

Key features of iManage Extract

Collect



Connect to iManage Work or other document repositories through API or simply upload documents using drag and drop

Cluster & Classify



Use Machine Learning to cluster and classify documents based on content

Workflow



Create workflows to manage training, review, and quality assurance processes

Search, Tag & Filter



Conduct bulk searches, locate and organise documents and extraction results

Extract



Use Pre-Trained extractors from the Extractor Library or easily train new extractors using the Extractor Toolbox to identify key information from documents

Export

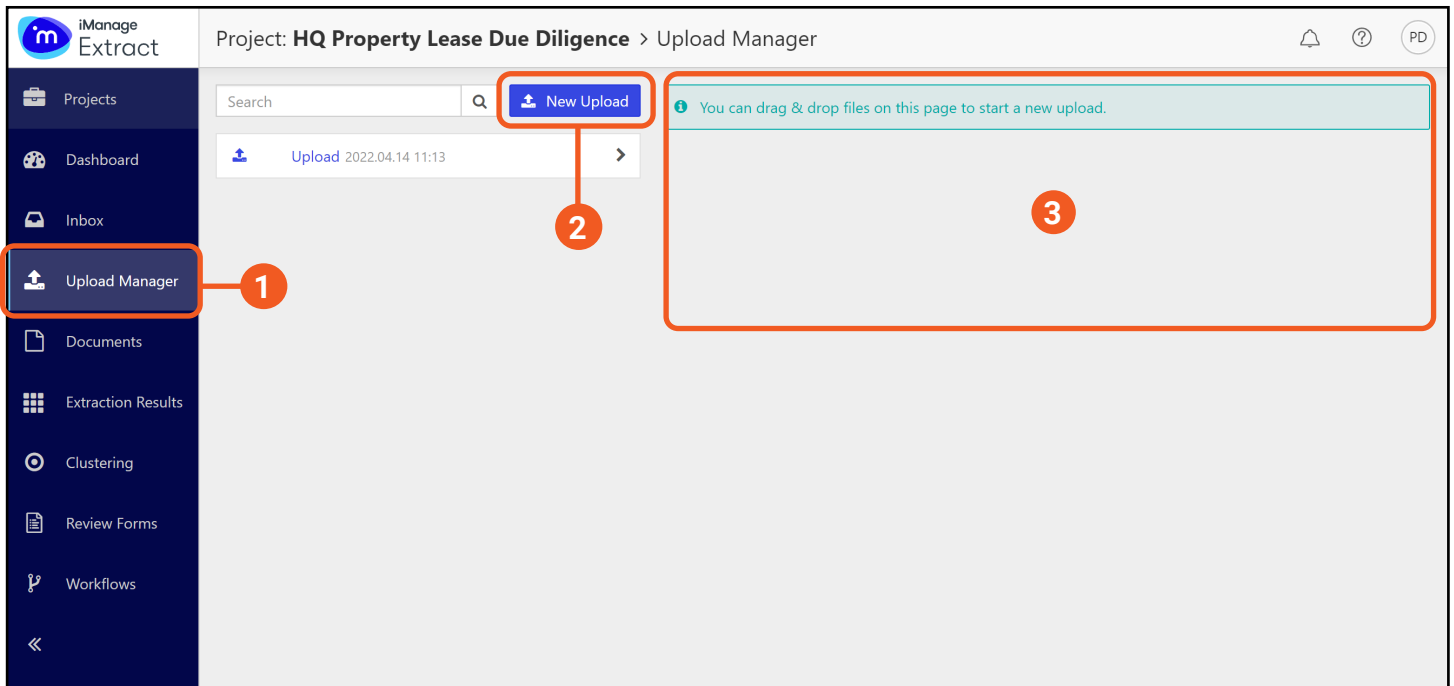


Export the extracted structured data and use it in a way that will bring you value



Collect

You can upload files into Extract projects by either using the **Upload Manager** feature in the Extract front-end or use the latest **Extract API** to connect your current documents repository system with Extract and ingest documents directly from there.



1. The Upload Manager tab

The **Upload Manager** tab is situated within a project, in the left navigation pane.

2. The New Upload button

Select **New Upload**, then select the files/folders/archives you to upload and select **Open**.

3. Drag and Drop uploading

Drag the files from your source on to the Extract Upload Manager interface to upload the files.

NOTE: Currently, we support uploading multiple file types, including simple folder, .zip, .pdf, .doc and others.

4. Extract API

Use the **Extract API** to upload documents or connect with your existent documents repository.

The Extract API is organized around **REST**. Our API has predictable resource-oriented URLs, accepts JSON-encoded request bodies, returns JSON-encoded responses, and uses standard HTTP response codes, authentication, and verbs.

NOTE: You can upload files to a project if you are either a **Superuser** or a user assigned the role of **Ingestor** in the project.



Cluster & Classify

Clustering in Extract is an unsupervised, machine learning technique that groups documents based on those which appear to originate from a similar template. You can create up to five cluster jobs per project. Leverage this classification feature to help you with your extraction strategy.

The screenshot displays the iManage Extract interface for a project named 'Commercial Contracts Due Diligence'. The left sidebar shows the 'Clustering' tab selected (1). The main area shows a 'Sample Clustering' job with 45 documents and 70% similarity (2). A button indicates 'Only 50 clusters shown' (3). The right side shows a grid of clustered documents, with one cluster labeled 'Simpson Associates' (4). A dialog box for creating a new cluster is open, showing fields for 'Cluster job name', 'Similarity threshold' (set to 75%), and 'Strategy' (set to Automatic).

1. The Upload Manager tab

The **Upload Manager** tab is situated within a project, in the left navigation pane.

2. Creating a new Cluster

In the **Documents** tab, select all the documents you would like to **Cluster & Classify** and select the **Cluster** button.

A dialog window will open where you will be able to give a **Cluster job name**, select **Similarity threshold** and **Strategy**.

3. Cluster View mode

You can choose to view the clustering job either in **Proportional** mode or **Equal** mode.

4. Clustered Documents

After the clustering job has been completed, you can see the number of documents that have clustered and classified and percentage of similarity threshold.

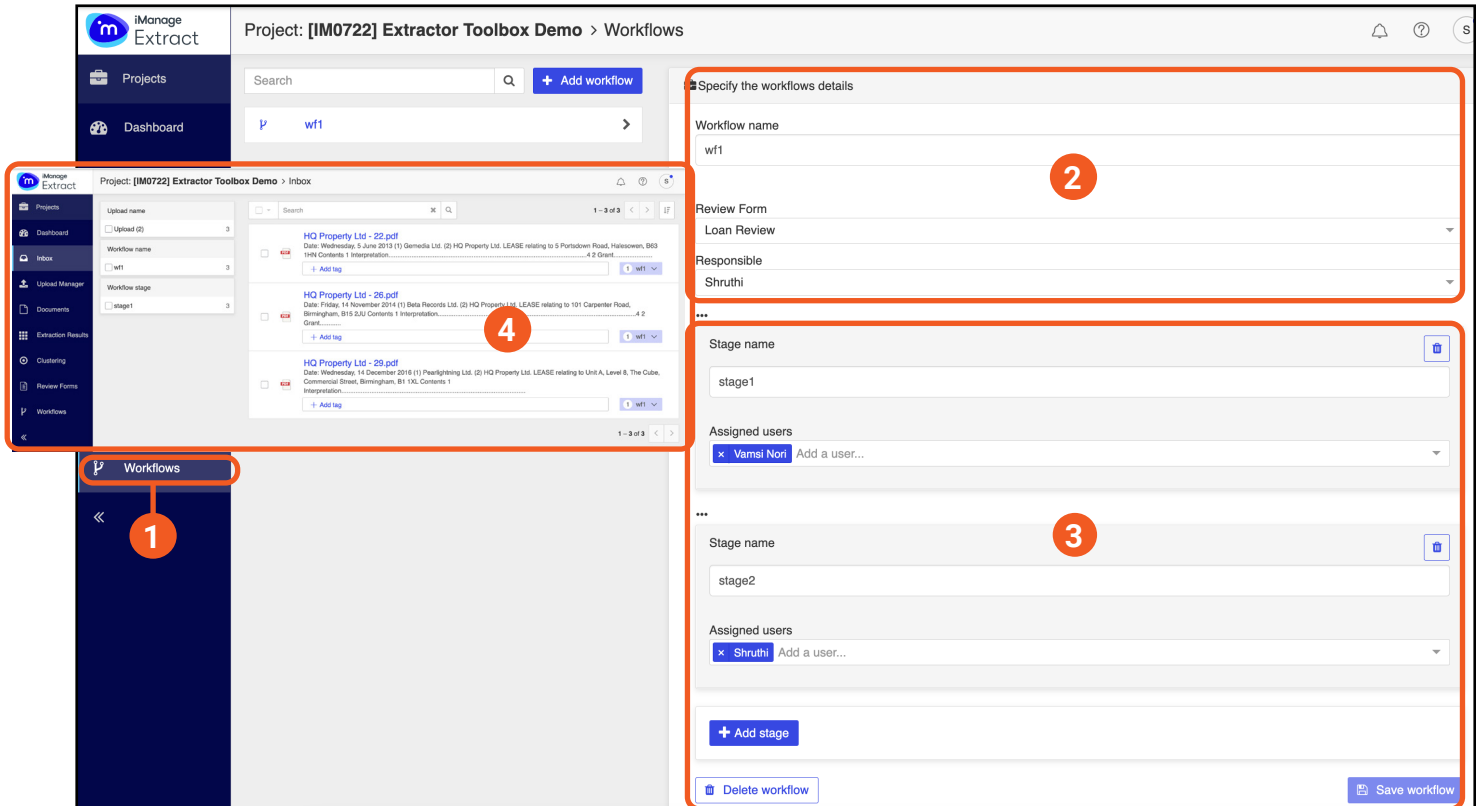
NOTE: You can view the Clustering page and perform clustering tasks if you are assigned the **Ingestor**, **Trainer** or **Superuser** roles.



Workflows

Workflow in **Extract** is the journey a document goes through before it is fully reviewed. Workflows have **stages**, and different users can be responsible for different stages of the workflow. For example, an employment contract document have to pass through various users, including the interviewer, recruitment manager, finance manager, and business head, in that particular order. At each stage the reviewer must **sign-off** on the document to move it to the next stage.

A workflow can be created per review team, per document type, per project. Workflows are manually created and a project can have more than one workflow.



1. The Workflows tab

The **Workflows** tab is situated within a project, in the left navigation pane. You can view, create and edit workflows from here.

2. Creating a new Workflow

After selecting **Add workflow** you will need to configure your new workflow. You can enter a **name**, assign a relevant review form from the **Review Form** drop-down list and select a user, who is responsible for adding documents to the workflow in the **Responsible** field.

3. Adding Stages

To add workflow stages, select **Add stage**. Each stage must be given a **name** and have users **assigned** to it. Multiple users can be assigned to a stage, and a user may be assigned to multiple stages. Assigning users to a stage enables them to use their **Inbox** to view only the documents, which are ready for them to review.

4. Inbox

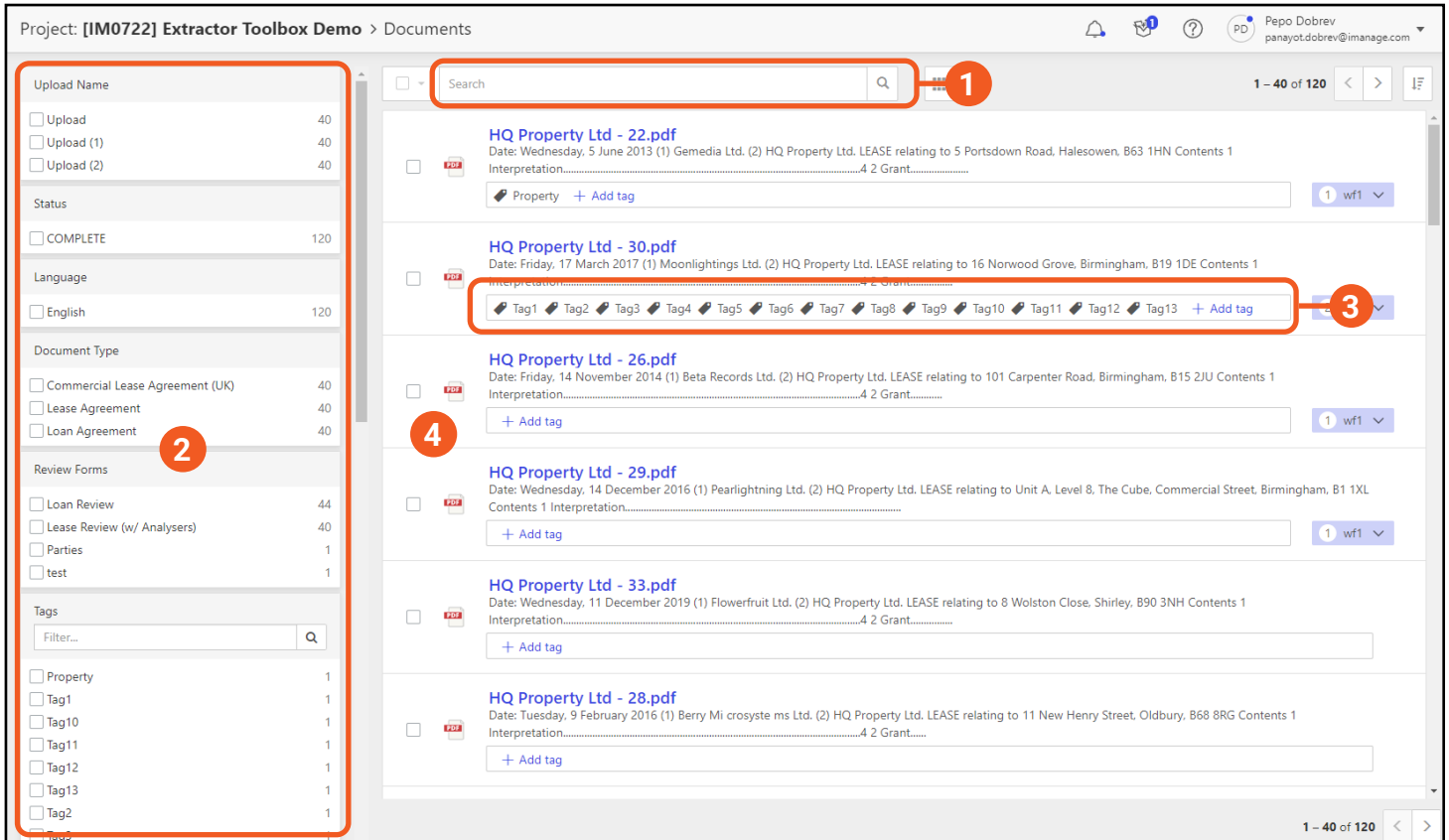
If you are an assigned user for a stage in a workflow, when the documents move to your workflow stage, in the **Inbox** you can view the documents to act upon. After you take the necessary action on the task, it moves to the next stage and the user responsible to act next receives the documents in their **Inbox**.

NOTE: Only users with **Admin** or **Superuser** role can perform the tasks described here. Users with **Reviewer** role can view the Inbox.



Search, Tag and Filter

From the **Documents** page, you can search for a document, run tasks on the documents, and export the extractions. You can also **search** the text and title of a document, **filter** out documents using facets and **tag** documents, so you can easily find them at a later date.



The screenshot shows the iManage Extractor Toolbox Demo interface. On the left, there is a sidebar with facets for filtering documents. These include Upload Name, Status, Language, Document Type, Review Forms, and Tags. The main area displays a list of documents, each with a title, date, and content summary. Red boxes and numbers 1-4 highlight specific features: 1. Search bar, 2. Facets, 3. Tagging options, and 4. Document list.

1. Search

The **Search** field enables you to search the text and title of a document. Boolean and Exact phrase searching are available. Group searching and Field based search is also supported.

**See Search Reference Table at the end of this guide*

2. Filter/Facets

Filter by **facets** and tags to organize documents and focus review. Facets appear only if there is at least one document that has a value in that Facet. Some useful facets include: **Document Type**, **Upload Name**, **Review Form**, **Tags**, **Governing Law**, **File Type** and others. Facets can also be added, changed, or removed using API.

3. Tagging

Documents can be tagged either **individually** or in **bulk**, and this adds a facet to enable documents to be filtered by the tag.

When documents are being ingested into iManage Extract during the Upload process, some tags are **auto-generated** by the system.

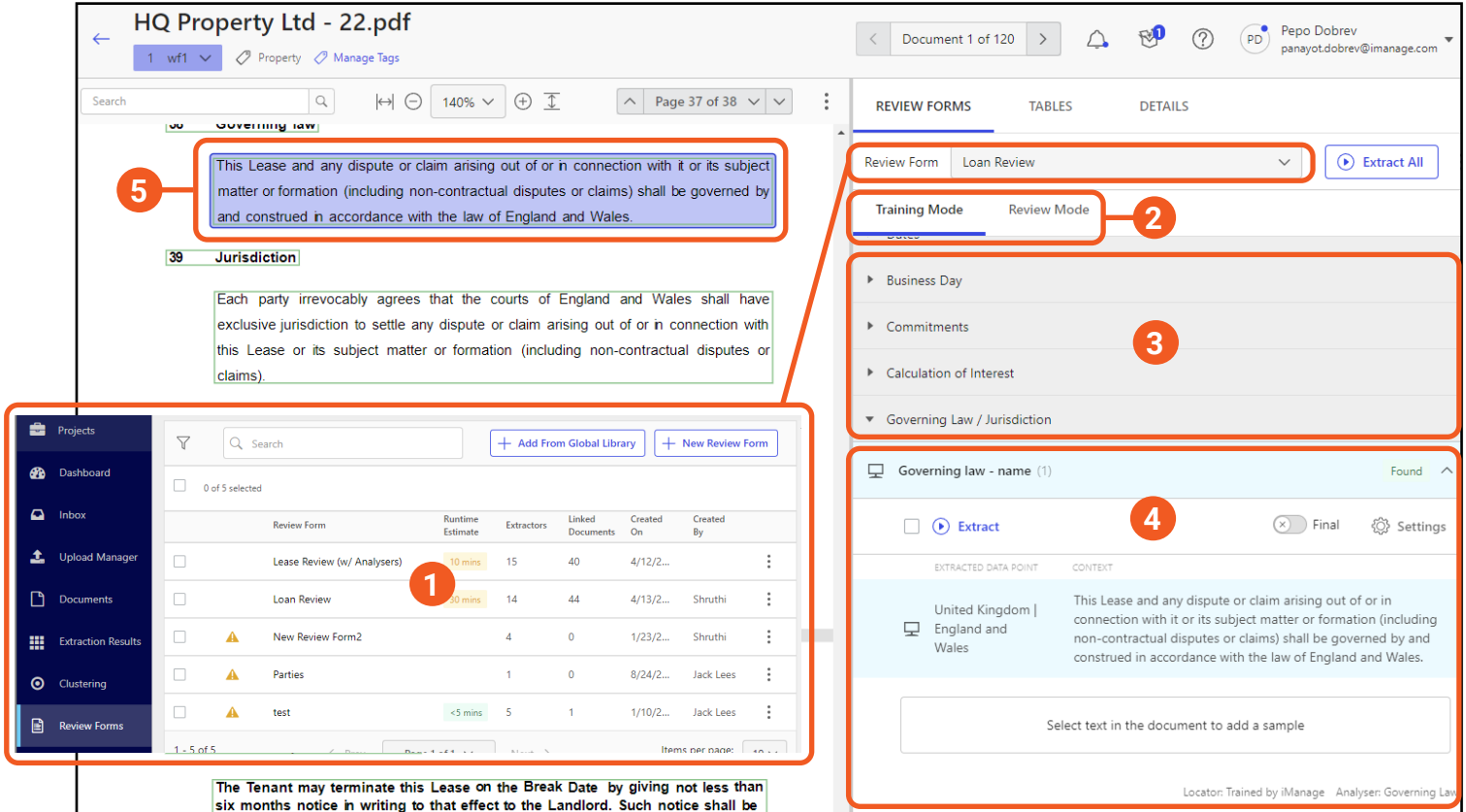
To tag a document, select the **Add tag** filed under the document, enter the tag text and press enter. When the page is refreshed, the tag appears in the Tags facet on the left.

To tag multiple documents at once, select the checkbox next to the relevant documents, or the **Select all** check box, and then select **Tag**.



Extract

The key feature of iManage Extract is the use of **AI, Machine Learning** and other techniques to **automate** the extraction of critical data points, creating **structured data** that enhances the quality of the services you provide. Extract manages this by utilising a tool called **extractor** which is responsible for finding the value of a data point in a document. Extractors can be trained by the user or available **pre-trained** by our internal iManage Lab Knowledge Engineers. Extractors are part of **Review Forms** which is a list of extractors that can be run on documents.



The screenshot illustrates the iManage Extract interface with the following components and numbered callouts:

- 1. Review Form Table:** A table listing review forms with columns: Review Form, Runtime Estimate, Extractors, Linked Documents, Created On, and Created By. It includes forms like 'Lease Review (w/ Analysers)', 'Loan Review', and 'New Review Form2'.
- 2. Review Form Selection:** A dropdown menu in the 'REVIEW FORMS' tab showing 'Review Form' and 'Loan Review'.
- 3. Review Mode Panel:** A panel on the right showing 'Training Mode' and 'Review Mode' tabs, with a list of categories like 'Business Day', 'Commitments', 'Calculation of Interest', and 'Governing Law / Jurisdiction'.
- 4. Extractor Details:** A detailed view of an extractor showing 'Governing law - name (1)' and 'United Kingdom | England and Wales' with a corresponding text snippet from the document.
- 5. Document Highlight:** A document page titled 'HQ Property Ltd - 22.pdf' showing a highlighted text snippet: 'This Lease and any dispute or claim arising out of or in connection with it or its subject matter or formation (including non-contractual disputes or claims) shall be governed by and construed in accordance with the law of England and Wales.'

1. Review Form

You can create a new review form, add extractors to it and link it to documents from the **Review Forms** tab. Review forms are project specific and are not available to other projects, however you can save a Review Form to the **Global Library** and then import it to other projects.

2. Training and Review Modes

Training mode is to train a machine learning model to find your data points within your documents by adding samples.

Review mode does not train the machine learning model. It can only **view** or **edit** already extracted data points or add additional data points without changing the extractor behaviour.

3. Extractors

View and **configure** extractors and their data points that are part of the selected Review Form.

4. Extractor Details

Run a single extractor, view extracted data points, add samples to train the model and configure a particular extractor.

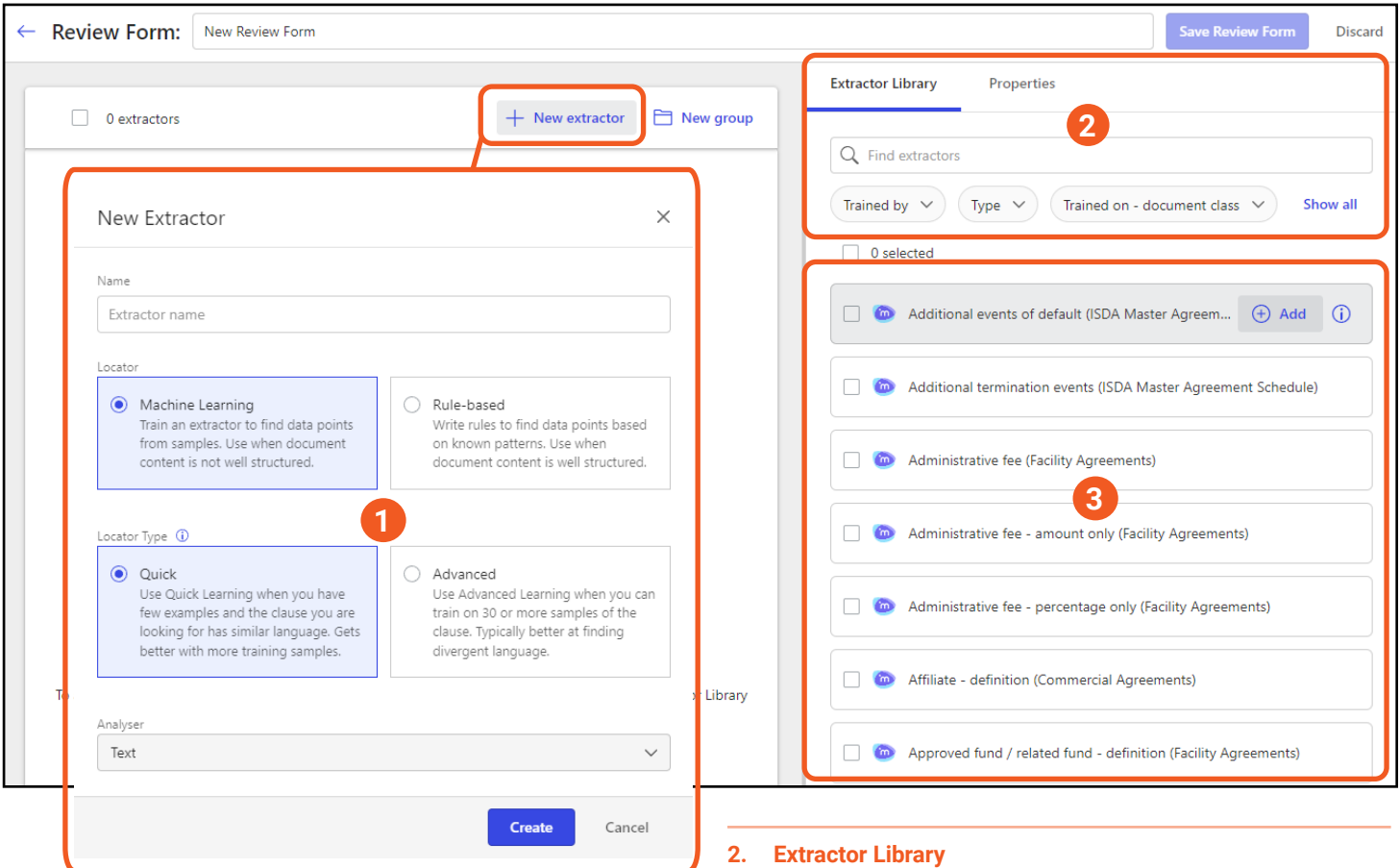
5. Highlight Data/View Extracted Data Point

Highlight the relevant data/text within a document while you have the corresponding extractor open on the review form panel to automatically populate as a sample for your extractor.



Extract Continued

In iManage Extract, you can either choose any of the hundreds readily available to you **pre-trained** extractors or choose to create your own from scratch. Custom extractors can utilise **Machine Learning** or be created using a set of rules - **Rules-based** extractors.



The screenshot shows the 'Review Form: New Review Form' dialog. On the left, the 'New Review Form' dialog is open, showing options for creating a new extractor. A red box labeled '1' highlights the 'New Extractor' dialog, which includes fields for 'Name' (Extractor name), 'Locator' (Machine Learning or Rule-based), 'Locator Type' (Quick or Advanced), and 'Analyser' (Text). A red box labeled '2' highlights the 'Extractor Library' on the right, which shows a list of pre-trained extractors. A red box labeled '3' highlights a specific extractor in the library, 'Administrative fee (Facility Agreements)'.

1. Extractor Types

Machine Learning - ML extractors allow end-users to "teach" Extract how to perform the data point extraction. The user provides a set of **samples** which are used by Extract to search for data points in previously unseen documents. There are two types of ML extractors - Quick Learning and Advanced Learning.

Quick Learning is the entry level machine learning method. You can start extracting data with just a single sample.

Advanced Learning is the bigger brother to Quick Learning. It is more sophisticated in the extraction process and it requires at least 30 samples.

Rules-based - These extractors don't require any samples to run. Instead they rely on user inputted search and regex queries to locate and analyse data points from the text of documents.

2. Extractor Library

The **Extractor Library** contains a wide range of pre-trained Machine Learning models, as well as any custom extractors that you have created. You can **sort, filter or view** information regarding your extractors here.

3. Pre-trained Extractors

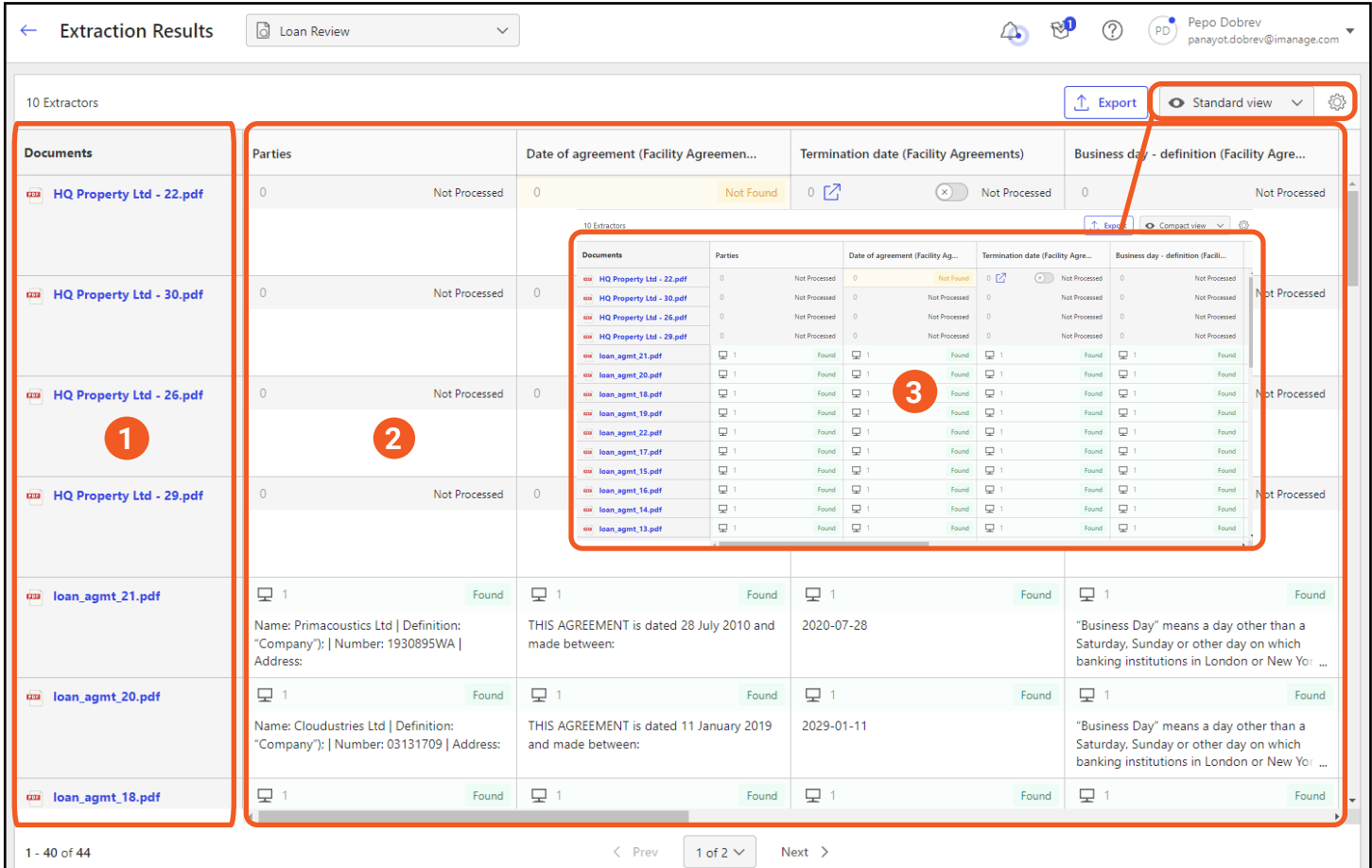
The **Pre-trained extractors** aim to provide users of Extract an extremely quick route into data extraction. They are available to use immediately after installation, so this means that there is little to no time before data extractions can start to happen.

You can use and view the properties of as many Pre-trained extractors as you wish, however you are not able to configure or re-train these extractors.



Extraction Results

The **Extraction Results** page lets you see the extractors and quickly review their extracted data points based on the Review Form all in one place. You are able to see if a data point or an extractor has been **Not Processed**, **Not Found**, **Found**, **Trained** or **Final**. There's also an indication if the data point has been extracted by the **machine**, added by the **user** or both **machine and human**. From this page, you can also **Export** your results.



Extraction Results

Loan Review

10 Extractors

Export Standard view

| Documents | Parties | Date of agreement (Facility Agreements) | Termination date (Facility Agreements) | Business day - definition (Facility Agreements) |
|--------------------------|-----------------|---|--|---|
| HQ Property Ltd - 22.pdf | 0 Not Processed | 0 Not Found | 0 Not Processed | 0 Not Processed |
| HQ Property Ltd - 30.pdf | 0 Not Processed | 0 Not Found | 0 Not Processed | 0 Not Processed |
| HQ Property Ltd - 26.pdf | 0 Not Processed | 0 Not Found | 0 Not Processed | 0 Not Processed |
| HQ Property Ltd - 29.pdf | 0 Not Processed | 0 Not Found | 0 Not Processed | 0 Not Processed |
| loan_agmt_21.pdf | 1 Found | 1 Found | 1 Found | 1 Found |
| loan_agmt_20.pdf | 1 Found | 1 Found | 1 Found | 1 Found |
| loan_agmt_18.pdf | 1 Found | 1 Found | 1 Found | 1 Found |

1 - 40 of 44

1. Documents

View the documents that data points have been **extracted** from. You can also quickly navigate to that document by selecting it.

2. Extracted Data Points

You can see an **overview** of the extracted data points across the documents and the name of the extractor. Hovering over a cell and selecting the data point will direct the user to the right document with the extractor open.

From here, you can also **finalize** an extractor. This prevents further modifications for the extractor and the data point.

3. Table Settings

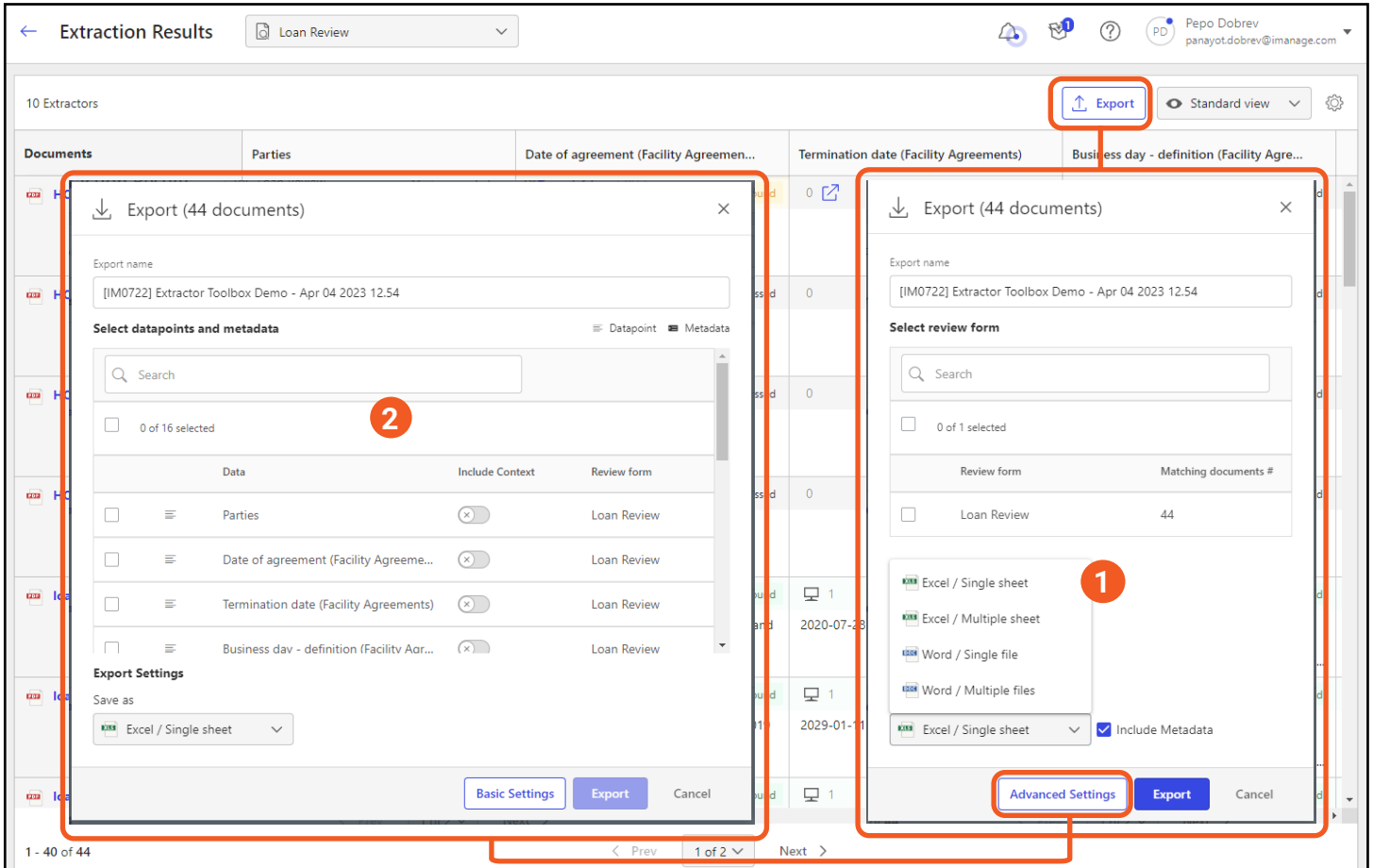
Selecting the **Cogs** icon on the right will let the user select only particular extractors to be shown and hide or show what columns to be displayed.

You can also change the view of the Extraction Results table. Choose between **Standard view** if you would like to view more details or **Compact view** if you would like to quickly review the status of each extractor.



Exporting

In **iManage Extract**, you have the option to **export** all your extracted data points in a structured way that can be used in various ways after to bring you value. You can export your results either from the **Documents** page, by selecting all the required documents and selecting **Export** or through the **Extraction Results** page by selecting the **Export** button on the top right and following the steps. Any export task is being saved in Extract for 30 days after the initial export, so you don't have to export multiple times which saves time on larger exporting tasks.



1. Basic Export

Basic Exporting allows you to select any Review Forms and export all the data points from all the extractor under that Review Form. Select **Include Metadata** check box if you want to include the document metadata in the output.

Results can be exported in either **Microsoft Excel** or **Microsoft Word**.

2. Advanced Export

You can select **Advanced Settings** if you would like to select or deselect particular extractors for your export. This feature gives the user more granular options on what results to be exported.

If you want to include the evidence of particular data point, enable the data point **Include context** slider.

NOTE: You can export data points only if you have the **Reviewer** or **Superuser** role.

| Search Operator | Search String | Result | Method of Operation | Description |
|-----------------|---|--|---|---|
| " " | "credit agreement" | The exact phrase credit agreement without stemming | A search for credit returns files containing the words credits, credited, and crediting. A search for "credit" enclosed in double quotes returns files with the exact term credit. A search for "credit agreement" returns files with the exact phrase credit agreement. | The double quotation marks operator locates files that contain an exact word or phrase. Use it to perform a literal search. You can use double quotation marks with any of the other operators. When you include search terms in quotation marks, iManage Insight removes any stop words (common words like <i>and, or, the</i>), but does not apply stemming . It matches only the exact word that appears within the quotation marks. |
| AND | amendment AND termination | The words amendment and termination | To find files that contain the words amendment and termination , enter amendment AND termination . | The AND operator locates files containing both terms on either side of it. |
| | "lease agreement" AND "maintenance obligation" | The exact phrases lease agreement and maintenance obligation (without stemming) | To find files that contain the phrases lease agreement and maintenance obligation , enter lease agreement AND maintenance obligation . | |
| OR | statutory OR sanctioned | The word statutory or the word sanctioned | To find files with one or both of the words statutory and sanctioned , enter statutory OR sanctioned . | The OR operator finds files that contain at least one of the terms on either side of it. It enlarges the search scope, and is normally used to look for terms that have similar meaning or refer to similar subjects. |
| NOT | "stock purchase agreement" NOT "asset purchase agreement" | The phrase stock purchase agreement but not the phrase asset purchase agreement | To find files that contain the phrase stock purchase agreement but not the phrase asset purchase agreement , enter: "stock purchase agreement" AND NOT "asset purchase agreement". The following example query does not return any results: NOT enactment. | The NOT modifier is used with the AND and OR operators to exclude files that contain a specified term. For the NOT modifier to work, you must first enter a positive search term. |
| ? | pr?scribe | The terms prescribe and proscribe | If you are uncertain whether the term is Prescribe or Proscribe, enter Pr?scribe to return files containing either term variation. | The question mark (?) is a wildcard operator that represents any single alphanumeric character. You can use more than one question mark within a term. Leading wildcards (? at the beginning of the word string) are not supported. |
| * | *ment | All words ending in ment e.g. judgement, amendment, enactment etc. | A search for s*tion returns all files containing words such as section, station, sanction, and so on. A search for *ment* returns files with judgement, amendment, enactment, and so on. A search for *ion matches the words termination, addition, limitation, and so on. | The asterisk (*) is a wildcard operator that represents any number of alphanumeric characters. When using the asterisk wildcard, try to narrow the potential scope of the wildcard entry as much as possible. |