

HULBEE ENTERPRISE SEARCH

User Manual v. 1.07.15.08



Table of Content

1		Introduction						
2		Se	arch s	system Hulbee Enterprise Search – start of work	4			
3		Home page						
4		Search query entry						
5	Search results page							
	5.	1.	Pag	ge functional elements	8			
		5.1	1.1.	Logotype	8			
		5.1	1.2.	Search box	8			
		5.1	1.3.	Profile	.8			
		5.1	1.4.	Toolbar	8			
	5.	2.	Mo	des of search results display	.9			
		5.2	2.1.	List view mode	.9			
		5.2	2.2.	Details view mode	.0			
6		Preview page						
7		Working with search results						
	7.	1.	Vie	wing the search results in browser1	.3			
	7.	2.	Des	sktop Manager utility1	.4			
8		lm	positi	ion of search results limitations1	.5			
	8.	1.	Filt	ers1	.5			
	8.	2.	Dat	taCloud1	.6			
9		Ad	lvance	ed search1	.6			
10	0		User	account1	.8			
	10	0.1.	F	Preferences	.8			
	10	0.2.	(Custom queries1	.9			
	10	0.3.	9	Stopwords1	.9			
1:	1		Searc	h query language1	.9			
12	2		Frequ	ently asked questions (FAQ)2	.1			
4	,		Class	om.				



1 Introduction

Hulbee Enterprise Search (HES) is a search system that provides an opportunity to find the necessary document in the shared data warehouse of an enterprise.

The present Manual is designed to introduce a user to the HES functionality, and its purpose is to help users organize more effective work.

The Manual describes both basic and special features of the search system, such as: digest , DataCloud, filters, advanced search, user's personal settings, search query language.

HES system user interface is browser-based. A user just needs to run a browser on any device and be within the same net with the search server to get access to search functions. To use all functions of HES while working with documents, a user will need to have a Windows-PC, running in the same enterprise domain as common documents of the enterprise.

Requirements to HES client side:

- Windows 7 SP1 or later, with installed updates.
- .Net Framework 4.5.1 or later.
- Latest browser version (FireFox or Chrome are recommended).

HES provides support and processing of different file formats. Formats that are supported¹ by the search engine HES are listed in the table below.

File format	Extension(s)	Extract texts	Extract meta tags	Extract attachments
Text	txt, rtf, doc/dot, odt, wri, sxw	V		
	docx/docm/dotx	V	V	
Publishing	pdf	V	V	
	хрѕ	V		
Hypertext	html, htm, xml	V		
	mht, shtml	V	~	
Table	xsl, xslt, xls, ods, csv	V		
	xlsx	V	~	
Presentation	pptx	V	V	

¹ The search engine doesn't just allow indexing meta information of files (file name, path, size, creation date, modification date), but also working with the content of the file – extraction of text and/or meta tags and/or unzipping archived files, that also contain other files. For a list of formats available to the user with a preview, see section 6 – page preview.

_



	ppt, pps, odp	V		
Graphic	bmp, jpg/jpeg, png, jfif, tif, tiff, jpe		~	
E-mail	msg, eml	V	~	V
Archiv	zip, rar, 7zip			V
Media	avi, mp3, mp4, wav, m4a, wma, wmv,ogg, flac, mkv, ape, mpc		~	
Source Code and Scripting	cs, vb, js, csproj, h, c, cpp, vbs,vcproj, vbproj, pl, sql, bat, cmd	~		
	CSS	~	~	

In order to process certain formats correctly, you must install the free Microsoft IFilter and Adobe.

The system does not process files that are protected by a password or files that exceed the maximum file size specified by the administrator.

The information provided in this Manual is applicable to the HES version 1.7 and it can be modified as new updates become available.

2 Search system Hulbee Enterprise Search – start of work

HES login is carried out by means of account of user, registered in the Active Directory of the enterprise.

Follow these steps to log in:

- 1. Open your browser and type the URL, given by an administrator of the enterprise, in the address bar (e.g.: http://es.company.com/).
- 2. Fill in the login form.
- 3. Confirm an entry.

The login form "Authorization" consists of two standard fields (username and password), entry confirm button "Sign in" and a prompt text for users, concerning possible variants of further actions.

Both fields of the form are mandatory and must be completed. In this respect two variants of log-in are possible: with domain name (e.g.: DOMAIN\user_name), without domain name (e.g.: user_name). In the latter case default domain name, selected by administrator, is used.

In the case of incorrect login or password entry, user receives the notification. In the case of correct login or password entry, user is referred to the HES home page (see 3).



In the version with a limited number of users, you will receive a notification after a failed login attempt. To resolve this issue, contact your network administrator.

Password aging is limited. For this reason, if a user continues his work after some interval of time, he is referred to the page of password entry before continuation of work. After confirmation of the password, the work is continued.

User can click one of the links in the footer of the page and choose language of the present User Manual.

3 Home page

It's possible to use simple and advanced search in the HES system. User can also manage and specify search results through the use of filters and special-purpose tool DataCloud (see 8.2).

On the HES home page one can see standard search box (Fig. 1 HES home page.Fig. 1, a). Using this box, one can carry out a simple search (see 4).

Additional functional elements on the home page:

- Advanced search icon (Fig. 1, b) "Advanced search" (see 9).
- Document type filters (Fig. 1, c) (see 8.1).
- User profile (Fig. 1, d), with the help of which a user is referred to the user account (see 10).
- DataCloud (Fig. 1, e).



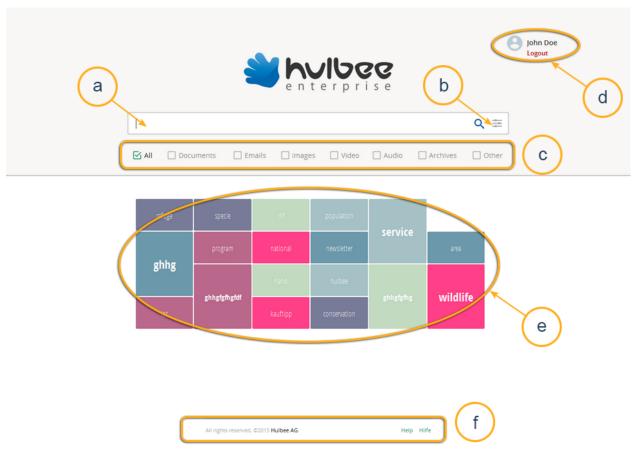


Fig. 1 HES home page.

The use of these functional elements is optional, but all these tools can help organize more effective search.

The footer of the home page (Fig. 1, f) is organized in the same way as footer of the page with the login form.

4 Search query entry

User query might be entered into the search box, located on the home page (see 3), on the page with the list of results (see 5) and on any other page of the HES search system. To start with, please put the cursor in the search bar.

Directly after beginning of entry, drop-down list – $Suggest^2$ (Fig. 2) with adequate variants appears. If there is already query entered before (for example on the page with the list of results), Suggest becomes available after continuation of typing or after placing the cursor in the entry field and clicking the down arrow on the keyboard.



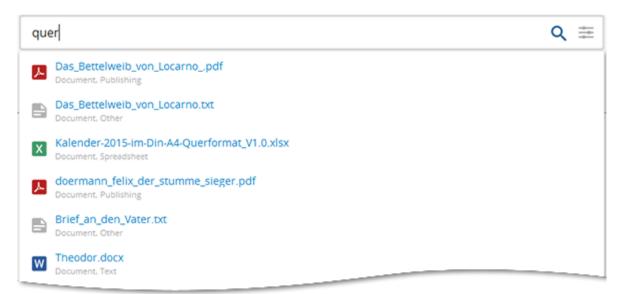


Fig. 2 Beginning of query entry. Suggest appearing.

The following actions are possible at this stage:

- 1. Choose one of the variants from the drop-down list, if it is suitable.
- 2. Ignore the prompt and continue entering your own query.

In the first case confirm the entry by clicking the chosen variant or press <Enter> key. After confirmation, a user is referred to the preview page (see 6).

In the latter case – press <Enter> and click on the loupe icon. In this case a user is referred to the search results page (see 5).

5 Search results page

A user is shifted to the search results page after the start of the search of query, entered by him (see 4). In the process of search, user permissions for documents are taken into account. In other words, user can see only accessed documents in the shared data warehouse.

By default, search results are displayed in the details view mode (see 5.2.2), sorted by relevance ranking, that is to say reflecting the degree of relatedness to the user's search query, from more suitable results to less suitable.

The number of search results is shown in the line "n items found" (Fig. 4, a). If nothing is found, a user sees the notification with the offer to enter another query and null field "0 items found" in the search results counter.



Search results page is divided into logical areas: filters area (Fig. 4, b), described in the section 8, and results display area. It also includes a defined set of functional elements, suitable for correction of search results and for further work with separate documents (see 7).

5.1. Page functional elements

The next sections include the description of search results page elements (hereafter referred to as page) and their functionality.

5.1.1. Logotype

Click on the HES "Hulbee enterprise" logotype shifts a user to the search home page (see 3).

5.1.2. Search box

A user can change the query entered earlier in the search box, specify it manually or with the help of DataCloud tool (see 8.2), enter new search query or activate the Suggest (see 4).

5.1.3. Profile

Profile shows user's personal information, and its fields are links. Click on the user name effects access to the personal account (see 10) link "Logout" shifts a user back to the login form (see 2) for the purpose of changing the profile.

5.1.4. Toolbar

Toolbar (Fig. 3) is designed for setting of search results display.

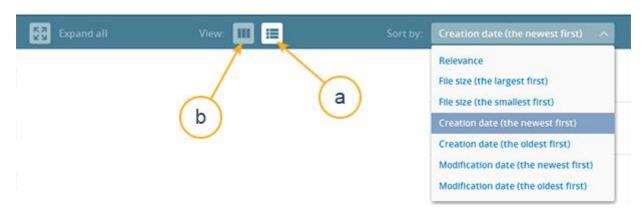


Fig. 3 Toolbar. Search results page.

One can find the following elements on the toolbar:

- "Collapse all/Expand all" option collapses/expands a piece of information, displayed in the search results for each particular document.
- "View" option allows a user to change modes of search results display. Two modes are available (Fig. 3, a, b): list (see 5.2.1) and details view (see 5.2.2).
- "Sort by:..." option helps to display the results by sorting them in a way chosen by a user: sorting by relevance ranking³, file size, file creation date or file modification date update. A user needs to choose one of the variants from the drop-down list (Fig. 3).



The functionality of the elements described in points one and three doesn't depend on the results display mode.

5.2. Modes of search results display

The search results can be displayed in two modes: list view mode and details view mode.

5.2.1. List view mode

This mode allows a user to view the search results quickly. In this case the information on each result is displayed by separate blocks (Fig. 4, c) and includes the name of a document, its location, creation date, size and tags (Fig. 4, d). The tag defines the file type; the second tag specifies the file type for documents, containing text. In the case of text files, the document fragment containing the query or its part is also shown.

Choose the mode "List view" (Fig. 3, a), using component "View" on the toolbar (see 5.1.4) to see listed results.

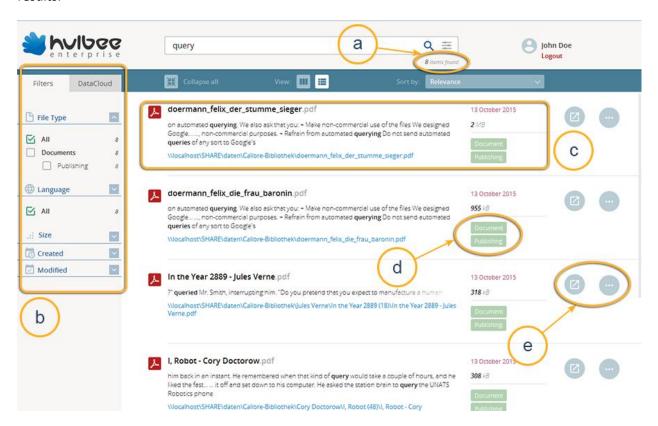


Fig. 4 Page with the list of search results. List view mode.

A user needs to carry out one of the following actions to get detailed information on each result:

- 1. Choose details view mode (see 5.2.2).
- 2. Open the preview page (see 6).



3. Open the document by use of Desktop Manager utility (see 7.2) or open folder with the document.

Special buttons-tools (Fig. 4, e) are designed for it. See section 7 for more detailed information on the work with search results.

In the list view mode a user can't see all the results at once. At the end of the list, a user must confirm the display of the next page with results, by click on the button at the foot of the page.

5.2.2. Details view mode

Choose the mode "Details view" (Fig. 3, b), using component "View" on the toolbar (see 5.1.4) to see detailed results.

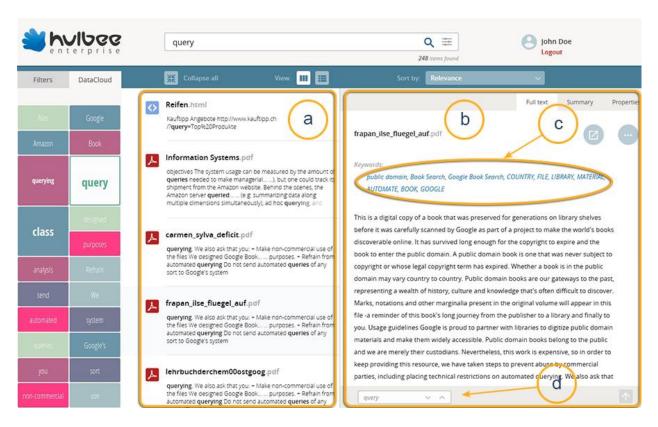


Fig. 5 Page with the list of search results. Details view mode, tab "Full text".

Details view mode combines the list of search results with the preview page. So, in this mode search results are displayed in two columns: the first (Fig. 5, a) represents a list and the second (Fig. 5, b) shows the detailed information on each result in common with the preview page (detailed description in the section 6).

The following actions are available in the details view mode:

- 1. Return to the list view mode (see 5.2.1).
- 2. Jump to the preview page (see 6).



3. Opening of the document by use of Desktop Manager utility (see 7.2) or opening of the folder with the document.

You may find detailed description of these items in the section 7 – working with search results.

For some documents (that contain texts), opened in the details view mode, a body of information could be shown in the tab "Full text" or in the tab "Summary". There are some tools, located at the foot of the page, for such documents navigation (Fig. 5, d). Their description is given in the section 6.

6 Preview page

Shift to the preview page may be caused by the following:

- 1. Selection of a document from Suggest (see 4).
- 2. Click on "Details" (Fig. 6) to see the information on a specific document (from the search results page).
- 3. Click on the block (Fig. 4, c) with information on a document (list view mode).



Fig. 6 Buttons with command options.

The preview page is divided into two columns: the first column represents a list of user's possible actions (Fig. 7, a) and metadata (file details) (Fig. 7, b), and the second replies for document contents.



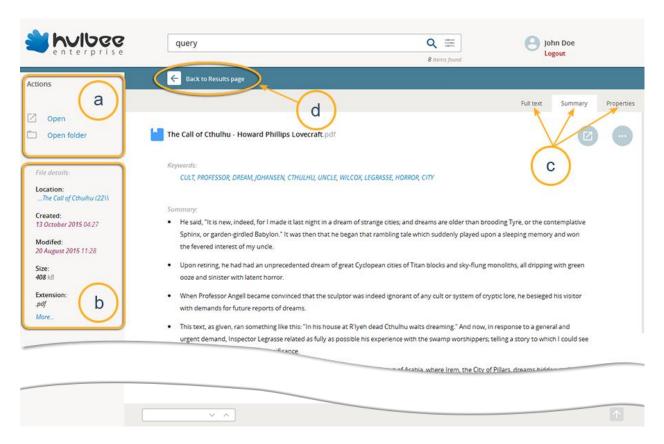


Fig. 7 Preview page. Tab "Summary".

Metadata is data about the location of the document, date of its creation, file modification date, size, extension, language of the document and its type. To see all properties of the document, click on the link "More..." at the end of the list with metadata. You can minimize the list by click on the link "Less...".

The second column represents the information on document contents and includes three tabs (Fig. 7, c):

- Tab "Full text" contains the name of the document, keywords line and text of the document in full or partially, depending on size and file type limits.
- Tab "Summary" also represents the name of the document and keyword line, but full text is replaced by digest – summary of the document in the form of paragraphs. Digest includes the most important sentences to make user understand what the document is about.
- Tab "Properties" gives information on document properties and, partly, duplicates metadata section.

Keywords on the preview page, as well as on the search results page (Fig. 5, c), are links. After the click on the keyword, the system begins to search for it in the text, displayed below. In such a manner keywords help to navigate through the text.

The division of the view area into three tabs is not available for graphic, audio and video files. Metadata tab and preview tab with image preview, internal audio- or video player are available for files of these



types on the preview page and in the details view mode (see 5.2.2). The list of formats for which the preview is available depends on codecs, supported by browser. Generally, this list includes common images (jpg, png, gif, tiff, bmp), mp3 sound files and mp4 video files with codec x264. In other cases, warning concerning inability to show this file type in the browser and suggestion to open it in the proper application on a computer for the continuation of work appears.

Navigation tools are located at the foot of the page (Fig. 7) and are available for documents, containing text. They include word or word group navigation and standard navigation button. A user needs to enter a word or group of words into the appropriate field and he will be shifted up or down the document contents. On default the field contains the text, entered by a user in query. In the process of keyword navigation, the field reflects the word, chosen by a user from the keyword list. To some degree, this navigation is the analogue of the system search, but in the frames of one document.

The preview page contains such functional elements as logotype, search box and profile (see 5.1.1-5.1.3).

The following actions are possible on the preview page:

- 1. Shift to the search results page (see 5) by click on the proper link (Fig. 7, d). In some cases (the direct forwarding of links sent, etc.), the link is missing or leads to the previous page.
- 2. Opening of the document by use of Desktop Manager utility (see 7.2) or opening of the folder with the document of your choice.

7 Working with search results

Documents, found by the HES search system after user's query processing, are available for further work with them, according to the user's access rights. They could be viewed in browser or opened by use of Desktop Manager utility (see 7.2) for processing in the proper application.

7.1. Viewing the search results in browser

Documents view is carried out in any Web browser available. For convenience in viewing, HES system provides a user with different modes of presentation of search results (list view mode, details view mode) and with different modes of presentation of document contents (metadata 4 , preview, full text), which were described above.

Here are HES functional elements, connected with viewing the search results and document contents in browser.





Fig. 8 Block with the information on document (list view mode). Buttons with command options.

Attention is directed to the result a user will obtain, using one or another tool.

- Click on the component "View" on the toolbar (see 5.1.4) changes the modes: list view mode or details view mode.
- Click on the block (Fig. 8, a), containing information on a specific document, shifts user to the preview (see 6) of this document.
- Click on the item "Details" (Fig. 8) shifts user to the preview page.

Opening of the chosen file out of browser is possible using Desktop Manager utility.

7.2. Desktop Manager utility

Desktop Manager utility is developed for reading of files or documents folders, if the opening of the file itself within HES is impossible: for example the file is zipped, file path is too long, etc. It provides an opportunity to work with the original file and not with its copy, downloaded via browser.

To use Desktop Manager do the following:

- Click on the item "Open Folder" (Fig. 8) opens the folder with file.
- Choose item "Open" (Fig. 8).
- Click on the button shown on the (Fig. 8, b).

First of all, the utility gets the username and checks the existence of the user with such name within the system. If the user is detected and has access to the opening file, Desktop Manager calls up an associated program, taking into account settings of the system concerning standard software. For example, *.doc, *.docx files can be opened in Microsoft Word, Open Office, *.txt files – in Notepad, etc. Referring to archives (*.rar, *.zip) and executable files (*.exe), explorer opens the folder, containing them.

Desktop Manager should be started and installed on the client computer individually for each user. Utility has autoloading and it's available right after user login. It doesn't have its own interface, so, if the program is started, user sees the hand icon in the Windows task bar (if you hover a cursor, you'll see the prompt).



If the utility is not installed (a user doesn't see the icon or he is offered to install Desktop Manager), do the following to install it:

- 1. Push the button "Download & Install" in the message window.
- 2. Confirm that you want to download the file by click on "Save File" and open it in the folder.
- 3. Run the executable file.
- 4. Follow the prompts of the installation wizard.

8 Imposition of search results limitations

Search results limitations mean the reduction of quantity of displayed results by elimination of results not matching definite criteria.

These criteria are established by a user with the help of filters (see 8.1), provided by the HES search system, and also with the help of special-purpose tool DataCloud (see 8.2).

There tools are available on the search results page (Fig. 4, b). DataCloud is also available on the HES home page (Fig. 1, e).

8.1. Filters

The possible ways of filtering search results are as follows: filter by type, language, size of the document, date of its creation or update. After the query execution, it is known how many files from the total quantity of search results match one or another criterion. Counters, located in the filters column, display this information.

You can filter search results both by one criterion and by several criteria at the same time.

To filter search results do the following:

- 1. Choose the necessary filter from the filters column.
- 2. Open the list of all possible variants, pushing the button near the filters names.
- 3. Choose the necessary item from the list.
- 4. Continue with another filter, if necessary.

The page reloads automatically and you see filtered results. After using filter, the quantity of search results changes. It is displayed by search results counters on the page and counters in the filters column.

Using filtering by type you may choose subtype for documents, containing text (for example spreadsheets, presentations, etc.).

Using filtering by file size and date you may see special entry fields. You'll have the possibility to choose unit of measurement of file size and to use calendar. Changes should be confirmed or cancelled, if necessary.



Click on the item "All" to see all search results or all the results, filtered by some criterion.

8.2. DataCloud

DataCloud is a set of tiles⁵ of different colors with keywords inside. The colors of tiles reflect semantic similarity of terms.

The HES search system has two DataClouds with different ways of filling.

The first DataCloud is available on the HES home page (see 3). Here, words inside tiles are custom queries (see 10.2), entered into the system by administrator and users. Each individual user doesn't have access to words, entered by other users. When finding lack of custom queries, the cloud adds keywords from documents, containing text. Its elements are ordered at random, occupying the designated area.

When you choose a tile, the search for your query starts and you jumps to the search results page (see 5).

The second DataCloud is located on the "DataCloud" tab on the search results page (Fig. 5). It consists of words-queries (white tiles) and words, which are the most frequently repeated words, chosen from the found documents, excluding stopwords (see 10.3).

By clicking on the query tile, you delete this term from the query and the search starts without it.

This DataCloud helps to specify the search result by adding new words to the entered query.

9 Advanced search

The advanced search offers an opportunity to use further search criteria, narrowing down search results.

To begin the advanced search click on the icon "Advanced Search", located in the search box on any HES page (see 3).

This icon is a link. It becomes active when you hover a cursor over it. You also see the prompt, describing its function.



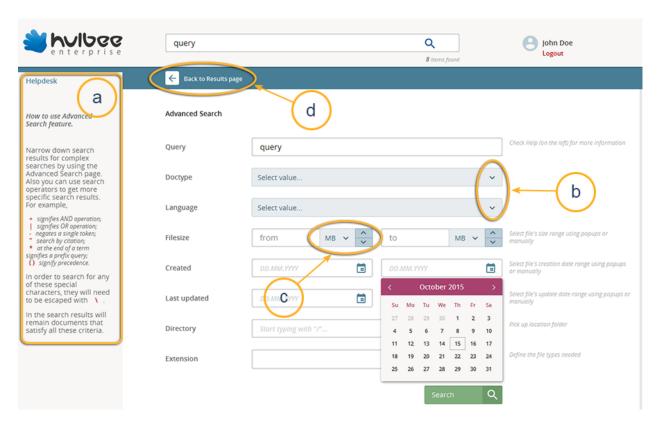


Fig. 9 Advanced search page.

The heading of the advanced search page contains the same functional elements (see 5.1.1-5.1.3) as other HES pages.

The page itself is divided into two columns. The first column contains support information (Fig. 9, a), concerning more effective use of the advanced search (see 11). The second is a form of the advanced search. You may find the following fields in it:

- Query.
- Doctype.
- Language.
- Filesize (you may indicate not the exact size, but a range).
- Created (you may indicate not the exact date, but a range of time).
- Last updated (you may indicate not the exact date, but a range of time).
- Location. File path should be entered in a standard way. Different level folders are separated by slashes. If you start typing with slash (double slash), proposal list will be available. When you type the name of file or folder not adding any symbols, you'll see the documents with this name in path.
- This field should be filled with extensions without leading dot. For example: pdf, txt, doc, etc. You may add more than one extension at the same time. Use <Enter> after adding every extension.



Select document type and language among the variants of drop-down list (Fig. 9, b). You may choose several variants.

There are spinners (Fig. 9, c) in the document size selection field and calendar in the date selection field.

While filling the form, pay attention to text prompts near its fields.

On the advanced search page, as well as on the preview page (see 6), there is a navigation link (Fig. 9, d), which shifts you to the search results page (see 5).

If a user has left the advanced search page, he has the possibility to return and edit or complete the previously entered data. They are not only stored, but also displayed on the search results page.

10 User account

Click on your name in any HES page profile to go to your personal account.

The profile contains the "Logout" link, which shifts you back to the login form (see 2) to change user account, and also user's personal information. It includes username, taken from enterprise Active Directory.

With the help of personal account, a user can manage his personal settings, provided by the HES search system. All variants of settings are listed in the left column and are called "My Settings".

10.1. Preferences

In the "Preferences" section you may choose the interface language. Select the necessary language from the drop-down list. Then save the setting.

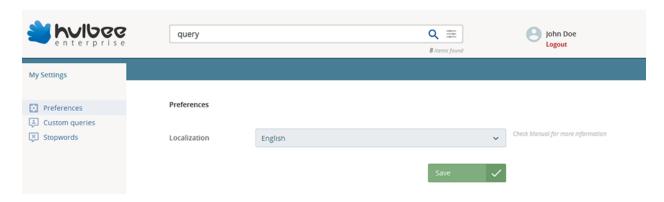


Fig. 10 User account. "Preferences" section.



10.2. Custom queries

"Custom queries" allow you to add keywords, which later will be reflected in the DataCloud on the home page, and which help to enter typical search queries quickly.

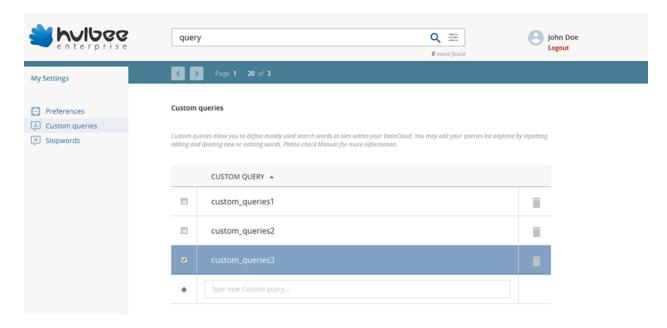


Fig. 11 User account. "Custom queries" section.

Using keyboard, you may type custom query in the line on the right. Then press <Enter> and continue typing, if necessary. You may add, edit and delete custom queries.

There is also custom query navigation - page scrolling (if you have more than one page).

If you enter more keywords than DataCloud can hold, part of them will not be displayed (at random).

10.3. Stopwords

Using "Stopwords" tab you may indicate words, which you don't want to see in DataCloud, located on the search results page. For example: the name of the user's company. It can be found almost in every document, for which reason it's useless for query specification.

Work with stopwords list is similar to the work with search query list in the "Custom queries" tab.

11 Search query language

Support information contains the list of search operators – special symbols and (or) keywords, added to the query for specifying search results or used as independent queries.



To see all variants of possible search queries, visit the advanced search page (see 9). They are displayed in the first column of the page.

Here are the examples of formation of queries through the use of search operators. Words for query example: "world", "peace", "war".

- world + peace operator "+" signifies "AND" operation and searches for documents containing words "world" and "peace" at the same time; the standard query with "+" is generally equivalent to the entered query: world + peace = world peace;
- world | peace operator "|" signifies "OR" operation and searches for at least one of these words:
- world -war operator "-" gives the possibility to show the results containing the word "world" and excluding the word "war";
- "world peace war" if the phrase is quoted, words in the search results will be ordered in the same way;
- "world peace war"~N the query in quotation marks is too limited because it is consecutive words in the document. By "~ N" the search will be extended, so that keyword matching results can still be found, where N stands for the maximum number of possible variations of the search term (insert, delete, rearranging of words) that can be inserted so an accurate search result can be displayed.
- hullbe~N fuzzy search; "N" stands for the maximum number of possible variations of the search term (insert, delete, rearranging of words) that can be inserted so an accurate search result can be displayed. For instance if we don't know exactly how to spell the brand "Hulbee", even when searching for "hullbe~2", the correct search result will be displayed;
- pe* operator "*" substitutes any unknown word ending in the query; in this case the search is carried out only by known fragment; for example: pe* - pen, person, peace, etc.;
- (world -war) | (world -peace) operator "()" is like mathematical parentheses; they arrange order of use of logical operators; used for complex query formation.

If a user wants to use one of the above mentioned symbols in the proper sense, "\" symbol should be typed first.

There are some more search queries, beginning with ":". All of them become available in the drop-down list after typing ":" symbol in the query line. You will also see a short description of such queries and their entry format.

- :all all files accessible for a particular user are shown; the same query returns, for example, all text files, if a limitation on type of displayed files is placed;
- :doctype: returns documents belonging to one or several doctypes, e.g.
 :doctype:Document,Image,Video;
- :lang: returns documents in one or several languages, e.g. :lang:en,de;
- size: returns documents whose size in bytes satisfies the condition, e.g. :size:0-1024;



- :created: returns documents whose creation date satisfies the condition, e.g.
 :created:01.01.2015-31.01.2015;
- :updated: returns documents whose modification date satisfies the condition, e.g.
 :updated:01.01.2015-31.01.2015;
- :extension: returns documents of one or several extension, e.g. :extension:.doc,.pdf;
- :path: returns documents located in a specific folder, e.g. :path:/public/docs.

All queries listed begin with a colon, this format should be maintained without spaces or punctuation marks at the end of the query.

12 Frequently asked questions (FAQ)

- 1. I can't log into system.
 - Try to log in, typing domain name before username.
 - Contact your system administrator.
- 2. I can't find the necessary document.
 - Check your right of access to this document.
 - Use "Advanced search".
 - Use query language.
- 3. Custom gueries are not shown in the Cloud.
 - Custom queries are shown in the Cloud at random, as the Cloud has limited area.
- 4. I can't open documents in Desktop Manager.
 - Check if Desktop Manager is installed for this user.
 - Check if Desktop Manager is started (if not start it manually).

13 Glossary

Digest – automatically formed text information, briefly describing the text. It consists of automatically generated keywords (word groups) and most information-rich sentences of the text.

Metadata – for HES it means the name of file, date of its creation and modification, size. For many formats meta tags are also available (author, type of camera, organization, etc.).



Relevance – the degree of relatedness of search results to the user's search query. The search engines use different methods for determining the relevance, which aim to match the query entered by the user. Accordingly, once sorted by relevance, the first results should be the most relevant for the user.

Suggest – drop-down list of documents, corresponding to the entered search query.

Spinner – a set of control elements, which help to increase or decrease the value in the entry field by means of the mouse.

Tile – the minimal DataCloud functional element.