

ConVIScope: Visual Analytics for Exploring Patient Conversations

Raymond Li* Enamul Hoque† Giuseppe Carenini* Richard Lester‡ Raymond Chau‡

* Department of Computer Science, University of British Columbia

† School of Information Technology, York University

‡ Department of Medicine, University of British Columbia

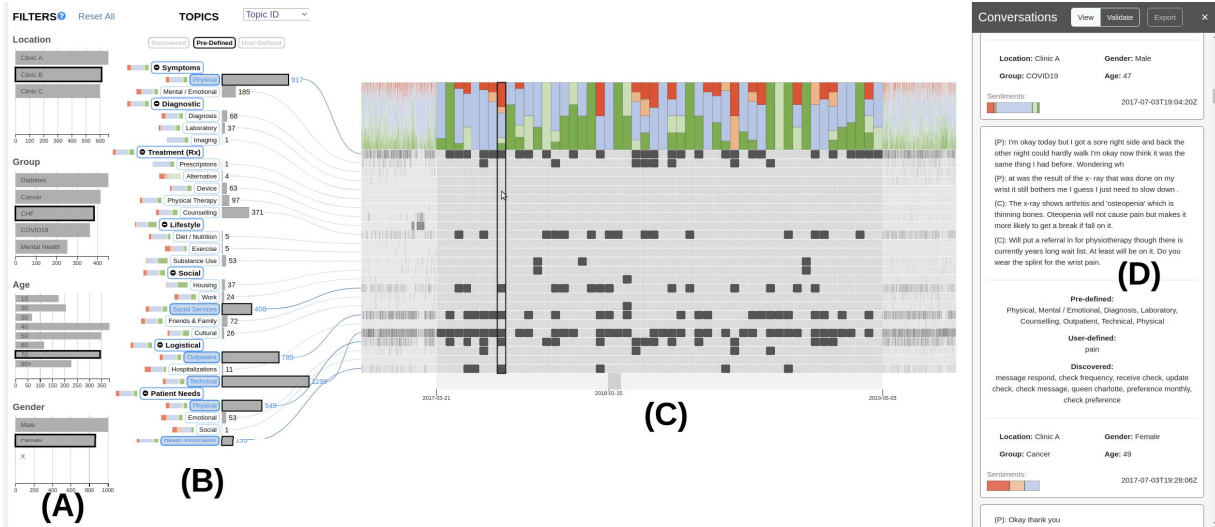


Figure 1: **ConVIScope** Views (see §4): Metadata (A), Topics (B), Analysis (C) and Conversations (D). Trend View shown in Fig. 2.

ABSTRACT

The proliferation of text messaging for mobile health is generating a large amount of patient-doctor conversations that can be extremely valuable to health care professionals. We present ConVIScope, a visual text analytic system that tightly integrates interactive visualization with natural language processing in analyzing patient-doctor conversations. ConVIScope was developed in collaboration with healthcare professionals following a user-centered iterative design. Case studies with six domain experts suggest the potential utility of ConVIScope and reveal lessons for further developments.

Index Terms: Human-centered computing—Visualization—Visualization design and evaluation methods

1 INTRODUCTION

Text messaging has become an important component of mobile health (mHealth) due to its wide adoption and affordable cost [17]. As seen during the COVID-19 pandemic, physicians and health systems worldwide are shifting towards virtualized treatment programs to minimize physical visits and reduce hospital transmissions [41]. In addition, studies involving texting-based mHealth programs have demonstrated their success at lifestyle modification [16], treatment adherence [28, 31], as well as disease monitoring [24].

The conversations generated from texting-based mHealth services can reveal valuable information about the patients and the healthcare quality. For example, at the micro-level health professionals may be interested to quickly get an overview of what patients are discussing about, such as a concern with medications or logistical problems.

*e-mail: {raymondli, carenini}@cs.ubc.ca

†e-mail: enamulh@yorku.ca

‡e-mail: {rlester, raychau}@mail.ubc.ca

At the macro-level, they may be interested to inform policy-making by knowing how conversations vary across different patient groups or to monitor what patients are complaining about in different regions. Visual text analytic solutions could be effective in supporting the above tasks by helping clinicians to discover interesting trends and patterns among the patient population, and assist the health care administrators in optimizing the delivery of services. However, most existing solutions applied Natural Language Processing (NLP) techniques on conversations extracted from online sources such as blogs [19, 20], forums [11, 12], and social media [36, 43, 44]. In contrast, visualizations that focus specifically to the healthcare domain primarily targeted healthcare records (e.g. [10, 25, 33]) or online forums of health communities [26] as opposed to text messages between patients and doctors. Finally, a few commercial mhealth platforms provide static charts for report generation but they do not support interactive visual exploration of a large set of messages [22].

In this paper, we conducted a design study with healthcare professionals to understand how to effectively combine NLP with visualization to support them in discovering useful trends and patterns from patient-doctor conversations and monitoring the delivery of health care services. By following visualization design study methodologies [30, 35], we first characterized the domain by interviewing experts to elicit their requirements. We then performed data abstractions on real datasets obtained from one of our collaborators who is a healthcare researcher. After an iterative design process, we developed **ConVIScope**: a visual text analytic system that supports the exploration and analysis of patient-doctor conversations. Our contributions include: (1) A requirement analysis from interviews with health professionals, which revealed the unique challenges in understanding patient-doctor conversations and guided us to identify user tasks and key design needs. (2) The design and implementation of ConVIScope to meet the identified user requirements. To the best of our knowledge, ConVIScope is the first visual text analytics framework designed specifically for monitoring and exploring a large collection of patient-doctor conversations. (3) An evaluation with six healthcare professionals which provides initial evidence of

the potential utility of ConVIScope and reveals lessons for further developing visual analytics systems for patient-doctor conversations.

2 RELATED WORK

Visualization for Health Care: Several early works on visualizing healthcare data mainly focused on electronic health records (EHR) data (e.g. [6,14,15,33,40,42]). For example, Lifelines [33] organized individual patient records in expandable facets, while Lifelines2 [40] provided a temporal summary of combined records. Lifeflow [42] used icicle trees to organize records into hierarchical structures to reveal general patterns and trends. Others have extracted and visualized text features from clinical data. For example, AdaptEHR [21] extracts medical concepts and maps them onto biomedical ontologies, while HARVEST [18] displays extracted concepts using a tag cloud and a timeline. RetainVis helps users interpret the prediction of neural models on medical records [25]. Unlike all the above works, which focus on summarizing and interpreting structured medical records, we focus on patient-doctor conversations.

Visual Text Analytics for Conversations: With the exponential growth of online conversations, there has been increasing interest in text analytics for such data. Several works focused on visualizing automatically extracted topics and sentiments from conversations [8,9,19,20]. For example, ConVis [19] shows the conversation structure and supports multi-faceted exploration using topics and authors. ConToVi helps to explore speaker behavior patterns in a multi-party conversation using animations [9], while others have focused on organizing a large number of topics into hierarchical tree structures [8,20]. Overall, although the above body of work targets multi-party conversations, they are not designed to deal with specific tasks of health professionals in exploring patient-doctor conversations, as we do in this paper.

In this regard, work studying the visualization of medical conversations is very limited [2,3,28]. For example, Discursis [1] has been applied to provide an overview of a patient-doctor conversation by visualizing the thematic contents as well as dynamics of turn-takings across speakers [2,3]. However, these techniques mainly focus on the conversational structure while ignoring content facets, that we found to be critical in our study. In contrast, VisOHC [26] does visualize sentiments and topics of conversation threads, but from health forums, not from patient-doctor conversations. Also, it only focuses on supporting online health community administrators. Finally, MedStory [38] aims to convey psychological and social aspects of a medical condition, by extracting and visualizing medical concepts and emotions. But again, its target are social media conversations and not patient-doctor messages.

3 TASKS AND DATA ABSTRACTIONS

3.1 Requirement Analysis

In order to gather an initial set of requirements, we performed semi-structured interviews with eight healthcare professionals including 3 practicing physicians, 4 health researchers, and 1 healthcare administrator. The interviews were open-ended in nature, where participants were asked about their needs for analysis and what they need to achieve with the insights derived from conversations with patients. In particular, the primary objectives of the healthcare professionals can be categorized into two scales. On a macro-scale, users require high-level insights across regions and demographics with the goal of designing more effective policies or supporting research objectives. On a micro-scale, users are interested in analyzing conversations from a particular patient group or clinic for identifying specific patient needs to improve the efficiency of the healthcare delivery. Critically, the two scales are not orthogonal, as micro-level analysis can provide context for the macro-level and vice versa. In summary, we organized these goals as three generic requirements on both scales. First, the insights should allow the users to understand and compare the needs of the patients across different regions and demog-

graphics (UR-1). Second, our interface should support evaluating the efficiency and quality of the healthcare services (UR-2). Lastly, it is necessary for users to explore, identify and explain emerging issues and problems faced by the patients (UR-3).

3.2 Task Model

Based on the identified user requirements, we derived a set of analytical tasks that helps the domain experts in understanding and extracting insights from the patient-doctor conversations.

T1. What is being discussed? Participants confirmed the importance of understanding the underlying topics being discussed in order to identify the patient needs (UR-1) and the emerging issues (UR-3) from the conversations. Often, they are interested in analyzing conversations with respect to predefined topics of medical relevance (e.g. treatment, prescriptions). Other times, they are interested in exploring more specific and/or unforeseeable emerging topics beyond the pre-defined ones (e.g. mental problems from COVID-induced lockdown). To support these analytical needs, we extract three types of topics (pre-defined, discovered, and user-defined).

T2. How are the attitudes being expressed? In addition to understanding what was discussed, users are also interested in how sentiments are expressed in conversations, which can be a very useful indicator of patient satisfaction with the quality and efficiency of their healthcare service (UR-2). For example, the attitudes expressed by patients while discussing logistic issues (e.g., outpatient, hospitalization) can provide insights on the efficiency of the healthcare service delivery. Moreover, visualizing the sentiment associated with a topic can help users to make policy decisions; for instance, addressing concerns about side-effects of a particular treatment.

T3. What are the trends for topics and opinions? Understanding the trend associated with topics such as diseases and symptoms can help the users determine whether there are certain problems or events causing the prevalence of such topics (UR-3). For instance, the user might find an increasing number of conversations mentioning financial troubles due to the pandemics. Also, opinion trends can assist the user in assessing whether some aspects of the healthcare quality or patient satisfaction are changing on an aggregated level as well as deriving possible explanations for such changes (UR-2).

T4. How do the conversations differ across different demographics and locations/clinics? Age plays an important role when considering phenotypic changes in health and disease [13], and different patient groups (ex. cancer patients) may have different needs or face different problems (UR-1). Additionally, users are also interested in differences between conversations across clinics and regions, because they can reveal and help explain emerging trends occurring on a macro scale (UR-1, UR-3), providing insights on how to improve health policies that are tailored to those differences.

3.3 Data Model

We pre-process the dataset by filtering out conversations containing less than 3 messages to focus on those that are long enough to perform meaningful studies. Each conversation is tagged with patient features such as location, patient group, age group, and gender. Additional metadata, including topics and sentiment, are derived from text content using the following NLP techniques.

Pre-defined Topics: To satisfy the need of organizing conversations based on general medical knowledge (T1), in close collaboration with health professionals, we define a set of hierarchically organized topic labels (e.g. logistics, treatment, social - see Fig. 1 (B)). We refer to these labels as pre-defined topics. These topics are assigned to conversations by supervised classifiers trained on human-annotated data. We used logistic regression with bag-of-word features for this task as it is a simple and interpretable classifier. Moreover, it has shown promising results in recent work on predicting patient needs from social media conversations, with limited training data [23,27]. We involved healthcare professionals and medical students for annotating 5K conversations to train the

classifiers. To measure the inter-rater reliability, we asked four annotators to label the same set of 200 conversations. The Cohen’s Kappa coefficient was 0.64, indicating a moderate level of agreement [29].

Machine-discovered Topics: In order to discover emerging topics beyond pre-defined topics (T1), we used the standard Latent Dirichlet allocation from the Gensim library [34] for unsupervised topic model [4] with $k = 3$ as the number of topics. The label for each topic comprises the five most likely words given that topic.

User-defined Topics: Predefined topics cover generic medical terms like diagnosis, treatment, exercise etc. However, during an analysis, users may also become interested in more specific topics. ConVIScope allows the user to specify such topics on the fly enabling them to find conversations that contain a user-provided phrase. Such a mechanism also finds conversations with semantically similar phrases to the given one, based on cosine similarity between pre-trained embeddings for their words [32].

Sentiment Analysis: To determine whether each message expresses a positive or negative attitude (T2), we use a lexicon-based method [39], which has been shown to deliver robust performance on conversational data [5]. We predict the sentiment value for each message which ranges from -2 to $+2$.

4 VISUALIZATION DESIGN

In order to support the user requirements and tasks identified above, we iteratively designed ConVIScope over a 9 month period based on biweekly consultations with a team of health care researchers led by an medical specialist. These meetings provided invaluable feedback on refining and revising design choices. The resulting final prototype (see Fig. 1) consists of multiple linked views including an overview summarizing the conversations with sentiment and topics using focus+context, and a detailed view showing the textual conversations. The system also supports filtering by multiple facets namely topics, time range, patient demographics, and locations.

4.1 Visual Encodings

The **Conversation Analysis View (C)** represents a visual summary of the entire set of conversations based on topics and sentiments (T1, T2). It allows the user to explore through the corpus on a conversational level as well as identifying temporal trends and emerging patterns (T3). Unlike ConVis [19] which visualizes a conversation as an indented tree and connects the topics with tree nodes, we use a compact heatmap design to encode the topic distributions, where each column represents a conversation, each row represents a topic (linked to the corresponding topic node in the topic hierarchy), and each cell indicates the presence of a topic using dark grey color. At the top of each column, we use a vertical stacked bar to encode the sentiment distribution of the corresponding conversation where green indicates positive and red indicates negative sentiment.

Overall, the Conversation Analysis View helps users identify how topics and sentiments are distributed across conversations and how they evolved over time in a space-efficient layout. With respect to MultiConVis [20] and VisOHC [26], which can only show a small number of conversations at a time, we improve scalability by designing a novel more compact representation that enables a focus+context approach. We opt for the focus+context technique to avoid the temporal (e.g. zoom) and spatial (e.g. overview+detail) separation by displaying the focus within the context in a single continuous display [7]. The focused window visualizes a constant number of conversations within a fixed window, while the surrounding context (at the left and right side of this window) summarizes the entire set of conversations by shrinking the resolution to fit all the columns inside the view. Finally, we include a separate **Trend View** that can be toggled to replace the Analysis View. This Trend View allows the user to explore how the parent (more general) topics evolve over time, by showing a histogram of the weekly volume of conversations for each topic (Fig. 2 E).

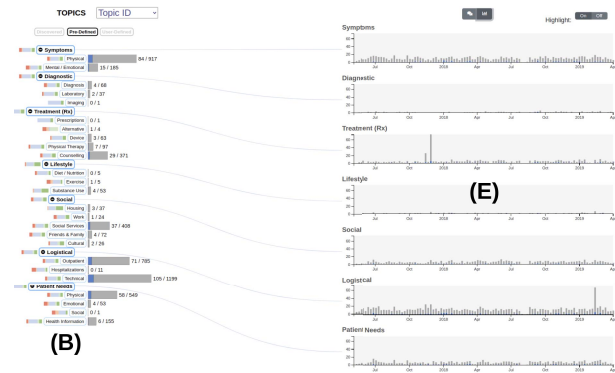


Figure 2: The Trend View (E) visualizes the number of conversations for each parent topic over each week.

In the **Metadata (A)** and **Topic View (B)** we support the visualization of patient features and topics (T1, T4) by using frequency charts to convey the aggregated count for all the conversations containing the attribute. For patient features, we used a bar chart to visualize the number of conversations for each attribute (e.g. location, age, gender), where the attributes are the metadata tags from our data provider. Similarly, the Topic View (B) presents the list of topics where each topic is associated with a bar indicating the frequency of that topic. We also visualize the aggregate sentiment distribution of each topic by placing a horizontal stacked bar to the left side of a topic node to provide a summary of the attitudes from patients encountering specific problems (T2). Topics are organized in a hierarchy using indentation and are connected to the corresponding row in the heatmap via subtle curved links [37]. In this way, the user can perceive which topics are being highly discussed (T1) and follow the associated rows in the Analysis View via the curved links.

The **Conversation View (D)** is a scrollable list that displays the actual text of the conversation. The sentiment distribution, patient features as well as the timestamp for the first message in the conversation are displayed along with the text.

4.2 User Interactions

Browsing through multiple granularity: The Analysis View initially provides an overview of all the conversations in the current dataset. Then, the user can create a focused window containing a constant number of conversations by selecting a region in the scrollbar below the heatmap (C). The user can move through the selection of a region in the scrollbar to change the focus which changes the context prior to and after the focused window. Clicking on any column representing a conversation causes scrolling to the relevant conversation in the Conversation View via a smooth animation. This idiom allows the visualization to provide information at multiple granularities, where the context represents the low-resolution summary, the focused window represents the high-resolution summary, and the most detailed view represents the actual conversation.

Faceted exploration: The user can perform faceted exploration by selecting attributes from the Metadata View and topic nodes in the Topic view. Hovering on any topic node results in highlighting the corresponding rows in the heatmap and the curved links that connect those rows with the hovered topic. Similarly, selecting a column in the heatmap results in highlighting relevant topic nodes in the Topic View for the corresponding conversation. Furthermore, features associated with that conversation are highlighted in the Metadata View by drawing a border around the corresponding rectangle in the bar charts. An example of this interaction is illustrated in Fig. 1, where the Metadata and Topic View respectively shows the corresponding patient features (‘Clinic B’, ‘CHF’, ‘Age 70-80’, ‘Female’) and the topics (‘Physical’, ‘Social Services’, ‘Outpatient’, etc.).

Filtering via Selection: To support the user in selecting the conversations pertaining to specific patient features and topics, we

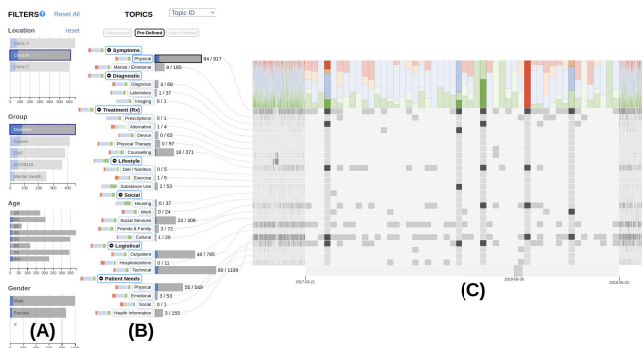


Figure 3: An example of cross-filtering.

employed a cross-filter approach. Specifically, whenever the user clicks on any item in the Metadata and Topic View (e.g. a location, an age group) the interface shows the proportion of conversations containing that selected item across other patient features and topics. For example, when the user selects ‘Female’ under the ‘Gender’ attribute, the proportion of conversations containing ‘Female’ patients are highlighted using the blue colored component of the associated bar of each attribute value. Additionally, the columns representing conversations that do not match the selected attributes are filtered-out (de-emphasized) in the Analysis View. Fig. 3 demonstrates an example of cross-filtering after the user selects three different criteria (“Diabetes patients with Physical Symptoms from Clinic B”).

Interactive Labeling of Topics: Since the predefined topics are classified by the supervised model, sometimes the predictions may be inaccurate. To build more accurate models for predicting predefined topics (T1, T2), we include a feature that allows the user to interactively revise a pre-defined topic assignment based on their domain expertise. By enabling the ‘Validate’ mode in the Conversation View (top (D)), the user can select whether they agree or disagree with the model prediction for a topic assignment based on the text content. At any time the user can export the revised model in a CSV file which can be used in future iterations of model training.

5 EXPERT CASE STUDIES

To assess the efficacy of ConVIScope, we performed case studies with 6 domain experts. The goal was to understand: i) whether ConVIScope help users to perform the tasks identified in Section 3; ii) which visualization features worked and did not work and iii) how to improve the system given the experts’ feedback.

Participants: We ran the study with six domain experts (2 male, 4 female, age range 27-50 years). All participants were health professionals: 2 practicing clinicians (C1, C2) and 4 health researchers (R1-R4). Among them, C1, R1-R4 also participated in the requirement analysis interviews. The participants held expertise in a variety of areas including infectious diseases, clinical psychiatry, public health, and physiology. The clinicians (C1, C2) were interested to use ConVIScope in improving the quality of practice as well as finding issues that patients are facing. Researchers (R1-R4) were more interested in various macro-aspects of patient conversations, such as studying the reasons for hospital admission and evaluating mHealth solutions for COVID-19 monitoring.

Procedure: The participants first went through short interviews, where we asked questions about their goals and the type of insights they hoped to get using ConVIScope. Then, they were provided with a 10-minute tutorial about how ConVIScope works. Participants then accessed ConVIScope to explore a dataset consisting of 5775 conversations between 03/2017 to 04/2021 collected from a real-world healthcare facility by our collaborators. On average, each conversation had 4.5 messages. The task was open-ended in nature where participants explored the dataset according to their own interests and they were free to use as much time as they needed. During the studies, we followed the think aloud protocol where we recorded

all interactions and responses suggesting potential use cases and design feedback. The studies ended with semi-structured interviews, where participants answered questions about the usefulness of individual visualization components and the extent to which their goals were satisfied. All studies were conducted online, where participants shared their screens. Each session lasted about an hour.

Interaction patterns: Reflecting the diversity of their goals and by leveraging the flexible interactions provided by ConVIScope, participants displayed a very rich and diverse set of exploration strategies. Half of them began their exploration using the Analysis View, while the other half began by interacting with the Metadata and Topics Views. In the Analysis View, participants mainly focused on the temporal trend of topics and sentiment by looking at the shrunken column in the surrounding context summaries, while using the focused window to select specific conversations for accessing their content in the Conversations View. In the Metadata and Topic Views, participants often used the cross-filter function, with researchers (e.g. R3 and R1) seemingly more interested in the aggregated counts of the selected conversations and correlations between topics and patient features (e.g. social issues for female cancer patients), while the two clinicians more frequently drilling down on the selected conversations in the Analysis View to determine whether particular topics were relevant for a patient group and then carefully reading the associated conversations.

Subjective feedback: Our analysis of the post-study interviews reveals that participants were impressed with ConVIScope. In particular, they highly appreciate the interface features for filtering conversations by topics and other facets (e.g. demographics and locations) and enabling them to verify what proportion of other facets belong to the filtered set of conversations through blue highlighting in the corresponding bar charts. For instance, R4 said, “*The number of (selected) conversations in the cross-filter can be used to supplement my research*”. Several participants also praised the heatmap component for providing an effective overview of the dataset: “*The (Analysis) View provides a good summary of the dataset, I can get a good idea of what’s being talked about without reading the conversations*” (C1). The ability to drill down to the raw messages was also critical: “*Conversation View is very important, where I verify whether the topics are actually being mentioned*” (R1).

Admittedly, participants also raised some questions and suggested improvements. Multiple participants expressed uncertainty regarding the meaning of each sentiment bin, “*It will be nice to have the criteria for each sentiment color to see how the machine is thinking.*” (R1), “*What exactly is a neutral conversation?*” (R3). Questions were also raised about the support for comparison. For instance, R4 said: “*Although it’s nice to compare conversations from two clinics by individually selecting them, it will be nice to distinguish between the two pooled results (compare the two selected clinics)*”.

6 CONCLUSION AND FUTURE WORK

We present ConVIScope, a visual analytic system that supports the exploration and analysis of patient-doctor conversations. Our system enables users to get an overview of a large set of conversations through multiple linked views, then filter through the set using topics and various metadata before narrow down to detailed messages. Results from our case study are encouraging, but multiple participants expressed the need for more interpretable sentiment analysis and for better support of comparison. These are the two main venues for future work. Further testing of the current prototype is ongoing at four different sites, where health care professionals can analyze their own data in a more ecologically sound setting.

ACKNOWLEDGMENTS

We thank Edward Chiu, Will Choi, Jodi Gunawan, Chris Lee, and Abhishek Singh for their efforts in developing the interface. This work was supported by Michael Smith Foundation for Health Research Award #17273, with match-funding from WelTel Inc.

REFERENCES

- [1] D. Angus, A. Smith, and J. Wiles. Conceptual recurrence plots: Revealing patterns in human discourse. *IEEE TVCG*, 18(6):988–997, 2011.
- [2] D. Angus, B. Watson, A. Smith, C. Gallois, and J. Wiles. Visualising conversation structure across time: Insights into effective doctor-patient consultations. *PLoS One*, 7(6):e38014, 2012.
- [3] R. Baker, D. Angus, E. R. Smith-Conway, K. S. Baker, C. Gallois, A. Smith, J. Wiles, and H. J. Chenery. Visualising conversations between care home staff and residents with dementia. *Ageing and Society*, 35(2):270, 2015.
- [4] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022, 2003.
- [5] G. Carenini, G. Murray, and R. Ng. Methods for mining and summarizing text conversations. *Synthesis Lectures on Data Management*, 3(3):1–130, 2011.
- [6] A. Chetta, J. M. Carrington, and A. G. Forbes. Augmenting ehr interfaces for enhanced nurse communication and decision making. In *Proc. Workshop on Visual Analytics in Healthcare*, pp. 1–6, 2015.
- [7] A. Cockburn, A. Karlson, and B. B. Bederson. A review of overview+detail, zooming, and focus+ context interfaces. *ACM Computing Surveys (CSUR)*, 41(1):1–31, 2009.
- [8] W. Dou, L. Yu, X. Wang, Z. Ma, and W. Ribarsky. Hierarchical topics: Visually exploring large text collections using topic hierarchies. *IEEE TVCG*, 19(12):2002–2011, 2013.
- [9] M. El-Assady, V. Gold, C. Acevedo, C. Collins, and D. Keim. Contovi: Multi-party conversation exploration using topic-space views. *Computer Graphics Forum*, 35, 2016.
- [10] A. Faiola, P. Srinivas, and J. Duke. Supporting clinical cognition: a human-centered approach to a novel icu information visualization dashboard. In *AMIA Annual Symposium Proc.*, vol. 2015, p. 560. American Medical Informatics Association, 2015.
- [11] S. Fu, Y. Wang, Y. Yang, Q. Bi, F. Guo, and H. Qu. Visforum: A visual analysis system for exploring user groups in online forums. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 8(1):1–21, 2018.
- [12] S. Fu, J. Zhao, W. Cui, and H. Qu. Visual analysis of mooc forums with iforum. *IEEE TVCG*, 23(1):201–210, 2016.
- [13] N. Geifman, R. Cohen, and E. Rubin. Redefining meaningful age groups in the context of disease. *Age*, 35(6):2357–2366, 2013.
- [14] M. Glueck, M. P. Naeini, F. Doshi-Velez, F. Chevalier, A. Khan, D. Wigdor, and M. Brudno. Phenolines: phenotype comparison visualizations for disease subtyping via topic models. *IEEE TVCG*, 24(1):371–381, 2017.
- [15] D. Gotz, F. Wang, and A. Perer. A methodology for interactive mining and visual analysis of clinical event patterns using electronic health record data. *Journal of biomedical informatics*, 48:148–159, 2014.
- [16] R. Haider, L. Sudini, C. K. Chow, and N. W. Cheung. Mobile phone text messaging in improving glycaemic control for patients with type 2 diabetes mellitus: a systematic review and meta-analysis. *Diabetes Research and Clinical Practice*, 150:27–37, 2019.
- [17] A. K. Hall, H. Cole-Lewis, and J. M. Bernhardt. Mobile text messaging for health: a systematic review of reviews. *Annual Review of Public Health*, 36:393–415, 2015.
- [18] J. S. Hirsch, J. S. Tanenbaum, S. Lipsky Gorman, C. Liu, E. Schmitz, D. Hashorva, A. Ervits, D. Vawdrey, M. Sturm, and N. Elhadad. Harvest, a longitudinal patient record summarizer. *Journal of the American Medical Informatics Association*, 22(2):263–274, 2015.
- [19] E. Hoque and G. Carenini. Convis: a visual text analytic system for exploring blog conversations. In *Proc. EuroVis*, pp. 221–230, 2014.
- [20] E. Hoque and G. Carenini. Multiconvis: A visual text analytics system for exploring a collection of online conversations. In *Proc. IUI*, pp. 96–107, 2016.
- [21] W. Hsu, R. K. Taira, S. El-Saden, H. Kangarloo, and A. A. Bui. Context-based electronic health record: toward patient specific healthcare. *IEEE Trans. Information Technology in Biomedicine*, 16(2):228–234, 2012.
- [22] S. J. Iribarren, W. Brown III, R. Giguere, P. Stone, R. Schnall, N. Staggers, and A. Carballo-Diegeuz. Scoping review and evaluation of sms/text messaging platforms for mhealth projects or clinical interventions. *Journal of Medical Informatics*, 101:28–40, 2017.
- [23] H. Jang, Y. J. Lee, G. Carenini, R. Ng, G. Campbell, and K. Ho. Neural prediction of patient needs in an ovarian cancer online discussion forum. In *Proc. Canadian AI*, pp. 492–497, 2019.
- [24] A. Kuehne, E. Lynch, E. Marshall, A. Tiffany, I. Alley, L. Bawo, M. Massaquoi, C. Lodesani, P. Le Vaillant, K. Porten, et al. Mortality, morbidity and health-seeking behaviour during the ebola epidemic 2014–2015 in monrovia results from a mobile phone survey. *PLoS Neglected Tropical Diseases*, 10(8):e0004899, 2016.
- [25] B. C. Kwon, M.-J. Choi, J. T. Kim, E. Choi, Y. B. Kim, S. Kwon, J. Sun, and J. Choo. Retainvis: Visual analytics with interpretable and interactive recurrent neural networks on electronic medical records. *IEEE TVCG*, 25(1):299–309, 2018.
- [26] B. C. Kwon, S.-H. Kim, S. Lee, J. Choo, J. Huh, and J. S. Yi. Visohc: Designing visual analytics for online health communities. *IEEE TVCG*, 22(1):71–80, 2015.
- [27] Y. J. Lee, H. Jang, G. Campbell, G. Carenini, T. Thomas, and H. Donovan. Identifying language features associated with needs of ovarian cancer patients and caregivers using social media. *Cancer Nursing*, 2021.
- [28] R. T. Lester, P. Ritvo, E. J. Mills, A. Kariri, S. Karanja, M. H. Chung, W. Jack, J. Habyarimana, M. Sadatsafavi, M. Najafzadeh, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in kenya (weltel kenya1): a randomised trial. *The Lancet*, 376(9755):1838–1845, 2010.
- [29] M. L. McHugh. Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282, 2012.
- [30] T. Munzner. A nested model for visualization design and validation. *IEEE TVCG*, 15(6):921–928, 2009.
- [31] B. K. Ngwatu, N. P. Nsengiyumva, O. Oxlade, B. Mappin-Kasirer, N. L. Nguyen, E. Jaramillo, D. Falzon, and K. Schwartzman. The impact of digital health technologies on tuberculosis treatment: a systematic review. *European Respiratory Journal*, 51(1), 2018.
- [32] J. Pennington, R. Socher, and C. D. Manning. Glove: Global vectors for word representation. In *In EMNLP*, pp. 1532–1543, 2014.
- [33] C. Plaisant, R. Mushlin, A. Snyder, J. Li, D. Heller, and B. Shneiderman. Lifelines: Using visualization to enhance navigation and analysis of patient records. In *Proc. AMIA Symposium*, pp. 76–80, 1998.
- [34] R. Rehurek and P. Sojka. Software framework for topic modelling with large corpora. In *Proc. of the LREC Workshop on New Challenges for NLP Frameworks*. Citeseer, 2010.
- [35] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE TVCG*, 18(12):2431–2440, 2012.
- [36] W. So, E. P. Bogucka, S. Scepanovic, S. Joglekar, K. Zhou, and D. Quercia. Humane visual ai: Telling the stories behind a medical condition. *IEEE TVCG*, 2020.
- [37] M. Steinberger, M. Waldner, M. Streit, A. Lex, and D. Schmalstieg. Context-preserving visual links. *IEEE TVCG*, 17(12):2249–2258, 2011.
- [38] N. Sultanum, M. Brudno, D. Wigdor, and F. Chevalier. More text please! understanding and supporting the use of visualization for clinical text overview. In *Proc. CHI*, pp. 1–13, 2018.
- [39] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede. Lexicon-based methods for sentiment analysis. *Computational Linguistics*, 37(2):267–307, 2011.
- [40] T. D. Wang, C. Plaisant, A. J. Quinn, R. Stanchak, S. Murphy, and B. Shneiderman. Aligning temporal data by sentinel events: discovering patterns in electronic health records. In *Proc. CHI*, pp. 457–466, 2008.
- [41] P. Webster. Virtual health care in the era of covid-19. *The Lancet*, 395(10231):1180–1181, 2020.
- [42] K. Wongsuphasawat, J. A. Guerra Gómez, C. Plaisant, T. D. Wang, M. Taieb-Maimon, and B. Shneiderman. Lifeflow: visualizing an overview of event sequences. In *Proc. CHI*, pp. 1747–1756, 2011.
- [43] Y. Wu, Z. Chen, G. Sun, X. Xie, N. Cao, S. Liu, and W. Cui. Stream-explorer: a multi-stage system for visually exploring events in social streams. *IEEE TVCG*, 24(10):2758–2772, 2017.
- [44] Y. Wu, S. Liu, K. Yan, M. Liu, and F. Wu. Opinionflow: Visual analysis of opinion diffusion on social media. *IEEE TVCG*, 20(12):1763–1772, 2014.