Using Tag Clouds as a Tool for Patients' Medical History Visualization and Record Retrieval

Daphne Foldes

A Thesis
in
The Department
of
Computer Science and Software Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Computer Science at
Concordia University
Montreal, Quebec, Canada

July 2015

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: DAPHNE FOLDES

Entitled: Using Tag Clouds as a Tool for Patients' Medical History Visualization and Record Retrieval

and submitted in partial fulfillment of the requirements for the degree of

Master of Computer Science (MCompSc)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Joey Paquet Chair

Dr. Thomas G. Fevens Examiner

Dr. Emad Shihab Examiner

Dr. Olga Ormandjieva, Dr. Kristina Pitula Supervisor

Approved by ________________________________
Chair of Department or Graduate Program Director

____
Dean of Faculty

Date ________________________________

ABSTRACT

Using Tag Clouds as a Tool for Patients' Medical History Visualization and Record Retrieval

Daphne Foldes

Reading through a patient's medical history can be challenging at best, let alone when the patient has a lengthy medical record. This can be a challenge when the patient is a relatively healthy person and even more challenging for a patient with a medical condition. Having a tool to quickly view and retrieve the pertinent elements of a person's medical record can be useful, especially in cases where a healthcare practitioner is treating a new patient. We propose a visualization tool that makes use of a tag cloud that would allow a healthcare practitioner to easily visualize and retrieve the essential elements of a patient's medical record. A prototype was created to run usability testing of the tag cloud tool and collect feedback from healthcare practitioners on the usefulness of such a tool in their day-to-day interactions with patients. Twelve paramedical practitioners tested and were questioned on the tool. The findings of this usability testing showed that such a visualization tool would be helpful to paramedical practitioners seeing a patient for the first time, as well as when dealing with patients who have a lengthy medical history.

DEDICATION AND ACKNOWLEDGEMENTS

Dedicated to Dr. T. Radhakrishnan who graciously agreed to take me on as a Master's student. His knowledge, experience, compassion, and insight allowed me to undertake this journey. Thank you Dr. Krishnan.

I would like to also thank my co-supervisors, Dr. Olga Ormandjieva and Dr. Kristina Pitula, for agreeing to take over for Dr. Radhakrishnan after his passing. Their guidance has allowed me to complete this thesis and fulfill my promise to complete my Master's. Thank you both for your time, knowledge, and guidance in completing this journey!

Table of Contents

Table of Figures

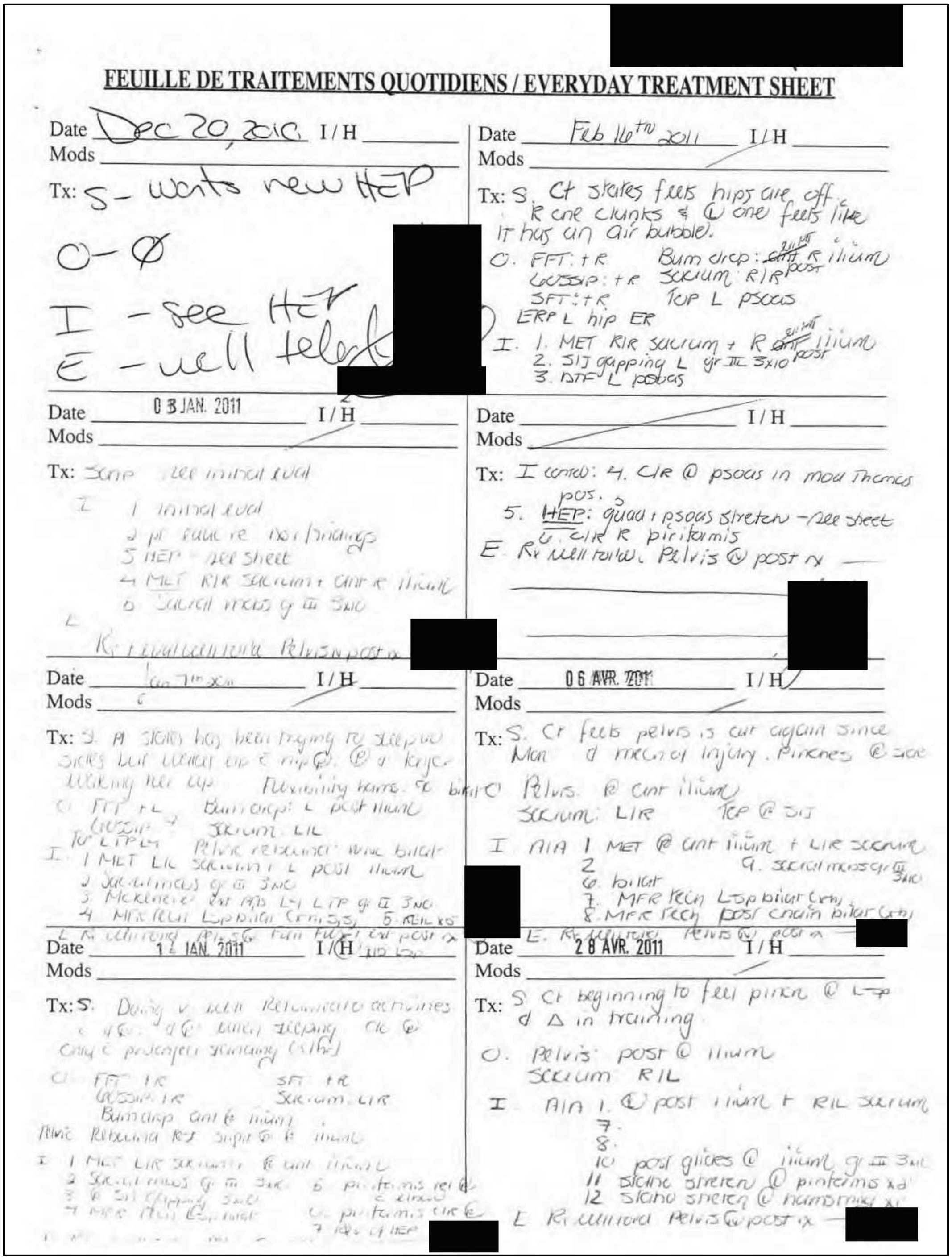
Table of Tables

1. INTRODUCTION

Electronic health records (EHRs) have great potential to improve safety, productivity, and efficiency in the medical field. With the arrival of EHRs in recent years, medical practitioners have been slowly transitioning from handwritten medical notes to entering electronic notes into these EHRs. This allows not only for having one central location for a patient's data, but also for sharing of patient history with other medical practitioners that are treating the same patient. Searching through a patient's medical history is also much easier when it is in electronic format as opposed to the paper format used for decades if not centuries. Being able to effectively visualize and search through this electronic patient history will become more important as more healthcare practitioners transition to EHRs.

1.1 PROBLEM STATEMENT

Currently paramedical practitioners, hereafter referred to as therapists, use handwritten notes on a sheet of paper that contains eight small square spaces, one square per visit, as the method for entering information in a patient's file. Multiple therapists work with the same file, so there is a variety of handwriting styles and shorthand used within the same patient file (see Figure 1.1). This can make it difficult to scan the patient's file for pertinent past medical history at each new visit, effectively taking away time from the visit that could be used to treat the patient. In the case of a Physiotherapist or an Athletic Therapist, a visit is typically thirty minutes. Osteopaths and Chiropractors provide hour-long treatments, but they could make better use of this time by treating their patient rather than searching for information among handwritten pages of treatment notes. This is even more important when the therapist is seeing a patient for the first time and has no prior knowledge of the patient's history or the content of the patient's file.



[redacted]

FEUILLE DE TRAITEMENTS QUOTIDIENS / EVERYDAY TREATMENT SHEET

Date Dec 20, 2010 I/H
Mods
Tx: S – wants new HEP
O – Ø
I – see HEP
E – well tol[redacted]

Date Feb 16th 2011 I/H
Mods
Tx: S. Ct states feels hips are off. R one clunks & L one feels like it has an air bubble.
O. FFT: + R Bum drop: sup/ant R ilium post
Gossip: + R Sacrum R/R
SFT: + R TeP L psoas
ERP L hip ER
I. 1. MET R/R sacrum + R sup/ant ilium post
2. SIJ gapping L gr III 3x10
3. DTF L psoas

Date 03 JAN. 2011 I/H
Mods
Tx: [illegible] see initial eval
I 1 initial eval
2 [illegible]
3 HEP – see sheet
4 MET R/R sacrum + ant R ilium
5 sacral mobs gr III 3x10
E
R [illegible] [redacted]

Date I/H
Mods
Tx: I contd: 4. C/R L psoas in mod Thomas pos.
5. HEP: quad + psoas stretch – see sheet
6. C/R R piriformis
E. Rx well tol. Pelvis L post rx
[redacted]

Date [illegible] I/H
Mods
Tx: S. Pt states has been trying to sleep on sides but [illegible] [illegible]
O FFT + L Bum drop L post ilium
Gossip + Sacrum L/L
I. 1 MET L/L sacrum + L post ilium
2 sacral mobs gr III 3x10
3 [illegible]
4 MFR [illegible] 5 [illegible]
E R [illegible]

Date 06 AVR. 2011 I/H
Mods
Tx: S. Ct feels pelvis is out again since Mon d mech of injury. Pinches L side
O Pelvis L ant ilium
Sacrum L/R TeP L SIJ
I AIA 1 MET L ant ilium + L/R sacrum
2 9. sacral mobs gr III 3x10
6. bilat
7. MFR tech L-sp bilat
8. MFR tech post chain bilat
E. R [illegible] [redacted]

Date 14 JAN. 2011 I/H
Mods
Tx: S. Doing v. well [illegible]
O FFT + R SFT + R
Gossip + R Sacrum L/R
Bum drop ant R ilium
I 1 MET L/R sacrum R ant ilium
2 sacral mobs gr III 3x10 5 piriformis rel
3 R SIJ gapping 3x10
4 MFR [illegible] 6 piriformis C/R
7 [illegible] HEP [redacted]

Date 28 AVR. 2011 I/H
Mods
Tx: S Ct beginning to feel pinch L L-sp d Δ in training
O. Pelvis: post L ilium
Sacrum R/L
I AIA 1. L post ilium + R/L sacrum
7.
8.
10 post glides L ilium gr III 3x10
11 static stretch L piriformis x2
12 static stretch L hamstring x2
E R [illegible] pelvis L post rx [redacted]

Figure 1.1 Example of a typical page found in a patient's file.

1.2 MOTIVATION

The medical community is under increasing pressure to use computer-based systems to support the clinical side of their practices. Between 2007 and 2010, the use of Electronic Health Records (EHRs) in Canada increased by 6% with 39% of physicians using EHRs to both enter and retrieve patient data in 2010 [1]. During the same period in the United States, use of EHRs increased by 16% with just over 50% of US physicians using an EHR in 2010 [2]. Outside of North America, use of EHRs is even more prevalent with data from 2006 showing that EHRs were already used by 79% of physicians in Australia, 89% in the United Kingdom, 92% in New Zealand, and 98% of physicians in the Netherlands are using EHRs for patient records [3].

This is an ideal time to transition from handwritten patient files to electronic files given the progressive adoption of EHRs in the last few years by the international medical community. Adding to this is society's increased dependence on technology in all areas of everyday life, not only in the medical community [4]. This increased use of technology makes it crucial to implement tools within EHRs that make data retrieval more efficient to ensure that patient care is not compromised.

The progressive implementation of EHRs across the medical community will also aid in the implementation in other healthcare fields. Not only will we find EHRs in hospitals and physicians' offices, but we will also see them adopted by therapists in the paramedical community, such as Osteopaths, Chiropractors, Physiotherapists, Athletic Therapists, and other specialists that require that they keep patient records.

1.3 RESEARCH QUESTIONS AND HYPOTHESES

This motivation leads us to two research questions:

1. How to visualize patient history more efficiently?
2. How to easily retrieve relevant patient information?

From these research questions, we hypothesize that

Hypothesis H1. The Tag Cloud allows more efficient visualization of patient history as compared to existing visualization methods.

Hypothesis H2. The Tag Cloud allows for easier retrieval of more pertinent patient information as compared to existing search methods.

Controlled experiments were designed and run to test the validity of hypotheses H1& H2 and answer the research questions on the effect of a tag cloud on the visualization and retrieval of patient health information in EHRs.

1.4 CONTRIBUTIONS

We propose a novel way of searching for patient information within an EHR. Using a tag cloud, as opposed to other possible search methods such as a search box or scrolling through the content of the EHR, will allow therapists to quickly locate information in the patient's file. The tag cloud used as a search tool would allow the therapist to have more time for treating the patient during the visit. We hypothesize that the tag cloud allows for more effective visualization of patient history and for easier retrieval of pertinent patient information as compared to existing search methods.

As a proof of concept, we built a prototype of tag cloud information retrieval for a patient file that would be stored in an EHR system. The prototype was used in an empirical study aimed to explore the opinions of paramedical practitioners about online access to EHRs and patient health information. We compared the usability of different search methods – tag cloud, search box, and scroll bar, to show that the tag cloud is the preferred tool for searching a patient's data. The quantitative and qualitative results gathered from usability testing, in comparing the tag cloud to the other search methods, showed a significant preference for using the tag cloud when having to navigate a patient's file for the purpose of finding pertinent information.

1.5 BASIC ASSUMPTIONS/LIMITATIONS

To focus our usability testing of the tag cloud as a visualization and retrieval tool, we used the following assumptions:

- A database exists as part of the EHR where the tags associated with each patient visit can be stored and retrieved to create the tag cloud.

- Tags are added to each patient visit entry at the same time that the notes for the visit are entered. The mechanism for entering the tags is not under consideration for the purpose of this thesis.

The following are the limitations that were taken into account:

- Known one-time only occurrences that are not beneficial to being tagged. Two examples of this would be “birth” and “death”, as each of these would only happen once in a patient’s history.
- Tags cannot show you the timeline of content, only points in time.

1.6 THESIS ORGANIZATION

This thesis is organized as follows: Chapter 2 surveys the related literature in Information Visualization, Tag Clouds and discusses the current use of Tag Clouds. Chapter 3 reviews EHRs and their current state, the methods that are currently used in EHRs, how Tag Clouds can be used as a Visualization Tool in EHRs, as well as describes the prototype created and used as a proof of concept for this thesis. The results obtained from the empirical study with therapists are presented in Chapter 4. Chapter 5 concludes with a summary of this thesis and outlines possible future work directions.

2. LITERATURE REVIEW

The body of literature pertaining to the emerging field of information visualization for EHRs is not overly large. This chapter provides an introduction to the field of information visualization and contains a critical review of the related literature concerning the use of information visualization methods in dealing with complex data. We first outline the origin of information visualization, and then discuss its application to medical systems, more specifically to EHRs. Following that, we examine the recent information visualization phenomenon known as tag clouds (a combination of information visualization and web design elements), discuss their current use and perceived advantages and drawbacks. To conclude, we present how tag clouds can be applied to improve the use of EHRs and current information visualization tools used within EHRs.

2.1 INFORMATION VISUALIZATION

2.1.1 What is Information Visualization?

On its own, visualization is something that "… can provide a qualitative overview of large and complex data sets, can summarize data, and can assist in identifying regions of interest and appropriate parameters for more focused quantitative analysis [5]." While information visualization is the graphical representation of data in such a way that information can easily be extracted by the user viewing this representation, its formal definition is "… the visual representation of large-scale collections of non-numerical information, such as files and lines of code in software systems, library and bibliographic databases, networks of relations on the internet, and so forth [6]."

2.1.2 Advantages of Information Visualization

The greatest advantage of Information visualization is the ability to identify and visually distill the most valuable and relevant information content of large data. The resulting visualization can therefore be very quickly interpreted when presented well [7].

The key benefits that information visualization provides to its users are [7]:

- It provides an ability to browse, and comprehend, huge amounts of data;
- Allows the perception of emergent properties that were not anticipated;

- Enables problems with the data itself to become immediately apparent;
- Facilitates understanding of both large and small-scale features of the data; and
- Facilitates hypothesis formation.

2.1.3 Some Information Visualization Methods

Simple x-y plots, timelines and LifeLines, Gantt and PERT charts are well-known information visualization techniques. Others such as time annotation glyphs or paint strips [8] are less known and used. For the most part, the current visualization methods mentioned above are useful for displaying medical data.

2.2 INFORMATION VISUALIZATION AND ELECTRONIC HEALTH RECORDS

Searching the various academic databases, we can find research on visualization and EHRs, but very little on visualization in EHRs.

Chittaro [9] touched briefly on the importance of visualization in EHRs, while Kosara and Miksch [20] discuss different visualization methods, which can be used for various types of data. Many authors discuss the use of timelines to visualize patient data over time [10][11][12]. A timeline is a visual representation of chronological events on a line (see Figure 2.1). Plaisant et al. [12] expanded on the timeline and brought forth the idea of "LifeLines" to concisely display a patient's health issues over time in a one-page summary. All of these bring important details to the area of EHR visualization and we will discuss each of these articles in more detail in the following sub-sections.

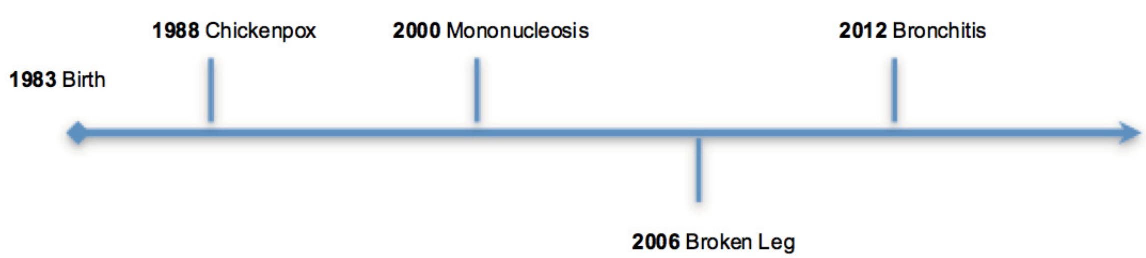


Figure 2.1 Simple example of a timeline.

2.2.1 Chittaro's "Information visualization and its application to medicine"

Although brief, Chittaro brings forth key points of good medical information visualization systems, which are [9]:

- "Visually present medical data in a more intuitive, easy to understand, easy to learn, easy to recognize, easy to navigate, easy to manage formats.
- Visually magnify subtle aspects of the diagnostics, therapeutic, patient management, and healing process, which otherwise could be difficult to notice.
- Prevent information overload and allow members of the clinical staff to master larger quantities of information".

Chittaro begins by reminding the reader that as humans we can only remember so much information at once, the medical community being no exception to the rule. As EHRs become more commonplace, processing and analyzing these large quantities of patient data will become increasingly difficult if EHRs do not take into account the human limitations and easily display this data to view, interact with and analyze. Chittaro's view of information visualization is that of a tool to achieve one of many goals in the implementation and use of an EHR [9]. The main goals are: 1) data that is easy to recognize and navigate, 2) visualizes subtle pieces of the patient's medical history, and 3) allows medical practitioners to work with large amounts of data without creating information overload [9].

2.2.2 Kosara and Miksch's "Visualization methods for data analysis and planning in medical applications"

Many phenomenon need to be captured in an EHR such as the healthcare practitioner's notes on the patient's visit, medical test results, scans and x-rays, as well as medication and other treatment options prescribed, all of which do not lend themselves to the same visualization methods. Kosara and Miksch [8] discuss different features and methods best suited to the different types of data within an EHR. There are three categories of data that are covered: measured data, incident and symptom data, and planning data.

"Measured data" is either continuous data, such as vital signs of a hospital patient, or data that is analyzed, like blood test results. "Incident and symptom data" pertains to the incidents and symptoms that a patient experiences over time. "Planning data" is used to plan the future treatment of a patient and is the most complex type of data to visualize, as it contains some elements of uncertainty [8]. The authors describe features needed for each of the described categories of data, as well as the visualization methods available to depict them. These data categories will be covered in greater detail in Chapter 3. Kosara and Miksch state that some data, such as cyclical data, still remains a challenge in terms of visualizing planning activities and that there is still a need for a method that would allow all of the information categories to be visible without creating information overload [8].

Many authors have worked on information visualization that revolve around the timeline concept. Plaisant et al. created LifeLines [12], Bui, Aberle, and Kangarloo envisioned a similar approach called TimeLine [10], while Hallet proposed the CLEF chronicle [11]. These are discussed in further detail below.

2.2.3 Plaisant's LifeLines

The main purpose of LifeLines is to visually display the key elements of a person's medical history on one screen, therefore enhancing the navigation and analysis of a patient's record. The LifeLines system is comprised of three concepts: facets, lines, and events. Facets refer to the different aspects of a patient's record, such as problems, allergies, diagnoses, medications, lab results, and so on. Lines are the visual timeline that is displayed starting on the first occurrence of an event, such as headache, penicillin reaction, seizure, or elevated LDL, for example. Visually, LifeLines is a screen with sections (the facets) that contain timelines (the lines) with labels (the events). Lines can be colored and differ in thickness to indicate severity, status, or other relevant information that needs to stand out. See Figure 2.2 for an example of LifeLines.

To view a particular event, one simply has to click on it to pull up the relevant information. As mentioned previously, each facet can be collapsed or expanded to allow the user to view the relevant information or to de-clutter the screen, since over time there will be more information as the patient's history accumulates. This last item is also one of the pitfalls of LifeLines. A patient will naturally accrue more data over time, which when all displayed on one screen defeats the purpose of why LifeLines was created. Plaisant et al. deal with this issue with the collapsible facets, as well as giving the user the ability to zoom in or out of a particular slice of time. Users are also given the ability to do textual searches of the system [12].

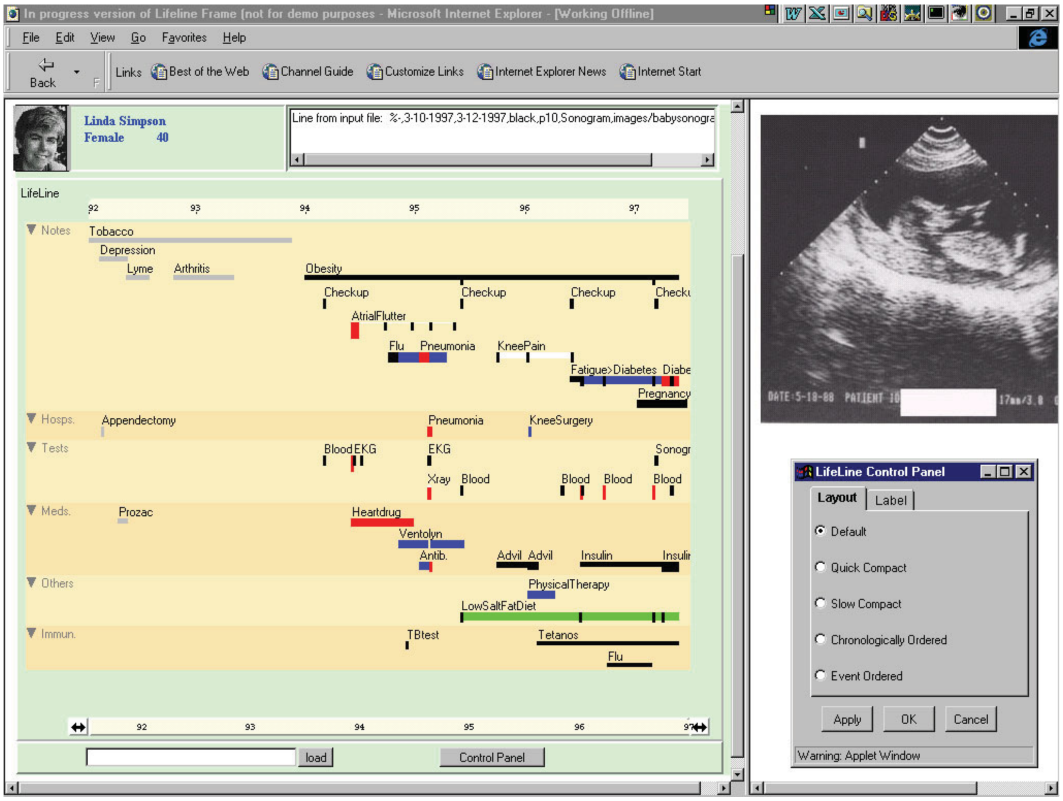


Figure 2.2 Example of LifeLines (Source: http://www.cs.umd.edu/hcil/lifelines/).

2.2.4 Bui, Aberle, and Kangarloo's TimeLine

The purpose of TimeLines is to bring data from various sources into one complete record for each patient. This data can either be viewed as a whole by the family physician or as a subset by a treating specialist. This flexibility of creating a customized user interface for one set of data depending on the information needed by the healthcare practitioner viewing the history is what sets apart the TimeLine system from Plaisant's LifeLines. The main elements of TimeLine are: patient details and encounter information, the medical problem list, graphical timelines, and a data viewer. These allow for the interface to display any kind of data, whether it be text, graphics, lab results, scans, or any other form of data that a healthcare practitioner might need to access in a patient's record [10].

2.2.5 Hallet's Clinical e-Science Framework (CLEF) Chronicle

Hallet's CLEF chronicle focuses on two requirements: providing a means to easily navigate a patient's complex record and to extract textual reports of the events of a given time span [11]. To satisfy these requirements, Hallet's interface resembles a pared down LifeLines in that there are only three sections of events (problems, treatment and investigations) and timelines for each event. Where Hallet differs is that her representation uses icons to indicate an occurrence of an event on a timeline (see Figure 2.3). CLEF chronicles also offer the possibility to zoom in and out to get a narrower or wider view of the patient's medical history. The user can then generate reports, either of a single event or multiple events over a timespan [11]. CLEF essentially allows users to extract specific sections of a patient's record, but it still constitutes a large amount of data to display to the user.

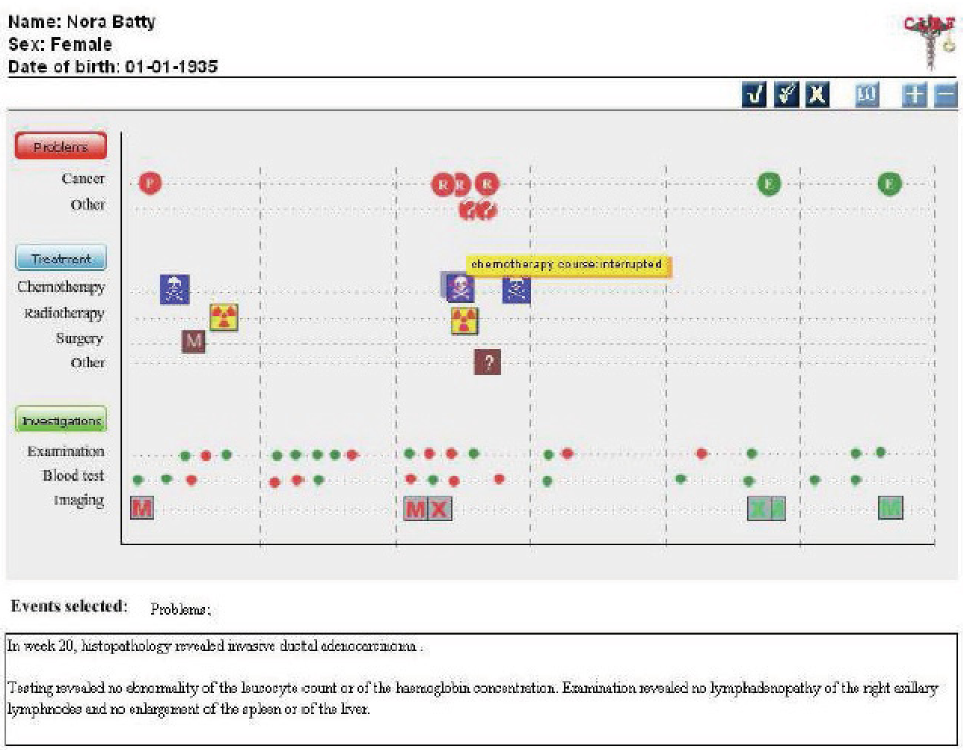


Figure 2.3 Example of Hallet's CLEF [11].

2.3 TAG CLOUDS

2.3.1 What is a Tag?

In layman's terms a tag is something used to label an item. In the technical world, more specifically within the area of computers and the web, a tag is defined as "… a nonhierarchical keyword or term assigned to a piece of information (such as an internet bookmark, digital image, or computer file). (This) helps describe an item and allows it to be found again by browsing or searching. Tags are generally chosen informally and personally by the item's creator or by its viewer, depending on the system [13].” Another way of looking at tags is that “Tagging refers to the addition of meaningful keywords to Web content (text, images, audio, video) for purposes of sharing, organization, and retrieval [14].”

2.3.2 What is a Tag Cloud?

A tag cloud is defined as “… a visual depiction of user-generated tags, or simply the word content of a site, typically used to describe the content of web sites. Tags are usually single words and are normally listed alphabetically, and the importance of a tag is shown with font size or color. (Making it) possible to find a tag alphabetically and by popularity. The tags are usually hyperlinks that lead to a collection of items that are associated with a tag [15].” In other words, a tag cloud is a visual representation of the most common tags found within a set of data and their relative importance in relation to all the other tags. Then, if one wants to see the content related to any given tag, they simply click on the tag that interests them and all associated entries with that tag are fetched for viewing. Visually, a tag cloud can look like Figure 2.4 or Figure 2.5 if one wants to use color.

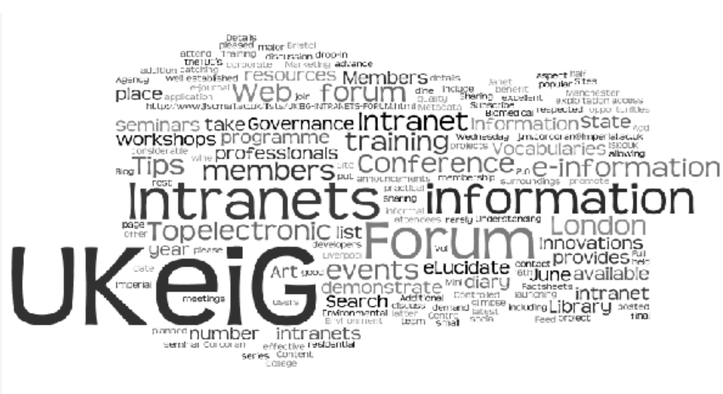


Figure 2.4 Example of a Tag Cloud (Source: http://www.rba.co.uk/tfttr/archives/2008/ukeigwordletagcloud.gif).

Figure 2.4 displays each tag in a different size and thus a different importance within the cloud. Although visually, we can spot the tags with greater importance, sometimes we may want to add color to the tags within the cloud, such as in Figure 2.5. In some cases the use of color has meaning, whereas in other instances, it is used purely to help differentiate one tag from another.

Figure 2.5 Example of a Color Tag Cloud (Source: http://www.useit.com/alertbox/wordle-word-cloud-donations.png).

There are several different ways of representing data using tag clouds. The tag cloud in Figure 2.4 is based on how often that tag was applied to a single item within the data set. The second and most common type is where each tag in the cloud represents how often a tag was applied to all items within the data set. Finally, tag clouds can also represent categories of data, where each tag is a category and the size of each tag represents the number of subcategories found within that tag [15]. To simplify things, the first type of tag cloud displays tags of an individual record within the whole data set, while the second type of tag cloud represents the tags used in all the records of a data set. A third type of tag cloud is when the tag is used to classify records within a data set to be of one group or another (see Figure 2.6).

Figure 2.6 Example of a Category Tag Cloud (Source: User created with http://tagcrowd.com/).

As previously mentioned, tag clouds can have different visual representations. In many cases, these visual representations have meaning and are not merely selected to beautify the tag cloud. Font size and thickness are used to show the tags that occur most frequently. Tags can either be listed alphabetically, as in Figures 2.7 and 2.8, to make finding specific tags easier, or they can be displayed with no particular order, as is seen in Figures 2.4 and 2.5. Color can also be used in tag clouds to either help separate each tag to make finding a specific tag easier, or an importance can be assigned to a specific color to make it stand out.

The order of the tags is also something that can be altered to make the retrieval of tags easier. The different ways that tags can be ordered are:

- Alphabetically, where the use of font size and weight are used to denote more frequently used tags;
- Alphabetically, where the font size remains the same, but color is used to highlight more important tags;
- In the order of frequency making use of color and font to indicate the importance of certain tags;
- By grouping those tags that have similar meaning close to each other and using font size, weight and color to indicate importance and frequency; and
- With no particular order whatsoever (as in Figures 2.4 and 2.5) and making use of font size, color, and weight to indicate importance and frequency [16].

Figures 2.7 and 2.8 contain the exact same tags with the exact same font sizing and tag placement. The difference between the two is that Figure 2.8 has the number of occurrences of each tag. In some cases this is preferable, as it gives the viewer the opportunity to give a quantitative importance to each tag, especially in cases where the difference in font size is minimal from one tag to another.

Figure 2.7 Typical Tag Cloud (Source: User created with http://tagcrowd.com/).

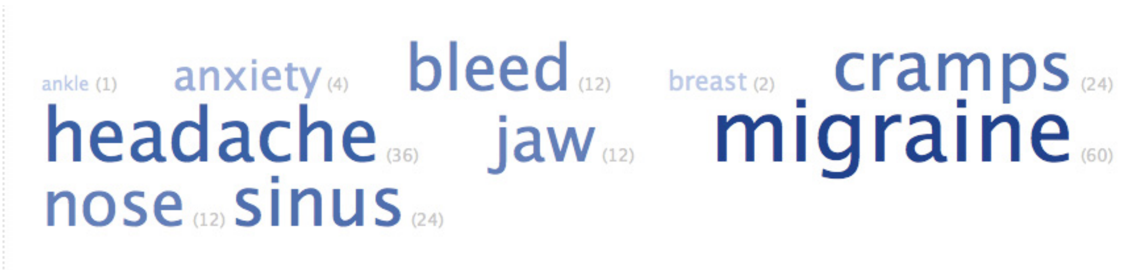


Figure 2.8 Tag Cloud with Weights (Source: User created with the aid of http://tagcrowd.com/).

2.3.3 How Tag Clouds Work

For each entry a person adds to the system, there is a field where they can type in keywords, which will become tags for that entry. Given that each tag in the cloud is linked to one or more entries that have been labeled with this particular tag, clicking on a particular tag brings up a list of entries that have that tag, similar to a search result list [17]. Depending on how the cloud is set up, what is seen and how it is displayed can vary. The simplest way of displaying the related entries is to list them chronologically, indicating the date they were entered into the system and displaying their title with a brief description of what they contain, as with a traditional search result.

2.4 TAG CLOUDS AS VISUAL INFORMATION RETRIEVAL INTERFACES

We will now review how tag clouds have been applied as a tool for information visualization and retrieval. Historically, tag clouds were used on websites and blogs to categorize and search through the content with Flickr being the first website on record to make extensive use of them [15]. It has also been noted that the first recorded sighting of a tag cloud in use may have been as early as the year 1992 [15]. Although tag clouds are still mainly used on blogs and websites, they have also been implemented in other contexts by researchers for the purpose of information visualization and retrieval.

Data repositories are one area where tag clouds have been implemented [18]. They have also been implemented for analyzing text within and across documents to quickly view what changes have been made since the previous publication [19]. An article in Computational Biology describes the use of tag clouds to make reading scholarly articles more dynamic and collaborative [20]. Another article mentions the use of tag clouds as a tool to help analysts search through large databases of documents to pinpoint the most relevant documents for their analysis [21]. In recent years, companies have applied tag clouds for advertising purposes, nick-naming the concept "tagvertising" [17]. A novel use of tag clouds has been to aid aviation mechanics search large operational documents from hand-held mobile devices to enable easy and quick retrieval of detailed information instead of carrying large manuals [22]. Another application of tag clouds was as part of a larger tool, SolarMap, which allows users to see how two, or more topics are possibly related to each other within a large set of documents [23].

One of the more in-depth uses of tag clouds has been in the library domain. Similar to the previous use of tag clouds for data repositories, a library is a repository of books, which can be daunting to search through. "Librarians have recognized the potential of Web 2.0 technologies and are moving quickly to adapt and use these tools for reference and user services [24]." This movement has been dubbed Library 2.0 (L2) in reference to the system being an offspring of the Web 2.0 movement. L2 is of special significance for the medical librarian, who is considered to be a leader in the technology field [14]. Given that "… secondary schools and colleges have already capitalized on Web 2.0 technologies and adapted teaching and learning techniques to appeal to digital natives, these adaptations will eventually trickle down to the training of scientists, healthcare practitioners, and allied health professionals, and ultimately affect medical library operations and facilities [14]."

Tag clouds have also been used in other novel ways. A company in Russia has made use of tag clouds in its search engine technology [25]. The most current application of tags and tag clouds was used in browsing content contained on a cellular phone. Researchers in Korea created a prototype where tag clouds were used to structure the content on cellular phones for ease of navigation and to "facilitate exploratory browsing [26]."

2.5 ACCEPTABILITY OF TAG CLOUDS BY USERS

Tag clouds are a useful tool because they allow users to quickly search for similar or related items in a data set. Since clicking on a tag within a tag cloud pulls up all related entries in the data set, a user can quickly scan the cloud for a particular word and select it to bring up the related entries.

There have been several research articles that have discussed and shown tag clouds to be a useful tool for narrowing down large sets of information. Sinclair and Cardew-Hall showed that tag clouds were preferred to search boxes when the user was not looking for something specific, but rather for general information on a topic. "Where relevant keywords are present, clicking on a tag is quicker and easier than entering query terms into a search box. It also provides a useful visual summary of the contents of the database [27]."

Knautz, Soubusta, and Stock [28] showed that their improved tool, tag clusters, which shows the links between the tags, was preferred over tag clouds for searching the web because of their ability to narrow search results by showing additional search possibilities after the initial web search. Tag clusters allow users to narrow their initial search results by means of both the vertices and edges of the cluster. Selecting the former adds this term to the initial search query, while the latter narrows the initial search using both terms at either end of the edge. The goal of their tag cluster, beyond showing that it is more useful than tag clouds, was that there is a need for users to be able to refine their search among large data sets in a simple manner, as opposed to using text searches making use of operators (AND, OR, etc.) to refine search results [28].

Hoeber and Liu's [29] research compared three methods to enhance personalized web search results: Tag Clouds, Term Histograms, and Term Lists. A term histogram is one where each bar in the histogram is for one term and shows the number of occurrences of that term. And a term list also is as the word suggests, a list of terms. Although their results ranked the tag cloud as the least preferred method chosen by participants, these results were for the specified context of personalized web search results [29].

Fung and Thandechteemapat [30] used tag clouds as one method to provide web search results as visual summaries of the contents of a website instead of the usual textual summary. The goal of the research was to show that in presenting information as a visual representation of the content of a website, users spent less time clicking on many web search results that would otherwise not be pertinent [30].

The articles mentioned in this section demonstrate that tag clouds are effective as a visualization tool in searching for information allowing users to easily retrieve pertinent information in Internet searches and data repositories. The research done was applied in the context of Internet search, not as part of an application, or as a means of reviewing previous content related to a particular element (for example, one patient file). Since these articles show that tag clouds can be useful in the context of Internet search and data repositories, there is a strong possibility of tag clouds also being useful in other contexts.

2.6 SUMMARY

Visualization is an important part of EHRs and their continued development, but the current tools used in EHRs could be enhanced to improve visualization. Of the current tools presented in this chapter, Plaisant's LifeLines is the tool that is most complete, however it has only been applied in the context of research. Although LifeLines is useful as a first-page overview of a patient's medical history, it is impractical when looking at a particular instance, as we need to go back to this first-page to see the summary.

Making use of the tag cloud can fill this gap and bring added functionality to the EHR without having to drastically alter the current EHR user interfaces. The tag cloud can provide the key elements of a good information visualization tool, such as helping to browse and comprehend huge amounts of data, facilitate understanding of both large and small-scale features of the data, and facilitate hypothesis formation [7].

Given the specialized search that healthcare practitioners must conduct in a patient's file, as well as the time limitations for performing such searches, the tag cloud solution proposed in this thesis will be shown to be preferred over other search methods.

3. APPLYING TAG CLOUDS TO ELECTRONIC HEALTH RECORDS

This chapter focuses on applying tag clouds to Electronic Health Records (EHRs). We will start by reviewing the current state of EHRs in Quebec and Canada, along with what visualization tools are currently found in these EHRs. We will then look at how tag clouds can be used as a visualization tool and how these can be used within an EHR for visualizing data. We will conclude this with a description of the prototype created and used for testing our hypotheses. The usability testing procedure and results will be covered in Chapter 4.

3.1 ELECTRONIC HEALTH RECORDS

Currently a very small amount of healthcare practitioners make use of EHRs not only in Quebec, but across Canada. The reasons for this are varied, concerns about privacy, implementation costs, and data security [31].

In 2010, the use of Electronic Health Records (EHRs) in Canada increased by 6% with 39% of physicians using EHRs to both enter and retrieve patient data [1]. During the same period in the United States, use of EHRs increased by 16% with just over 50% of US physicians using an EHR in 2010 [2]. Outside of North America, use of EHRs is even more prevalent with data from 2006 showing that EHRs were already used by 79% of physicians in Australia, 89% in the United Kingdom, 92% in New Zealand, and 98% of physicians in the Netherlands are using EHRs for patient records [3]. In Quebec the EHR project was initiated in 2006 and scheduled to be completed in 2010. It was then reassessed in 2011 where it was concluded that another five years would be needed to complete the project [4].

3.1.1 Categories of Data in EHRs

As was seen in Section 2.2.2, Kosara and Miksch [8] indicated that the different data categories that can be found in an EHR are:

- Incident and Symptom Data, which is textual
- Measured Data, which is continuous
- Planning Data, which is used to plan future treatment.

Examples of EHR data within the different data categories can be seen in Table 3.1 below.

Table 3.1 Sample EHR data and associated Data Categories

Data Category	EHR Data	Example(s)
Incident & Symptom Data	Anatomical / Clinical Terms	Anterior, Lumbar, L5, C3
	Symptoms	Seizure, Panic Attack, Acute Pain
Measured Data	Images / Scans	CT Scan, MRI Scan, X-Ray
	Test / Lab Results	Platelet Count, Cholesterol Levels
Planning Data	Treatment	Therapeutic Steps, Medication

3.1.2 Current Visualization Tools in EHRs

In [32], the authors review what data EHRs are likely to contain such as notes, diagnostic images, laboratory results, and biometric information and that currently EHRs visualize this data in the form of images, graphs, and text. As well, the authors mention that this data is displayed across multiple screens as opposed to on one screen where the clinician using the EHR could see the overview of the data.

West, Borland, and Hammond [33], reviewed the scholarly literature for innovative visualization techniques and came up with several tools that have been created to visualize data in EHRs, some of which were covered in Chapter 2, most notably Plaisant's LifeLines [12]. They concluded after this review that "few techniques have been found to effectively and efficiently display the large and complex data in EHRs [33]," leaving the door open for further research in this area of visualization in EHRs.

3.2 TAG CLOUDS AS AN INFORMATION VISUALIZATION TOOL

As noted in Section 2.2, little research has been conducted on visualization in EHRs and what visualization research has been conducted has revolved around the timeline concept. Based on the available research, no research has been done on tag clouds in the context of visualization within EHRs. One article mentions the use of tag clouds to display results of research pertaining to visualizing annotation schemes of transcribed verbal medical narratives, but did not delve into their use within an EHR, merely using them to show how they captured cognitive reasoning of physicians [34].

The use of tag clouds is a novel tool to add to the research field of information visualization in EHRs. Tag clouds are a visual representation to quickly view key elements of a patient's medical history that would otherwise not be seen without doing in-depth navigation of the data. This also avoids information overload at the same time that it provides ample information in a visual format that is easy to understand, learn, recognize, navigate, and manage. These are some of the key points of a good medical visualization system mentioned by Chittaro [9]; tag clouds clearly meet these criteria. Being simple to use, all a healthcare practitioner needs to do is add the tags related to the medical history entry they write after a patient visit and they will be added to the cloud. As well, since they can be placed on any screen, adding them as a visual cue and search aid can allow the healthcare practitioner to quickly pull-up any related entries in a patient's EHR.

The research to date on tag clouds relates to their use in search engines, commercial web sites, blogs, and on the different algorithms that can be used to generate tag clouds. The closest article related to tag clouds and the medical domain discusses their use as a tool to compare two versions of a medical document to find what items have changed since the previous version's release [19]. For these reasons, and because they address the design principles discussed in Kosara and Miksch's paper, we believe that the use of tag clouds in EHRs would be of benefit to health care practitioners when looking up their patients' history.

3.2.1 Why Tag Clouds would be a useful tool for EHRs

As EHRs become more prevalent, the amount of data they contain on any given patient will also grow. Chittaro [9] brought forth that since humans have natural limitations on the amounts of data that they can process at any given time, if we want healthcare practitioners to be able to take full advantage of medical data systems, then we will need to consider effective presentation and interaction with the data contained in these systems [9].

In Chapter 2 we reviewed the three key elements Chittaro indicated are needed for good information visualization in medical applications [9], namely:

1. Visually easy to understand, learn, and navigate,
2. Visually magnify subtle elements of the data, and

3. Prevent information overload.

These three key elements are applicable to EHRs because we want to not only store the data for each patient, but we want to be able to easily retrieve it for future use.

While Grams stated in his article on EHRs that an EHR "… must offer a one page summary screen that is graphical and highly interactive that allows the total content of medical knowledge about a patient to be displayed [35]." Plaisant's LifeLines satisfies this statement, but it would also be useful to have a more concise summary that can be referenced when working in other sub-sections of the EHR, so that the user does not need to go back to the whole summary. Tag clouds lend themselves ideally to this task and do so in a novel way for EHRs. Since tag clouds can occupy less space on the screen, they can easily be added to each page, so that a healthcare practitioner can have access to a pared-down summary to quickly and efficiently look-up related information in the patient's record. The same tag cloud would be presented on each page and its scope would be global to the patient, thus showing all pertinent data in the patient's health record.

3.2.2 LifeLines versus Tag Clouds

If we compare Plaisant's LifeLines [12] and tag clouds, the most obvious similarity is that both are visualization tools for large amounts of data. The most obvious similarity is that both are visualization tools for large amounts of data. Both LifeLines and tag clouds use size to indicate the relative amount of times data occurs. In LifeLines it is the Lines and in the tag cloud it is the tags that grow bigger the more often the data they represent appears.

A LifeLine Event is similar to a tag in the tag cloud in that both are a label for a specific item, but the difference is that in LifeLines an Event is a given point in time, whereas a tag is an agglomeration of multiple items that can span over time. Tags, like Events in LifeLines, are a hyperlink to additional in-depth details of the data they are representing.

Unlike the Facets of LifeLines, a tag cloud is only one instance. To get similar functionality to the Facets of Plaisant's LifeLines, one would have to use multiple tag clouds, each representing some aspect of a patient's record, to be similar. As mentioned with respect to Plaisant's LifeLines, as a patient's file grows, the visual representation can clutter the screen and act as a distraction [12]. Since tag clouds are only one instance, they are less likely to clutter the screen because the tag cloud will use the allotted space on the screen where it was placed based on the boundaries set when implementing the tag cloud.

While Plaisant's LifeLines does provide additional information on many elements within a patient's file by means of Facets, a tag cloud is more efficient since all information that has been labeled (i.e. tagged) is in one visual representation. Users do not have to look at multiple visualizations of the data (i.e. Facets); there is only one visualization containing all major elements of the patient's medical history (i.e. the tag cloud). Having to focus on only one location for this information is more effective, thus requiring less time to scan for relevant information. It also means less noise on the screen vying for the healthcare practitioner's attention, as opposed to the Facets of Plaisant's LifeLines. A tag cloud is also more efficient than LifeLines because each tag is linked to multiple entries spanning over time, whereas an Event in LifeLines only encompasses a single point in time. Thus, selecting an Event in LifeLines will bring the user to only one entry, whereas a tag cloud can bring up as many entries as have been tagged.

3.2.3 Applying Tag Clouds to EHRs

To apply tag clouds in an EHR, we assume that there is a database as part of the EHR where the tags selected for each patient visit can be stored and retrieved to create the tag cloud. The tags would then be added to each patient visit entry at the same time that the notes for the visit are entered. As there are several ways to implement the mechanism to enter the tags, this implementation aspect would be a decision left up to the developer of the system based on the needs of the clinic requesting the EHR.

3.2.3.1 Advantages of Tag Clouds as Information Visualization Tool

While reading through a patient's EHR, the tag cloud allows the healthcare practitioner to quickly see if the current symptom the patient is talking about has occurred previously by scanning the cloud for this word. If this particular symptom has occurred often, we know from the description of the tag cloud, that it will be larger and possibly bolded. By clicking on the particular tag for that symptom, the healthcare practitioner can bring up the details of all other occurrences of this symptom. The tag cloud also allows the healthcare practitioner to quickly search for other related symptoms to see if they too have occurred previously and to see if they have some correlation with the current reason for the patient's visit.

3.2.3.2 Limitations of Tag Clouds as Information Visualization Tool

There are, however, some limitations that need to be taken into consideration when it comes to using tag clouds. The most noteworthy of these being:

- Known one-time only occurrences are not beneficial to being tagged. Two examples of this would be "birth" and "death", as each of these would only happen once in a patient's history.
- Tags cannot show you the timeline of content, only points in time.

3.3 TAG CLOUD TOOL PROTOTYPE

To test our hypotheses, we created a prototype with which to conduct usability testing. In Figure 3.1 below we can see the context diagram for the prototype showing what actions are possible with the prototype. The prototype created allows users to log in and view the content of one patient file and search for content in that patient file. Note only one patient file was implemented for the purpose of usability testing. The prototype also allows for entering new visits for this patient, but this functionality was not part of the usability testing, so will not be discussed further.

This section will cover the prototype requirements, prototype and tag data, the tag cloud selected for the prototype, and we will end with a review of the implemented prototype.

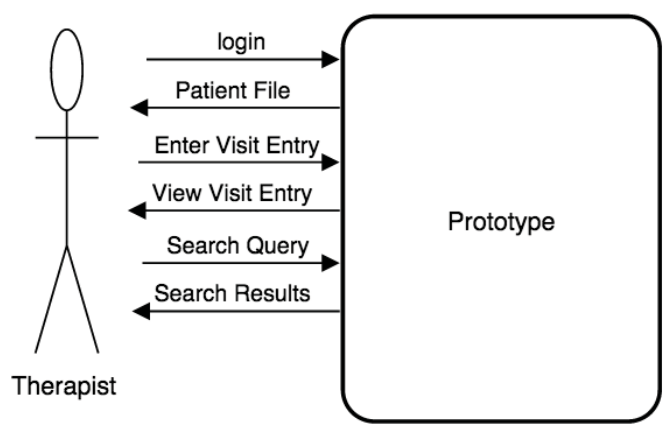


Figure 3.1 Context Diagram of prototype.

3.3.1 Prototype System Requirements

The following were the requirements that were selected as being necessary to build the prototype based on conversations with an Osteopath and Physiotherapist:

1. Separate Entries
    a. Each entry, which is a patient visit with a therapist, must be contained in a separate entry, so that it can be viewed without any noise from other entries.
2. Reverse Chronological ordering of entries
    a. Entries must be ordered in reverse chronological order, so that the most recent entry is at the top and the oldest at the bottom, as this is how paper entries are currently ordered.
3. Scrolling Capabilities
    a. The system must allow users to scroll through entries, as opposed to displaying them on several pages.
4. Search Capabilities
    a. The prototype must contain a search box that allows users to enter terms with which they can find entries containing those terms.
    b. The search functionality must also inform the user when the term cannot be found in the entries.
5. Tagging Capabilities
    a. The prototype must allow for tagging entries with tags, whether they are single words or a combination of words (i.e. a term).
6. Tag Cloud
    a. There must be a tag cloud to display the tags entered in the entries.
    b. Individual words, as well as a combination of words (i.e. a term), must be allowed as a tag.
    c. The tags in the tag cloud must be clickable and display the related entries when selected.
    d. Tags in the tag cloud must grow based on the number of occurrences of each tag in the patient file.

3.3.2 Prototype Data

The data used to create the prototype was the present author's own therapy record, which avoided any issues associated with access to data, as well as data security and privacy concerns. All identifying information, whether the author's or that of the different therapists, was changed to keep personal data confidential. The record contained a combination of entries from the various therapists the present author had seen over the last five years, which gave a record with sufficient data to be able to build a prototype that allowed for use of search capabilities, scrolling, and generating a tag cloud with a variety of tags.

3.3.2.1 Tag Data

Tags for each patient visit entry were selected based on terms highlighted by a Physiotherapist not participating in the usability testing. This ensured that the tag cloud was populated with tags that participants would be familiar with and expect to see when searching through a patient's file.

The data tagged in the prototype is incident and symptom data, as this is the data that is available in the patient record used to create the prototype. The prototype allows for entering terms, either single or multiple words, as tags, which are then used to generate the tag cloud. At the moment the tag cloud is restricted to showing incident and symptom data; future enhancements are discussed in section 3.3.3.8 of the current chapter.

3.3.3 Prototype

To quickly create a prototype that had all the requirements described in the previous section, a WordPress blog account [36] was used, as it contained all the basic requirements needed to replicate the elements of an EHR that were needed for usability testing: separate entries, reverse chronological ordering of entries, all entries on one main page with scrolling, tagging and tag cloud tool, as well as search functionality. Other functionality came as part of using WordPress, but it did not impede the usability testing of the prototype.

3.3.3.1 Tag Cloud

The tag cloud used in the prototype was the tag cloud widget that comes as a standard option in WordPress, as it contained the necessary functionality sought to test the tool.

The tag cloud widget from WordPress displays tags in alphabetical order and each tag grows based on the number of times it is present in the patient's file.

3.3.3.2 Prototype Screens

We will now look at the different screens of the prototype. Below are the screenshots from the prototype, along with a description of the elements within the prototype. It should be noted that screenshots of the prototype will look different when viewing search results for different search terms used in the various tasks.

3.3.3.3 Main Screen

Figure 3.2 below shows the screen that appears once the therapist logs into the EHR prototype. The different sections of this screen are labeled in Figure 3.2 below and described as such:

1. Patient's Name
2. Patient's File Number
3. Most recent Patient Visit Entry
    a. Date of Visit and Visit Type
    b. Date entry was entered into the system
    c. Therapist's Initials
    d. Therapy Type
    e. Date of Visit
    f. Therapist's Notes
    g. Quick Links
        i. Posted in – goes to the detailed entry (ies) for that category [*WordPress functionality – not part of the usability testing*]
        ii. Tagged – Displays tags entered for that visit and that populate the Tag Cloud. Selecting one of the tags from this section will bring up the search results screen for the tag selected, just like selecting the tag in the tag cloud.
        iii. Leave a Comment [*WordPress functionality – not part of the usability testing*]
4. Previous Patient Visit Entry (beginning of)
5. Search Box
6. Tag Cloud

    a. Previously selected tags in purple, as per web browser settings.

7. Categories [*WordPress functionality – not part of the usability testing*]
    a. Section that can be used as additional tags within the system that do not appear within the tag cloud. Basic entries were created to avoid an empty section on the screen:
        i. Therapist's Initials
        ii. Type of Entry
            1. Initial Evaluation
            2. Follow-Up Evaluation
            3. Follow-Up Call
            4. Progress Report
8. Calendar [*WordPress functionality – not part of the usability testing*]
    a. Any dates where an entry exists for a patient visit would be highlighted and clickable. Selecting a clickable date would bring up the same types of search results as can be seen in *Figures 3.3* and *3.4* below.

Figure 3.2 Main screen of prototype after log in.

3.3.3.4 Search Box Results

After entering a search term and selecting the Search button, the search results are displayed. Figure 3.3 below shows the screen with the search results for the search term *anterior* entered into the search box. The different sections of this screen are labeled in Figure 3.3 below and described as such:

1. Search Term used to pull up search results
2. Most recent visit entry synopsis that matches search term; clicking on title of search result will bring up the detailed entry (example in Figure 3.5)
3. Next visit entry that matches search term
4. Label that appears when synopsis of visit entry is longer than the space allocated on screen
    a. Clicking on this label will also bring up the detailed entry
5. Next visit entry that matches search term

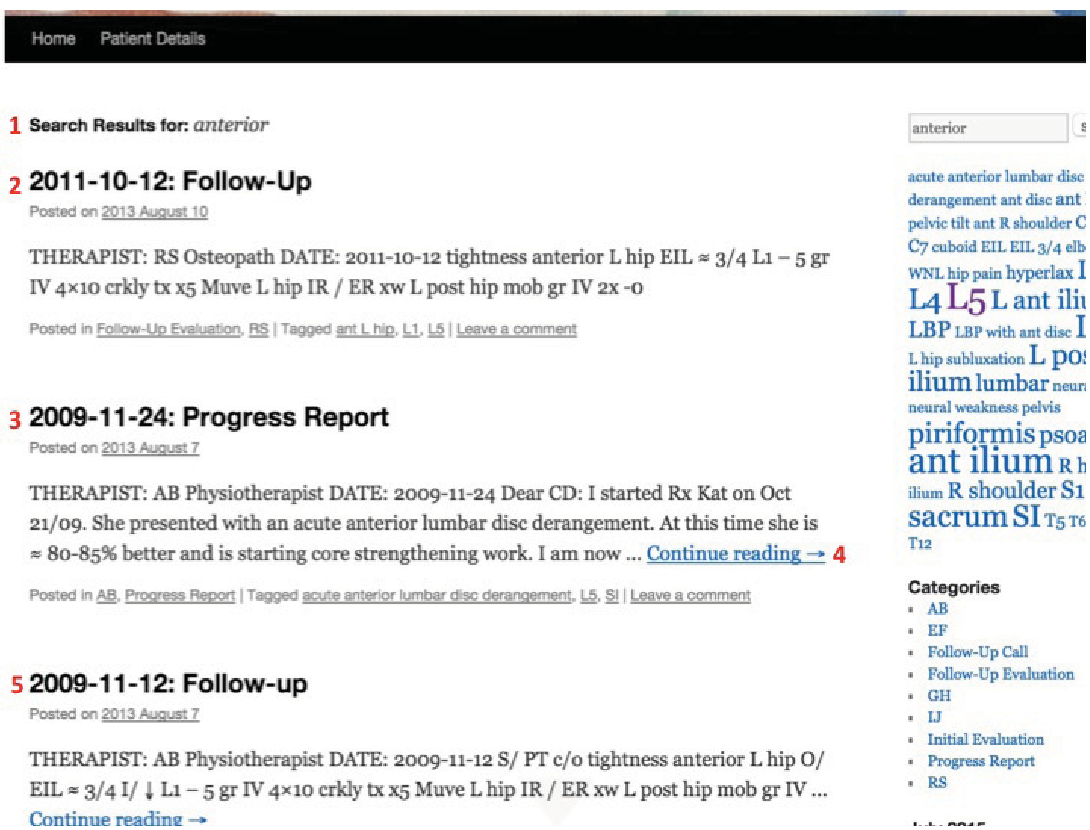


Figure 3.3 Search Box results screen for search term *anterior*.

3.3.3.5 Tag Cloud Tag Selection Results

Once a tag is selected in the tag cloud, the search results for entries that have that tag are displayed. Figure 3.4 below shows the screen with the search results for the selected tag *L5*. The different sections of this screen are labeled in Figure 3.4 below and described as such:

1. Tag selected in tag cloud to pull up search results
2. Most recent visit entry synopsis that matches selected tag
    a. Clicking on title of search result will bring up the detailed entry (example in Figure 3.5)
3. Label that appears when synopsis of visit entry is longer than the space allocated on screen
    a. Clicking on this label will also bring up the detailed entry
4. Next most visit entry that matches tag selected
5. Next most visit entry that matches tag selected

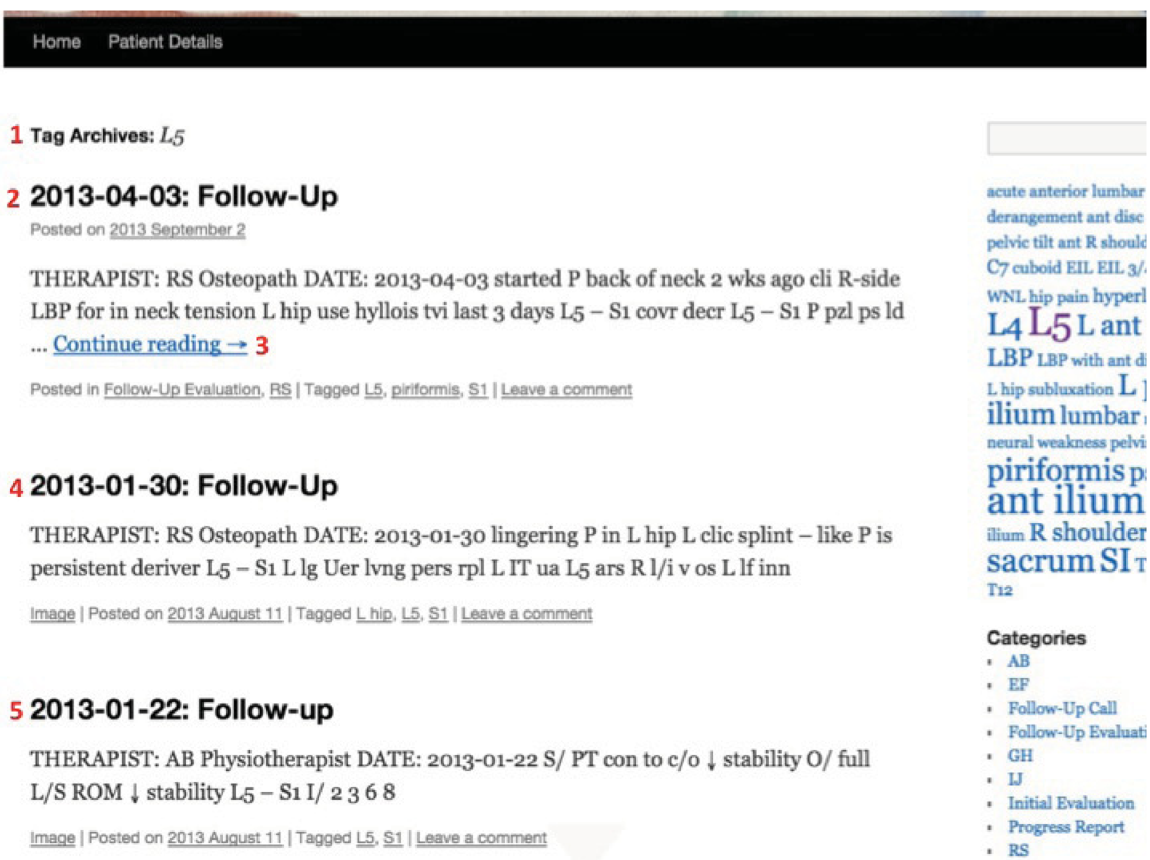


Figure 3.4 Tag Cloud tag selection results screen for tag *L5*.

3.3.3.6 Detailed Entry after selecting synopsis from either Search Box Results or Tag Cloud Tag Selection Results

Figure 3.5 below shows the screen with the detailed entry once a search result for either the search term or tag is selected. The different sections of this screen are labeled in Figure 3.5 below and described as such:

1. Previous (detailed) entry
2. Next (detailed) entry
3. Detailed visit entry of selected Search Box, or Tag Cloud, search results synopsis entry

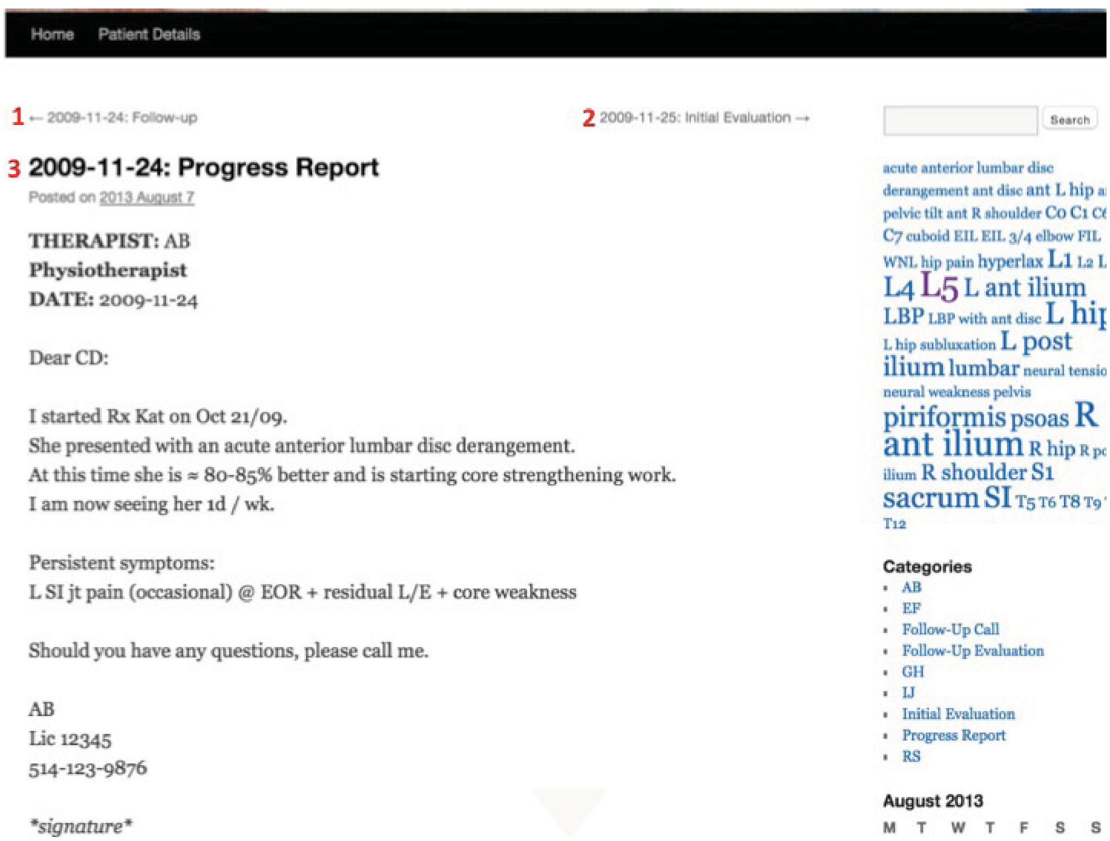


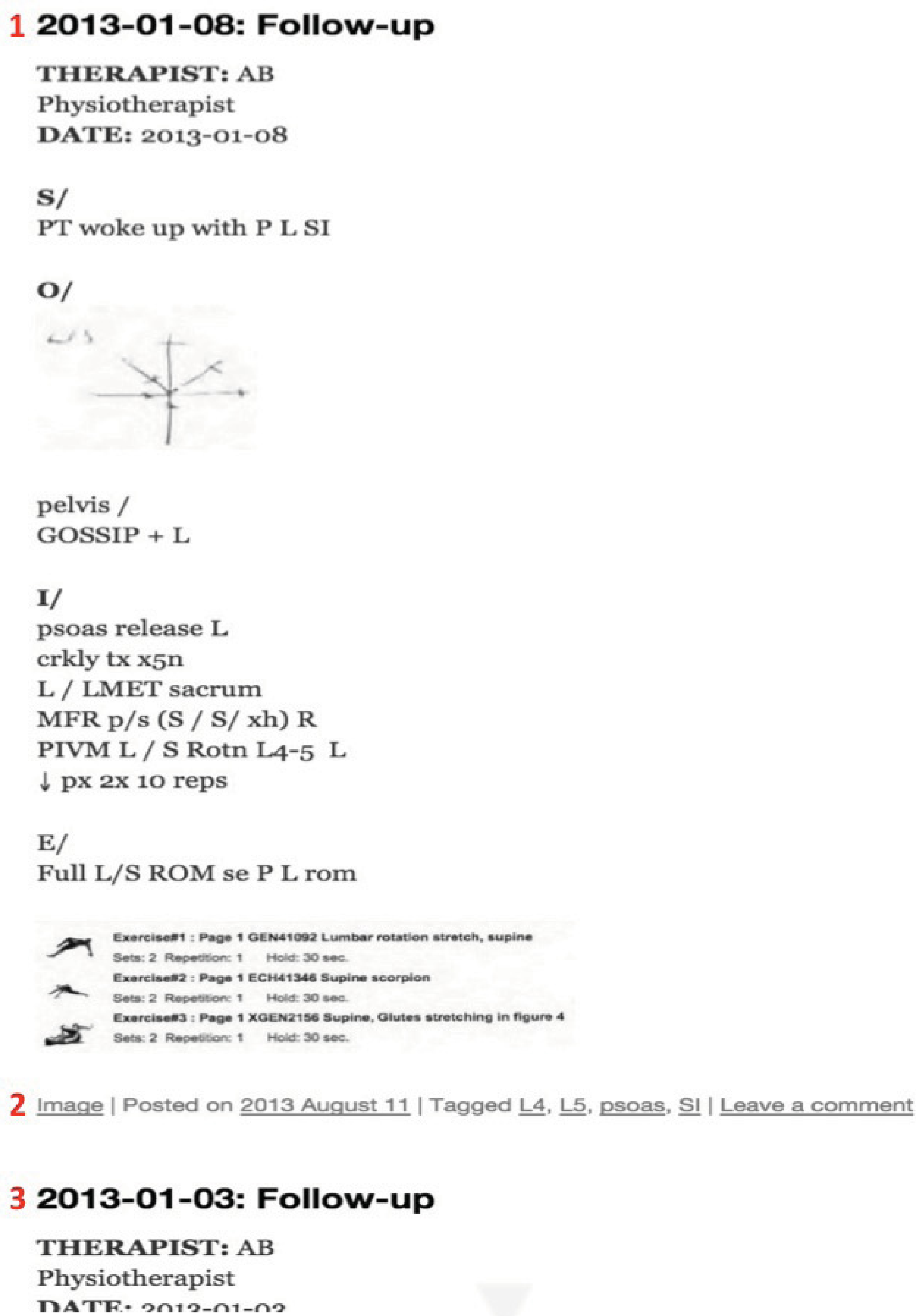
Figure 3.5 Detailed entry for patient visit after selecting result from either Search Box or Tag Cloud.

3.3.3.7 Scrolling View of Patient Record

Figure 3.6 below shows a snapshot of the scrolling view of the patient record. The different sections of this screen are labeled in Figure 3.6 below and described as such:

1. Most recent Patient Visit Entry
   a. Date of Visit and Visit Type

    b. Date entry was entered into the system
    c. Therapist’s Initials
    d. Therapy Type
    e. Date of Visit
    f. Therapist’s Notes
2. Quick links
    a. Image – brings up a search result page with all patient visit entries that have at least one image [*WordPress functionality – not part of the usability testing*]
    b. Posted on – goes to the detailed entry (ies) for that date [*WordPress functionality – not part of the usability testing*]
    c. Tagged – Displays tags entered for that visit and that populate the tag cloud. Selecting one of the tags from this section will bring up the search results screen for the tag selected, just like selecting the tag in the tag cloud.
    d. Leave a Comment [*WordPress functionality – not part of the usability testing*]
3. Previous Patient Visit Entry (beginning of)



1 **2013-01-08: Follow-up**

THERAPIST: AB
Physiotherapist
DATE: 2013-01-08

S/
PT woke up with P L SI

O/

pelvis /
GOSSIP + L

I/
psoas release L
crkly tx x5n
L / LMET sacrum
MFR p/s (S / S/ xh) R
PIVM L / S Rotn L4-5 L
↓ px 2x 10 reps

E/
Full L/S ROM se P L rom

Exercise#1 : Page 1 GEN41092 Lumbar rotation stretch, supine
Sets: 2 Repetition: 1 Hold: 30 sec.
Exercise#2 : Page 1 ECH41346 Supine scorpion
Sets: 2 Repetition: 1 Hold: 30 sec.
Exercise#3 : Page 1 XGEN2156 Supine, Glutes stretching in figure 4
Sets: 2 Repetition: 1 Hold: 30 sec.

2 Image | Posted on 2013 August 11 | Tagged L4, L5, psoas, SI | Leave a comment

3 **2013-01-03: Follow-up**

THERAPIST: AB
Physiotherapist
DATE: 2013-01-03

Figure 3.6 Snapshot of scrolling view of patient record (one full entry with beginning of previous entry).

3.3.3.8 Missing Elements and Future Enhancements

The prototype described in the previous sections, as mentioned, was created in the simplest way possible that would allow for usability testing. This means that certain functionality is missing in the prototype, which could be implemented if further, more in-depth usability testing is needed in the future.

Some of these missing elements are:

- The prototype was created using only one file, thus if additional patient files were to be added, the functionality to navigate to other patient files would need to be added to the prototype.
- The prototype used a web-based tool, which does allow for switching between multiple users of the system by means of the login functionality, but if the tool were implemented as part of a desktop application, the ability to switch between therapists from within the application would need to be implemented.

Future enhancements that could be tested in further research are:

- The use of several tag clouds, if one so desires, to display various different areas that may be of interest to the healthcare practitioner, such as one for symptoms, one for medication, one for lab results, one for allergies, or any other area that the healthcare practitioner would like to have access to while searching various areas of a patient's record. Understandably, displaying too many of these will over crowd the display and defeat the intended purpose of using tag clouds, so care must be taken to use but a few.
- There is currently no way to tag, nor search, the visual content (i.e. scanned images) within the prototype. Only entries, which can contain visual content, can be tagged and searched. This functionality would also be a consideration for future enhancements.

3.3.3.9 Constraints

The main constraints of developing a tag cloud in an EHR are:

- **The politics and economics of implementing EHRs**: This still remains the biggest constraint in developing any EHR, let alone specific functionality like a tag cloud.

- **Incorporating tags into the database:** Once this constraint is resolved, then it becomes an issue of dealing with the database. Where in the database the tags will be stored, as well as how they will be linked to entries and retrieved.
- **The mechanism for entering the tags into the EHR:** There are several methods that could be implemented to allow users to enter tags such as manual entry of tags, drop-down selection of pre-determined tags, or pre-filled tags based on content entered. Each method mentioned above has advantages and disadvantages, leaving the choice of method to those implementing the EHR.
- **Lexicon of clinical terminology:** Each clinic implementing a tag cloud as part of their EHR will have to put in place some governance as to the terminology used for tagging. This will ensure that each person, whether entering entries into the EHR or retrieving information from the EHR, will be working from the same set of tags. This will ensure the most pertinent results are returned when using the tag cloud to retrieve information.

3.4 SUMMARY

EHRs are slowly being implemented across North America and widely used in many other areas of the world. The current information visualization tools available in EHRs allow for information viewing and retrieval but not in a manner that allows easy viewing or retrieving of the pertinent information. Previous research has come up with alternate information visualization tools, but they too do not completely solve the problem. Tag clouds have been proposed as a novel way to solve this issue and a prototype was described that allows for usability testing of this information visualization tool. In the next chapter we will cover the usability testing of the tag cloud to validate its ability to easily display and retrieve pertinent data in a patient's file, as well as discuss the results of the usability testing.

4. EXPERIMENT AND RESULTS

This chapter contains the background information on the user population size, as well as the chosen questionnaires for usability testing. We then describe the experiment's usability testing protocol, followed by a discussion of the results obtained. We finish with the issues that were encountered during the usability testing.

4.1 BACKGROUND OF PROTOCOL DESIGN

The following section explains the rationale for the small sample size used in the usability testing of the prototype, as well as explaining the choice of questionnaires used for the usability testing. The application of these questionnaires, the Usability Metric for User Experience (UMUX) and the System Usability Scale (SUS), will also be described.

4.1.1 User Population Size

Twelve participants were recruited for the usability testing of the prototype. Although twelve is a small sample size, recruiting participants in these fields is difficult given the demands on their time. As will be seen in the following section, both the UMUX and SUS were used to evaluate the system. It has been shown that using SUS allows for a fairly confident measure of usability of a system with a small sample (8-12 users) [37] and in [38] it was shown that UMUX is an equivalent replacement for SUS. Therefore, we can be confident that the small sample of users for this experiment is sufficient.

4.1.2 UMUX

UMUX was chosen due to its reliability, as well as its short length of four Likert-scale questions (see Appendix B). This ensured that the usability testing was kept within an acceptable time frame, while still gathering valuable data.

UMUX is a standard set of four Likert-scale questions:

1. [This system's] capabilities meet my requirements.
2. Using [this system] is a frustrating experience.
3. [This system] is easy to use.
4. I have to spend too much time correcting things with [this system].

Where [this system] is replaced with the system or interface component being tested. The questions use a Likert-scale response with a range of 1 through 7, where 1 is Strongly Disagree and 7 is Strongly Agree. Questions are keyed to alternate between positive and negative questions to avoid acquiescence bias. Once data is collected, they need to be recoded, with the following scoring:

- Odd questions are scored as [response – 1]
- Even questions are scored as [7 – response]

This removes the positive/negative keying of the items and allows a minimum score of zero.

Each individual UMUX item has a range of 0 – 6 after recoding, giving the entire four-item scale a preliminary maximum of 24. To get a value on a scale of 100, a participant's UMUX score is the sum of the four items divided by 24, and then multiplied by 100. These scores across participants are then averaged to find a mean UMUX score. It is this mean score and its confidence interval that become the task's UMUX metric.

4.1.2 SUS

The SUS was also chosen for its relatively short length of ten Likert-scale questions (see Appendix C). The ten standard SUS questions are:

1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

The questions use a Likert-scale response with a range of 1 through 5, where 1 is Strongly Disagree and 5 is Strongly Agree. As with UMUX, SUS questions are keyed to alternate between positive and negative questions to avoid acquiescence bias. Once data is collected, it needs to be properly recoded, with the following scoring:

- Odd questions are scored as [response – 1]
- Even questions are scored as [5 – response]

As with the UMUX questions, we recode the positive/negative keying of the items, which allows a minimum score of zero. To calculate the SUS score, first sum the score contributions from each item. Then multiply the sum of the scores by 2.5 to obtain the overall value of SUS, ranging from 0-100.

The SUS provides a single number representing a composite measure of the overall usability of the system being tested. Although based on 100, this is not a percentage and the resulting value must be turned into its percentile value. This is done using the s-curve seen below in Figure 4.1. It is to be noted that an average SUS score is 68, which translate into a 50^{th} percentile value [39].

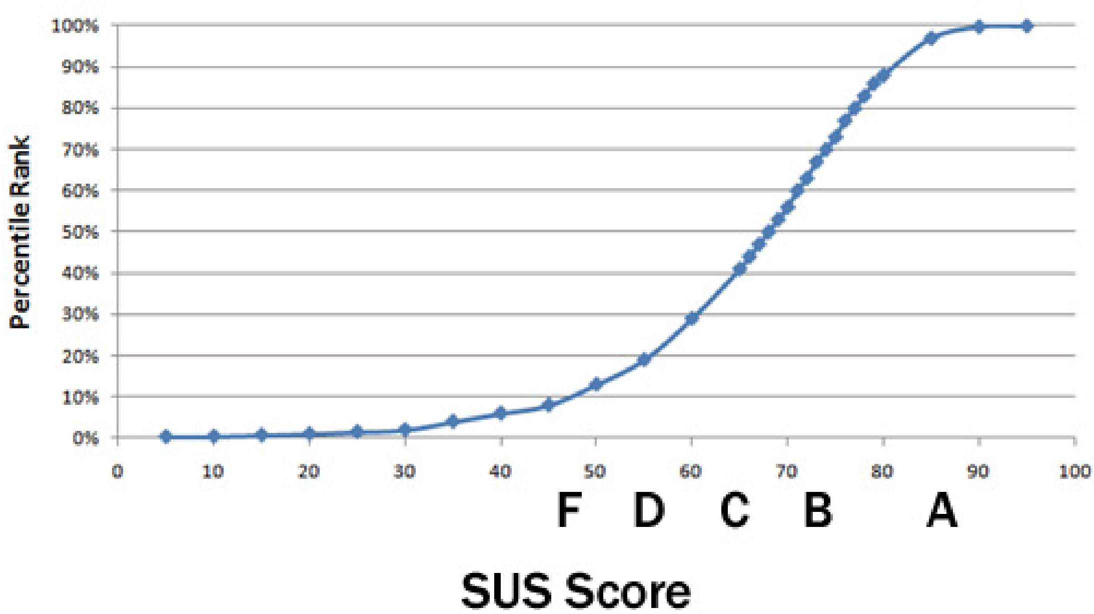


Figure 4.1 SUS S-Curve Percentiles (http://www.measuringu.com/images/sus-curve.png).

4.2 EXPERIMENT

In this section we summarize the experiment and usability testing protocol. The detailed usability Testing Protocol used during usability testing sessions can be found in Appendix A.

4.2.1 User Population

Participants were Physiotherapists, Osteopaths, Athletic Therapists, and Chiropractors. Usability testing was performed with twelve participants, where seven where Physiotherapists, three were Osteopaths, and there was one each of Athletic Therapist and Chiropractor. Participants were recruited either through the clinics where the researcher is a patient or through friends, family, and colleagues with acquaintances working in the aforementioned areas of occupation. All participants currently use paper patient records.

4.2.2 Usability Testing Session

The usability testing was conducted individually with each participant at the clinic in which they work, to ensure the least amount of time was taken away from their workday. Each session lasted between 20 and 45 minutes and included an introduction, completion of three tasks and associated questions, as well as a final debrief questionnaire. Notes were taken throughout the usability testing sessions, as well as audio recordings, of the comments verbalized by participants during the session.

The following disclaimers were provided to each participant prior to starting the scenario:

> *“Please keep in mind that this is a prototype, so not all functionality is present.”*
>
> *“As we tried to make things as realistic as possible, as in some of the entries being entered by a third person who is not the therapist (example: administrative assistant), you may see some entries that have been left blank for later entry or there may be typos in cases where the person entering the notes was unable to properly decipher the writing and did not get a chance to have it clarified.”*

Usability testing was done using one scenario with three tasks, each task using a different tool to search through the data (patient’s record).

The scenario given to each participant prior to completing each task was the following:

One of your colleagues has referred one of their patients to you, as they are leaving on a month-long vacation. The patient in question has come in with excruciating pain in her lower back and is having trouble sitting down. This is the first time you see this patient and you need to understand their medical history.

The three search tasks were:

1. Search Box
   - Use the Search Box to search for information within the patient's record.
2. Tag Cloud
   - Use the Tag Cloud to search for information within the patient's record.
3. Scrolling
   - Scroll through the patient record to search for information within the patient's record.

It should be noted that we were unable to run experiments with the LifeLines and CLEF tools because they are not publicly and freely available to researchers.

In an attempt to avoid the pitfall of participants remembering the last thing they saw, the above order of the tasks was chosen. Although the usual methodology of task order is to randomize the tasks for each participant, it was decided to make each participant follow the same order of tasks to ensure that the same conditions were re-created for each participant.

After each task was completed, users answered the Usability Metric for User Experience (UMUX) questionnaire [38].

The questions for each task were:

- TASK 1 – Search Box
  - The Search Box capabilities meet my requirements.
  - Using the Search Box is a frustrating experience.
  - The Search Box is easy to use.

    - I have to spend too much time correcting things with the Search Box.
- TASK 2 – Tag Cloud
    - The Tag Cloud capabilities meet my requirements.
    - Using the Tag Cloud is a frustrating experience.
    - The Tag Cloud is easy to use.
    - I have to spend too much time correcting things with the Tag Cloud.
- TASK 3 – Scrolling
    - The Scroll Bar capabilities meet my requirements.
    - Using the Scroll Bar is a frustrating experience.
    - The Scroll Bar is easy to use.
    - I have to spend too much time correcting things with the Scroll Bar.

At the end of each task, participants were asked if they had any additional comments that they wanted to share beyond what they had already given while performing the task. This allowed the collecting of valuable, rich information on the different tools used.

Once all three tasks had been performed, the System Usability Scale (SUS) questionnaire was used as a post-questionnaire to get a global rating of the system [40].

4.2.3 Confounds

- One of the participant sessions was not recorded due to the recording device malfunctioning during that particular usability testing session. All precautions were taken thereafter to ensure that such a situation did not occur again.
- Some participants were not familiar with working on a Mac. This was covered at the beginning of the session and explanations were given to these participants prior to starting the first task.
- The disclaimers read to the participants may have influenced their expectations regarding their experience with the system.

4.3 RESULTS

The key results obtained from the usability testing are discussed below. A detailed Usability Report with a more in-depth analysis can be found in Appendix D.

4.3.1 Summary of Results

- Overall, 83% (10/12 participants) had a very positive reaction to the tag cloud tool.
- The tag cloud was preferred by 50% of participants, while 25% of participants preferred the search box and 25% the scrolling. This supports H1: The Tag Cloud allows more efficient visualization of patient history as compared to existing visualization methods.
- The UMUX mean scores gave results of 81 for the tag cloud, followed by the search box with 71, then scrolling with 68. This supports H2: The Tag Cloud allows for easier retrieval of more pertinent patient information as compared to existing search methods.
- The system as a whole, that is the prototype with all search methods, had a mean SUS score of 80, which places it at the 90^{th} percentile.

4.3.2 Detailed Results

The collected data is not normally distributed, probably due to its small size. Further work is required to improve the results by collecting data from a larger sample of therapists; this is covered in the future work section.

Using a T-distribution to account for the small sample size, the results show that when asked which search method they preferred, the participants' answers to the UMUX showed that the tag cloud was preferred, followed by the Search Box, and then Scrolling.

We can see the results obtained from usability testing with the twelve participants using a 95% t-distribution confidence interval in Table 4.1 and Figure 4.2.

Table 4.1 UMUX Results for the 3 Tasks using a 95% t-distribution confidence interval.

	Lower Confidence Interval	Mean	Upper Confidence Interval
Tag Cloud	66	81	96
Search Box	56	71	86
Scrolling	50	68	86

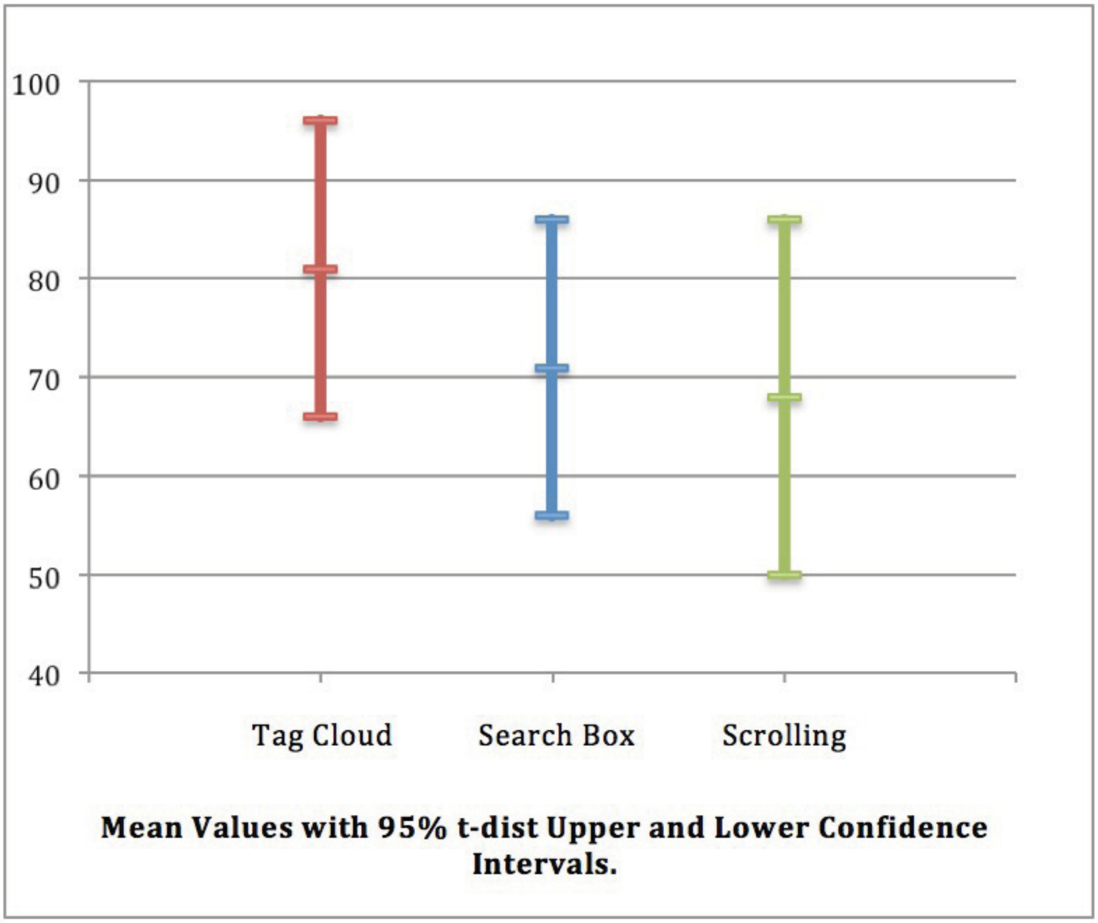


Figure 4.2 Mean Values with 95% t-distribution Upper and Lower Confidence Intervals for UMUX Results for the 3 Tasks.

If we remove the lowest and highest scores and calculate the mean and 95% t-distribution confidence interval for each task with the remaining ten (10) users, we get the results seen in Table 4.2 and Figure 4.3 below. In removing the lowest and highest scores for each of the three tasks, we can see that the confidence interval for the tag cloud task is tighter and greater than the mean of each of the other two tasks.

Table 4.2 UMUX Results for the 3 Tasks removing lowest and highest score and using a 95% t-distribution confidence interval.

	Lower Confidence Interval	Mean	Upper Confidence Interval
Tag Cloud	76	85	94
Search Box	57	73	89
Scrolling	51	69	87

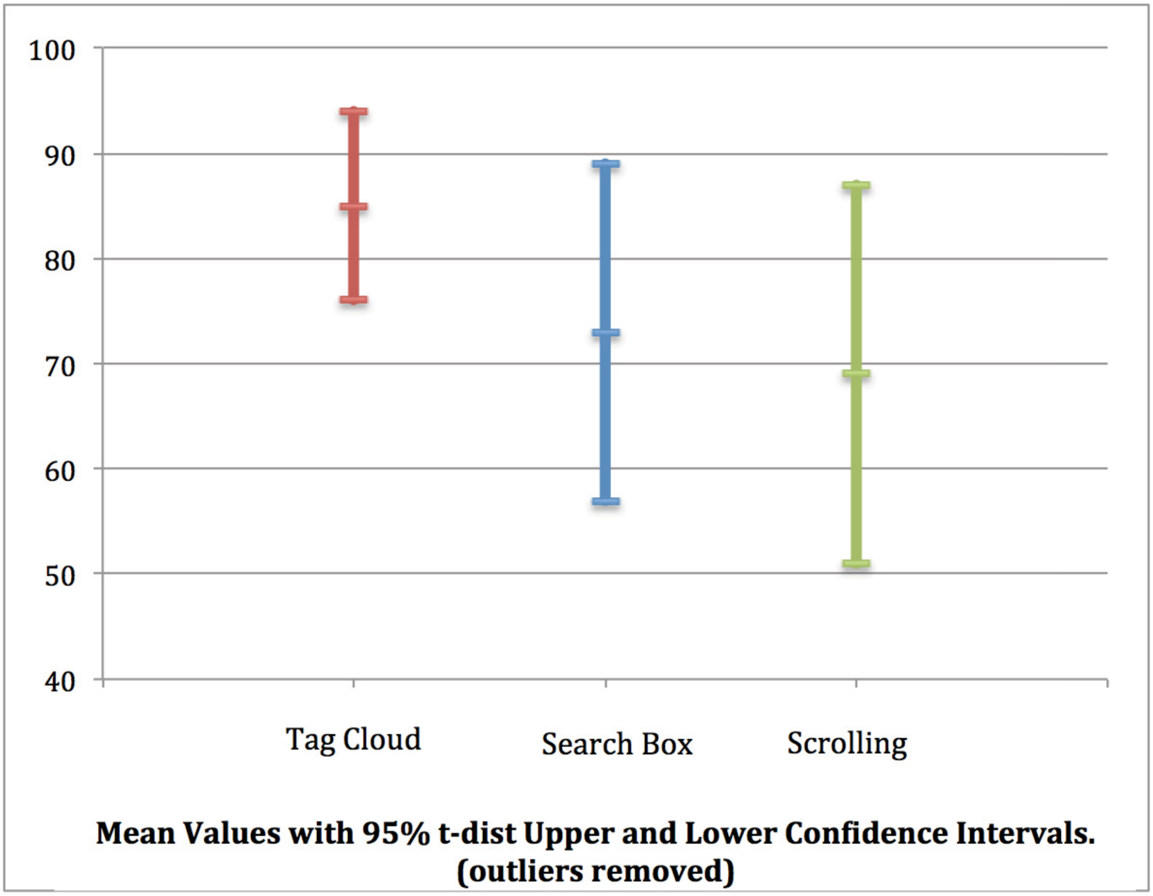


Figure 4.3 Mean Values with 95% Upper and Lower Confidence Intervals for UMUX Results for the 3 Tasks removing lowest and highest score.

Due to the small sample size, to see a greater difference between the tag cloud results and the other two search methods, we used a 70% t-distribution confidence interval. The results with the 70% t-distribution can be seen in Table 4.3 and Figure 4.4.

Table 4.3 UMUX Results for the 3 Tasks using a 70% t-distribution confidence interval.

	Lower Confidence Interval	Mean	Upper Confidence Interval
Tag Cloud	74	81	88
Search Box	63	71	79
Scrolling	59	68	77

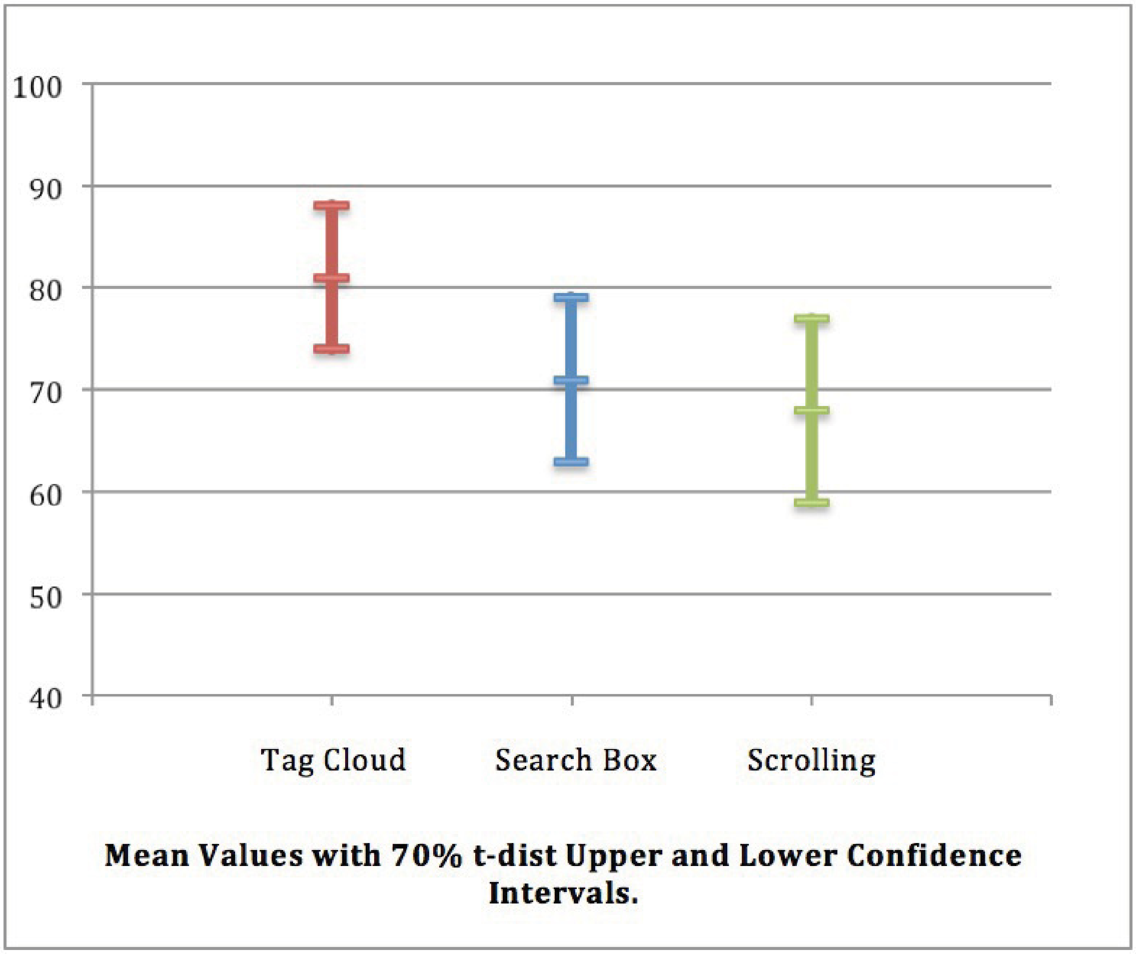


Figure 4.4 Mean Values with 70% t-distribution Upper and Lower Confidence Intervals for UMUX Results for the 3 Tasks.

If we once again remove the lowest and highest scores and calculate the mean and use a 70% t-distribution confidence interval for each task with the remaining ten (10) users, we get the results seen in Table 4.4 and Figure 4.5 below.

Table 4.4 UMUX Results for the 3 Tasks removing lowest and highest score and using a 70% t-distribution confidence interval.

	Lower Confidence Interval	Mean	Upper Confidence Interval
Tag Cloud	81	85	89
Search Box	65	73	81
Scrolling	60	69	78

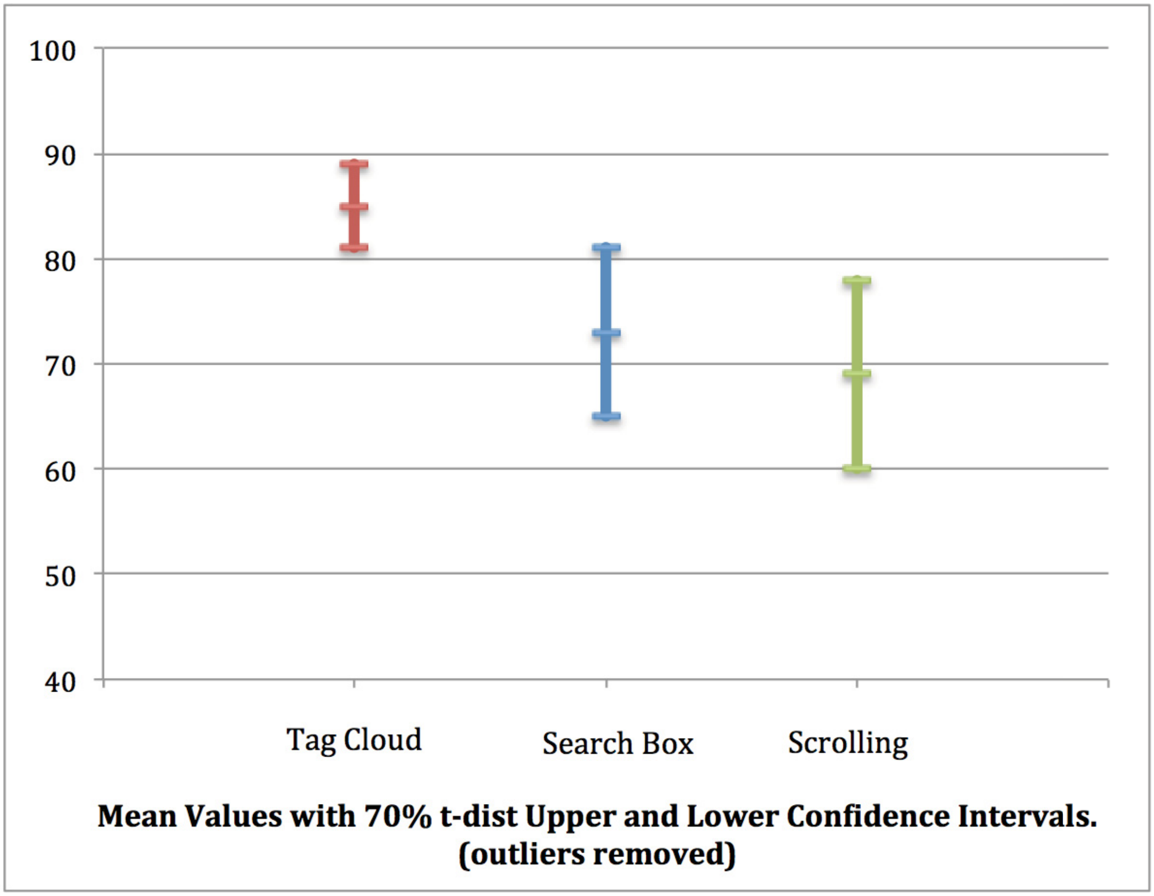


Figure 4.5 Mean Values with 95% Upper and Lower Confidence Intervals for UMUX Results for the 3 Tasks removing lowest and highest score and using a 70% t-distribution confidence interval.

In removing the lowest and highest scores for each of the three tasks and using a 70% t-distribution confidence interval, we can see that the confidence interval for the tag cloud distinguishes itself from the other two tasks' confidence interval. These results show that the tag cloud is the preferred search method by the participants.

4.3.3 Participant Comments

Participants were encouraged to talk aloud while performing the tasks and in doing so many of them had comments that corroborated the results seen above.

The most comments made were during the tag cloud task, as a majority of the participants had a strong positive reaction to the tag cloud. Comments received for this task included:

> *"This tag cloud is nice"*

"I like that! I like that!"

"I much prefer this as opposed to putting it in a search box, simply because I can just click on it instead of typing."

"I like the idea that it gets bigger when it's recognized more often. I think that's very valuable. I mean it's easy; it's the first thing you see. So I think it's a very quick tool to see if what the person comes in presenting with, especially since you've never seen this person before, that it's a very quick thing. That, 'oh ya this is either something really new or this is a possible exacerbation of something they've had in the past' and that's nice to know. It's valuable for the patient."

"Ok, so this kind of just gives you a synopsis of what's like reoccurring kind of dysfunction with the person"

"Like that is a synopsis of what the patient has had."

"This is cool."

"What I like more about the cloud, is that if you're looking for something in particular, you can find it right away."

Comments on the other tasks were few and not as enthusiastic. As can be seen below, only two participants made comments on the search box and only four for the scrolling.

Of the Search Box:

"I'm not a huge fan of the search box."

"I'm not a fan of search boxes."

Of the Scrolling:

"I like this one for its ease of use, for the first few. If the person's been in 60 times to the clinic, it's a disadvantage. It's just nice to see past history."

"Love how it reads, it's fine, very clear, it's just I have to scroll"

"Wouldn't really feel the need to look that far back"

"Where could see being useful [...] comments [...] kind of neat to see that way"

Comments verbalized by participants during the usability testing, as well as comments and discussions held after testing sessions were complete, indicated that the tag cloud was the preferred tool by the majority of participants. Many participants were excited to see such a tool and wanted to make use of it in their every day work as soon as possible.

4.3.4 Final Debrief Questionnaire

As a final debrief questionnaire, the SUS was used to evaluate the entire prototype, not just the different search methods tested. The results for the SUS gave a mean SUS score of 80, which is equivalent to the 90th percentile for SUS. An average SUS score being 68 (50th percentile), this means that the usability of the prototype is considerably above average.

Once again, if we remove the lowest and highest scores and calculating with the remaining ten (10) users, we get a SUS score of 88, which is equivalent to the 98th percentile. Thus showing the usability of the prototype to be considered extremely high by the majority of participants.

It should be noted that although we did find a high SUS score for the prototype, it might be due to the fact that therapists are still working with paper files. In other words, because of the arduous task of searching through paper records, the high rating might be due to simply having something that is better than paper records. This notion was not tested as part of the research presented in this thesis.

4.4 TESTING OBSERVATIONS AND REMARKS

There were a few difficulties when evaluating the prototype. The data entry of the patient file was done in a realistic manner, meaning that it was entered by a third party, as a therapist's receptionist might be tasked with doing. This meant that there were some typos in the patient's file, as well as instances where the notes for the patient visit had not been entered yet thus creating an empty entry, which although it did not hinder the usability testing, it did mean that there were some questions asked by the participants that would not have otherwise been asked had all entries been completed in the patient's file.

Another difficulty was the shorthand used by one therapist is not necessarily the same as the shorthand used by all therapists. In some areas of practice, such as in Physiotherapy there is a standard shorthand that is used, but others, such as Osteopathy, Athletic Therapy, and Chiropractic, do not have a standard shorthand. This meant that some therapists had additional questions related to the content of certain entries.

A related difficulty to common shorthand is a need for a common lexicon for tagging entries. Although the entries were tagged from the same set of tags, there is a need for governance of terminology in any system that is created to ensure that all users are tagging and searching on the same set of terms.

Some users mentioned the order of tasks might have been better if scrolling was first, as it would have helped them with the search and tag cloud tasks, as they would have known the terminology used in the file, making searching more productive. In a real world setting this would not have been needed, as users would have already been familiar with the terminology given that they all work within the same clinic and make use of the same file when treating a given patient and a common set of terminology is a cornerstone in any good data management system, not only in EHRs.

Demographic differences were not readily available, as the sample size did not provide enough participants to get a representative amount of users in each demographic category, namely when comparing differences between therapist types or age brackets.

4.5 SUMMARY

Based on the results obtained using UMUX, we have demonstrated that tag clouds are a useful tool for naive users who are looking for historical information on a new patient. This, combined with the positive comments from participants to the tag cloud tool, warrant further investigation of tag cloud use in EHRs.

5. CONCLUSION AND FUTURE WORK

In this chapter, we provide a summary and a conclusion of our research work. We will also suggest some research directions to be undertaken in the near future.

5.1 SUMMARY AND CONCLUSION

With increased pressure on the medical community to transition from paper patient records to EHRs, finding ways to allow for efficient information visualization and retrieval of the information within these EHRs is crucial. In this thesis, we reviewed information visualization methods and tools for EHRs, the current state of EHRs, tag clouds as an information visualization tool, and applying tag clouds to EHRs. We proposed a new tag cloud-based EHR information visualization approach and tool, created a prototype, and conducted usability testing on the tag cloud tool. We compared the usability of different search methods – tag cloud, search box, and scroll bar, to show that the tag cloud is the preferred tool for searching a patient's data.

Current information visualization tools for EHRs are efficient, but not comprehensive. These tools could be enhanced to offer improved information visualization in EHRs. Plaisant's LifeLines covers most of the key elements of a good information visualization tool for EHRs, visually easy to understand, learn, navigate, magnifies subtle elements of data, and prevents information overload, but requires the user to navigate back to the main screen each time they want to see the overview. Reviewing tag clouds and their current use for Internet searches, as well as in data repositories, we proposed a tag cloud based EHR visualization tool that allows a healthcare practitioner to easily visualize and retrieve the essential elements of a patient's medical record.

After conducting usability testing on a tool's prototype containing a tag cloud for EHR information visualization and information retrieval, the results of the usability testing indicated that tag clouds are deemed useful by therapists searching for pertinent information in a patient's EHR. The usability testing also showed that the search box and scrolling were not as preferred by participants for searching for information in the patient's file.

As mentioned in Chapter 3, the SUS post-questionnaire results were likely high because therapists currently work with paper records, thus any system that allows them to search the content of a patient's EHR is likely to be well received.

The preliminary results from the usability testing conducted have shown tag clouds to be an effective tool in an EHR for therapists seeing a patient for the first time. The tag cloud allows them to quickly find pertinent historical information that would be beneficial to the patient's treatment at the time of their visit. This allows the therapist to spend more time treating the patient, as they can quickly go to the pertinent past visits to find the information they need for the current visit. In conclusion, the results suggest that integrating tag clouds into an EHR would be beneficial to both therapist and patient alike whether it is a first visit or subsequent visit.

5.2 FUTURE WORK

There are several areas where the research presented in this thesis could be expanded in the future. Some of these include expanding to other areas, medical and non-medical, onto mobile devices, testing with different demographics, using multiple tag clouds, as well as research on how best to tag entries.

Given the benefit of tag clouds with therapists shown in this thesis, testing with a larger group to ensure that the results are valid across demographics within the therapy field would be recommended. As well, testing with patient files of varying sizes to see if similar results are seen for small, medium, and large patient files. Then testing with other healthcare practitioners, such as nurses, doctors, veterinarians, pharmacists, or even clinical research labs, would be interesting to see if different contexts of use would influence the results seen in this thesis. Further research with these users warrants looking into as an extension to the research conducted and presented in this thesis. Expanding the present research to non-medical fields would also warrant investigation to see if tag clouds can be extended to other software applications that contain large datasets.

An interesting area of research that would be beneficial to investigate is testing the tag cloud tool on mobile devices, whether tablets or phones. Mobile devices are portable, thus much easier to carry around making them important tools in future EHRs. Exploring information visualization tools that work on these devices is an important area that should not be neglected.

Another area that would require further investigation is extending tag cloud visualization of the current thesis' research focused on symptoms to other data categories, such as measured data and planning data as seen in Chapters 2 and 3, that might lend themselves well to being presented in a tag cloud. This research would investigate whether it is best to combine all of these data categories into one tag cloud or if it would be beneficial to have multiple tag clouds, one per data category. Examples of this might be a tag cloud for each of Diagnosis, Treatment, Medical History, Medication, Therapist Type, or any other category that might benefit from being presented in its own tag cloud.

With regards to the tags themselves, further investigating would be needed to see how tagging is to be done within an EHR when entering the visit information. Some additional topics to cover would be testing different algorithms for the tag cloud generation, how to handle synonyms, shorthand, or different languages such as French in the case of EHRs in Quebec. These questions would require further investigation to ensure that the tag cloud built from the identified tags contains the most pertinent tags that will allow for the best results when retrieving information from the tag cloud.

Taking the topic of the tag cloud one step further, investigating alternate forms of tag clouds such as tag clusters [28] or dynamic tag clouds that combine the timeline and tag cloud would be interesting. Testing these alternate forms of tag clouds to see if they would be preferred to the search methods tested in the usability testing of this thesis would allow to see what visualization methods are selected as most useful by users.

Part of testing any of the above-mentioned forms of tag clouds, or different tag cloud algorithms, should also include testing different text search algorithms at the same time. Testing search accuracy as part of these future tests would also be advisable to ensure that results take into account all aspects of search.

6. REFERENCES

[1] S. C. Biro, D. T. Barber and J. A. Kotecha, "Trends in the use of electronic medical records," Canadian Family Physician, vol. 58, pp. e21, 2012.

[2] C. P. D. Hsiao and E. M. P. H. Hing. Use and characteristics of electronic health record systems among office-based physician practices: United States, 2001–2013. NCHS Data Brief 143(January 2014), 2014.

[3] J. Besner, "When it comes to electronic health-care records, we're so last century," (Health Council Canada), [online] 2006, http://www.healthcouncilcanada.ca/dl_check.php?id=104&tp=1 (Accessed: 10 October 2014).

[4] D. Thompson, "U.S. Hospitals Triple Use of Electronic Health Records: Report," (US News: Health), [online] 2013, http://health.usnews.com/health-news/news/articles/2013/07/08/us-hospitals-triple-use-of-electronic-health-records-report (Accessed: 10 October 2014).

[5] G. G. Grinstein and M. O. Ward. "Introduction to data visualization," in *Information Visualization in Data Mining and Knowledge Discovery*, 1st Edition, San Francisco: Morgan Kaufman Publishers Inc., 2001, pp. 21-45.

[6] R. Spence, *Information Visualization: An Introduction*, 3rd Edition, New York: Springer International Publishing, 2014.

[7] C. Ware, *Information Visualization: Perception for Design*. 2nd Edition, San Francisco: Morgan Kaufmann, 2004.

[8] R. Kosara and S. Miksch, "Visualization methods for data analysis and planning in medical applications," Int. J. Med. Inf., vol. 68, pp. 141, 12/18, 2002.

[9] L. Chittaro, "Information Visualization and its Application to Medicine", Artificial intelligence in medicine, Vol 22(2), pp. 81-88, 2001.

[10] A. A. T. Bui, D. R. Aberle and H. Kangarloo, "TimeLine: Visualizing Integrated Patient Records," in IEEE Transactions on Information Technology in Biomedicine, Vol. 11, pp. 462-473, 2007.

[11] C. Hallett, "Multi-modal presentation of medical histories," in Proceedings of the 13th International Conference on Intelligent User Interfaces, Gran Canaria, Spain, 2008, pp. 80-89.

[12] C. P. Phd, C. Plaisant, R. Mushlin, A. Snyder, J. Li, D. Heller, B. Shneiderman and K. P. Colorado, "LifeLines: Using visualization to enhance navigation and analysis of patient records," in Proceedings of the 1998 American Medical Informatic Association Annual Fall Symposium, 1998, pp. 76-80.

[13] Tag (metadata), Wikipedia, [online] 2013, http://en.wikipedia.org/wiki/Tag_(metadata) (Accessed: 5 February 2013).

[14] E. Connor, "Library 2.0," Med. Ref. Serv. Q., vol. 26 (01/02), pp. 5-23, 2007.

[15] Tag cloud. Wikipedia, [online] 2013, http://en.wikipedia.org/wiki/Tag_cloud (Accessed: 9 February 2013).

[16] V. Friedman, "Tag Clouds Gallery: Examples And Good Practices," (Smashing Magazine), [online] 2007, http://www.smashingmagazine.com/2007/11/07/tag-clouds-gallery-examples-and-good-practices/ (Accessed: 11 February 2013).

[17] B. S. Bulik and A. Klaassen, "Tag Clouds," Advert. Age, vol. 77, pp. 19-19, 2006.

[18] A. Hubmann-Haidvogel, A. Scharl and A. Weichselbraun, "Multiple coordinated views for searching and navigating Web content repositories," Inf. Sci., vol. 179, pp. 1813-1821, 2009.

[19] D. Gill and A. Griffin, "Good Medical Practice: what are we trying to say? Textual analysis using tag clouds," Med. Educ., vol. 44, pp. 316-322, 2010.

[20] D. Shotton, K. Portwin, G. Klyne and A. Miles, "Adventures in Semantic Publishing: Exemplar Semantic Enhancements of a Research Article," PLoS Computational Biology, vol. 5, pp. 1-17, 2009.

[21] C. Görg, Z. Liu, J. Kihm, J. Choo, H. Park and J. Stasko, "Combining Computational Analyses and Interactive Visualization for Document Exploration and Sensemaking in Jigsaw," in IEEE Transactions on Visualization and Computer Graphics, vol. PP, pp. 1-1, 2013.

[22] O. Thiele, H. Knapp, M. Schader and N. Prat, "MobileAERO: Using tag clouds for mobile knowledge management," in Seventh International Conference on Networking, 2008, pp. 686-691, 2008.

[23] N. Cao, D. Gotz, J. Sun, Y. Lin and H. Qu, "SolarMap: Multifaceted visual analytics for topic exploration," in IEEE 11th International Conference on Data Mining (ICDM), 2011, pp. 101-110, 2011.

[24] M. S. Wood, "Medical Librarian 2.0: Use of Web 2.0 Technologies in Reference Services", Medical reference services quarterly, Vol. 26 (supplement no. 1), pp. 1-3, 2007.

[25] Anonymous, "From Russia with results," KM World, vol. 15 (03), pp. 34-34, 2006.

[26] K. Seongwoon, L. Inseong, L. Kiho, S. Jung, J. Park, K. Yeun Bae, K. Sang Ryong, and K. Jinwoo, "Mobile Web 2.0 with Multi-Display Buttons," Commun ACM, vol. 53 (01), pp. 136-141, 2010.

[27] J. Sinclair and M. Cardew-Hall, "The folksonomy tag cloud: when is it useful?" Journal of Information Science, Vol. 34, pp. 15-29, 2008.

[28] K. Knautz, S. Soubusta and W. G. Stock, "Tag clusters as information retrieval interfaces," in 43rd Hawaii International Conference on System Sciences (HICSS), 2010, pp. 1-10.

[29] O. Hoeber and Hanze Liu, "Comparing tag clouds, term histograms, and term lists for enhancing personalized web search," in IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010, pp. 309-313.

[30] C. Fung and W. Thanadechteemapat, "Discover information and knowledge from websites using an integrated summarization and visualization framework," in Third International Conference on Knowledge Discovery and Data Mining WKDD, 2010, pp. 232-235.

[31] D. A. Ludwick and J. Doucette, "Adopting electronic medical records in primary care: Lessons learned from health information systems implementation experience in seven countries," Int. J. Med. Inf., Vol. 78, pp. 22, 2009.

[32] B. J. Lesselroth and D. S. Pieczkiewicz, "Data visualization strategies for the electronic health record," in Advances in Medicine and Biology, 2012, pp. 107.

[33] V. L. West, D. Borland and W. E. Hammond, “Innovative information visualization of electronic health record data: A systematic review,” Journal of the American Medical Informatics Association, Vol. 22(2), pp. 330-339, 2015.

[34] W. McCoy, C. O. Alm, C. Calvelli, R. Li, J. B. Pelz, P. Shi and A. Haake, "Annotation schemes to encode domain knowledge in medical narratives," in Proceedings of the Sixth Linguistic Annotation Workshop, Jeju, Republic of Korea, 2012, pp. 95-103.

[35] R. Grams, "The “New” America Electronic Medical Record (EMR)—Design Criteria and Challenge," J. Med. Syst., vol. 33 (12), pp. 409-411, 2009.

[36] WordPress, WordPress, [online] 2013, https://wordpress.com/ (Accessed: 3 August 2013).

[37] T. S. Tullis, and J. N. Stetson, "A Comparison of Questionnaires for Assessing Website Usability," (Comcast), [online] 2004,

http://home.comcast.net/~tomtullis/publications/UPA2004TullisStetson.pdf (Accessed: 22 April 2015).

[38] K. Finstad, "The Usability Metric for User Experience," Interact Comput, vol. 22, pp. 323-327, 2010.

[39] System Usability Scale, Usability.gov, [online] 2015, http://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html (Accessed: 5 May 2015).

[40] J. Sauro, "Measuring Usability With The System Usability Scale (SUS)," (Measuring Usability), [online] 2011, http://www.measuringusability.com/sus.php (Accessed: 5 May 2015).

7. APPENDICES

A. USABILITY TESTING PROTOCOL

Information Access by Naïve Users Testing Protocol

Session Script, Data-Collection Sheets, and Participant Forms

Daphne Foldes, Master's Student

Computer Science & Software Engineering Department, Concordia University

[Version 2 | 2013-08-11]

Information Access by Naïve Users – Testing Protocol

This document is the testing protocol for the Master's project Information Access by Naïve Users, which Daphne Foldes, Master's student at Concordia's Computer Science Department, will conduct with up to 10 participants. This document contains all the materials needed for the observation sessions, including the session set-up checklist, the facilitator's script, data-collection sheets, and participant forms.

Session Set-Up Checklist

Before a session begins, perform all set-up tasks for the session as follows:

COMPUTER	
Start Computer	
Set Up Voice Recorder	
Do Test Recording	
Verify Sound in Test Recording	
Set Prototype to Start Page	
PARTICIPANT PAPERWORK	
Consent Form	
Introduction	
Pre-Questionnaire	
Script and Questionnaire	
Post-Questionnaire	
Debrief, Gift Card and Receipt of Compensation Form	

Participant Paperwork

Consent Form

CONSENT TO PARTICIPATE IN INFORMATION ACCESS BY NAÏVE USERS

I understand that I have been asked to participate in a research project being conducted by *Daphne Foldes* of the *Computer Science & Software Engineering Department* of Concordia University ([redacted]*)* under the supervision of *Prof. Thiruvengadam Radhakrishnan* of the *Computer Science & Software Engineering Department* of Concordia University ([redacted]*).*

A. PURPOSE

I have been informed that the purpose of the research is as follows:

To see if using a tool that allows to see at a glance the main items in a patient's history of visits would be useful to practitioners seeing a patient for the first time, either as a referral or taking over the patient's file. The goal is to prove, or disprove, that this tool is more useful than other methods of searching for the same information in a patient's file.

B. PROCEDURES

- *I understand that an interview will be conducted at my place of employment at a time that is mutually agreeable to both myself and the principal investigator.*
- *I understand that during this interview, I will be asked to complete tasks using a prototype of the aforementioned tool while answering questions asked by the principle investigator both while performing these tasks, as well as after completing all of these tasks.*
- *I understand that my participation will be recorded, both orally and on paper, by the principal investigator to ensure accuracy of the answers provided for the purpose of extracting statistics to complete the goal of the study.*
- *I understand that the recorded voice and verbal statements from my participation in this study will be used to extract the aforementioned statistics, as well as provide quotes to support the statistics.*
- *I understand that the entire interview process will last between 30 and 45 minutes, and not longer than an hour, depending on the richness of the answers provided to the questions asked by the principal investigator.*
- *I understand that my name will not be associated with any of the answers given to the questions asked by the principal investigator, nor in the quotes that may be used.*
- *I understand that my age, gender, and area of practice will be noted and used for statistical purposes in conjunction with the answers given to the questions asked during the interview.*

C. RISKS AND BENEFITS

- *I understand that there are no risks in my participating in this study; I will be viewing and using a prototype while answering questions, both during and after the testing, related to this prototype.*

D. CONDITIONS OF PARTICIPATION

- I understand that I am free to withdraw my consent and discontinue my participation at any time without negative consequences.

- I understand that in the eventuality that I should choose to withdraw my consent and discontinue my participation at any point during the process, any data collected up until that point will still be used in this research.

- I understand that my participation in this study is CONFIDENTIAL (i.e., the researcher will know, but will not disclose my identity).

- I understand that the statistical data and verbal statements collected from this study may be published.

- I understand that the content, data and prototype, shown to me as part of this study is PROPRIETATRY and CONFIDENTIAL and as such that I will not divulge to anyone any information regarding what I have seen or discussed during this study. This will apply from when I sign this consent form up to, and including, the date of publication of the results of this study.

I HAVE CAREFULLY STUDIED THE ABOVE AND UNDERSTAND THIS AGREEMENT. I FREELY CONSENT AND VOLUNTARILY AGREE TO PARTICIPATE IN THIS STUDY.

NAME (please print) ____

SIGNATURE ____

If at any time you have questions about the proposed research, please contact the study's Principal Investigator.

Daphne Foldes
Master's Student
Computer Science & Software Engineering Department
Concordia University
[redacted]

Or

Prof. Thiruvengadam Radhakrishnan
Computer Science & Software Engineering Department
Concordia University

[redacted]

If at any time you have questions about your rights as a research participant, please contact the Research Ethics and Compliance Advisor, Concordia University, 514.848.2424 ex. 7481 ethics@alcor.concordia.ca

Introduction

Hello [*name*] and thank you for accepting to participate in this research study.

My name is **Daphne** and I will be accompanying you throughout the interview, which should last approximately **45 minutes**.

I will be showing you a **prototype of a patient record system** on which I will be asking you to execute **a few tasks**, as well as **asking for your feedback**. Once we complete the tasks, there will be a few additional questions to answer.

First and foremost, I am looking to get **your feedback** on the screens and the **tools** that are part of the prototype. There are **no right or wrong answers**, so please do not be shy with your comments.

As indicated on the consent form that you have signed previously, the **session will be recorded** so that your answers can be accurately compiled into the study's results.

I want you to be comfortable in giving feedback, as we are evaluating the prototype, not you.
Please feel free to give me your **honest response to the prototype**.

Is everything clear for you up until now?
Do you have **any questions before we start**?

Great! Let's start the session and the recording.

Pre-Questionnaire

Before we get started with the prototype, I have a few questions for you.

Demographic Information

Pre-Q 1
Your gender is?

	Male
	Female

Pre-Q 2
What is your main occupation?

	Physiotherapist
	Osteopath
	Chiropractor
	Athletic Therapist
	Other (Specify)

Pre-Q 3
What age bracket do you belong to?
(Show cue card with options)

	18-24
	25-29
	30-34
	35-39
	40-44
	45-49
	50-54
	55-59
	60-64
	65-69

Script and Questions

PLEASE NOTE: *The Scenario and tasks wording may vary from what is indicated on the paper based on the interaction with the participant. What is written down within here is a guideline and should be followed as closely as possible, but human nature being as it is, participants may not interact in such a way that will allow for following the script 100%. Thus, there may be some deviation from what is in the script, but the general goal will remain the same and the questions asked will remain the same.*

Today we're going to run through one **(1) scenario** that will have three **(3) tasks** to complete.

For this scenario and the tasks, we will be using the **prototype** that is **on the computer in front of you**.

Please keep in mind that this is a prototype, so not all functionality is there.

The scenario that we will see will **revolve around searching through a patient record file** for information on that patient.

During the entirety of the session, please **feel free to talk out loud** and **let me know what you are doing & thinking**.

SCENARIO
One of your colleagues has referred one of their patients to you, as they are leaving on a month-long vacation. The patient in question has come in with excruciating pain in her lower back and is having trouble sitting down. This is the first time you see this patient and you need to understand their medical history.

Task A - SEARCH BOX SEARCH (i.e. make user use search box to search)

Goal: Make the user use the Search Box to search the patient's record for relevant information, so that they can see if they like and/or prefer this method of searching the patient's record.

Re-iterate scenario if needed
One of your colleagues has referred one of their patients to you, as they are leaving on a month-long vacation. The patient in question has come in with excruciating pain in her lower back and is having trouble sitting down. This is the first time you see this patient and you need to understand their medical history.

Task: To understand the patient's history, you need to find all relevant information within their record. This time we will use a specific method, the search box, to search the patient's record. Please walk me through your thought process and your reasoning while using the search box to search the patient's record.

Observed Results

Completed	**Completed with Assist**		**Not completed**
	General Assist		
	Specific Assist		
	# of Assists		

Q A1
The Search Box capabilities meet my requirements.

	1- Strongly Disagree
	2
	3
	4
	5
	6
	7 – Strongly Agree

Q A2
Using the Search Box is a frustrating experience.

	1- Strongly Disagree
	2
	3
	4
	5
	6
	7 – Strongly Agree

Q A3
The Search Box is easy to use.

	1- Strongly Disagree
	2
	3
	4
	5
	6
	7 – Strongly Agree

Q A4
I have to spend too much time correcting things with the Search Box.

	1- Strongly Disagree
	2
	3
	4
	5
	6
	7 – Strongly Agree

Comments

Task B - TAG CLOUD SEARCH (i.e. make user use tag cloud to search)

Goal: Make the user use the Tag Cloud to search the patient's record for relevant information, so that they can see if they like and/or prefer this method of searching the patient's record.

Re-iterate scenario if needed
One of your colleagues has referred one of their patients to you, as they are leaving on a month-long vacation. The patient in question has come in with excruciating pain in her lower back and is having trouble sitting down. This is the first time you see this patient and you need to understand their medical history.

Task: This time we will use the tag cloud to search the patient's record. Please walk me through your thought process and your reasoning while using the tag cloud to search the patient's record.

Observed Results

Completed	**Completed with Assist**		**Not completed**
	General Assist		
	Specific Assist		
	# of Assists		

Q B1
The Tag Cloud capabilities meet my requirements.

	1- Strongly Disagree
	2
	3
	4
	5
	6
	7 – Strongly Agree

Q B2
Using the Tag Cloud is a frustrating experience.

	1- Strongly Disagree
	2
	3
	4
	5
	6
	7 – Strongly Agree

Q B3
The Tag Cloud is easy to use.

	1- Strongly Disagree
	2
	3
	4
	5
	6
	7 – Strongly Agree

Q B4
I have to spend too much time correcting things with the Tag Cloud.

	1- Strongly Disagree
	2
	3
	4
	5
	6
	7 – Strongly Agree

Comments

Task C - SCROLL SEARCH (i.e. eye-balling patient record)

Goal: Make the user use the Scroll Search to search the patient's record for relevant information, so that they can see if they like and/or prefer this method of searching the patient's record.

Re-iterate scenario if needed
One of your colleagues has referred one of their patients to you, as they are leaving on a month-long vacation. The patient in question has come in with excruciating pain in her lower back and is having trouble sitting down. This is the first time you see this patient and you need to understand their medical history.

Task: This time we will scroll the patient's record to search. Please walk me through your thought process and your reasoning while scrolling to search the patient's record.

Observed Results

Completed	**Completed with Assist**		**Not completed**
	General Assist		
	Specific Assist		
	# of Assists		

Q C1
The Scroll Bar capabilities meet my requirements.

	1- Strongly Disagree
	2
	3
	4
	5
	6
	7 – Strongly Agree

Q C2
Using the Scroll Bar is a frustrating experience.

	1- Strongly Disagree
	2
	3
	4
	5
	6
	7 – Strongly Agree

Q C3
The Scroll Bar is easy to use.

	1- Strongly Disagree
	2
	3
	4
	5
	6
	7 – Strongly Agree

Q C4
I have to spend too much time correcting things with the Scroll Bar.

	1- Strongly Disagree
	2
	3
	4
	5
	6
	7 – Strongly Agree

Comments

Post-Questionnaire

Now to complete the interview, I have a few additional questions.

Post-Q 1
I think that I would like to use this system frequently

	1- Strongly Disagree
	2
	3
	4
	5 – Strongly Agree

Post-Q 2
I found the system unnecessarily complex

	1- Strongly Disagree
	2
	3
	4
	5 – Strongly Agree

Post-Q 3
I thought the system was easy to use

	1- Strongly Disagree
	2
	3
	4
	5 – Strongly Agree

Post-Q 4
I think that I would need the support of a technical person to be able to use this system

	1- Strongly Disagree
	2
	3
	4
	5 – Strongly Agree

Post-Q 5
I found the various functions in this system were well integrated

	1- Strongly Disagree
	2
	3
	4
	5 – Strongly Agree

Post-Q 6
I thought there was too much inconsistency in this system

	1- Strongly Disagree
	2
	3
	4
	5 – Strongly Agree

Post-Q 7
I would imagine that most people would learn to use this system very quickly

	1- Strongly Disagree
	2
	3
	4
	5 – Strongly Agree

Post-Q 8
I found the system very cumbersome to use

	1- Strongly Disagree
	2
	3
	4
	5 – Strongly Agree

Post-Q 9
I felt very confident using the system

	1- Strongly Disagree
	2
	3
	4
	5 – Strongly Agree

Post-Q 10
I needed to learn a lot of things before I could get going with this system

	1- Strongly Disagree
	2
	3
	4
	5 – Strongly Agree

Debrief

Thank you for participating in this study.

Your **participation and feedback is greatly appreciated** and **very valuable** to the goals of this study.

I'd like to remind you that the **contents of this study are confidential** and that you not discuss anything that you saw here today.

As a token of my appreciation, please accept this **$20 Starbucks gift card**.

Hand over envelope with gift card.

Please **verify** that the **contents of the envelope** are correct and if you can then **sign** the following **form acknowledging** that you have **received the gift card**.
Hand over Receipt of Compensation form for signature.

We have come to the end of the interview.
Do you have **any questions before we leave**?

Thank you, once again, for your participation!

B. USABILITY METRIC FOR USER EXPERIENCE (UMUX)

1. [This system's] capabilities meet my requirements.

Strongly Disagree	1 2 3 4 5 6 7	Strongly Agree

2. Using [this system] is a frustrating experience.

Strongly Disagree	1 2 3 4 5 6 7	Strongly Agree

3. [This system] is easy to use.

Strongly Disagree	1 2 3 4 5 6 7	Strongly Agree

4. I have to spend too much time correcting things with [this system].

Strongly Disagree	1 2 3 4 5 6 7	Strongly Agree

Once data are collected, they need to be properly recoded, with a method that borrows from the SUS.

Odd items are scored as [response – 1]
Even items are scored as [7 – response]

As with the SUS, this removes the positive/negative keying of the items and allows a minimum score of zero.

Each individual UMUX item has a range of 0 – 6 after recoding, giving the entire four-item scale a preliminary maximum of 24.

To achieve parity with the 0–100 range provided by the SUS, a participant's UMUX score is the sum of the four items divided by 24, and then multiplied by 100.

These scores across participants are then averaged to find a mean UMUX score.
It is this mean score and its confidence interval that become the application's UMUX metrics for a system's usability tracking and goal-setting.

C. SYSTEM USABILITY SCALE (SUS)

1. I think that I would like to use this system frequently

Strongly Disagree 1 2 3 4 5 Strongly Agree

2. I found the system unnecessarily complex

Strongly Disagree 1 2 3 4 5 Strongly Agree

3. I thought the system was easy to use

Strongly Disagree 1 2 3 4 5 Strongly Agree

4. I think that I would need the support of a technical person to be able to use this system

Strongly Disagree 1 2 3 4 5 Strongly Agree

5. I found the various functions in this system were well integrated

Strongly Disagree 1 2 3 4 5 Strongly Agree

6. I thought there was too much inconsistency in this system

Strongly Disagree 1 2 3 4 5 Strongly Agree

7. I would imagine that most people would learn to use this system very quickly

Strongly Disagree 1 2 3 4 5 Strongly Agree

8. I found the system very cumbersome to use

Strongly Disagree 1 2 3 4 5 Strongly Agree

9. I felt very confident using the system

Strongly Disagree 1 2 3 4 5 Strongly Agree

10. I needed to learn a lot of things before I could get going with this system

Strongly Disagree 1 2 3 4 5 Strongly Agree

The SUS scale is generally used after the respondent has had an opportunity to use the system being evaluated, but before any debriefing or discussion takes place.

Respondents should be asked to record their immediate response to each item, rather than thinking about items for a long time.

All items should be checked. If a respondent feels that they cannot respond to a particular item, they should mark the centre point of the scale.

Scoring SUS
SUS yields a single number representing a composite measure of the overall usability of the system being studied.
Note that scores for individual items are not meaningful on their own.

To calculate the SUS score, first sum the score contributions from each item.
Each item's score contribution will range from 0 to 4.
For items 1,3,5,7,and 9 the score contribution is the scale position minus 1.
For items 2,4,6,8 and 10, the contribution is 5 minus the scale position.
Multiply the sum of the scores by 2.5 to obtain the overall value of SU.
SUS scores have a range of 0 to 100.

D. USABILITY REPORT

USABILITY REPORT

Information Access by Naïve Users

Introduction

Summary

Usability testing for the Tag Cloud Tool Prototype was conducted from November 2013 through May 2014. The testing was conducted at a mutually agreed upon location with each participant. Participants were recruited through physiotherapy/osteopathy clinics, as well as through acquaintances of friends and family. The participant profile is provided in section 2.1.1. Each session lasted between 20 and 45 minutes and included an introduction, completion of three (3) tasks and associated questions, and a final debrief questionnaire.

The majority of participants completed the session without difficulty. Two participants had some difficulty with the technology, owing to not being familiar with it, but were still able to complete the session.

- Overall, 83% (10/12 participants) had a very positive reaction to the tag cloud tool.
- The tag cloud was preferred, with a mean score of 81%, followed by the search box with 71%, then scrolling with 68%.
- The system as a whole, that is the prototype with all search methods, had a mean SUS score of 80, which places it at the 90^{th} percentile.
- All but two (2) participants required specific help with some of the tasks.
- The following are issues that were noted by the majority of participants:
    - As there are multiple types of therapists entering data into the patient's file, there needs to be a standardized set of terminology to ensure that all of the pertinent data can be retrieved when searching.
- Participants provided numerous suggestions for improving the prototype. We also observed some elements, which could be improved. These are listed in the "Recommendations" section at the end.

Goals of the Test

The goal of the testing described in this document was too see if using a tool that allows to see at a glance the main items in a patient's history of visits would be useful to therapists seeing a patient for the first time, either as a referral or taking over the patient's file. The goal is to prove, or disprove, that this tool is more useful than other methods of searching for the same information in a patient's file.

Participant Profiles

- Participants were evenly distributed along gender.
    - 7 males (50%)
    - 7 females (50%)
- The main occupation of participants were as follows:
    - 7 participants were Physiotherapists (58.33%)
    - 3 were Osteopaths (25%)
    - 1 was an Athletic Therapist (8.33%)
    - 1 was a Chiropractor (8.33%)

- Participants' ages varied, as follows:
  - 1 participant was between18-24 (8.33%)
  - 2 were between 25-29 (16.66%)
  - 4 were between 30-34 (33.33%)
  - 3 were between 35-39 (25%)
  - 1 was between 40-44 (8.33%)
  - 1 was between 55-59 (8.33%)
- Participants were recruited either through the clinics where the researcher is a patient or through friends, family, and colleagues with acquaintances working in the aforementioned areas of occupation.

Session Schedule and Tasks Performed by Participants

Each session with a participant followed this approximate schedule of events:

- Participants were contacted and given a brief description of the research and asked if they would agree to participate. If they agreed, a testing session was scheduled.
- Participants were greeted and provided with a consent form, which included and asking if they agreed to audio taping of the session and non disclosure of confidential information on the prototype.
- Participants were asked 3 demographic questions prior to starting the testing.
- An introduction to the goal of the testing, prototype, and testing procedure was given.
- Participants were then asked to complete the tasks. A scenario was read to the participant and they could ask questions before proceeding.
- After completing each task, participants were asked a set of four (4) Likert-scale questions. Participants were also asked if they had any comments they would like to provide pertaining to that task.
- After completing all the tasks, participants were debriefed to get their general impressions of the prototype.
- Due to problems with the recording device, one participant's session was not recorded.

Participants were informed that for the purposes of this test that the tester was playing the role of the patient that they were to be treating in this session. Participants were instructed that if they had questions they would ask the patient as part of a session, that they could, and should, ask them to ensure that the information retrieved was as realistic as possible. Participants were also encouraged to talk aloud during the completion of each task, so that the tester could gather information on possible problems, improvements, or elements that should be kept in the prototype. All participants completed the three (3) tasks in the same order to ensure that the results were being compared based on the same criteria.

Task 1

One of your colleagues has referred one of their patients to you, as they are leaving on a month-long vacation. The patient in question has come in with excruciating pain in her lower back and is having trouble sitting down. This is the first time you see this patient and you need to understand their medical history.

To understand the patient's history, you need to find all relevant information within their record. We will be using the prototype that is on the computer in front of you. This time we will use a specific method, the search box, to search the patient's record. Please walk me through your thought process and your reasoning while using the search box to search the patient's record.

Task 2
The scenario is the same as on the last task:
One of your colleagues has referred one of their patients to you, as they are leaving on a month-long vacation. The patient in question has come in with excruciating pain in her lower back and is having trouble sitting down. This is the first time you see this patient and you need to understand their medical history.

For this task, we will use the tag cloud to search the patient's record. Please walk me through your thought process and your reasoning while using the tag cloud to search the patient's record.

Task 3
Once again, the scenario is the same as on the last two tasks:
One of your colleagues has referred one of their patients to you, as they are leaving on a month-long vacation. The patient in question has come in with excruciating pain in her lower back and is having trouble sitting down. This is the first time you see this patient and you need to understand their medical history.

For this third, and final task, we will scroll the patient's record to search. Please walk me through your thought process and your reasoning while scrolling to search the patient's record.

Confounds

- One (1) of the participant sessions was not recorded due to the recording device malfunctioning during that particular testing session. All precautions were taken thereafter to ensure that such a situation did not occur again.
- Some participants were not used to working on a Mac. This was covered at the beginning of the session and explanations were given to these participants prior to starting the first task.
- The following disclaimers were provided to each participant prior to starting the tasks:
    - Please keep in mind that this is a prototype, so not all functionality is present.
    - As we tried to make things as realistic as possible, as in some of the entries being entered by a third person who is not the therapist (example: administrative assistant), you may see some entries that have been left blank for later entry or there may be typos in cases where the person entering the notes was unable to properly decipher the writing and did not get a chance to have it clarified.

Results

Summary

During the testing sessions, most (8/12) participants asked questions they would have asked the patient in a real life situation. This helped make the testing closer to a real-life scenario.

Application Ratings

While completing each task, participants were asked to talk out loud about what they were searching for and how they were trying to search for this information while completing the task. Once each task was completed, participants were asked to answer four Likert-scale questions on the interface used to complete the task. Participants were also asked to provide any additional comments they had above and beyond those they mentioned while talking aloud during the task completion.

Rating Scheme

Task Questions - Usability Metric for User Experience (UMUX)
After completing each of the three tasks, participants were asked the following four (4) questions from the UMUX:

1. [This system's] capabilities meet my requirements.
2. Using [this system] is a frustrating experience.
3. [This system] is easy to use.
4. I have to spend too much time correcting things with [this system].

Where [this system] was replaced with the interface being tested in the task; Search Box, Tag Cloud, or Scrolling. The questions used a Likert-scale response with a range of 1 through 7, where 1 was Strongly Disagree and 7 was Strongly Agree. Once data are collected, they need to be properly recoded, with the following scoring:

Odd items are scored as [response – 1]
Even items are scored as [7 – response].

This removes the positive/negative keying of the items and allows a minimum score of zero.

Each individual UMUX item has a range of 0 – 6 after recoding, giving the entire four-item scale a preliminary maximum of 24.

To get a value on a scale of 100, a participant's UMUX score is the sum of the four items divided by 24, and then multiplied by 100.

These scores across participants are then averaged to find a mean UMUX score. It is this mean score and its confidence interval that become the task's UMUX metric.

Final Debrief - System Usability Scale (SUS)

After completing all three tasks, participants were asked the following questions from the System Usability Scale (SUS) as a debrief questionnaire:

1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

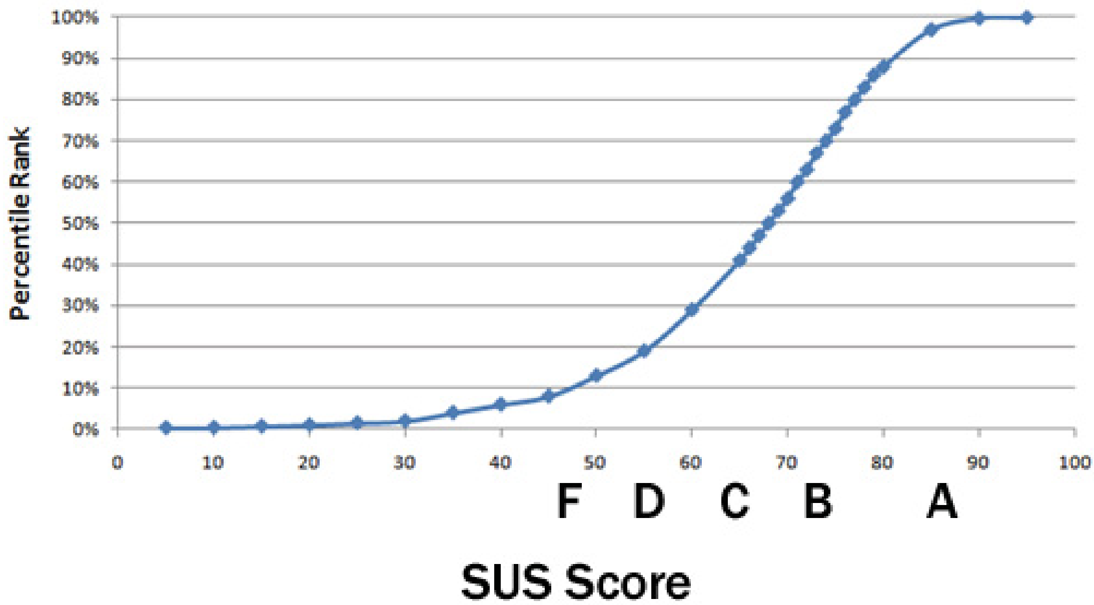
The questions used a Likert-scale response with a range of 1 through 5, where 1 was Strongly Disagree and 5 was Strongly Agree. The Likert-scale for the final debrief is on a scale of 5, as this questionnaire was done separately and its results not compared to the individual task results. Once data are collected, they need to be properly recoded, with the following scoring:

Odd items are scored as [response – 1]
Even items are scored as [5 – response].

As with the UMUX questions, we recode the positive/negative keying of the items and allows a minimum score of zero.

Each individual UMUX item has a range of 0 – 4 after recoding. To calculate the SUS score, first sum the score contributions from each item. Then multiply the sum of the scores by 2.5 to obtain the overall value of SUS; the range will be from 0-100.

The SUS provides a single number representing a composite measure of the overall usability of the system being tested. Although based on 100, this is not a percentage and the resulting value must be turned into its percentile value. This is done using the below s-curve. It is to be noted that an average SUS score is 68, which translate into a 50^{th} percentile value.



Summary of Results

Task Results

1. Mean scores for each task across all participants showed that the Tag Cloud was the preferred method of searching the patient's file.
   - Search Box mean 71%
   - Tag Cloud mean 81%
   - Scrolling mean 68%

2. Recalculating the mean by removing the outliers (lowest and highest score for each task) also resulted in the Tag Cloud being the preferred search method.
   - Search Box mean 73%
   - Tag Cloud mean 85%
   - Scrolling mean 69%

3. The 95% t-distribution confidence intervals for the three tasks were wide, but remained in line with the searching task preferences seen with the means.
   - Search Box confidence interval +/- 15.39
   - Tag Cloud confidence interval +/- 14.70
   - Scrolling confidence interval +/- 17.66

	Upper CI	MEAN	Lower CI
Search Box	86	71	56
Tag Cloud	96	81	66
Scrolling	86	68	50

4. An interesting observation was that when the 95% t-distribution confidence intervals were calculated for the mean scores with the outliers removed, the confidence interval for the Tag Cloud became much tighter, whereas the confidence intervals for the other two search methods remained almost identical.
    - Search Box confidence interval +/- 15.62 (versus +/- 15.39)
    - Tag Cloud confidence interval +/- 8.91 (versus +/- 14.70)
    - Scrolling confidence interval +/- 18.21 (versus +/- 17.66)

	Upper CI	MEAN	Lower CI
Search Box	89	73	57
Tag Cloud	94	85	76
Scrolling	87	69	51

5. Mean times for each of the three tasks showed interesting results. Each task's mean time diminished from one task to the other. Although the hope was to see a clear indication of one task taking less time to complete than the other two, the results showed that the first task took the most time (07:25), the last task took the least amount of time (03:39), and the second task took approximately the average time (05:09) of the first and third tasks combined.
    - A possible explanation for this would be that as participants perform a task, they are at the same time becoming familiarized with the content of the patient's file. Based on this assumption, we cannot infer from the task times which task is preferred or most efficient.
    - Another explanation could be that, as mentioned in section 2.1, 8/12 participants had questions throughout each task related to treating the patient. Answers received when completing one task, would mean not needing to ask this question again on the subsequent tasks. This could also explain why task times may not be an accurate reflection of which task tool is preferred.

Overall System Results

The SUS questionnaire was used a final debrief for the prototype as a whole. The SUS mean score for the overall system evaluation was 80, putting it in the 90th percentile. An average SUS score is 68, which is equivalent to the 50th percentile.

Detailed Results

The tables below give the participants' responses to the UMUX questionnaire for each task. We start with a comparison table and then break down the results per task. After each task table, a more detailed breakdown of their responses, as well as any comments they had while completing the tasks, including observations about their experience, are described.

Comparison of Task Results

Task Comparison	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	UMUX Mean	Std Dev	CI
Search Box	33.33	95.83	54.17	95.83	50.00	95.83	75.00	50.00	45.83	95.83	79.17	83.33	71	23.20	15.39
Tag Cloud	87.50	20.83	83.33	87.50	79.17	91.67	91.67	54.17	87.50	100.00	95.83	91.67	81	22.15	14.70
Scrolling	100.00	58.33	29.17	87.50	100.00	54.17	37.50	100.00	41.67	46.00	87.50	75.00	68	26.61	17.66

Search Box

Search Box		P1		P2		P3		P4		P5		P6		P7		P8		P9		P10		P11		P12					
Q1	[score – 1]	3	2	6	5	5	4	6	5	3	2	7	6	6	5	5	4	3	2	6	5	6	5	5	4				
Q2	[7 – score]	5	2	1	6	4	3	1	6	6	1	2	5	3	4	4	3	7	0	1	6	4	3	1	6	Raw			
Q3	[score – 1]	5	4	7	6	5	4	7	6	4	3	7	6	6	5	4	3	4	3	7	6	7	6	7	6	Score	UMUX	Std	
Q4	[7 – score]	7	0	1	6	5	2	1	6	1	6	1	6	3	4	5	2	1	6	1	6	2	5	3	4	Mean	Mean	Dev	CI
	Total		8		23		13		23		12		23		18		12		11		23		19		20	17			
	UMUX		33.33		95.83		54.17		95.83		50		95.83		75		50		45.83		95.83		79.17		83.33		71	23.20	15.39

The following participant comments and observations made during testing of the Search Box task.

- ➢ Two (2) participants specifically mentioned that using the Search Box frustrated them.
- ➢ Three (3) participants commented on not knowing what terminology was used in the file, so searching became difficult due to this.
    - One (1) participant mentioned seeing the words in the tag cloud (before knowing what the tag cloud was or performing the task for the tag cloud), which helped them with using the search box.
- ➢ Two (2) participants indicated that they would not use the Search Box, as it is not something that they would normally use to search through a patient record.
- ➢ One (1) participant noted that using the Search Box did not help in finding the most recent visit/treatment the patient had received.

Participant Quotes

- *"I'm not a huge fan of the search box."*
- *"I'm not a fan of search boxes."*
- *"Trying to find last treatment with search not easy."*
- *"[...] don't know if search is complete using search because don't know if is everything in file."*

Tag Cloud

Tag Cloud		P1		P2		P3		P4		P5		P6		P7		P8		P9		P10		P11		P12					
Q1	[score – 1]	5	4	2	1	6	5	5	4	5	4	6	5	7	6	2	1	5	4	7	6	7	6	5	4				
Q2	[7 – score]	2	5	5	2	2	5	2	5	3	4	1	6	2	5	7	0	2	5	1	6	2	5	1	6	Raw			
Q3	[score – 1]	7	6	2	1	6	5	7	6	6	5	6	5	6	5	7	6	7	6	7	6	7	6	7	6	Score	UMUX	Std	
Q4	[7 – score]	1	6	6	1	2	5	1	6	1	6	1	6	1	6	1	6	1	6	1	6	1	6	1	6	Mean	Mean	Dev	CI
	Total		21		5		20		21		19		22		22		13		21		24		23		22	19			
	UMUX		87.50		20.83		83.33		87.50		79.17		91.67		91.67		54.17		87.50		100		95.83		91.67		81	22.15	14.70

The following are participant comments and observations made during testing of the Tag Cloud task.

- ➢ Four (4) participants would like to see multiple tag clouds, such as per category of therapy, treatment, past medical history, imaging (i.e. MRI scans, X-rays, etc.), test results.
- ➢ Five (5) of the participants made reference to the fact that they found it easier and faster to use the tag cloud to pull up past visits because they could easily spot the keywords/key terms that appeared in the patient's file most often.
- ➢ Ten (10) participants had highly positive reactions to the tag cloud task.
- ➢ All twelve (12) participants had the most comments while performing the tag cloud task than the other two tasks (search box, scrolling).

Participant Quotes

- *"I like the idea that it gets bigger when it's recognized more often. I think that's very valuable. I mean it's easy; it's the first thing you see. So I think it's a very quick tool to see if what the person comes in presenting with, especially since you've never seen this person before, that it's a very quick thing. That, 'oh ya this is either something really new or this is a possible exacerbation of something they've had in the past' and that's nice to know. It's valuable for the patient."*
- *"Ok, so this kind of just gives you a synopsis of what's like reoccurring kind of dysfunction with the person."*
- *"Like the way it looks, really nice."*
- *"Good for person in rush to find information."*
- *"Like that is a synopsis of what the patient has had."*
- *"[I] never used [a tag cloud] before, I like this."*
- *"I like that! I like that!"*
- *"I much prefer this as opposed to putting it in a search box, simply because I can just click on it instead of typing."*
- *"This is cool."*
- *"What I like more about the cloud, is that if you're looking for something in particular, you can find it right away."*
- *"This tag cloud is nice."*

Scrolling

Scrolling		P1		P2		P3		P4		P5		P6		P7		P8		P9		P10		P11		P12		Raw Score Mean	UMUX Mean	Std Dev	CI
Q1	[score – 1]	7	6	4	3	2	1	6	5	7	6	2	1	6	5	7	6	5	4	5	4	5	4	5	4				
Q2	[7 – score]	1	6	2	5	6	1	1	6	1	6	7	0	7	0	1	6	5	2	6	1	2	5	3	4				
Q3	[score – 1]	7	6	5	4	3	2	5	4	7	6	7	6	2	1	7	6	4	3	7	6	7	6	6	5				
Q4	[7 – score]	1	6	5	2	4	3	1	6	1	6	1	6	4	3	1	6	6	1	7	0	1	6	2	5				
	Total		24		14		7		21		24		13		9		24		10		11		21		18	16			
	UMUX		100		58.33		29.17		87.50		100		54.17		37.5		100		41.67		46		87.50		75.00		68	26.61	17.66

The following are participant comments and observations made during testing of the Scrolling task.

- Seven (7) participants stated that scrolling through the whole record would be time consuming, especially if the record contained lots of information / visits.
  - One (1) participant specifically mentioned that the computers at their clinic were old and slow, meaning that trying to scroll through the whole file would take even more time.
- Three (3) participants quickly scrolled the record down to the bottom and right back up without taking much time, if any, to read the content in the file before stating that they were ready for the questionnaire.
- Two (2) participants mentioned that starting with the scrolling task would have been preferred, as they would have had a better idea of what terms to use in the search box task.
  - It is to be noted that the order of the tasks was specifically chosen with the scrolling to be the last task to ensure that this particular familiarity with the patient file was not there when performing the other two tasks.
- Two (2) participants commented specifically with reference to the tag cloud while performing the scrolling task.

Participant Quotes

- *"I like this one for its ease of use, for the first few. If the person's been in 60 times to the clinic, it's a disadvantage. It's just nice to see past history."*
- *"Love how it reads, it's fine, very clear, it's just I have to scroll."*
- *"Wouldn't really feel the need to look that far back."*
- *"Where could see being useful [...] comments [...] kind of neat to see that way."*
- *"Would like to go to bottom with button or button to reverse order of entries."*
- *"Nice to have it all laid out there to see history and progression"*

SUS (Debrief) Questionnaire

The table below gives the participants' responses to the SUS questionnaire used as a debrief questionnaire after completing the three tasks. As with the above tables, a more detailed breakdown of the participants' responses, as well as any comments they had while completing the debrief questionnaire, including observations about their experience, are described.

Debrief		P1		P2		P3		P4		P5		P6		P7		P8		P9		P10		P11		P12				
Q1	[score – 1]	3	2	5	4	4	3	2	1	3	2	3	2	4	3	2	1	4	3	5	4	4	3	5	4			
Q2	[5 – score]	1	4	2	3	2	3	2	3	1	4	1	4	2	3	4	1	1	4	1	4	1	4	1	4			
Q3	[score – 1]	5	4	3	2	5	4	4	3	4	3	5	4	5	4	3	2	4	3	5	4	4	3	5	4			
Q4	[5 – score]	1	4	1	4	1	4	2	3	1	4	2	3	1	4	1	4	1	4	1	4	2	3	1	4			
Q5	[score – 1]	5	4	2	1	5	4	4	3	5	4	5	4	3	2	4	3	5	4	5	4	5	4	4	3			
Q6	[5 – score]	3	2	2	3	2	3	2	3	1	4	1	4	3	2	3	2	1	4	1	4	1	4	1	4			
Q7	[score – 1]	5	4	4	3	4	3	3	2	5	4	5	4	4	3	2	1	4	3	4	3	5	4	5	4			
Q8	[5 – score]	3	2	1	4	2	3	3	2	1	4	1	4	2	3	4	1	3	2	1	4	1	4	1	4	Raw		
Q9	[score – 1]	5	4	3	2	4	3	3	2	4	3	4	3	5	4	2	1	4	3	5	4	4	3	5	4	Score	SUS	
Q10	[5 – score]	1	4	2	3	2	3	2	3	1	4	2	3	2	3	4	1	2	3	1	4	2	3	1	4	Mean	Mean	Percentile
	Total		34		29		33		25		36		35		31		17		33		39		35		39	32		
	SUS		85		72.5		82.5		62.5		90		87.5		77.5		42.5		82.5		97.5		87.5		97.5		80	90th

The following are participant comments and observations made during the debriefing questionnaire.

- Three (3) participants had no comments to add from what they had already provided on the individual task.
- Five (5) participants gave enthusiastic comments on the system and would like to have such a system in place and at their disposal.
- One (1) participant commented on Question 1 (*I think that I would like to use this system frequently*) prior to answering that they are used to quickly flipping through the paper chart currently used, but could see that a system like the prototype would be good for large files, as it is more organized.

Participant Quotes

- *"System has great potential."*
- *"Great that could have system with all in one place, especially with multi-disciplinary therapists."*
- *"Great potential for tool. Really great, since patients don't remember things accurately."*
- *"Super user friendly."*
- *"Neat, like tag cloud."*
- *"Like that have alternative ways to find information."*
- *"Appeals to different ways of working of different therapists."*

Observations

- Most participants (9/12) commented on a need for common terminology to be used in the system to ensure that when searching, everyone knows what terms to search on.
- There was a general frustration among most participants with the search box (i.e. spelling term differences).
- Many participants had multiple questions to ask the patient as they were searching for relevant information in the file. This was useful for completing the tasks, but did increase testing time for 3/12 participants.
- A few of the participants remarked that they were uncertain if they had found all the information they needed with the search box or if there was more information to be found.
- A strong majority of participants (8/12) were very enthusiastic when performing the tag cloud task.
- All participants gave suggestions on improving the prototype to ensure that it was efficient in retrieving information.
- All participants had an easy time using the tag cloud.
  - Only three (3) participants needed an explanation on what a tag cloud was and how it worked.
  - One (1) participant did not know what a tag cloud was, but instinctively knew how to use it.
- Five (5) participants commented that scrolling was not the most efficient when there were a large amount of entries in a patient's file. They all stated, in some form, that they would not scroll very far in cases with many entries.
- Two (2) participants referred back to the tag cloud while performing the scrolling task, giving positive feedback on using the tag cloud.
- The two (2) oldest participants seemed less comfortable using the prototype and had some minor difficulties in performing the tasks. They required some specific assistance, but this did not impede on the task completion.

Recommendations

This section lists the suggestions offered by participants when either completing a task or commenting on their answers to the questions.

Search Box

- Predictive typing to help with spelling and terms in the patient file
- Indicator in search results if it is the wrong word that was typed in (i.e. "did you mean '[term]'?") or if the word just does not exist in the file (i.e. "there are no entries for the word '[term]'").
- Highlighting the search term in the search results.
- Results returning only the entries' sub-section that referenced the term searched.

Tag Cloud

- One tag cloud per category (example: Medical History, Precautions, etc.).
- Having a tag that says "past history" or "past medical history" or "current problem" would be of great benefit.

- Having tag clouds per therapist type (example: Physiotherapy, Osteopathy, Athletic Therapy, etc.) to find out what was done in a specific type of therapy by other therapists of that type.
- Make sure keywords used to tag are the same by all therapists.
    - One (1) participant suggested having a drop-down box with the terms that can be selected.
- Highlighting the search term in the search results. (Note: this was also suggested in the Search Box task.)

Scrolling

- Highlight each therapist's name a specific color in the records, so that can quickly see in each entry which therapist looked after the patient on that particular visit.
- Having a quick link to items.
- Adding a mechanism to reverse the order of the entries (i.e. chronological versus reverse chronological).
- Having an expand/collapse for each entry in the file, so that scrolling would take less time and could only have to read the entries that wanted to read.

System as a Whole

- Ensure that all therapists use the same terminology, especially for tags.
- Remove all non-relevant content, such as the header in the file, so that more of the patient's record is visible on the page.

Recommendations for Future Enhancements

1. Ensure that there is no shorthand in the entries.
    a. Write out full term.
    b. Use both full term and shorthand as tags to enhance search results.
2. Have multiple tag clouds per category.
    a. Diagnosis
    b. Treatment
    c. Medical History
    d. Therapist Type
3. Enhance the search box functionality.
    a. Predictive typing.
    b. Highlight search term in search results.
4. Add functionality to the scrolling.
    a. Reverse order of entries (i.e. chronological , reverse chronological).
    b. Make entries collapsible / expandable.

Location of Audio Recordings

The audio recordings are stored on the researcher’s mobile phone. at Acme headquarters in Montreal, in the Usability Test room. One of the sessions was not recorded due to the recording device malfunctioning.