Visualizing Semantic Structure of a Clinical Text Document

Jonah Kenei, Elisha Opiyo, and Robert Oboko

Abstract — The increasing use of Electronic Health Records (EHRs) in healthcare delivery settings has led to increase availability of electronic clinical data. They generate a lot of patients' clinical data each day, requiring physicians to review them to find clinically relevant information of different patients during care episodes. The availability of electronically collected healthcare data has created the need of computational tools to analyze them. One of the types of data which doctors have access to is clinical notes that resides in electronic health records. These notes are useful as they provide comprehensive information about patients' health histories with many practical uses. For example, doctors always review these notes during care episodes to appraise themselves about the health history of a patient. These reviews are currently manual where a doctor reads a patient's chart while looking for specific clinical information. Without the proper support, this manual process leads to information overload and increases physician cognitive workload. Current electronic health records (EHRs) do not provide support to help physicians reduce cognitive workload when completing clinical tasks. This is especially true for long clinical documents which require quick review at the point of care. The growing amount of clinical documentation available in EHRs has arose the need of tools that support synthesize of information in EHRs. The use of visual analytics to explore healthcare data is one such research direction to address this problem. However, existing visualization techniques are mainly based on structured electronic health record and rarely support therapeutic activities. Therefore, visualization of unstructured clinical records to support clinical practice is required. In this paper we propose a unique approach for graphically representing and visualizing the semantic structure of a clinical text document to aid doctors in reviewing electronic clinical notes. A user evaluation demonstrates that the proposed method for visualizing and navigating a document's semantic structure facilitates a user's document information exploration.

Index Terms — Semantics, visualization; Electronic health records.

I. INTRODUCTION AND MOTIVATION

Electronic health records (EHRs) are now widely used in many healthcare facilities which have resulted in unprecedented availability of longitudinal patients' clinical data [1]. They have enhanced the ability to collect a large amount of patient-specific health information over long periods of time. This information plays a critical role in providing the best evidence for decision making for any given patient. However, the availability of more information in

Published on December 4, 2020.
Jonah Kenei, School of Computing and Informatics, University of Nairobi, Kenya.
(e-mail: Jonah.kenei@gmail.com)
Elisha Opiyo, School of Computing and Informatics, University of Nairobi, Kenya.
(e-mail: opiyo@uonbi.ac.ke)
Robert Oboko, School of Computing and Informatics, University of Nairobi, Kenya.
(e-mail: roboko@uonbi.ac.ke)

EHRs is only useful when information contained in them can be properly extracted and understood [2].With increasingly available clinical documentation, physicians are presented with more and more patient information in EHRs. However, one major challenge is how to summarize and present this information at the point of care [3] during care episodes. This is an instance of a general problem that confronts computing where we are being confronted with increasing size of datasets and the challenge of processing these datasets [4]. Therefore, for the benefits of EHRs to have a positive impact, physicians must be able to use the available data effectively as they are readily available. Diagnosis of a medical condition and treatment is largely dependent on being able to efficiently and effectively identify the critical medical information in such documents.

Clinical documents usually contain information often needed by physicians in the course of patient care and this vital clinical information is recorded using a natural language such as English. The unstructured nature of these documents limits their use in decision-making [5] because it is not structured via pre-defined data models or schema. Therefore it is difficult to search for information. In contrast, structured data usually have clearly defined data types and pattern which makes it easily searchable [6]. Although, unstructured texts store a lot of valuable medical information it lacks common structure which makes it hard for analysis [7].Therefore, clinical notes are currently underused compared to structured data, due to their high-dimensional and sparse nature [8].

Unstructured data is often easier to read by humans, but it is much difficult for computers [9]. Even though human can easily read unstructured text, the amount of text can sometimes be overwhelming for physicians to read manually [9] due to information overload problem [10] and cognitive overload [11]. Browsing lengthy text documents is usually a very time-consuming task and this is especially true with long clinical records. Finding key information in this case often requires one to manually read through the text. Research has shown that, current electronic health records (EHRs) do not offer support to help physicians reduce their cognitive workload with little support for recognizing clinically relevant patterns in patients' clinical records [12]. They are yet to provide support to physicians in their patient care activities. The amount of narrative text available can sometimes outweigh the time one have to actually read where users must look for certain information from growing volumes of information. In healthcare, there is a continuous need for clinicians to review multiple EHR clinical documents during a typical out-patient visit [13].

Therefore, the problems of using this data have emerged as an active area in clinical informatics research [11]. The medical record chronologically documents the care of the patient and is an important element contributing to high

quality care. It is well known that, EHRs do not support the cognitive processes of physicians [14]. Visualizing a document's content has been proven to be useful in aiding information retrieval and reducing cognitive load associated with reviewing large amounts of electronic clinical documents within time-constrained clinical encounters [15]. The expectation is to enhance and promote better clinical decision making [15] and help reduce the time required to analyze and understand clinical data [18]. However, existing visualization techniques are mainly based on structured electronic health record [16] and are rarely used to support diagnostic and therapeutic activities [17]. For clinical visualizations, much of the needed information is usually in unstructured notes rather than structured data fields. Therefore, visualization based on unstructured patient records to support clinical practice is required.

The classic examples are tag clouds, often used as a visual representation of text data [17] and timelines [18]. Timeline is the most widely used approach which is designed to retrieve and present clinical data in a hierarchical and timeline based structure where clinicians can browse chronologically through existing data [18].It emphasizes the significance of time in the clinical data and tries to organize and display clinical data in an interactive timeline using visual enhancements to facilitate readability [19]. Timeline structure, however, is only one of the ways in which people can naturally conceptualize a document.

In a typical clinical narrative text, concepts are usually semantic units such as symptoms, diagnosis, and treatment etc.

Therefore, another way in which people can conceptualize a document is in terms of its semantic structure. In talking about patient clinical histories, physicians often make references such as "symptoms" or "the diagnosis." These are references to a document's semantic structure. By glancing through a visual representation of these semantic units, a person can be able to quickly glean the overall semantic structure of a document, allowing one to quickly access a section of interest. This paper introduces an approach for visualizing a document's semantic structure using a cluster map.

First, we describe how the semantic structure of a clinical document is produced through analysis of documentation formats. Second, we explain the choice of semantic classes as the basis of our classification. Third, we describe how the cluster map is used to organize the semantic content of a document. Fourth, we discuss a user evaluation. We conclude with a summary of contribution and directions for future work and help to represent this information in a human and usable form. The tool comprises two modules, a classification component and a visualization component for the graphical presentation of clinical information. The paper introduces an approach that organizes classified sentences semantically and displays them spatially for reading.

II. BACKGROUND

In recent years, electronic health records have gain wide acceptance in many healthcare delivery settings [20]-[22] and their use is now a standard part of our contemporary medical practice [23], [24]. Physicians therefore routinely collect and document care of their patients using electronic health records

and thus generating digital health records of patients during one or more clinical encounters in any care delivery setting [25]. Medical practitioners are increasingly using EHRs to support decision-making in patient treatments however, this is hampered by the overwhelming amount of information available, and insufficient time and tools on the clinician's part to locate and synthesize the available information. EHRs were designed with the aim of improving healthcare, like clinical decision-making and information exchange [26]-[28]. One of its distinct advantages is making patients' medical history available to physicians, which in paper records took considerably more time to obtain [29].

They provide snapshots of a patient's clinical history which have created unprecedented opportunities to investigate clinical events of patients using data driven approaches and machine learning [30]. For individual patients, patient trajectories can help individual care; across a population, they can provide important information for secondary purposes such as population health management and inform healthcare policies [30]. When health care providers have access to complete and accurate information, patients receive better medical care; however without an easy access to this information, making decisions may be challenging and time consuming [29].

Electronic health records (EHRs) can improve the ability to diagnose diseases and reduce or even prevent medical errors, improving patient outcomes [31]. However, it has created challenges as well as opportunities for improving health care [22]. One challenge which primary care physicians encounter daily is the cognitive overload, due to the form of electronic health record documentation [32]. Therefore information retrieval remains a challenge [32], [33], leading to challenges in reviewing clinical notes during care episodes [34].

Electronic health records contain an abundance of valuable information [35] that can be used to guide patient care [30] and clinical research [36]-[39]. Automatic recognition of medical concepts in unstructured text is an important component of many clinical and research applications, and its accuracy has a large impact on electronic health record analysis [40].

However, the large volume of information embodied in these records also renders access to relevant information a time-consuming and inefficient process [33]. The challenges experienced in navigating, retrieving, and reviewing paper-based clinical records by physicians still remain unsolved with current EHR systems [22]. While EHRs are effective in collecting patient clinical information, the increasing amount of clinical information and presenting all of them on one screen can lead to information overload [41]. The information overload associated with volumes of large amounts of clinical narratives has threatened the effective use of EHRs [42]. The presence of extraneous data makes it difficult for health care providers to identify relevant information [43]. In the recent past, clinical care and research rely increasingly on digitized patient information for patient-level healthcare management and research respectively [39] e.g. where doctors increasingly use patient's clinical history and consult past similar cases for clinical comparisons and past outcomes to provide targeted healthcare [44]. There has been an ever-increasing amount of

patient records and yet the capacity of extracting useful information from these records remains limited.

In addition, there are many challenges associated with EHR use such as prolonged documentation time, usability difficulties, lack of cognitive support etc. [45]. Inefficient navigation in electronic health records has been shown to increase users' cognitive load, which may increase potential for errors, reduce efficiency, and increase fatigue [46]. Clinicians' time constraint is one of the main barriers to using electronic health resources at the point of care. As a result, the utility of available data in electronic health records in terms of understanding and improving health care is yet to be achieved [47]. The data stored in electronic health records is underused data resource for biomedical research [48]. Therefore, there is need to make medical documents more understandable to doctors.

The use of Electronic Health Records (EHRs) was expected to improve medical decision-making by giving physicians access to stored medical information [49]. For example, effective management of chronic diseases requires health care providers to be able to identify all relevant information at the point of care in a way that supports care coordination, planning and clinical decision making [50]. However, chronic illness care presents a challenge in identifying what information to present and how to present it. This challenge is in part due to the large amount of information generated for a patient with one or more chronic illnesses [51]. The longitudinal nature of the interactions by the chronic illness population with the health care system also includes multiple settings and multiple providers and spans years.

Many studies such as [29] have pointed out the need to improve EHR note design and presentation to support or improve clinicians' note review experience. A key challenge that faces users with a large unstructured clinical report presented in the conventional page view format and relying on keyword querying, is finding out particular information. Two alternatives exist:

   i.   Read through the document to find out the required information.
   ii.   Search for information using key word search to find specific information.

The amount of time and effort required in both cases depends on the length of the report. The user's productivity in this task is determined by their reading speed, their ability to structure effective queries and the effectiveness of the underlying search mechanism. Current search methods are usually keyword-based, and as such, do not incorporate semantics into the search algorithm. The results are either a wealth of information which contains a high degree of irrelevant hits, or at the other extreme, no results at all, because the specific query words were not found although their semantic equivalents do appear quite often in the text. The unprecedented growth in the amount of clinical information accessible by physicians highlights the need for more efficient and more powerful data exploration tools.

III. RELATED WORK

The need of information visualization tools within electronic health records is an active research [52] area in the field of health informatics and general computing. Patients' health records collected over time represent a patient's clinical history and healthcare professionals rely on this information for diagnosis and treatment [53]. We summarize past work on clinical data visualization techniques of electronic health records (EHRs).

There are a number of approaches in the literature for visualizing patients' medical records. Some of the research works have proposed techniques for visualizing clinical documents which mainly targets patients' medical histories available in electronic health records. Few examples in this direction include [54]-[56]. Research studies such as [57]-[59] developed tools that use text visualization to summarize and extract important information from clinical text. West, Borland and Hammond [54] extensively studied the application of visualization techniques in health data and concluded that there are still few techniques which have been found effective and efficient in displaying large and complex clinical data to improve presentation, analysis, and understanding of data [60]. A number of approaches visualizing patient clinical records have been proposed. The most widely used technique uses the concept of timeline in organizing patient records along the time axis. This has given rise to LifeLines [61] and its variants such as LifeLines2 [62]. It distinguishes records using facets, such as problems, symptoms, tests/results, diagnosis, treatments and medications, etc. In addition, colours are used to show different types of information facets and severity of an event [63]. Lifelines displays information facets in a screen distinguished by alternating background colours, stories or aspects are lines; periods correspond to changes of size or color along the line, while discrete events are marked by icons [64]. From the literature, clinical data visualization techniques vary in their strategies. The most predominantly used visualization techniques are time-based visualizations techniques which use graphical representations of data collected over time. The use of graphs has been the widely used technique to illustrate comparisons, trends, and associations in data with an aim making it easily understood [65]. In healthcare domain, these graphs represent time as the horizontal axis to display various types of data [65] and there are several examples of such visualization tools that have been developed mainly for use with temporal data. Examples include LifeLines [66], [67] KNAVE and VISITORS [68]-[70] which are commonly used. The LifeFlow system is designed to support an interactive exploration of event sequences using visualization techniques.

IV. OBJECTIVE

The objective of this study is to propose a technique for identifying and presenting information within clinical notes to support doctors and researchers respectively in seeking information in narrative clinical text. The purpose of our study was to evaluate how text visualization can assist clinical decision making process of a physician when analyzing the clinical history of a patient.

V. Preliminaries

We begin with our notational definitions. In what follows, S denotes a set of input sentences in a clinical document, C denotes a set of possible classes (semantic classes) of S, and c ∗ (s) is the correct labeling of s. For each pair of sentence s and class c ∈ S, we compute a vector-valued feature representation f(s, c). Let $S = \{s_1, \dots \dots \dots s_n\}$ be a sequence of sentences in a patient medical chart. Given a set of relevant sentence classes $C = \{c_1, \dots c_n\} C$ that describe a patient medical history, our task is to extract sentences in each class from the patient medical chart S. I.e., we would like to classify each sentence s_i into one of the classes c_i. Each of the sentences in the clinical notes has features which can help us classify it. We assume that there are no irrelevant sentences in our classification task. It is also possible that, depending on specific cases, a sentence may belong to only one category. A semantic class contains sentences that share a semantic feature. For example, sentences can describe symptoms or treatment. According to the meaning of the sentence, they are categorized into different semantic classes.

VI. Challenges of Unstructured Data

Simple charts and graphs have been used to display specific types of healthcare data to quickly determine the need for appropriate interventions [54].Clinical records by nature contain longitudinal data of patient visits over time, with records of changing problems, medications, treatments, and responses related to evolving health status. From the literature, many health visualization techniques represent few facets of information using graphical representations such as bar charts and scatter plots. Although they are useful for simple perceptual and cognitive tasks, they are inadequate for dealing with tasks that involve exploration of various facets of information [71]. Therefore, there is need for visualizations techniques that encode many facets of data and support exploration of huge amount of data and more complex tasks [71]. During every patient encounter healthcare provider rely on hundreds of discrete data on which they base their decisions in order to provide effective treatment. As care teams assess, diagnose, and treat patients, they are creating and adding to the patients' medical record with a range of data such as discrete lab results, qualitative descriptions, transcripts of their opinions and decisions and more. It is no surprise the industry is facing a monumental challenge: how to unlock all of these unstructured clinical data for new insights into innovation and improving patient care.

VII. Semantic Structure of a Document

In this paper we are concerned with retrieving clinical information, organizing, and finally displaying information. Grouping similar items (e.g. sentences) into logical groups is a natural cognitive human process. When a person is presented with a set of items, one naturally groups and organizes these items within his or her mind. This tendency has been imitated in machine learning using unsupervised clustering algorithms. In this paper, we present a visual system that uses sentence classification and visualization to aid data exploration in a natural way. Classifying important sentences or phrases from

text according to their semantic classes and visualizing them is our proposed technique. We used classification to determine which semantic class, a sentence belongs to. Deducing the semantic structure of a document takes place in two stages. First, each sentence in the document is classified into one of the five basic semantic classes of symptoms, diagnosis, disease, treatment, and personal identifying information. Secondly, sentences are grouped into semantic groups, thus rendering the semantic "structure" of the document.

A. Sentence-level Classification

The ability to group information into categories based on some similarity of meaning is an important human cognitive behaviour [72]. We introduce a method for classifying sentences of patients' medical charts into a number of semantic classes. In selecting classes to represent the semantics of patient medical charts, we started by asking doctors the kinds of information that they would want to explore during care episodes in a patient's clinical chart. A user reviewing this data might be interested in some information, such as symptom, condition, previous diagnosis etc. The notes are usually documented using SOAP format. The SOAP (Subjective, Objective, Assessment and Plan) is an acronym representing a documentation method used by healthcare professional to write clinical notes. A SOAP note is information about the patient, which is written or presented in a specific order, which includes certain components [73], [74]. SOAP notes are used for admission notes, medical histories, and other documents in a patient's chart. Doctors usually document their interactions with their patients hence producing a series of clinical notes which are used to provide better care for patients. These notes are often documented using SOAP standard and feature four distinct sections: Subjective, Objective, Assessment, and Plan. SOAP is a documentation method employed by physicians to write out clinical notes in a patient's chart. SOAP standard can be organized into four SOAP sections [75] as follows:

i. **Subjective**: Patients verbally express symptoms and observations. This also includes details about medication history, family history etc.

ii. **Objective**: Observations that can be measured in different ways, such as physical examination, test result, blood pressure, height, weight, and other vital signs.

iii. **Assessment**: a list of diagnoses regarding a patient's condition.

iv. **Plan**: This refers to how a doctor is going to address a patient's problem. Follow-up directions for the patient, such as medications, treatment plan etc.

During consultation, a physician usually starts by noting down patient symptoms in order to understand the patient's clinical status (Subjective). Then, the physician documents observations which can be measured i.e., quantifiable data and scientific evidence experienced by the patient (Objective). The physician will then use subjective and objective information to come up with a possible diagnosis (Assessment) and, finally, gives a plan to treat the underlying disease/condition (Plan) [76]. From the above, we used the above SOAP framework for clinical sentences classification and developed an automated classifier that classifies sentences in clinical documents. We

identified familiar terms which can be mapped to SOAP framework as shown in the table below:

TABLE 1: MAPPING OF SOAP TERMS TO FAMILIAR CLINICAL TERMS

Familiar Term	SOAP Term
Symptoms	S: subjective
Diagnosis	O: objective
Assessments/Conclusions	A: Assessment
Treatments	P: plans

We will begin our exploration into the semantics of sentences/phrases based on the above table. Semantics is defined as the study of meaning in language. In addition to the above semantic classes, we introduced personal identifying information (PII) class to represent the demographic information which is usually part of the documentation.

Our approach is to classify sentences according to the above taxonomy. We focused on the meaning(s) attributed to clinical sentence/phrase meaning. We focus on classifying sentences into five relevant classes describing a patient care episode:

(1) Symptoms;
(2) Diagnostic;
(3) Disease;
(4) Treatment;
(5) Personal identifying information.

In this work, we seek to build a classifier using training data from the semantic labels already created. The purpose of classification is to facilitate grouping of semantically similar sentences to meet user interests in information seeking and understanding. Classification is done using a deep learning model trained using a large corpus of real-world clinical records. This novel method takes into account the semantic meaning of a whole sentence. Consider the text, "Patient has high fever." Our approach is able to classify this as a symptom sentence. The advantage of such an approach is that we are now able to classify text reliably enough to be able to provide the semantic content of large text documents at arbitrary levels of granularity. Our approach is unique as most text classification techniques are based on word and/or phrase analysis of the text.

B. Providing Semantic Structure

Our concept of visualizing sentences in semantic groups' combines sentence classification and cluster map generation.

To provide a semantic structure, sentences are put into semantic groups producing meaningful clusters, and also providing meaningful topic cues which would make analysis of the clusters much easier than not having any labels or topics. We organize sentences of narrative text into a smaller number of meaningful semantic units which are presented as a cluster map. Classification of texts in a clinical document is used as a precursor to create a semantic structure. Semantically classified sentences are grouped into five predefined semantic classes based on their classifications. We group sentences based on their labels which have been assigned by the classifier. We assumed complete independence of the semantic classes. We then display them using a cluster map, where each cluster represents a unique semantic class with sentences for that particular class. A cluster map displays group of sentences in their respective cluster nodes hence allowing users to view, explore and visualize information in groups that would have otherwise

been hidden in a linear continuous text. It quickly aggregates information into smaller groupings of information and thus providing a better navigation and reading of the available text.

C. Distinguishing Semantic Classes Using Clours

We highlight sentences of different semantic classes using different colours. This was motivated by the need for a representation that can distinctly and visually codify five semantic classes and be identifiable even in their own clusters. We selected the following colours for different semantic classes.

TABLE 2: MAPPING OF CLINICAL TERMS TO COLOURS

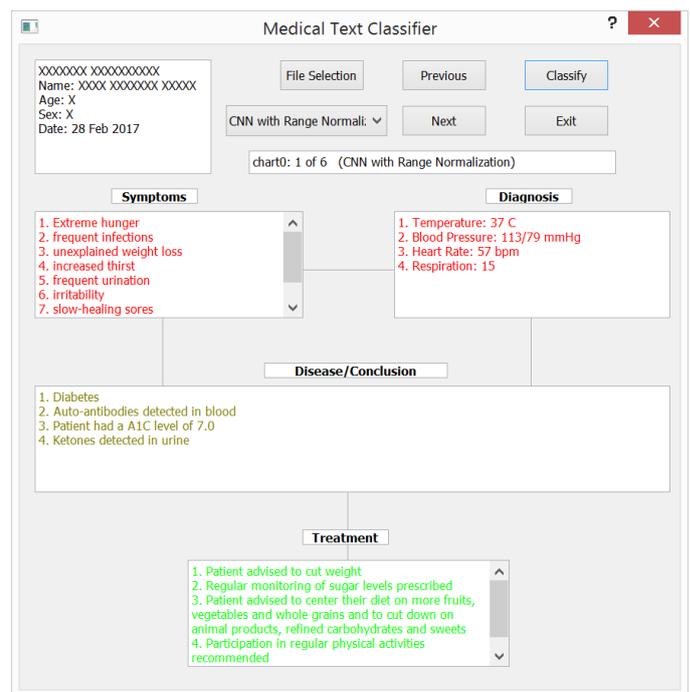| Semantic class	Colour
Symptoms	*Red*
Diagnosis	*Red*
Disease	*Purple*
Treatment	*Green*
PII	*Black*

Fig. 1. Sample cluster map.

VIII. RESEARCH METHODS

Using text classification and visualization techniques in order to represent long textual documents graphically and to enable users to explore data at different levels of detail was our main objective. As a proof of concept, we built a prototype of the classification and visualization for a patient medical chart stored in Electronic Health Records (EHRs). The prototype was used in an empirical study aimed to explore the opinions of physicians about usability and ease of access to relevant patient health information. The quantitative and qualitative results gathered from usability testing, showed a significant preference for using visualization tool when having to navigate a patient's medical chart for the purpose of finding pertinent information.

IX. RESEARCH SETTING

A. System Model

A prototype was developed to carry out usability testing of the proposed visualization tool and collect feedback from physicians on the usefulness of such a tool in their daily encounters with patients. Here we describe the study design, data and task choices, participant cohort, and study methods. Our proposed model classifies and visualizes any text as a cluster map. The resulting map can be used to get a quick visual summary of the underlying text and thus help user in reading the most relevant information. We formulated our evaluation to study four research questions:

i. How do physicians rate the usability of visual presentation of real patients' medical charts during care episodes?

ii. Do visual representations of patients' medical charts appeal to physicians - specifically, do they impact accuracy, timely information seeking?

B. Dataset

Physicians usually use a variety of clinical documentation during care episodes. In this paper, we used one commonly used document, patient medical chart which describes patient medical history. In this research, we used patient medical chart as our clinical document source. For evaluation purpose, we used a dataset of 4998 semi-synthetic medical charts developed from an original set of clinical documents obtained from (http://www.mtsamples.com/). The website provides a large collection of publicly available transcribed medical records. We retrieved 4998 medical charts since these charts contain very rich medical history information. We identified four SOAP sections and added another section for which includes demographic information about the patient as well. The data consist of human-generated transcripts with clinical notes. Figure 1.1 shows the distribution of sentences in the dataset

C. Design Science Research

The research questions posed above are practical problems faced by physicians during healthcare delivery. Physicians need to assess the usability, information access and information usability of our proposed visualization. Both tasks are different facets of the general problem of evaluation of visualization techniques. In this paper we adopted design science research approach. This is a research technique with the objective of developing new or better ways to achieve human goals [77], [78]. The human goals are related to relevant practical problems (or tasks) faced by humans in a specific environment. It attempts to solve problems and challenges by creating new and innovative artifacts [78]. Thus, design science implies the existence of a relevant practical problem for which it tries to find a solution. The solution is usually referred to as "artifact" [77], [78].

According to March and Smith [77], design science research has two main activities:

a) Build the solution and
b) Evaluate the solution.

Building is the process of constructing an artefact for a specific purpose; evaluation is the process of determining how well the artifact performs [77]. In the build solution activity, a researcher develops an innovative and valuable solution (artifact). Usually the solution developed by the researcher is based on existing knowledge of the problem-domain, methodology, and technology [78]. In evaluate the solution activity, the utility of the developed artifact is assessed [77]. The evaluation is however a difficult process, due to the fact that "performance is related to intended use, and the intended use of an artifact can cover a range of tasks" [77]. The evaluation criteria and metrics may differ for each intended task and context of use of the artifact. Therefore, the criteria and metrics must be determined (developed and/or specified) for the artifact in each particular environment (context of use) for which the artifact is evaluated. The employed metrics are justified by using natural science approaches (e.g., data collection and analysis). As March and Smith point out, the evaluation of the constructed artifact is important in judging the research effort of building the artifact. Design artifacts are constructs models, methods, and instantiations. Purposeful artifacts are built to address heretofore unsolved problems [78]. Hevner and others make clear that the goal of design-science research is utility, and a proposed solution (artifact) addresses important unsolved problems in unique or innovative ways or solves problems in more effective or efficient ways [78].

Design science research (DSR) approach was to design a high-fidelity prototype. The prototype was evaluated using the System Usability Scale (SUS) framework to evaluate the overall usability of the system (n = 12). The second evaluation was conducted by designing a running prototype and evaluating the design using a questionnaire to evaluate the system's usefulness and information quality, and interface quality (n = 50).

X. EVALUATION

We adopted the evaluation framework for evaluating visual Data-mining tools proposed in [79] by Marghescu and Rajanen to evaluate the developed artifact by taking into consideration three levels of analysis:

i. Information Completeness - evaluating if the required information to make clinical decision are complete in the visual summary.

ii. Utility - evaluating the capability of the visualization model to summarize a chart and help a user to review the clinical documents.

iii. Usability of the visualization model- Evaluating the usability is conducted in order to find out whether the users of the system consider the system easy to use and learn, accurate, effective and efficient.

For each of the three levels, we identified and described the corresponding attributes. In this study, our hypothesis for this study was that applying text visualization to clinical document synthesize improves their usability, information and utility.

In order to get end user perspective we used two sets of questionnaires; System Usability Scale (SUS) to evaluate the usability regarding the visualization model and 7 Likert scale questionnaire to evaluate the user perception on the ease of access of information and utility of the information obtained from the visualization model. The model was demonstrated to

participants to facilitate easier evaluation which allowed them to conduct evaluations.

TABLE 3: ELEMENTS OF ANALYSIS ATTRIBUTES

Usability		
	1.	Ease of use
	2.	Learnability
	3.	Accuracy
	4.	Efficiency
Information	1.	Relevant information
	2.	Accuracy
	3.	Completeness
	4.	Visibility
	5.	Accessibility
	6.	Satisfying
	7.	Novelty
Utility	1.	Richness
	2.	Accuracy
	3.	Clarity
	4.	Understandability
	5.	Fast
	6.	Present ability
	7.	Privacy
	8.	Research

A. Usability Evaluation Using (SUS) Scale

Usability is defined as the effectiveness, efficiency, and satisfaction of a system. Our proposed model needs to work (effective), work well (efficient), and not cause any unnecessary frustration (satisfying). To gain insight into the usability of our prototype, we conducted a formative user study with 12 participants, who reviewed our prototype usability in summarizing patients' clinical charts. The study used usability questionnaires called; the System Usability Scale (SUS) which is a widely used and validated instrument [80] and a popular measure of perceived usability [81]. In order to evaluate the usability of the clinical document visualization prototype, the System Usability Scale (SUS) was used to assess subjective reactions to the prototype. SUS is a usability evaluation tool developed by J. Brooke [82] which is widely accepted technique for evaluating usability of software systems. It evaluates effectiveness, efficiency, and satisfaction parameters of usability. It is a questionnaire made up of 10 questions which are rated on a 5-point scale rating [83]. It is given to the participants who then score the questions after using the system that is being evaluated. Each question was scored with a value ranging from *one* to *five* i.e. pparticipants ranked each question from 1 to 5 based on how much they agreed with different statements. A score of one means strongly *disagree* and a score of five means strongly *agree*. From these values a SUS-score was calculated. The SUS is calculated as follows; for odd number questions, the scored values are reduced by 1 and for even number questions, the scored values are reduced by 5. All scores are then added up and multiplied by 2.5. This results in a value ranging from 0 to 100. It is a widely used tool for measuring usability because it is cost-effective. Using a 10-question Likert scale has the advantage of giving a global view of subjective assessments of usability. It is sometimes referred to as a 'quick and dirty' usability scale, because of its easiness of use [83].

B. Information Access and Utility of the Evaluation

To determine the usefulness of the clinical visualization model, we designed another questionnaire which evaluates the usefulness of each element of the visualization model. This was arrived at after taking into account the limitation of SUS score, which only gives an overall performance score without considering each element of the system. The two useful elements in the proposed evaluation were easy access to information in patients' medical charts and the utility of the visualized information in adding the users in analyzing clinical documents.

For this questionnaire we had 50 participants. Each question had 1-7 Likert scale answering possibility, where 1 is totally disagree/not useful at all, and 7 is totally agree/extremely useful. Below are the results of this questionnaire.

TABLE 4: 7-POINT LIKERT SCALE

Consideration	Value
7	Strongly Agree
6	Agree
5	Moderately Agree
4	Slightly Agree
3	Slightly Disagree
2	Moderately Disagree
1	Strongly Disagree

Likert scales do a non-comparative scaling technique where respondents are asked to evaluate a single trait. Respondents are asked to indicate their level of agreement with a given statement by way of an ordinal scale. 5-point scale is the mostly used though some practitioners advocate the use of 7 and 9-point scales which have the advantage of adding additional granularity. Others include 4-point scale and even-numbered scales which are used to produce an ipsative (forced choice) measure where no indifferent option is available. Research has shown that Likert scales with seven response options are more reliable than equivalent items with greater or fewer options.

C. Data Collection

In this paper we used a System Usability Scale (SUS) to gather usability rating from users. We also administered a 7-Likert scale questionnaire to users to evaluate information access and information usability of the visualization model.
Both the questionnaires were made online for users to evaluate. Follow-ups were made to ensure that respondents' response was high for the study. In regard to questionnaire used, it had two sections i.e. Section A that covered the ease of access of information and Section B that gathered information on utility of the information obtained from the visualization model.

D. Sampling Strategy and Selection of Participants

We adopted purposive sampling strategy to select study participants, for the evaluation of the model. The participants were contacted in person and were requested if they were willing to participate. The participation was voluntary. During the process, a set of inclusion and exclusion criteria was made as follows:

i) The inclusion criteria for evaluation required participants to be professional doctors currently using any form of electronic health systems, and are documenting patient health data using any form of EHR.

ii) Individuals who did not meet the above inclusion criteria were excluded from the study.

Accordingly, participants were recruited to evaluate the prototype using questionnaire. A purposive sample of twelve doctors was selected for review of clinical charts. Patients' charts selected for this study had descriptions of complex conditions and multiple comorbidities.

E. Task

During evaluation, participants were asked to use clinical document visualization model to perform a common clinical task, to review some of the patients' medical charts by visualizing a given patient's medical chart and reviewing the patient's clinical history.

XI. RESULTS

A. The Result of SUS Scale Scores Visualization Tool Usability Study

TABLE 5: SUS RESULTS

Participant	Questions										SUS Score
	1	2	3	4	5	6	7	8	9	10	
1	5	1	5	1	4	1	4	1	4	2	90.0
2	3	2	4	2	5	2	5	2	4	3	75.0
3	5	2	3	2	4	2	4	1	5	1	82.5
4	4	1	4	2	4	3	3	1	4	2	75.0
5	5	2	4	1	3	1	4	2	4	1	82.5
6	2	2	3	3	2	2	3	3	3	2	52.5
7	4	2	2	3	4	2	4	2	2	1	65.0
8	4	2	3	2	5	1	2	2	3	3	67.5
9	5	2	4	1	4	2	4	1	4	1	85.0
10	4	1	2	2	4	3	3	2	3	2	65.0
11	4	2	4	2	4	2	2	2	4	2	70.0
12	3	2	3	1	3	1	4	1	4	1	77.5
AVERAGE											74.0

The SUS-score results show a good average score of 73.96 in favour of the visualization prototype which indicates better usability. An average SUS-score of 73.96 is considered good rating according to the SUS score scale below proposed IN [82] by Bangor.
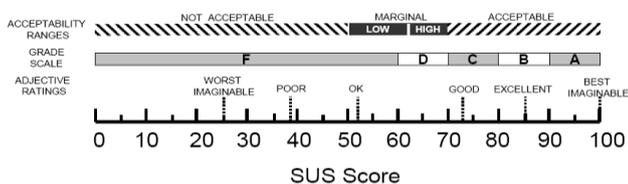

Fig. 2. SUS score scale.

B. Information Access Results

The results above depict the opinions of the participants regarding the quality of information obtained from the visualization model. Among the positive features, we observed includes the good relevancy of visualized information where 85.2% of the respondents agree that Information obtained from the visualization model is relevant to their work. Other positive ratings are accuracy of information (54%), completeness of information (62%), ease access of information (62.40%), subjective satisfaction of using the prototype (60%) and whether the approach is a creative way of summarizing clinical documents (86.8%). Majority of the

respondents (86.80%) agreed that the approach of visualizing clinical documents is a creative way of tackling information overload affecting healthcare delivery.
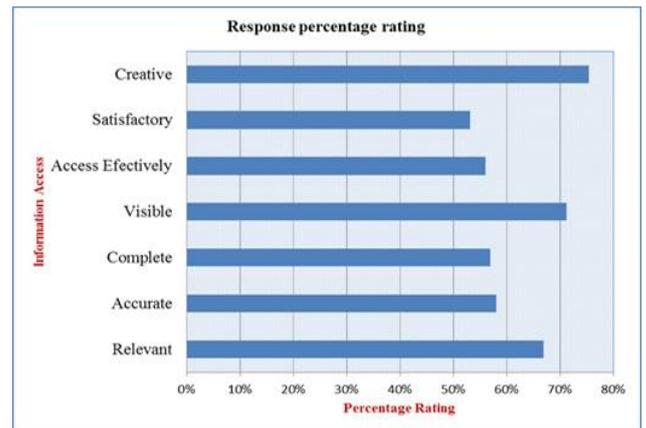

Fig. 3. Information access results.

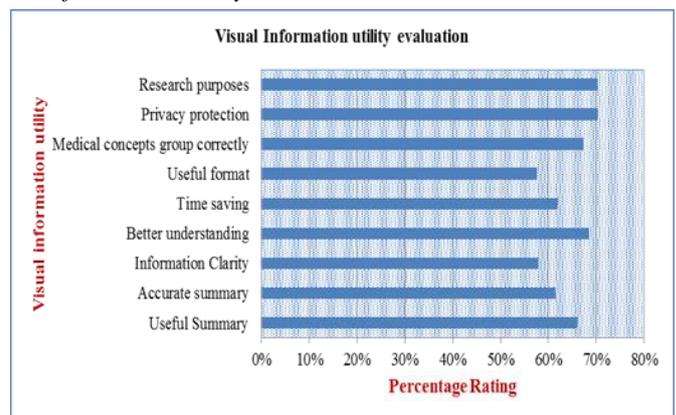### C. Information Utility Results

Fig. 4. Information utility results.

The results above depict the opinions of the participants regarding the utility of information obtained from the visualization model. We can conclude that the information obtained from the visualization model is helpful and useful in clinical chart analysis. It is also relevant, accessible, accurate, complete, satisfactory, and creative for most of the respondents.

XII. DISCUSSIONS

In this paper we assessed the feasibility of automatically generating visual summaries for a particular patient clinical chart which is usually made up of sentences/phrases describing different aspects of a patient clinical history. In today's era, when the size of digital clinical data is increasing exponentially, there is need to retrieve and present concise summaries of the available information. However, most of this data is available in unstructured form and creating summaries from text is a tedious task. One has to read the whole information document to be able to come up with a summary representation of the text. Therefore, the wealth of digitally stored clinical data available in electronic health records (EHRs) today increases the demand for effective tools to retrieve and manage relevant data. Keyword searches over this digital data easily retrieve lists of several hundreds of

documents. Visualizing a document's structure is proven to be useful in aiding information retrieval.

The results of any patient encounter process are predominantly documented in natural language as clinical notes. These notes contain content such as symptoms, diagnosis, and treatment. For the subsequent use of documented clinical notes, it is important to be able to differentiate between patient relevant information and other auxiliary content. Therefore, one has to read these clinical notes to manually identify each information element of a clinical document. However, this manual task is time-consuming and error-prone. In this paper, we have presented an approach to automatically classify and visually present relevant elements of a patient clinical chart such as "symptoms", "diagnosis", "disease" or "treatment".

XIII. Conclusion

The growing number of electronic clinical records available in electronic health records (EHRs) entails an increasing demand by users for interfaces that facilitate the simplified finding and comprehension of relevant information. In this paper, we have presented a prototype that incorporates classification and visualization to promote the reviewing of different clinical narrative text and helps to present this information in an easy to use way using a cluster map. The prototype comprises two modules, sentence classification and a visual presentation of classified sentences using a cluster map. An evaluation of the proposed model with domain experts revealed that visualization of clinical documents has potential to be used in analysing clinical documents. The evaluation of the proposed model for its intended purpose in a healthcare delivery setting, achieved good ratings from the respondents. In conclusion, we can conclude that automatic classification and visualization of clinical documents to help doctors in health care delivery is feasible and effective.

Acknowledgement

References

[1] Wang, Y., Tafti, A.P., Sohn, S., & Zhang, R. (2019). Applications of Natural Language Processing in Clinical Research and Practice. *NAACL-HLT*.

[2] Meystre, S., Savova, G., Kipper-schuler, K., & Hurdle, J. (2008). Extracting information from textual documents in the electronic health record: a review of recent research. *Yearbook of medical informatics*, 128-44 .

[3] Laxmisan, A., McCoy, A. B., Wright, A., & Sittig, D. F. (2012). Clinical Summarization Capabilities of Commercially-available and Internally-developed Electronic Health Records. *Applied Clinical Informatics*, *3*(1), pp. 80–93. http://doi.org/10.4338/ACI-2011-11-RA-0066.

[4] Katal, A., Wazid, M., & Goudar, R.H. (2013). Big data: Issues, challenges, tools and Good practices. 2013 Sixth International Conference on Contemporary Computing (IC3), 404-409.

[5] Jonnalagadda, S. R., Del Fiol, G., Medlin, R., Weir, C., Fiszman, M., Mostafa, J., & Liu, H. (2013). Automatically extracting sentences from Medline citations to support clinicians' information needs. *Journal of the American Medical Informatics Association: JAMIA*, *20*(5), 995–1000. https://doi.org/10.1136/amiajnl-2012-001347.

[6] Taylor, C., 2020. Structured Vs. Unstructured Data. [online] Datamation.com. Available at: <https://www.datamation.com/big-data/structured-vs-unstructured-data.html> [Accessed 2 August 2020].

[7] Sun, W., Cai, Z., Li, Y., Liu, F., Fang, S., & Wang, G. (2018). Data Processing and Text Mining Technologies on Electronic Medical Records: A Review. Journal of Healthcare Engineering, 2018.

[8] Huang, K., Altosaar, J., & Ranganath, R. (2019). ClinicalBERT: Modeling Clinical Notes and Predicting Hospital Readmission. ArXiv, abs/1904.05342.

[9] Barbulescu, M., Grigoriu, R., Halcu, I., Neculoiu, G., Sandulescu, V. C., Marinescu, M., & Marinescu, V. (2013). Integrating of structured, semi-structured and unstructured data in natural and build environmental engineering. In 2013 11th RoEduNet International Conference (pp. 1–4). IEEE. doi:10.1109/RoEduNet.2013.6511738.

[10] Farri, O., Pieczkiewicz, D.S., Rahman, A., Pakhomov, S.V., Adam, T., & Melton, G.B. (2012). A Qualitative Analysis of EHR Clinical Document Synthesis by Clinicians. *AMIA ... Annual Symposium proceedings. AMIA Symposium, 2012*, 1211-20.

[11] Farri, O., Monsen, K.A., Pakhomov, S.V., Pieczkiewicz, D.S., Speedie, S.M., & Melton, G.B. (2013). Effects of time constraints on clinician-computer interaction: A study on information synthesis from EHR clinical notes. Journal of biomedical informatics, 46 6, 1136-44.

[12] Pollack, A.H., & Pratt, W. (2020). Association of Health Record Visualizations With Physicians' Cognitive Load When Prioritizing Hospitalized Patients. *JAMA Network Open, 3*.

[13] Farri, O. (2012). Understanding Clinician Information Demands and Synthesis of Clinical Documents in Electronic Health Record Systems.

[14] Bui, A.A., Aberle, D., & Kangarloo, H. (2007). TimeLine: Visualizing Integrated Patient Records. IEEE Transactions on Information Technology in Biomedicine, 11, 462-473.

[15] Farri, O., Rahman, A., Monsen, K. A., Zhang, R., Pakhomov, S. V., Pieczkiewicz, D. S., … Melton, G. B. (2012). Impact of of a Prototype Visualization Tool for New Information in EHR Clinical Documents. Applied Clinical Informatics, 3(4), 404–418. http://doi.org/10.4338/ACI-2012-05-RA-0017.

[16] Ruan, W., Appasani, N., Kim, K., Vincelli, J., Kim, H., & Lee, W. (2018). Pictorial Visualization of EMR Summary Interface and Medical Information Extraction of Clinical Notes. *2018 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, 1-6.

[17] Deng, Y., & Denecke, K. (2014). Visualizing Unstructured Patient Data for Assessing Diagnostic and Therapeutic History. *Studies in health technology and informatics, 205*, 1158-1162.

[18] Ledesma, A., Bidargaddi, N., Strobel, J.E., Schrader, G., Nieminen, H., Korhonen, I., & Ermes, M. (2019). Health timeline: an insight-based study of a timeline visualization of clinical data. BMC Medical Informatics and Decision Making, 19(170).

[19] Bui, A., Aberle, D.R., Kangarloo, H., (2007). TimeLine: Visualizing Integrated Patient Records. IEEE Transactions on Information Technology in Biomedicine 11(4), pp.462-473.

[20] Interactive Visualization and Exploration of Patient Progression in a Hospital Setting.

[21] Dornan, L., Pinyopornpanish, K., Jiraporncharoen, W., Hashmi, A.H., Dejkriengkraikul, N., & Angkurawaranon, C. (2019). Utilisation of Electronic Health Records for Public Health in Asia: A Review of Success Factors and Potential Challenges. *BioMed Research International, 2019*.

[22] Boonstra, A., Versluis, A., & Vos, J.F. (2014). Implementing electronic health records in hospitals: a systematic literature review. *BMC Health Services Research, 14*.

[23] Wolfe, L., Chisolm, M.S., & Bohsali, F. (2018). Clinically Excellent Use of the Electronic Health Record: Review. *JMIR Human Factors, 5*.

[24] Pivovarov, R., & Elhadad, N. (2015). Automated methods for the summarization of electronic health records. Journal of the American Medical Informatics Association: JAMIA, 22(5), 938–947. https://doi.org/10.1093/jamia/ocv032.

[25] Schopf, T.R., Nedrebø, B., Hufthammer, K.O., Daphu, I.K., & Lærum, H. (2019). How well is the electronic health record supporting the clinical tasks of hospital physicians? A survey of physicians at three Norwegian hospitals. BMC Health Services Research, 19.

[26] Menachemi, N., & Collum, T.H. (2011). Benefits and drawbacks of electronic health record systems. Risk Management and Healthcare Policy, 4, 47 - 55.

[27] Bates, D.W., Teich, J.M., Lee, J., Seger, D.L., Kuperman, G.J., Ma'Luf, N., Boyle, D., & Leape, L. (1999). Research Paper: The Impact of Computerized Physician Order Entry on Medication Error Prevention.

Journal of the American Medical Informatics Association : JAMIA, 6 4, 313-21.

[28] Bates, D.W., Leape, L., Cullen, D.J., Laird, N., Petersen, L.A., Teich, J.M., Burdick, E., Hickey, M., Kleefield, S., Shea, B., Vliet, M.V., & Seger, D.L. (1998). Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. JAMA, 280 15, 1311-6.

[29] HealthIT. "Improved Diagnostics & Patient Outcomes." HealthIT.gov. http://www.healthit.gov/providers-professionals/improved-diagnostics-patient-outcomes. Updated June 2019. Accessed May 2020.

[30] Landi, I., Glicksberg, B., Lee, H., Cherng, S.T., Landi, G., Danieletto, M., Dudley, J., Furlanello, C., & Miotto, R. (2020). Deep representation learning of electronic health records to unlock patient stratification at scale. NPJ Digital Medicine, 3.

[31] Koopman, R.J., Steege, L.M., Moore, J.L., Clarke, M.A., Canfield, S.M., Kim, M.S., & Belden, J.L. (2015). Physician Information Needs and Electronic Health Records (EHRs): Time to Reengineer the Clinic Note. *Journal of the American Board of Family Medicine : JABFM, 28 3*, 316-23.

[32] Mujjiga, S., Krishna, V., Chakravarthi, K., & Vijayananda, J. (2019). Identifying Semantics in Clinical Reports Using Neural Machine Translation. *AAAI*.

[33] Reis, J.C., Bonacin, R., & Perciani, E.M. (2016). Intention-Based Information Retrieval of Electronic Health Records. *2016 IEEE 25th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, 217-222.

[34] Hultman, G.M., Marquard, J.L., Lindemann, E., Arsoniadis, E.G., Pakhomov, S.V., & Melton, G.B. (2019). Challenges and Opportunities to Improve the Clinician Experience Reviewing Electronic Progress Notes. *Applied clinical informatics, 10 3*, 446-453.

[35] Reichert, D., Kaufman, D., Bloxham, B., Chase, H., & Elhadad, N. (2010). Cognitive analysis of the summarization of longitudinal patient records. AMIA ... Annual Symposium proceedings. AMIA Symposium, 2010, 667–671.

[36] Nordo, A.H., Levaux, H.P., Becnel, L.B., Gálvez, J., Rao, P.K., Stem, K., Prakash, E., & Kush, R.D. (2019). Use of EHRs data for clinical research: Historical progress and current applications. Learning Health Systems, 3.

[37] Juhn, Y.J., & Liu, H. (2019). Natural language processing to advance EHR-based clinical research in Allergy, Asthma, and Immunology. The Journal of allergy and clinical immunology.

[38] Meystre, S.M., Lovis, C., Bürkle, T., Tognola, G., Budrionis, A., & Lehmann, C.U. (2017). Clinical Data Reuse or Secondary Use: Current Status and Potential Future Progress. Yearbook of medical informatics, 26 1, 38-52.

[39] Beresniak, A., Schmidt, A., Proeve, J., Bolaños, E., Patel, N., Ammour, N., Sundgren, M., Ericson, M., Karakoyun, T., Coorevits, P.L., Kalra, D., Moor, G.J., & Dupont, D. (2016). Cost-benefit assessment of using electronic health records data for clinical research versus current practices: Contribution of the Electronic Health Records for Clinical Research (EHR4CR) European Project. Contemporary clinical trials, 46, 85-91.

[40] Arbabi, A., Adams, D.R., Fidler, S., & Brudno, M. (2019). Identifying Clinical Terms in Medical Text Using Ontology-Guided Machine Learning. JMIR Medical Informatics, 7.

[41] Hultman, G.M., Marquard, J.L., Lindemann, E., Arsoniadis, E.G., Pakhomov, S.V., & Melton, G.B. (2019). Challenges and Opportunities to Improve the Clinician Experience Reviewing Electronic Progress Notes. Applied clinical informatics, 10 3, 446-453.

[42] Su, Y., Chao, C., Hung, L., Sung, S., & Lee, P. (2020). A Natural Language Processing Approach to Automated Highlighting of New Information in Clinical Notes. Applied Sciences, 10, 2824.

[43] Clarke, M.A., Belden, J.L., Koopman, R.J., Steege, L.M., Moore, J.L., Canfield, S.M., & Kim, M.S. (2013). Information needs and information-seeking behaviour analysis of primary care physicians and nurses: a literature review. *Health information and libraries journal, 30(3)*, 178-190.

[44] McEnery, K. W. 2018. *Reference Guide in Information Technology for the Practicing Radiologist*. American College of Radiology.

[45] Ajami S, Bagheri-Tadi T. Barriers for adopting electronic health records (EHRs) by physicians. Acta Informatica Medica 2013;21:129. pmid:24058254
View ArticlePubMed/NCBIGoogle Scholar.
Roman, L.C., Ancker, J., Johnson, S., & Senathirajah, Y. (2017). Navigation in the electronic health record: A review of the safety and usability literature. *Journal of biomedical informatics, 67*, 69-79.

[46] Reisman, M. (2017). EHRs: The Challenge of Making Electronic Data Usable and Interoperable. *P & T : a peer-reviewed journal for formulary management, 42 9*, 572-575.

[47] Jensen, P., Jensen, L. & Brunak, S. Mining electronic health records: towards better research applications and clinical care. *Nat Rev Genet* 13, 395–405 (2012). https://doi.org/10.1038/nrg3208.

[48] Ben-Assuli, O., Ziv, A., Sagi, D., Leshno, M., & Ironi, A. (2014). Improving Diagnostic Accuracy Using EHR in Emergency Departments: A Simulation-Based Study. Journal of biomedical informatics, 55, 31-40.

[49] Aleksi, D.S., Rajkovi, P., Vukovi, D., Jankovi, D., & Milenkovi, A. (2017). Data summarization method for chronic disease tracking. *Journal of Biomedical Informatics, 69*, 188-202.

[50] Samal, L., Wright, A., Wong, B., Linder, J., & Bates, D. (2011). Leveraging electronic health records to support chronic disease management: the need for temporal data views. *Informatics in primary care, 19 2*, 65-74.

[51] Lesselroth, B.J., & Pieczkiewicz, D.S. (2011). Data Visualization Strategies for the Electronic Health Record.

[52] Ledesma, A., Bidargaddi, N., Strobel, J., Schrader, G., Nieminen, H., Korhonen, I., & Ermes, M. (2019). Health timeline: an insight-based study of a timeline visualization of clinical data. *BMC Medical Informatics and Decision Making, 19*.

[53] West, V., Borland, D. and Hammond, W. (2014). Innovative information visualization of electronic health record data: a systematic review. *Journal of the American Medical Informatics Association*.

[54] Roque, F., Slaughter, L.A., & Tkatsenko, A. (2010). A Comparison of Several Key Information Visualization Systems for Secondary Use of Electronic Health Record Content. In: Proceedings of NAACL HLT Workshop on Text and Data Mining of Health Documents, pp.1–8.

[55] Rind, A., Wang, T. D., Aigner, W., Miksh, S., Wongsuphasawat, K., Plaisant, C., and Shneiderman, B (2013). Interactive Information Visualization to Explore and Query Electronic Health Records. Foundations and Trends in Human–Computer Interaction, 5(3), pp.207-298.

[56] Sultanum, N., Singh, D.A., Brudno, M., & Chevalier, F. (2018). Doccurate: A Curation-Based Approach for Clinical Text Visualization. IEEE transactions on visualization and computer graphics.

[57] Trivedi, G., Pham, P., Chapman, W.W., Hwa, R., Wiebe, J., & Hochheiser, H. (2018). NLPReViz: an interactive tool for natural language processing on clinical text. Journal of the American Medical Informatics Association: JAMIA, 25(1), pp. 81-87.

[58] Yerebakan, H.Z., Shinagawa, Y., Bhatia, P.S., & Zhan, Y. (2018). Document Representation Learning for Patient History Visualization. COLING.

[59] West, V. L., Borland, D., & Hammond, W. E. (2015). Innovative information visualization of electronic health record data: a systematic review. Journal of the American Medical Informatics Association : JAMIA, 22(2), pp. 330–339. http://doi.org/10.1136/amiajnl-2014-002955.

[60] Plaisant, C., Mushlin, R., Snyder, A., Li, J., Heller, D., & Shneiderman, B. (1998). LifeLines: using visualization to enhance navigation and analysis of patient records. *Proceedings. AMIA Symposium*, 76-80.

[61] Wang, T.D., Plaisant, C., Quinn, A.J., Stanchak, R., Murphy, S., & Shneiderman, B. (2008). Aligning temporal data by sentinel events: discovering patterns in electronic health records. *CHI*.

[62] Zhang, Z., Ahmed, F., Mittal, A., Zhao, R., Viccellio, A., & Mueller, K. (2011). AnamneVis: A Framework for the Visualization of Patient History and Medical Diagnostics Chains.

[63] Plaisant, C., Milash, B., Rose, A., Widoff, S., & Shneiderman, B. (1996). LifeLines: visualizing personal histories. *CHI '96*.

[64] West, V. L., Borland, D., & Hammond, W. E. (2015). Innovative information visualization of electronic health record data: a systematic review. *Journal of the American Medical Informatics Association : JAMIA, 22(2)*, 330–339. https://doi.org/10.1136/amiajnl-2014-002955.

[65] Plaisant, C., Mushlin, R., Snyder, A., Li, J., Heller, D., & Shneiderman, B. (1998). LifeLines: using visualization to enhance navigation and analysis of patient records. Proceedings. AMIA Symposium, 76-80.

[66] Plaisant, C., Milash, B., Rose, A., Widoff, S., & Shneiderman, B. (1996). LifeLines: visualizing personal histories. CHI '96.

[67] Martins, S.B., Shahar, Y., Goren-Bar, D., Galperin-Aizenberg, M., Kaizer, H., Basso, L., McNaughton, D., & Goldstein, M. (2008). Evaluation of an architecture for intelligent query and exploration of time-oriented clinical data. Artificial intelligence in medicine, 43 1, 17-34.

[68] Klimov, D., Shahar, Y., & Taieb-Maimon, M. (2010). Intelligent visualization and exploration of time-oriented data of multiple patients. Artificial intelligence in medicine, 49 1, 11-31.

[69] Klimov, D., Shahar, Y., & Taieb-Maimon, M. (2009). Intelligent interactive visual exploration of temporal associations among multiple time-oriented patient records. Methods of information in medicine, 48 3, 254-62.

[70] Ola, O., & Sedig, K. (2016). Beyond simple charts: Design of visualizations for big health data. Online Journal of Public Health Informatics, 8.

[71] Sahu, M., & Maringanti, H.B. (2016). Cognitive Architectural Model for Solving Clustering Problem. FICTA.

[72] Andrus, M.R., Mcdonough, S.K., Kelley, K., Stamm, P.L., McCoy, E., Lisenby, K.M., Whitley, H., Slater, N.A., Carroll, D., Hester, E., Helmer, A., Jackson, C.W., & Byrd, D. (2018). Development and Validation of a Rubric to Evaluate Diabetes SOAP Note Writing in APPE. American Journal of Pharmaceutical Education, 82.

[73] Lisenby, K.M., Andrus, M.R., Jackson, C.W., Stevenson, T., Fan, S., Gaillard, P., & Carroll, D. (2018). Ambulatory care preceptors' perceptions on SOAP note writing in advanced pharmacy practice experiences (APPEs). Currents in pharmacy teaching & learning, 10 12, 1574-1578.

[74] [4] Cresswell, S. (2016, October 13). Understanding SOAP format for Clinical Rounds. Retrieved from https://www.gapmedics.com/blog/2015/01/02/understanding-soap-format-for-clinical-rounds/.

[75] Mowery, D., Wiebe, J., Visweswaran, S., Harkema, H., & Chapman, W. (2012). Building an automated SOAP classifier for emergency department reports. *Journal of biomedical informatics, 45 1*, 71-81.

[76] March, S. T. and Smith, G. F. (1995) Design and natural science research on information technology, in Decision Support Systems, 15(4): 251-266.

[77] Hevner, A. R., March, S. T., Park, J., and Ram, S. (2004) Design science in information systems research. In MIS Quarterly 28 (1):75-105.

[78] Norman, K.L. & Schneiderman, B. (2017) QUIS: Questionnaire for UserInteraction Satisfaction. Vol. 2017 University of Maryland.

[79] Marghescu, D., & Rajanen, M. (2004). Evaluating the Quality of Use of Visual Data-Mining Tools.

[80] Lewis, J.R. (2018). The System Usability Scale: Past, Present, and Future. *International Journal of Human–Computer Interaction, 34*, 577 - 590.

[81] Brooke, J. (1996) SUS - A quick and dirty usability scale. Usability evaluation inindustry, 189(194), 4-7.

[82] Bangor, A., Kortum, P. & Miller, J. (2009) Determining what individual SUSscores mean: adding an adjective rating scale. J. Usability Studies, 4(3),114-123.

[83] Brooke, J. (2013) SUS: a retrospective. J. Usability Studies, 8(2), 29-40.