

Short Instruction for Search and Find with InfoCodex

Start URL on the demo server: yourservername/infocodex6.html e.g. <https://demo.infocodex.com>

After the login you will see the search mask where you can select a document **collection**.

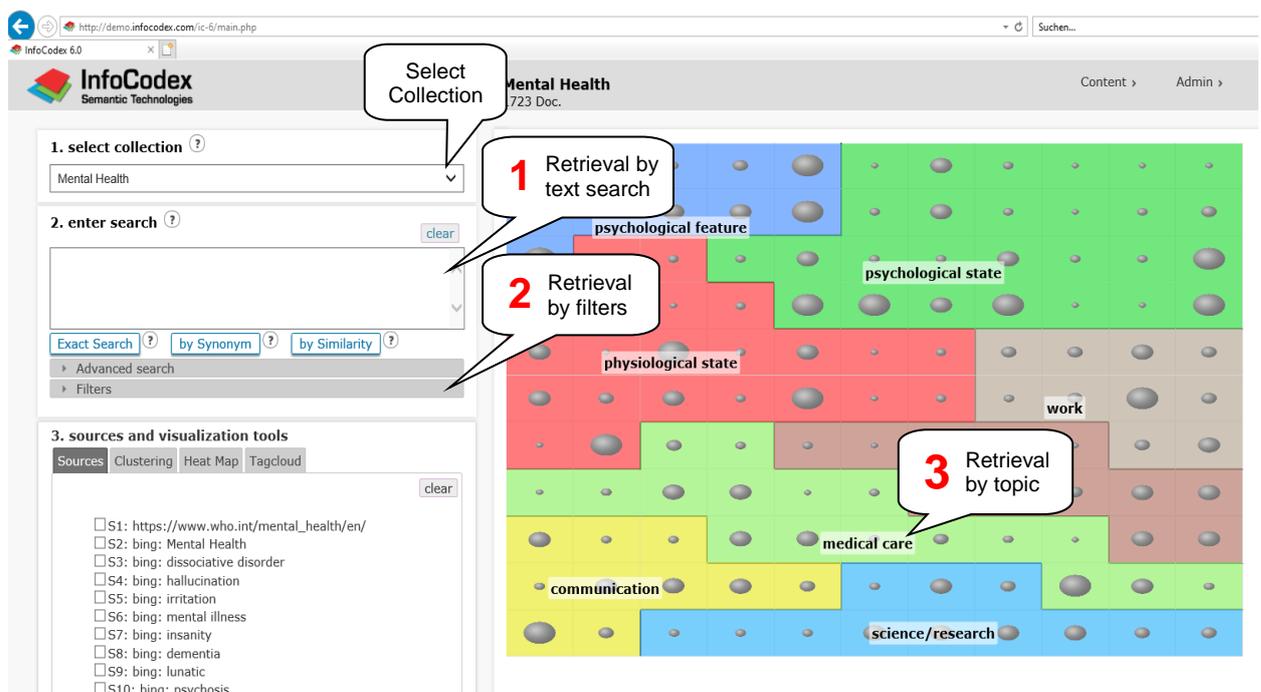
A collection is a set of documents gathered from various document sources: File server, Internet (web sites, results of Internet search engines), e-mails and local folders.

InfoCodex supports different search strategies (isolated or combined):

- 1. Search with keywords or any free text in three different ways:**
 - Exact search corresponding to the traditional full text search
 - Synonym search semantic and cross-lingual search (EN/DE/FR/IT/ES)
 - Similarity search thematic comparison between the query text and the existing documentsAdditionally, search by document properties such as author, date range or document source.
- 2. Retrieval by filters**

In addition to the standard filters (sources, persons, organizations, locations), the optional setting of other filters is supported (customer IDs, concept filters etc.).
- 3. Retrieval by thematic topics (clusters or heat map)**

When documents are imported into a collection, InfoCodex's text analysis module understands the content of the documents well enough so that it can automatically categorize the documents by thematic topics and arrange them into logically structured "book shelves". This enables a targeted search which is particularly helpful in the case of large result lists.

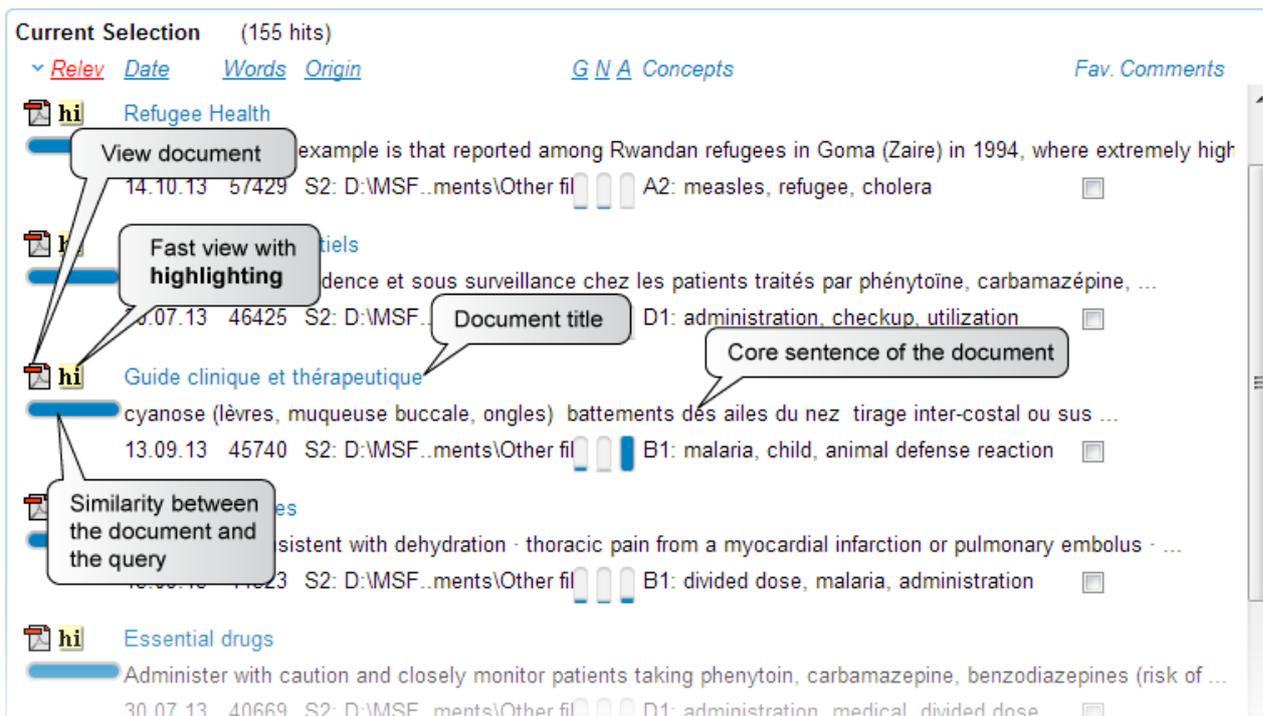


The screenshot displays the InfoCodex 6.0 search interface. On the left, there are three main sections: '1. select collection' with a dropdown menu showing 'Mental Health'; '2. enter search' with a text input field and a 'clear' button; and '3. sources and visualization tools' with tabs for 'Sources', 'Clustering', 'Heat Map', and 'Tagcloud'. Below these are search method buttons: 'Exact Search', 'by Synonym', and 'by Similarity'. A list of sources (S1-S10) is shown at the bottom left. On the right, a 'Heat Map' visualization shows a grid of colored cells representing thematic topics. Three callout boxes point to specific features: '1 Retrieval by text search' points to the search input; '2 Retrieval by filters' points to the search method buttons; and '3 Retrieval by topic' points to the 'Heat Map' visualization. The visualization includes labels for 'psychological feature', 'psychological state', 'physiological state', 'work', 'medical care', 'communication', and 'science/research'.

Search Result List

The found documents are listed in the lower part of the screen. The assessment of the hits is facilitated by the following means:

- The documents are ordered by **similarity** with the query text (objective ranking on the basis of the patented similarity measure of InfoCodex).
- The similarity is graphically visualized with the horizontal bar: 
- For each document, the **title** (in blue) and the "**core sentence**" of the document are displayed.
- This information is supplemented by the **document date**, the **number of words**, and the three most important **concepts** (main thematic topics).
- With a click on the document title the automatically generated **abstract** (summary) and the 18 most relevant **descriptors** (keywords) are shown. This facilitates the browsing of search result lists and saves time in finding the relevant documents.
- To show the original document click on the document symbol 
- With the symbol **hi** a fast HTML view of the document with **highlighted** search terms is displayed.
- By clicking the similarity bar  you will get a **visual representation of the similarity** between the document and the query text.
- The symbol ***2** below a similarity bar indicates that two very similar documents are summarized into a **document family**
→ substantial reduction of document lists, because multiple (almost identical) documents occupy only one position in the search result list.



The screenshot shows a search result list with the following elements and callouts:

- Current Selection (155 hits)**: At the top left.
- Sorting options**: *Relev*, *Date*, *Words*, *Origin*, *G N A Concepts*, *Fav. Comments*.
- Document 1**: Title "Refugee Health", date "14.10.13", words "57429", source "S2: D:\MSF...ments\Other fil", concepts "A2: measles, refugee, cholera". Callout: "View document".
- Document 2**: Title "tiels", date "0.07.13", words "46425", source "S2: D:\MSF...". Callout: "Fast view with highlighting".
- Document 3**: Title "Guide clinique et thérapeutique", date "13.09.13", words "45740", source "S2: D:\MSF...ments\Other fil", concepts "B1: malaria, child, animal defense reaction". Callouts: "Document title" (pointing to the title), "Core sentence of the document" (pointing to the core sentence).
- Document 4**: Title "es", date "13.09.13", words "45740", source "S2: D:\MSF...ments\Other fil", concepts "B1: malaria, child, animal defense reaction". Callout: "Similarity between the document and the query" (pointing to the blue bar).
- Document 5**: Title "Essential drugs", date "30.07.13", words "40669", source "S2: D:\MSF...ments\Other fil", concepts "D1: administration, medical, divided dose".