

An integrated modular approach for Visual Analytic Systems in Electronic Health Records

Muhammad Sheraz Arshad Malik¹, Dr.Suziah Sulaiman²

Department of Computer & Information Science
Universiti Teknologi PETRONAS
Bandar Seri Iskandar, 31750, Tronoh, Perak, Malaysia.

Abstract— Latest visual analytic tools help physicians to visualize temporal data in regards to medical health records. Existing systems lack vast support in the generalized collaboration, a single user-centered and task based design for Electronic Health Records (EHR). Already existing frameworks are unable to mentor the interface gaps due to problems like complexity of data sets, increased temporal information density and no support to live databases. These are significant reasons for a single model to comply the end user requirements. We propose an integrated model termed as CARE 1.0 as a future Visual analytic process model for resolving these kinds of issues based on mix method studies. This will base on different disciplines of HCI, Statistics as well as Computer Sciences. This proposed model encompasses the cognitive behavioral requirements of its stake holder's i.e. physicians, database administrators and visualization designers. It helps in presenting a more generalized and detailed visualization for desired medical data sets.

Keywords- Visual analytic Systems; EHR; Information visualization; CARE 1.0.

I. INTRODUCTION

Visual Analytic System (VA) is a combination of automated analysis techniques with information visualizations for an effective knowledge derivations, relationship of data and decisions on very large and complex data sets[5][6][7]. Visual analytic systems help in relating the information to a simple, easy and understandable form for corrugated, adjunct, multi-dimensional and complicated data. Most salient features within a VA system are earlier detection of expected data anomalies, easier understanding of results and accelerate efficiency of decision support for various data sets belonging in medical to engineering domains.

VA systems play a vital role in presenting the Electronic Health Records (**EHR**) for improving health care operations in various countries around the globe. These tools enables the users, designers and collaborators to better understand the temporal and non-temporal queries , differential analysis and validated decisions As data repositories increased both in sizes and intangibilities, existing visual analytic applications reflect poor implementation of Information visualization in terms of exploratory analysis, effective user cognitive representation and user driven process modeling[8][9].

This research work focuses on temporal categorical EHR data including past hospital visits of patients, diagnosis and drugs care plan etc.

Same strategy can be used for temporal numerical data as that of blood pressure readings of patient, pulse rates, heart beats etc. Current EHR systems mostly deal with data entry, retrieval, availability and query based numerical representations. But these systems are lacking the abstract temporal analysis, lesser information density as well as poor exploratory processes. Existing visualization tools Midgaard [14], Lifelines [12][13] , Web-based interactive Visualization Systems[11] and VIE-VISU [9] help physicians to represent EHR data using interactive visualization for different particular scenarios as like.

As in medical health records, visualization application like lifeline2 [12] already tried to solve the various temporal data problems by providing colored triangular dots on a screen. But due to variance in temporal data i.e. both categorical and numerical, it is hard to generalize the requirements of physicians at user interface level. Major reasons are misalignment data formats, mismatch between stake holders requirements and limited scope of datasets. These effects on data integration for analytical representation of information as well as co relation with active live data base systems. As there is a clear difference arises between the user requirements understanding between physicians, data base administrators and visual analysts based on different temporal queries. This leads in generating a potential gap to attain a fully functional analytic system required for easy understandability of information from EHR database.

In this work, a detail studies have been carried out on different available existing VA frameworks to identify user interface problems that hinders understanding of visualization. Misalignment in representation of data fields, results in temporal data exploration issues within existing EHR systems. This provokes in reduced capability in formulation of representation of analyzed data based on the user cognitive experiences, queries input and results exploration. So resulting information details lack with reference to end user perspective as in case of doctors who required particular information for a particular portion of existing record. Existing Visual analytic systems do not provide any kind of live or dynamic support for online databases. This is most importantly needed for latest kind of applications based on reducing the time factor to minimize the information processing for analytical representation in front of physicians.

Our model proposes a simplified and dynamic integrated VA solution from different disciplines e.g medical data base

sets, HCI, Information visualization and Statistical analysis and databases etc for improving health care interactive facilities. This will help physicians to improve the diagnostic approaches for best patient care facilities as well as other users for oncoming needs and trends based on different temporal data repositories. This will not only leverage the service standards in medical fields but also in other disciplines of social sciences for future research and exploration. In addition, it will work to help designers to align a visual analytic system based on collaboration of adjunct user's requirements who can be from different disciplines as well as formulate a base for researchers in information visualization.

This model integrates the stake holders inputs, their information exploring trends for both temporal and non-temporal queries that will help to process the information both offline and online. The resulting framework will help to determine the requirements for a multi varied application tool. This model will facilitate designers for the knowledge management of physicians and medical specialists to align data representation on the basis of their desired EHR data sets. In a normal VA system, physicians demand a visualization based on their requirements using queries, backend queries are prepared in SQL or any database language by DBA and design of visualization is controlled by IV designer. So integration of these stake holders is considered very much important within a single model that is our real research motivation.

A systematic literature review methodology is used to present related work section in this paper regarding user interface problem in existing VA systems. Proposed Model section is going to explain the details of the various portions of this framework. Conclusions and future work sections represent perspective ongoing development in it.

II. RELATED WORK

“A picture is worth a thousand of words”[3] really depicts the values of pictures in any information communication processes and facilitate the understanding of information using pictorial shapes.

Different VA systems frame works studies have been carried out in past few years in Information visualization to better visualize the patient data records. LifeLine is one of the first implemented frameworks that presents patients data like problems, allergies, diagnosis, labs, imaging, medications, and immunizations in the form of lines. Thickness and color of lines represent different conditions of severity as well as termination of any symptoms [13]. Temporal and casual relationships within these facets are not focused within this framework. This was further addressed in using object based scenario within visual representations using Prefuse Toolkit [4]. Dendrogram, Knowledge tables/ Hierarchies, Scatterplot/histograms and Parallel coordinates are used in HCE 3.0 tool to understand the variability in data using clustering algorithmic techniques. This focused on realizing the importance of multi dimensional data visualizations based on clustered gene technology [1] to segregate it on basis of similar gene groups.

Visual data analysis is used to manage meta-knowledge to handle vast amounts of extracted knowledge. Existing

graphical approaches representing association rules in data sets are not offering global and detailed views at same time and that is crucial in HCI [15] proposed prototype CBVAR using FEV (Fish Eye View) focuses on using reduction of association rule sets including generic base sets plus generic association rules in XML file format. This prototype is not providing enough support for user interface contexts for processing user information. However, there is another approach of using Visual analytic systems was proposed on time oriented data concept using computational analysis methods on diversified data[20]. This proposition, still requiring the implementation in task oriented domains as in medical data not all temporal queries can be mapped as time oriented data and it may differ in representing characteristics of time. Similar but different in approach concept in the form of data aesthetics that could be intrinsic or extrinsic were used on VA systems with termed as data focus and mapping formats for representation [21]. This work represented the role of data aesthetics in information visualization representations in VA systems but it lacks the knowledge encompassing aesthetic factors validation.

A sketching-oriented design is proposed for information visualization tools like InfoViz [22]. According to Craft, cartoon like representation is a better and more visualized idea within the visualization schemas to present the complex 2D and 3D designs. A varied set of sketches should be taken regards to be taken and compared as with reference to visualize the temporal categorical data e.g. structure of kidneys elongations or in case of skull 2D views in cases of MRI. Thus a varied set of change in shapes can be presented in the form of already available set of sketches and then it can better be configured with existing databases by designers to match the most suitable data set presentation based on that using any information visualization kit mostly used 2D visualization kits.

InfoVizModel, is another emerging framework is another approach to address the solution of representation of web based data using IV models using Information Architecture (IA)[19]. This work closely related to our present work as it is strongly works in division of small level and big level architecture and currently used as a part for information retrieval from CNKI, Baidu and Google. In this system, information architecture is divided into four phases' navigation, organization, labeling and retrieval where each phase works independently as well as dependently. Former architecture focuses on creation of and management of personal information content and later is responsible for the construction of a webplatform and web content.

Another approach to visualize the web based information was presented in the form of separate APIs in the form of separate stages where one API translates and keeps records of other API[18]. In this way only those icons and data facets are represented that are required by the user as active and inactive icons category. This approach is mostly and widely adopted by Google, Yahoo and MSN web portals in these days for representation of different data sets.

Two stage Visual Analytic framework was proposed by [23] based on combination of different disciplines to solve

heterogeneous data visualization issues. This model tried to sort out major problems in user interface issues, data transformations by using HCI, Computer Sciences, and Mathematics etc based two stages. This work is also an extension of nested model design [16][14][17] and contextual design as combination in separate stages. Designers can derive tangible artifacts based on two requirements a) these should be concrete basis for practical VA system b) Artifacts must be usable by users without introducing cognitive overhead. Primary stage have domain observation, analytical requirements that are further divided into analysis base on each layer i.e user, context,task and organization. Secondary stage refers to user centric refinement that is referring to logging of information about how data generated. This further leads to user pattern analysis and customization that is substituting usage collection, annotation tracking and content sharing and interaction logging.

Last but not the least most closest and up to date work that actually urges to work visualization analytics was the LifeLine2 model as one of the latest approach to address EHR data representation issues. LifeLine2 is providing a model based solution is proposed as one step next version to previous LifeLine due to its increased performance, temporal event based data visualizations and a bit approach to user alignment issues. It does not focus on temporal numerical data so only temporal categorical data e.g past hospital visits, diagnosis etc not the blood pressure readings or such other data sets [1][9][12][8]. One of the key features are granularities with the data sets by keeping them as events to better represent using triangular dots and a graphical user interface based on time stamps.

All the above solutions presented till yet are lacking still lots of areas of research that leads to poor exploratory processes, incomplete information representation and differential visualization. This results in frustration and un necessary delay in representations different perspectives in cases of a complete temporal health record data base that is very much important to improve our existing health care systems not in Malaysia but in the whole world.

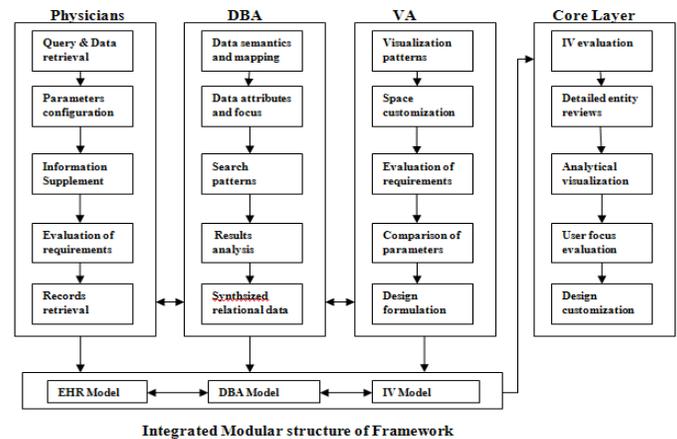
Our conceptual integrated model will focus on solution side of these issues, as in all previous models by the help of integration of VA, Database and physicians query modular structures together used within the latest models to better grasp the visualization for its stake holders.

III. PROPOSED INTEGRATED MODEL CARE 1.0

The newly proposed model is comprised of three major portions that are pertaining to its three major stake holder's i.e. Data base administrators & holders for data sets, Physicians query structures and finally the VA designers using IV modules.

This model is integration of three base models along with one core layer that is providing a base line for visualization on the basis of most important stake holder i.e physicians. As with the previous related works of different researchers in this domain that focuses on each individual portion on single model structure [12][16][21]. This work goes into an extension of these previous work to accumulate the positive

aspects of covering predicted problems by physicians with temporal data i.e both categorical and numerical data. Temporal categorical data sets are tested already been used at limited way to explore the use of IV [12] but still it lacks the ability to cover support to backtrack history of an event of interest in results exploration.



Three components are described separately as mentioned in the above model.

A. EHR Model:

Each EHR data set model associates with the *type of query* for the data it requires e.g. in case of finding the number of patients entered in an Intensive Care Unit (ICU) suffering heart attacks would have a previous history of visiting hospitals, medicine dose recommendations, symptoms and any possible smoking habits or not. So if a physician realizes to utilize the previous record then it's easier to understand the patient history and potential reasons for the attack causes. Parameters configuration, information supplement, evaluation of requirements and records retrieval within the available datasets classification are associated components within this model. As correct parameters settings are required based on personal experience, training and background of a physician and specialist to help in determining the clearest data set in a complex dataset. Information supplement is any additional info associated within a desire query and evaluation of requirements is focusing on the validity of commended operation. Each of these factors effect on the usability of other portions of the model and dependent on each other and if there is any issue occurs that can be tracked on that portion of the model easily.

B. DBA Model:

Data semantics and mapping is mostly carried out on the basis of the temporal queries in already existing VA applications but still these mappings are not considered with reference to search patterns and data focus[9][14][21] i.e either intrinsic or extrinsic. This integrated modular framework presenting the same mapping with relation to search patterns like index searching within the desired datasets selection based on physicians query. As the medical terms, is a different area of understanding for Database teams and people so there is a mutual framework of understanding can be developed in conjunction with continuous feedback and data linkage with previous phase. So database administrator gets

query request clearly understandable to the context what kind of exact data modules are required based on all these previous inter related activities and this will yield a more closer relational data form. This data form will then be testified both at physicians requirements analysis side and after matching to the right demand set it will be processed to formulate a visualization.

C. IV Model:

Each visualization model comprises of its various components that are developed on the basis of various directly and indirectly related factors for its resulting visualization could be a tree, histogram, lines, bars or any match of colors etc. IV model within this framework is comprised of its influencing factors that are characterized as patterns, space, evaluation of requirements, comparison of parameters and design evaluation that are interlinked with each other. Various VA models tried to follow either a few components or fewer integration [4][8][16] but as temporal data sets generate complex variable visualizations that integrations at a wider range is required. Current VA model used within this proposed work is using all these features as integration in a patterned way so that even physicians requirements should not be unaddressed and in the same way it also measures the validity of a visualization existence for a given data set based on its relationship within its own and associated parameters. Most significant feature within this area is the collaboration of this phase with DBA phase as visualization will directly be impacted by the data it requires to formulate that is already associated with resulting query.

D. Core layer:

Core layer is the most integrated and vibrant portion of this model. This is comprising of components that work on validation and testing side of the resultant data, its visualization, analysis linkage and user focus disciplines by the interconnection of all three modular layers of the integrated model. This portion of the model facilitates its all stake holders on user specific areas of exploration, design customization based on different data sets and anomalies. These features help to co relate the inputs and outputs of each portion of the stake holders and provides an inter linkage between them thus trying to remove errors, measure the level of flexibilities within a given data sets and all possible visualizations with maximized level of analysis and co relations of entities.

Normally this portion will help not only the physicians but also the health policy makers to identify the trends analysis of various health issues e.g outbreak of any disease, drug effects on a particular age group in a geographical area or multi root cause analysis of symptoms, diagnosis and suggested measures.

IV. CONCLUSION AND FUTURE WORK

Information visualization in EHRs is one of the significant areas of interest but due to lack in collaboration of its different stake holders that are from different backgrounds it leaves a wider gap for its complete implication. This conceptual model tries to focus on removing the gaps as created by the requirements of physicians that need a kind of visualization

that can help them to fulfill their diagnostic requirements flexibly. While DBAs and Visualization designers are two other different disciplines groups that leave the potential gap of data evaluation based on former's requirements.

This model tried to bring the three stake holders work together by adjuncting their outputs and validating their inputs. Still there are further studies required at micro level conjunctionalities of mapping of user inputs, design restructuring flexibilities side effects and encompassing the variabilities in EHR datasets with reference to exploration spectrum.

ACKNOWLEDGEMENT

Authors will like the opportunity to say thanks to Computer and Information Sciences Department of Universiti Teknologi PETRONAS for providing technical support. Moreover, special thanks to Dr. Suziah Bt.Sulaiman for her guidance and tiring efforts at every instance of this piece of research proposal. We are thankful to The Universiti Teknologi PETRONAS for providing logistic and financial support for this study.

REFERENCES

- [1] J. Seo and B. Shneiderman, "A knowledge integration framework for information visualization," in *From Integrated Publication and Information Systems to Virtual Information and Knowledge Environments*, H. Matthias, et al., Eds., ed: Springer-Verlag, 2005, pp. 207-220.
- [2] P. Hastreiter and T. Ertl, "Integrated registration and visualization of medical image data," in *Computer Graphics International*, 1998. Proceedings, 1998, pp. 78-85.
- [3] D. Pfitzner, et al., "A unified taxonomic framework for information visualization," presented at the Proceedings of the Asia-Pacific symposium on Information visualisation - Volume 24, Adelaide, Australia, 2003.
- [4] J. Heer, et al., "prefuse: a toolkit for interactive information visualization," presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, Portland, Oregon, USA, 2005.
- [5] D. B. N. Christopher, A. Harle and R. Padman, "AN information visualization approach to classification and assessment of diabetes risk in primary care," in *Proceedings of the 3rd INFORMS Workshop on Data Mining and Health Informatics*, J. Li, D. Aleman, R. Sikora, eds., 2008.
- [6] G. A. Daniel Keim, Jean-Daniel Fekete, Carsten Görg, Jörn Kohlhammer, and Guy Melancon., "Visual Analytics: Definition, Process, and Challenges in Information Visualization." vol. 4950. Heidelberg: Springer-Verlag, Berlin, 2008.
- [7] J. C. Granda, et al., "Design Issues in Remote Visualization of Information in Interactive Multimedia E Learning Systems," in *Visualisation*, 2008 International Conference, 2008, pp. 70-76.
- [8] W. Xiaoyu, et al., "A two-stage framework for designing visualanalytics system in organizational environments," in *Visual Analytics Science and Technology (VAST)*, 2011 IEEE conference on, 2011, pp. 251-260.
- [9] W. Horn, Popow, C., and Unterasinger, L., "Support for fast comprehension of ICU data: Visualization using metaphor graphics.," *Method Info. Med*, vol. 40(5);, pp. 421-424, 2001.
- [10] D. L. McGuinness, et al., "Towards Semantically Enabled Next Generation Community Health Information Portals: The PopSciGrid Pilot," in *System Science (HICSS)*, 2012 45th Hawaii International Conference on, 2012, pp. 2752-2760.
- [11] D. S. Pieczkiewicz, Finkelstine, S. M., and Hertz, M. I., "Design and evaluation of a web-based interactive visualization system for lung transplant home monitoring data.," *Proc. Am. Med. Inform. Assoc. Ann. Symp.*, pp. 598-602, 2007.

- [12] T. D. Wang, et al., "Visual information seeking in multiple electronic health records: design recommendations and a process model," presented at the Proceedings of the 1st ACM International Health Informatics Symposium, Arlington, Virginia, USA, 2010.
- [13] C. Plaisant, Mushlin, R., Snyder, A., Li, J., Heller, D., and Shneiderman, B., "Using visualization to enhance navigation and analysis of patient records.," Proc. Am. Med. Inform Assoc., pp. 76–80, 1998.
- [14] R. Bade, Schelchweg, S., and Miksch, S., "Connecting timeoriented data and information to a coherent interactive visualization.," ACM Int Conf on Human Factors in Comp Syst., vol. Proc 22nd, New York, NY, USA., pp. 105–112, 2004.
- [15] O. Couturier, et al., "A scalable association rule visualization towards displaying large amounts of knowledge," in Information Visualization, 2007. IV '07. 11th International Conference, 2007, pp. 657-663.
- [16] T. Munzner, "A Nested Model for Visualization Design and Validation," Visualization and Computer Graphics, IEEE Transactions on, vol. 15, pp. 921-928, 2009.
- [17] H. C. Purchase, et al., "Theoretical Foundations of Information Visualization," in Information Visualization, K. Andreas, et al., Eds., ed: Springer-Verlag, 2008, pp. 46-64.
- [18] K. Matsui, et al., "A Proposal of Framework for Information Visualization in Developing of Web Application," in Applications and the Internet (SAINT), 2011 IEEE/IPSJ 11th International Symposium on, 2011, pp. 457-462.
- [19] W. Jiaxin, "WIVF: Web information visualization framework based on information architecture 2.0," in Computer and Automation Engineering (ICCAE), 2010 The 2nd International Conference on, 2010, pp. 734-738.
- [20] W. Aigner, et al., "Towards a conceptual framework for visual analytics of time and time-oriented data," presented at the Proceedings of the 39th conference on Winter simulation: 40 years! The best is yet to come, Washington D.C., 2007.
- [21] A. Lau and A. V. Moere, "Towards a Model of Information Aesthetics in Information Visualization," in Information Visualization, 2007. IV '07. 11th International Conference, 2007, pp. 87-92.
- [22] B. Craft and P. Cairns, "Directions for Methodological Research in information Visualization," in Information Visualisation, 2008. IV '08. 12th International Conference, 2008, pp. 44-50.
- [23] W. Xiaoyu, et al., "A two-stage framework for designing visual analytics system in organizational environments," in Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on, 2011, pp.251-260.