



CHANL Workbench User Guide (V 1.2)



Tianrun Cai, M.D.

Document Version: V 1.1
Software Version: 2019-Fa-1

Credits

Software Designer

Tianrun Cai, M.D.

Software Developer

Tianrun Cai, M.D.

Software Graphics

Tianrun Cai, M.D.

Software Contributor

Bryan Cai, PhD (Stanford University)

User Guide Author

Tianrun Cai, M.D.

User Guide Editor

Amanda King, MS (T.H. Chan Harvard School of Public Health)

Jacqueline Honerlaw, MPH (VA Boston Healthcare System)

Splash Graphics

Tianrun Cai, M.D.

Acknowledgement

Tianxi Cai, ScD

Katherine Liao, M.D., MPH

About the Author

Tianrun Cai, M.D., serves as a health informatician within the Division of Rheumatology, Immunology, and Immunity at Brigham and Women's Hospital. Initially trained in medicine, he transitioned to full-time research in health informatics. Over many years, he has specialized in designing and developing software and algorithms aimed at translating complex and unstructured clinical data from narrative reports into structured and analyzable formats. Among his notable contributions are the development of several NLP software applications for various purposes:

- EXTEND (Extraction of EMR Numerical Data), designed to extract numerical data such as vital signs and ejection fraction from electronic medical records.
- NICE (NLP Interpreter for Cancer Extraction), tailored for extracting cancer-related information such as clinical cancer stage, TNM stage, and historical data.
- The SMALL algorithm, an Ensemble Machine Learning tool crafted to assist with eligibility screening for patient recruitment in clinical trials.

Furthermore, he possesses expertise in GUI (Graphical User Interface) development using WxPython, encompassing both back-end and front-end software design, prototyping, testing, and debugging. His diverse projects span across different medical disciplines, including cardiology, radiology, and rheumatology, demonstrating a commitment to enhancing the efficiency of medical research through innovative approaches in natural language processing (NLP).

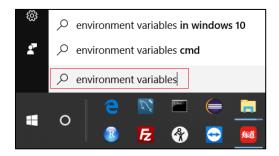
Section I. Set Environment Variables

I-1. Open Environment Variable Editor

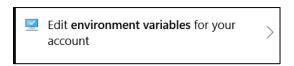
Note: The icons and menus are from Windows 10. In different versions of Windows such as Windows server or Windows 7, 8 some icons, menus or positions might be different from those displayed below.

I-1-1. From "Start Menu"

- 1. Click the search button on the "Start" menu or taskbar.
- 2. Search for "environment variables"



3. Click on the option "edit environment variables for your account" to open the environment variable editor.

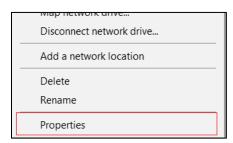


I-1-2. From "This PC"

1. Right click "This PC",



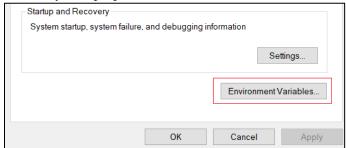
2. choose "Properties" on the pop-up menu



3. then click "Advanced system settings" in the "System" window.

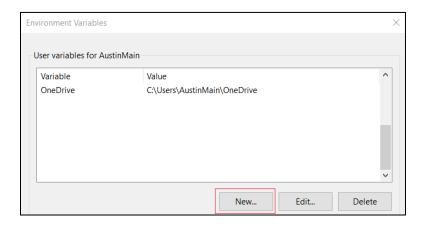


4. In the "System properties" window, on the "Advanced" tab, click "Environment Variables".

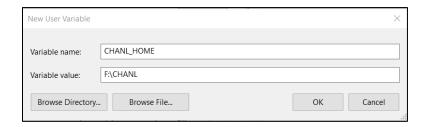


I-2. Create a New Environment Variable

1. Click the button "New"



2. Then enter "CHANL_HOME" for the variable name and enter in the file path of the CHANL folder in the "Variable value" field. E.g. "F: \CHANL"



3. Log out or sign out of the current Windows user (instead of "disconnect"), then log-in again.

I-3. Launch CHANL

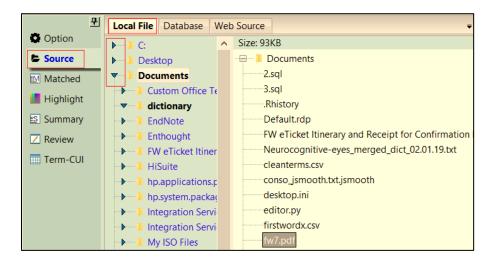
Navigate to the CHANL folder and double click the icon CHANL or click on the shortcut on the desktop to launch CHANL.

Section II. Load Source Data

Note: Currently, there are two options available: Local file and Database.

II-1. Load Data from Local Files

1. Select "Source" in the middle of the CHANL interface, click the "Source" tab, and click the "Local file" tab, browse folders by clicking or double click folder names.



2. In a target folder choose a file to load to CHANL.

Currently, CHANL supports plain text file such as '.txt', '.csv', '.py' etc. and PDF files. CHANL does not support images, '.doc' files and zipped files.

The file size allowed is 40Mb. If the size of a file is greater than 40Mb, just the top 20,000 rows will be displayed. This feature could be used to check a sample of huge txt files.

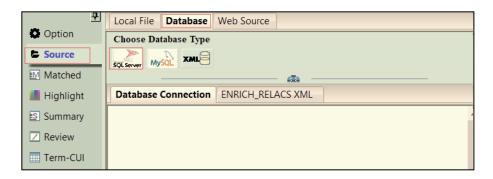
For PDF files, CHANL will automatically ignore non-text content and non-ascii characters.

II-2. Load Data from Database

II-2-1. Database Connection

II-2-1-1. MSSQL

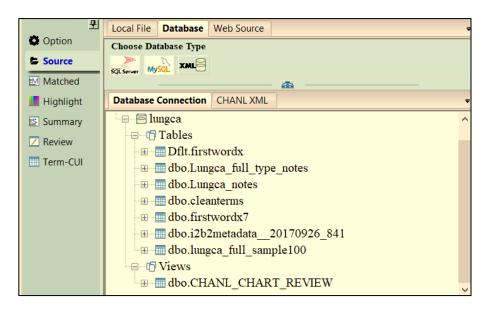
1. Select the "Source" tab in the middle of the CHANL interface, click the "Database" tab, then choose a database type.



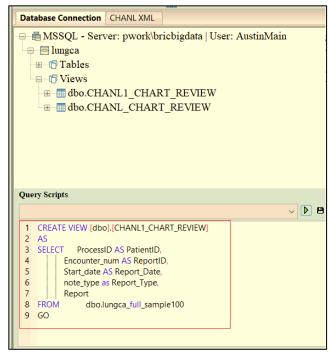
2. Click , and the MSSQL database connection window will pop up.



- 3. Provide "Server" name, then choose "Authentication Type". If "SQL Server Authentication" is chosen, please enter "User Name", "Password" and "Database" name. If "Windows Authentication" is chosen, "User name" and "Password" are not needed.
- 4. Click "OK" to launch the connection to the target database. Once the connection is built, see the example below:



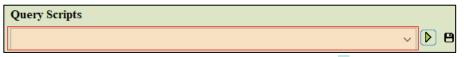
5. In order to have better performance, we recommend building a view to customize the order of column names by "PatientID", "ReportID", "Report_Date", "Report_Type", "Report". We can create a view in CHANL directly using the example SQL script below (also see example SQL scripts in Appendix 2.), when a connection to a specific database is built. Please re-connect the database after creating a new view.



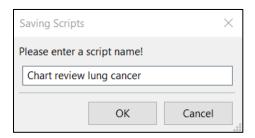
6. Enter or choose SQL scripts in the "Query Scripts" area then click to query patient data to load to CHANL. See an example below:

select * from dbo.CHANL CHART REVIEW where processID in (1,3,100)

We provide a few query examples which are listed in the dropdown menu below.

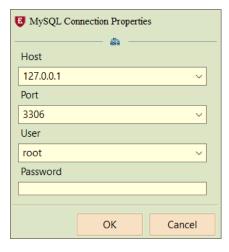


7. Scripts can be saved to CHANL for later use by clicking , then enter a name for scripts on the popup window and click "OK":

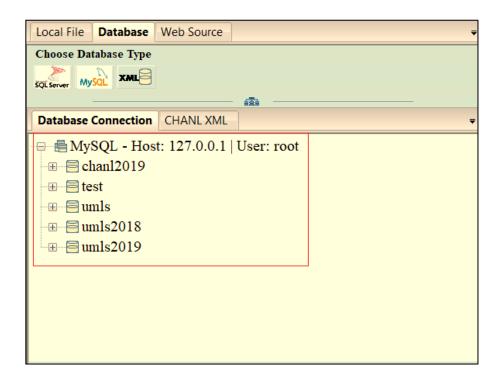


II-2-1-2. MySQL

1. The MySQL database connection window will pop up if we click the button



- 2. Enter information for "Host", "Port", "User" and "Password".
- 3. Click "OK" to launch the connection to MySQL database. For MySQL connection, we don't need to specify the database name.
- 4. After the connection is built, unlike the MSSQL database connection in CHANL, the MySQL connection will show different databases. See below:



5. Browse a database and choose a target table to query data and load to CHANL. Note: the gramma of query scripts in MySQL is different from MSSQL. Please change the scripts to fit the relevant database.

II-2-1-3. XML Database

II-2-1-3-1. Generating XML Database File

Here XML databases are XML files saved in a local or shared drive. In order to import XML files to CHANL interface as data sources, we need to convert plain text files to a specific format of XML files shown below:

```
▼<CHANL_xml>

▼<ROW>

<COLUMN NAME="PatientID"> 10204 / COLUMN>

<COLUMN NAME="ReportID"> 38383032 / COLUMN>

<COLUMN NAME="Report_Date"> 2010-05-06 00:00:00 / COLUMN>

<COLUMN NAME="Report_Type"> LNOK / COLUMN>

<COLUMN NAME="Report"> Subject: [Radiology Report:Scanned] [report_end] </COLUMN>

</ROW>
```

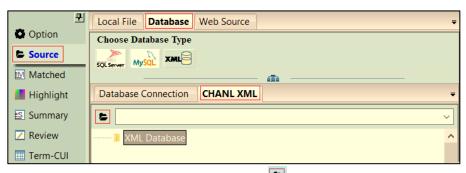
The format of xml files should include the root "CHANL_XML", the first level of node "ROW" and the second level of node "COLUMN NAME". Each "ROW" contains information for each note. Five "COLUMN NAME" values are needed for CHANL including "PatientID" (patient identifier), "ReportID" (note identifier), "Report_Date" (date for the note), "Report_Type", and "Report" (the content of a note). Please see the script in Python in **Appendix 1**. to create XML files using database or local txt files.

When generating XML database, we recommend not exceeding 100 patients for each file.

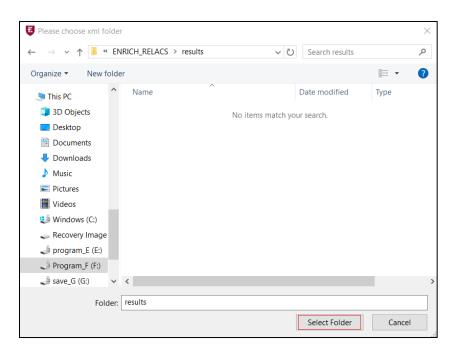
II-2-1-3-2. Load the Folder of XML

1. Select "Source" in the middle of the CHANL interface, choose the tab "Database" and click the "CHANL XML" tab.

2.



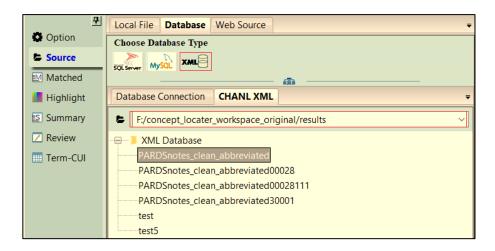
3. Click "Select Folder" on the popup window to browse the computer to select a folder which contains xml files.



4. The path of the folder will be shown F:/concept_locater_workspace_original/results

II-2-1-3-3. Load XML Files

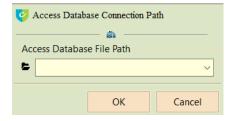
Choose a file under MML Database by clicking the file name and clicking to build the connection to the XML database.



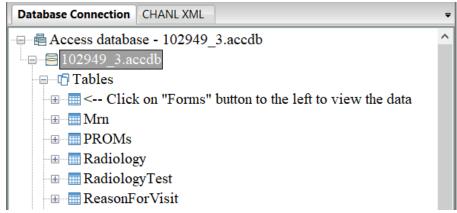
II-2-1-4. Access Database

1. Click the button to open a window for choosing an Access database file path.





- 2. Click the button on the popup window to choose Microsoft Access database file with extension name ".mdb" or ".accdb".
- 3. Click "OK" to connect the database. The tables and columns for each table will be listed in the tab "Database Connection"

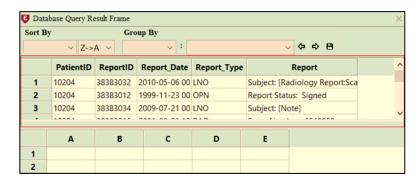


4. Data query gramma is similar as MSSQL. Please read the relevant part for MSSQL above.

II-2-2. Data Selection

II-2-2-1. Query Result Display

The data will be displayed in the upper grid in a popup window when it's loaded from a database. Note: this popup window will always stay on top of opened windows.

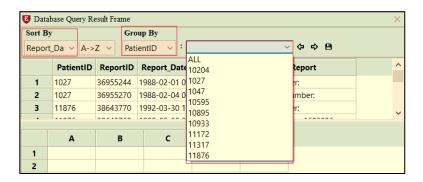


II-2-2-2. Sort Data

Data can be sorted by one of the columns loaded by ascending $(A \rightarrow Z)$ or descending $(Z \rightarrow A)$ order.

II-2-2-3. Data Grouping

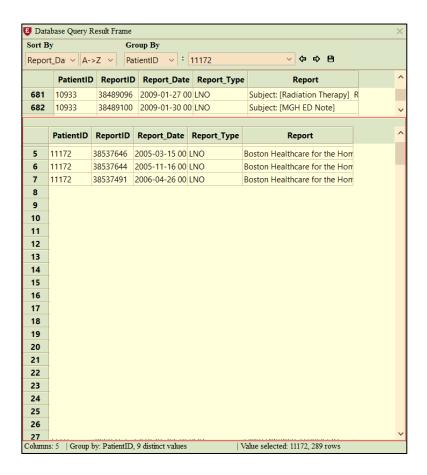
Before loading data to the display panel, we need to group data by a distinct value of a certain column. For example, group data by distinct "PatientID" values.



II-2-2-4. Select Data to Load

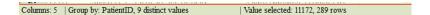
Choose a value of "PatientID" from a dropdown menu beside "PatientID"

Then all data for the specific value chosen above will be loaded to the lower grid in the popup window.



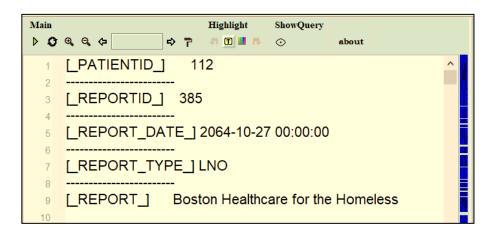
II-2-2-5. Data Information Display

With column number in the database table chosen, "Group by" column name. The count of the distinct value of the "Group by" column name, the selected value and the count of the rows will be displayed on the bottom of the window.



II-2-2-6. Load Data to Main Panel

When the data is selected in step II-2-2-4, the data will be reformatted and loaded to CHANL.



II-2-2-7. Automatic Hiding

Drag the title bar of the "Database Query Result Frame" to the edge of the screen (left, right and top), then move the mouse off. The frame will be automatically hidden.



II-2-2-8. Recall the Data Frame

Put the mouse on the edge of the screen when the frame is hidden or click the button on the main panel, and the frame will be displayed again.



Section III. Data Search Options

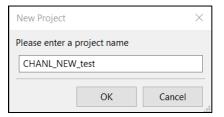
III-1. Project

III-1-1. Create a Project

1. Select the "Option" tab in the middle of the CHANL interface, then click on the top right.

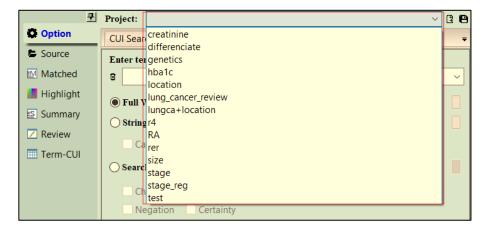


2. Enter a project name in the popup window to create a new project.



III-1-2. Load a Project

1. From the project dropdown menu select a project to load.



III-1-3. Save a Project

1. A project will be saved automatically after running any search, but you can save a project with another name by clicking on the top right.

III-1-4. Delete a Project

1. Select a project, right click the project name, then click "Delete this project" on the popup menu to delete an existing project.

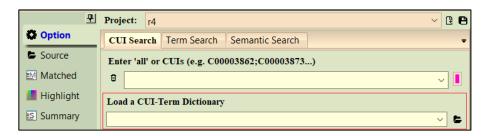


Note: Before doing any type of search, notes need to be loaded and a project needs to be selected or created.

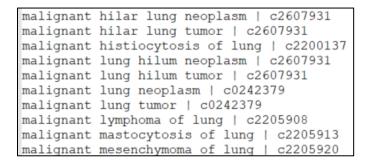
III-2. CUI Search

III-2-1. Load a Dictionary

1. On the "CUI search" tab, click to browse the computer to load a pre-built CUI-term dictionary.

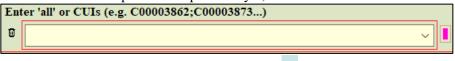


2. The format of the dictionary should be a term and a CUI in each line separated by "|". See the example below:



III-2-2. Enter CUIs

1. Enter a CUI or multiple CUIs separated by ";" or enter "all" to search all concepts in the dictionary.



2. Select a CUI from the dropdown list, then click on the left to delete the CUI from the list.

III-2-3. Select Color

1. Click on the "CUI Search" tab to select a color from a popup window below for CUI Search.



III-2-4. Run Search

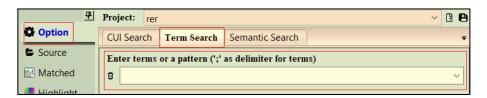
Click on the "Main" panel to run search.



III-3. Term Search

III-3-1. Enter Terms

1. Enter a term or pattern or multiple terms separated by ";" in the space below.

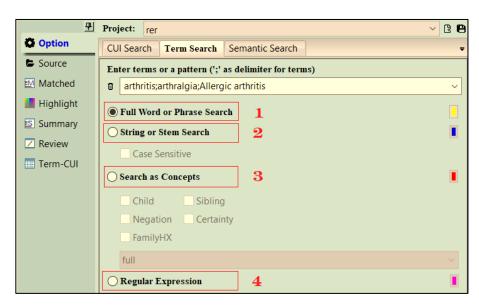


III-3-2. Delete Saved Terms

1. Choose a record from the dropdown list then click to delete the record from the list.

III-3-3. Choose Term Search Type

There are four options for performing term searches: Full Word or Phrase Search, String or Stem Search, Concept Search and Regular Expression.



III-3-3-1. Full Word or Phrase Search

This setting will search a term entered as a word with a boundary and ignore case.

E.g. If we search a word "arthralgia" in two sentences:

- a. he had arthralgias yesterday. (No match)
- b. he had arthralgia yesterday. (1 match)

III-3-3-2. String or Stem Search

This setting will not consider boundaries. Case will be considered if Case Sensitive is checked. E.g. If we search a word "arthralgia" in two sentences below:

- c. he had arthralgias yesterday. (1 match)
- d. he had arthralgia yesterday. (1 match)

III-3-3-3. Concept Search

With this setting, CHANL will search a term entered as well as all the synonyms of the term. Concept Search is not case sensitive. E.g. If we search a word "arthralgia" as a concept, CHANL will include all the synonyms and lexical variants such as "joint pain", "painful joint", "painful joints", "arthralgias", etc. in the search list. There are more options when performing Concept Search.

III-3-3-3-1. Child Concept

If Child is checked, all the sub-concepts of the concept entered will be included in the search list. E.g. If we search a word "arthralgia" as a concept including sub-concepts, in addition to all synonyms of arthralgia, all synonyms of sub-concepts such as "shoulder pain", "hip pain", "joints stiff" will also be included in the search list and highlighted.

III-3-3-3-2. Sibling Concept

If Sibling is checked, all the direct sibling concepts of a concept entered will be included in the search list and highlighted.

III-3-3-3. Negation Information

If Negation is checked, all the terms or phrases with negative meanings such as "no", "hasn't", and "deny" will be included in the search list and highlighted.

III-3-3-3-4. Certainty Information

If Certainty is checked, all terms or phrases about certainty such as "likely" and "possible" will be included in the search list and highlighted.

III-3-3-3-5. FamilyHX

If FamilyHX is checked, all terms or phrases about family members will be included and highlighted.

III-3-3-4. Regular Expression Search

CHANL can also search terms using Regular Expression. Please check the menu here for generating a regular expression.

III-4. Semantic Type Search (In development)

III-5. Similarity Search (In development)

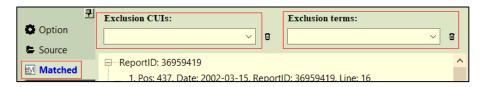
III-6. Color Customization

For each search option, there is a button • for choosing a color for highlighting.



III-7. Term and Concept Exclusion

When performing a concept search or CUI search, we can exclude certain terms or concepts.

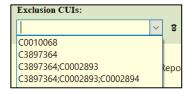


III-7-1. Enter Exclusion CUIs or Terms

Before running a search, please choose the tab "Matched" in the middle of the CHANL interface. Enter CUIs separated by ";" or terms separated by ";". All the terms of the CUIs will be excluded from the search list.

III-7-2. Choose CUIs or Terms

Click the relevant dropdown list to choose a record.



III-7-3. Delete CUIs or Terms

Choose a record from the dropdown list, then click follow to delete a record from the list.

Section IV. Searching Result Representation

IV-1. Main Panel

IV-1-1. Run a Search

After choosing or creating a project and loading notes to the main panel, click on the main panel to run a search to highlight target terms.

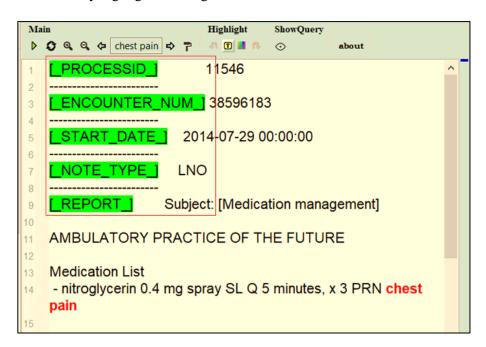


IV-1-2. Reset Highlighted Terms

Click to remove the highlighting for terms.

IV-1-3. Column Name Highlighting

When loading data from a database query or correctly formatted XML file, the column names will be automatically highlighted with green.



IV-1-4. Zoom In and Zoom Out

Click to increase the font size of the loaded notes. Click to decrease the font size of the loaded notes.

IV-1-5. Matched Location

Click to iterate the matched terms in the notes in a backward direction and click to iterate the matched terms in the notes in a forward direction. The notes will scroll to where the current term is located. The current term will be shown like this:

IV-1-6. Format Report

Click the button to format the report by replacing multiple spaces with a single space, multiple returns by a single return. (Note: reformatting is not applied to the leading spaces of each line).

IV-1-7. Manually Highlighting

1. Click the toggle button to enable "Manual Highlighting". Then click the button to choose a color for manual highlighting.



- 2. On the note, hold the left button on the mouse and select text to highlight.
- 3. Click the button to undo a highlight and click to redo a highlight when they are not greyed out.

IV-1-8. Recall Query Result Frame

Click the button to recall the data query frame if the data is loaded from a database query or the CHANL format of XML files.

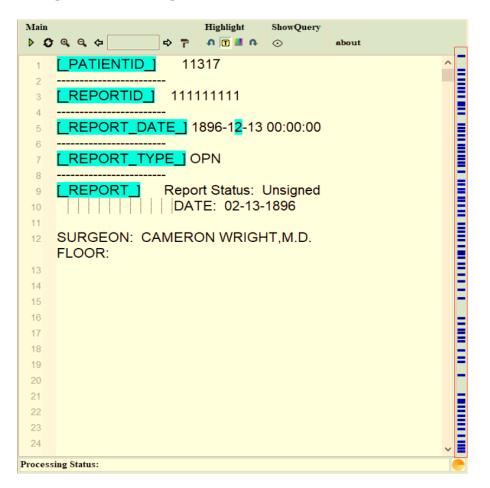
IV-1-9. About CHANL

Click the button about to view the version and Author information of the software.

IV-1-10. Switch Report

IV-1-10-1. Locate Report Switching Bar

When data is loaded from a database query or XML file, there will be an extra information bar beside the scroll bar on the main panel. The information bar consists of many blue squares —. Each blue square stands for a report.



IV-1-10-2. Display Report Header

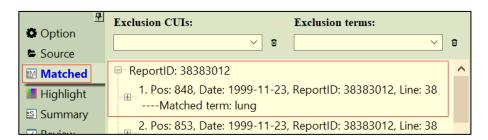
Hover the mouse on a blue square , and there will be a tooltip window displaying the header information of the report.

IV-1-10-3. Switch Report

When clicking any of the blue squares, the displayed content will switch to the report clicked.

IV-2. Detailed Matching Information

The detailed information for matching terms in notes will be displayed in the tab Matched. The order of the matched information listed in this tab will be the same as the information highlighted in the main panel. When clicking on a certain matched item, the content in the main panel will switch to the matched position.



IV-2-1. Information Listed

The information in each item listed in the tab includes index, the matched sequence position, report date, reportID, line number and matched term. Matching information for different reports will be grouped for each report and listed under different nodes if the initial data is from a query or XML file.

IV-2-2. Collapse and Expand Matching Information

Click [™] to expand or [□] to collapse the report node.

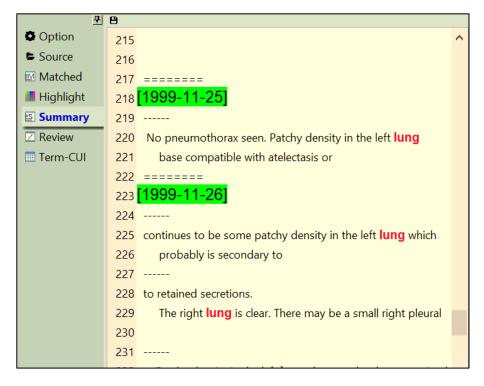
IV-2-3. Display Matching Sentence

Click ⊕ to expand an item to display the matched sentence. Click ⊨ to collapse the item.

IV-3. Search Result Summary

IV-3-1. Summary Data Format

In the tab Summary, the matched sentences will be listed as a group for a distinct date. The dates will be ascending. Different dates are separated by ======= and different sentences are separated by



IV-3-2. Color

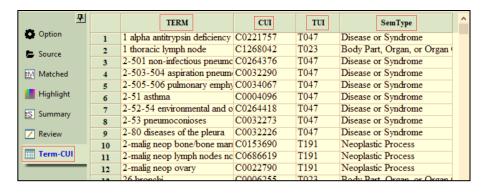
Dates are highlighted with green. Matched terms will be highlighted with red. Negation terms will be highlighted with blue and family members will be highlighted with brown.

IV-3-3. Reflect to the Main Panel

Double click each sentence in the summary tab, the main panel will switch to the relevant position where the sentence is located.

IV-4. Concept Dictionary

IV-4-1. Data Format



IV-4-2. Sort data

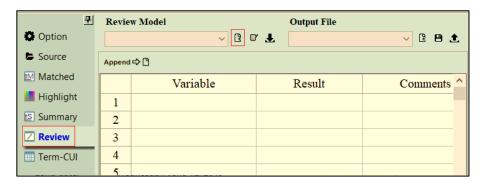
Click on each column name, and data will be re-sorted by relevant column ascendingly.

Section V. Perform Review

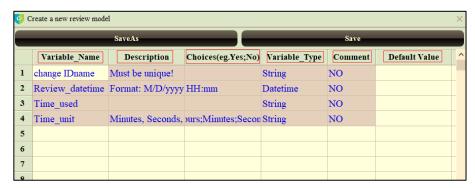
V-1. Review Model

V-1-1. Create a New Review Model

Choose the tab Review in the middle of the CHANL interface, click to start creating a new review model.



A popup window will appear:



V-1-1-1. Variable Name

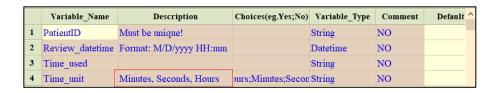
In the review model popup window, please enter variable names (e.g. "PatientID", "Diagnosis") for chart review.



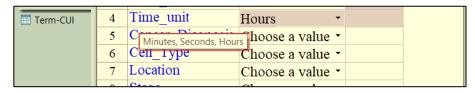
There are 3 default variable names in each model: Review_datetime, Time_used, Time_unit. The values for these variables will be filled by CHANL automatically if users don't actively enter values.

V-1-1-2. Description

A value in the column "Description" will be displayed when the mouse is hovered over a variable to remind reviewers about the detail of a variable. E.g. Minutes, Seconds, Hours in the "Description"



will be shown in a review model when the mouse is hovered over the variable "Time unit".

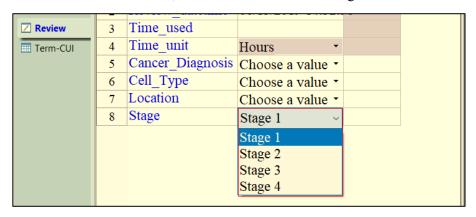


V-1-1-3. Choices

We can enter choices in the column "Choices" if there is a list of predefined values for a variable. For example, choices for cancer stage could be "Stage 1;Stage 2;Stage 3;Stage 4". The delimiter must be ";".



In the review model loaded, we can choose one of the stages for the variable "Stage".



V-1-1-4. Variable Type

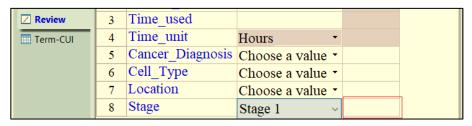
Please define a variable type for each variable. When a variable type is defined as "Integer", a string such as "stage 1" will be invalid. Please define the type as "String" if a string value is a variable. There are 8 types on the list including 'Integer', 'Float', 'String', 'Date', 'Datetime', 'Boolean', 'currency' and 'String(unit)'. Please choose one from the dropdown list.

V-1-1-5. Comment

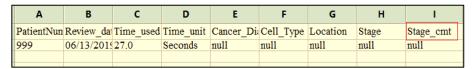
"Comment" is enabled by choosing "YES" in the dropdown list, a variable of "*_Cmt" will be automatically generated in the result file for the relevant variable and the "Comment" area will be editable. E.g. In the model creation window, "YES" is chosen for the "Comment" for the variable "Stage".

	Variable_Name	Description	Choices(eg.Yes;No)	Variable_Type	Comment	Default ^
1	PatientNum	Must be unique!		String	NO	
2	Review_datetime	Format: M/D/yyyy	HH:mm	Datetime	NO	
3	Time_used			String	NO	
4	Time_unit	Minutes, Seconds,	ours;Minutes;Secon	String	NO	Hours
5	Cancer_Diagnosis	0 - Negtive;1 - De	0;1;9	Integer	YES	
6	Cell_Type		l;Adenocarcinoma;	String	YES	
7	Location		ht upper;Right mid	String	YES	
8	Stage	Could be TNM sta	1;Stage 2;Stage 3;S	String	YES	

The cell of the "Comment" for the variable "Stage" will be editable on the review model.



A variable "Stage cmt" will be generated in the review result file automatically.

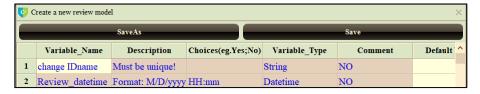


V-1-1-6. Default Value

If a value appears often for a variable, we can set a default value for this variable. Eg. If "Stage 1" is very common for the variable "Stage", we can set "Stage 1" as the default value for the variable "Stage". However, we can still select another value from the dropdown list.

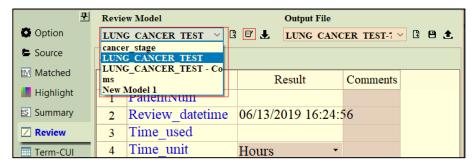
V-1-1-7. Save a Model

Click the button "Save" to save a model.



V-1-2. Edit a Review Model

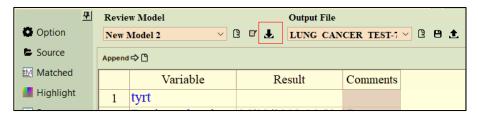
Choose a model from the dropdown list and click between to enter the review model editor window.



After modification, click the button "Save" to save the change or click "Save As" to save as another model name. The review model needs to be re-loaded to make it effective by selecting the model from the dropdown list.

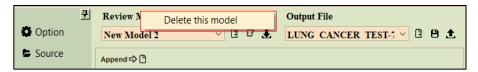
V-1-3. Import a Review Model

Click the button **to browse the computer to import a CHANL review model.**



V-1-4. Delete a Review Model

Choose a review model from the dropdown list and right click on the model name. Click the popup menu "Delete this model" to delete an existing model from the list.



V-2. Record Review Results

V-2-1. Record Results

The first row is for PatientID which must be filled. All other fields will be filled as 'null' automatically if there are no values entered. Please enter values with the correct value type.

V-2-2. Enter or Select Comment

If the field of comment is not disabled or greyed out, we can enter or select comments using the CHANL manual highlighting tool.

Before using the CHANL highlighting tool, we need to select the field of comment, then click the toggle button on the main panel to enable "Manual Highlighting". The text we select in the notes will be copied to the comment field.

V-2-3. Review Time

V-2-3-1. CHANL Timer and Clock

Every time we append results to a file, there will be a CHANL clock generated

12:33:26 00:00:52. It displays the real computer time and the time when the result appending occurred. Double clicking the CHANL icon can close the clock. It will appear again when clicking the button "Append". The clock can be moved by dragging.

V-2-3-2. Time Records

Review_timeadate will always be filled automatically. The field for "Time_used" will be calculated and filled automatically along with the time unit if no values are entered. The automatic time will start from the time a new review model was loaded or last time the result was successfully appended.

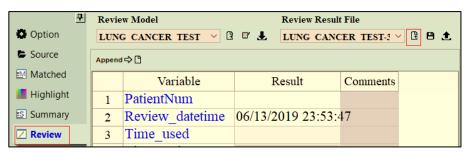
V-2-4. Append Results

Click the button Append to append results to a review result file. The result will be displayed in the "Review Result Display" sheet which will be described later.

V-3. Review Result Files

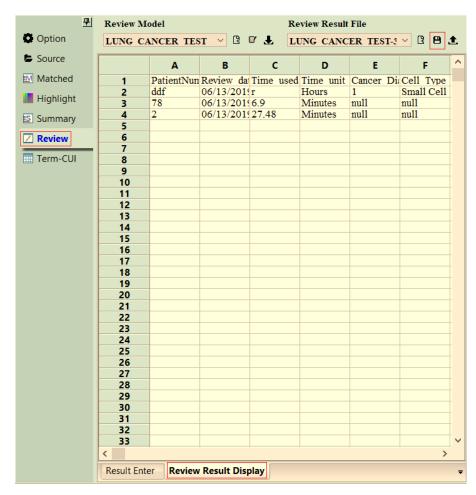
V-3-1. Create a Review Result File

Choose a Review model first. Click the tab "Review" then click the button in the area "Review Result File" to create a new review result file for the model chosen. If the review model is "Lung_Cancer_test" and a name entered for result file is "11", the final result file name is "lung_cancer_test_11".



V-3-2. Save Change to a Review Result File

When successfully appending data to a result file, data will be automatically saved. But we can also go to the "Review Result Display" tab to manually change the result and click the button to save the modification.



V-3-3. Export a Review Result File

We can export CHANL review result to a CSV file with delimiter "|" by clicking the button ...



V-3-4. Delete a Review Result File

Load a result file, then right click the name of the file loaded to delete it.

