Biomedical Document Retrieval System Using Spacetree Diagram

by

Nurdini Binti Nordin

Dissertation submitted in partial fulfillment of

the requirement for the

Bachelor of Technology (Hons)

(Business Information System)

MAY 2012

Universiti Teknologi PETRONAS

Bandar Seri Iskandar

31750 Tronoh

Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

Biomedical Document Retrieval System Using Spacetree Diagram

By

Nurdini Binti Nordin

A project dissertation submitted to the

Information System Programme

Universiti Teknologi PETRONAS

in partial fulfillment of the requirement for the

BACHELOR OF TECHNOLOGY (Hons)

(BUSINESS INFORMATION SYSTEM)

Approved by,

(Ms ELAINE CHEN YOKE YIE)

UNIVERSITI TEKNOLOGI PETRONAS

TRONOH, PERAK

May 2012

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

_____

NURDINI BINTI NORDIN

Abstract:

This project aim is to provide an application that support visualization of information from biomedical documents available online. The large numbers of potentially related documents retrieved by search engines ie. Google or PubMed are being represented in a non-interactive and organized form. This conventional representation of information did not help in providing meaningful information for the user in helping them to navigate through the co-related documents retrieved from the same field in biomedical studies. Information visualisation technique can further provide user with the relationship among the biomedical documents retrieved from the user search results. By forming the relationship between the biomedical documents retrieved, user can have an overview of the connection between the information resides in multiple documents. The graphical visualization are to be in the form of tree diagram whereby user can see the expansion of the information being provided by each of the biomedical documents retrieved in a guided approach.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Project Background:

Natural Language Processing (NLP) field of study is the pioneer in the development of studies regarding human-computer language interaction. The further development in NLP fields produced useful tools such as automatic text summarization, automatic text translation into another human language, information extraction from text and information retrieval. With the advancement of technology, millions of documents are being made available online. Biomedical documents are one of the most rapidly growing in number for the documents made available online. Information retrieval system has been a handy tool for user in searching through the mass of documents available online. Search engines such as Google, PubMed for biomedical documents are the readily example of information retrieval tools. The system received user input i.e: keyword and will performed the algorithm programmed in retrieving the documents related to the keyword.

The systems for retrieval of biomedical documents which are made available online are mainly focusing on retrieving the potentially useful documents from queries made by the user. The results retrieved are driven by text manipulation ie: the number of occurrences or similarity between strings of citation found in the documents compared to user's search input. These query results are being accompanied by automatic text summarization result found in the documents by referring to user's input. Most of the query results are being represented in lists which are ranked according to date of publication instead of relevancy of the information contained in the documents. The representations of the information extracted are stand-alone and portrays no relationship between one another.

Information visualization approaches in visualising biomedical information retrieved from online text and search engines query results has been researched to test the theory of

information visualisation for biomedical documents can be developed to portrays the relationship between the biomedical documents. Throughout the years, researches done on visualising biomedical information had comes up with different forms of graphical visualisation to represent biomedical information from text documents. Further researches on visualising biomedical text are focussing on providing relationships between the biomedical documents in the form of graphical visualization.

Problem Statement

Most documents retrieval system works to only suggest potentially useful documents and provides no meaning for the automatic text summarization driven. The query results did not portray the relationship between the documents retrieved hence increases the risk of user overlooking important information. The usual representations of the retrieved results are in lists form which did not provide easy navigation throughout the retrieved results. Graphical representation i.e: mind maps, will provide better representation and easier navigation through the biomedical documents retrieved.

Significant of the Project

This project aim is to develop a web application that will manage the query results retrieved into a visual representation of the relationship among the biomedical documents retrieved via the search engines. This application will enable the citation of text retrieved to be represented visually in a mind map that clearly portrays the relationship among the biomedical information retrieved. The information visualization is also aimed to provide easier navigation for the user through large numbers co-related biomedical documents retrieved.

Objectives of Study

The objectives of the projects are as follows:

1. The objectives for this project is to develop a web application that can provide a visual representation of the biomedical documents query results retrieved from databases or online search engines.
2. To visualize the relationship between biomedical documents retrieved from database.

Table 1.0

The Relevancy of the Project

The application that is going to be developed from this project is aimed to aid a simple search to be of a more informative and easier navigation throughout the large numbers of related biomedical documents. The visualization of the biomedical document's query results shall provide an easier understanding regarding the searched topics for the user. The graphical visualization will help user to get clear overall view of the information retrieved from their search. The relationship formed among the biomedical documents will allow user for a guided and easier navigation though the topic.

Chapter 2 : Literature Review

Literature Review

Information Visualization for Biomedical Text is a research field that has been given lots of attention by researches all over the world. Most of the researches are intrigued to find ways to manage the large amount of document available online [1],[2],[3] and means of presenting the information by using visual aids [1],[2],[4],[3],. The researchers are driven by theory of transforming automatic text summarization into semantic information through query results from large databases. Another supporting theory for this project is information visualization as the best solution in representing the structured management of the large number of documents retrieved.

The researches on this subject are being done[1] by two different approaches 1) focussing on one particular field in biomedical field[1] [1],[2],[4] and 2) involves all biomedical documents available from database i.e PubMed [5],[6],[3]. Among the area of studies being focus in the previous research papers for Information Visualization for Biomedical Text are therapeutic studies [1], adverse drug effects [2] and clock genes studies [4].

The similarities which can be found from the past researches are on the approach in managing and extracting valuable information from a large bank of biomedical documents. There are several issues being highlighted from the past works. The first concern is on how to effectively manage the results of the query retrieved in order to represent only the information that are relevant for the user[1] ,[6],[4]. The next concern is to represent the query results in a visual form that can provide useful information regarding the relationship between the documents retrieved from the query results [6]. This concern has heightened the need to enables the retrieved citation from the documents to be shown in a clear relationship between one another [3].

The methodologies used in the past works in this research field are using an almost similar sequence of methods in transforming the information retrieved into a visual representation. These researches will retrieve query results of citation or text summarization from online search engine to be process for relationship identification among the citation before being represented visually i.e : mind maps [1], [4], bar charts [2].

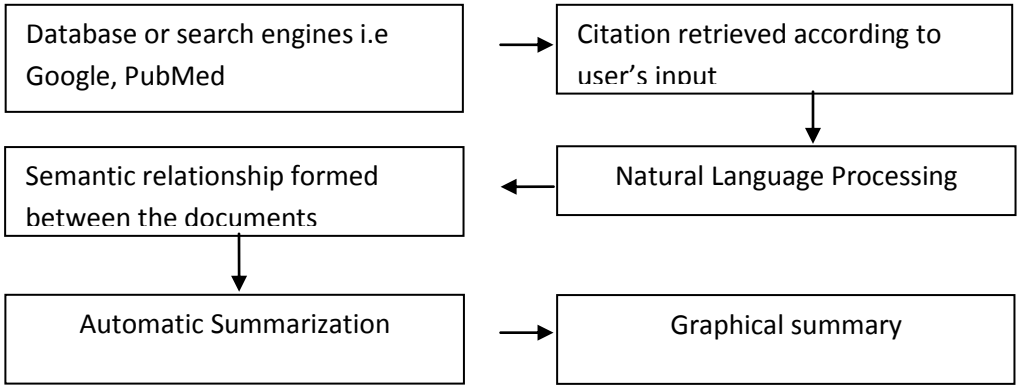


Figure 1.0 – General steps in visualizing biomedical information [5].

However each researches have a distinctive variation in several steps, as can be found in research paper titled "*A New NLP System for Biomedical Text Analysis"* [5] the researcher were focusing on refining the algorithm so as to produced a solid text summarization that contains only facts by eliminating other ancronym before the information is being represented in text format [5]. The representation of this research are not through graphical aids, instead it's a summarization of text which contains only facts from the scientific biomedical reports or documents [5]. This research paper offers useful techniques used in extracting relevants facts from text documents hence it is included in this literature review.

In the research paper titled '*The NIH Visual Browser: An Interactive Visualization of Biomedical Research'* [3] the emphasis of the research was to come out with a more interactive information visualization application [3]. The information retrieved are being represented in a big mind map of connecting dots or lines [3]. Although the visualization of the information adopted here did not gives off the major idea in the first glance, the key features of this research are on the deep zooming, selection, color-schemed code, full text query and multi level labelling features that were embedded in the application [3]. This research was focussing on the interactive functions for the web application develop to access the database. The interactive features of this research outcomes suggested that information visualization with interactive functions is a good combination in generating semantic knowledge from explicit texts.

As being reported in the research paper titled '*A concept-driven biomedical knowledge extraction and visualization framework'* [6] the application BioKEVi was designed to represent extracted queries with graphical visualization aids according to the level of depth for the topic aquired by the user [6]. This research were focusing on visualising the relationship between the documents with similar biomedical informations. Clear representation of the relationship among the biomedical text are crutial in this research as it is the key concept to support easier navigation [6] for the user in going through the biomedical documents retrived. This research paper reports on the methods of forming the relationship among the similar biomedical documents. This research paper represents the relationship in the form of mind map with relationship keyword in between the connecting lines. Thus the mind map is made easier to navigate from one document content to another.

Information visualization for biomedical text had as well been research in providing a Decision Suport System (DSS). In the research paper titled '*Data visualization speeds review of potential adverse drug events in patients'* [2] information visualization is being implemented in DSS to expediate the reviewing of potential drugs effects on patient[2].

The system develop from this research will present a comparison of adverse drug effect into a graphical visualisation [2]. The graphical visualisation adopted in this research are mostly bar-charts as the medical practitioner need to be able to decide which adverse drug effects are more likely to occurs based on the height of the bar charts elements. The implementation of graphical visualization in the DSS system to speed the time for medical prectitioner to come to a decision before prescribing drugs to their patient[2]. This research paper suggested that information visualization can provide better repersentation of information from biomedical documents.

In research paper titled '*Degree centrality for semantic abstraction summarization of therapeutic studies'* [1] the research was done with the aimed to produce a system that can manage the large amount of citation retrieved through search engines[1] . The research had focused on visualising the citation of biomedical documents retrived from medical library i.e: PubMed [1]. These citation of biomedical text retrieved are being grouped together according to its relevancy before being projected in a mind map graphical visualization. The lines connecting the mind map are being embedded with the link to the actual document the citation was retrieved [1]. This research suggested that by only a keyword, the system will be able to generate a mind map that can help user to narrow down their searches to the specifics subjects and leads to discoveries of new information.

A more organized research following '*Degree centrality for semantic abstraction summarization of therapeutic studies'* [1] had been reported in another research paper titled '*Semantic MEDLINE: An advanced information management application for biomedicine'* [4] which were as well published on 2011 but in a another publication. In '*Semantic MEDLINE: An advanced information management application for biomedicine'*[4] , the research had been focussing on studies of Clock Genes field in biomedical [4]. This research had develop a web protal that act as an application that will retrieved query results from search engines i.e: Google or PubMed [4]. The citation retrieved will be be connected based on fact-driven before being presented in a connecting graphs that resembles a mind map [4]. The

link to the actual documents will be embedded in the lines connecting the graphs. An added key feature of text preview has been embedded in the graphical visualization [4]. This research has suggested that forming relationships by analysing the citation has resulted in higher relevancy of the text being relate in the mind map.

From this literature review, all the cited works being mentioned are related to the this FYP objectives. Although some of the cited works did not fulfilled the aimed to visualized information [5], the methods and steps reported to achive the goal of this projects are helpful as references for the next phase of this project. These research papers had as well serve as a proof that information visualization for biomedical documents is possible. Contribution of information visualization in providing better representation of information from biomedical documents had as well been proven from these research papers reviewed.

Chapter 3: Methodology

Methodology

This project will be developed using Rapid Development methodology. This project requires detailed analysis from previously similar or in the same field research papers. Finding from these research papers might changes the detailed problem indentified in the early development of this project, hence by adapting rapid development methodology this project will be flexible for changes during the development cycle.

Rapid methodology emphasized on the user's feedback regarding the prototype which will provide useful inputs for the development of this project

Project Gantt chart

Final Year Project Semester 1 Timelines

No	Details/ Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Selection of Project Topic														
2	Preliminary Research Work														
3	Submission of Extended Proposal Defence														
4	Proposal Defence														
5	Project work continues														
6	Submission of Interim Draft Report														
7	Submission of Interim Report														

Figure 2.0

Phase 1 - Week 1 to 5

- Project title selection.
- Develop understanding on Information Visualization for text documents field of study.
- Develop understanding on the overview of the state of biomedical documents available online i.e: the numbers of published research published online and format of the publication.
- Selecting research papers related to the project title.
- Literature review analysis for the past research papers selected.
- Drafting Extended Proposal.

Week 6

- Submission of Extended Proposal.

Phase 2 – Week 7 to 8

- Preparation for Proposal Defence
- Planning phase for the project should be finalised, although further changes can still be made from time to time in accommodating the project development.
- Identify the sources, tools & equipments needed to develop the application.
- Designing the application blueprint.

Week 9

- Oral presentation for proposal defence.

Phase 3 – Week 10 to 12

- Preparation of interim report.

Week 11

- Submission of draft for interim report.

Week 14

- Submission of interim report.

FYP 2 Timelines

No	Details/Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Project Work Continues															
2	Submission of Progress Report															
3	Project Work Continues															
4	Pre-EDX															
5	Submission of Draft Report															
6	Submission of Dissertation (soft bound)															
7	Submission of Technical Paper															
8	Oral Presentation															
9	Submission of Project Dissertation (Hard Bound)															

Figure 3.0

Phase 4 – Week 1 – 12

- Development of the web application for information visualization tools for biomedical documents.
- Testing of prototype.

Week 8

- Submission of progress report.

Week 11

- Pre- EDX

Week 12

- Submission of draft report.

Week 13

- Submission of dissertation
- Submission of technical paper

Week 14

- Oral presentation

Week 15

- Submission of project dissertation (Hard Bound)

Tools Required

- PHP
- mySQL
- CSS

More tools and equipments shall be identified later on during the planning and design phase of the web application.

Chapter 4: Result and Discussion

4.1 Information Visualization Techniques

Information visualization can be done using several different techniques. There aren't specific techniques to be used or specific guidelines that specify which visualization is the best to be used in certain field. Hence, the visualization techniques are not limited and are open for more discoveries from researchers.

From previous studies, there are several visualization techniques that have been a favorite in visualizing large volume of data. As the studies developed, researchers are focusing on the alternative to transform hierarchical approach into a more guided alternative for user in dealing with the visualization.

Graphical trees visualization has been one of the most used visualization in forming a hierarchical representation of data. Tree-mapping allow information to be hierarchically organized and presented in a co-related tree structure.

1. Hyperbolic Trees

   Tree-map is a static visualization of the information with quite a number of limitations especially in screen viewing aspect. Hyperbolic Tree is the dynamic form of tree-mapping that offers an effective way to display complex trees clearly with screen-size no longer the limitation.

   Previous works suggested that hyperbolic tree visualization is suitable for data types that are hierarchical in nature and mostly used to display web information. Hypertree are also often used to better visualize the content and relationship between the information residing in the large volume of data. Hyperbolic trees are being represented in hyperbolic browser in where user can view the generated hyperbolic trees with features such as zooming and re-direct to another page for links applied from the ‘‘nodes’’ in the

hyperbolic-tree. The browser enables user to zoom or focus on a certain ‘‘nodes’’ and its related ‘‘nodes’’.

Hyperbolic tree are constructed by determining the root at the centre which also serve as the main topic being visualize. The tree was then being developed by adding ‘‘nodes’’ to further elaborate the information that were being visualize.

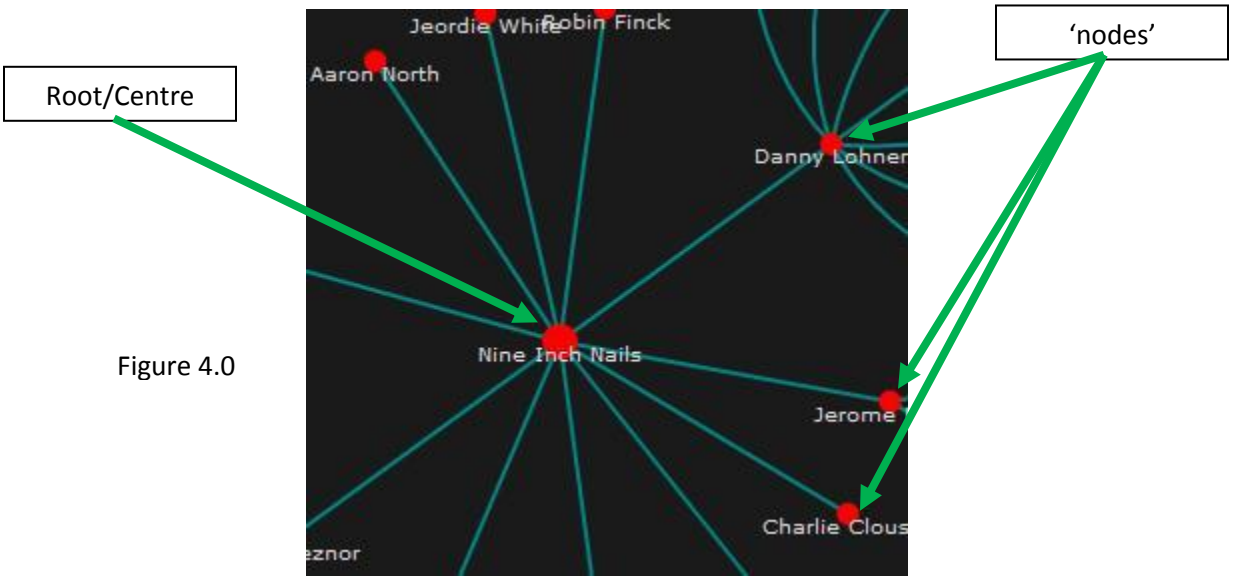


Figure 4.0

The positioning of ‘root’ and ‘nodes’ in hyperbolic tree diagram.

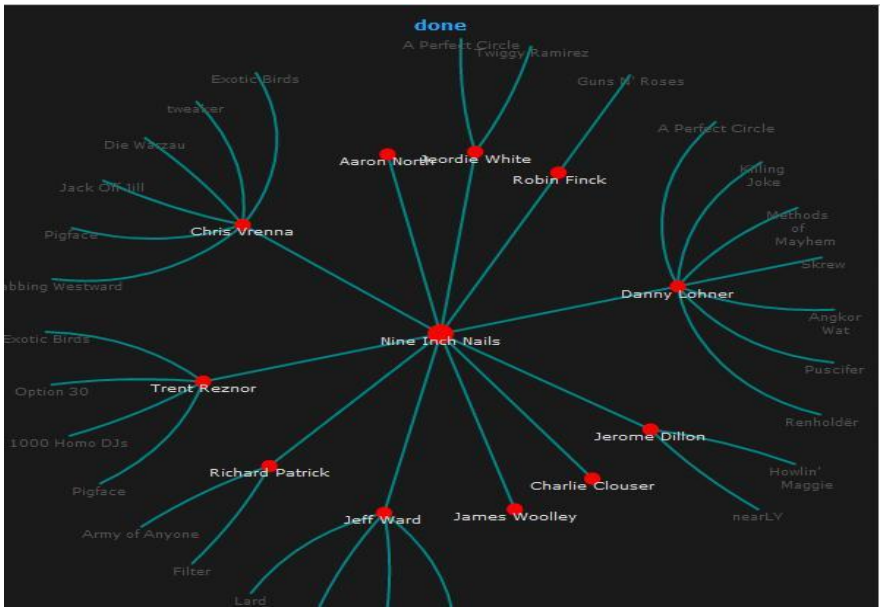


Figure 5.0

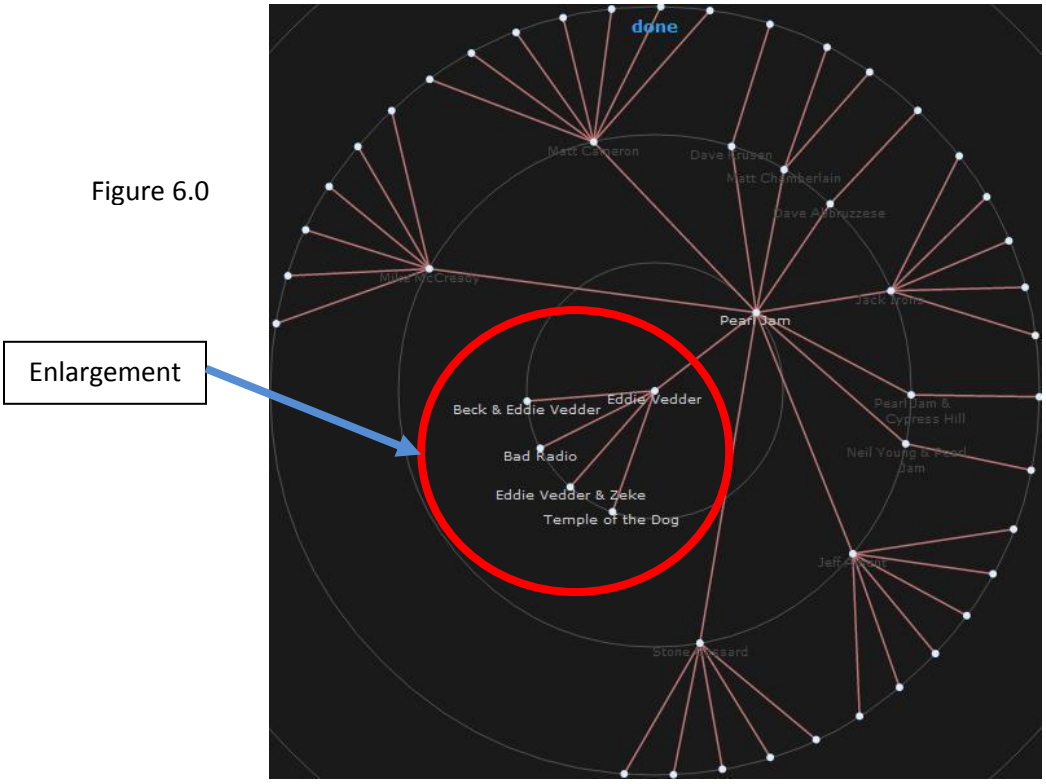
The illustration of hyperbolic tree diagram with ‘root’ and ‘nodes’.

Despite the great flexibility in arranging the 'nodes' using hyperbolic technique, there are several issues concerning the techniques. Among the issues are regarding the arrangement of the 'nodes'. Some 'nodes' that are very close or directly link to root might get positioned very far from the root due to spacing arrangement of the 'nodes'. Other than the tree diagram would be not be nice for the view, it can increase the risk of user to missed the 'nodes' all together due to the farther position from the 'root'. The feature of focusing and zooming is helpful in allowing user to better see the relation between 'nodes' up close, however it as well limits user's view of the entire hyperbolic tree diagram.

2. Radial tree diagram

Radial tree diagram is similar in structure with hyperbolic tree diagram, however radial tree diagram is emphasizing on the importance of ability to view the whole hierarchy even after focusing on one single node.

Figure 6.0



Above illustration for radial tree shown that the node being focused on is being bring to the centre and highlighted, however the overall hierarchical relation is still visible as radial tree diagram restricted the span of the nodes of same hierarchy.

Radial tree diagram are constructed with the aim to fit everything on a single screen view. The nodes are being positioned within the given sectors which are the remaining space left after the "first nodes", nodes that are directly attached to "root" are forms. This construction style makes sure that the nodes will not overlap with one another.

As per its purpose, radial tree have the limitation in term of spacing and focusing features. The focusing can be done, but with limited enlargement as being illustrated in the figure above. Radial tree diagram has been suggested to be used with user's already setting their focus on the subject they are working on as the nodes growth could be out of control hence resulting in poor quality of information visualization.

3. Spacetree diagram

Spacetree diagram is a simpler version of both hyperbolic tree diagram and radial tree diagram. Spacetree is the conventional nodes linked together forming a tree diagram. It is the first form of tree diagram before later on being developed as either hyperbolic or radial tree diagram to accommodate the limitation of viewing.

Spacetree can clearly show the relation between the nodes and the parent nodes as it is very directed in construction. The branches are predefined and are not being enabled for expansion. This way the navigation is made more clear and guided. User can get the general idea of the information supplied without having problems of having too many branches of same level.

The construction of spacetree starts with the subject/title also known as the parent node at either end of the space (top, bottom, left, right) and the nodes are being expanded in the provided spaces. Spacetree browser allows the nodes that user didn't wish to view to be hidden and concentrates on the nodes required by the users.

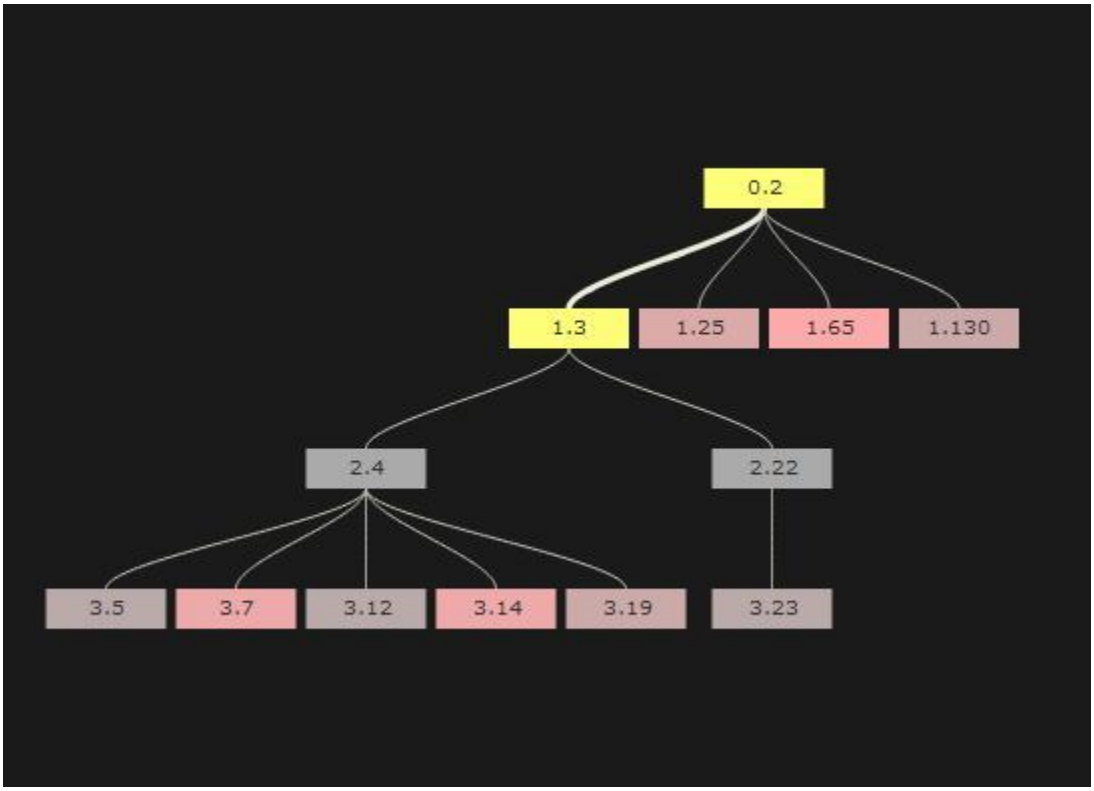


Figure 7.0

The figure above illustrates the hierarchical organization of the spacetree. The last nodes can be expanded provided the user chose to view it as such.

4.2 Hierarchical Visualization – Spacetree Diagram

This project had decided to pursue spacetree diagram as the visualization technique of choice. Spacetree diagram is the first form of hierarchical tree diagram representation. Spacetree diagram will be able to clearly shown the hierarchical relationship of the visualization intended for this project.

Spacetree diagram is easily expandable without having to make changes to the hierarchical representation of the visualization. The expandable ability is important as the end product is designed to accommodate to large amount of document. The relationship between the documents shall be made possible only if the tree diagram is able to expand according to needs.

Although it has not been proven in previous studies, the success of manipulating spacetree diagram as information visualization might determine the relevancy of using tree diagram as the visualization technique in visualizing biomedical documents information. The other visualization techniques, which are the revolution of spacetree diagram, will also be possible as the best visualization representation for biomedical documents when spacetree diagram is a success.

4.3 Application algorithm

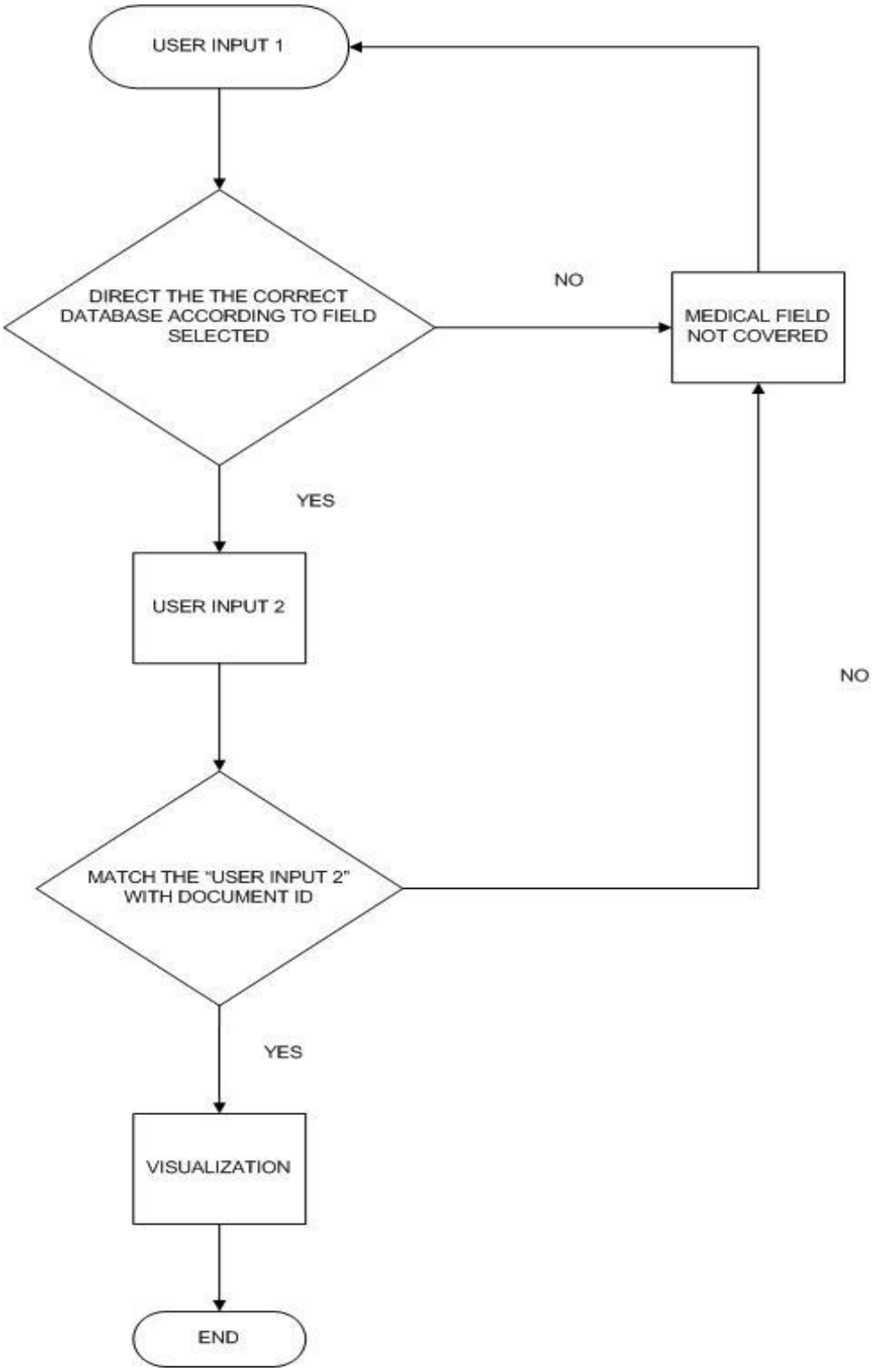


Figure 8.0

User will trigger the application by providing two inputs through the dropbox menu in the interfaces. USER INPUT 1 will be used to determine the selected medical field (i.e: Alzheimer's, Dementia) in which the user in interested in. If the said medical field in not covered in the database, user will be prompted to re-select the available medical field provided.

USER INPUT 2 will determine the area of the medical field that the user is interested to learned more about .(i.e: Symptoms, Treatment). If the area of the medical field is not covered in the database, the system will direct the user back to the search options. Once the area of medical field is selected, USER INPUT 1 & 2 will do the matching processes with the tags which also served as the document's ID.

Once the matching is done, the result will be represented as tree diagram visualization. User will be able to see the information resides in the documents available through the connecting nodes that show the relationship of each documents to the area of medical field intended by the user.

4.4 Information Visualization using spacetree diagram application

User will be presented with the first page of the web application, in tab "Home" in where they will find the description regarding the intention and usage of the web application. User guide is being provided in the box section on the left of the page. The same user guide will be appearing at all user accessible pages in the application. The purpose is to make the instruction available throughout the navigation of the web application, except for the visualization page. The other tab "Articles" is the tab in where user can later find the list of all articles available in the database of this web application.

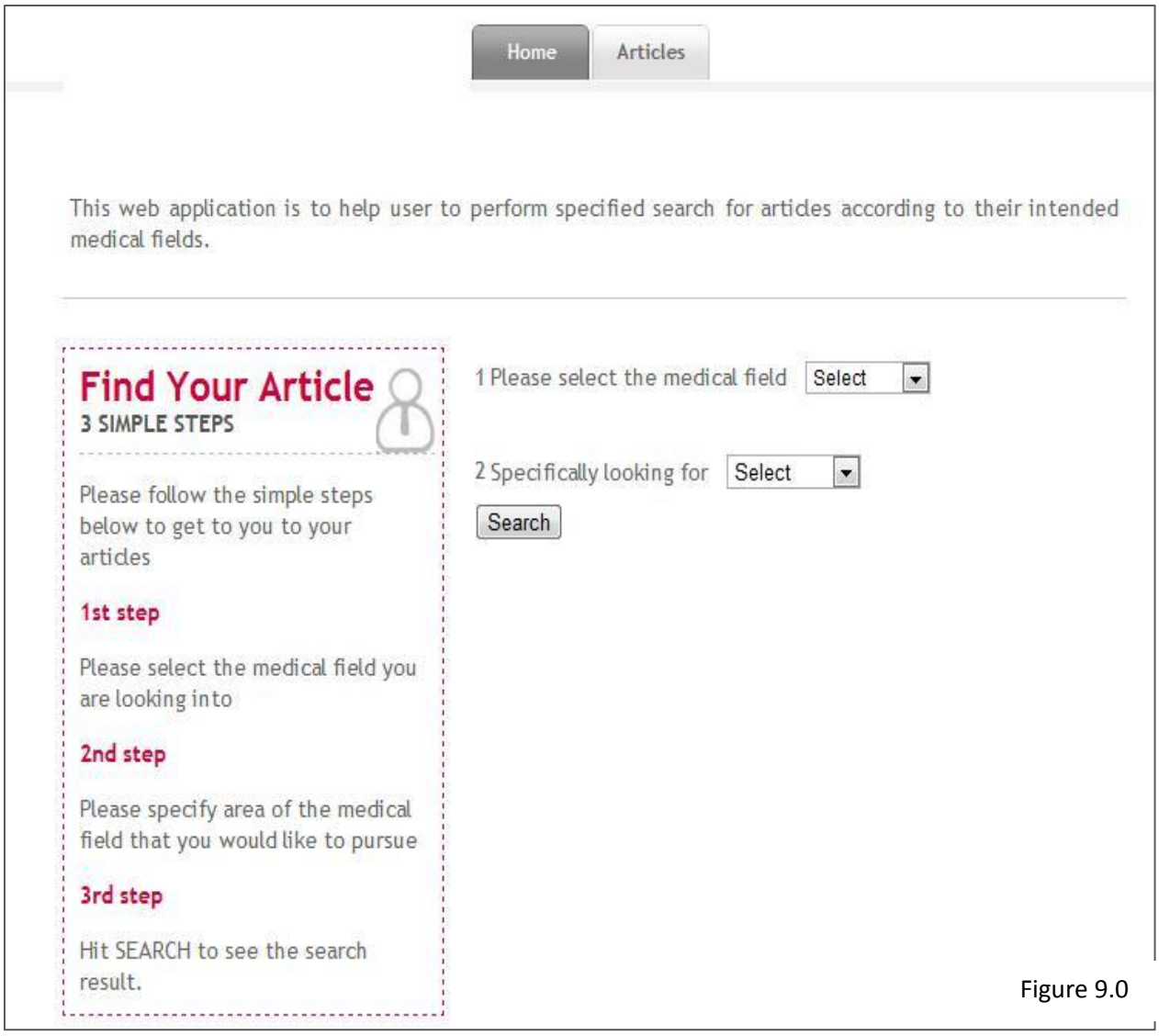


Figure 9.0

The first step is for user to select the medical field they intended to search for. As for this project, there has been two medical field made available 1) Alzheimer and 2) Dementia. This input from user is necessary to make sure the search is on the correct medical area.

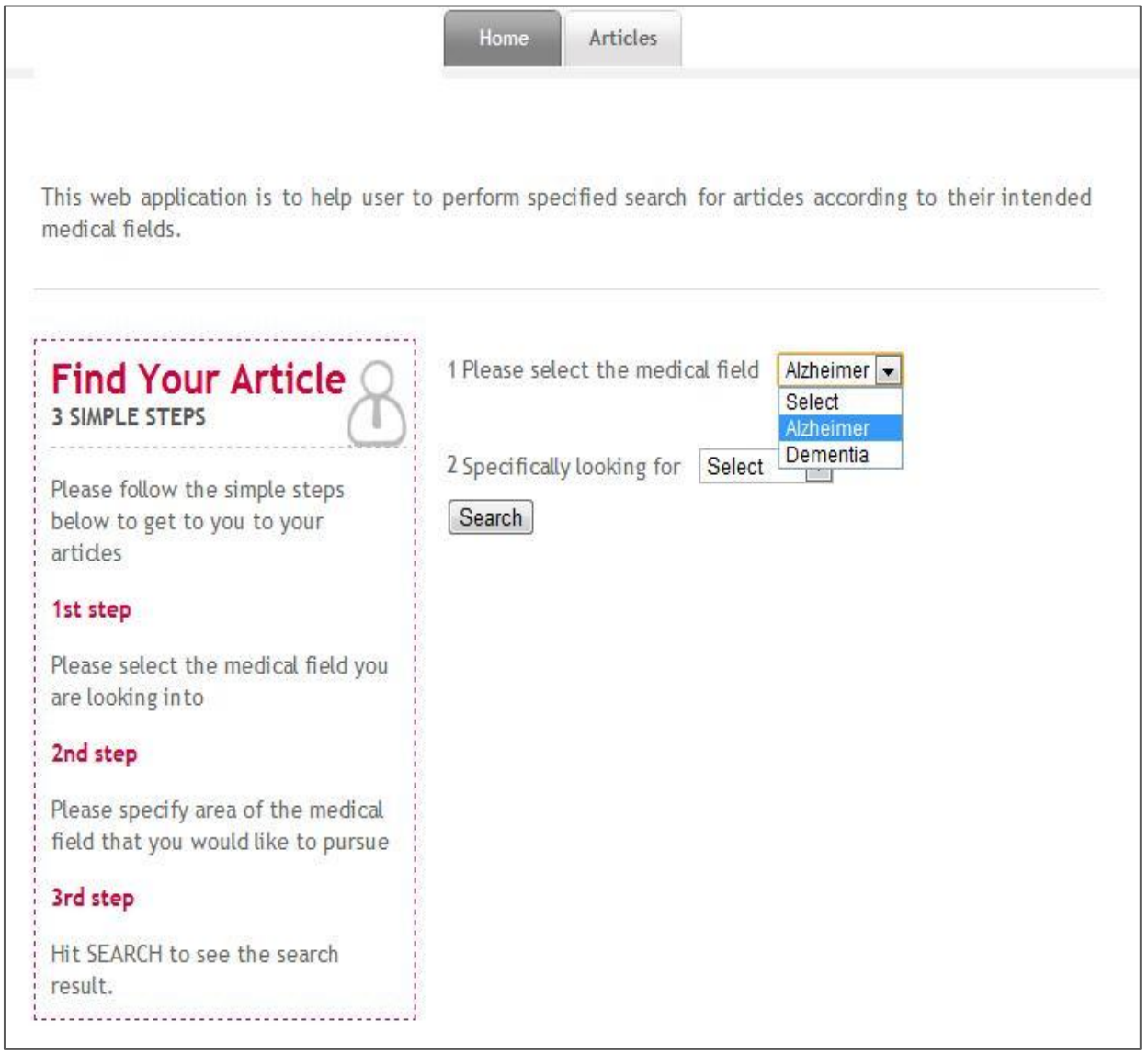


Figure 10.0

The next step is for user to select from the list available the specific area in the medical field that they are really looking for. The options made available in this project are 1) Symptoms 2) Prevention and 3) Treatment. User input for this part is important in providing users with articles on relevant to the area of their interest.

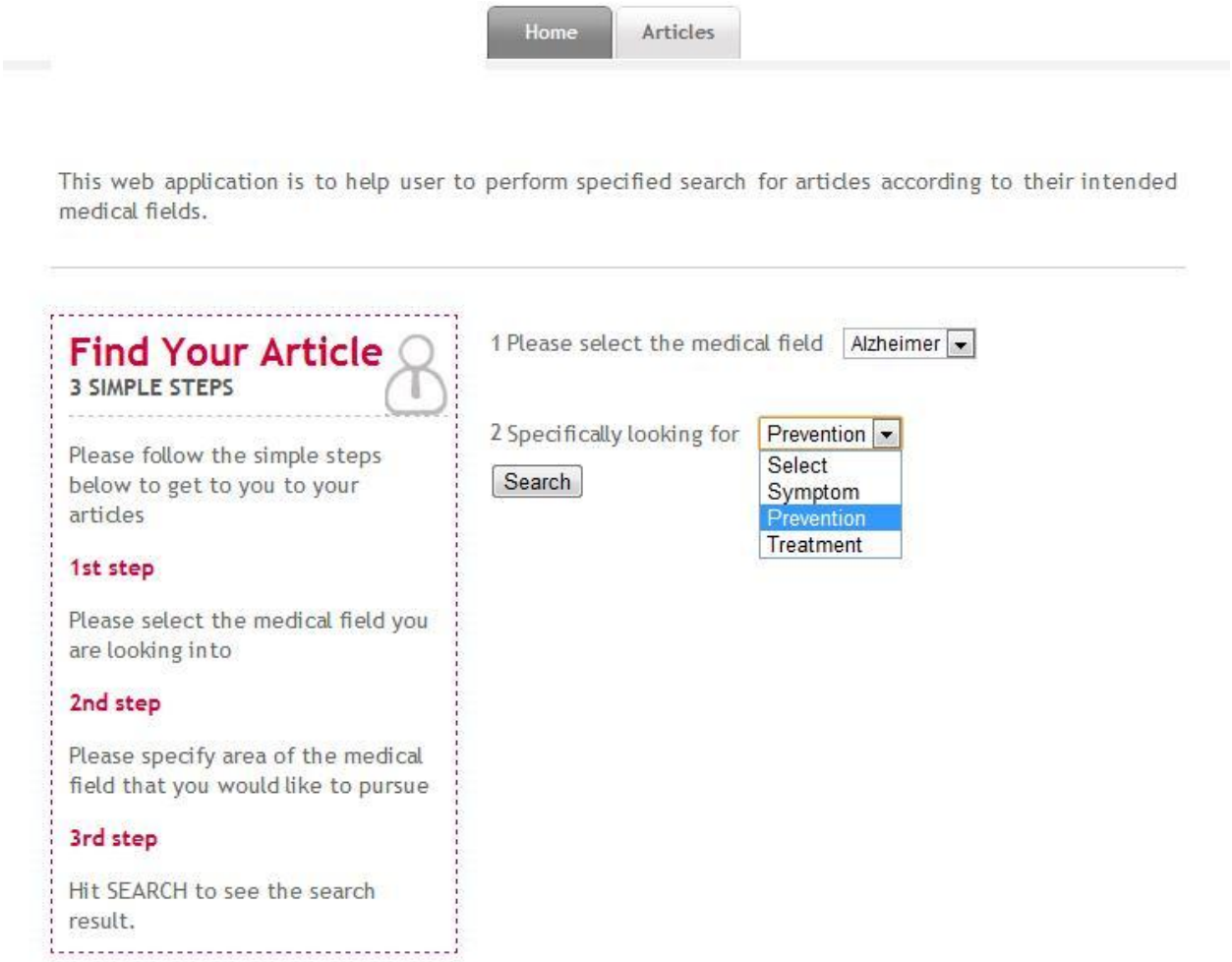


Figure 11.0

After the selections are made, user will need to click on the "Search" button to proceed with the search result.

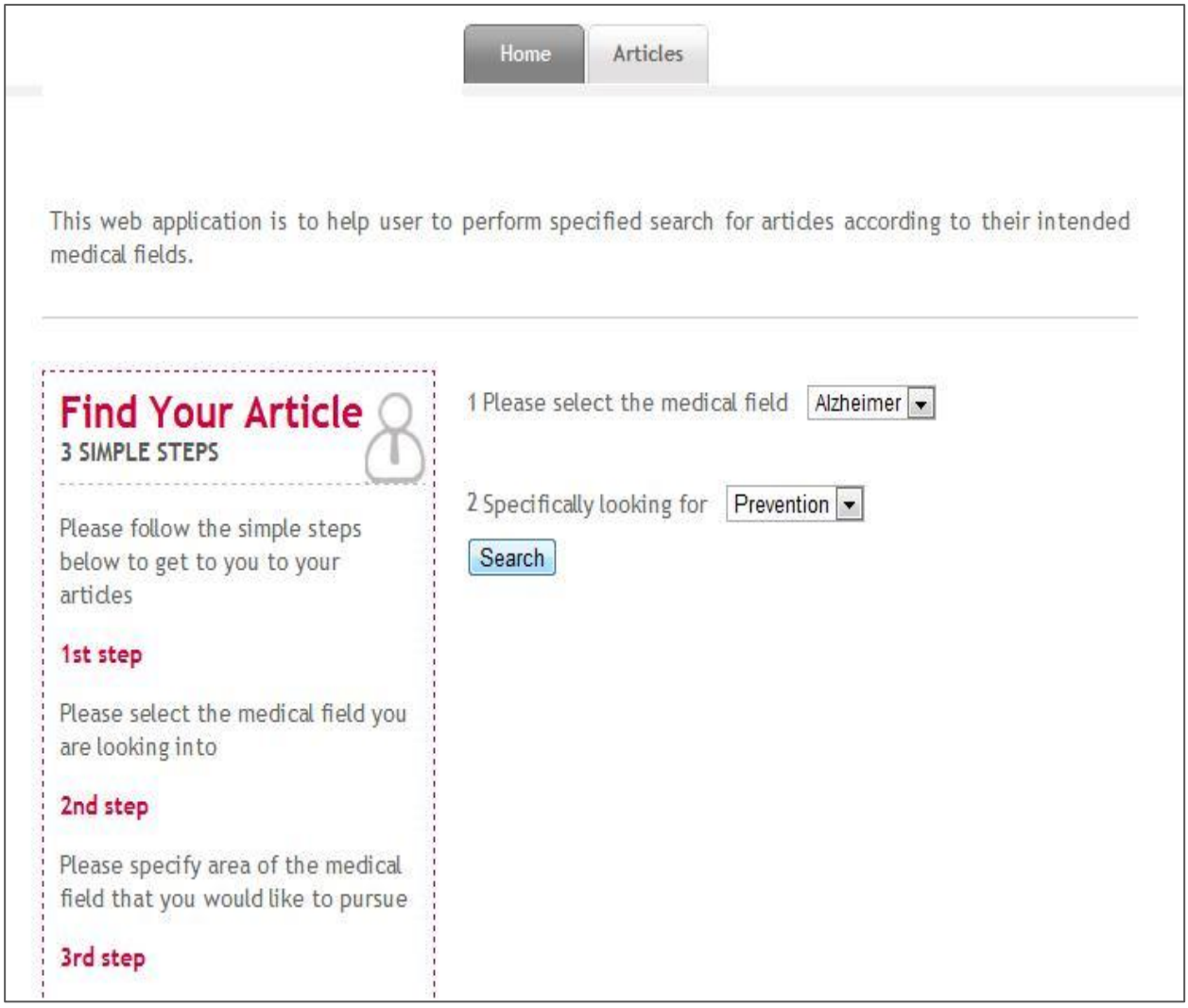


Figure 12.0

The search result will be presented as been shown in the figure below. The tree diagram is being represented in horizontal hierarchy. The root is the medical field being selected by the user in the previous page which is “Alzheimer”. The second branch is the area of the medical field which the user is specifically looking for which is “Prevention”. The third branches are the grouping of the suggested articles under specific prevention subject that can be found in the area of the medical field.

The visualization manage to portray the information resides in the bundle of articles available in the database by helping to navigate using through the document with the help of the grouping done by the third branch. From the figure below, it can be seen that the visualization had manage to present the articles related to prevention of Alzheimer’s illness in three category, which is prevention through “Assesment”, “Neuronal SIRT1” and “Neuroimaging”. The visualization had managed to portray the preview of the information in the articles.

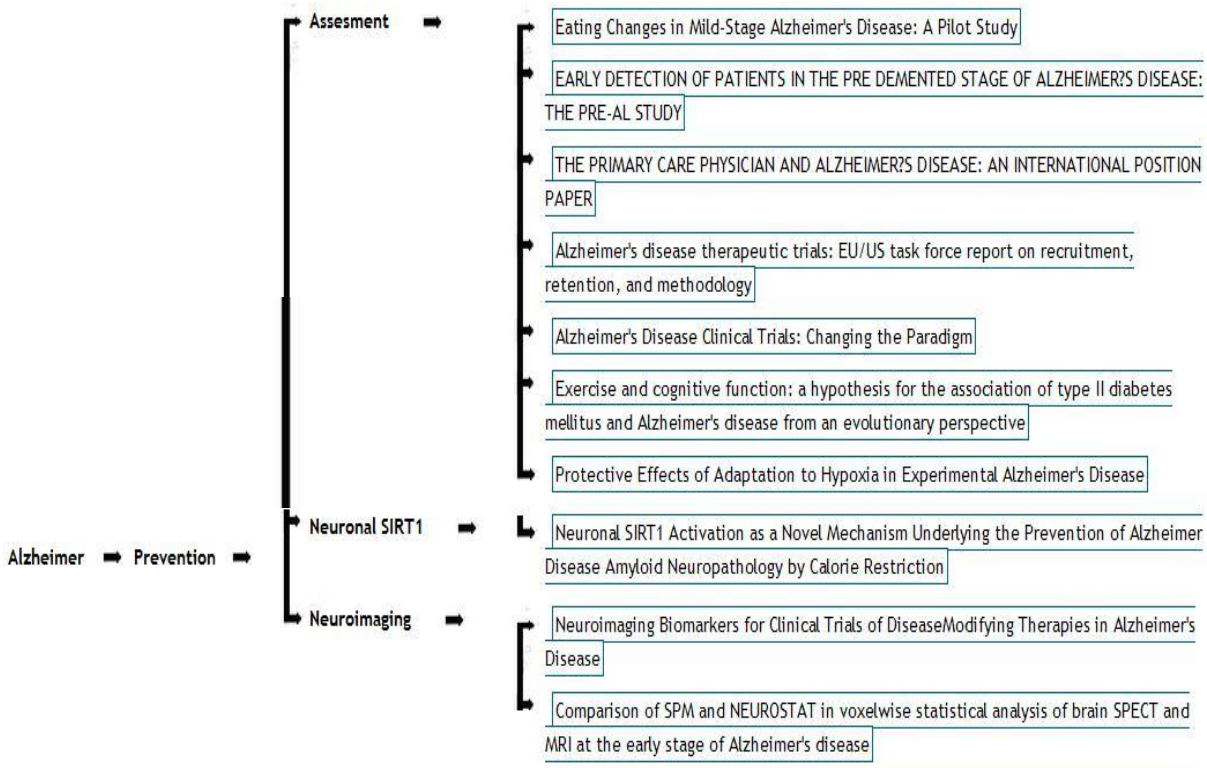


Figure 13.0

In the figure below, the first article's title is being highlighted as it is link to another page that will provide more detailed information regarding the content of the article.

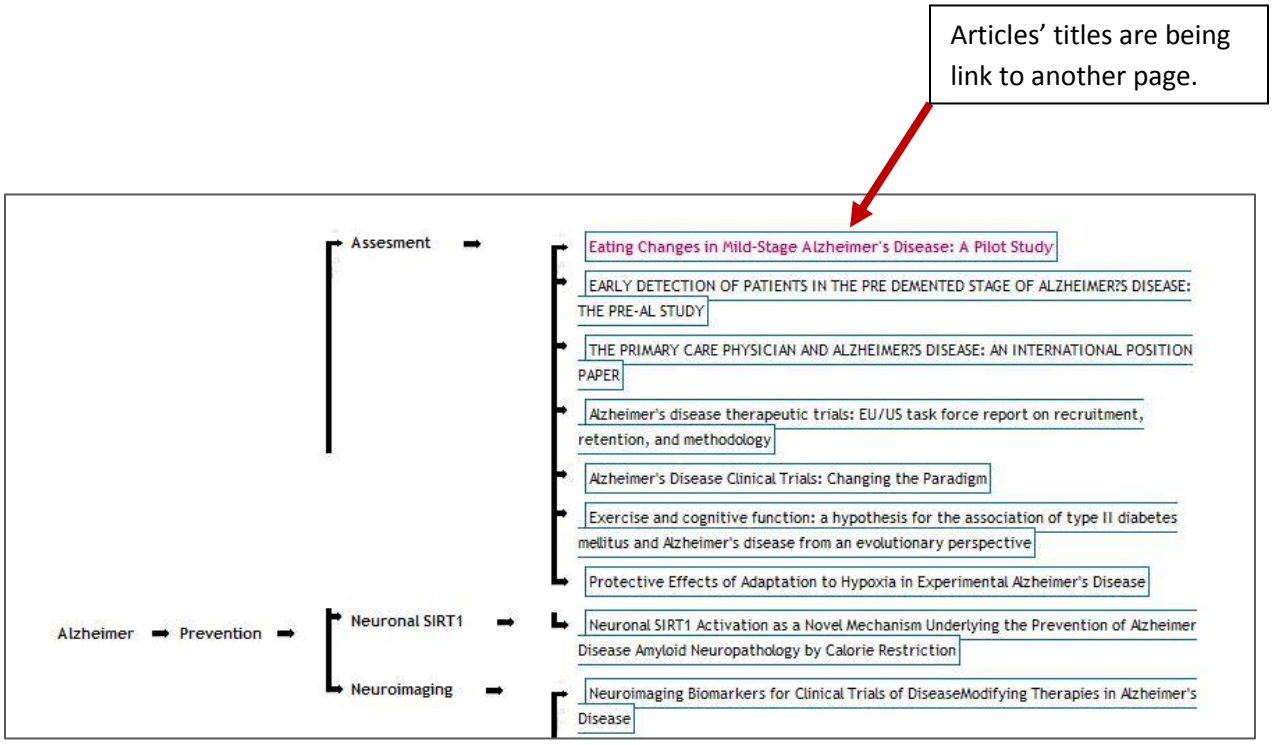


Figure 14.0

The page with more information on the articles, directed after user click.

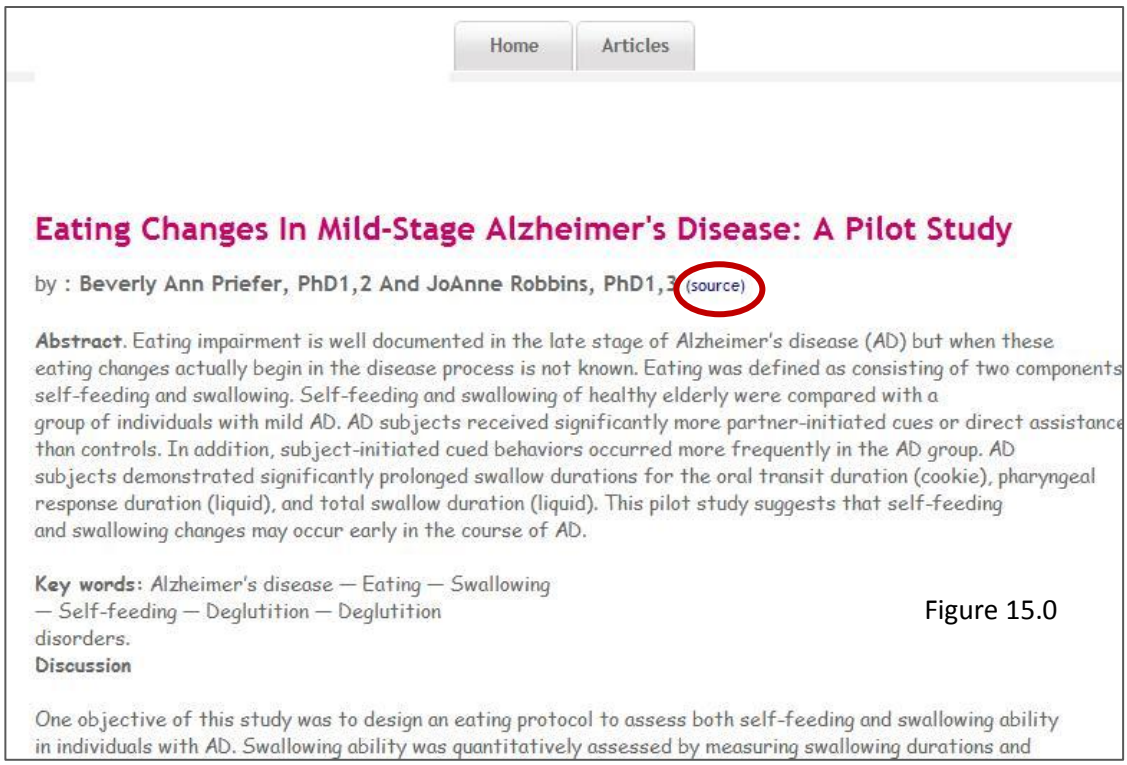


Figure 15.0

In Figure 15.0 above, the red circle is indicating the link for user to click to get to the original articles which are available online. Most biomedical documents are protected under the copyright act regardless of the country of origin; it is unethical to store the articles in the web application database without consent from the publisher.

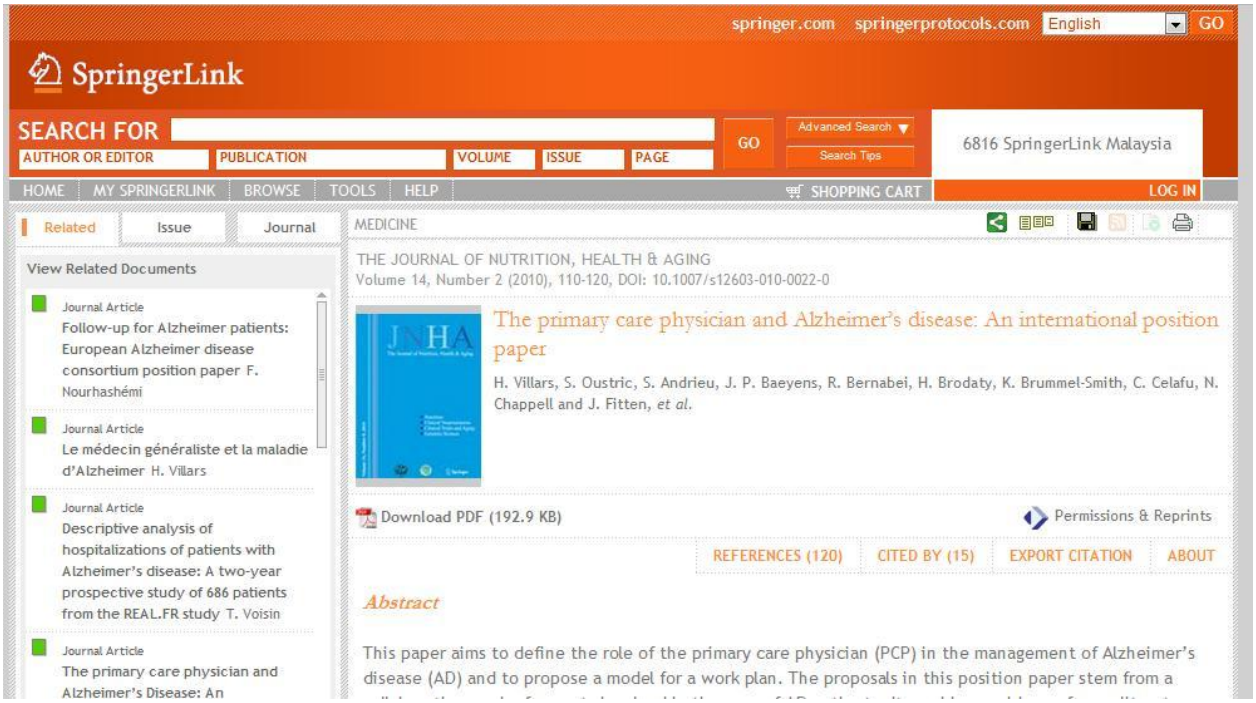


Figure 16.0

Example of webpage being link when user clicks "Source"

4.5 User Advantage

User testing session had been carried out to further test the objectives which concerning user's benefits from using the web application. The testing was done by demonstrating the product to user and allowing them to work the application hands on. The results of the testing are shown as below. The questionnaires scripts are provided in the appendices as proofs. A total of 10 respondents had participated in the product testing and questionnaires. The respondents are medical students.

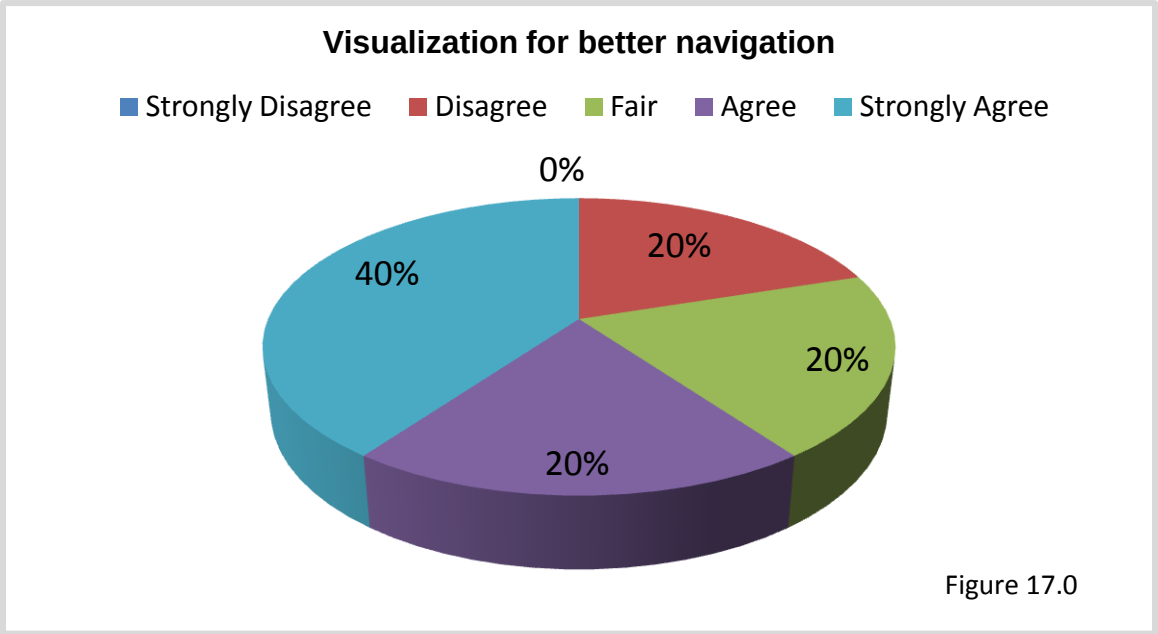


Figure 17.0

User testing result had showed that 40% of the respondents strongly agree that the visualization helps in providing better navigation through the search results. Another 20% agrees that visualization did help in navigation while another 20% are neutral about having visualization in providing better navigation in going through search results. The rest 20% disagree that visualization helps in navigation. None of the respondent strongly disagrees.

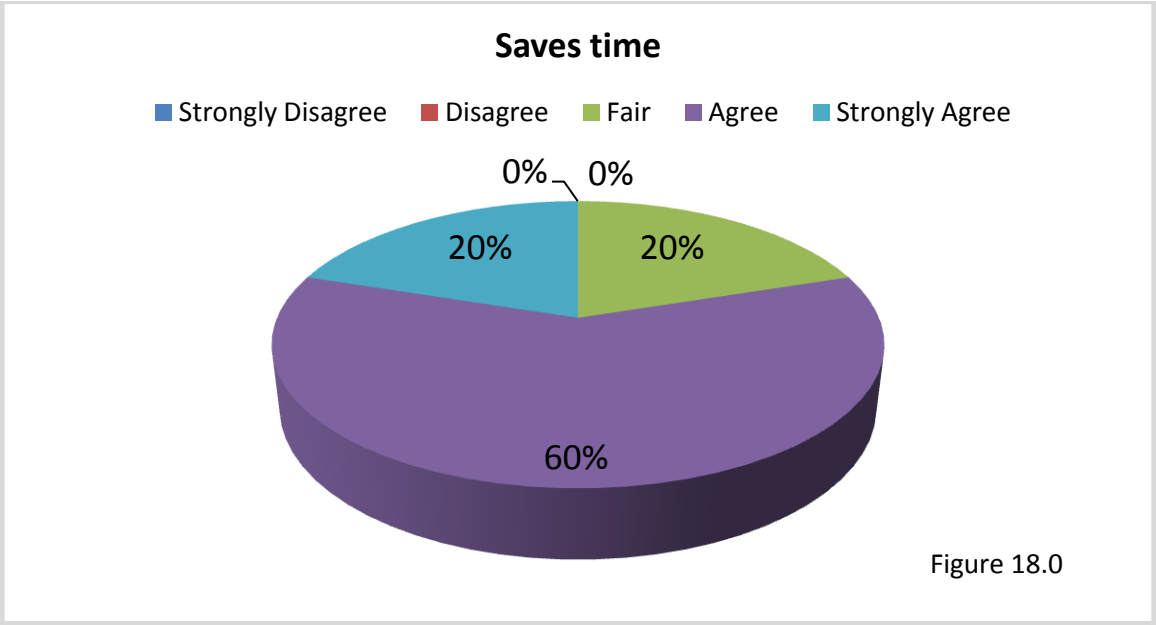


Figure 18.0

The total of 60% of the respondents strongly agrees that the visualization representation approach help in saving their time to get to the correct articles that they are looking for. Another 20% agrees and the rest 20% are neutral regarding the visualization approach in helping them saved time in looking for the articles. None of the respondents disagree that visualization of the search results had helped in saving their time. The respondents find it helpful that the search process is specified, it helps them to browse through the possible suggestions of articles much faster.

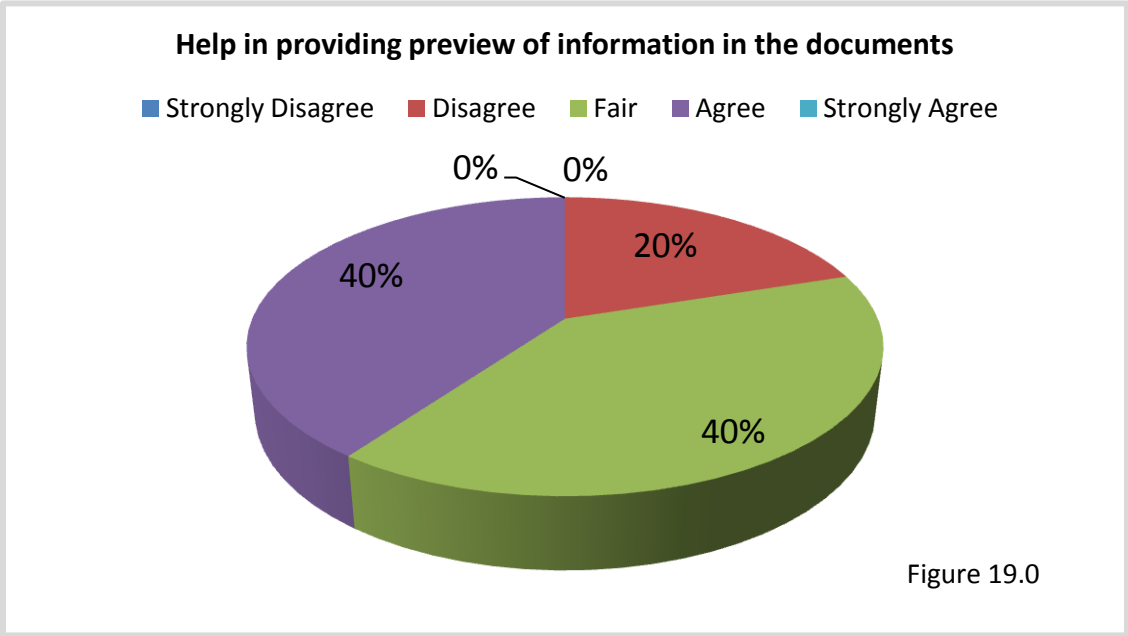


Figure 19.0

A total of 40% of all respondents agrees that the visualization helps in providing previews of the information contains in the documents retrieved. Another 40% found it fair that the web application did provide information and the rest 20% disagrees. The main reason for the lack of positive responses lies on the limitation of advanced medical knowledge by the author and the limitation of resources for biomedical texts.

Chapter 5: Conclusion and Recommendation

The objective of this project is to provide a guided visualization for users in navigating through the biomedical documents suggested by the retrieval system.

Hence the information visualization techniques has been thoroughly studied and several most relevant visualization techniques that applies to this projects has been reported in the as result and discussion for this progress report. One of the visualization has been chosen as the visualization to be developed in completing this project and testing the value of information visualization for biomedical documents search results.

The visualization technique chosen is spacetree diagram, the first form of hierarchical tree diagram. Spacetree diagram built is simple but compact, it supports the expansion of branches and able to portray clear hierarchy structure. The project has been carried out to develop a spacetree diagram visualization that manages to provide general information for the user without them needing to read the articles.

The algorithm of the project has undergone some changes but the general idea and flow of the application is the same. The end product has been developed using PHP, Cascading Styling Sheets and mySQL as the database. The software used might not be the most difficult and advanced but it was suitable in helping this project to achieve its objectives.

Due to time and work-power constraint, only one visualization technique is able to be tested in visualizing biomedical documents search results. As being reported in the result and discussion part of this project, there are several more visualizations that have the potential to be developed as the next visualization representation. This project recommendation for future works would be for other visualization technique to be developed and presented as another useful information visualization tools, not only limited for biomedical documents but for any other industry as well.

For future development of this project, the author would like to make few recommendations on the features to be improves. The visualization can be further developed by completing the other form of visualization techniques. The improvements can be detected through the comparisons of all the visualization techniques which had been reported in the discussion part of this report.

For this project the visualization only able to provided suggestions for articles related to the specified search, for future developments more typed of media such as videos, audios and photos are for consideration to be included in the visualization. This is to accommodate to more wide users, as reading is not everyone forte, other types of medias might me for helpful for them in browsing through the search results.

The project needs expertise in biomedical field and multimedia field to be able to progress in a more promising way. The information previewed in the visualisation is relatively simple and general due to the author limitations in biomedical fields. The help of expertise in cracking the code of biomedical text would be very helpful in providing a more detailed information preview for the biomedical text suggested in the visualization. The project is in need of more resources for biomedical text in order to be able to provide more information for the users. The involvement of experts and larger biomedical library will results in a large cost incurred. Thus this development is only possible with large funding.

The other alternative is to convert this web application into some sort of reading trackers for medical practitioners. This way, the issues of limited resources will be eliminated as the medical practitioners are responsible for their own resources. However, more features need to be included for the web application to be fully controlled by the users, such as adding more friendly features and more clear instruction on the usage of the web application.

References

[1] H. Zhang, M. Fiszman, D. Shin, C. M. Miller, G. Rosemblat and T. C. Rindflesch, "Degree centrality for semantic abstraction summarization of therapeutic studies," *J. Biomed. Inform.,* 2011.

[2] J. D. Duke, X. Li and S. J. Grannis, "Data visualization speeds review of potential adverse drug events in patients on multiple medications," *J. Biomed. Inform.,* vol. 43, pp. 326-331, 2010.

[3] B. Herr, E. M. Talley, G. A. P. C. Burns, D. Newman and G. LaRowe, "The NIH visual browser: An interactive visualization of biomedical research," in *Information Visualisation, 2009 13th International Conference,* 2009, pp. 505-509.

[4] T. C. Rindflesch, H. Kilicoglu, M. Fiszman, G. Rosemblat and D. Shin, "Semantic MEDLINE: An advanced information management application for biomedicine," *Information Services and use,* vol. 31, pp. 15-21, 2011.

[5] R. P. Futrelle, J. Satterley and T. McCormack, "NLP-NG — A new NLP system for biomedical text analysis," in *Bioinformatics and Biomedicine Workshop, 2009. BIBMW 2009. IEEE International Conference on,* 2009, pp. 296-301.

[6] M. Abulaish and L. Dey, "A concept-driven biomedical knowledge extraction and visualization framework for conceptualization of text corpora," *J. Biomed. Inform.,* vol. 43, pp. 1020-1035, 2010.

[7] Indiana University, "Radial Tree," vol. 2003, 2004, 2004.

[8] Jepson Herbaria of the University of California, "Deep Green Hyperbolic Tree," vol. 2010, .

[9] Lamping John, Rao Ramana, Pirolli Peter, "Hyperbolics Trees," vol. 2011, March 15, 2011.

[10] Lamping John, Rao Ramana and Pirolli Peter, "A Focus Context Technique Based on Hyperbolic Geometry for Visualising Large Hierarchies," vol. 1995, 1995.

[11] lewellen, "Tree Drawing : The Radial Tree," 2008-08-31, 2008.

[12] N. Sheth and Q. Cai, "Visualizing mesh dataset using radial tree layout," 2003.

[13] P. Catherine, G. Jesse and B. B. Benjamin, "SpaceTree: Supporting Exploration in Large Node Link Tree, Design Evolution and Empirical Evaluation," .

APPENDICES

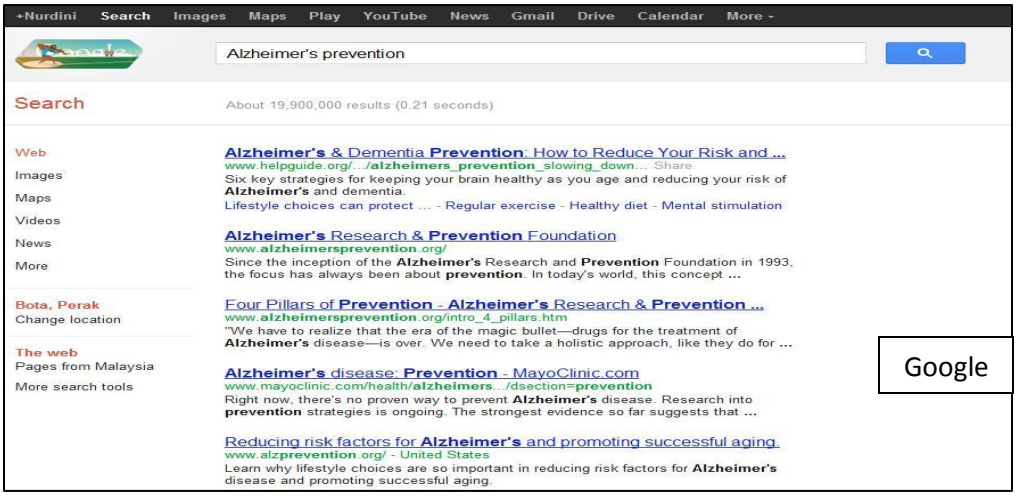
Appendix 1: Questionnaires

Information Visualization for Biomedical Text

Occupation: **Describe how you are related to medical field:**

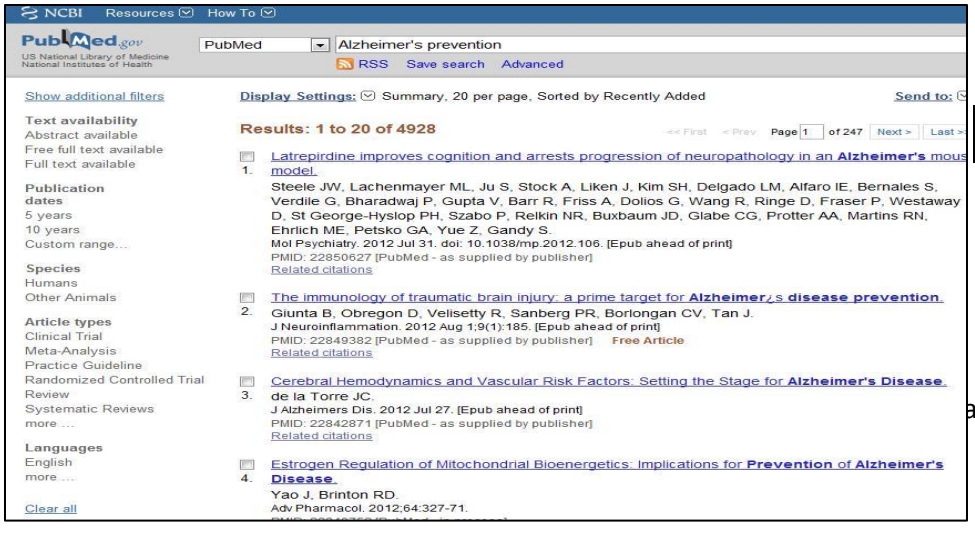
This questionnaire is designed to test the effectiveness of information visualization approach used in this project.

Search result representation by online search result:



Google

Bing



PubMed

Please refer to the examples provided in the first page of this questionnaire.

a. General question

1. Are you familiar with the representation of search result provided by online search result such as Google, Bing and PubMed?

Yes	
No	

2. The search result representation provides a preview of the content of the documents suggested.

Strongly Disagree	Disagree	Fair	Agree	Strongly Agree

3. It is easy to navigate through the search result to find documents with the same specific topic related i.e: Alzheimer's prevention by Neuroimaging.

Strongly Disagree	Disagree	Fair	Agree	Strongly Agree

The questions below are to be answered with close reference to the visualization presented by the web application developed for this project.

b. Visualization for better navigation.

1. The spacetree diagram representation helps in assisting user to navigate through the search results in a more guided manner than the traditional listing representation.

Strongly Disagree	Disagree	Fair	Agree	Strongly Agree

2. The grouping presented after the 3rd branch of the spacetree diagram help user to see the connection of the documents suggested by the search results.

Strongly Disagree	Disagree	Fair	Agree	Strongly Agree

c. Help in providing preview of information in the search result and documents suggested.

1. The spacetree diagram representation allow user to have a preview of the specified search they would like to pursue (i.e: symptoms, prevention)

Strongly Disagree	Disagree	Fair	Agree	Strongly Agree

2. The abstract, introduction and excerpt provided in the web application help to provide more detailed information for the user.

Strongly Disagree	Disagree	Fair	Agree	Strongly Agree

d. The specified search and spacetree diagram visualization approaches help in reducing time used for user to find their intended articles.

1. The specified search approaches lessen the time taken for user to find the article related to their search interest.

Strongly Disagree	Disagree	Fair	Agree	Strongly Agree

2. The spacetree diagram visualization guides user to find their articles faster.

Strongly Disagree	Disagree	Fair	Agree	Strongly Agree

Biomedical Document Retrieval System Using Spacetree Diagram

Nurdini Binti Nordin
Business Information System
Universiti Teknologi PETRONAS
Bandar Seri Iskandar, Perak
Malaysia.

Abstract This project aim is to provide an application that support visualization of information from biomedical documents available online. The large numbers of potentially related documents retrieved by search engines ie. Google or PubMed are being represented in a non-interactive and organized form. This conventional representation of information did not help in providing meaningful information for the user in helping them to navigate through the co-related documents retrieved from the same field in biomedical studies. Information visualisation technique can further provide user with the relationship among the biomedical documents retrieved from the user search results. By forming the relationship between the biomedical documents retrieved, user can have an overview of the connection between the information resides in multiple documents. The graphical visualization are to be in the form of tree diagram whereby user can see the expansion of the information being provided by each of the biomedical documents retrieved in a guided approach.

1.0 Problem Statement

Most documents retrieval system works to only suggest potentially useful documents and provides no meaning for the automatic text summarization driven. The query results did not portray the relationship between the documents retrieved hence increases the risk of user overlooking important information. The usual representations of the retrieved results are in lists form which did not provide easy navigation throughout the retrieved results. Graphical representation i.e: mind maps, will provide better representation and easier navigation through the biomedical documents retrieved.

2.0 Project Background Natural Language Processing (NLP) field of study is the pioneer in the development of studies regarding human-computer language interaction. The further

development in NLP fields produced useful tools such as automatic text summarization, automatic text translation into another human language, information extraction from text and information retrieval. With the advancement of technology, millions of documents are being made available online. Biomedical documents are one of the most rapidly growing in number for the documents made available online. Information retrieval system has been a handy tool for user in searching through the mass of documents available online. Search engines such as Google, PubMed for biomedical documents are the readily example of information retrieval tools. The system received user input i.e: keyword and will performed the algorithm programmed in retrieving the documents related to the keyword.

The systems for retrieval of biomedical documents which are made available online are mainly focusing on retrieving the potentially useful documents from queries made by the user. The results retrieved are driven by text manipulation ie: the number of occurrences or similarity between strings of citation found in the documents compared to user's search input. These query results are being accompanied by automatic text summarization result found in the documents by referring to user's input. Most of the query results are being represented in lists which are ranked according to date of publication instead of relevancy of the information contained in the documents. The representations of the information extracted are stand-alone and portrays no relationship between one another.

Information visualization approaches in visualising biomedical information retrieved from online text and search engines query results has been researched to test the theory of information visualisation for biomedical documents can be developed to portrays the relationship between the biomedical documents. Throughout the years, researches done on visualising biomedical information had comes up with different forms of graphical visualisation to represent biomedical information from text documents. Further researches on visualising biomedical text are focussing on providing relationships between the biomedical documents in the form of graphical visualization.

3.0 Objectives of Study The objectives for this project is to develop a web

application that can provide a visual representation of the biomedical documents query results retrieved from databases or online search engines and o visualize the relationship between biomedical documents retrieved from database

4.0 Literature Review Information Visualization for Biomedical Text is a research field that has been given lots of attention by researches all over the world. Most of the researches are intrigued to find ways to manage the large amount of document available online [1],[2],[3] and means of presenting the information by using visual aids [1],[2],[4],[3],. The researchers are driven by theory of transforming automatic text summarization into semantic information through query results from large databases. Another supporting theory for this project is information visualization as the best solution in representing the structured management of the large number of documents retrieved.

The researches on this subject are being done[1] by two different approaches 1) focussing on one particular field in biomedical field[1] [1],[2],[4] and 2) involves all biomedical documents available from database i.e PubMed [5],[6],[3]. Among the area of studies being focus in the previous research papers for Information Visualization for Biomedical Text are therapeutic studies [1], adverse drug effects [2] and clock genes studies [4].

The similarities which can be found from the past researches are on the approach in managing and extracting valuable information from a large bank of biomedical documents. There are several issues being highlighted from the past works. The first concern is on how to effectively manage the results of the query retrieved in order to represent only the information that are relevant for the user[1] ,[6],[4]. The next concern is to represent the query results in a visual form that can provide useful information regarding the relationship between the documents retrieved from the query results [6]. This concern has heightened the need to enables the retrieved citation from the documents to be shown in a clear relationship between one another [3].

The methodologies used in the past works in this research field are using an almost similar sequence of methods in transforming the information retrieved into

a visual representation. These researches will retrieve query results of citation or text summarization from online search engine to be process for relationship identification among the citation before being represented visually i.e : mind maps [1], [4], bar charts [2].

contains only facts from the scientific biomedical reports or documents [5]. This research paper offers useful techniques used in extracting relevants facts from text documents hence it is included in this literature review.

In the research paper titled '*The NIH Visual Browser: An Interactive Visualization of Biomedical Research'* [3] the emphasis of the

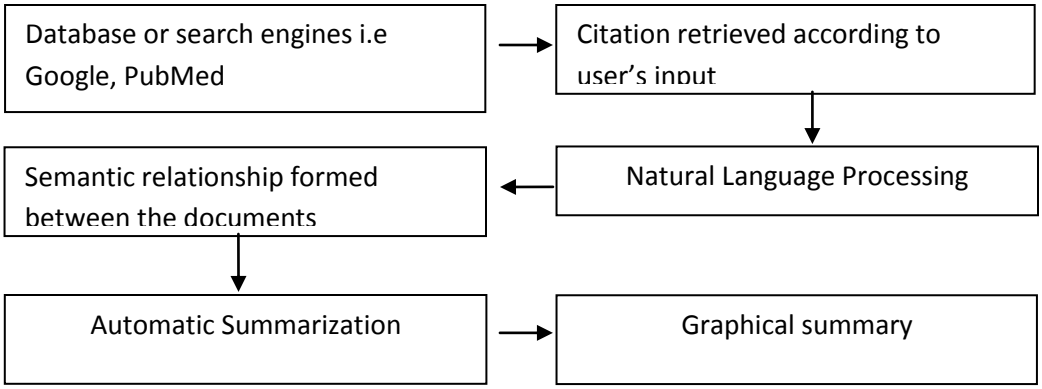


Figure 1.0 – General steps in visualizing biomedical information [5].

However each researches have a distinctive variation in several steps, as can be found in research paper titled "*A New NLP System for Biomedical Text Analysis"* [5] the researcher were focusing on refining the algorithm so as to produced a solid text summarization that contains only facts by eliminating other ancronym before the information is being represented in text format [5]. The representation of this research are not through graphical aids, instead it's a summarization of text which

research was to come out with a more interactive information visualization application [3]. The information retrieved are being represented in a big mind map of connecting dots or lines [3]. Although the visualization of the information adopted here did not gives off the major idea in the first glance, the key features of this research are on the deep zooming, selection, color-schemed code, full text query and multi level labelling features

that were embedded in the application [3]. This research was focussing on the interactive functions for the web application develop to access the database. The interactive features of this research outcomes suggested that information visualization with interactive functions is a good combination in generating semantic knowledge from explicit texts.

As being reported in the research paper titled '*A concept-driven biomedical knowledge extraction and visualization framework'* [6] the application BioKEVi was designed to represent extracted queries with graphical visualization aids according to the level of depth for the topic aquired by the user [6]. This research were focusing on visualising the relationship between the documents with similar biomedical informations. Clear representation of the relationship among the biomedical text are crutial in this research as it is the key concept to support easier navigation [6] for the user in going through the biomedical documents retrived. This research paper reports on the methods of forming the relationship among the similar biomedical documents. This research paper represents the relationship in the form of mind map with relationship keyword in between the connecting lines. Thus the mind map is made easier to navigate from one document content to another.

Information visualization for biomedical text had as well been research in providing a Decision Suport System (DSS). In the research paper titled '*Data visualization speeds review of potential adverse drug events in patients'* [2] information visualization is being implemented in DSS to expediate the reviewing of potential drugs effects on patient[2]. The system develop from this research will present a comparison of adverse drug effect into a graphical visualisation [2]. The graphical visualisation adopted in this research are mostly bar-charts as the medical practitioner need to be able to decide which adverse drug effects are more likely to occurs based on the height of the bar charts elements. The implementation of graphical visualization in the DSS system to speed the time for medical prectitioner to come to a decision before prescribing drugs to their patient[2]. This research paper suggested that information visualization can provide better repersentation of information from biomedical documents.

In research paper titled '*Degree centrality for semantic abstraction summarization of therapeutic studies'* [1] the research was done with the aimed to produce a system that can manage the large amount of citation retrieved through search engines[1] . The research had focused on visualising the citation of biomedical documents retrived from medical library i.e: PubMed [1]. These citation of biomedical text retrieved are being grouped together according to its relevancy before being projected in a mind map graphical visualization. The lines connecting the mind map are being embedded with the link to the actual document the citation was retrieved [1]. This research suggested that by only a keyword, the system will be able to generate a mind map that can help user to narrow down their searches to the specifics subjects and leads to discoveries of new information.

A more organized research following '*Degree centrality for semantic abstraction summarization of therapeutic studies'* [1] had been reported in another research paper titled '*Semantic MEDLINE: An advanced information management application for biomedicine'* [4] which were as well published on 2011 but in a another publication. In '*Semantic MEDLINE: An advanced information management application for biomedicine'*[4] , the research had been focussing on studies of Clock Genes field in biomedical [4]. This research had develop a web protal that act as an application that will retrieved query results from search engines i.e: Google or PubMed [4]. The citation retrieved will be be connected based on fact-driven before being presented in a connecting graphs that resembles a mind map [4]. The link to the actual documents will be embedded in the lines connecting the graphs. An added key feature of text preview has been embedded in the graphical visualization [4]. This research has suggested that forming relationships by analysing the citation has resulted in higher relevancy of the text being relate in the mind map.

From this literature review, all the cited works being mentioned are related to the this FYP objectives. Although some of the cited works did not fulfilled the aimed to visualized information [5], the methods and steps reported to achive the goal of this projects are helpful as references for the next phase of this project. These research papers had as well serve as a proof that information visualization for biomedical documents is possible.

Contribution of information visualization in providing better representation of information from biomedical documents had as well been proven from these research papers reviewed.

5.0 Methodology This project will be developed using Rapid Development methodology. This project requires detailed analysis from previously similar or in the same field research papers. Finding from these research papers might changes the detailed problem indentified in the early development of this project, hence by adapting rapid development methodology this project will be flexible for changes during the development cycle. Rapid methodology emphasized on the user's feedback regarding the prototype which will provide useful inputs for the development of this project.

6.0 Results & Discussion

6.1 Information Visualization Techniques

Information visualization can be done using several different techniques. There aren't specific techniques to be used or specific guidelines that specify which visualization is the best to be used in certain field. Hence, the visualization techniques are not limited and are open for more discoveries from researchers.

From previous studies, there are several visualization techniques that have been a favorite in visualizing large volume of data. As the studies developed, researchers are focusing on the alternative to transform hierarchical approach into a more guided alternative for user in dealing with the visualization.

Graphical trees visualization has been one of the most used visualization in forming a hierarchical representation of data. Tree-mapping allow information to be hierarchically organized and presented in a co-related tree structure.

4. Hyperbolic Trees

Tree-map is a static visualization of the information with quite a number of limitations especially in screen viewing aspect. Hyperbolic Tree is the dynamic form of tree-mapping that offers an effective way to display complex trees clearly with screen-size no longer the limitation.

Previous works suggested that hyperbolic tree visualization is suitable for data types that are hierarchical in nature and mostly used to display web information. Hypertree are also often used to better visualize the content and relationship between the information residing in the large volume of data. Hyperbolic trees are being represented in hyperbolic browser in where user can view the generated hyperbolic trees with features such as zooming and re-

direct to another page for links applied from the ‘‘nodes’’ in the hyperbolic-tree. The browser enables user to zoom or focus on a certain ‘‘nodes’’ and its related ‘‘nodes’’.

Hyperbolic tree are constructed by determining the root at the centre which also serve as the main topic being visualize. The tree was then being developed by adding ‘‘nodes’’ to further elaborate the information that were being visualize.

The positioning of ‘root’ and ‘nodes’ in hyperbolic tree diagram.

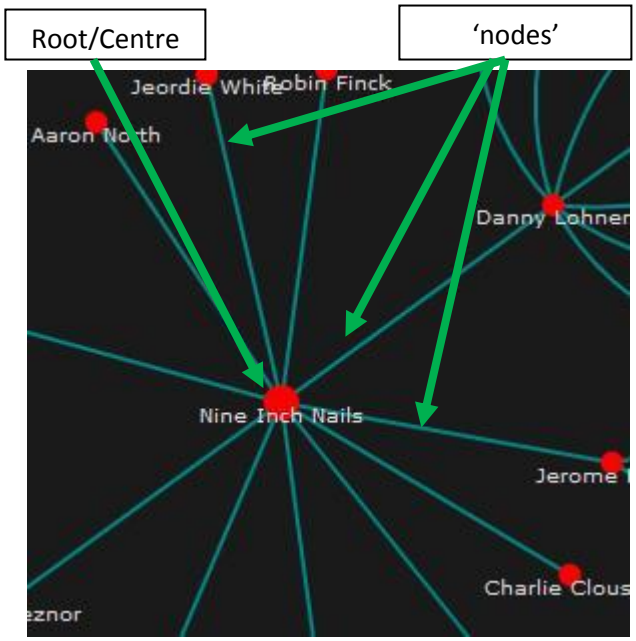
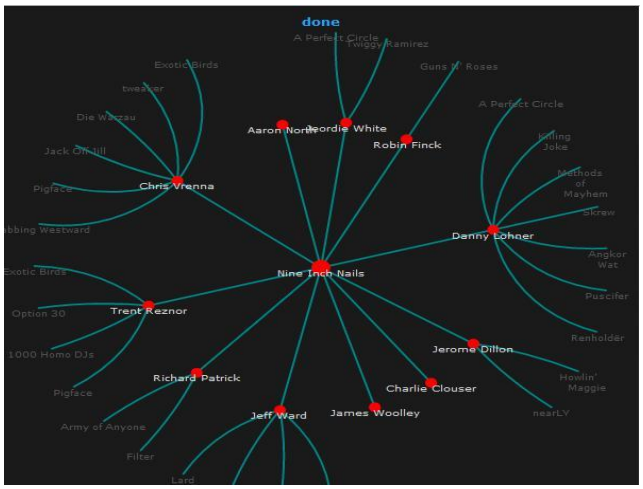


Figure 2.0

The illustration of hyperbolic tree diagram with ‘root’ and ‘nodes’.

Figure 3.0



Despite the great flexibility in arranging the ‘nodes’ using hyperbolic technique, there are several issues concerning the techniques. Among the issues are regarding the arrangement of the ‘nodes’. Some ‘nodes’ that are very close or directly link to root might get positioned very far from the root due to spacing arrangement of the ‘nodes’. Other than the tree diagram would be not be nice for the view, it can increase the risk of user to missed the ‘nodes’ all together due to the farther position from the ‘root’. The feature of focusing and zooming is helpful in allowing user to better see the relation between ‘nodes’ up close, however it as well limits user’s view of the entire hyperbolic tree diagram.

5. Radial tree diagram

Radial tree diagram is similar in structure with hyperbolic tree diagram, however radial tree diagram is emphasizing on the importance of ability to view the whole hierarchy even after focusing on one single node.

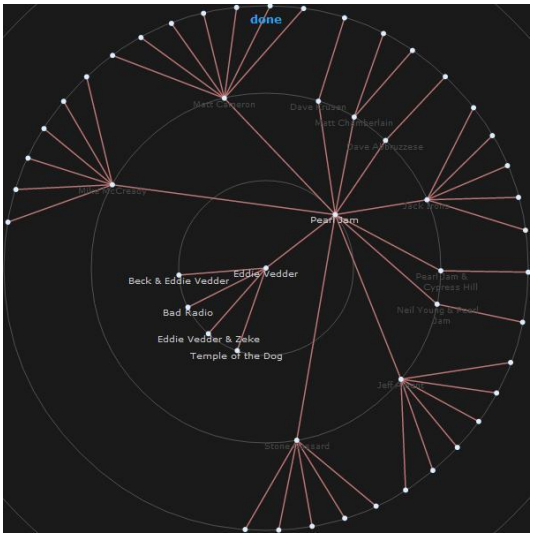


Figure 4.0

Above illustration for radial tree shown that the node being focused on is being bring to the centre and highlighted, however the overall hierarchical relation is still visible as radial tree diagram restricted the span of the nodes of same hierarchy.

Radial tree diagram are constructed with the aim to fit everything on a single screen view. The nodes are being positioned within the given sectors which are the left after the "first nodes", nodes that are directly attached to "root" are forms. This construction style makes sure that the nodes will not overlap with one another.

As per its purpose, radial tree have the limitation in term of spacing and focusing features. The focusing can be done, but with limited enlargement as being illustrated in the figure above. Radial tree diagram has been suggested to be used with user's already setting their focus on the subject they are working on as the nodes growth could be out of control hence resulting in poor quality of information visualization.

6. Spacetree diagram

Spacetree diagram is a simpler version of both hyperbolic tree diagram and radial tree diagram. Spacetree is the conventional nodes linked together forming a tree diagram. It is the first form of tree diagram before later on being developed as either hyperbolic or radial tree diagram to accommodate the limitation of viewing.

Spacetree can clearly show the relation between the nodes and the parent nodes as it is very directed in construction. The branches are predefined and are not being enabled for expansion. This way the navigation is made more clear and guided. User can get the general idea of the information supplied without having problems of having too many branches of same level.

The construction of spacetree starts with the subject/title also known as the parent node at either end of the space (top, bottom, left, right) and the nodes are being expanded in the provided spaces. Spacetree browser allows the nodes that user didn't wish to view to be

hidden and concentrates on the nodes required by the users.

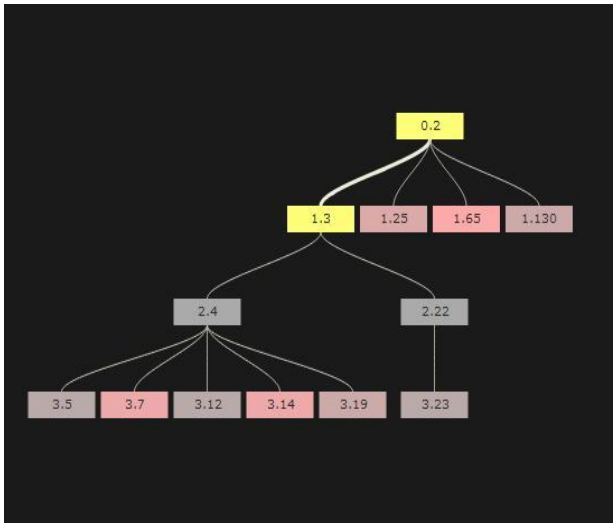


Figure 5.0

The figure above illustrates the hierarchical organization of the spacetree. The last nodes can be expanded provided the user chose to view it as such.

6.2 Hierarchical Visualization – Spacetree Diagram

This project had decided to pursue spacetree diagram as the visualization technique of choice. Spacetree diagram is the first form of hierarchical tree diagram representation. Spacetree diagram will be able to clearly shown the hierarchical relationship of the visualization intended for this project.

Spacetree diagram is easily expandable without having to make changes to the hierarchical representation of the visualization. The expandable ability is important as the end product is designed to accommodate to large amount of document. The relationship between the documents shall be made possible only if the tree diagram is able to expand according to needs.

Figure 7.0

Although it has not been proven in previous studies, the success of manipulating spacetree diagram as information visualization might determine the relevancy of using tree diagram as the visualization technique in visualizing biomedical documents information. The other visualization techniques, which are the revolution of spacetree diagram, will also be possible as the best visualization representation for biomedical documents when spacetree diagram is a success.

6.3 Application algorithm

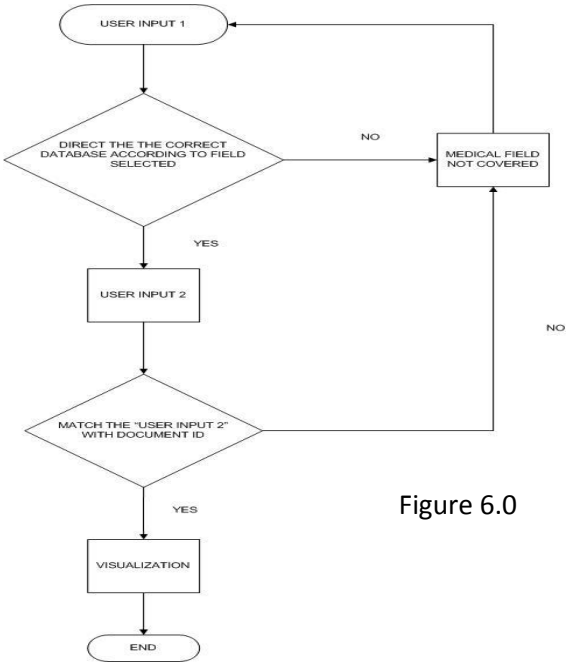


Figure 6.0

User will trigger the application by providing two inputs through the dropbox menu in the interfaces. USER INPUT 1 will be used to determine the selected medical field (i.e: Alzheimer's, Dementia) in which the user in interested in. If the said medical field in not covered in the database, user will be prompted to re-select the available medical field provided.

USER INPUT 2 will determine the area of the medical field that the user is interested to learned more about .(i.e: Symptoms, Treatment). If the area of the medical field is not covered in the database, the system will direct the user back to the search options. Once the area of medical field is selected, USER INPUT 1 & 2 will do the matching processes with the tags which also served as the document's ID.

Once the matching is done, the result will be represented as tree diagram visualization. User will be able to see the information resides in the documents available through the connecting nodes that show the relationship of each documents to the area of medical field intended by the user.

6.4 User Advantage

User testing session had been carried out to further test the objectives which concerning user's benefits from using the web application. The testing was done by demonstrating the product to user and allowing them to work the application hands on. The results of the testing are shown as below. The questionnaires scripts are provided in the appendices as proofs. A total of 10 respondents had participated in the product testing and questionnaires.

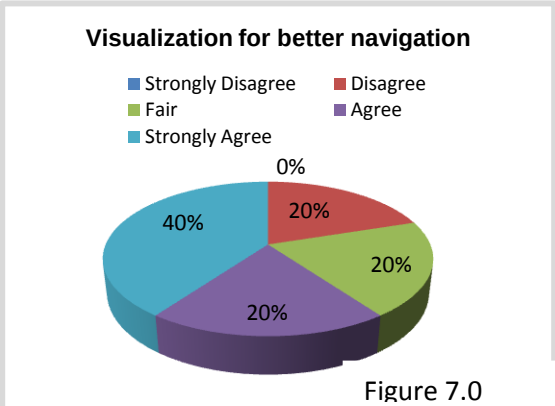


Figure 7.0

User testing result had showed that 40% of the respondents strongly agree that the visualization helps in providing better navigation through the search results. Another 20% agrees that visualization did help in navigation while another 20% are neutral about having visualization in providing better navigation in going through search results. The rest 20%

disagree that visualization helps in navigation. None of the respondent strongly disagrees.

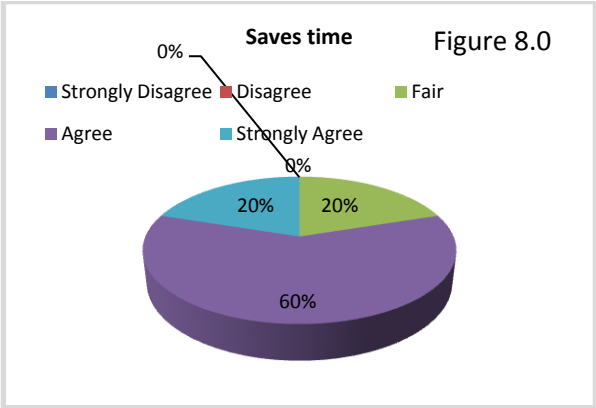
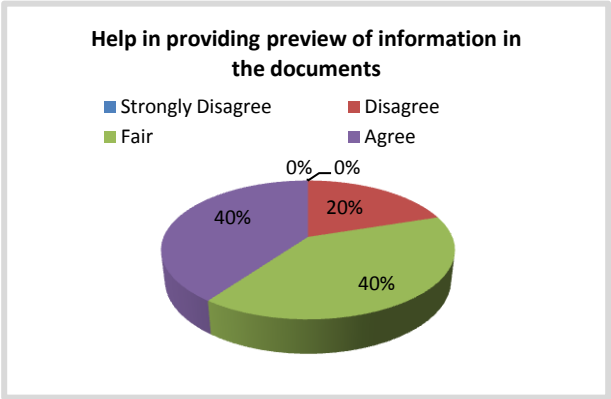


Figure 8.0

The total of 60% of the respondents strongly agrees that the visualization representation approach help in saving their time to get to the correct articles that they are looking for. Another 20% agrees and the rest 20% are neutral regarding the visualization approach in helping them saved time in looking for the articles. None of the respondents disagree that visualization of the search results had helped in saving their time. The respondents find it helpful that the search process is specified, it helps them to browse through the possible suggestions of articles muchfaster.



A total of 40% of all respondents agrees that the visualization helps in providing previews of the information contains in the documents retrieved. Another 40% found it fair that the web application did provide information and the rest 20% disagrees. The main reason for the lack of positive responses lies on the limitation of advanced medical knowledge by the author and the limitations of resources for biomedical texts.

Figure 9.0

7.0 Conclusion and Recommendations

The objective of this project is to provide a guided visualization for users in navigating through the biomedical documents suggested by the retrieval system.

Hence the information visualization techniques has been thoroughly studied and several most relevant visualization techniques that applies to this projects has been reported in the as result and discussion for this progress report. One of the visualization has been chosen as the visualization to be developed in completing this project and testing the value of information visualization for biomedical documents search results.

The visualization technique chosen is spacetree diagram, the first form of hierarchical tree diagram. Spacetree diagram built is simple but compact, it supports the expansion of branches and able to portray clear hierarchy structure. The project has been carried out to develop a spacetree diagram visualization that manages to provide general information for the user without them needing to read the articles.

The algorithm of the project has undergone some changes but the general idea and flow of the application is the same. The end product has been developed using PHP, Cascading Styling Sheets and mySQL as the database. The software used might not be the most difficult and advanced but it was suitable in helping this project to achieve its objectives.

Due to time and work-power constraint, only one visualization technique is able to be tested in visualizing biomedical documents search results. As being reported in the result and discussion part of this project, there are several more visualizations that have the potential to be developed as the next visualization representation. This project recommendation for future works would be for other visualization technique to be developed and presented as another useful information visualization tools, not only limited for biomedical documents but for any other industry as well.

For future development of this project, the author would like to make few recommendations on the features to be improves. The visualization can be further developed by completing the other form of visualization techniques. The improvements can be detected through the comparisons of all the visualization techniques which had been reported in the discussion part of this report.

For this project the visualization only able to provided suggestions for articles related to the specified search, for future developments more typed of media such as videos, audios and photos are for consideration to be included in the visualization. This is to accommodate to more wide users, as reading is not everyone forte, other types of medias might me for helpful for them in browsing through the search results.

The project needs expertise in biomedical field and multimedia field to be able to progress in a more promising way. The information previewed in the visualisation is relatively simple and general due to the author limitations in biomedical fields. The help of expertise in cracking the code of biomedical text would be very helpful in providing a more detailed information preview for the biomedical text suggested in the visualization. The project is in need of more resources for biomedical text in order to be able to provide more information for the users. The involvement of experts and larger biomedical library will results in a large cost incurred. Thus this development is only possible with large funding.

The other alternative is to convert this web application into some sort of reading trackers for medical practitioners. This way, the issues of limited resources will be eliminated as the medical practitioners are responsible for their own resources. However, more features need to be included for the web application to be fully controlled by the users, such as adding more friendly features and more clear instruction on the usage of the web application.

8.0 References

[1] H. Zhang, M. Fiszman, D. Shin, C. M. Miller, G. Rosemblat and T. C. Rindflesch, "Degree centrality for semantic abstraction summarization of therapeutic studies," *J. Biomed. Inform.,* 2011.

[2] J. D. Duke, X. Li and S. J. Grannis, "Data visualization speeds review of potential adverse drug events in patients on multiple medications," *J. Biomed. Inform.,* vol. 43, pp. 326-331, 2010.

[3] B. Herr, E. M. Talley, G. A. P. C. Burns, D. Newman and G. LaRowe, "The NIH visual browser: An interactive visualization of biomedical research," in *Information Visualisation, 2009 13th International Conference,* 2009, pp. 505-509.

[4] T. C. Rindflesch, H. Kilicoglu, M. Fiszman, G. Rosemblat and D. Shin, "Semantic MEDLINE: An advanced information management application for biomedicine," *Information Services and use,* vol. 31, pp. 15-21, 2011.

[5] R. P. Futrelle, J. Satterley and T. McCormack, "NLP-NG — A new NLP system for biomedical text analysis," in *Bioinformatics and Biomedicine Workshop, 2009. BIBMW 2009. IEEE International Conference on,* 2009, pp. 296-301.

[6] M. Abulaish and L. Dey, "A concept-driven biomedical knowledge extraction and visualization framework for conceptualization of text corpora," *J. Biomed. Inform.,* vol. 43, pp. 1020-1035, 2010.

[7] Indiana University, "Radial Tree," vol. 2003, 2004, 2004.

[8] Jepson Herbaria of the University of California, "Deep Green Hyperbolic Tree," vol. 2010, .

[9] Lamping John, Rao Ramana, Pirolli Peter, "Hyperbolics Trees," vol. 2011, March 15, 2011.

[10] Lamping John, Rao Ramana and Pirolli Peter, "A Focus Context Technique Based on Hyperbolic Geometry for Visualising Large Hierarchies," vol. 1995, 1995.

[11] lewellen, "Tree Drawing : The Radial Tree," 2008-08-31, 2008.

[12] N. Sheth and Q. Cai, "Visualizing mesh dataset using radial tree layout," 2003.

[13] P. Catherine, G. Jesse and B. B. Benjamin, "SpaceTree: Supporting Exploration in Large Node Link Tree, Design Evolution and Empirical Evaluation

This page is purposely left blank.