

# Visit Notes Analysis Module User's Guide

Kavya Katipally Advisor: Dr. Barry Levine

#### **TABLE OF CONTENTS**

1. Overview	3
1.1 Named Entity Recognition of Visit Notes	3
1.2 OpenMRS Visit Notes Analysis Module	4
1.3 Named Entity Recognition Algorithm	4
1.4 Limitations	5
2. Visit Notes Analysis Module Installation	6
2.1 Requirements	6
2.2 Pre-Installation	6
2.3 Installation	6
3. Visit Notes Analysis User Interface	8
3.1 Visit Notes Analysis User Interface – Page 1	9
3.1.1 Visit Dates Chart	9
3.1.2 Date Range Slider, Entity Type Selector, Number of Entities Drop down	10
3.1.3 Word Cloud	10
3.1.4 Search Bar	11
3.2 Visit Notes Analysis User Interface – Page 2	12
3.2.1 Search History	13
3.2.2 Search bar and date pickers	14
3.2.3 Heat map	16
3.2.4 Visit Note List	19
3.2.5 Visit Note Rendering	25
3.2.6 Admin Notification	32

# 1. Overview

## 1.1 Named Entity Recognition of Visit Notes

Physicians have a limited amount of time to get familiar with patient history and visit notes have been identified as a primary source of patient history information. Electronic medical records (EMRs) have been widely adopted in the last few years but they do not provide strong support for unstructured data such as visit notes.

The first step in extracting information from visit notes is to identify key terms such as problems, treatments and tests. Named Entity Recognition (NER) identifies Named Entities, which are elements within unstructured text that fall within predefined conceptual categories. Tagged named entities in this module fall in one of three categories – Problems, Treatments or Tests.

Here's an example of text with named entities identified:

ILLNESS: The patient is a 47 year old male, with end stage liver disease secondary to hepatitis C cirrhosis diagnosed about 5 years prior to admission. The patient had undergone treatment with interferon and Ribavirin. He had been admitted to the St. Margaret 's Center for Women & Infants multiple times early in 2013 for management of encephalopathy and ascites. The patient had been discharged from the St. Margaret 's Center for Women & Infants on 2013-05-06, but was readmitted on 2013-05-09 when noted to have worsening renal function. The patient 's serum creatinine on the day of discharge, on 2013-05-06, was 1.9, but was noted to increase to 3.2 on 2013-05-08, and was further elevated to 3.6 on 2013-05-09. The patient was admitted with concern for hepatorenal

Fig 1: Sample visit note with named entities

## 1.2 OpenMRS Visit Notes Analysis Module

The goal of this module is to help physicians get a quicker understanding of patient history. It helps them identify the most frequently occurring terms in the visit notes within various time periods and entity types. They can then analyze the timelines of search terms as well as those of related terms. They can further refine their search or filter visit notes and identify the context in which an entity occurred in a note. Thus, the physician can look at higher level summarized data and then dig deeper and identify the specific context in which an entity was noted.

# 1.3 Named Entity Recognition Algorithm

The NER algorithm used in the OpenMRS Visit Notes Analysis module has two steps – a string matching step which identifies entities occurring in the OpenMRS concept dictionary and a second step which runs an NER tool called BANNER. BANNER uses the machine learning algorithm Conditional Random Fields (CRF) to locate named entities.

For the first step, a mapping is required between the OpenMRS concept classes and the entity types being tagged by the module (Problems, Treatments and Tests). The algorithm looks at all concepts across different classes as well as all synonyms associated with the concepts. Synonym lists may include words in many languages and so words from other languages can also be tagged.

The second step with BANNER identifies entities which were not in the OpenMRS concept dictionary. In case of a collision between entities found in the two steps, the one identified in the first step takes precedence.

# 1.4 Limitations

As part of the development of the first version of this module, four NER systems were evaluated. The measurements done in that study showed that of the entities tagged by BANNER, about 80% are accurate. It was also found that about 70% of all entities are found by BANNER. This performance is good, but it should be noted that the results from this model are not perfectly accurate.

# 2. Visit Notes Analysis Module Installation

# 2.1 Requirements

JDK 7, OpenMRS Standalone 2.3.1, at least 2 GB available RAM.

#### 2.2 Pre-Installation

Ensure that server JVM has at least 2GB of heap space available by making the following change in openmrs-standalone-runtime.properties:

vm\_arguments=-Xmx2048m -Xms1024m -XX:PermSize=256m -XX:MaxPermSize=256m -

XX:NewSize=128m

#### 2.3 Installation

Please go to Manage Modules page and select the .omod file using 'Add or Upgrade Module' button.

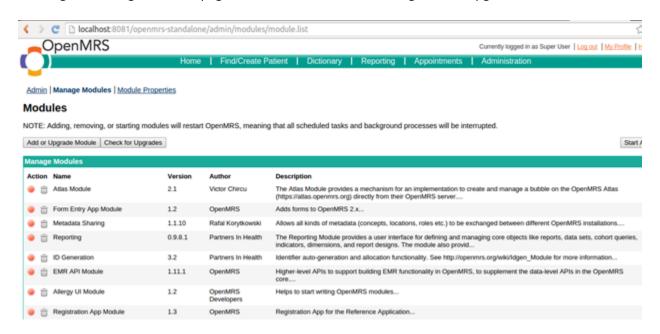


Fig 2: Upload the module using the OpenMRS Manage Modules page

To use the module, go to the Patient page and add some Visit Notes. Then select the 'Visit Notes Analysis' link under 'General Actions' to access the module.

# 3. Visit Notes Analysis User Interface

To go to the Visit Notes Analysis module, go to the Patient Dashboard. Under 'General Actions' to the right, click the 'Visit Notes Analysis' link. The figure below shows the location of the link on the dashboard:

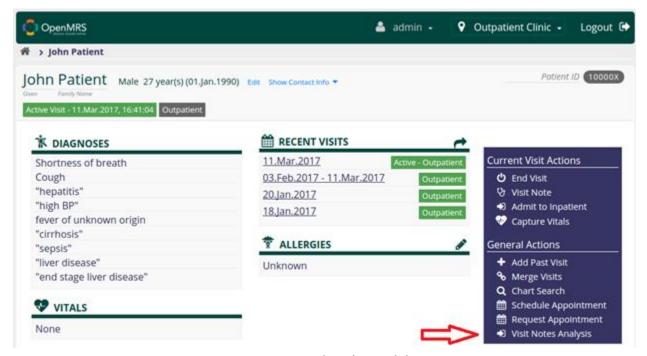


Fig 3: Finding the module

The user interface has two pages. Screenshots and corresponding details are mentioned in the following sections.

# 3.1 Visit Notes Analysis User Interface - Page 1

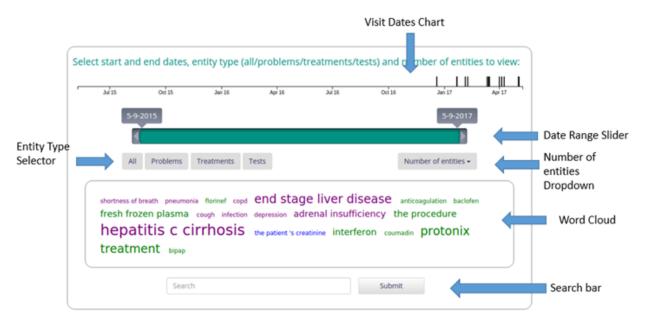


Fig 4: Visit Notes Analysis UI - Page 1

The components of page 1 of the user interface are discussed below -

#### 3.1.1 Visit Dates Chart

The chart at the top displays all visit notes available for this patient as vertical bars. The chart covers 2 years by default, which matches the slider below. If the slider dates are modified, the chart domain gets updated automatically. Hovering over a vertical bar displays some details regarding the visit note and clicking it shows the visit note rendering on page 2.

The figure below shows a change in the chart domain when the slider is moved. We hover over a vertical bar and see a tooltip with the date.

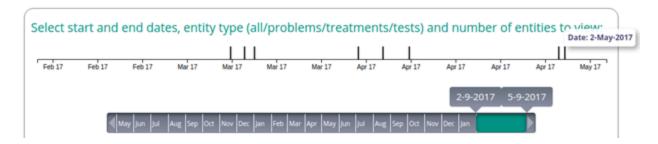


Fig 5: Page 1 - Change in chart domain upon moving slider

#### 3.1.2 Date Range Slider, Entity Type Selector, Number of Entities Drop down

The date range slider has a minimum range of 3 months, maximum range of 2 years and by default, the start date is two years ago and the end date is today. The date range slider, entity type buttons and the number of entities drop down menu can be changed to update the word cloud.

The figure below shows the date range slider with a range of 3 months, the Entity Type Selector with 'Problems' selected, and the options shown for the Number of Entities Drop down menu:



Fig 6: Page 1 – Date Range Slider, Entity Type Selector and Number of entities drop down menu

#### 3.1.3 Word Cloud

The word cloud displays the most frequent entities present in visit notes within time range selected in the date range slider. The entities are of the entity type selected and the number of entities displayed comes from the number of entities drop down menu. The entities in the word cloud are color coded –

problems are shown in purple, treatments in green and tests in blue. The font size is proportional to the frequency of occurrence of the entity during the time range selected.

#### 3.1.4 Search Bar

The user can click particular entities in the word cloud to add them to the search bar and hit 'Submit' to go to page 2. Users can also directly type into the search bar. Cases such as duplicate terms, too much whitespace and too many commas etc are handled automatically.

The figure below shows the word cloud when date range is from 2-9-2017 to 5-9-2017, entity type selected = 'Problems', Number of entities = 10. The entities 'end stage liver disease' and 'shortness of breath' in the word cloud are clicked and added to the search bar.

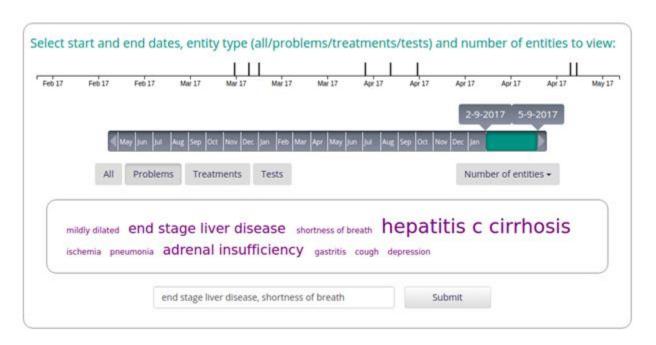


Fig 7: Page 1 – Word Cloud and Search Bar

# 3.2 Visit Notes Analysis User Interface - Page 2

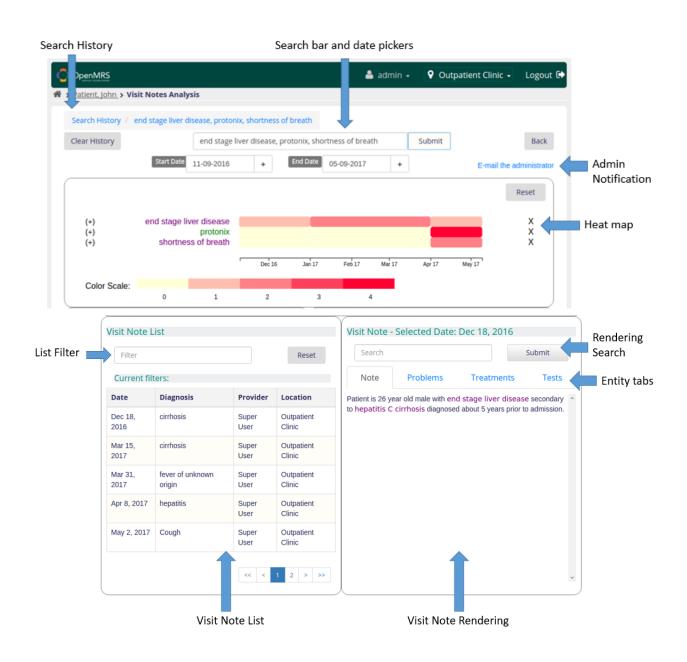


Fig 8: Visit Notes Analysis UI - Page 2

## 3.2.1 Search History

Breadcrumbs are displayed at the top of page 2 of the UI. This functionality is used to store search history and it can handle multiple search terms and date ranges. The user can go back to previous search terms/date range by clicking on the corresponding term.

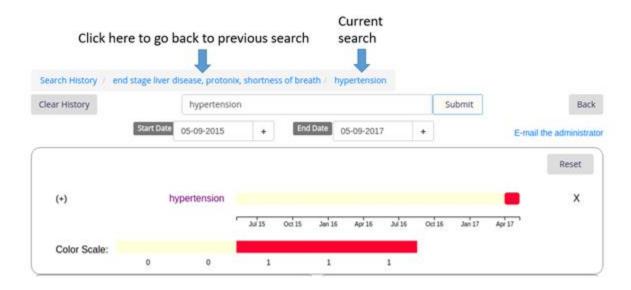


Fig 9: Page 2 – Search History

After clicking the previous search terms, the corresponding search terms (end stage liver disease, protonix, shortness of breath) and start and end dates (11-09-2016, 05-09-2017) are re-submitted:

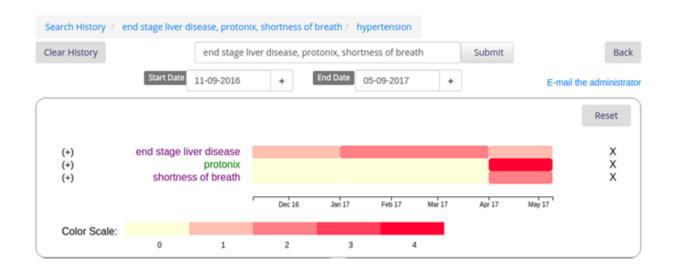


Fig 10: Page 2 – Re-submitting from Search History

The 'Clear History' button is used to clear the Search History.



Fig 11: Page 2 – Clear History

## 3.2.2 Search bar and date pickers

The search bar at the top of page 2 is similar to the one on page 1. Start and end dates in the date pickers default to dates from the slider on the previous page. The minimum date range allowed is 3 months and maximum date range allowed is 2 years.

When the end date is 05-09-2017, since the maximum date range allowed is 2 years, the start date cannot be before 05-09-2015.

We see below that start dates before 9<sup>th</sup> May 2015 are greyed out and the user cannot select them.

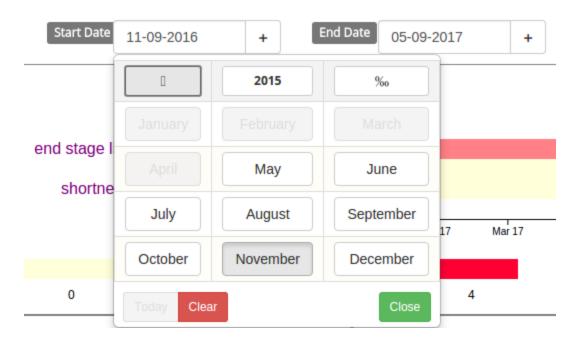


Fig 12: Page 2 – Maximum range of date pickers

Similarly, for end date = 05-09-2017, since the minimum date range is 3 months, the start date cannot go beyond 02-09-2017. We see below that dates after  $9^{th}$  February 2017 are also greyed out.

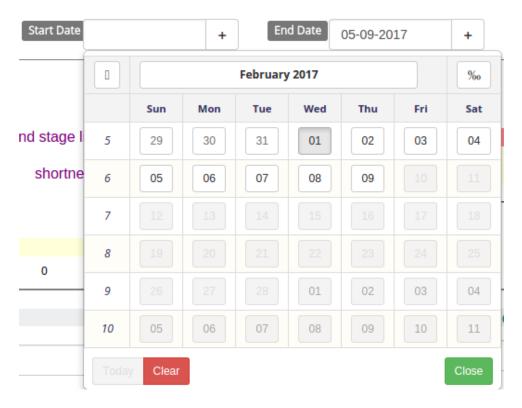


Fig 13: Page 2 – Minimum range of date pickers

#### 3.2.3 Heat map

The heat map section displays the timelines of occurrence of each of the search terms within the date range selected. No duplicate terms are displayed. A uniform color scale is applied across the entire heat map, including rectangles corresponding to quarters in a year, and vertical bars corresponding to individual notes (which can be seen upon hovering over a rectangle).

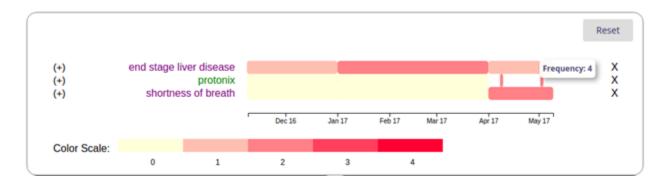


Fig 14: Page 2 – Heat map

In case the number of search terms exceeds the limit (set to three), only the search terms are displayed in the heat map. However, if the number of search terms is less than the limit, related terms can be seen by clicking the (+) toggle symbol displayed beside each search term. This is shown in the figure below.

The most frequent five problems, five treatments and five tests from visit notes with the search term are considered 'related' to the particular search term.

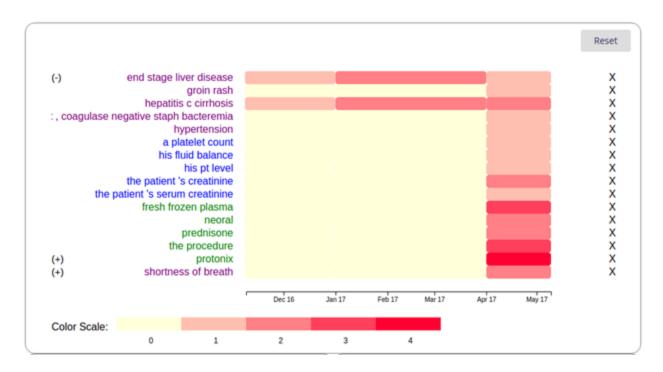


Fig 15: Page 2 - Heat map with related terms

Terms can be removed from the visualization by clicking the 'X' symbol to the right. In case the 'X' symbol is clicked on a search term, the search term selected and its related terms are all deleted from the heat map. In the figure below, 'shortness of breath' and all of its related terms and all treatment (green) related terms of 'end stage liver disease' have been removed.

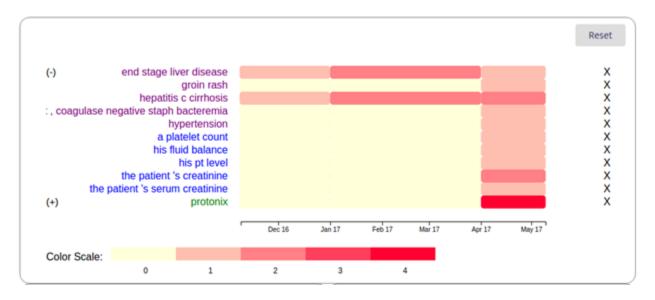


Fig 16: Page 2 – Removing terms from the Heat map

Clicking the 'Reset' button seen on the top right of the heat map restores all deleted terms and resets the heat map to the original display. Here's a screenshot after 'Reset' is clicked from the case above.

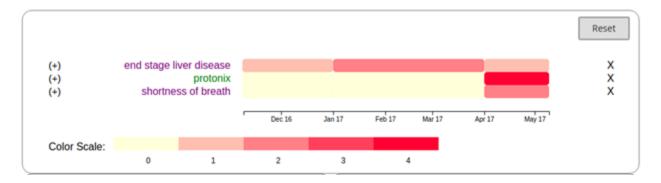


Fig 17: Page 2 – Resetting the Heat map

Hovering over any rectangle displays a tooltip with the total frequency of the corresponding entity in that time period. The vertical bars seen are the individual visit notes with the corresponding entity. The color scale applied to both the rectangles (covering a quarter of a year) and the vertical bars (single note) is the same and follows the color scale legend.



Fig 18: Page 2 – Hovering over a Heat map rectangle

Hovering over a vertical bar shows a tooltip with details of the particular visit note.

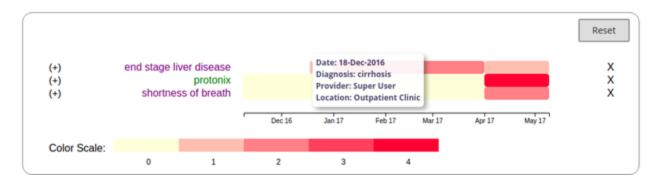


Fig 19: Page 2 – Hovering over a Heat map vertical bar

Right clicking any entity in the heat map adds it to the search bar. Here 'hepatitis C cirrhosis' and 'fresh frozen plasma' were clicked and added to the search bar at the top of page 2.



Fig 20: Page 2 – Right click heat map entity to add to search bar

#### 3.2.4 Visit Note List

The bottom left section of page 2 has the Visit Note List. When search terms are submitted, all notes displayed on the heat map (including those for related terms) are listed in the Visit Note List section.

There are no duplicate notes in the list and five notes are shown on each page of the list. For each note, the Date, Diagnosis, Provider and Location details are mentioned.

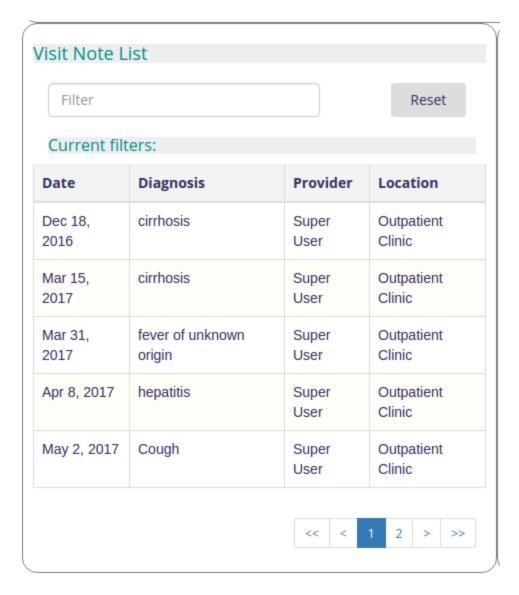


Fig 21: Page 2 – Visit Note List

The Visit Note List section can be filtered in a few ways. Left clicking any entity on the heat map adds it to the filter input section of the Visit Note List. The user can also manually enter anything into this input box to filter the visit notes. This is a very general filter that looks for the entered value across all the data for the visit note, including the Date, Provider, Location, Diagnosis and the associated problems, treatments and tests. Also note that this filtering is done incrementally so with each extra letter added by the user, the filtering is repeated.

The example below shows that just one note is displayed when filtered by '2016', since all other notes for this search occur in 2017.

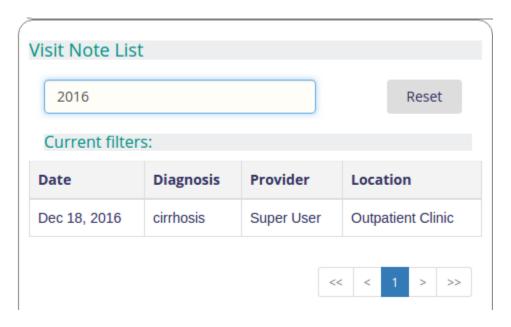


Fig 22: Page 2 – Visit Note List filter manually

Another example is shown below. 'Protonix' is left clicked on the heat map. This automatically scrolls the page down to the Visit Note List section and shows the two notes that contain 'Protonix' as a treatment.

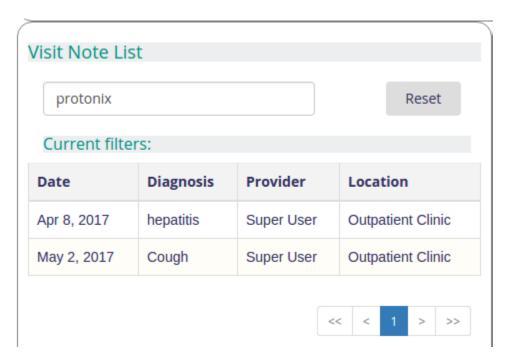


Fig 23: Page 2 – Left click heat map entity to filter Visit Note List

Another way to filter the Visit Note List is shown in the figure below. Each rectangle or vertical bar in the heat map has corresponding start and end dates and an associated entity. Clicking either a rectangle or a vertical bar filters for visit notes satisfying these criteria: The visit date has to fall in between the start and end dates of the filter, and the filter entity has to be present in the problem, treatment or test list of the note. The filters currently being applied are shown above the note list. In the example below, we click the note with date = 8<sup>th</sup> Apr 2017 and entity = 'fresh frozen plasma'. Hence, filter start date = filter end date = 8<sup>th</sup> Apr 2017 and filter entity = 'fresh frozen plasma'. Note that the filter start and end dates are the same since we clicked a single note.

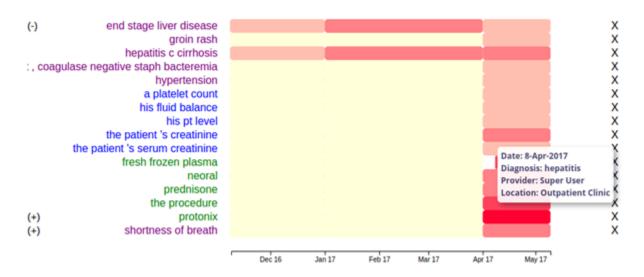


Fig 24: Page 2 – Click heat map vertical bar

Now, all notes in the Visit Note List with the date in between the filter start and end dates and 'fresh frozen plasma' as a problem, treatment or test are displayed. The only note satisfying these criteria corresponds to the vertical bar clicked. The start date filter, end date filter and entity filter being applied currently are displayed above the Note table.

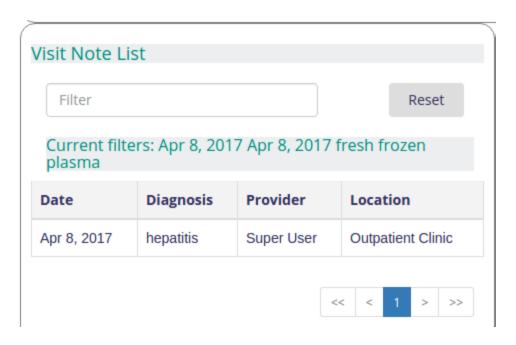


Fig 25: Page 2 – Click heat map vertical bar to filter Visit Note List

Let's consider another case where we click the rectangle corresponding to the second quarter of 2017. Here, filter start date =  $1^{st}$  Apr 2017, filter end date =  $9^{th}$  May 2017, filter entity = 'protonix'.

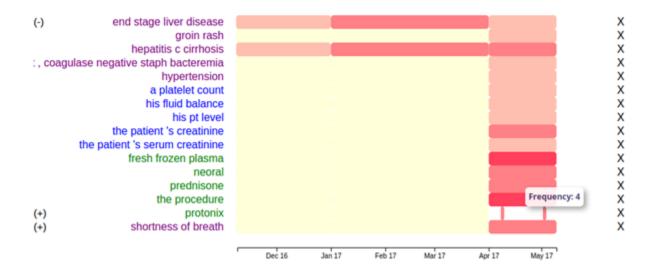


Fig 26: Page 2 – Click heat map rectangle

Only notes that have the date in between the filter start and end dates and contain 'protonix' as a problem, treatment or test are displayed in the Visit Note List. The only notes that satisfy these criteria

that the two corresponding to the vertical bars seen above. The filter criteria are displayed above the table.

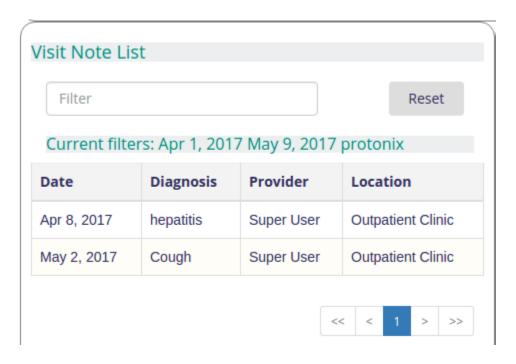


Fig 27: Page 2 – Click heat map rectangle to filter Visit Note List

The 'Reset' button in the Visit Note List section can be clicked to remove all filters. After clicking 'Reset', all original visit notes from the heat map are displayed once again in the list as seen below. The current filters section is now empty.

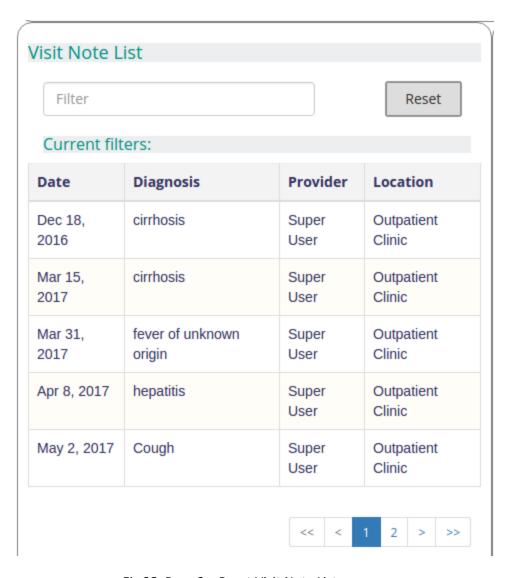


Fig 28: Page 2 – Reset Visit Note List

#### 3.2.5 Visit Note Rendering

The bottom right section of page 2 has the Visit Note Rendering section. This section shows the entire Visit Note with all entities highlighted. Problems are shown in purple, treatments in green and tests in blue. This section can be updated by clicking the Visit Note List row corresponding to the note of interest. The date of the note displayed is shown at the top.

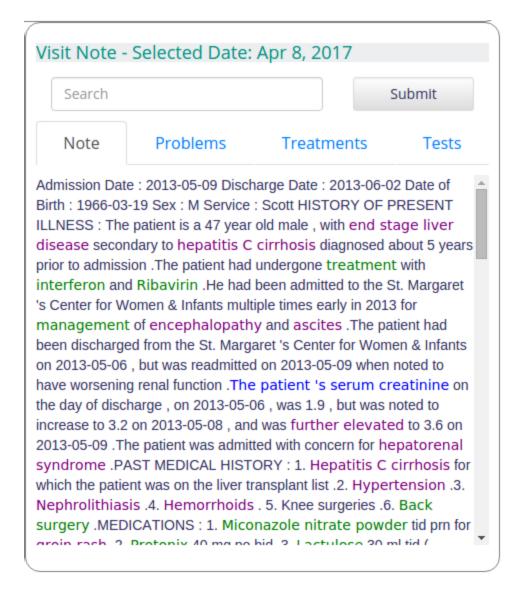


Fig 29: Page 2 -Visit Note Rendering

All problems, treatments and tests present in the note are displayed separately in the other tabs.

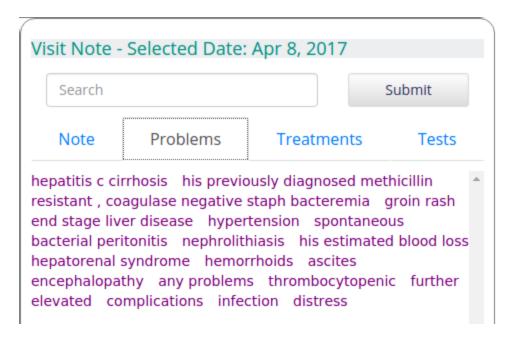


Fig 30: Page 2 – Visit Note Rendering Problems tab

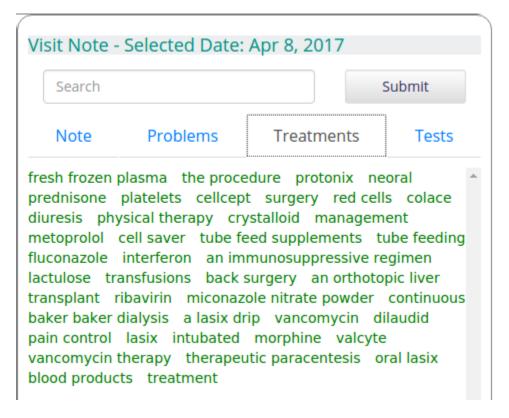


Fig 31: Page 2 – Visit Note Rendering Treatments tab



Fig 32: Page 2 – Visit Note Rendering Tests tab

In the Visit Note Rendering section, any of the problem, treatment or test entities in the tabs can be clicked to be added to the search bar. Entities can also be added manually to the search field, but please note that the entire entity of interest has to be added for it to be found. For example, if the entity is 'a platelet count', entering just 'platelet' will not work. The entire entity 'a platelet count' has to be added for it to be located in the note. Here 'fresh frozen plasma' is clicked, so it's added to the Search field of the rendering section.

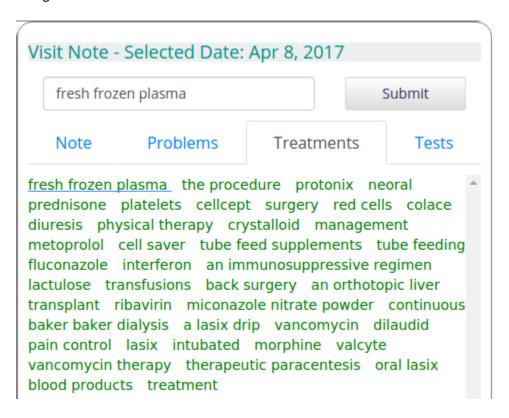


Fig 33: Page 2 – Click entity in Visit Note Rendering to add to Search

Upon hitting 'Submit' the location of the entity is highlighted in the note.

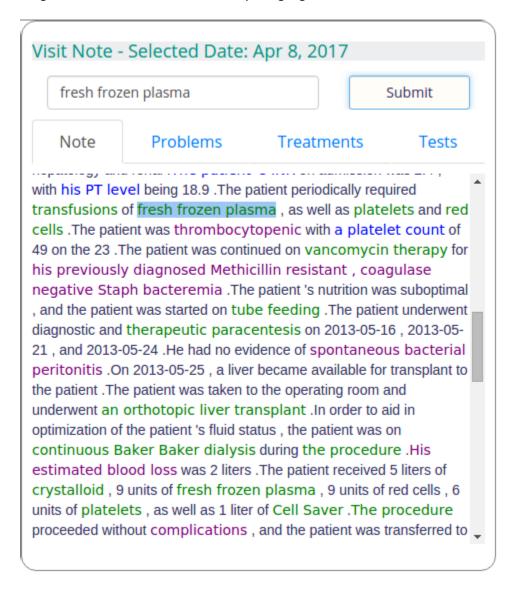


Fig 34: Page 2 – Search for entity in Visit Note Rendering

If a term is present multiple times in the same note, hitting 'Submit' again scrolls the note to the next location. 'Fresh frozen plasma' is present three times in this note, so submitting the entity once again takes us to the next location.

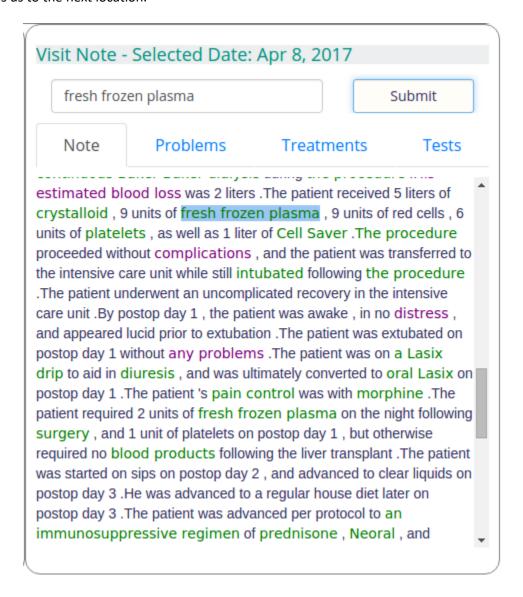


Fig 35: Page 2 – Search for multiple locations of entity in Visit Note Rendering

This highlighting loops around, so if 'Submit' is clicked when the user has viewed the last location of the entity, the note is scrolled back to the first location of the entity in the note.

#### 3.2.6 Admin Notification

This button can be used to notify the administrator regarding errors in tagging. Once the user notices errors made by the current model, clicking on the link prepares an email with the currently selected note. The administrator will be able to use the standalone training application to make the requested changes and train a new NER model.

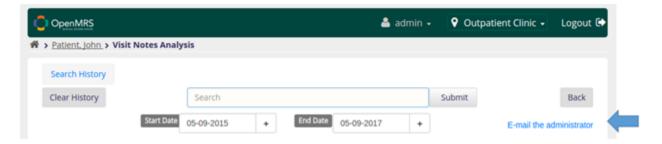


Fig 36: Page 2 – Admin Notification