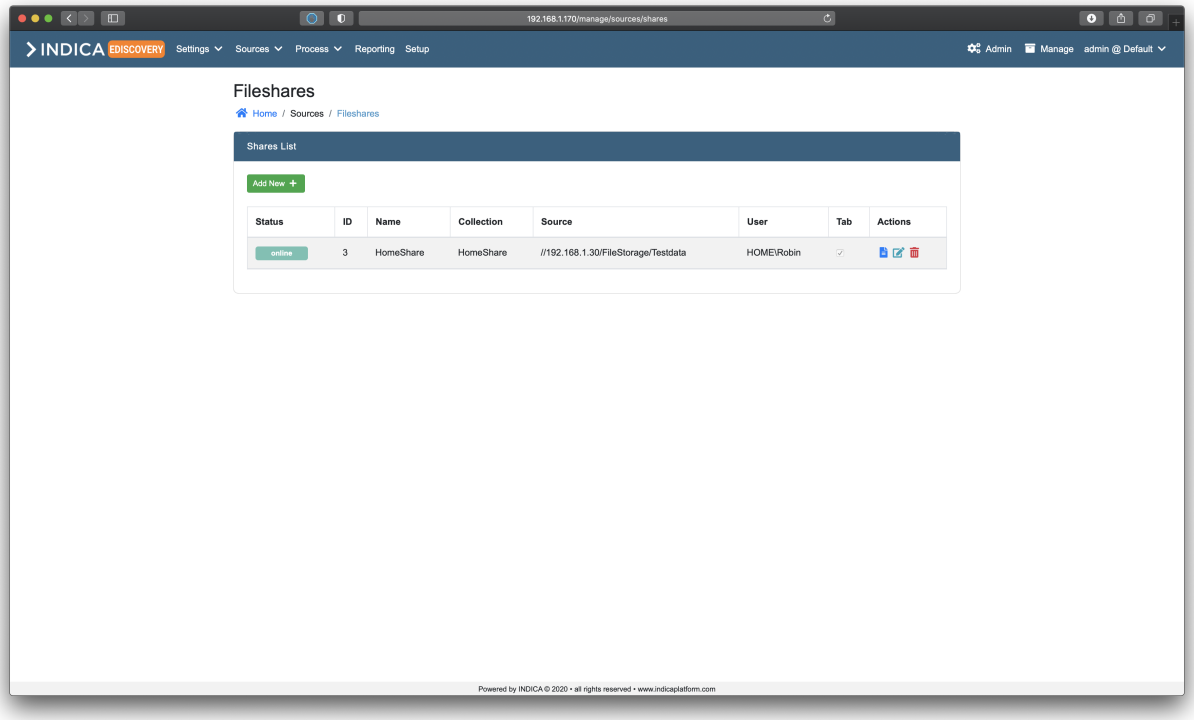
- **Enable POP3?**
  Whether to enable or disable the POP3 settings
- **Enable TLS/SSL?**
  Whether to enable or disable TLS/SSL.

- **POP3 Server**
  The server name of the POP3.
- **POP3 Port**
  The port of the POP3.
- **POP3 User**
  The POP3 username.
- **POP3 Password**
  The POP3 password corresponding with the user in the previous option.
- **Delete mails?**
  Whether you wish to delete the emails or not.



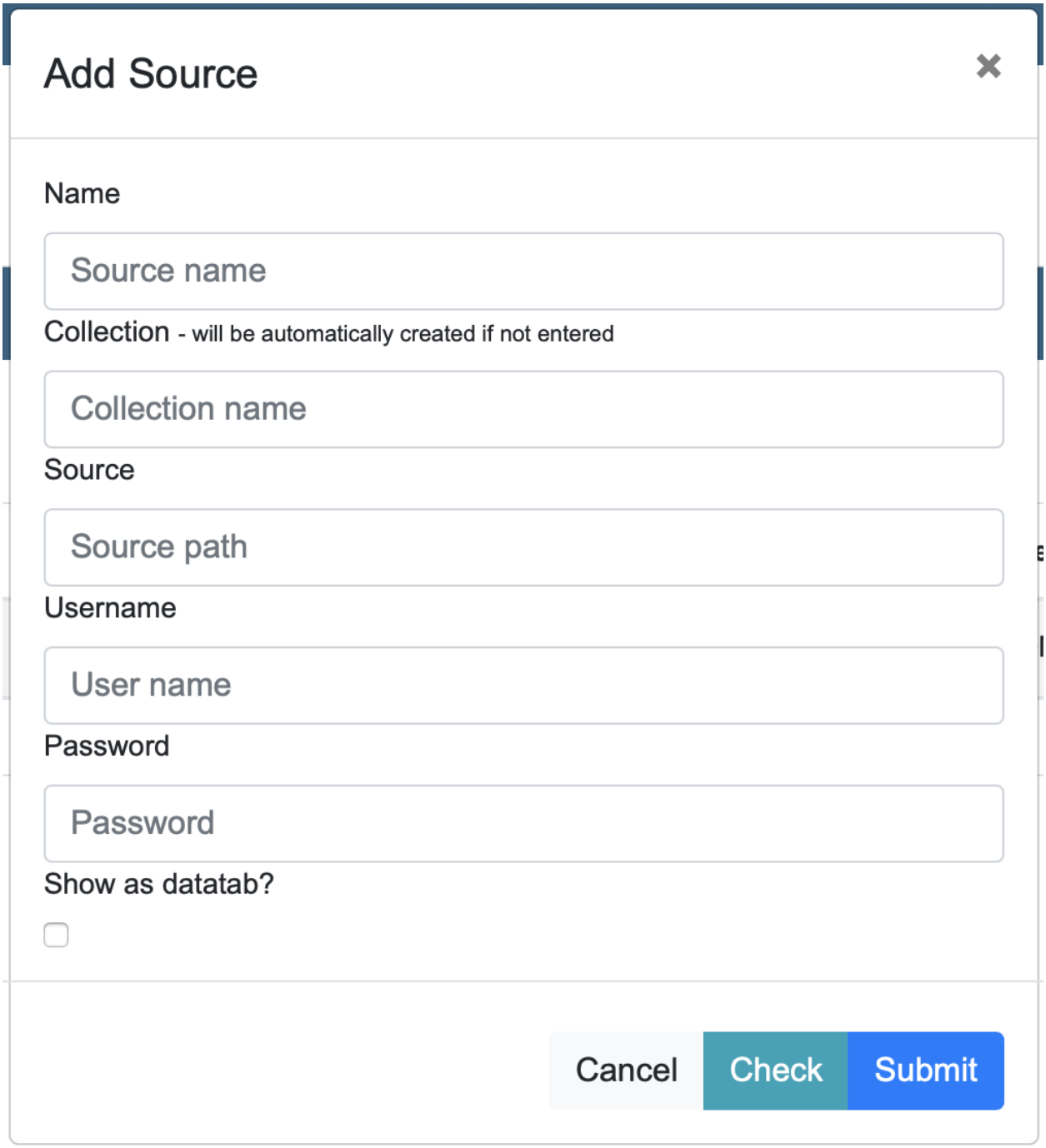
21.4 Fileshares

INDICA can connect to several types of fileshares. It can connect to any fileshare that can be mounted using CIFS. INDICA also supports mounting DFS setups, however, in most cases a direct share/server path is required to properly read Access Control Lists.



21.4.1 Adding a fileshare

To add a fileshare to INDICA, click the green “Add New” button in the fileshares list. A new window will open:



This window asks for the following information:

- **Name**
  Name of the fileshare;
- **Collection**
  Name of the collection to add the fileshare to. Pick an existing collection to add this share to it, or enter a new collection to create a new one;
- **Source**
  Full path to the share;

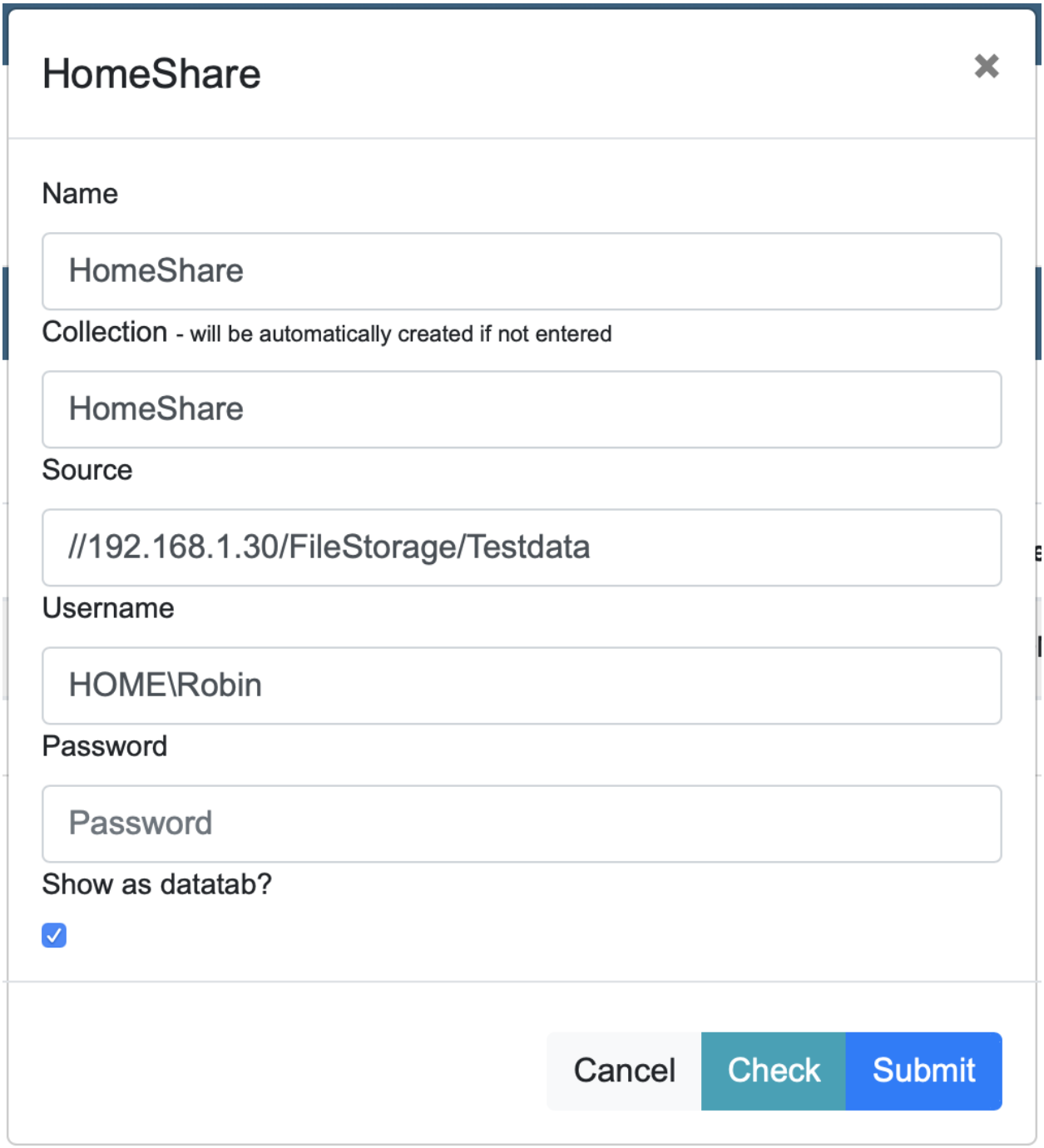
- **Username**
  Username of a user that INDICA can use to connect to the share. This is usually a special service account;
- **Password**
  Password of the user;
- **Show as datatab?**
  Whether this share will be displayed as a datatab in the navigation bar.

When all information is filled out, click the "Check" button to test if INDICA can find the fileshare and is able to authenticate. If everyhting is working, it returns "Succes". Then click the "Submit" button to add the fileshare and begin the indexing process.

Note: INDICA will, by design, mount the added share read-only. This means that INDICA will not be able to modify the source data in any way.

21.4.2 Editing a fileshare

To edit a fileshare, click the paper icon (Edit) in the actions column. A window will open:



This window is identical to the "Add fileshare" window, with the exception that all information is already filled in. Here you can edit the information needed, fill in the password of the user again, and save the changes by clicking "Submit".

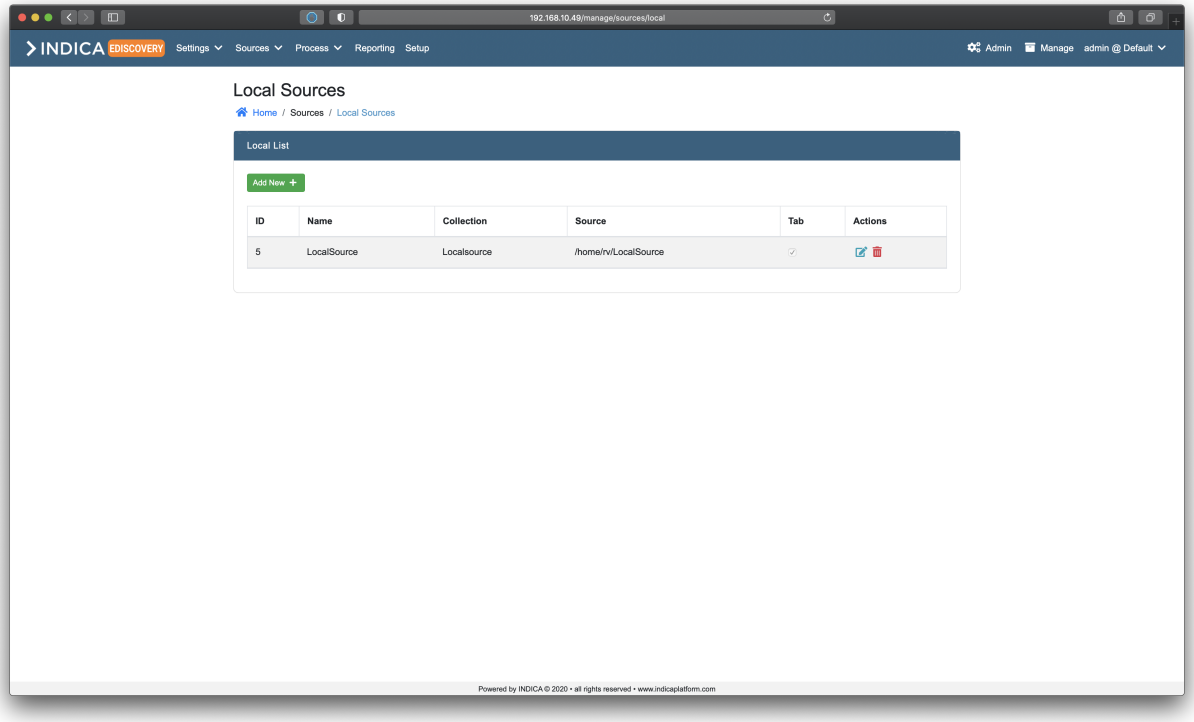
21.4.3 Deleting a fileshare

Removing a fileshare can be done by clicking the red trash-bin icon (delete) in the "Actions" column of the table.

> **Warning:** Removing a fileshare will remove all corresponding documents from the index. This is irreversible, and can only be undone by re-indexing the fileshare.

21.5 Local

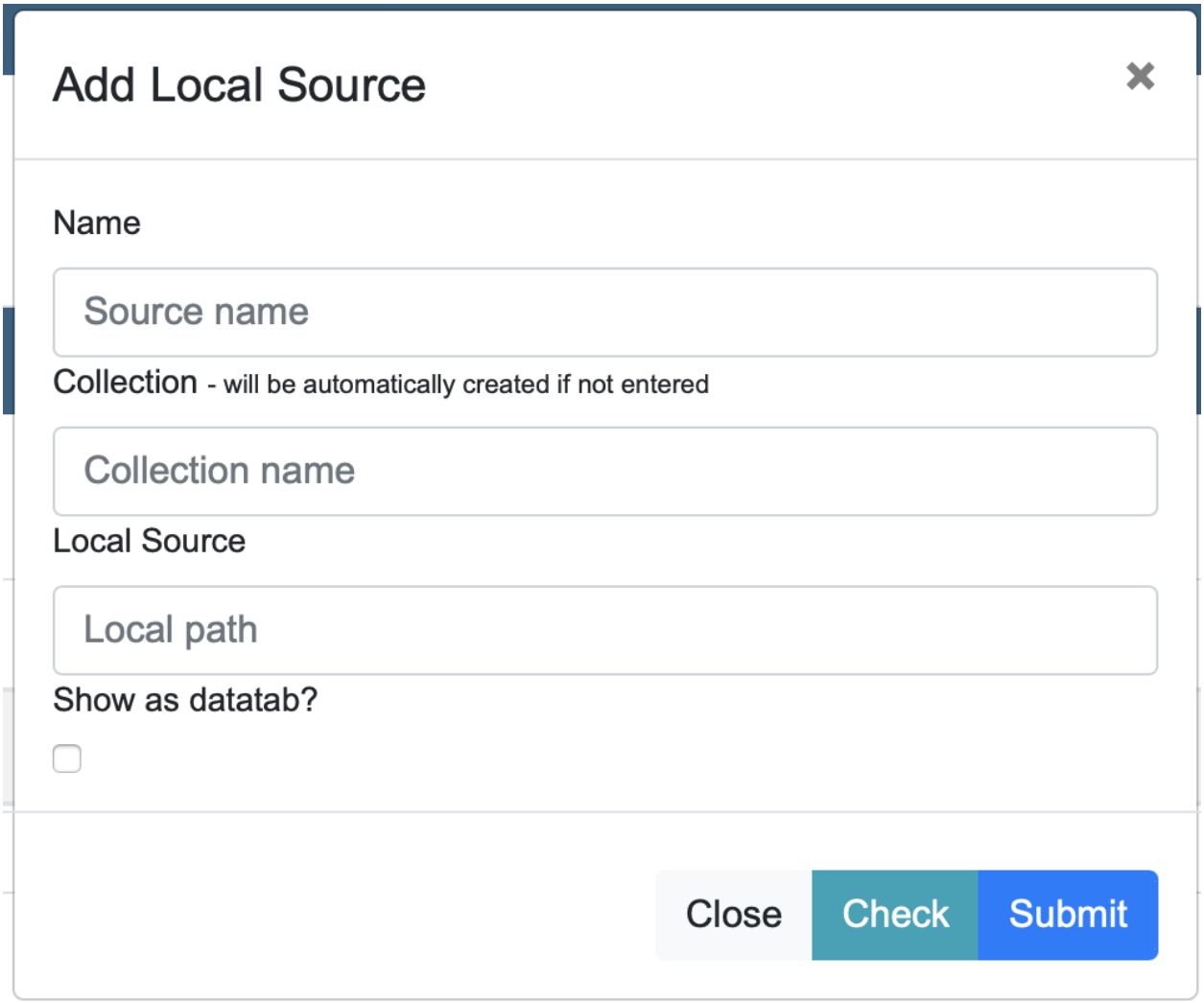
It is also possible to index a local source. This means that documents stored on the server can be added to the index as well. This allows for upload of documents to the server, and indexing the folder where they are stored.



21.5.1 Adding a local source

Adding a local source to INDICA is easy. The only piece of information to be known is the full path of the folder on the appliance.

To add a local source, click the green "Add New" button. A window will open:



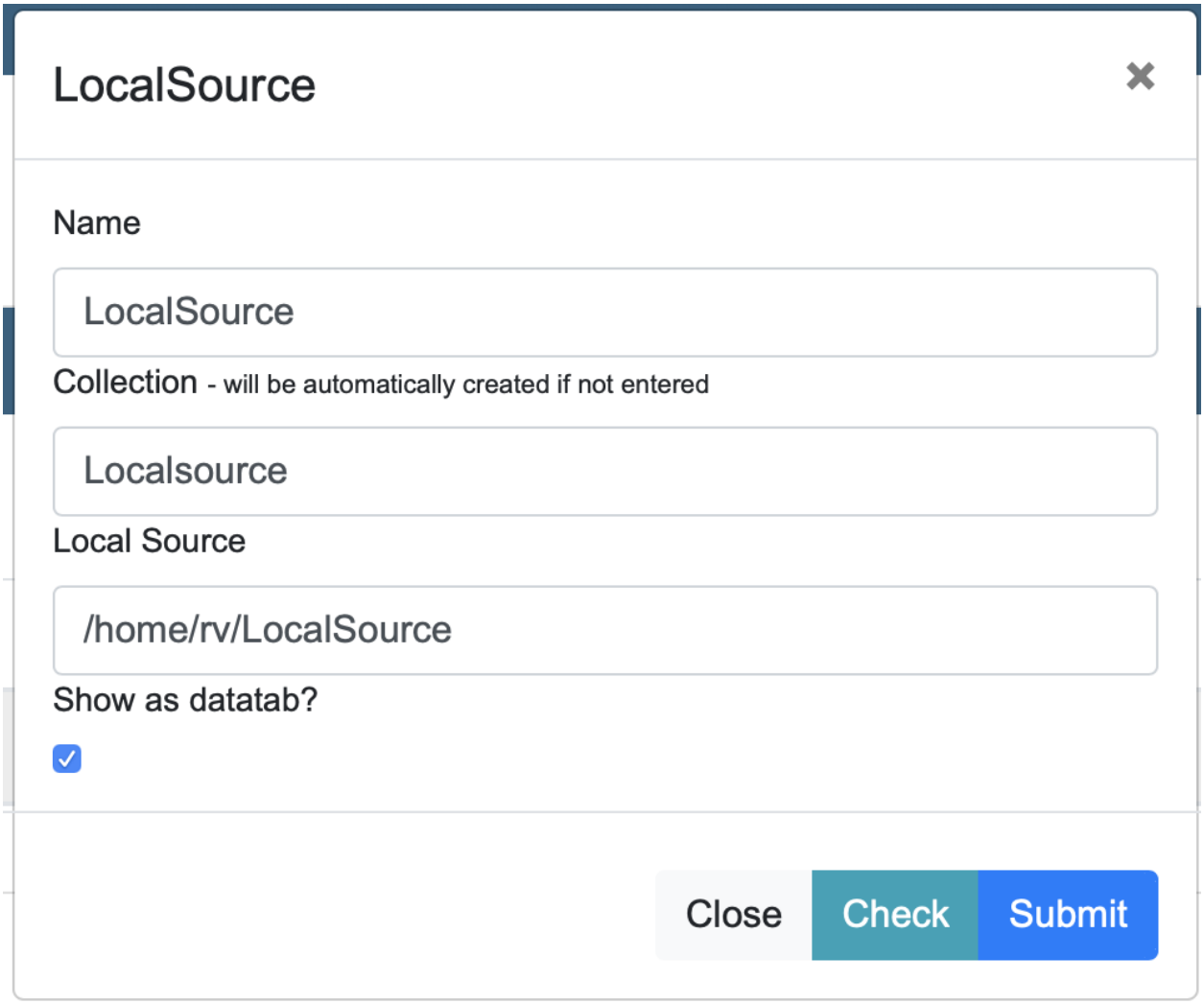
Fill in the following information:

- **Name**
  Pick a name for the local source;
- **Collection**
  Name of the collection to add the local source to. Pick an existing collection to add this source to it, or enter a new collection to create a new one;
- **Local Source**
  Full path of the folder on disk;
- **Show as datatab?**
  Whether this share will be displayed as a datatab in the navigation bar.

When all information is filled in, click the "Check" button. INDICA will check if the directory exists and if the proper permissions are in place. It should return "Success" if everyting is working. Click the "Submit" button to add the local source to the appliance. The indexing of the source will begin shortly.

21.5.2 Editing a local source

To edit a local source, click the blue paper icon (Edit) in the actions column. A window will open:



This window is identical to the "Add Local Source" window, with the exception that all information is already filled in. Here you can edit the information needed, and save the changes by clicking "Submit".

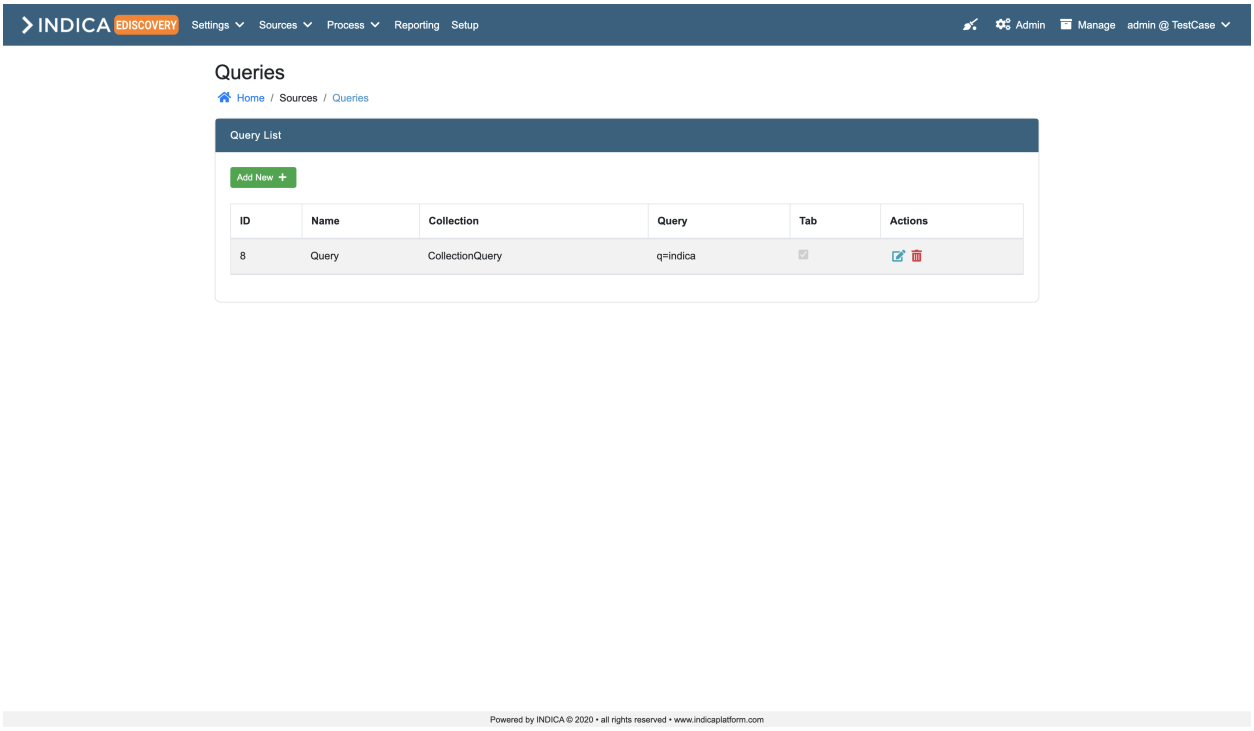
21.5.3 Deleting a local source

Removing a local source can be done by clicking the red trash-bin icon (delete) in the "Actions" column of the table.

Warning: Removing a local source will remove all corresponding documents from the index. This is irreversible, and can only be undone by re-indexing the local source.

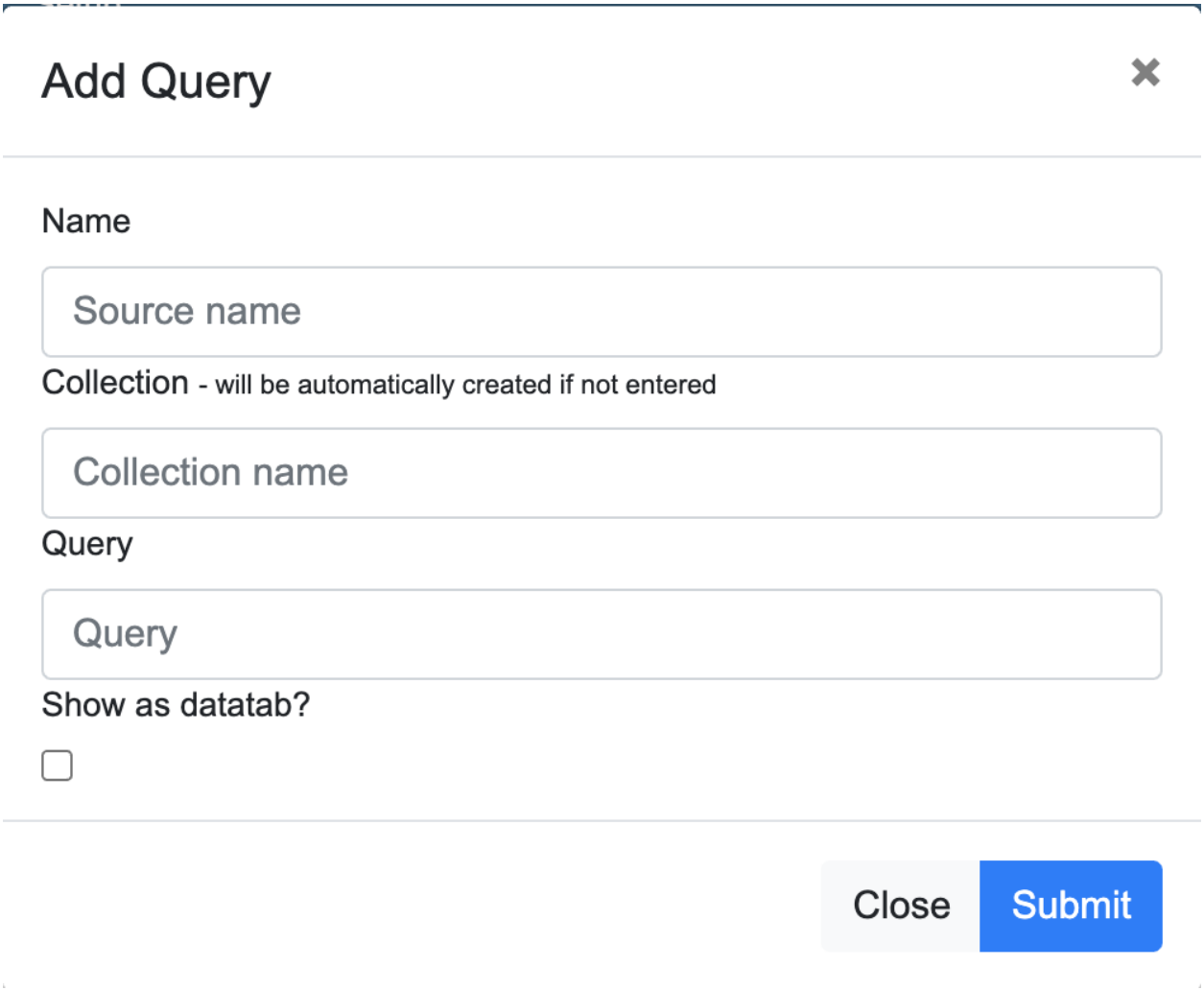
21.6 Queries

Within INDICA it's possible to add queries to your existing collections. This makes it easier to immediately filter on specific data within that collection. To add a query to a collection, within the manage interface, click on *Sources* in the menu and select *Queries*. This brings you to the following page:



21.6.1 Adding Queries

To add a new query, click on the *Add new* button. This will open a window:

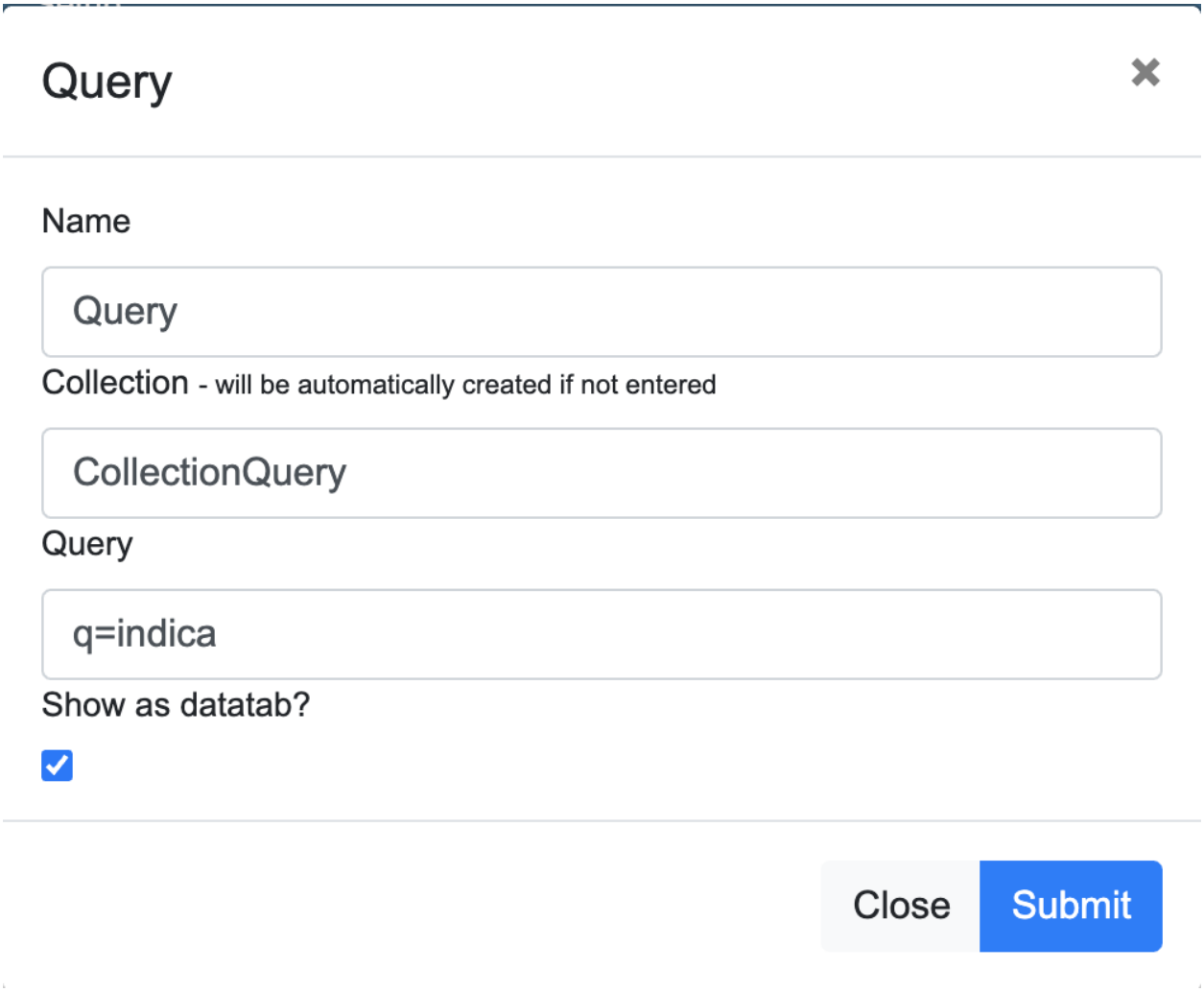


Here you can fill in the following information:

- **Name**
  Pick a name to identify your query with
- **Collection**
  Name of the collection to add the query to. Pick an existing collection to add this source to it, or enter a new collection to create a new one;
- **Query**
  Which query you wish to add
- **Show as datatab?**
  Whether this query will be displayed as a datatab in the navigation bar in the search interface.

21.6.2 Editing Queries

To edit a query, click the blue pen with paper icon (Edit) in the actions column. This will open a window:



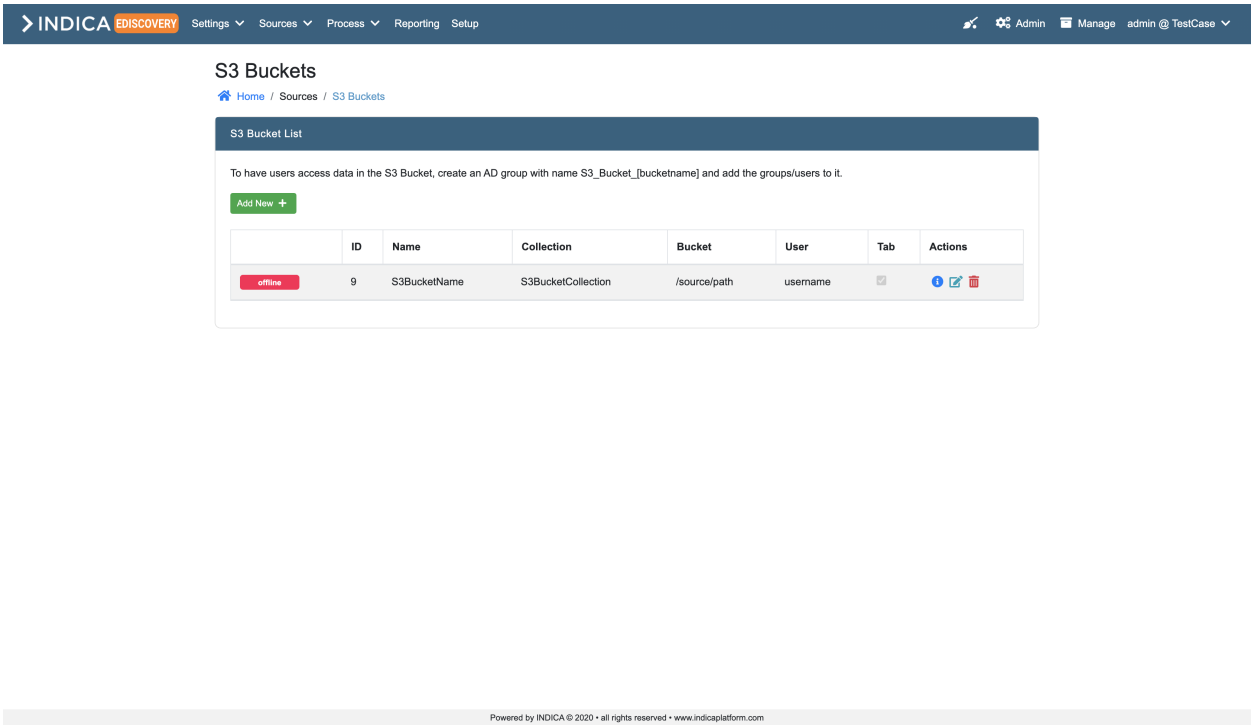
This window is identical to the "Add Queries" window, with the exception that all information is already filled in. Here you can edit the information needed, and save the changes by clicking "Submit".

21.6.3 Deleting Queries

Removing a query can be done by clicking the red trash-bin icon (delete) in the “Actions” column of the table.

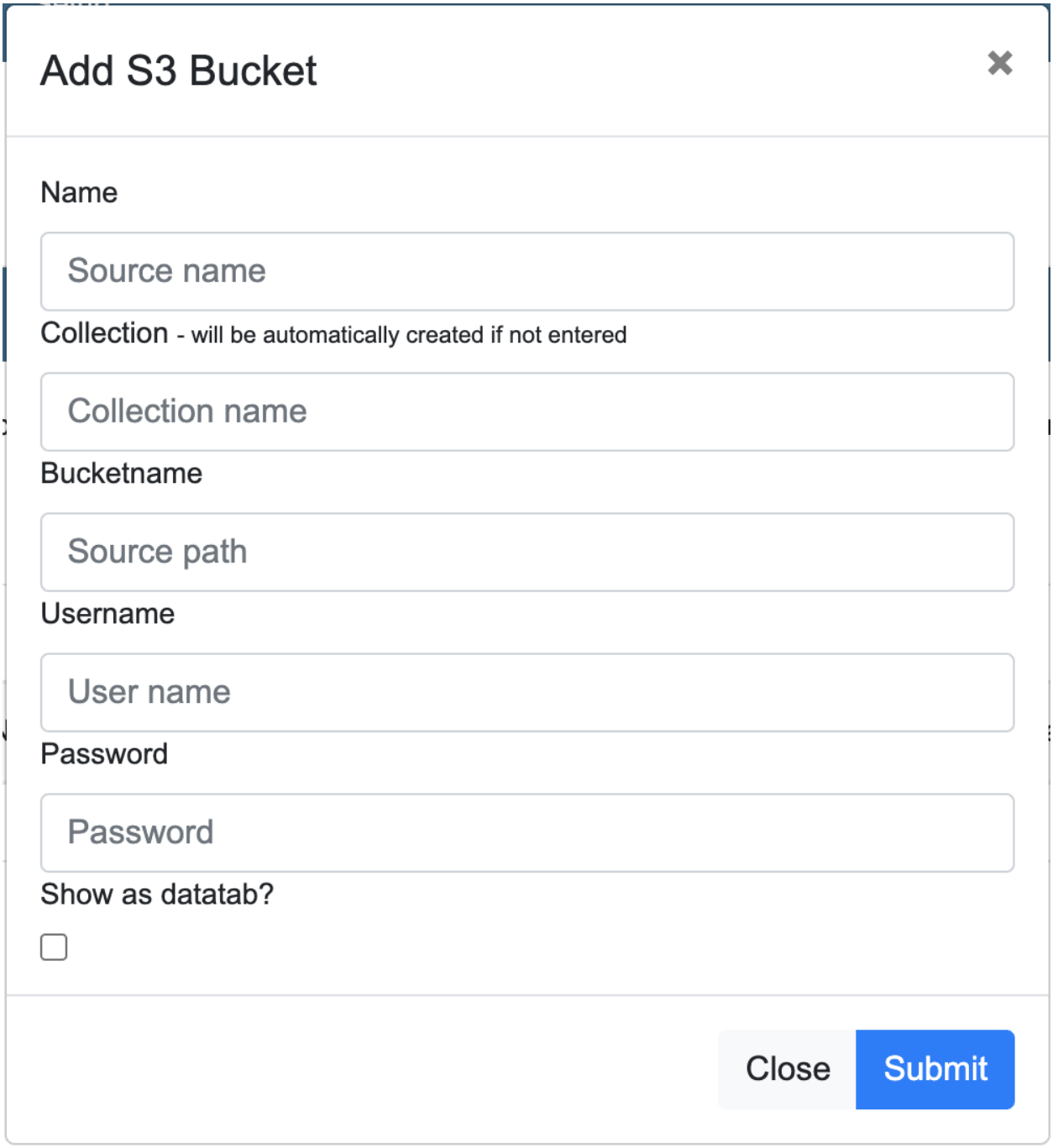
21.7 S3 Buckets

With INDICA it’s also possible to connect S3 Buckets as sources. To have users access data in the S3 Bucket, create an AD group with name S3_Bucket_[bucketname] and add the groups/users to it. You can add S3 buckets by going to the manage interface, clicking on *Sources* in the menu and navigating to *S3 Buckets*.



21.7.1 Adding S3 Buckets

To add a new S3 bucket, click on the *Add new* button. This will open a window:



Here you can fill in the following information:

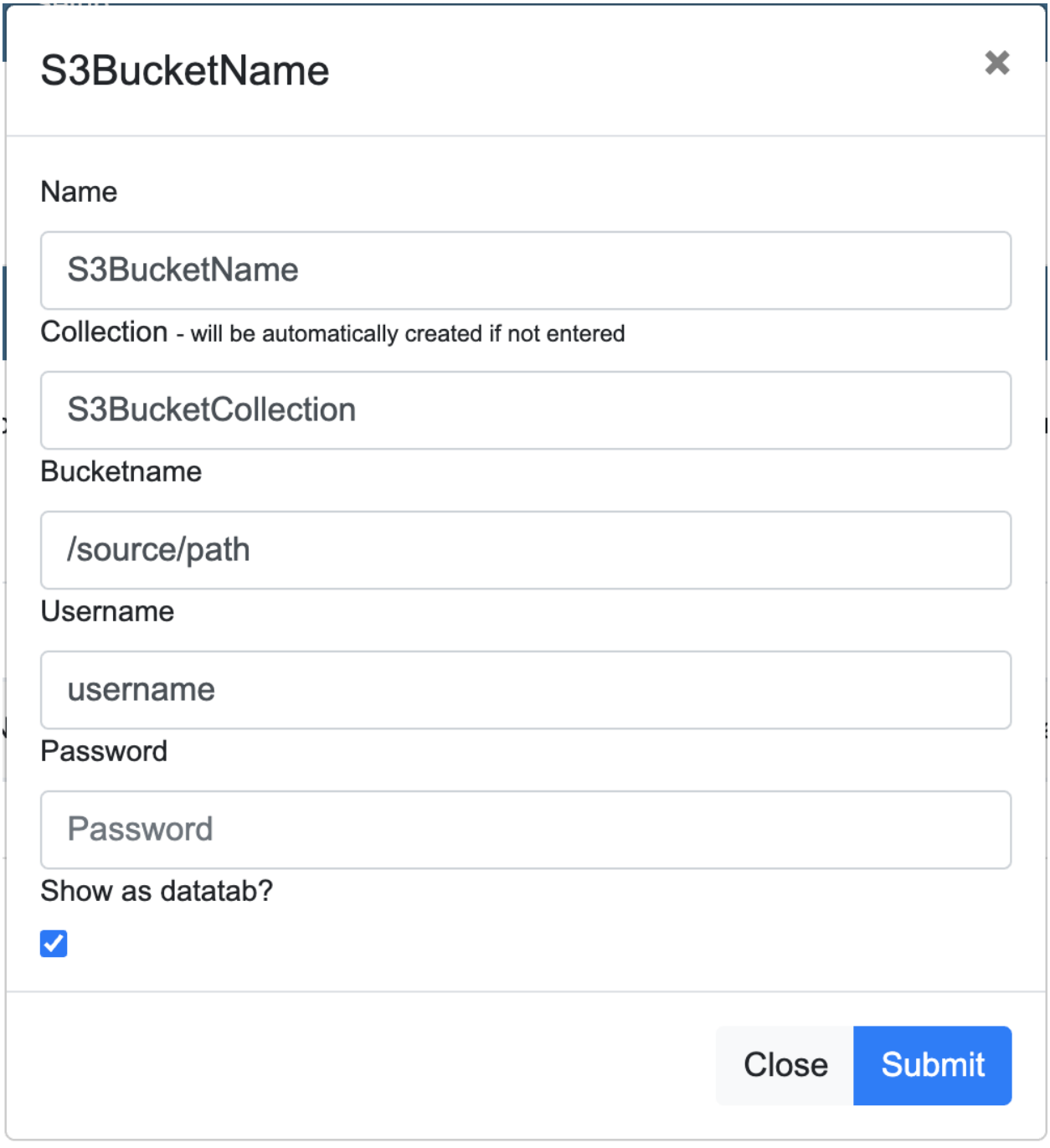
- **Name**
  Pick a name to identify your S3 Bucket with
- **Collection**
  Name of the collection to add the S3 Bucket to. Pick an existing collection to add this source to, or enter a new collection to create a new one.
- **Bucketname**
  Fill in the source path to your S3 bucket.
- **Username**

Fill in the username for the S3 bucket.

- **Password**
  Fill in the password for the S3 bucket.
- **Show as datatab?**
  Whether this query will be displayed as a datatab in the navigation bar in the search interface.

21.7.2 Editing S3 Buckets

To edit a S3 bucket, click the blue pen with paper icon (Edit) in the actions column. This will open a window:



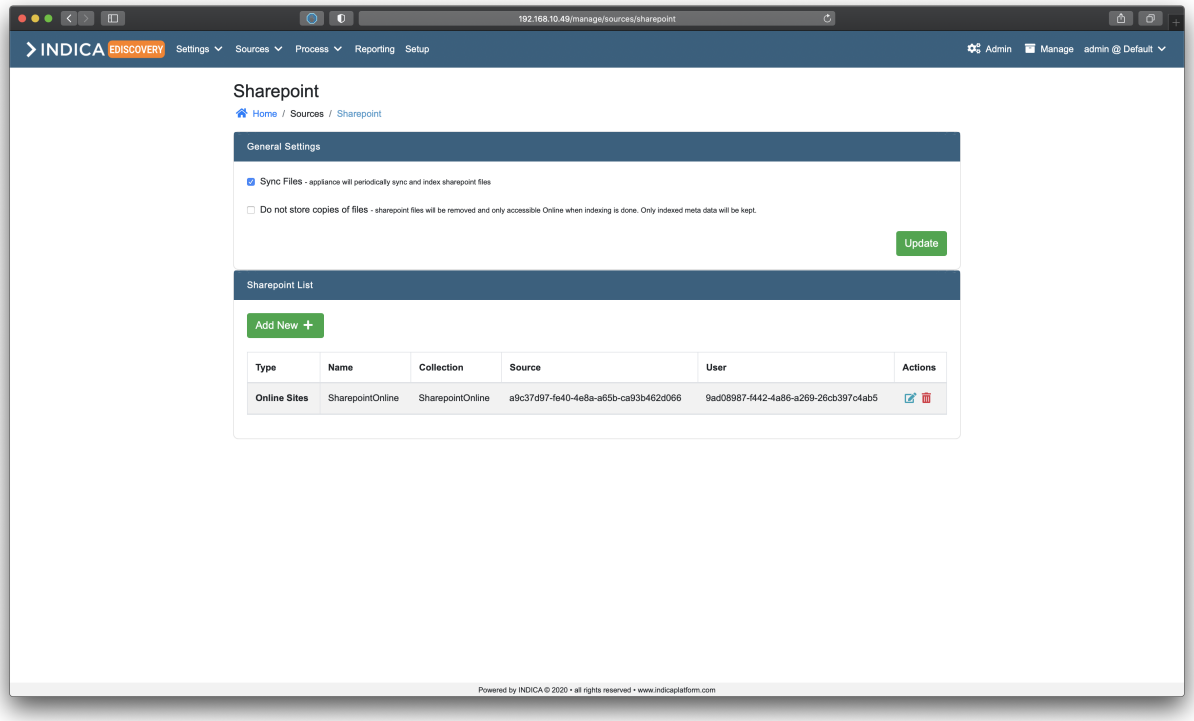
This window is identical to the "Add Queries" window, with the exception that all information is already filled in (except the password field). Here you can edit the information needed, re-enter your password and save the changes by clicking "Submit".

21.7.3 Deleting S3 Buckets

Removing a S3 bucket can be done by clicking the red trash-bin icon (delete) in the "Actions" column of the table.

21.8 SharePoint

INDICA can connect to SharePoint as well. It is possible to index online sites, on-premise sites, and online personal drives.



21.8.1 General Settings

There are two ways to index files stored in sharepoint:

- **Sync Files**
  The INDICA appliance will periodically synchronise, download and index sharepoint files. A copy of the files will be kept on disk. Please be advised that sufficient disk space is needed for this option.
- **Do not store copies of files**
  Sharepoint files will be removed after indexing and are only accessible Online when indexing is done. Only indexed meta data will be kept. In some situations, previewing a document will fall back to plain text when a document is unavailable.

Pick one of the two options, and click te "Update" button for the changes to take effect.

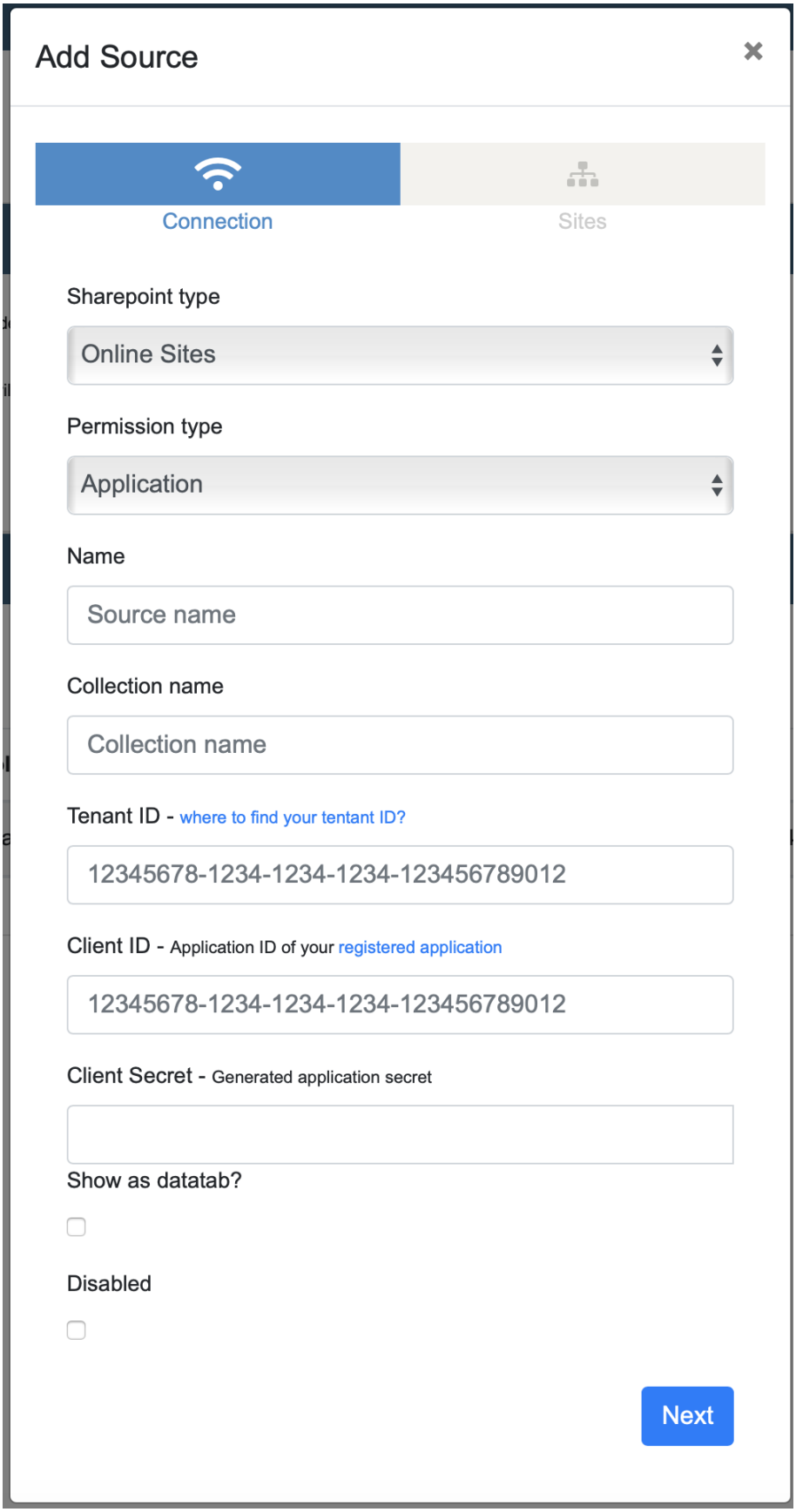
21.8.2 Adding a sharepoint source

Preparation

When using Application permissions, some extra preparation needs to be done before adding the source. The steps needed to take can be found on the INDICA support knowledge-base .

Adding a source

Adding a sharepoint source can be done by clicking the “Add New” button in the sharepoint sources list table. A new window will open:


Add Source
Connection
Sites
Sharepoint type
Online Sites
Permission type
Application
Name
Source name
Collection name
Collection name
Tenant ID - where to find your tentant ID?
12345678-1234-1234-1234-123456789012
Client ID - Application ID of your registered application
12345678-1234-1234-1234-123456789012
Client Secret - Generated application secret
Show as datatab?
Disabled
Next

The first step is picking what kind of sharepoint will be added as a data source. A choice can be made between:

- Online sites;
- On-premise sites;
- Online personal drives.

Depending on the chosen option, the form will ask for specific information. The following paragraphs will describe the three options in more detail.

Online sites

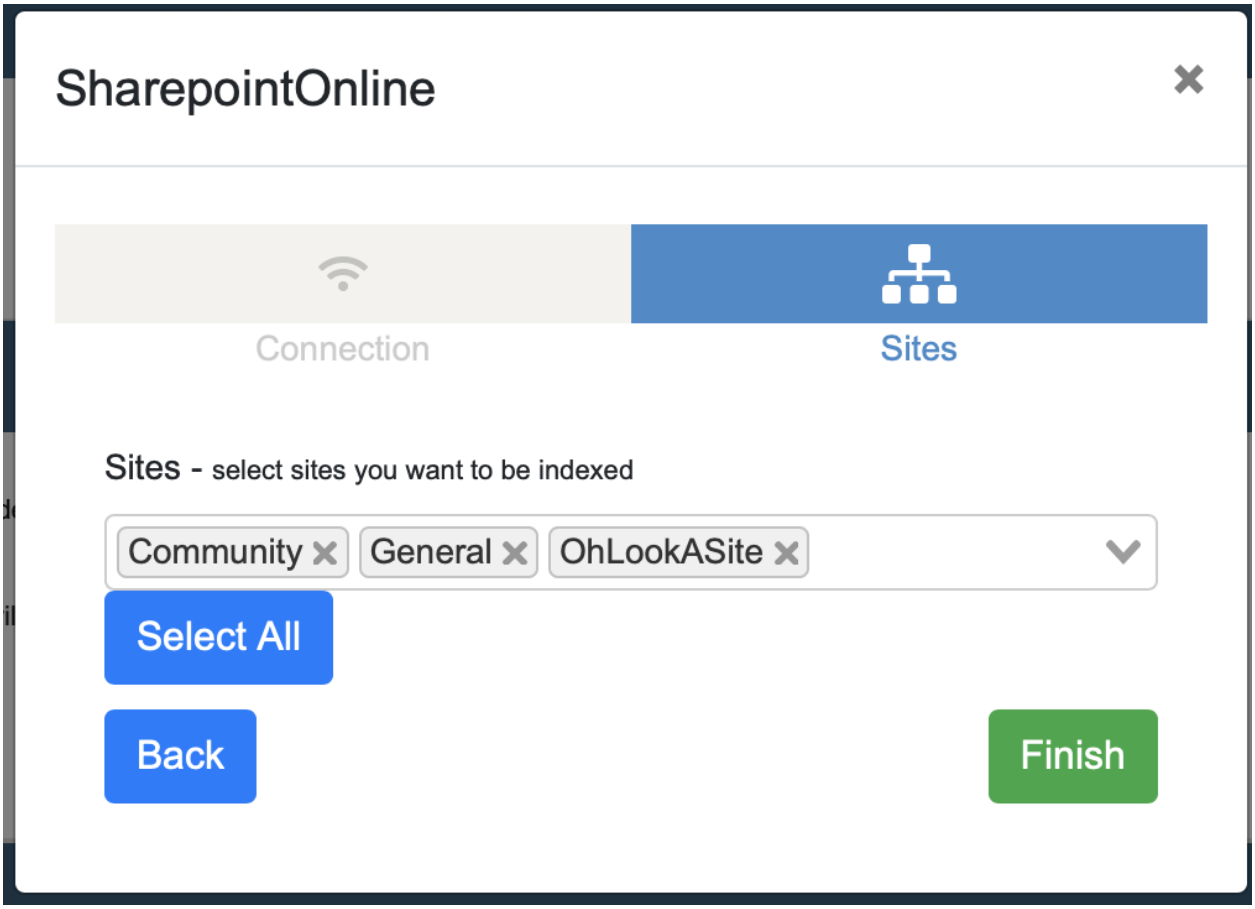
When "Online sites" has been chosen, the following information is needed:

- **Permission type**

  This can be Application, or user delegated.
  When "Application" is chosen, INDICA will use pre-configured API permissions to connect to sharepoint as an application.
  When "User delegated" is chosen, INDICA will inherit the current Microsoft session. This requires authenticating with the chosen user by Microsoft.

- **Name**
  Name of the sharepoint connection. This can be chosen freely.
- **Collection name**
  Name of the collection to add the local source to. Pick an existing collection to add this source to it, or enter a new collection to create a new one;
- **Tenant ID**
  Tenant ID is the ID used by Microsoft as a unique identifier of your organisation.
- **Client ID**
  Application ID of your registered application
- **Client Secret**
  Token that will be used to authenticate as an application by Microsoft.
- **Show as datatab?**
  Whether this source will be displayed as a datatab in the navigation bar.
- **Disabled**
  Allows you to enable or disable this source.

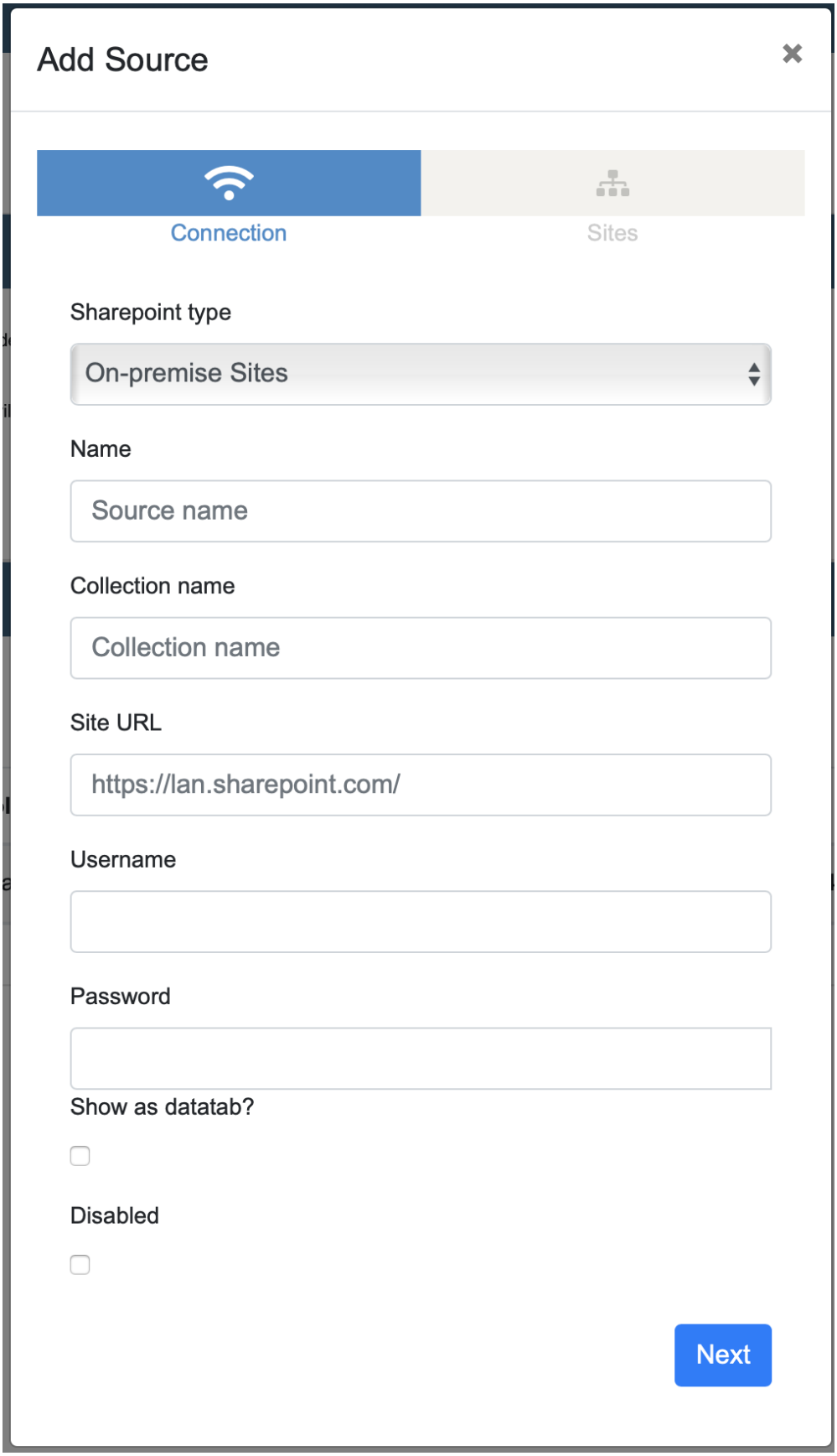
After filling out the required information, click the "Next" button. INDICA will now communicate with sharepoint and check if the settings are correct. If this is succesful, you will be presented with the following screen:



Here you have the possibility to select the sharepoint sites to index. It is also possible to select all sites with the "Select All" button. When the sites have been chosen, click "Finish" to add this source to INDICA. Downloading and indexation of the sites will begin shortly.

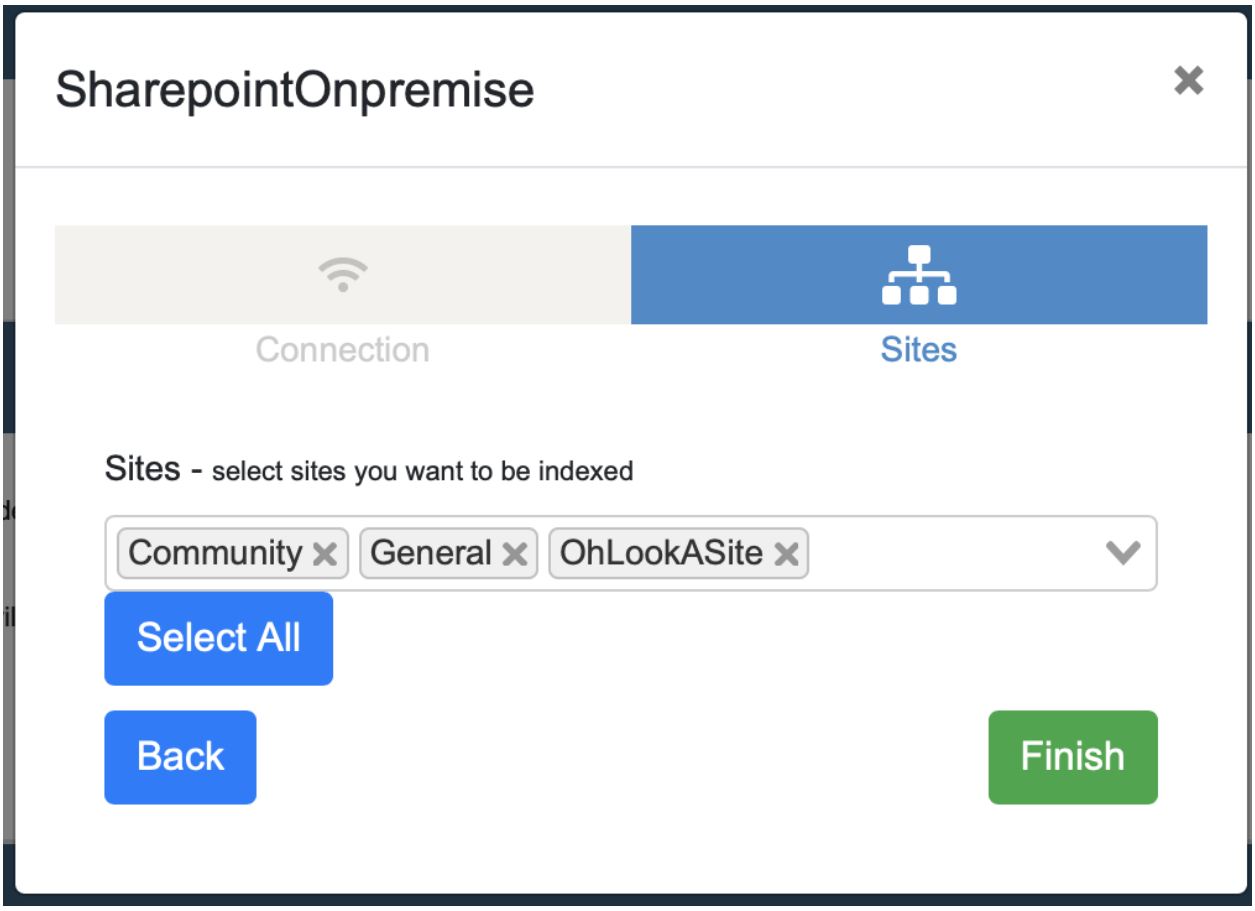
On-premise sites

The needed information for on-premise sites is slightly different. If this option has been selected, the following information is needed:


Add Source
Connection
Sites
Sharepoint type
On-premise Sites
Name
Source name
Collection name
Collection name
Site URL
https://lan.sharepoint.com/
Username
Password
Show as datatab?
Disabled
Next

- **Name**
  Name of the sharepoint connection. This can be chosen freely.
- **Collection name**
  Name of the collection to add the local source to. Pick an existing collection to add this source to it, or enter a new collection to create a new one;
- **Site URL**
  Full URL of the sharepoint instance;
- **Username**
  Username of the user that will be used for authenticating;
- **Password**
  Password of the corresponding user;
- **Show as datatab?**
  Whether this source will be displayed as a datatab in the navigation bar.
- **Disabled**
  Allows you to enable or disable this source.

After filling out the required information, click the "Next" button. INDICA will now communicate with sharepoint and check if the settings are correct. If this is succesful, you will be presented with the following screen:



Here you have the possibility to select the sharepoint sites to index. It is also possible to select all sites with the "Select All" button. When the sites have been chosen, click "Finish" to add this source to INDICA. Downloading and indexation of the sites will begin shortly.

Online personal drives

The needed information for online personal drives is the same as for online sites. If this option has been selected, the following information is needed:

- **Permission type**

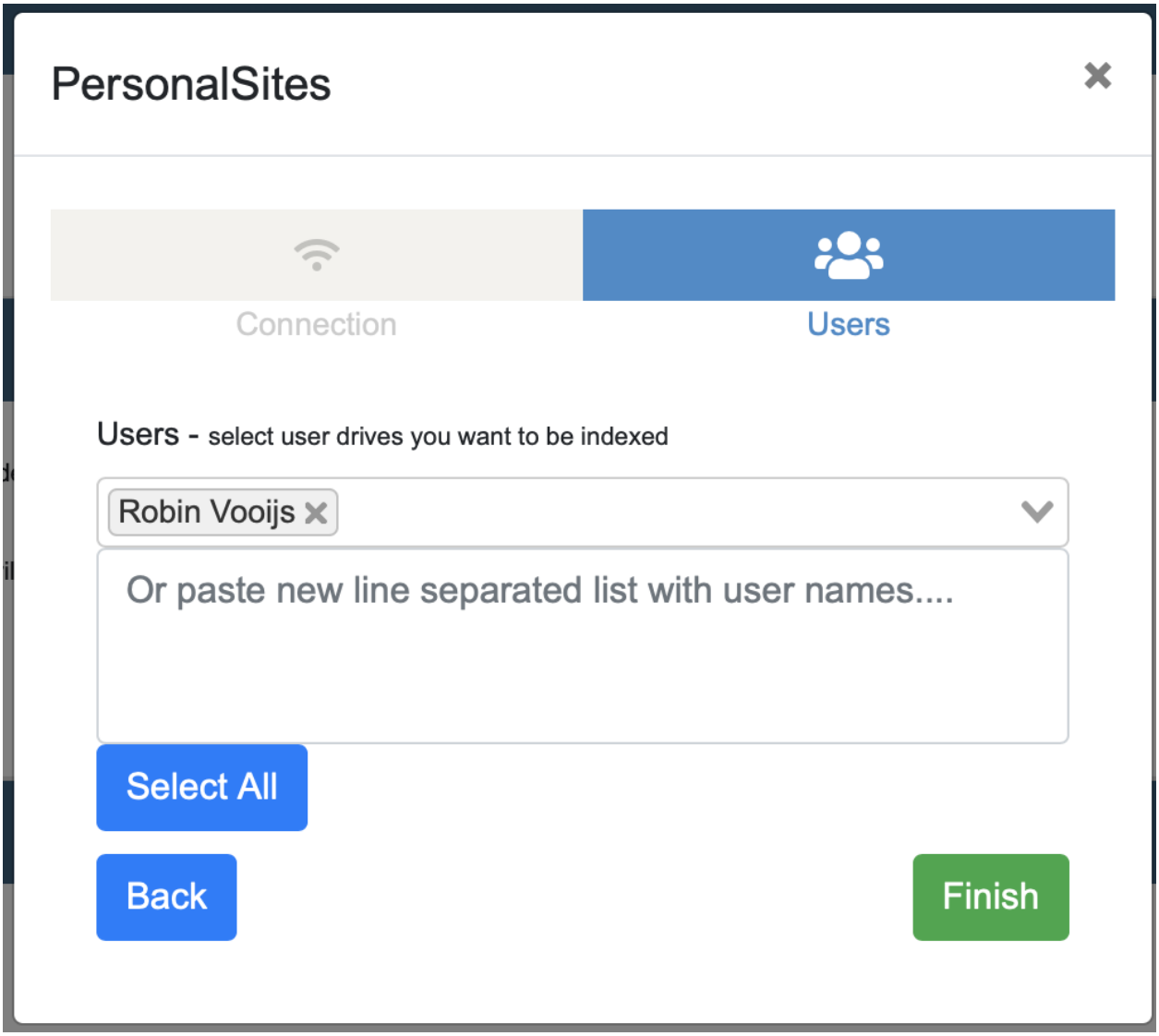
  This can be Application, or user delegated.

  When "Application" is chosen, INDICA will use pre-configured API permissions to connect to sharepoint as an application.

  When "User delegated" is chosen, INDICA will inherit the current Microsoft session. This requires authenticating with the chosen user by Microsoft.

- **Name**
  Name of the sharepoint connection. This can be chosen freely.
- **Collection name**
  Name of the collection to add the local source to. Pick an existing collection to add this source to it, or enter a new collection to create a new one;
- **Tenant ID**
  Tenant ID is the ID used by Microsoft as a unique identifier of your organisation.
- **Client ID**
  Application ID of your registered application
- **Client Secret**
  Token that will be used to authenticate as an application by Microsoft.
- **Show as datatab?**
  Whether this source will be displayed as a datatab in the navigation bar.
- **Disabled**
  Allows you to enable or disable this source.

After filling out the required information, click the "Next" button. INDICA will now communicate with sharepoint and check if the settings are correct. If this is succesful, you will be presented with the following screen:



Here you can pick the personal sites of the users from the dropdown list. It is also possible to paste a list (new-line separated) with the usernames of the personal sites that need to be indexed. Alternatively, you can click the "Select All" button to select all the users.

When done selecting users, click "Finish" to save the source to the appliance. The indexation of the personal sites will begin shortly.

21.8.3 Editing a sharepoint source

Editing a sharepoint source can be done by clicking the blue "paper" icon (edit) in the "Actions" column of the table. A new window will open with all the information already filled in. If needed, the information can be changed and saved right away. If changes need to be made to the list of sites or users, it is possible to go to the next step and edit the list of users/sites. Click "Finish" to save the new settings.

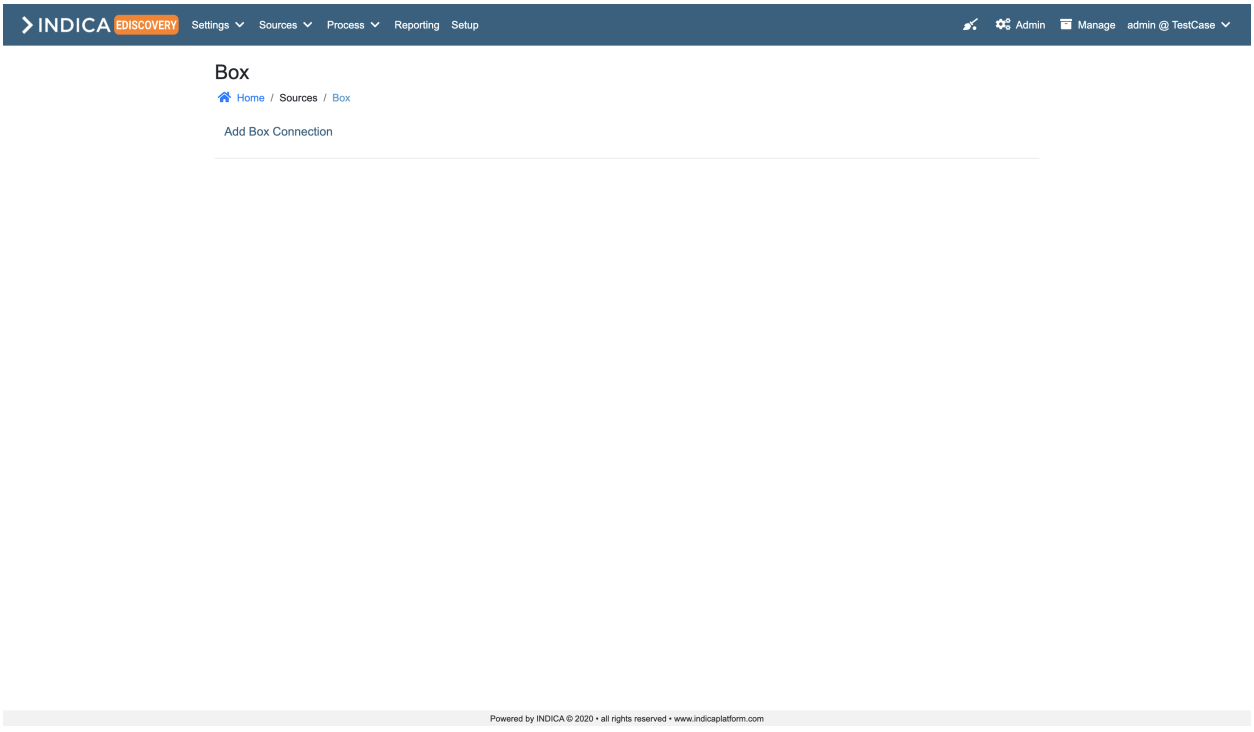
21.8.4 Deleting a sharepoint source

Deleting a SharePoint source can be done by clicking the red 'trashbin' icon in the "Actions" column of the table. It will ask for a confirmation before deleting the source.

> **Warning:** Removing a SharePoint source will remove all corresponding documents from the index. This is irreversible, and can only be undone by re-indexing the SharePoint source.

21.9 Box

INDICA can connect to Box.com sources as well. To set up your Box connector, head over to the manage interface, click on *Sources* and navigate to *Box*.



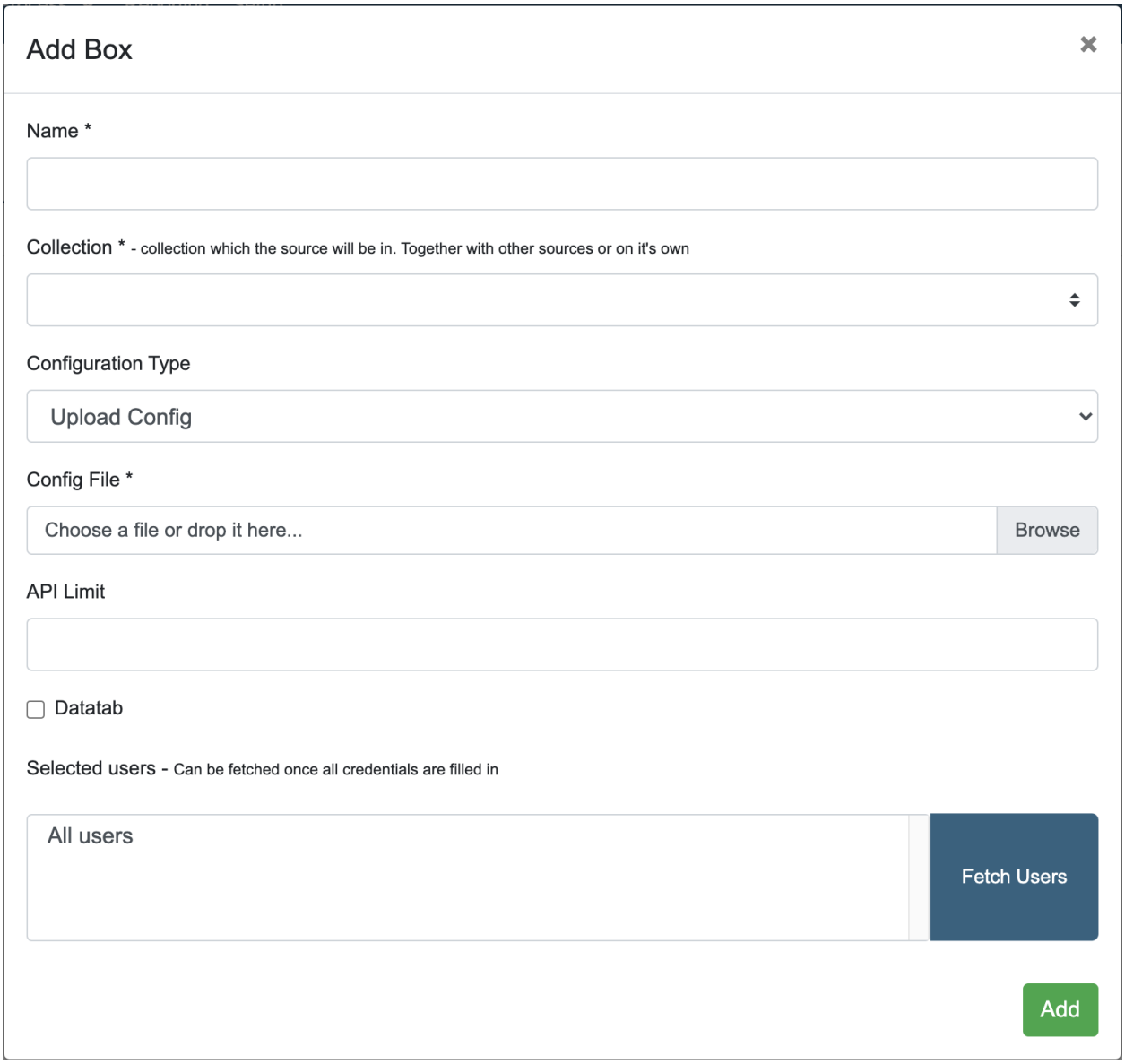
21.9.1 Adding a box source

Box Preparation

Before you can add a Box source to INDICA, you'll need to prepare your Box app first. The steps needed to take can be found on the INDICA support knowledge-base under box connector .

Connecting Box to INDICA

After walking through the steps in *Box Preparation* you can connect your Box to INDICA. Adding a box connector can be done by clicking on the *Add Box Connector* button. This will open a new window:



Here, you can fill in the following information:

- **Name**
  Name of the Box connection. This can be chosen freely.

- **Collection**
  Name of the collection to add the Box to. Pick an existing collection to add this source to it, or enter a new collection to create a new one.
- **Configuration Type**

  You can pick between two options here. This will change your form depending on which one you pick.

  - **Upload config**
    - **Config File**
      Choose between your files or drop your config file here.
  - **Manual**
    - **Client ID**
      Fill in your Box Client ID.
    - **Client Secret**
      Fill in your Box Client Secret.
    - **Enteprise ID**
      Fill in your Box Enteprise ID.
    - **Public Key ID**
      Fill in your Box Public Key ID.
    - **Private Key File**
      Choose the file from your computer where your private key file is located.
    - **Private Key Passphrase**
      Fill in your private key passphrase.
- **API limit**
  Fill in the limit for the API here.
- **Datatab**
  Whether this source will be displayed as a datatab in the navigation bar.
- **Selected users**
  Select the users you wish to index the files from. The users can be fetched once the credentials are filled in correctly.

21.9.2 Editing a box source

Editing a box source can be done by clicking the blue "paper" icon (edit) in the "Actions" column of the table.

21.9.3 Deleting a box source

Deleting a box source can be done by clicking the red 'trash bin' icon in the "Actions" column of the table.

CHAPTER

TWENTYTWO

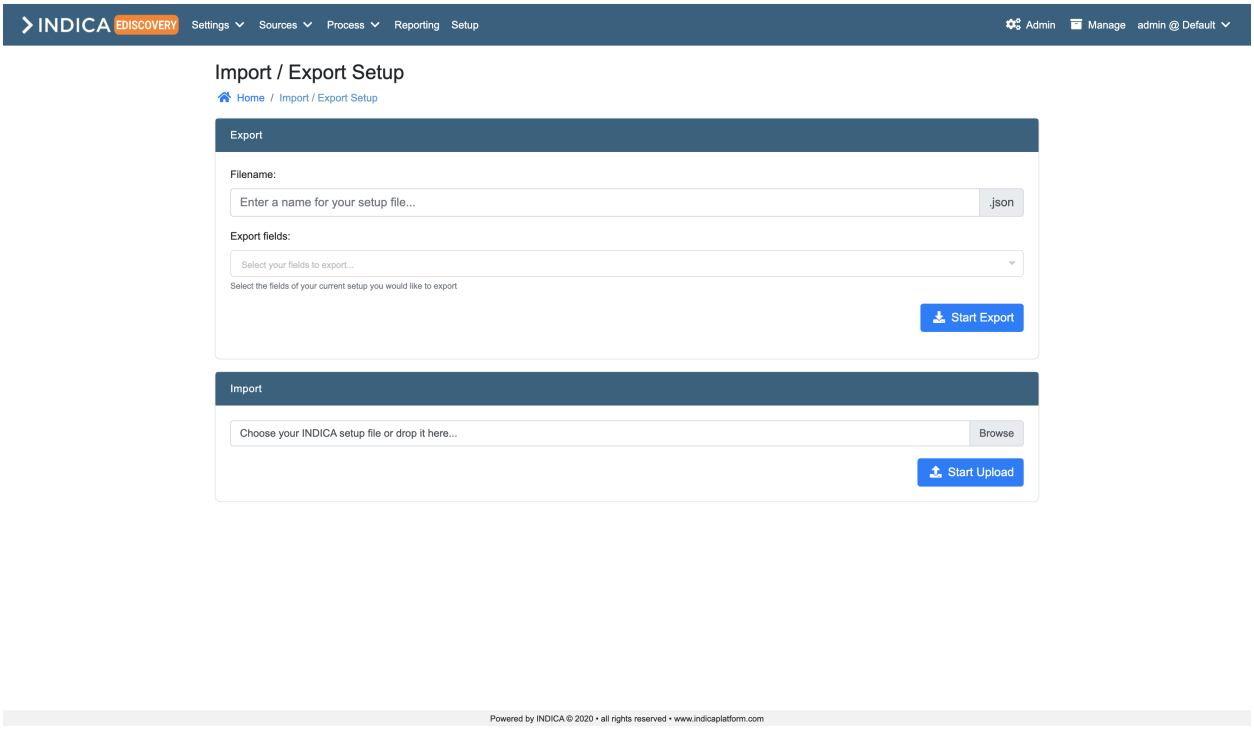
MANAGER SETUP

Table of Contents

- *Manager Setup*
  - *Export*
  - *Import*

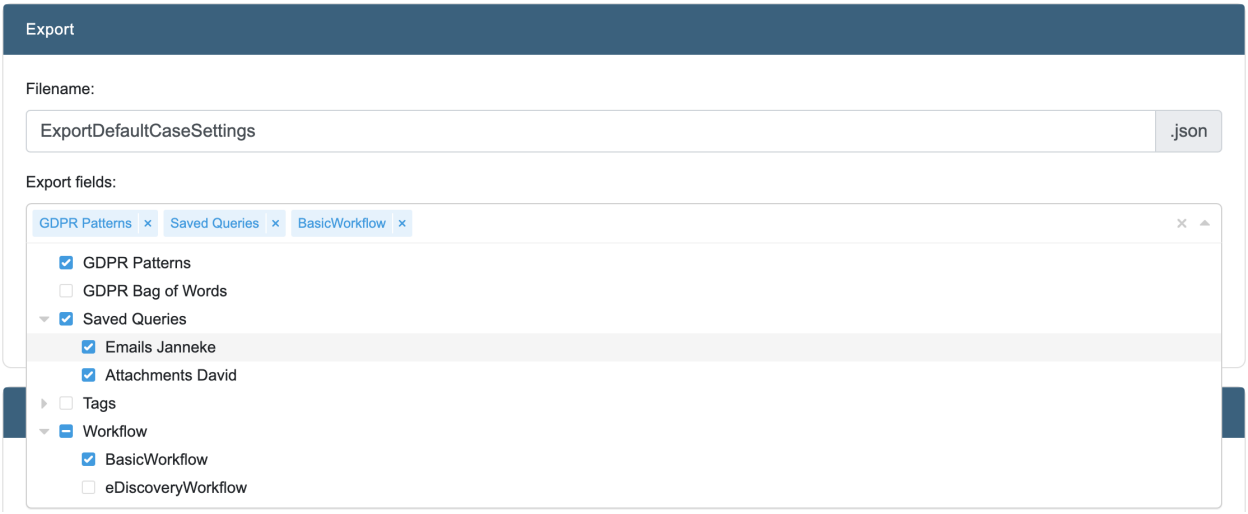
With INDICA it's possible to export and import your case settings. This can be useful if you want to reuse some of the settings in any of your new cases for a quick setup. The following settings are available for exporting and importing:

- GDPR Settings (if GDPR is enabled)
- GDPR Bag of Words (if GDPR is enabled)
- Saved Queries
- Tags
- Workflow



22.1 Export

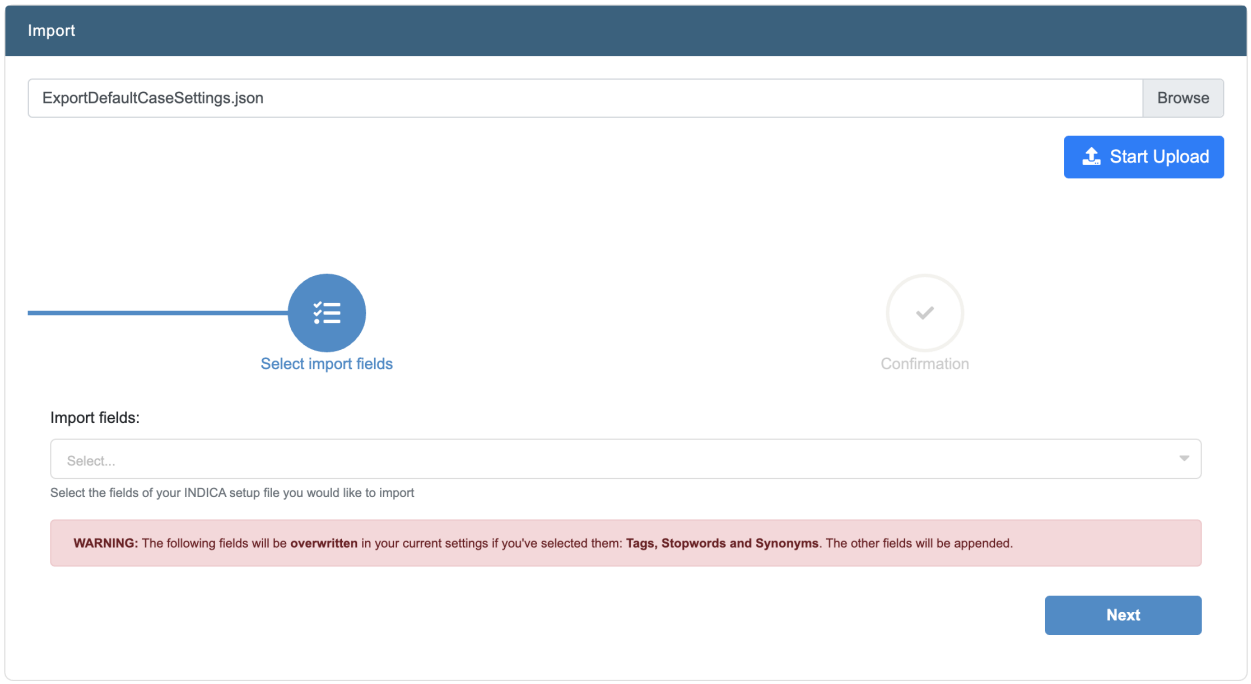
Within this panel it's possible to export your settings. Choose a name for your file and select the settings you wish to export. For your Saved Queries, Tags and Workflows it's also possible to select specific items from the list instead of export all of them. Select the settings you wish to export and click on the "Start Export" button. This will download a .json file with your settings.



22.2 Import

To import settings in your new case, upload your exported settings .json file. The file will be processed and gives you the option to review the imported file. From here on you can choose to import all settings within the file, or only import a selected few. This makes it easier to reuse settings you need only for that case.

Caution: An important thing to keep in mind, is that **tags will be overwritten**. This means that any tags existing in your case will be overwritten by the import file (if you choose to import the tags from the file, of course). All the other settings will be appended.


Import
ExportDefaultCaseSettings.json
Browse
Start Upload
Select import fields
Confirmation
Import fields:
Select...
Select the fields of your INDICA setup file you would like to import
WARNING: The following fields will be overwritten in your current settings if you've selected them: Tags, Stopwords and Synonyms. The other fields will be appended.
Next

CHAPTER

TWENTYTHREE

LOGIN AND CREDENTIALS

Table of Contents

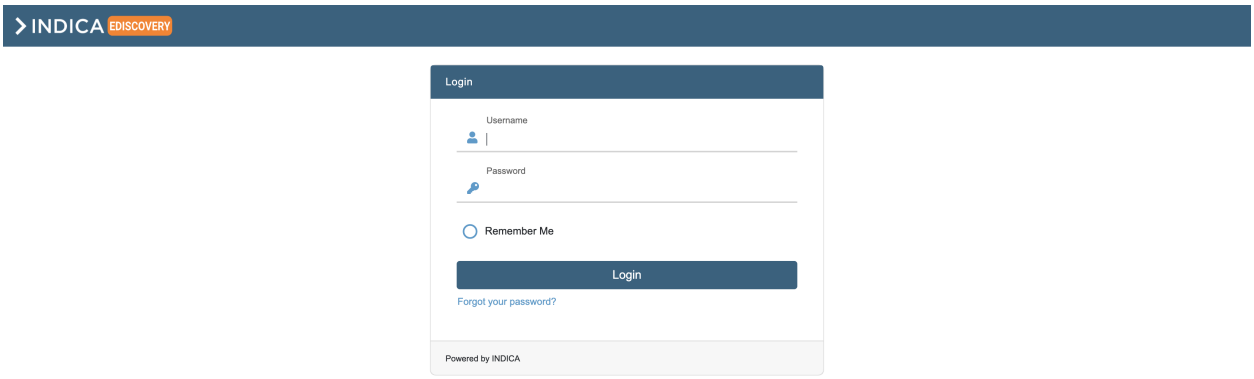
- *Login and Credentials*
  - *Two-Factor Authentication*

When navigating to your INDICA platform, you can log in using the information provided by your system administrator. Confirm your credentials with the INDICA administrator. We recommend you use either one of these browsers to access INDICA:

- Chrome
- Safari
- Mozilla Firefox

Caution: We specifically **do not recommend** you to use Internet Explorer.

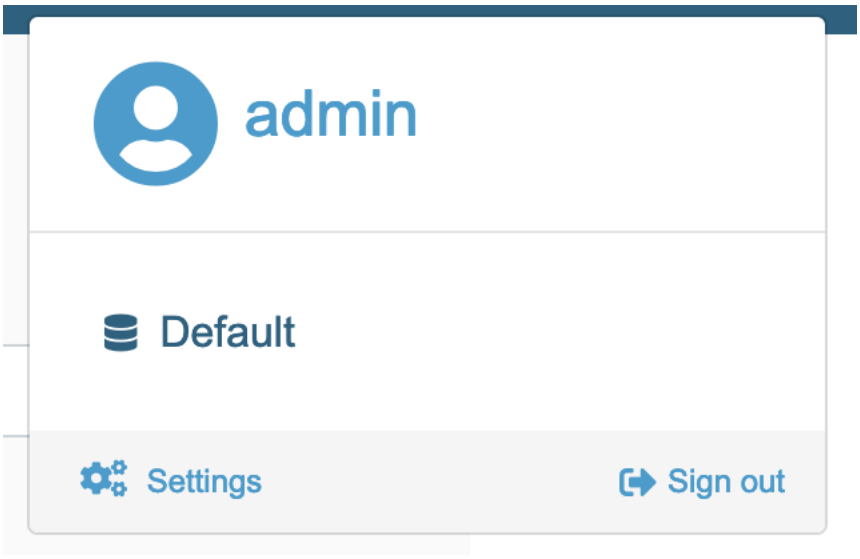
See image below for an example of the login page.



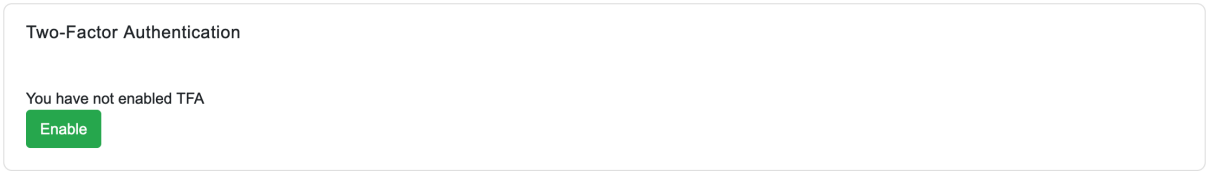
In case you experience any troubles when logging in or displayed data significantly differs from your previous experiences, please log off and clear your browser cookies or alternatively open INDICA in an Incognito window.

23.1 Two-Factor Authentication

Users may also enable two-factor authentication. This can be done by navigation to the user settings from the case dropdown in the upper right corner.



Once in the user settings, two-factor authentication can be enabled. This will require the use of a smartphone with an authenticator app. We recommend the **Authy** and **Google Authenticator** apps. All that is left to do is to follow the instructions shown and validate two-factor authentication.

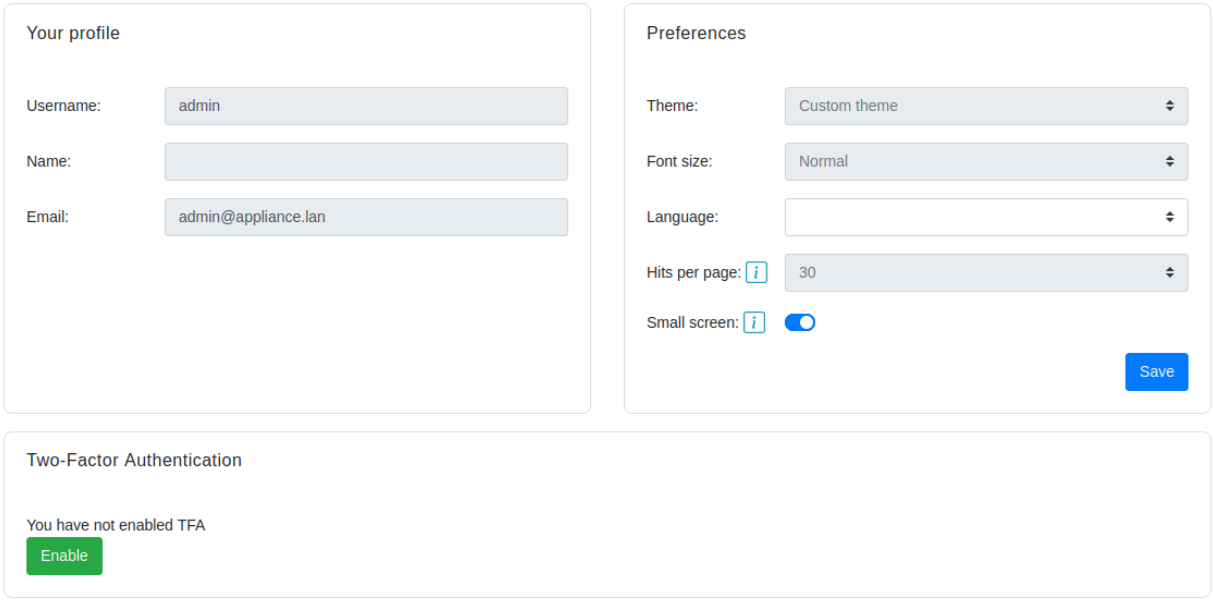


Warning: Make sure to write down the recovery codes shown and put them in a safe place. If access to the authenticator app is lost and 2-factor authentication is enabled, recovery codes are the only way to log in.

CHAPTER

TWENTYFOUR

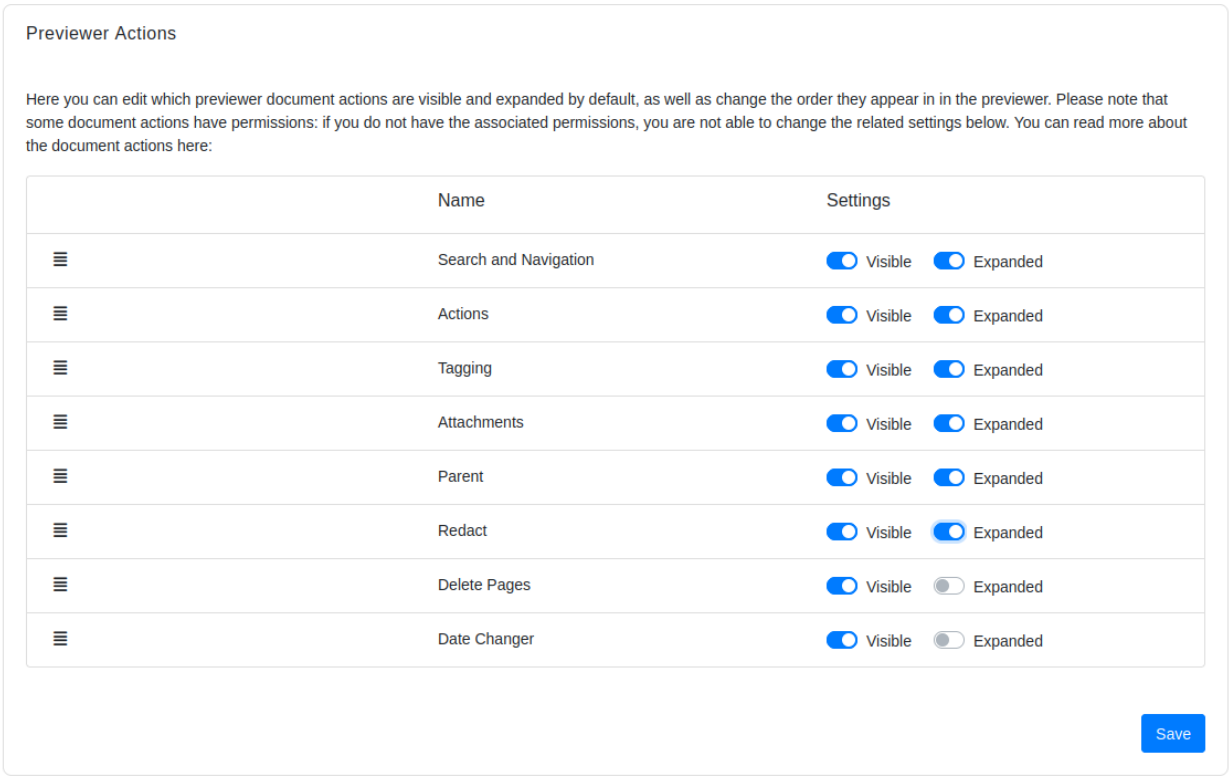
USER SETTINGS



Every user can change certain settings related to their user account. These settings are outlined in the image above. The user can change the main language of the interface, and select whether or not you are using INDICA on a small screen.

One of the main settings is the ability to set two-factor authentication. The user requires an authenticator app (we recommend Authy) to enable two-factor authentication and in order to log in once two-factor authentication is enabled.

The user can also change the order in which previewer document actions are shown in the previewer, by dragging and dropping actions and changing whether or not they are visible or expanded by default. Some previewer document actions, such as tagging an redaction, have user permissions attached to them. If the user does not have the necessary permissions, they are not able to change the order and visibility of that specific document action. You can read more about these document actions here: *Document Actions*.


Previewer Actions
Here you can edit which previewer document actions are visible and expanded by default, as well as change the order they appear in in the previewer. Please note that some document actions have permissions: if you do not have the associated permissions, you are not able to change the related settings below. You can read more about the document actions here:
Name
Settings
Search and Navigation
Visible
Expanded
Actions
Visible
Expanded
Tagging
Visible
Expanded
Attachments
Visible
Expanded
Parent
Visible
Expanded
Redact
Visible
Expanded
Delete Pages
Visible
Expanded
Date Changer
Visible
Expanded
Save

CHAPTER

TWENTYFIVE

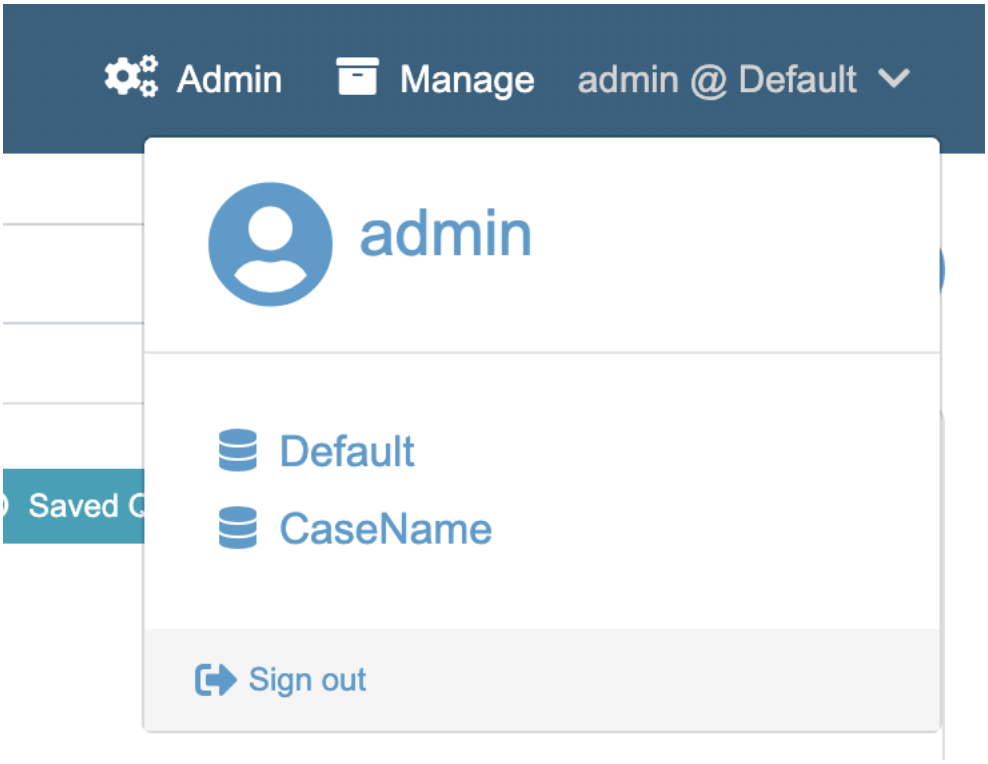
SWITCHING CASES

Table of Contents

- *Switching Cases*

Switching between different cases is easy in INDICA. Just click on the *username@casename* at the right most item in the menu at the top to switch to the case you need to work on. See the image below for an example.

Note: Depending on your rights, you might not see the "Admin" and "Manage" menu items.



CHAPTER

TWENTYSIX

FILTERING AND DATERANGE

Table of Contents

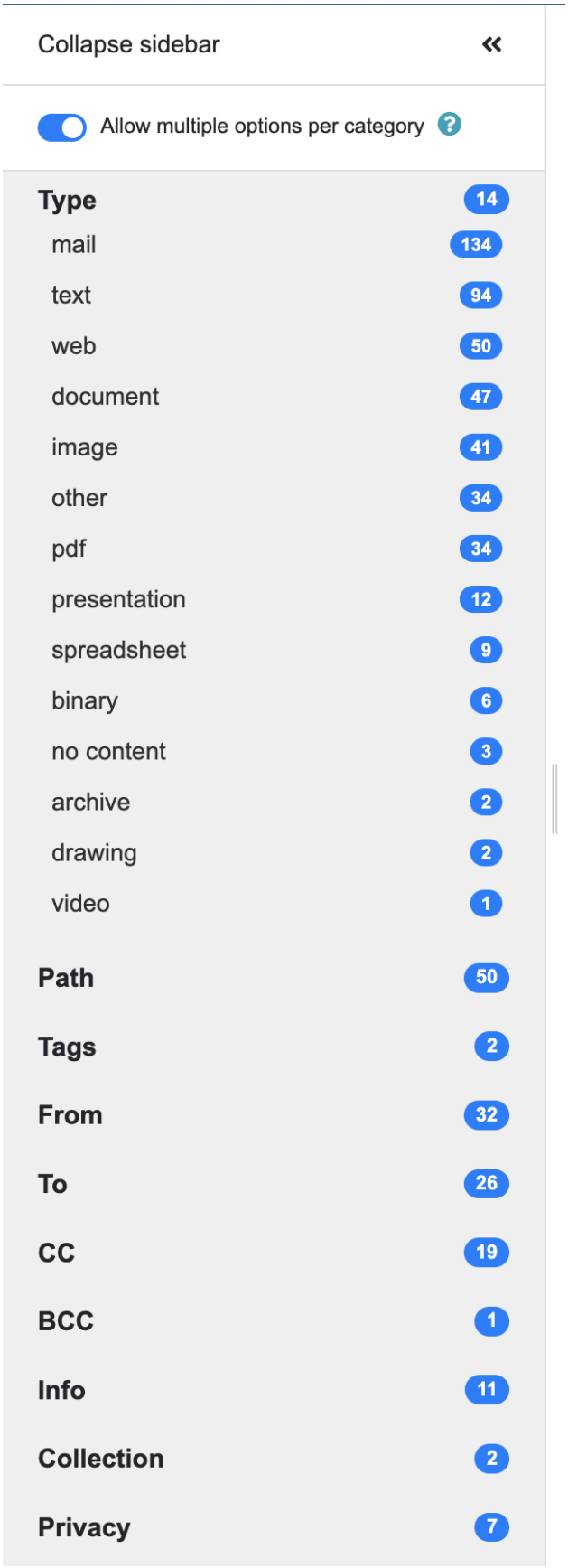
- *Filtering and Daterange*
  - *Sidebar*
    - *Expanding and collapsing the sidebar*
  - *Selecting multiple filters in one category*
  - *Daterange*

26.1 Sidebar

On the left side of the search overview in INDICA you can see most of the filtering options. These filters are categorized in a few categories:

- Type
- Path
- Tags
- Filter
- From
- To
- CC
- BCC
- Info
- Collection
- Access
- Dupes

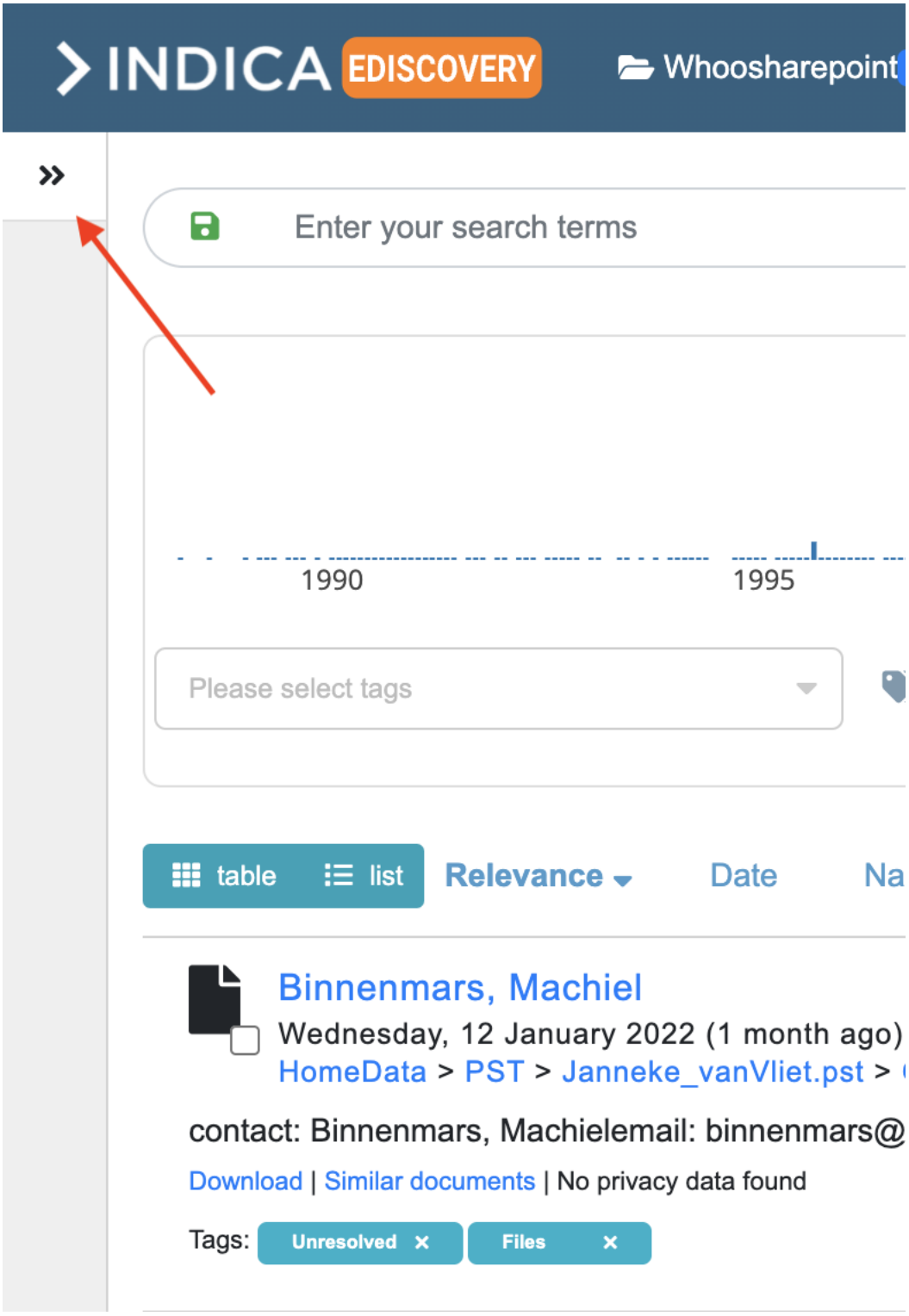
Your categories may vary depending on what is in your dataset. The number next to the filter shows the number of results within that filter.


Collapse sidebar
Allow multiple options per category
Type 14
mail 134
text 94
web 50
document 47
image 41
other 34
pdf 34
presentation 12
spreadsheet 9
binary 6
no content 3
archive 2
drawing 2
video 1
Path 50
Tags 2
From 32
To 26
CC 19
BCC 1
Info 11
Collection 2
Privacy 7

26.1.1 Expanding and collapsing the sidebar

Within the sidebar there is a button to collapse the sidebar. See the thumbnail below.

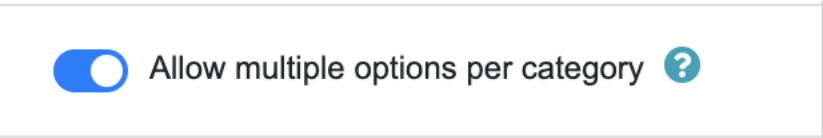
When clicking on the button, the sidebar will collapse. This will give you more screen space. To uncollapse the sidebar, just click the >> icon in the top of the sidebar. See the image below.



Tip: Since INDICA v12.0.0 it's also possible to resize the sidebar. Just next to the sidebar - in the middle of the screen - you will see the two stripes. If you click and drag your cursor over the screen it will resize to your preference.

26.2 Selecting multiple filters in one category

Sometimes you might want to select multiple filters in one category (e.g. the "Type" category). You are able to select multiple filters in one category by enabling the switch named "Allow multiple filters per category" at the top of the sidebar.



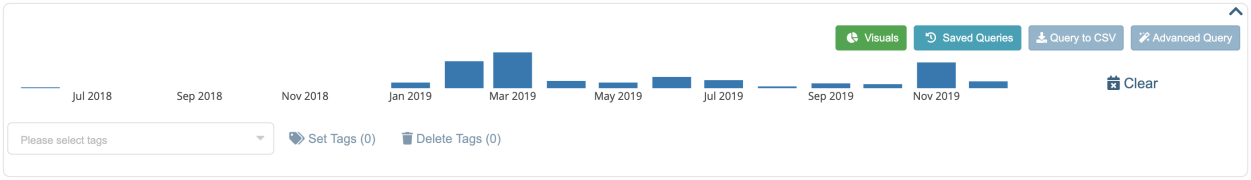
Eg. Let's say you decide to select two filters in the category Type: image and pdf.

- **When the switch is enabled** When selecting the image and pdf filters with the switch enabled, the search will give results for both image document OR pdf documents.
- **When the switch is disabled** When trying to select both types with the switch disabled, you'll notice that this is not possible. Trying to do this would result in documents that are *both* an image document AND a pdf document, which is not possible.

Tip: Whether the switch is enabled or disabled will be saved for your user. This means that if you enable the switch, this will be enabled until you decide to disable it.

26.3 Daterange

To set a date range, simply click and drag your mouse over the timeline. It immediately sets the filter the moment you let go of your mouse. To clear the date range filter, click on the clear button next to the timeline, or the yellow badge with the 'yyyymmdd' text underneath the search bar.



CHAPTER

TWENTYSEVEN

GDPR DASHBOARDING

The INDICA GDPR dashboard is a powerful tool to discover, manage and help clean up privay-sensitive information. It consists of the following pages:

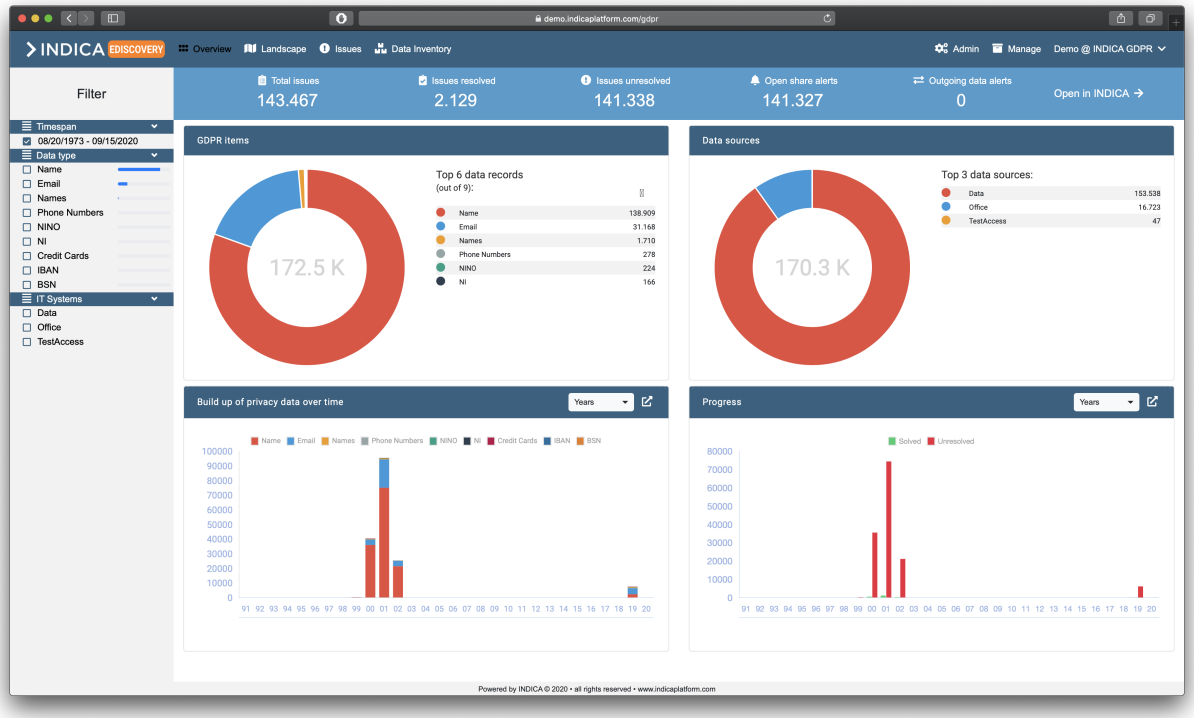
- **Overview**
  Gives a general overview of the number of privacy issues in the indexed data;
- **Landscape**
  Gives insight in the number of users that are able to access certain data items;
- **Issues**
  List of all the GDPR issues that are unresolved. This helps with picking up issues;
- **Data Inventory**
  Gives insight in the defined business rules and their corresponding data.

Every page in the GDPR dashboard features a sidebar where filters can be applied. These filters are:

- **Timespan**
  Filter on a timespan;
- **Data Type**
  Filter on one or more data types;
- **IT Systems**
  Filter on one or more IT systems;
- **Risk Levels (only on the "Issues" page)**
  Filter on one or more Risk Levels.

27.1 Overview Page

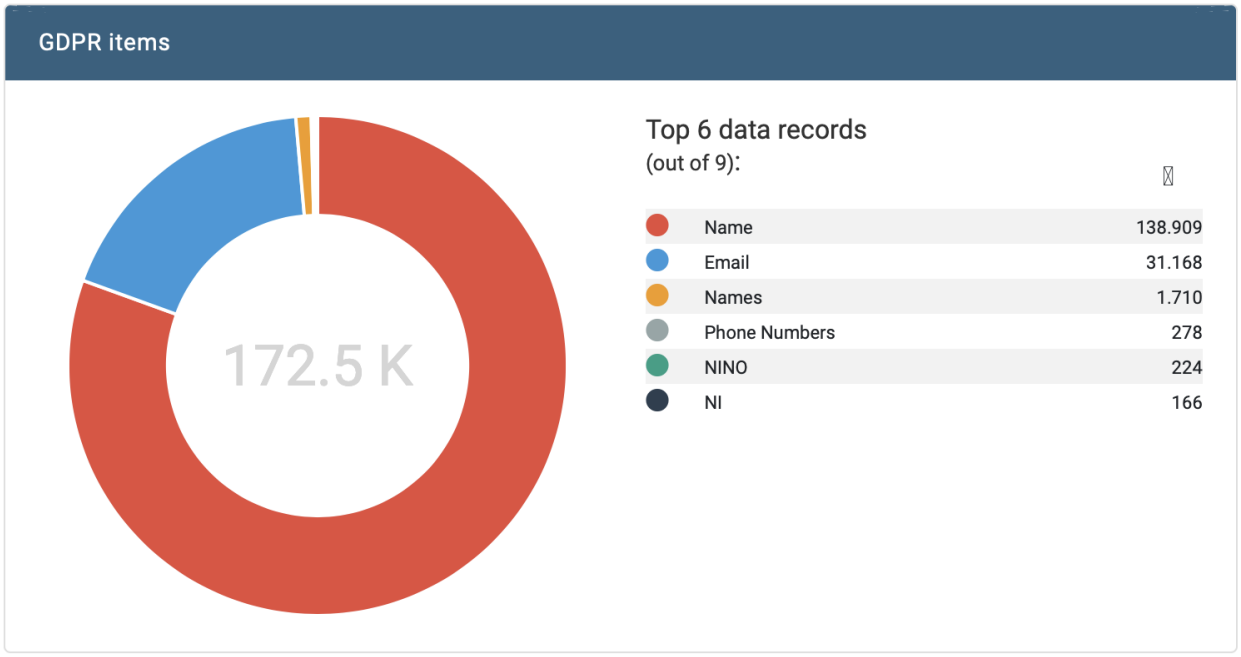
The overview page gives a general overview of the number of issues, how they are distributed, and from which data-sources they come.



Above the graphs, five counts are displayed:

- **Total issues**
  This is the number of documents that did/do contain one or more issues;
- **Issues resolved**
  This is the number of documents that used to be issues but are resolved;
- **Issues unresolved**
  This is the number of documents that currently have one or more issues;
- **Open share alerts**
  This is the number of open shares;
- **Outgoing data alerts**
  This is the number of times data was sent out via email.

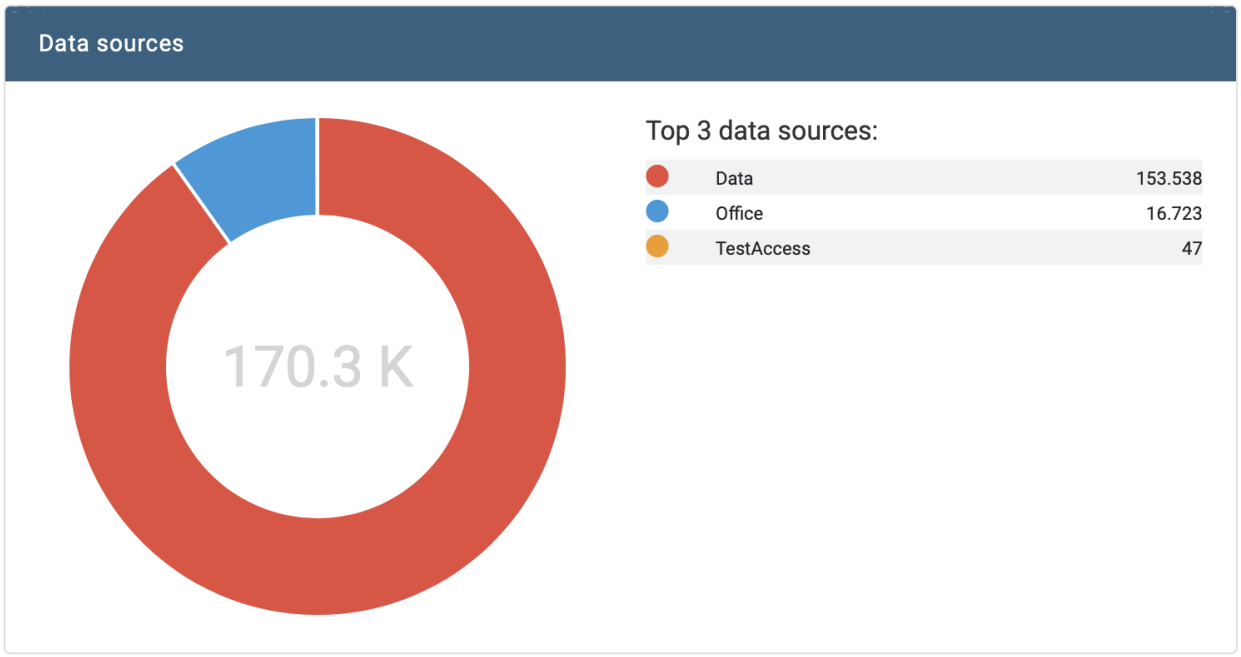
27.1.1 GDPR Items Graph



This graph shows the distribution of all found GDPR items by type. The colors show what kind of privacy token was found, and the number of times it was found.

This graph is interactive, meaning that it is possible to click one of the categories and directly see all corresponding files.

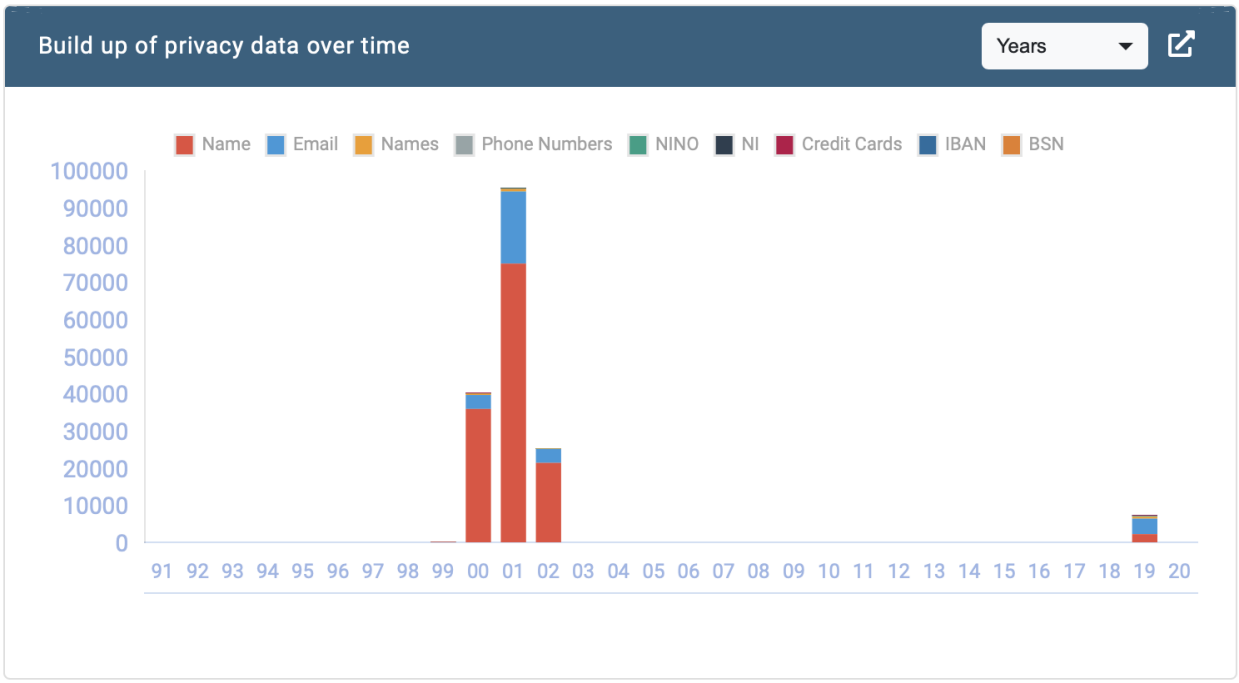
27.1.2 GDPR Sources Graph



This graph shows the distribution of found GDPR items over the connected datasources. This makes it possible to quickly get a grasp of where most items come from.

This graph is interactive, meaning that it is possible to click one of the categories and directly see all corresponding files.

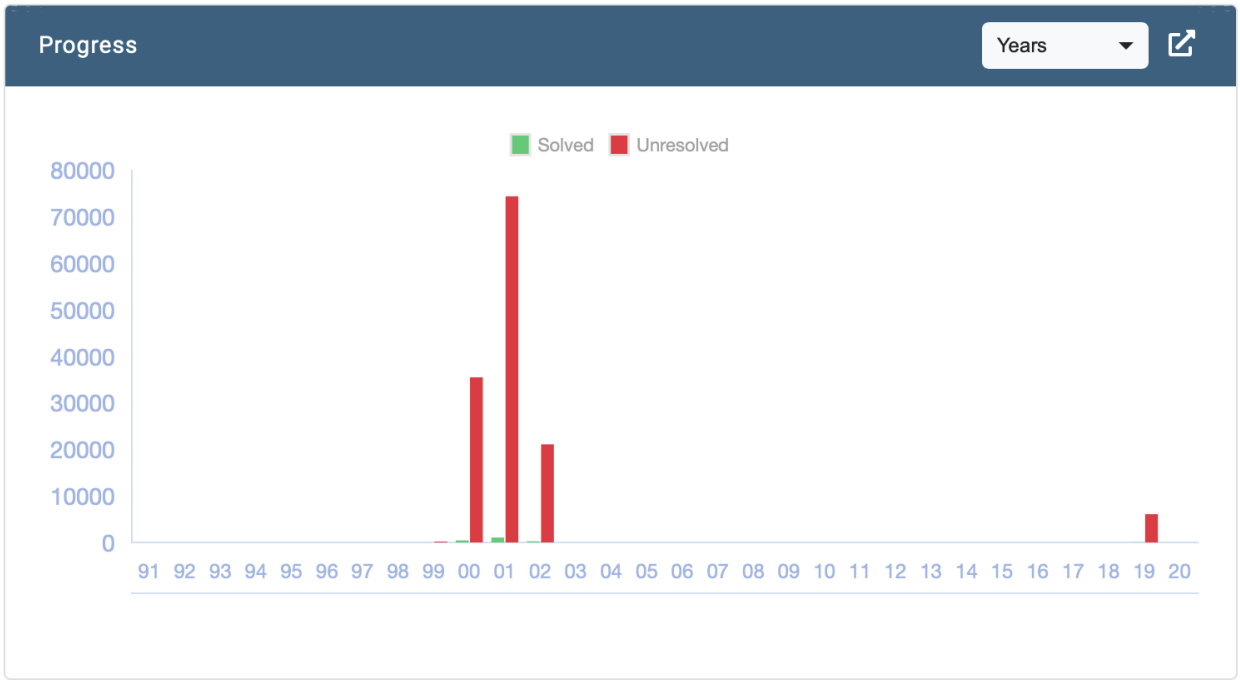
27.1.3 GDPR Buildup Over Time Graph



The "Build up of privacy data over time" graph shows how many privacy-sensitive documents have been created in the past. This graph can be switched between 'Years' and 'Months' for a more detailed view. It is also possible to 'pop-out' this graph for a bigger image. This can be done by clicking the icon in the top right corner of the card.

This graph is interactive, meaning that it is possible to click one of the categories and directly see all corresponding files. This also takes the date-range into account.

27.1.4 GDPR Progress Graph

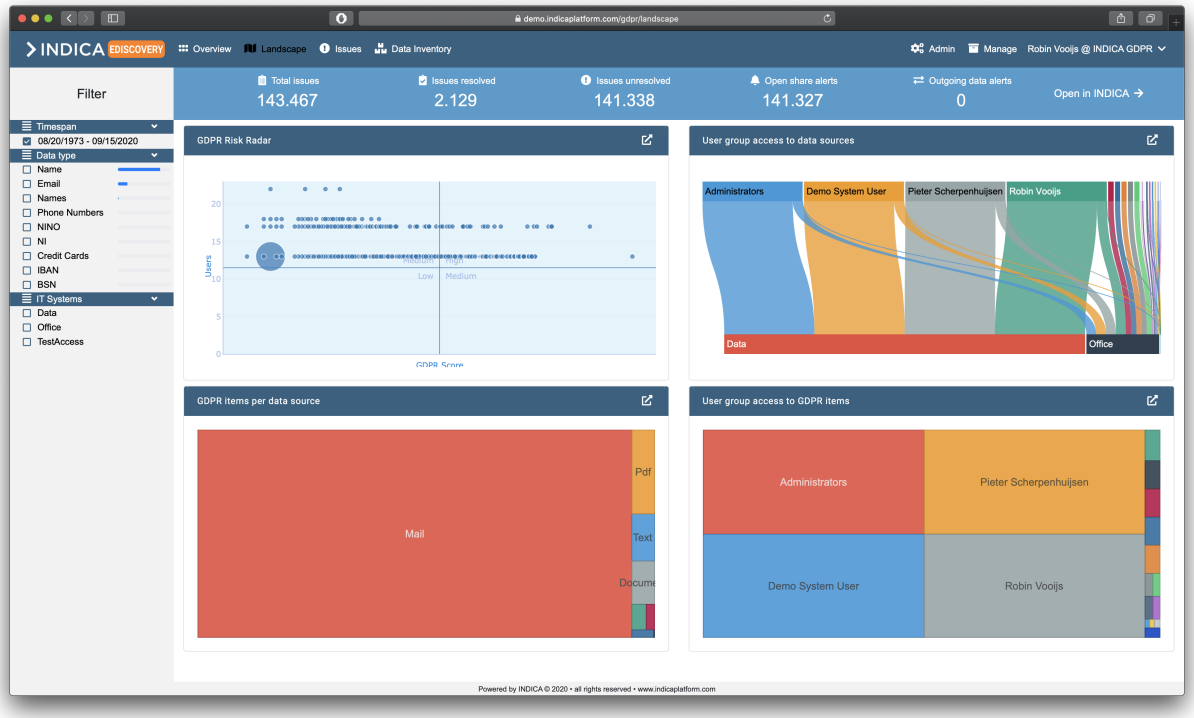


This graph shows you the progress that has been made with the cleanup of issues. When more work is done, the red bars become smaller and the green bars will grow until evetually all issues are resolved. This graph can be switched between 'Years' and 'Months' for a more detailed view. It is also possible to 'pop-out' this graph for a bigger image. This can be done by clicking the icon in the top right corner of the card.

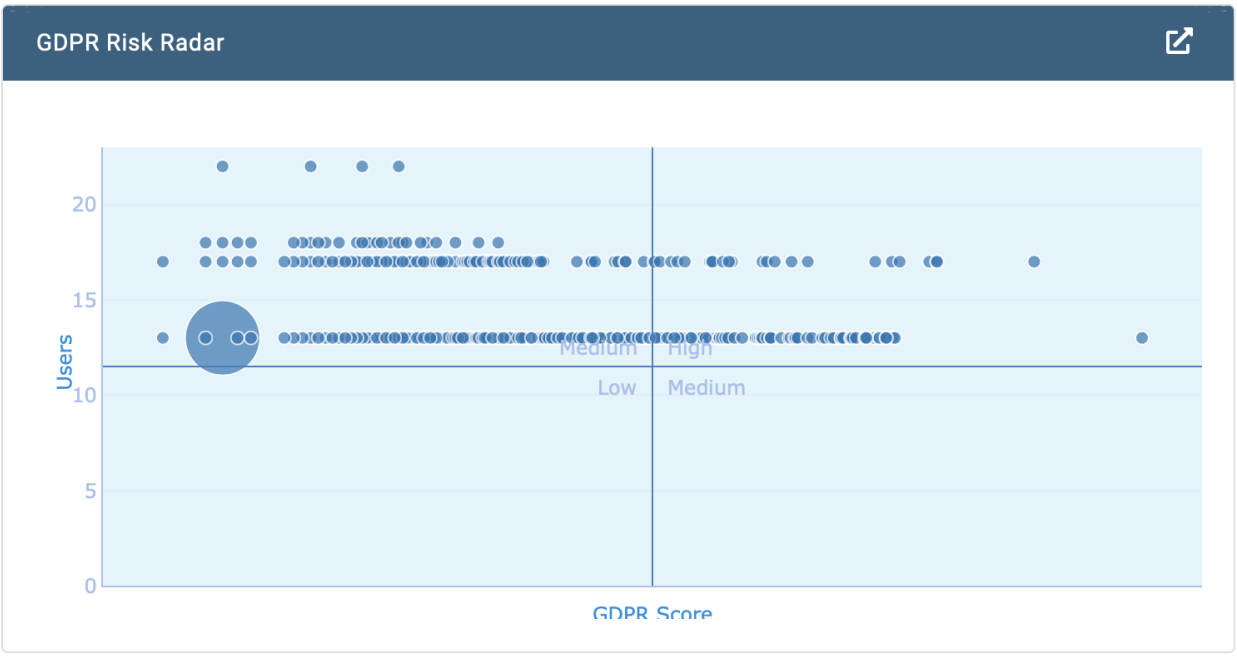
This graph is interactive, meaning that it is possible to click one of the bars and directly see all corresponding files. This also takes the date-range into account.

27.2 Landscape Page

The Landscape page is a powerful tool to get a feel for where all the privacy-sensitive information resides and how many people can access it. It gives the possibility to quickly scope out documents that need action soon, and helps with prioritising the data cleanup effort.



27.2.1 GDPR Risk Radar Graph



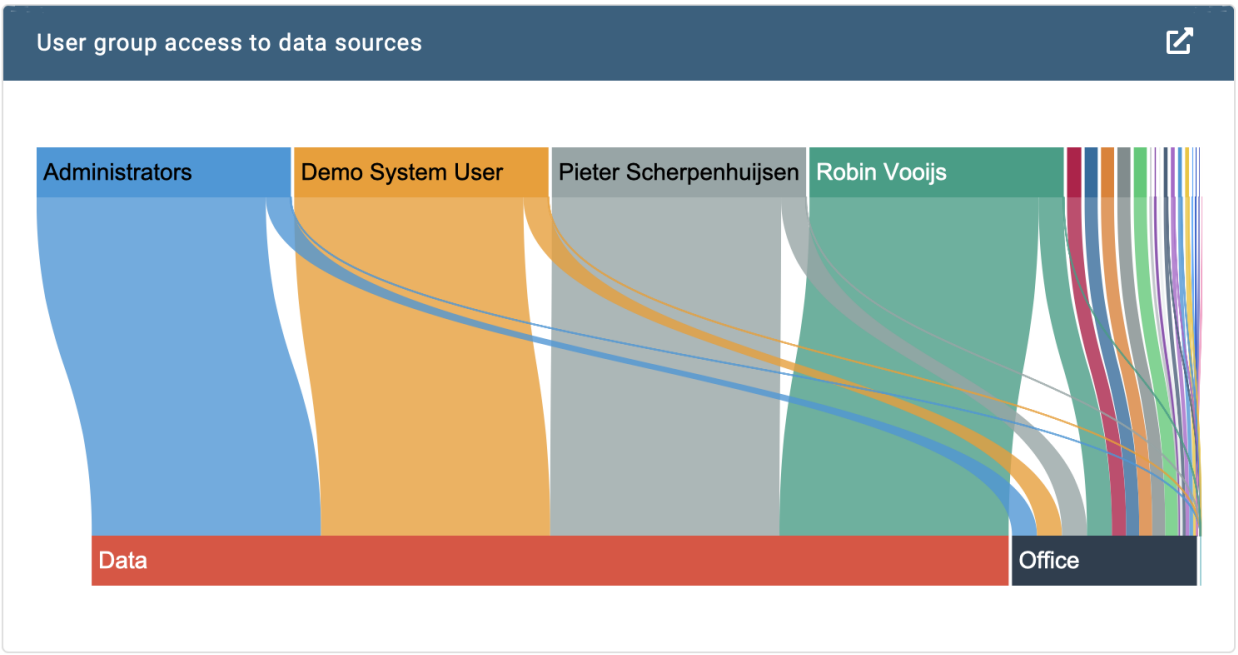
The Risk Radar combines two of the most powerful pieces of information about the indexed data; the found tokens in the data and the number of people who can access the data. This information is displayed in this graph. The Y-axis shows the number of people who have access to the data, and the X-axis shows the GDPR score. The higher the score and the more people can access the data, the higher the risk for such documents is. The document will then be displayed

as a dot in the rop-right corner. The bigger the dot, the more documents fall into that category. During the cleanup process, it is worthwile to focus on the documents in the 'High' quadrant of the graph.

It is also possible to 'pop-out' this graph for a bigger image. This can be done by clicking the icon in the top right corner of the card.

This graph is interactive, meaning that it is possible to click one of the dots and directly see all corresponding files.

27.2.2 GDPR User Group Access To Data Sources Graph

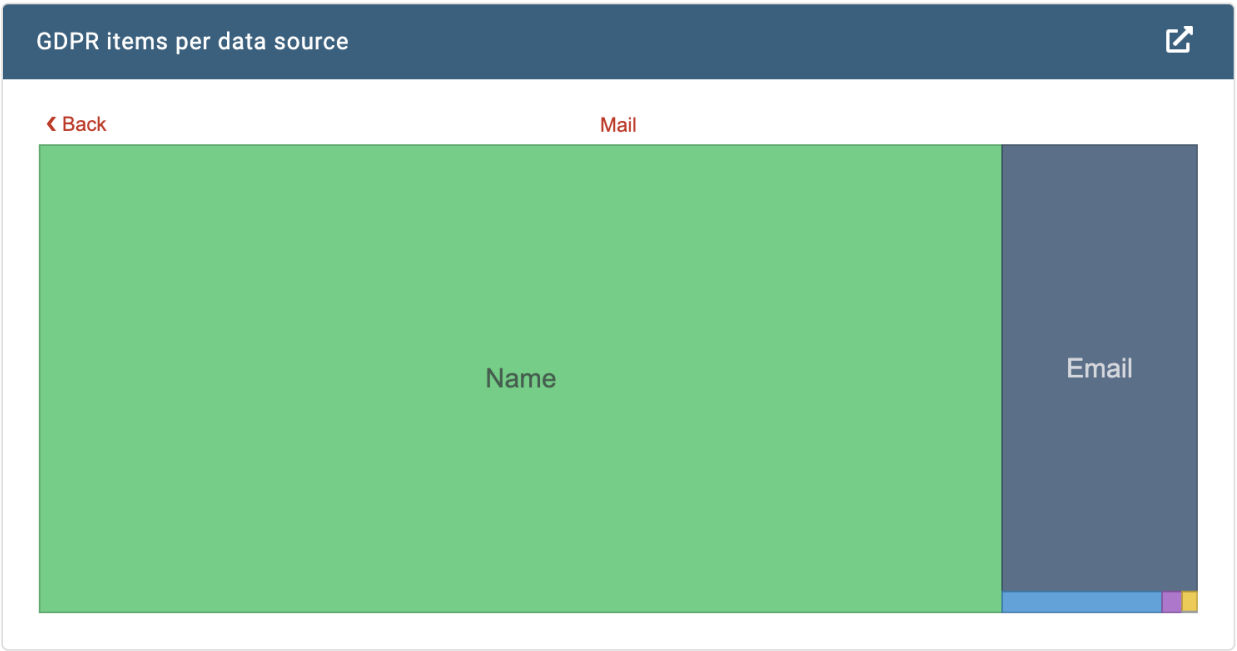


The "User Group Access To Data Sources" graph visualises the users and groups that have access to the connected data sources. The users and groups are displayed on the top, and the data sources are displayed on the bottom. The arrows between the users and datasources visualise the flow (access) of the data. The bigger the arrow, the more documents the user can see in that datasource.

It is also possible to 'pop-out' this graph for a bigger image. This can be done by clicking the icon in the top right corner of the card.

27.2.3 GDPR Items Per Data Source Graph

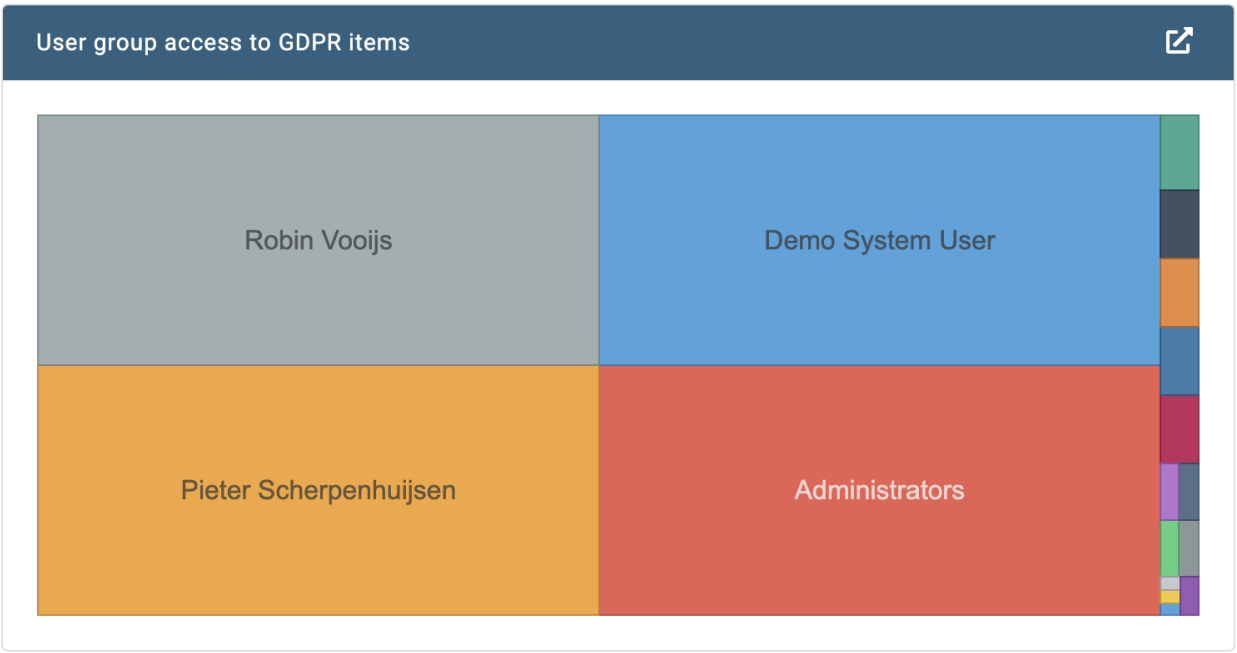
This graph shows the number of GDPR items per data source type. You can click a data source type and see how many GDPR items are from that document type:



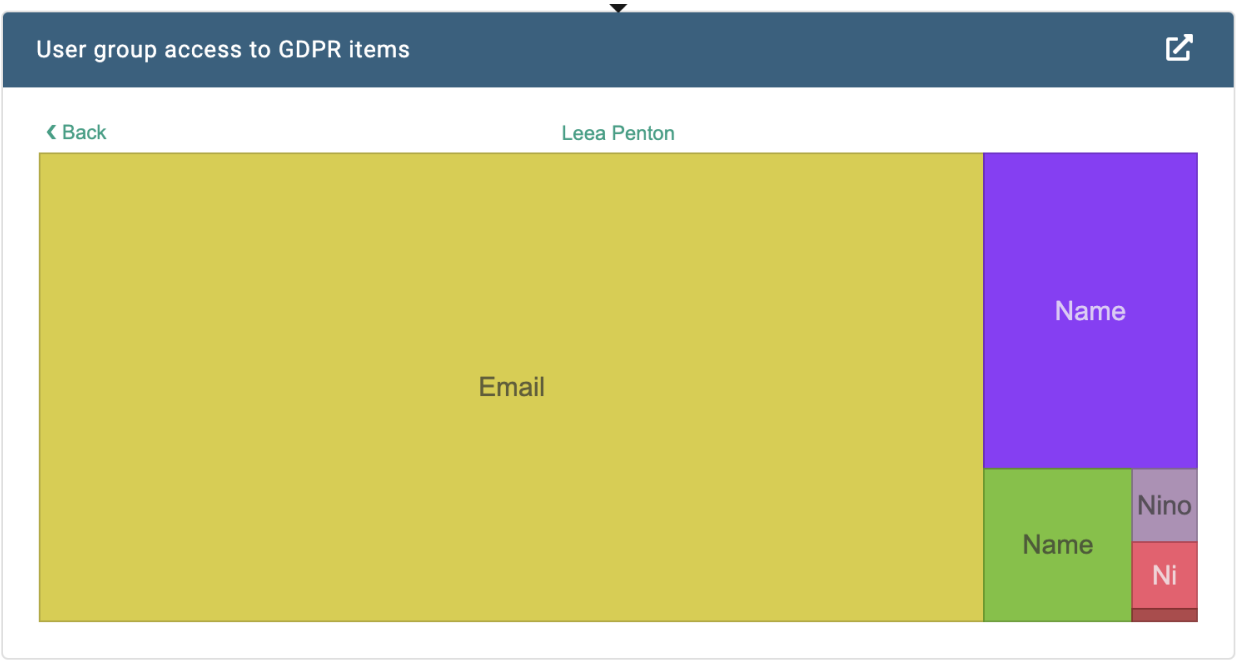
The bigger the part of the graph is for that data source type and GDPR category, the more GDPR items are in that category.

It is also possible to 'pop-out' this graph for a bigger image. This can be done by clicking the icon in the top right corner of the card.

27.2.4 GDPR User Group Access To GDPR Items Graph



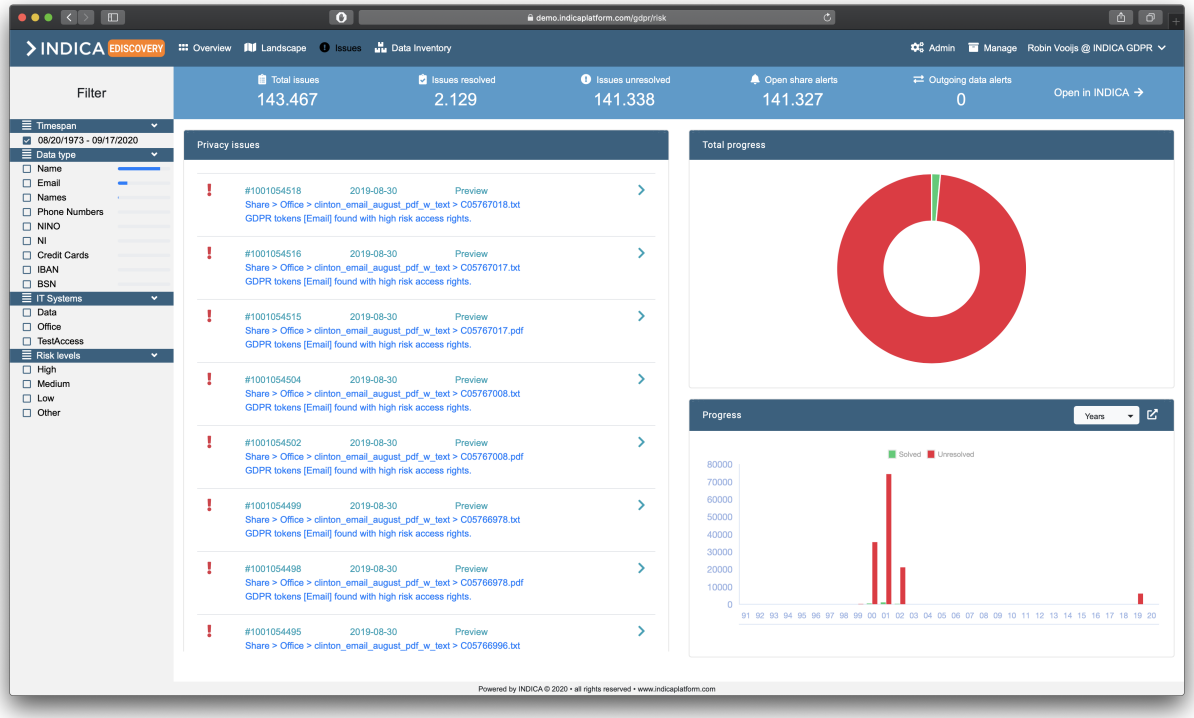
This graph helps to get a feeling for who (which group) can access what tokens. The bigger the user/group is in the graph, the more items this user/group can access. Clicking one of the users/groups reveals how many of what type of GDPR item can be accessed:



It is also possible to 'pop-out' this graph for a bigger image. This can be done by clicking the icon in the top right corner of the card.

27.3 Issues Page

The Issues page sums up all the issues that are found in the indexed data. This page orders them by severity, and helps prioritise the cleanup effort.



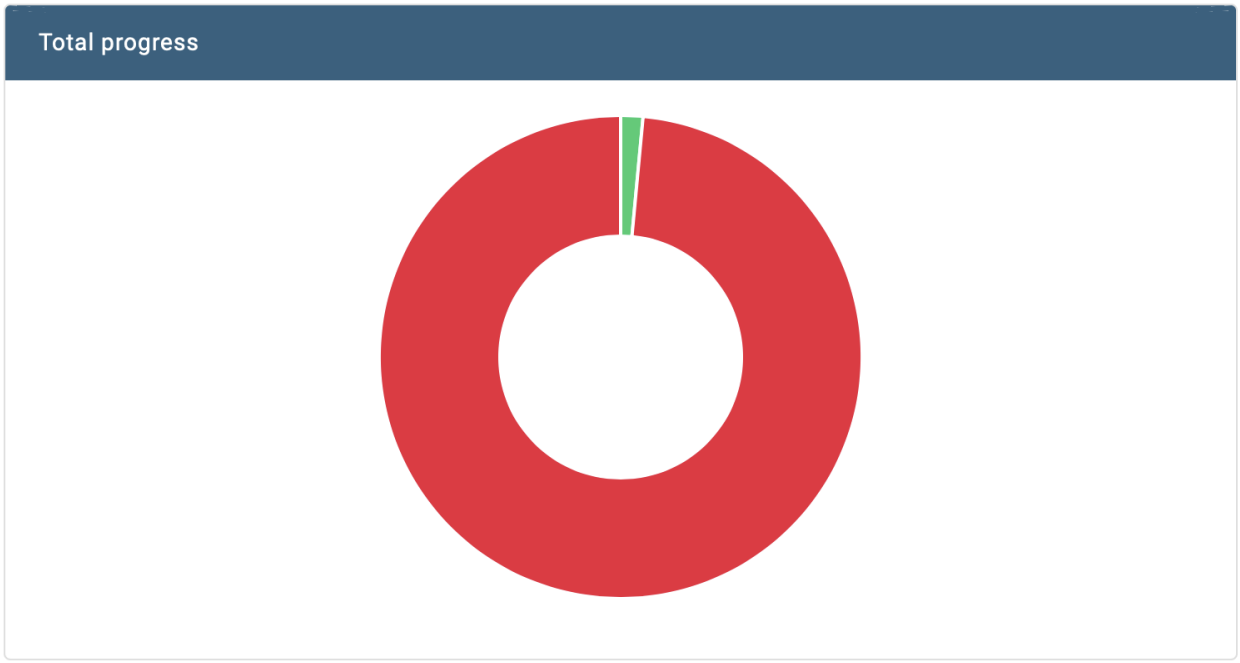
27.3.1 Privacy Issues List

The Issue List helps pick up issues one by one. It gives several pieces of information:

- **Severity**
  Other, Low, Medium, High;
- **Document ID**
  Internal document ID; assigned during indexation;
- **Date**
  Date on which the document was last modified;
- **Preview**
  Hover over this link and a preview opens containing GDPR score, summary, users with access, and the found tokens;
- **Filepath**
  Location of the file in the source data;
- Reason why it is flagged.

Clicking the ">" arrow will open the document in the Search interface, so action can be taken immediately.

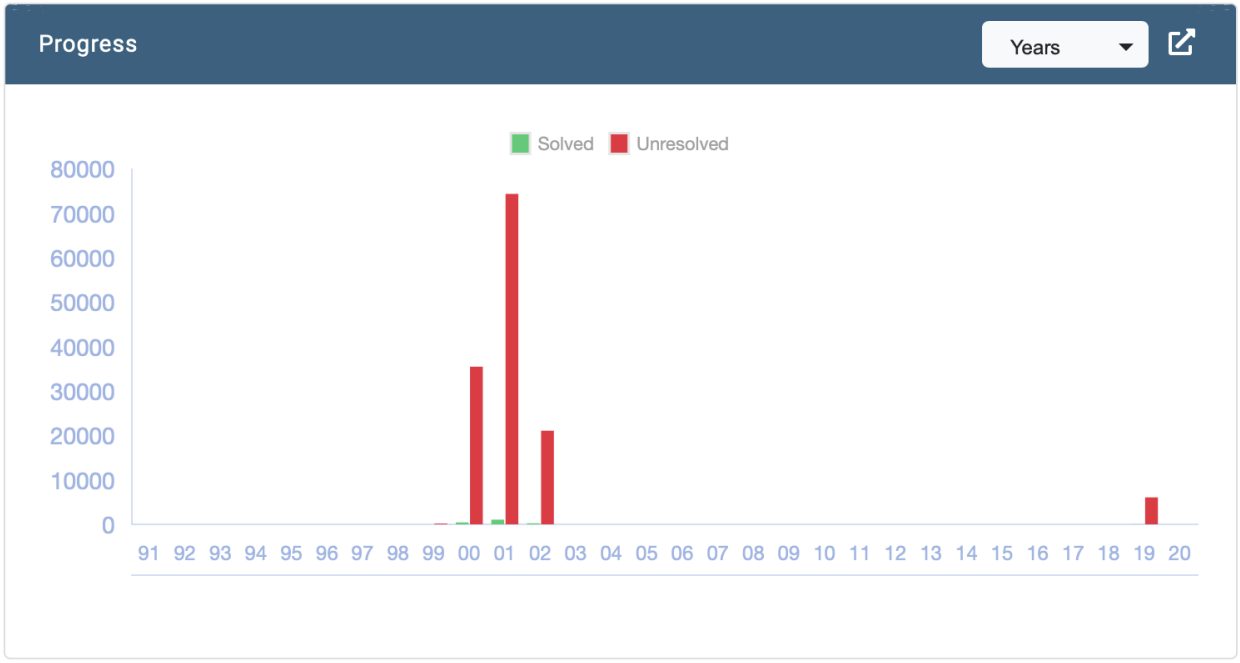
27.3.2 Total Progress Graph



This graph gives insight in the total progress that has been made with the data cleanup efforts. The bigger the green part, the more issues are resolved.

Hovering over one of the parts of the graph will reveal the counts.

27.3.3 Progress over time Graph



This graph shows you the progress that has been made with the cleanup of issues. When more work is done, the red bars become smaller and the green bars will grow until evetually all issues are resolved. This graph can be switched between 'Years' and 'Months' for a more detailed view. It is also possible to 'pop-out' this graph for a bigger image. This can be done by clicking the icon in the top right corner of the card.

This graph is interactive, meaning that it is possible to click one of the bars and directly see all corresponding files. This also takes the date-range into account.

27.4 Data Inventory Page

The Data Inventory Page shows you an overview of records found within a business rule. These business rules are added by case managers in the GDPR settings of the case.

Let's say that for example, the Customer Service Department of your company isn't allowed to have sensitive information of their customers after an x amount of time. This might be a rule added by your case manager. The amount of records (data) that fall within that category is shown here. If you click on the eye-icon, you'll be taken to the INDICA search page with these records as search results.

ID	94	95	96	97	98	99	100	102
CONTROLLER	Internal audit	Internal audit	Internal audit	Internal audit	Internal audit	Internal audit	Internal audit	
BUSINESS PROCESS	Client intake	Client intake	Client intake	Client administration	Client billing	Client billing	Client admin	Mailings
PERSONAL DATA TYPES	NAME	EMAIL	TEL	NINO	NI	CC	IBAN	EMAIL
PROCESSING ACTIVITY	Contacting	Contacting	Contacting	Billing	Billing	Billing	Client admin	Product Updates
CATEGORIES OF DATA SUBJECTS	Customers	UNKNOWN	Customers	Customers	Customers	Customers	Customers	Customers
CATEGORIES OF PERSONAL DATA								
IT ASSET	Client CRM	Client CRM	Client CRM	Client CRM	Client CRM	Client CRM	Client CRM	Client CRM
PURPOSES OF PROCESSING								
LEGAL GROUND		1.e						1.d
PROCESSOR								
INTERNATIONAL TRANSFER	0	0	0	0	0	0	0	0
RETENTION TERM	0	0	0	0	0	0	0	12
DATA TYPE CLASSIFICATION								
TECHNICAL MEASURES	Authentication	Authentication	Authentication	Authentication	Authentication	Authentication / encryption	Authentication	
ORGANISATIONAL MEASURES								
RECORDS	1187	0	142	132	101	17	3	0
OPEN								

CHAPTER

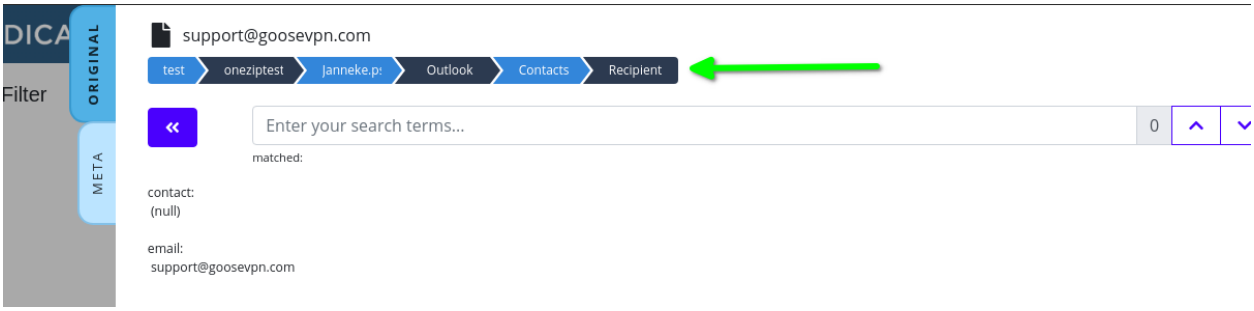
TWENTYEIGHT

PREVIEWER

Here you can find many of the tools that INDICA has to offer. Main purpose is to audit the documents.

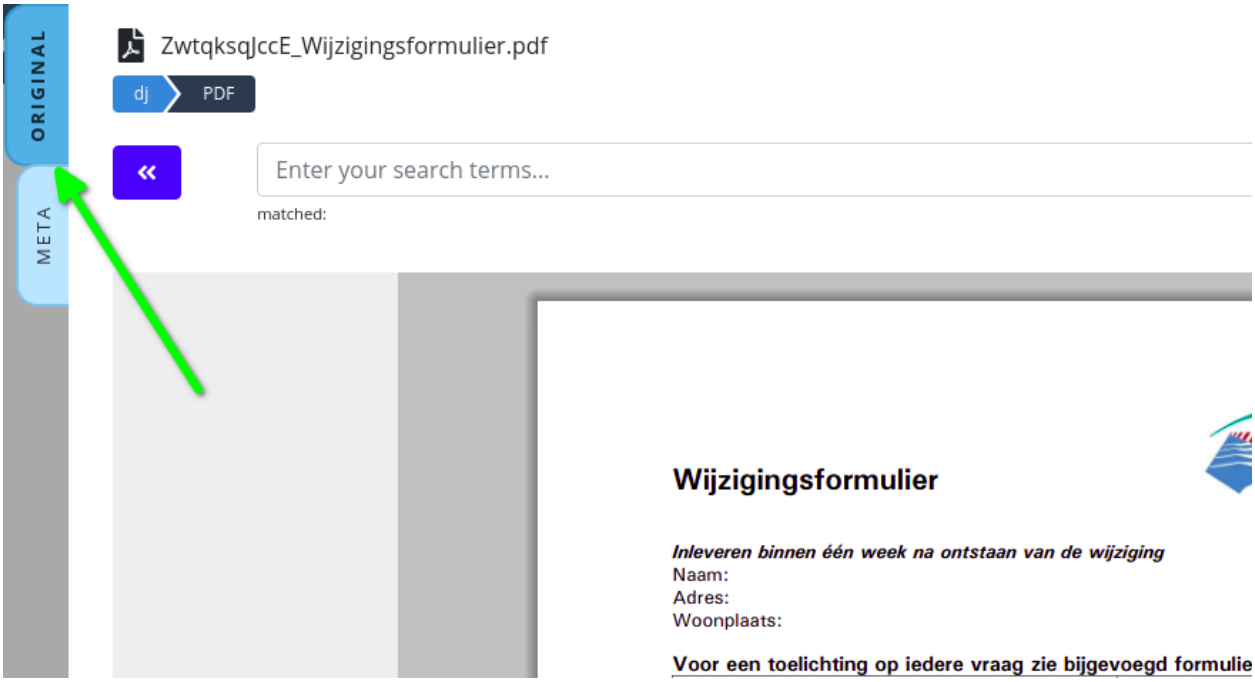
28.1 Breadcrumbs

Breadcrumbs show the system path of the source document is located. You can also click on breadcrumbs to create a filter query of the specific path.



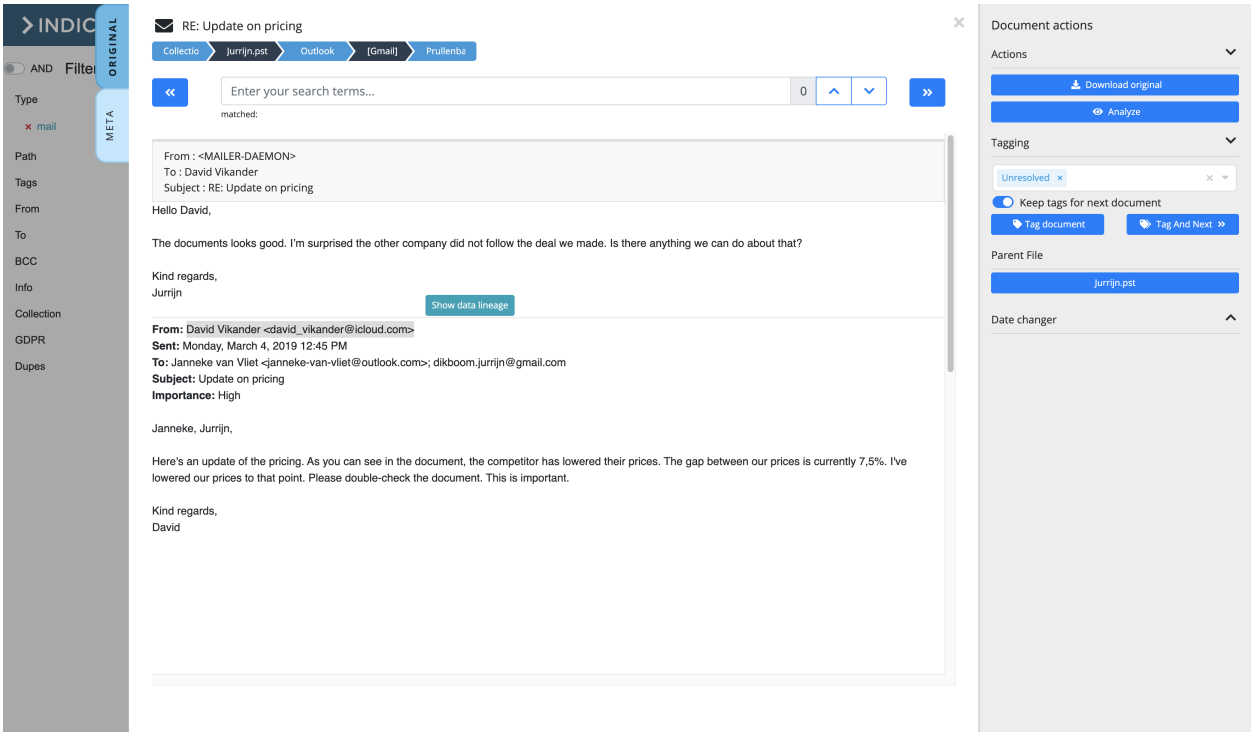
28.2 Switching Views

Depending on the document type preview has multiple ways of displaying different content. To switch between these views you have to click on one of the sidebar buttons in the previewer sidebar.



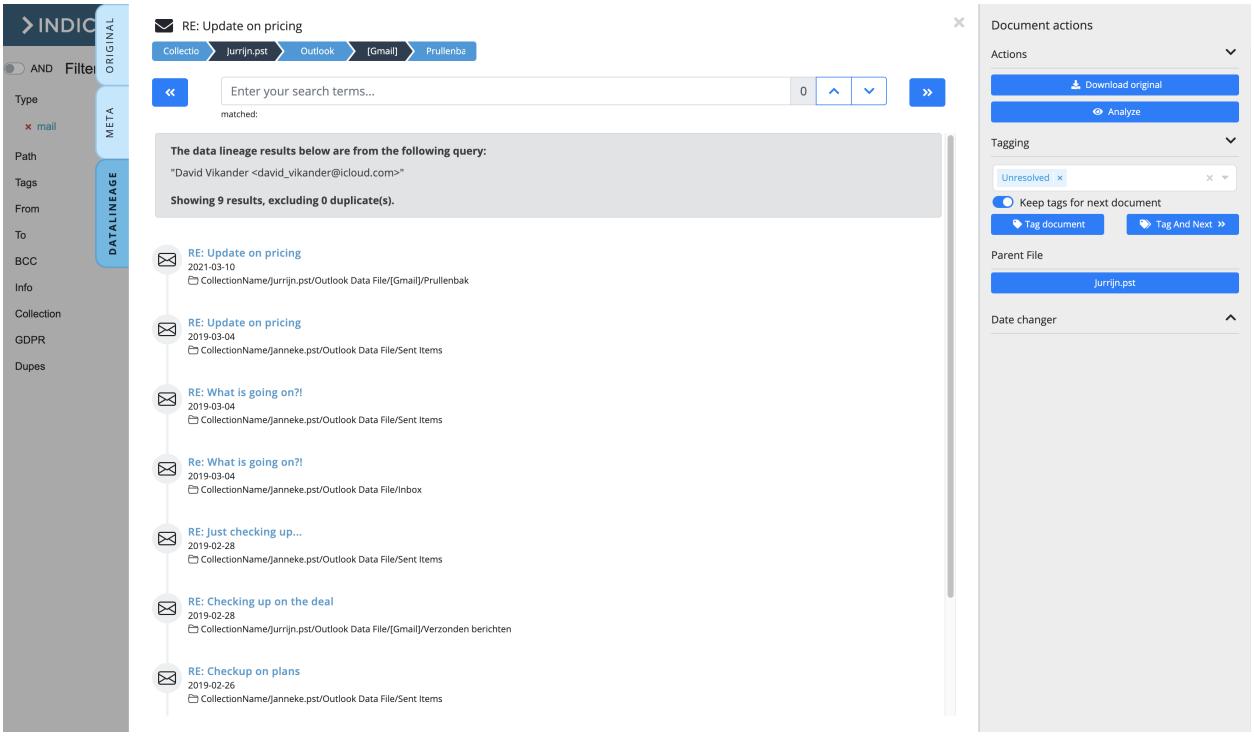
28.3 Data Lineage (Tab)

A new feature in INDICA is the possibility to find the "single source of truth" when dealing with duplicate information. When this feature is enabled in the manage interface you can select a part of text within documents. A button "Show data lineage" will show up, see example below:



Clicking on this button will open a new tab in the previewer called "Data Lineage". This will give an overview of the

query results for the previously selected text. Within this view you can click on the titles of documents to open them in another browser tab.



Note: Due to the way iframes work the button will show up at the top of the preview. This means that selecting text in (for example) some e-mails the button will show up at the top of the preview.

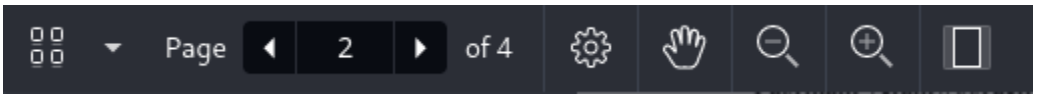
28.4 PDF Controls

When you open a PDF some extra control option will appear.

1) Allows to collapse and hide thumbnails on the left side
2) Lets you rotate the document clockwise or counterclockwise
3) Lets you zoom in/out
4) Resets the zoom level

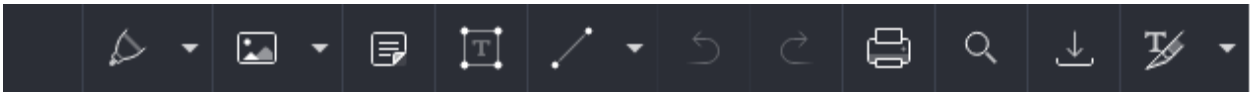
28.4.1 Advanced PDF Viewer

Note: The Advanced PDF Viewer is not available on all INDICA appliances.



From left to right:

1. Lets you view thumbnails of the document. You can also view the outline, annotations and bookmarks of the document.
2. Lets you navigate the pages of the document.
3. Lets you customise the document's page layout.
4. Lets you enter Pan mode.
5. Lets you zoom in and out of the document.
6. Fits the document to the width of the previewer.



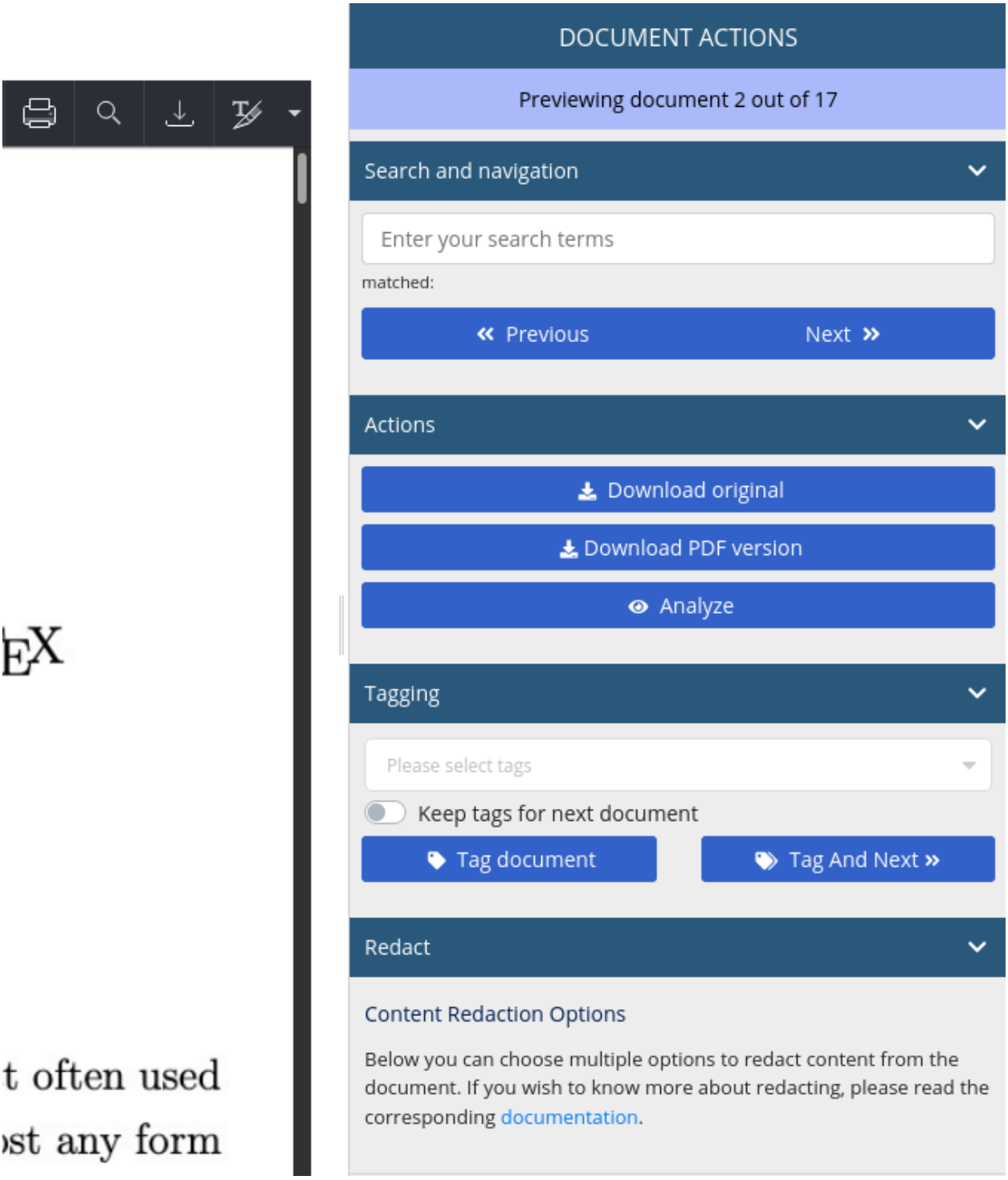
From left to right:

1. Drawing on top of the document, including highlighting.
2. Lets you insert images and some pre-set stamps.

3. Add notes to the document.
4. Add a text box to the document.
5. Draw lines on the document.
6. Un-do and re-do changes.
7. Print the document.
8. Search inside the document.
9. Export the document as a PDF.
10. Redaction tool. You can redact manually highlighted text or get a free-form redaction box to select an area.

28.5 Document Actions

Documents will have different actions you can take, each of them will be described in detail below.



1. “Search and Navigation” - search through the document content, navigate to the next and previous documents.
2. “Actions” - download the original, non-redacted version of the document, download the PDF version of the file, restore the original file and analyze the document.
3. “Redact” - redaction tools for files. Read more: *Redacting*.
4. “Delete Pages” - ability to delete pages from your file. Requires the advanced PDF viewer. Read more: *Page Deletion (Advanced PDF Viewer)*.

5. “Date changer” - allows to change the date of the document. This can cause the document to appear in a different place in search results.
6. “Tagging” - Allows to tag individual documents. Read more: Tagging individual items.
7. “Attachments” - if the document is an e-mail, it may have attachments. The link to those attachments is provided in this section.
8. “Parent” - if the document is an e-mail attachment or part of an archive, the link to the parent is provided in this section.

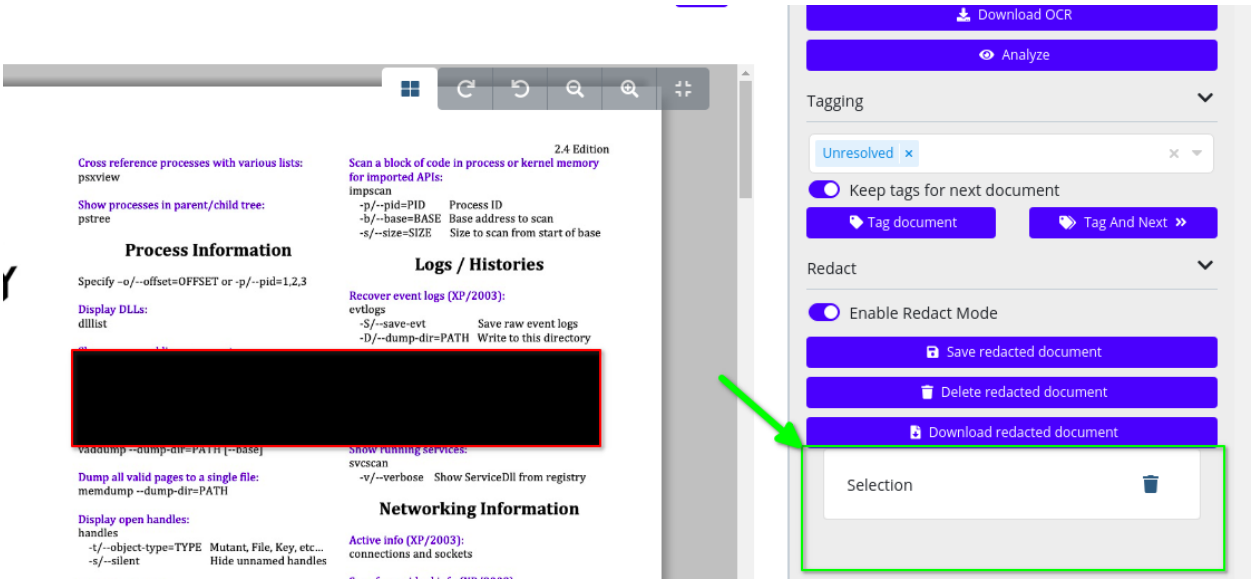
28.5.1 Advanced PDF Viewer

Note: The Advanced PDF Viewer is not available on all INDICA appliances.

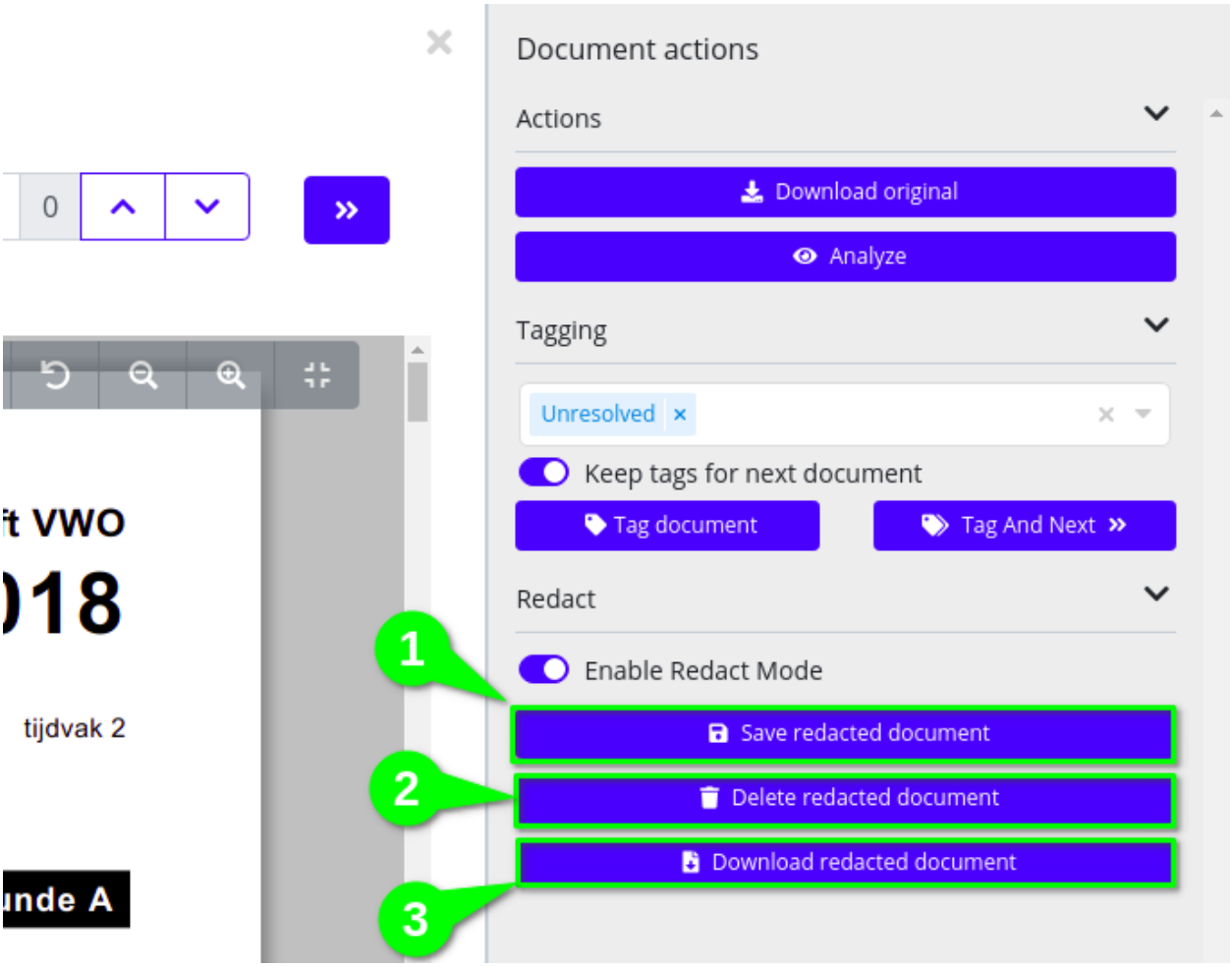
The document actions for the previewer stay the same, for the exception of the redaction tools. The search function is also moved to the document actions. Read more about the redaction tools here: *Redacting (Advanced PDF Viewer)*.

28.6 Redacting

Redacting allows to hide parts of the document and download the redacted version.



You can start redacting by first enabling Redact Mode with “Enable Redact Mode” switch. Once Redacting is enabled you can drag your cursor to mark areas to hide. All the newly created selection will show up below the action buttons. You are able to delete them if you click the “recycling bin” icon.



Once redacting is done you can take few actions:

1) You can save the document, so when the document is opened next time you also see the redacted version.

Note: Original non redacted version of the document is still kept and available for preview. When a redacted version is available, you will be shown the redacted version.

---

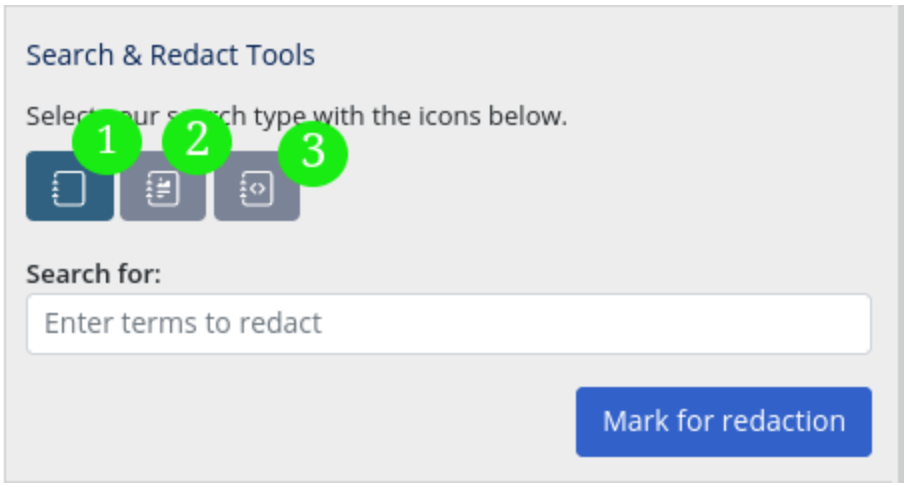
2) You can delete the redacted version of the document.

3) You can download the redacted version of the document.

28.7 Redacting (Advanced PDF Viewer)

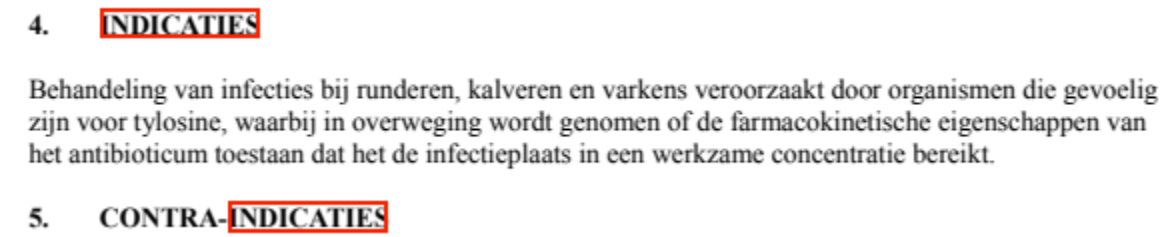
Note: The Advanced PDF Viewer is not available on all INDICA appliances.

In the advanced PDF viewer, there are multiple redaction options available. They are listed below.

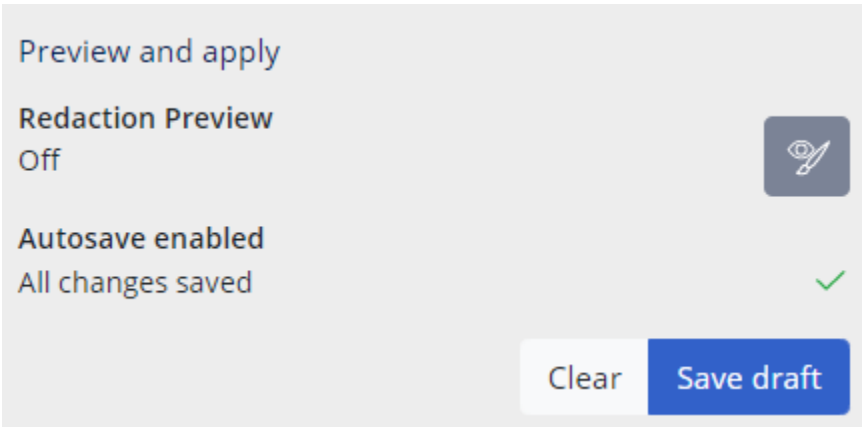


1) Text - You can enter terms to search for in the document. If found, they will be marked for redaction.
2) Preset patterns - You can select preset patterns to mark for redaction in the document, e.g. IBAN, telephone numbers, and BSN numbers.
3) Custom regex - You can enter your own regex pattern. If matches are found, they will be marked for redaction.

When words are marked for redaction, they are outlined with a red box, shown below.



From here, there are multiple options.



You can preview your new redactions at any time during redaction. Any redactions marked with a red box will fill. You can preview your redactions by clicking the button with the “eye” icon.

4. [redacted]

Behandeling van infecties bij runderen, kalveren en varkens veroorzaakt door organismen die gevoelig zijn voor tylosine, waarbij in overweging wordt genomen of de farmacokinetische eigenschappen van het antibioticum toestaan dat het de infectieplaats in een werkzame concentratie bereikt.

Saving a draft will not apply the redactions to the document, but saves what has been marked for redaction. On refresh, or upon closing the document and opening it again, you will see the words marked previously for redaction.

When the "eye" button is pressed, there is an option to apply the redactions to the document. Once the redactions are applied, you are not able to remove them or change them individually.

After applying redactions, the page will refreshed automatically. The changes are saved on the original tab and a concept tab is added to the previewer. On this new tab you can add or alter redactions.

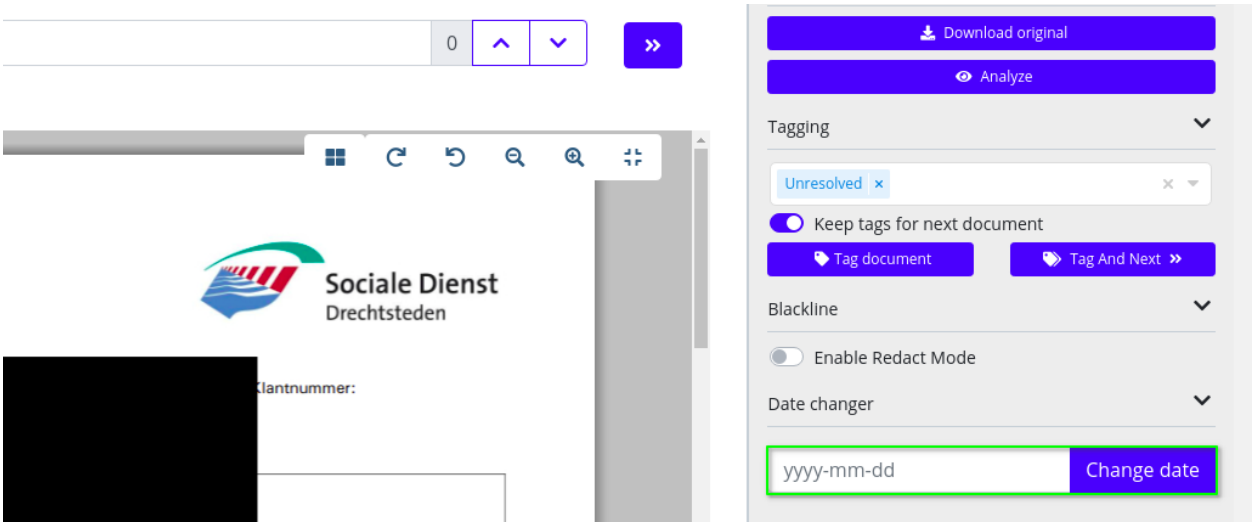
You can remove all the redactions at once by pressing the "restore original preview" button. This will remove all the saved redactions and place the original document back into the viewer.

Autosave drafts Case managers can enable this function. When autosave is enabled, the changes on drafts are saved automatically. A confirmation message appears underneath 'Autosave enabled'. When autosave is disabled, you will have to save manually. After saving manually, a confirmation message appears at the top of the page.

Note: The autosave function will only save drafts. Applying redactions permanently to the document still needs to be done manually by clicking "Apply redactions".

28.8 Date-changer

Allows changing creation date for any of the documents. This is especially useful when a document has been digitalised, and has lost its original creation date.



28.9 Page Deletion (Advanced PDF Viewer)

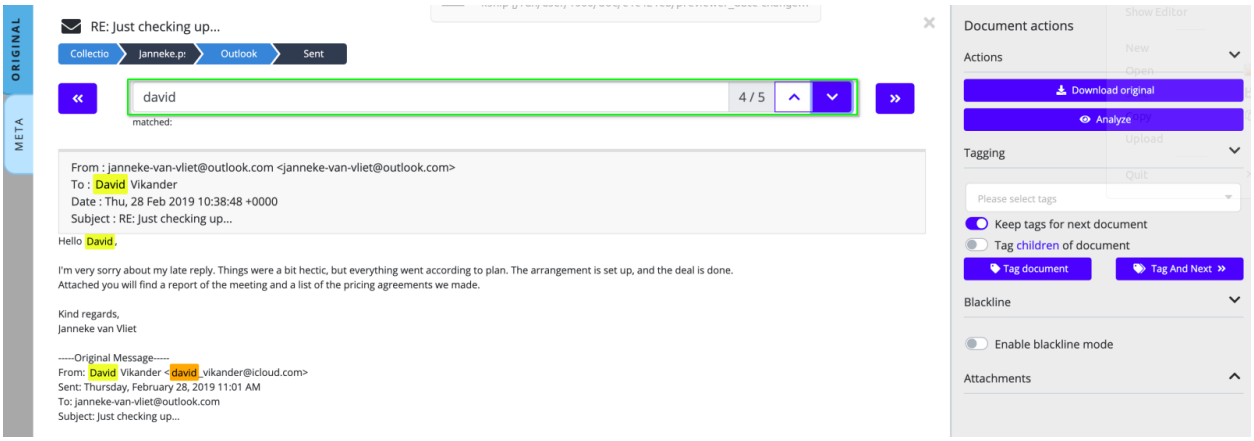
Note: The Advanced PDF Viewer is not available on all INDICA appliances.

The user can choose to delete pages from a specific document in INDICA. This, like redaction, is reversible – however, one cannot reverse only page deletion or only redaction , it is done at the same time. Only a range of pages can be deleted. If a document has concepts on a page that is deleted, the concepts are deleted as well. If a redacted and a concept version exists, deletion can only happen on the concept version, and the redacted version does not update unless the concepts are applied to the document

28.10 Searching in documents

It's also possible to search within the preview of the document. Searching in a document will highlight the keywords within the document.

The highlighted keywords will be yellow. The current highlighted keywords, which starts with the first match, will be orange. Clicking on the up- and down buttons next to the search bar will bring you to the next highlighted keyword. See image below for an example.



CHAPTER

TWENTYNINE

QUERYING

Table of Contents

- *Querying*
  - *Writing queries*
    - *Search By*
    - *Exact match*
    - *Terms and Phrases*
    - *Boolean Operators*
    - *Grouping*
    - *Range Searches*
    - *Wildcards*
    - *Fuzzy Searches*
    - *Word Distance Searches*
    - *Boosting terms*
  - *Advanced query builder*
  - *Escaping special characters*

INDICA allows you easily create powerful queries. The creation of queries can be done in multiple ways:

- Writing queries by hand
- Using the advanced query builder

29.1 Writing queries

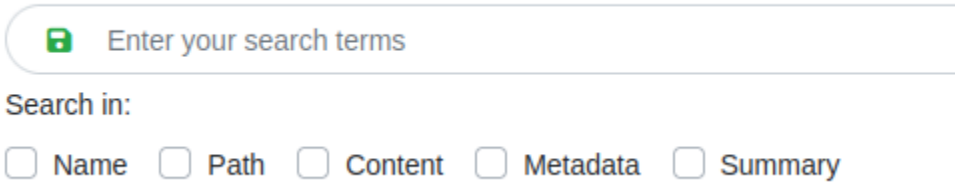
INDICA provides a rich query language, giving its users a flexible and advanced way to search through large amounts of documents.

The search bar allows multiple ways of querying. It is possible to type in search terms and manipulate the way the query interpreter deals with the query.

INDICA allows the following query manipulations:

29.1.1 Search By

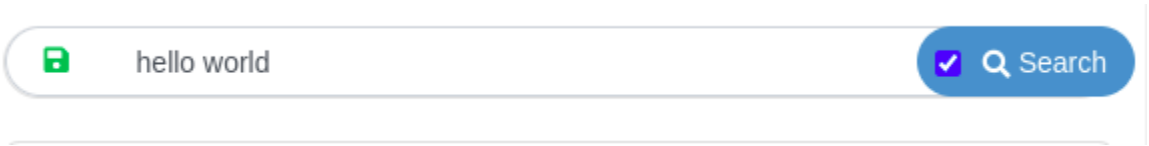
Users can search for keywords inside the name, path, content, metadata or summary of the document.



29.1.2 Exact match

Exact match checkbox allows for quick term or phrase searches. It only supports simple text and Filters, no query modifiers such as Boolean Operators, Fuzzy Searches, Wildcards, etc. are supported . After the checkbox is ticked, whole text will be treated as a single search term. Any advanced terms will be ignored!

See example below



29.1.3 Terms and Phrases

A query is broken up into terms and operators. There are two types of terms: Single Terms and Phrases.

A single term is a single word such as `test` or `hello`.

A phrase is a group of words surrounded by double quotes such as `"hello test"`.

29.1.4 Boolean Operators

Boolean operators allow terms to be combined through logic operators. INDICA supports the following operators:

- **AND**
  The `AND` operator matches documents where both terms exist anywhere in the text of a single document. The symbol `&&` can be used in place of the word `AND`.
  To search for documents that contain “jakarta apache” and “Indica search” use the query `"jakarta apache" AND "INDICA search"`
- **“+”**
  The + operator requires that the term after the + symbol exist somewhere in a field of a single document.
  To search for documents that must contain “jakarta” and may contain “INDICA” use the query `+jakarta INDICA`
- **OR**
  The `OR` operator is the default conjunction operator.
  This means that if there is no boolean operator between two terms, the `OR` operator is used.
  The `OR` operator links two terms and finds a matching document if either of the terms exist in a document.

The symbol `||` can be used in place of the word `OR`.

To search for documents that contain either "jakarta apache" or just "jakarta" use the query `"jakarta apache" jakarta` or `"jakarta apache" OR jakarta`.

- **NOT**

  The `NOT` operator excludes documents that contain the term after `NOT`.

  The symbol `!` can be used in place of the word `NOT`.

  To search for documents that contain "jakarta apache" but not "INDICA search" use the query `"jakarta apache" NOT "INDICA search"`

  > **Note:** The NOT operator cannot be used with just one term. For example, the following search will return no results: `NOT "jakarta apache"`

- **"-"**

  The `-` operator excludes documents that contain the term after the `-` symbol.

  To search for documents that contain "jakarta apache" but not "INDICA search" use the query `"jakarta apache" -"INDICA search"`

Important: Boolean operators **must be** ALL CAPS

29.1.5 Grouping

INDICA supports using parentheses to group clauses to form sub queries. This can be very useful if you want to control boolean logic for a query. To search for either "jakarta" or "apache" and "website" use the query `(jakarta OR apache) AND website`. This eliminates any confusion and makes sure you that "website" must exist and either term "jakarta" or "apache" may exist.

29.1.6 Range Searches

Range Queries allow one to match documents whose field(s) values are between the lower and upper bound specified by the Range Query. Range Queries can be inclusive or exclusive of the upper and lower bounds. Sorting is done lexicographically.

Example: `date:[2002-01-01 TO 2003-01-01]`

This will find documents whose date fields have values between 2002-01-01 and 2003-01-01, inclusive.

To perform a Range Query and exclude the upper and lower bound, use `{` and `}` instead of `[` and `]`. The inclusive and exclusive range characters can also be mixed:

Example: `date:{2002-01-01 TO 2003-01-01]`

29.1.7 Wildcards

Wildcard searches can be done with two operators:

- **"?"**

  A question mark will allow different characters at that specific position in the word.
  Example: `?est`
  INDICA will return words like "rest", "test", "nest"

- **"*"**

  An asterisk places after (part of) the word, will enforce the presence of that (part of) the word in the results and then expand the word.
  Example: `test*`
  INDICA will return words like "testing", "testresult", "testcase"

29.1.8 Fuzzy Searches

A 'tilde' (~) after the term will imply that the term might be spelled in different ways. The similarity can be changed by adding a value between 0 and 1. A value closer to 1 will return words that are very similar. The value 0.5 cannot be used.

29.1.9 Word Distance Searches

Distance searches use the 'tilde' ~ as well. It can be used in the following format:

```
"testing Europe"~10
```

This will search for the two terms "testing" and "Europe" with a maximum of 10 words in between. It will return results like:

```
Testing some pieces advanced software is done routinely in Europe
```

and

```
The developer is doing some testing in Europe
```

29.1.10 Boosting terms

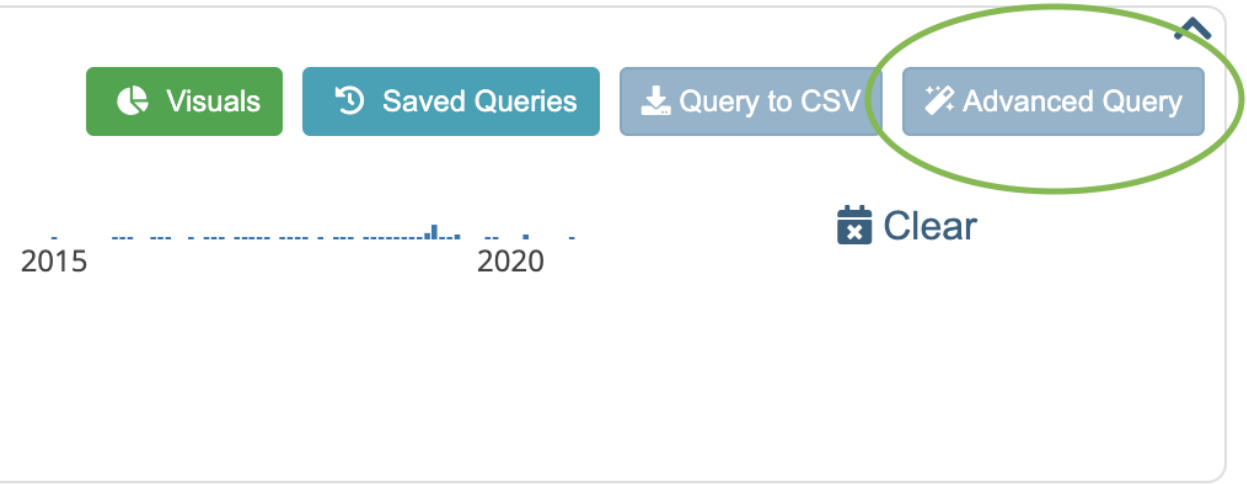
INDICA provides the relevance level of matching documents based on the terms found. To boost a term use the caret `^` symbol with a boost factor (a number) at the end of the term you are searching. The higher the boost factor, the more relevant the term will be.

Boosting allows you to control the relevance of a document by boosting its term. For example, if you are searching for `jakarta apache` and you want the term "jakarta" to be more relevant, boost it using the `^` symbol along with the boost factor next to the term. You would type: `jakarta^4 apache`. This will make documents with the term jakarta appear more relevant.

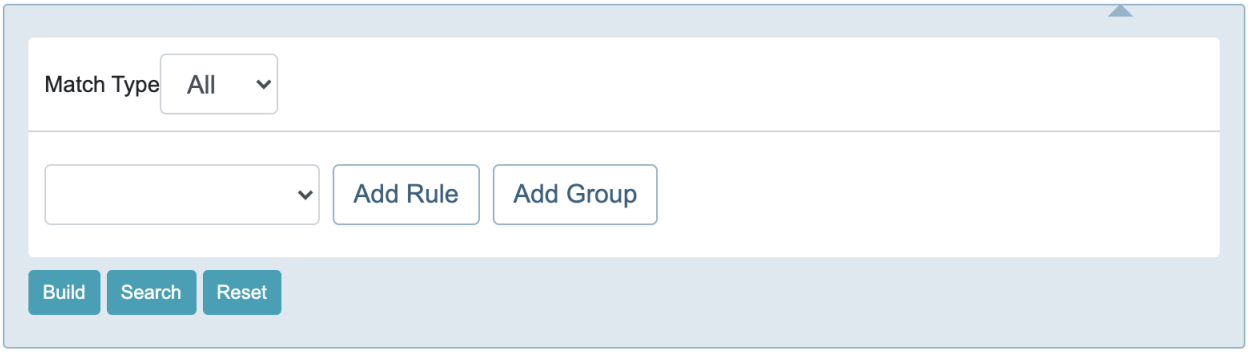
You can also boost Phrase Terms as in the example `"jakarta apache"^4 "Indica search"`. By default, the boost factor for each term or phrase is 1. Although the boost factor must be positive, it can be less than 1 (e.g. 0.2)

29.2 Advanced query builder

The advanced query builder is a visual tool to create powerful queries. The tool can be found when clicking the "Advanced Query" button next to the "Query to CSV" button.



A new panel will open:



A few options can be filled in here:

- **Match Type**
  The "Match Type" determines that either "All" or "Any" condition of the query needs to be true. Either all conditions in the query need to be true, or any condition needs to be true.
- **Add rule**

  It's possible to filter on a few different items:

  - Type
  - Path
  - Tags
  - Info
  - Collection
  - Workflow
  - GDPR

  - Dupes
  - Size
  - Document ID
  - Content
  - File name
  - Date (yyyymmdd)

Select a filter within the input field left of the *Add Rule* button which you wish to query on, click on add rule to add this to the query.

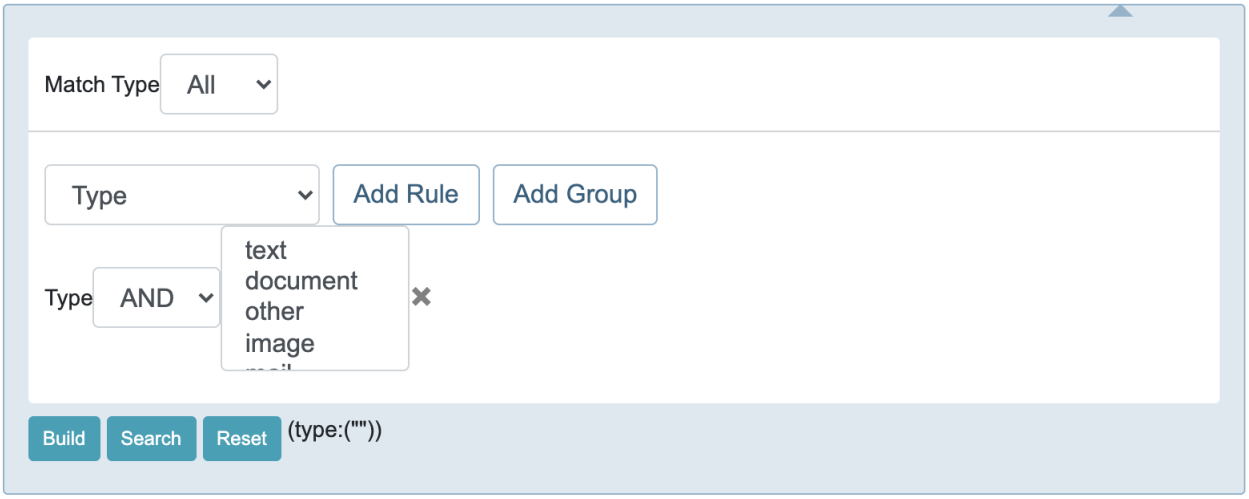
The rule will be added underneath your latest rule. Here you can also choose per rule if you wish to add this as an 'AND' or 'OR' query. Clicking on the 'X' next to the item will delete the rule.

- **Add group**
  The *Add Group* button is only available at the top of the Advanced Query Builder next to the *Add Rule* button. This will add a new group underneath the existing group. Groups can be visually distinguished by the new 'Match Type' filter and the horizontal line underneath it.
- **Build**
  The query can always be previewed next to the buttons at the bottom. With the *Build* button however, you're able to build the query in the search bar at the top. Please be aware that this building is necessary if you wish to save the query!
- **Search**
  Execute the search with the previewed query.

  Note: The query needs to be build in order to save it correctly as a saved query!

- **Reset**
  Reset the whole query. This will empty both the search input field, search results and the advanced query preview.

Use the drop-down menu to select a field. Click "Add Rule" to add a new rule for that field.



In case multiple options are available for the selected field, a list will be shown. Use CTRL+click to select multiple options. A choice can also be made between "All" to query on all the selected conditions, or "OR" to query on any of

the selected conditions.

29.3 Escaping special characters

In some cases, the search term may use one of the Query Engine's special characters. If that's the case, the character needs to be 'escaped' so the engine will interpret the character as a literal character.

The following characters are special characters:

```
+ - && || ! ( ) { } [ ] ^ " ~ * ? : \
```

To escape these character place a backslash ("\") before the special character.

CHAPTER

THIRTY

QUERY TO CSV

Table of Contents

- *Query to CSV*
  - *Exporting to CSV from the front-end*
  - *Exporting to CSV and XLS(X)*

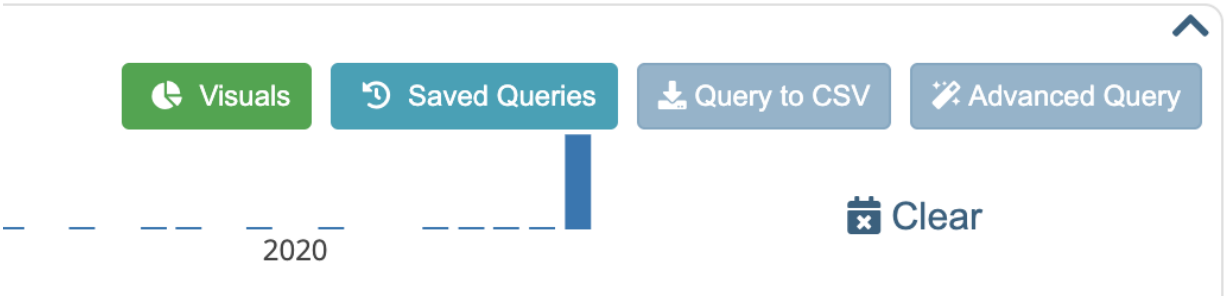
INDICA has a really powerful feature to export query results to CSV.

This can be done in two ways:

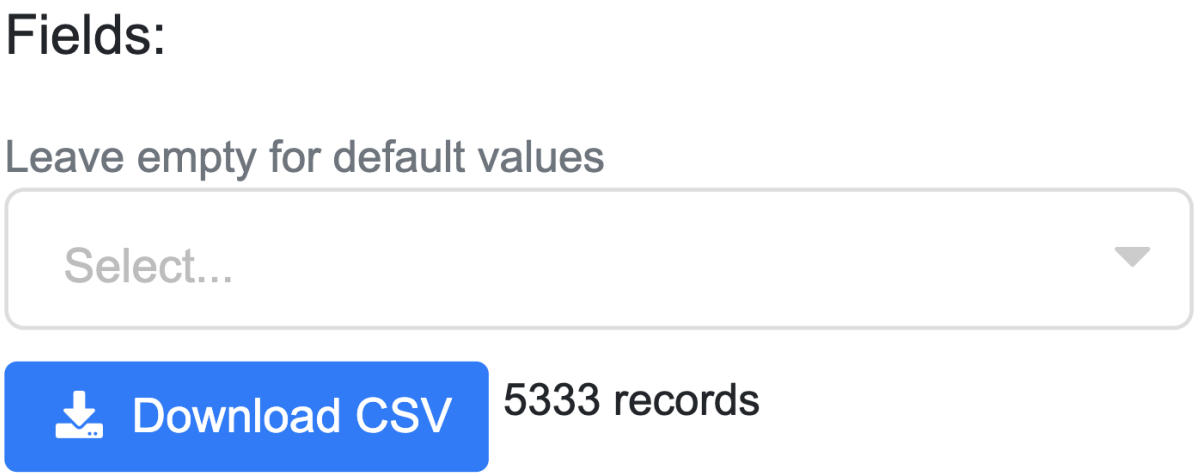
- Export to CSV directly from the front-end based on the active query;
- Export to CSV and XLS(X) on the management page.

30.1 Exporting to CSV from the front-end

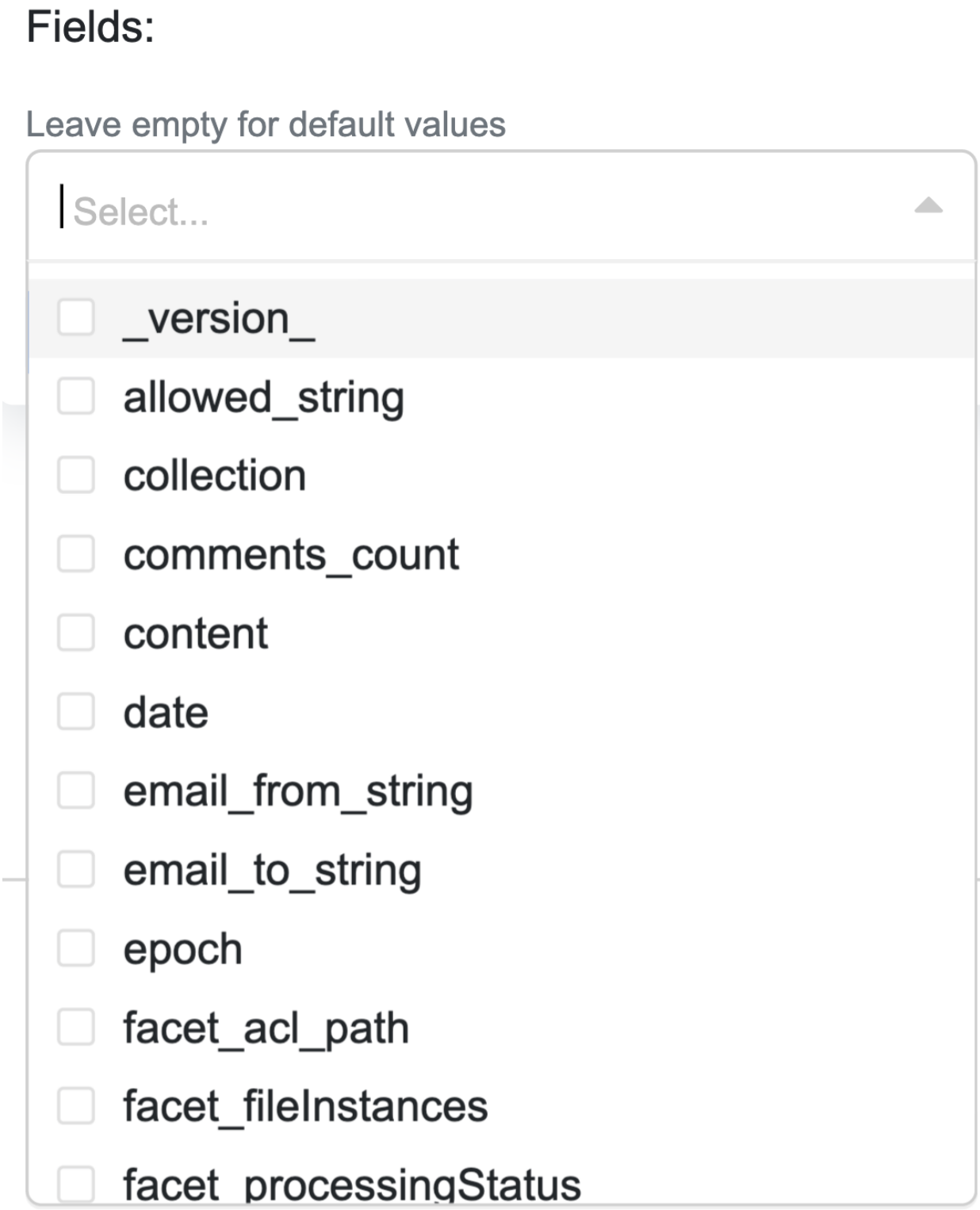
On the front-end, there's a button called "Query to CSV".



When clicking this button, a box will open:



Here, you can select the fields that you would like to export:



Simply select the fields that you want to export. It is also possible to type in the first few characters of the field name to filter the list. When the desired fields are selected, click the "Download CSV" button.

INDICA will now begin gathering and ordering the required fields of all the files. After that, a CSV will be created on

disk and downloaded.

Attention:

- Creating large exports (>50.000 files) can be really slow and may fail.
- Creating and downloading larger exports may be done on the management page.
- For system administrators, it is also possible to create exports via the command-line interface.

30.2 Exporting to CSV and XLS(X)

This can be done by case managers. See *Manager Process Settings* for more information.

CHAPTER

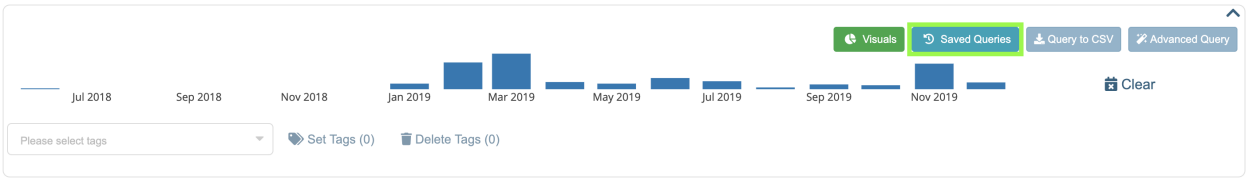
THIRTYONE

SAVED QUERIES

Table of Contents

- *Saved Queries*

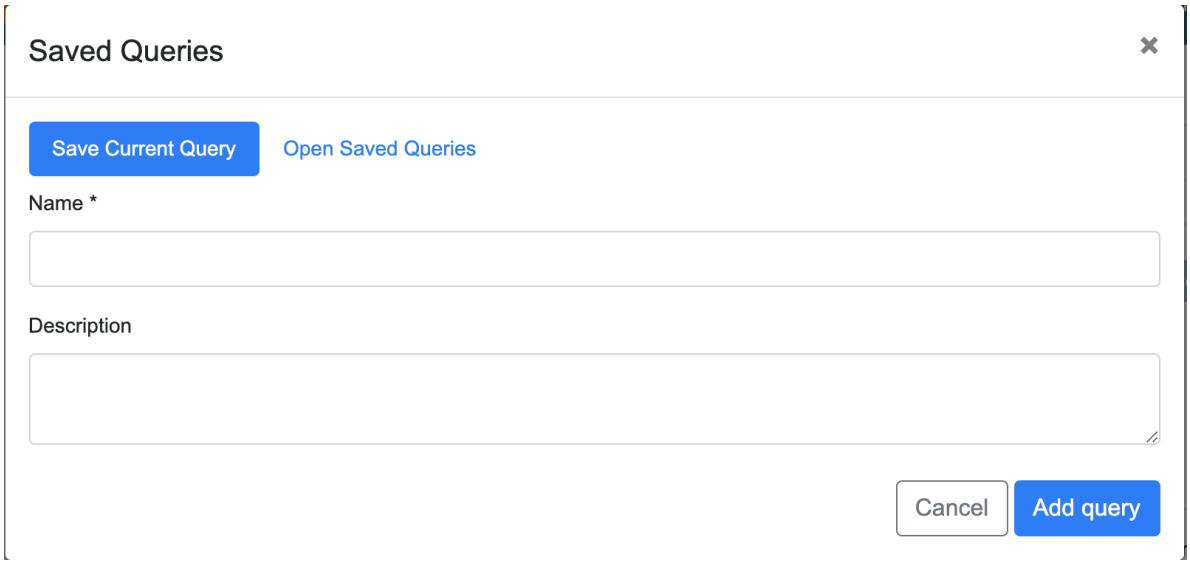
Save your queries for example to review the data in them at a later point in time, or to filter the data in your case. To save your query click on the button "Saved Queries" while you have an active query.



Once you've clicked the saved queries button, a modal will open with two menu options.

- **Save Current Query**

  Saves your current query and filters. Add a name to identify the query later on.
  There's also a possibility to add a description if you wish to do so. Adding a description is not required.

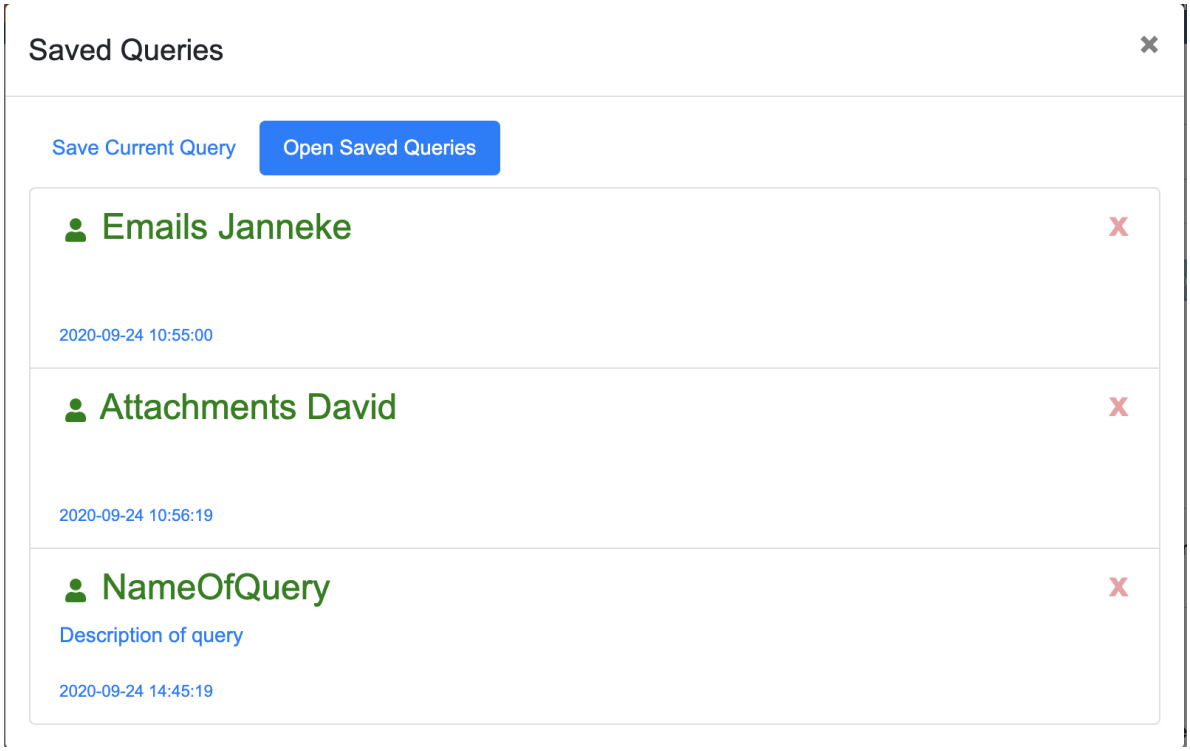


- **Open Saved Queries**

Here you can open your already saved queries. It gives an overview of the name and description of the query.

If you hover over the saved query it will also show the terms and filters that the query has.

Clicking on the item brings you to the search overview with the active query.



CHAPTER

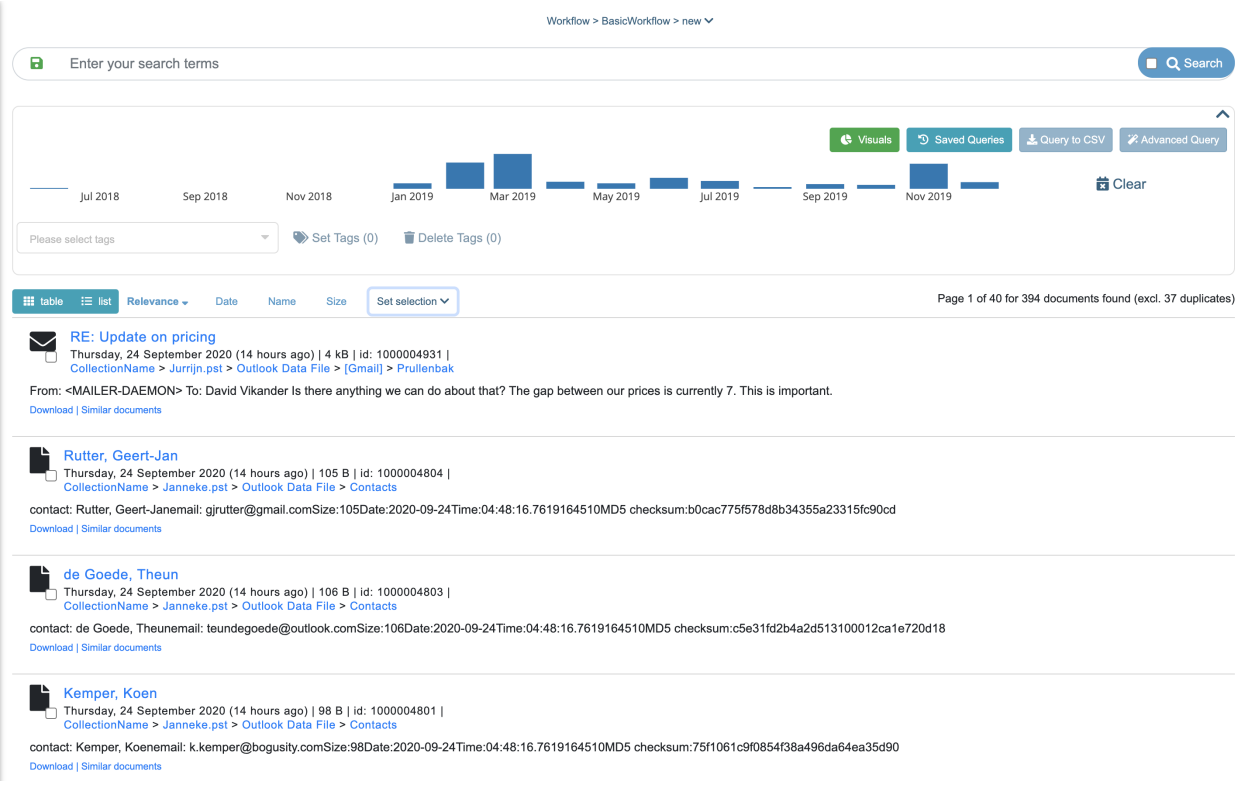
THIRTYTWO

SELECTING AND DE-SELECTING

Table of Contents

- *Selecting and de-selecting*
  - *Select documents on this page*
  - *Select all query results*
  - *Deselect all*

Apart from selecting documents by hand by clicking on the checkbox next to a document, it's also possible to select and deselect multiple documents at once. Just above the search results there's a button with an option to 'Set Selection'.



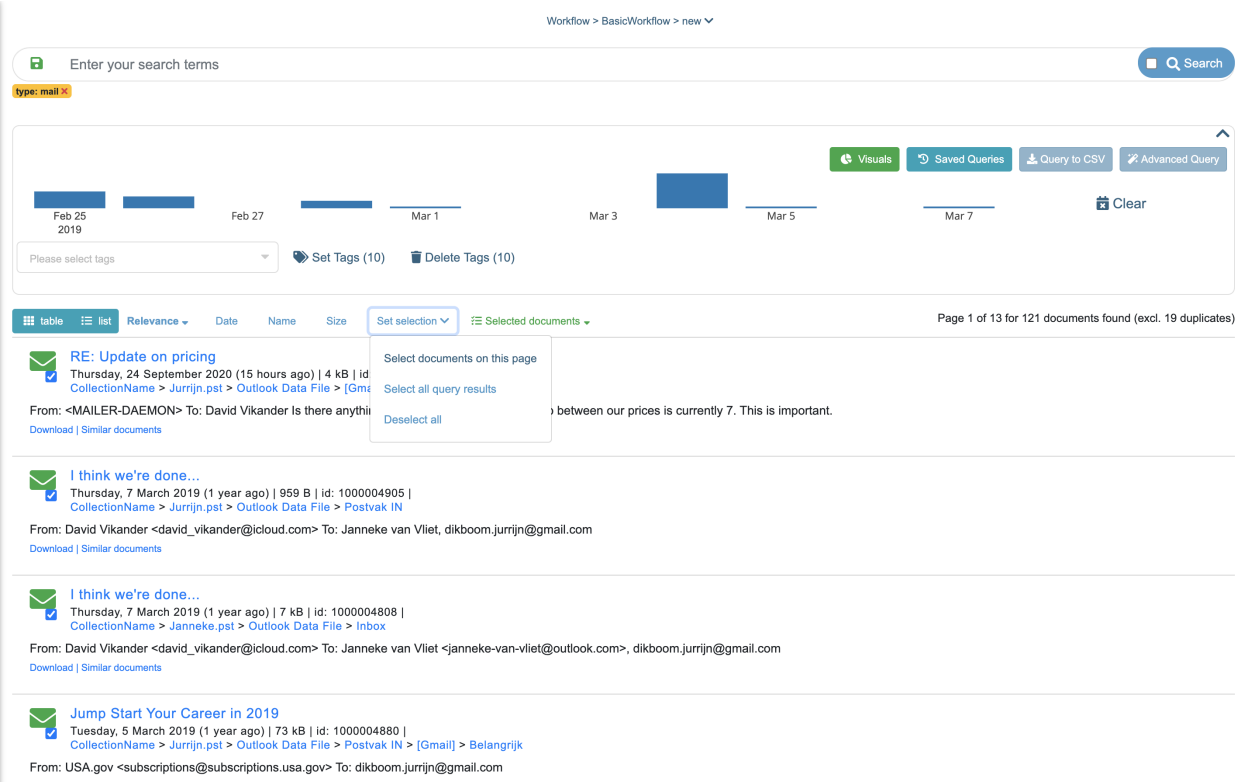
This comes in handy, especially when using the workflow to move multiple documents at once, or for tagging multiple documents at the same time. Once you click 'Set Selection' there are three options:

- Select documents on this page

- Select all query results
- Deselect all

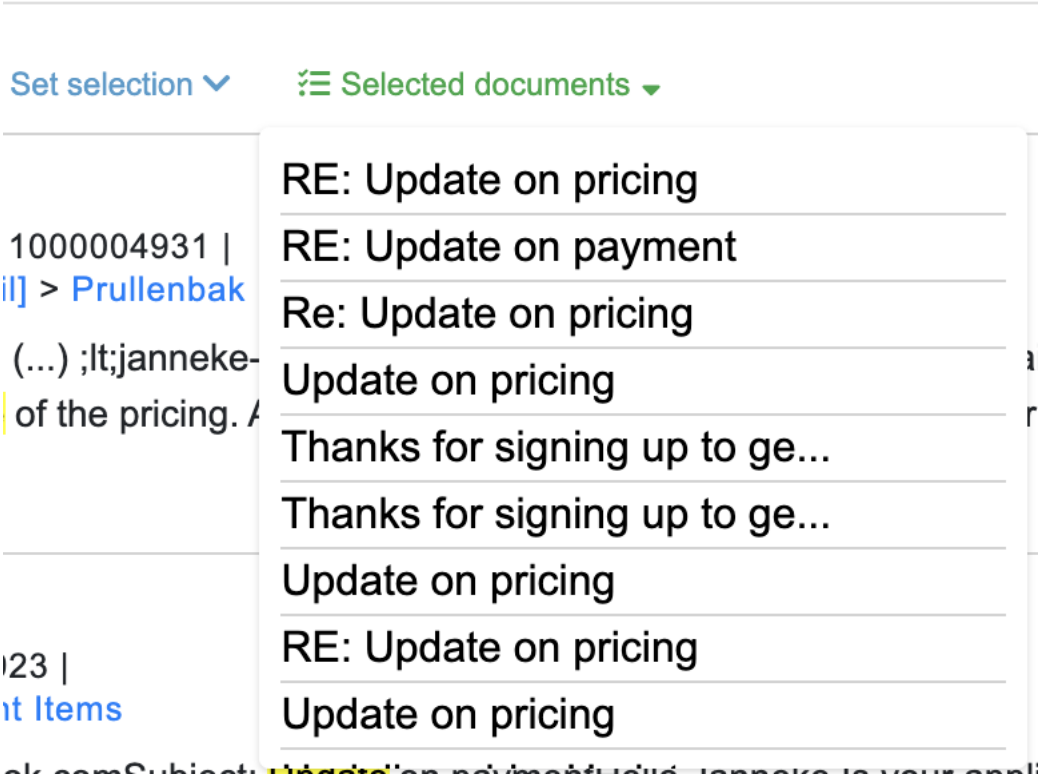
32.1 Select documents on this page

This option selects all the documents that are shown on the page at that point in time. The icons of the selected documents will turn green and the checkboxes will be checked. See example below.



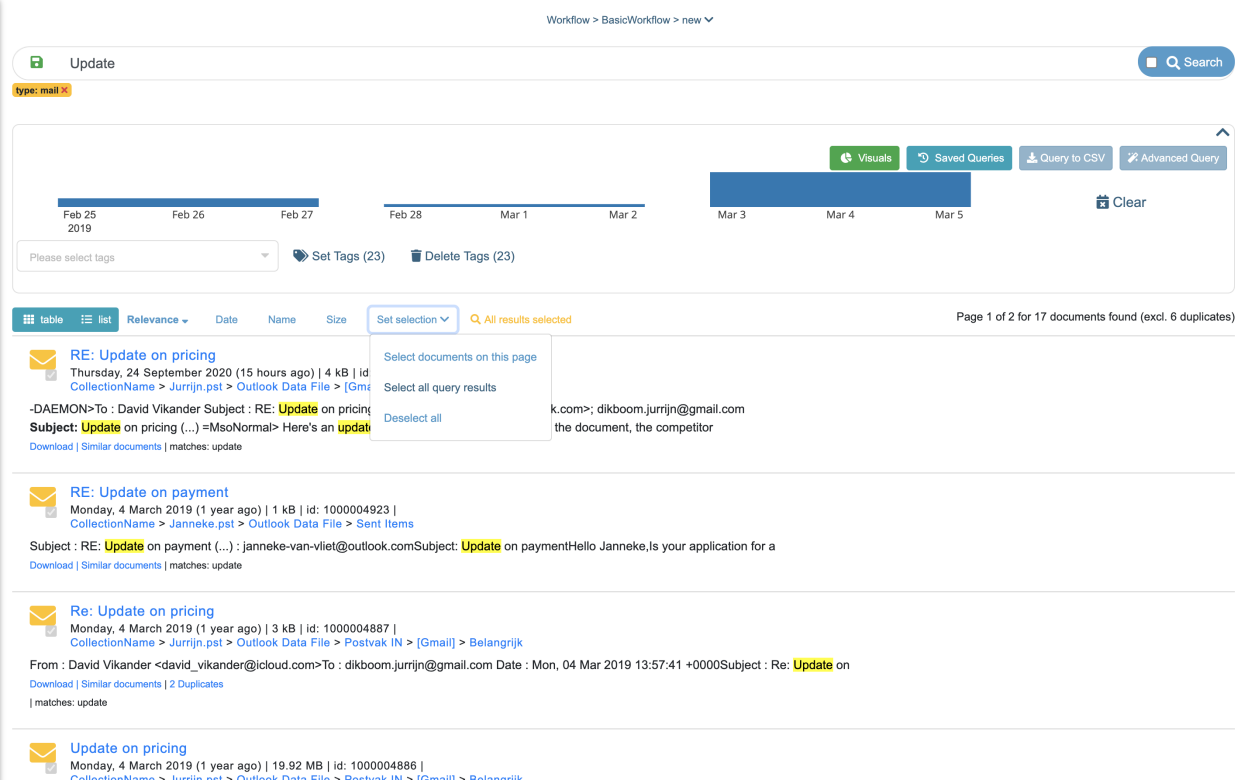
Note: If you select the documents on a page, it will only select on that specific page. If you decide to go to the next page, *the selection will be undone*.

To have a quick overview of which documents are selected there's also a small overview of the selected documents. See image example below.



32.2 Select all query results

This option selects **all** the documents results from an executed query. The icons of the selected documents will turn orange and the checkboxes will be checked. See example below.



32.3 Deselect all

The deselect all option will deselect all documents previously selected. The icons will turn back to their normal color and the checkboxes will be unchecked.

CHAPTER

THIRTYTHREE

SORTING

Table of Contents

- *Sorting*

Within the search overview there are a few easy sorting options.

- Relevance

  Sort the documents by relevance. Relevance is set during the indexation process. This filter is set by default at descending.

- Date

  Sort the documents by date. Ascending or descending.

- Name

  Sort the documents by name. Ascending or descending.

- Size

  Sort the documents by size. Ascending or descending.

CHAPTER

THIRTYFOUR

TABLE & LIST VIEW

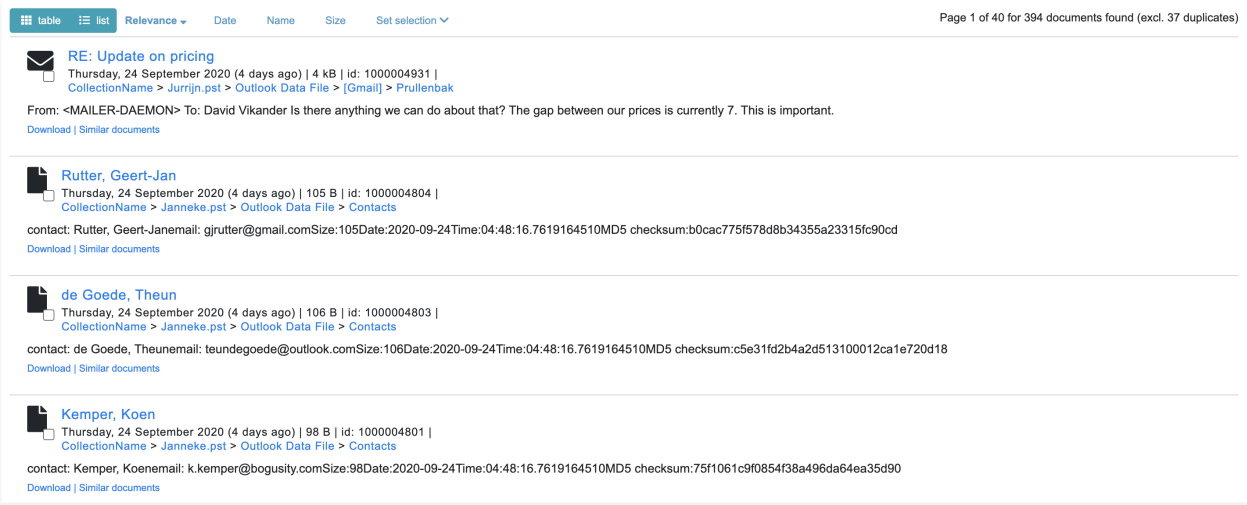
Table of Contents

- *Table & List View*
  - *List view*
  - *Table view*
  - *Table view (with flipped table)*

Since version 11.0.0 of INDICA it's possible to switch the document results between a list view (as you're used to) and a table view. List view is the standard view when opening INDICA. If you want to switch between the two, you can click on the button for 'table' or 'list' just above the search results.

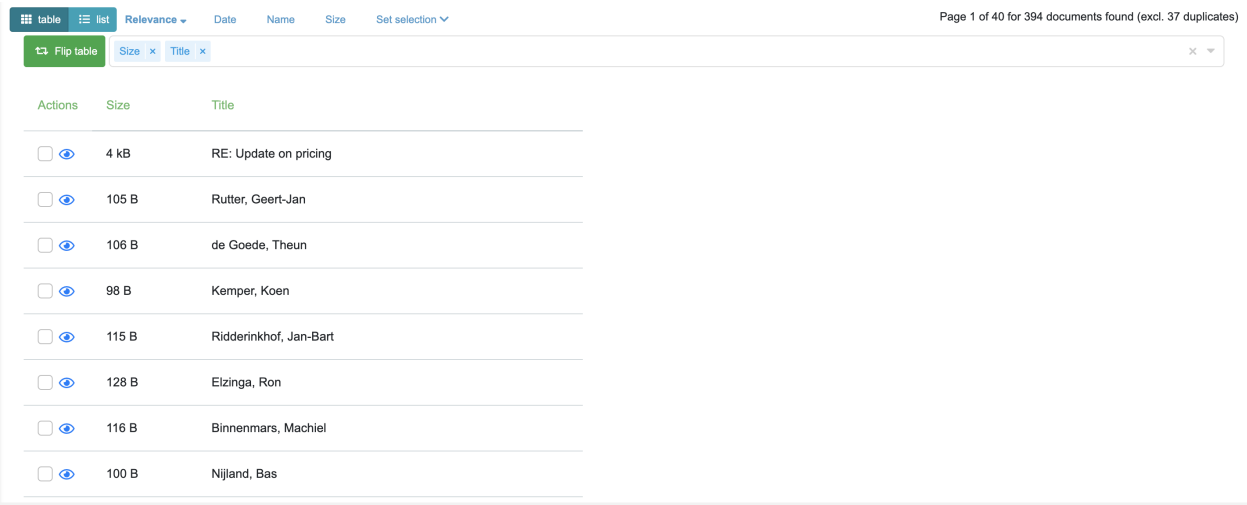
34.1 List view

If you've used INDICA before, the list view is what you're probably used to. Here you'll see the data listed per item, the information provided per item is set up by case managers.



34.2 Table view

The table view is new in our platform. The first time you click on the table view, it will only show the *size* and the *title* of the document. In table view you can view all the information just as in list view. For this to happen, click the dropdown above the table and select the information you want to view. Clicking on the eye icon will open the previewer for you.



34.3 Table view (with flipped table)

If you have a lot of items selected in your table view, it might be hard to have a good overview of the item. There's also a possibility to flip the table if needed. This will give a vertical overview per item, instead of the normal horizontal view. See image below for an example.

CHAPTER

THIRTYFIVE

TAGGING

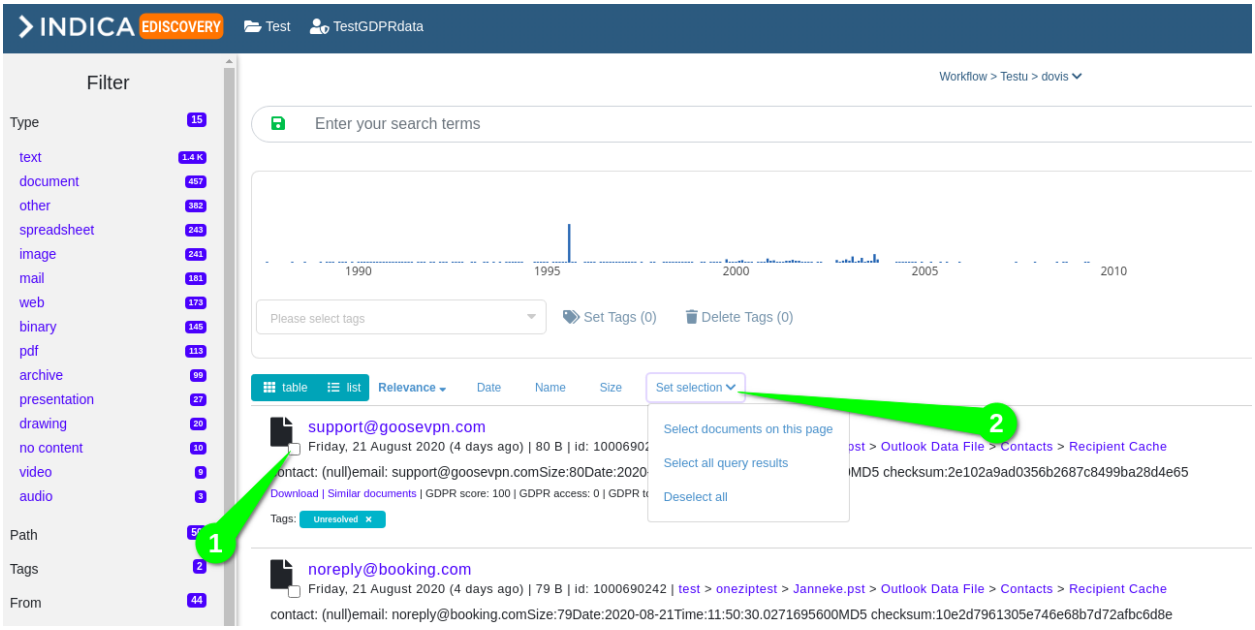
Table of Contents

- *Tagging*
  - *Selecting items for tagging*
  - *Tagging selected items*
  - *Tagging individual items*

Tagging function is available in both main search page and on every individual item that you *preview*. There are multiple ways you can start tagging:

35.1 Selecting items for tagging

- Tag individual items by clicking on the checkbox (1).
- Mass tag multiple items by selecting "Set selection" dropbox (2).

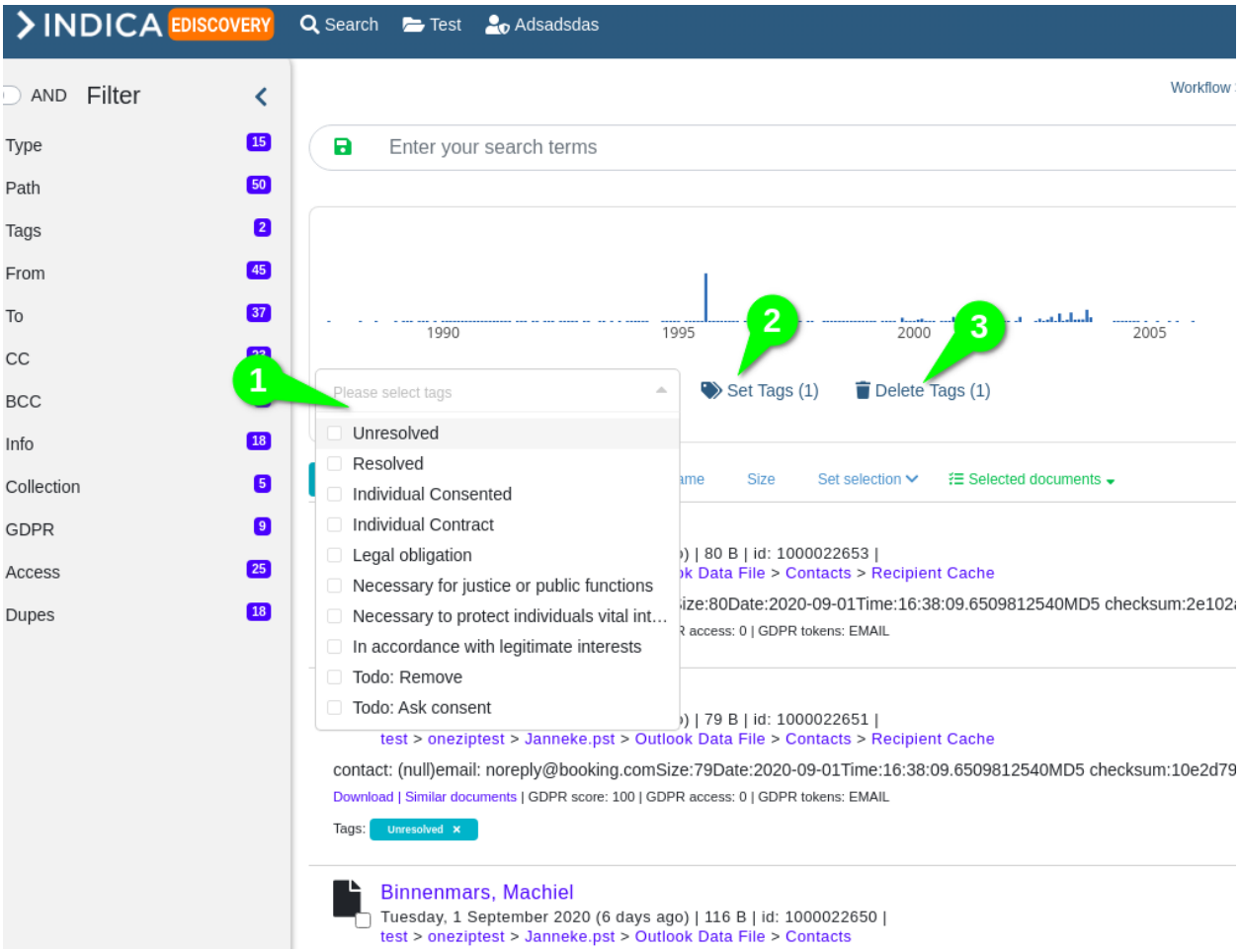


- "Select documents on the page" - will select all the documents on the current page.

- “Select all query results” - will select everything that was found based on current query.
- “Deselect all” - deselects everything you have selected.

35.2 Tagging selected items

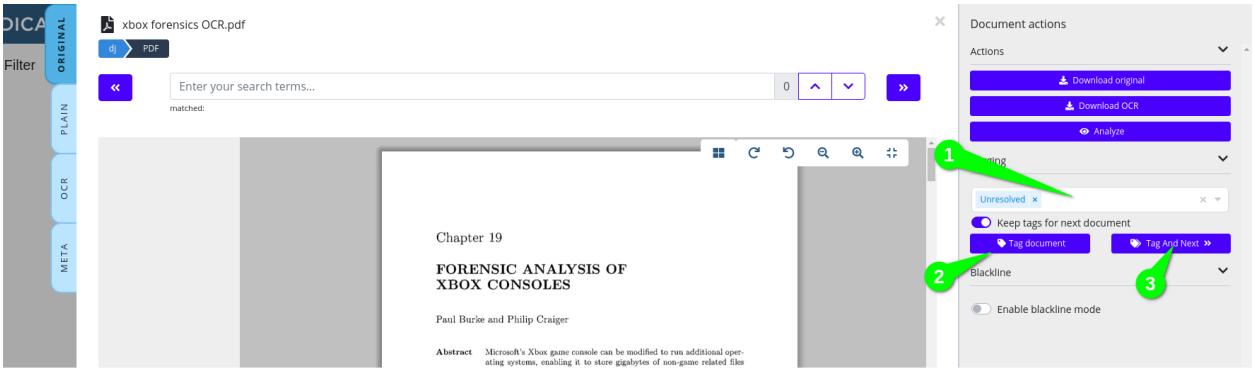
Once you have selected the items you can take further actions.



First you select which tags you want to add/remove (1). Then you either Set the tags (2) or Delete the tags (3). It’s important to note that having a lot of selected items takes longer to process the set or delete actions. The process will run in the background and you will be notified once it is done.

35.3 Tagging individual items

Sometimes you might want to tag individual items when auditing. To tag individual items you can still use the mass tagging tools, but a more informed decision can be made if you open the document *preview*



Once you have opened the *preview* you can select tags (1) on the right side of the window. Once the tags are selected you can either "Tag document" (2) and stay on the same document, or "Tag and next" (3) to tag the document and move to the next one.

CHAPTER

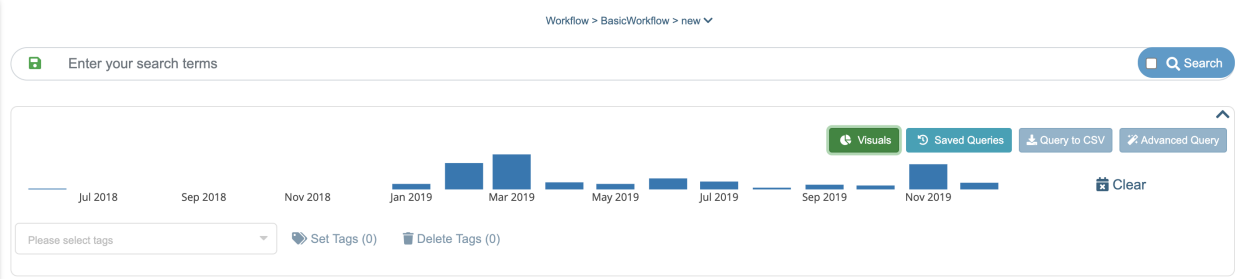
THIRTYSIX

VISUALS

Table of Contents

- *Visuals*
  - *Location*
  - *Per Country*
  - *Per type*
  - *Email*
  - *NLP Graph*

INDICA offers a visual representation of your data. To access the visuals overview click on the "Visuals" button in the search overview.



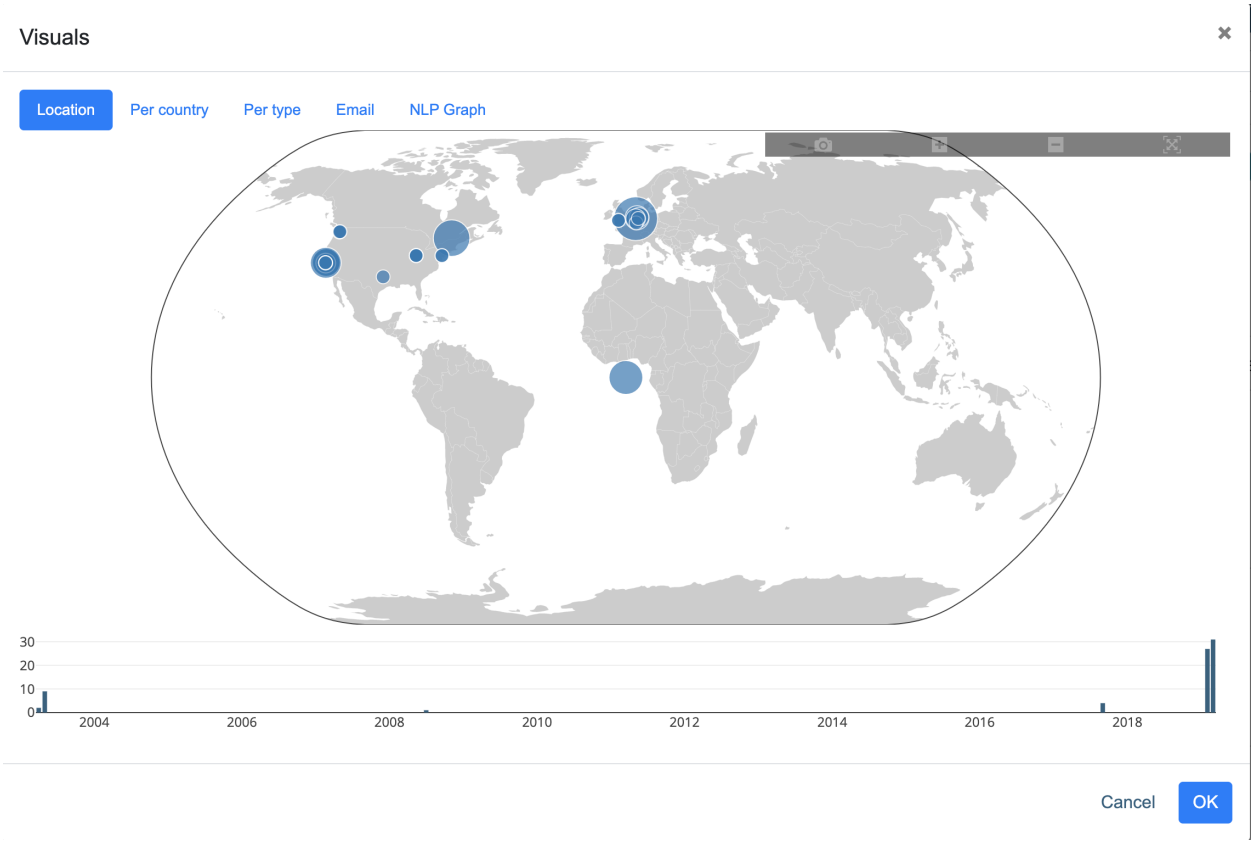
There are a few tabs you can click in the screen that opens. They will be discussed in the chapters below.

The graphs shown in the *Location* and *Per Country* tab also have a timeline underneath. Clicking and dragging your mouse over this timeline will set an extra filter on the graph. This updates the results in the graph to reflect only the emails sent in that specific time range.

Tip! To zoom the timeline back out again, double click on the timeline to reset

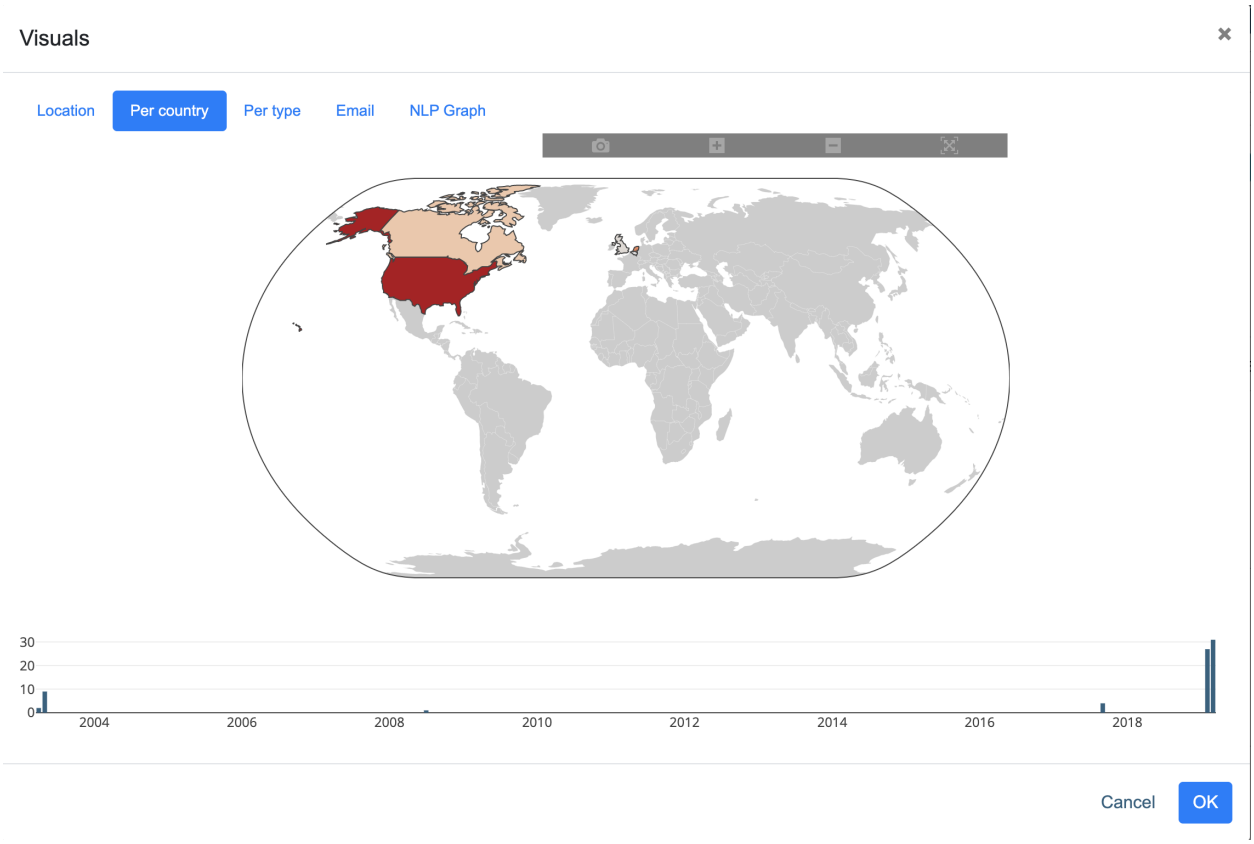
36.1 Location

This tab shows a visual representation of email communication around the world. Clicking on an item within the graph will filter the search results to emails sent from that specific IP address.



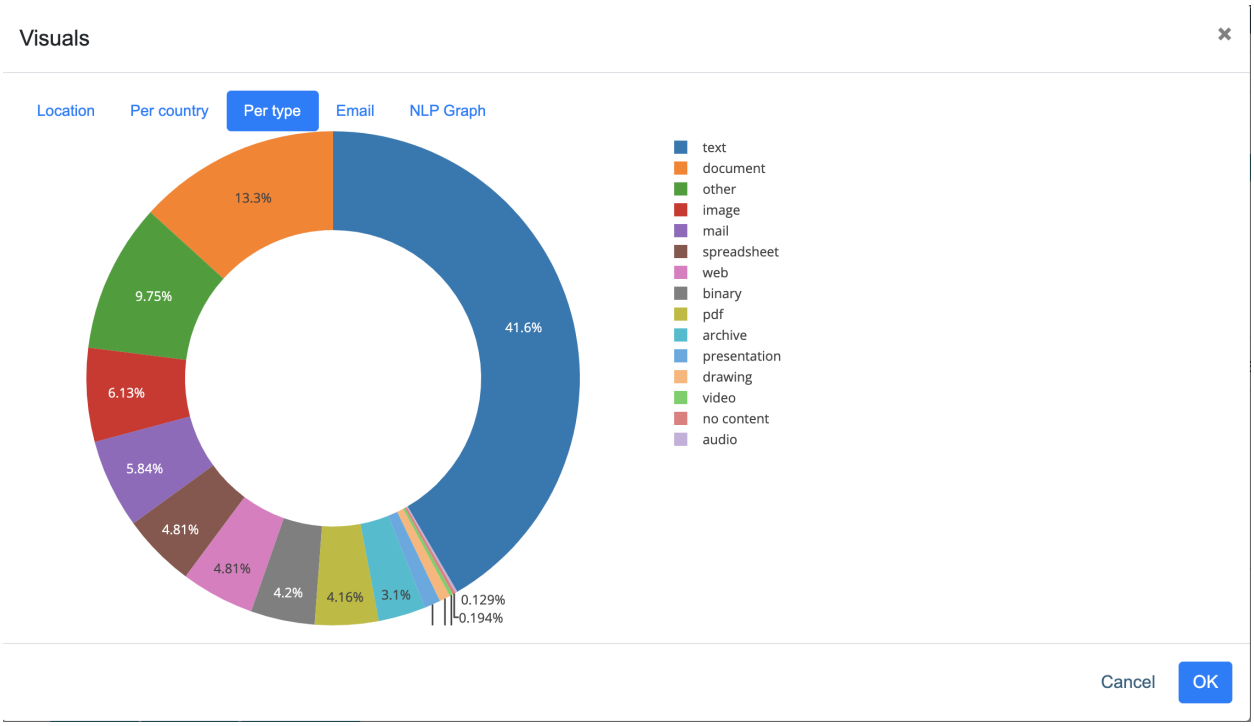
36.2 Per Country

This tab shows a visual representation of email distribution per country.



36.3 Per type

This tab shows a visual representation of the available data types in the dataset.



36.4 Email

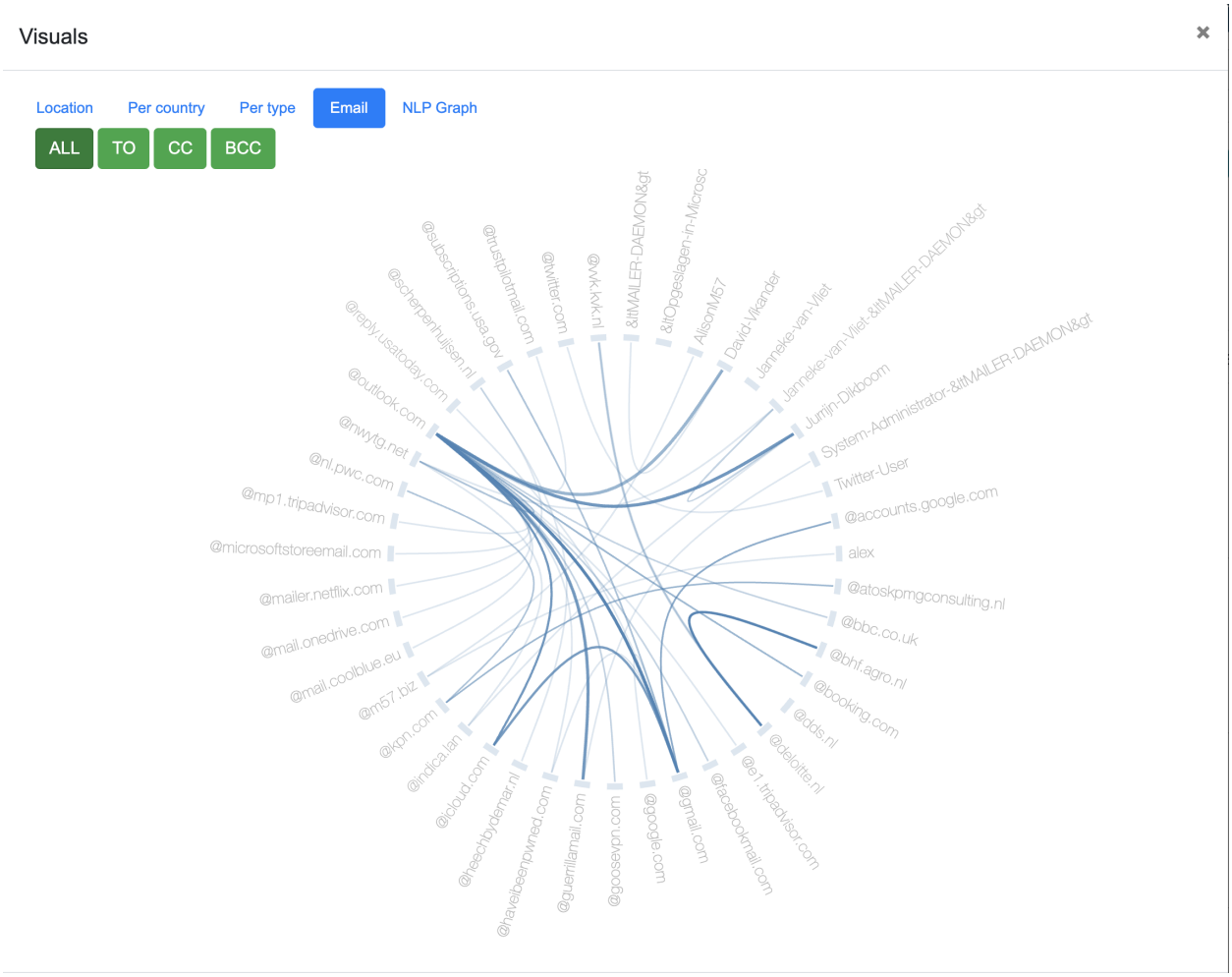
This tab shows a visual representation of email communication between different persons. There's also a possibility to filter these from 'ALL' to 'TO', 'CC' or 'BCC'. Hovering over items will show you the email addresses they've been in contact with.

- **Green lines**

  Shows the node source(s). So the email account(s) that have sent email(s) to that specific email address.

- **Red lines**

  Shows the node target(s). So the email account(s) that specific email address has sent email(s) to.



36.5 NLP Graph

This tab shows the NLP graph. It is possible to visually see documents, their Natural Language keywords, and the connections between them. This enables you to quickly see and discover connections between documents, based on the content of the documents.

In the graph, the number of documents is limited to 20. The blue circles are documents, the green circles are the NLP keywords. The green circles with a red border do not have a connection or more documents - other than its parent.

The displayed documents can be manipulated by applying a query to your data and opening the graph again.

The plot can be manipulated by dragging and dropping the circles, or by dragging the background to pan around. Using the mouse wheel will zoom in and out.

In the top left corner, some controls are provided:

- Draggable scrollbar

  This increases and decreases the 'gravity' between nodes.

- Filter icon

  This will filter out all the nodes that do not connect to more than one document.

- Wand icon

  This enables automatic zooming. When nodes are not moving, it will zoom in and out automatically.

- Magnifying glass

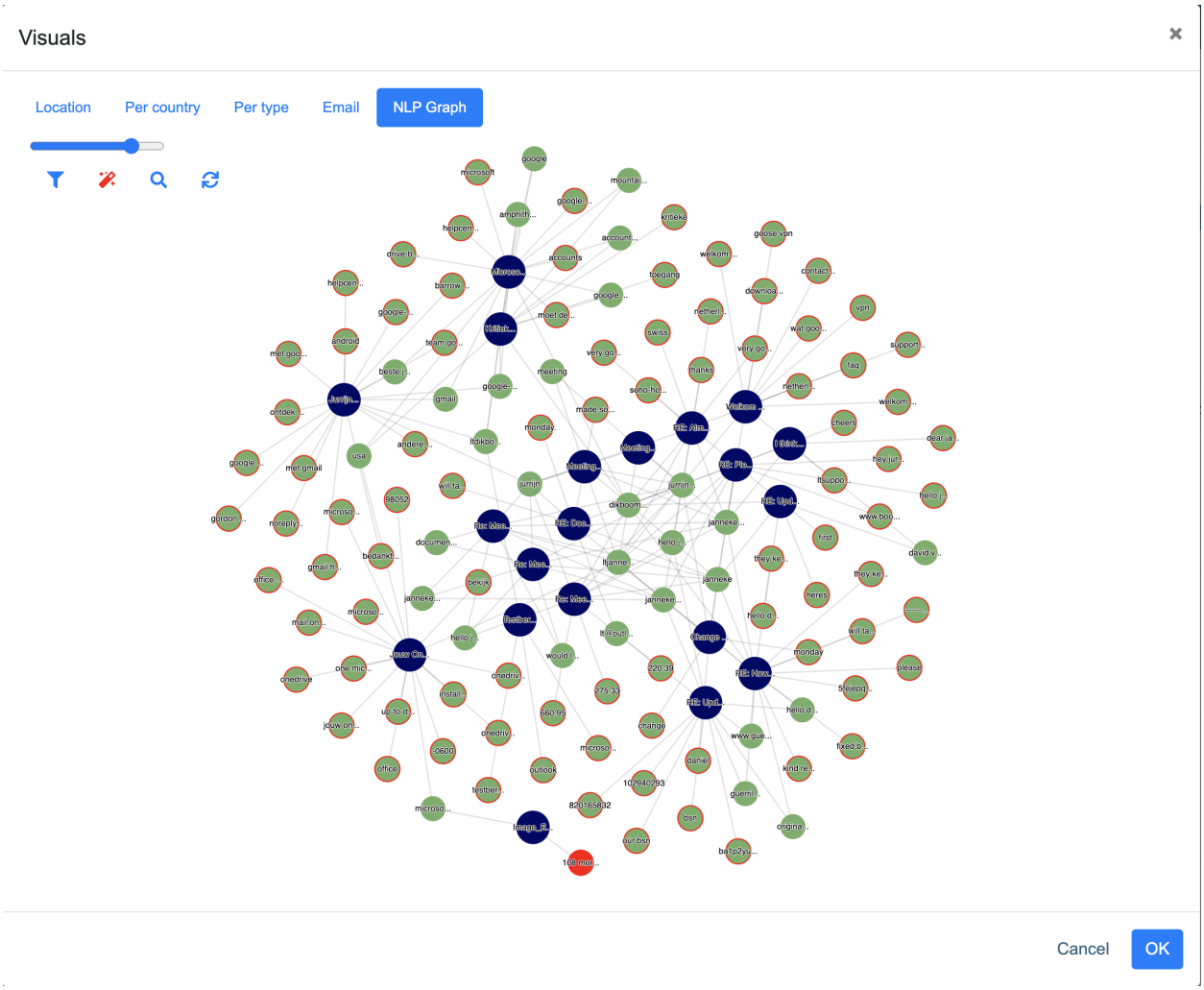
  Clicking this button will zoom in or out so all nodes are visible.

- Reset icon

  Once one or more nodes are clicked, the selection can be reset with this button.

Double clicking a node will open the INDICA interface with a query applied to search for that node. This can be a keyword, or a document.

Filters are currently not applied to this graph.



CHAPTER

THIRTYSEVEN

WORKFLOW

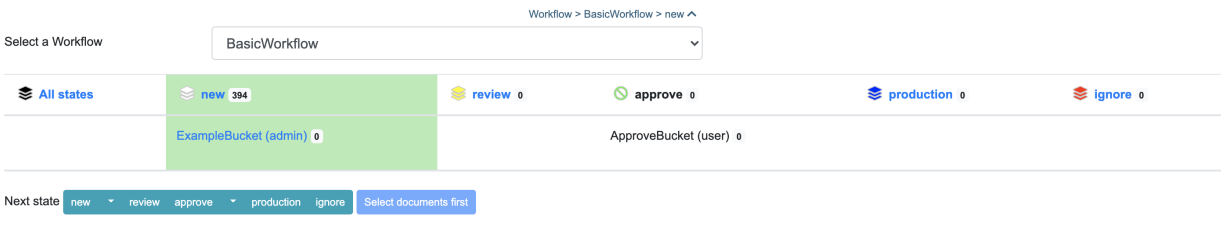
Table of Contents

- *Workflow*
  - *States*
  - *Buckets*
  - *Moving documents*

If there's an active workflow within your case you'll see an extra item above your search bar. This item can be uncollapsed by clicking on it, giving you a full overview of the workflow.

The overview of the workflow shows the different states and assigned buckets to users. States might be accessible to multiple users (depending on the workflow settings), but buckets are user-specific. There are three reasons why you might not see the Workflow section in your workspace:

1. The section is collapsed on your screen, please press the Workflow button.
2. There is no workflow set up for your case.
3. You do not have access to the workflow.



37.1 States

You can see if a state is accessible to you by the icon next to the state name. If the icon is a stack of layers, you can access this state. If the icon is a prohibition sign, you do not have access to that particular state.

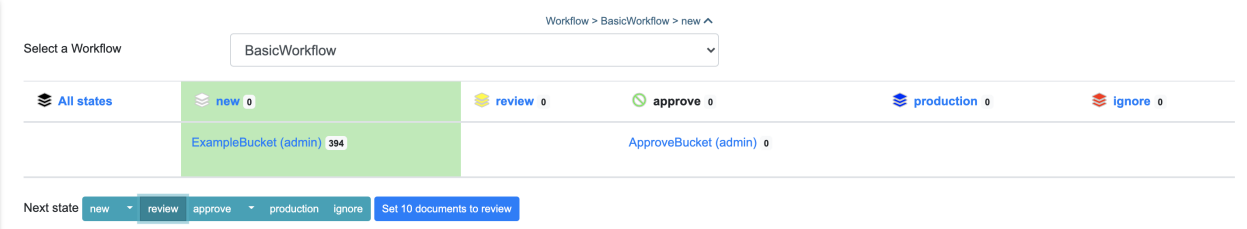
Which state is the active state can be recognized by the green background on that state. Clicking on a state will set the state to active, which sets a filter on the data to the items in that state.

37.2 Buckets

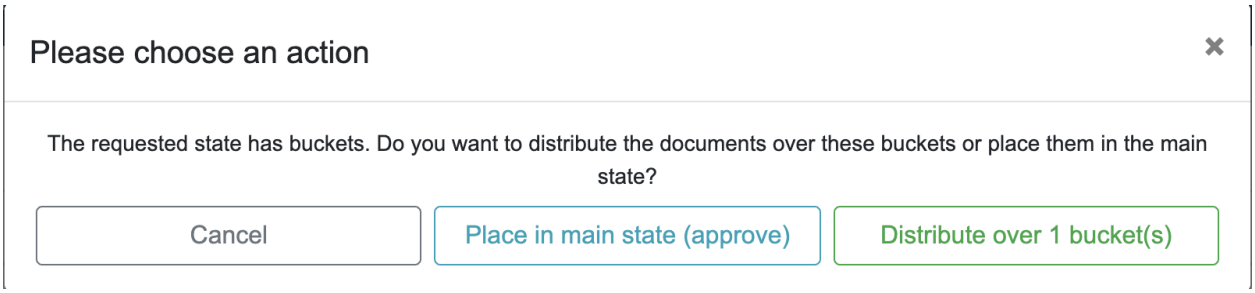
A workflow may have a state with (multiple) bucket(s). Unlike workflow states, buckets are user specific and are only accessible to that specific user. A case manager or other reviewers can set documents to your bucket for you to review.

37.3 Moving documents

Moving documents from one state (or bucket) to another is quite simple in INDICA. Select the documents you wish you move. The button next to the different states will then change from 'Select documents first' to 'Select next state'. Once you've selected a state or bucket, the button will turn blue and show 'Set *x* documents to *statename*'. Once this button is clicked, the documents will be moved to your selected state. See image below for an example.



If you've selected a state which also has buckets, you will get a prompt in which you will be asked if you want to move the documents to the specific state, or if you wish to distribute the documents over the different buckets.



CHAPTER

THIRTYEIGHT

CONTACTING INDICA

Table of Contents

- *Contacting INDICA*

In case the available documentation does not provide the answer to your question, please don't hesitate to get in touch with us.

INDICA Support

INDICA Knowledge Base

INDICA Website