TIEVis: a Visual Analytics Dashboard for Temporal Information Extracted from Clinical Reports

Robin De Croon
robin.decroon@kuleuven.be
KU Leuven
Leuven, Vlaams-Brabant, Belgium

Artuur Leeuwenberg
t.leeuwenberg@gmail.com
University Medical Center Utrecht
Utrecht, Netherlands

Jan Aerts
Marie-Francine Moens
Vero Vanden Abeele
Katrien Verbert
firstname.lastname@kuleuven.be
KU Leuven
Leuven, Vlaams-Brabant, Belgium

ABSTRACT

Clinical reports, as unstructured texts, contain important temporal information. However, it remains a challenge for natural language processing (NLP) models to accurately combine temporal cues into a single coherent temporal ordering of described events. In this paper, we present TIEVis, a visual analytics dashboard that visualizes event-timelines extracted from clinical reports. We present the findings of a pilot study in which healthcare professionals explored and used the dashboard to complete a set of tasks. Results highlight the importance of seeing events in their context, and the ability to manually verify and update critical events in a patient history, as a basis to increase user trust.

CCS CONCEPTS

• **Information systems → Decision support systems**.

KEYWORDS

information visualization, temporal information extraction

ACM Reference Format:

Robin De Croon, Artuur Leeuwenberg, Jan Aerts, Marie-Francine Moens, Vero Vanden Abeele, and Katrien Verbert. 2021. TIEVis: a Visual Analytics Dashboard for Temporal Information Extracted from Clinical Reports. In *26th International Conference on Intelligent User Interfaces (IUI '21 Companion), April 14–17, 2021, College Station, TX, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3397482.3450731

1 INTRODUCTION

Crucial knowledge is often embedded in natural language texts, necessitating human reading and interpretation. However, the vast growth of potentially important information makes it impossible to manually access and process all relevant knowledge residing in these texts. This is also a key issue in the medical domain [4, 10]. It remains a challenge to gain insight in structures hidden in reports, which hinders data-driven patient profiling and stratification.

IUI '21 Companion, April 14–17, 2021, College Station, TX, USA

ACM ISBN 978-1-4503-8018-8/21/04.
https://doi.org/10.1145/3397482.3450731

Incorporation of semantic relationships between words and sentences, including temporal and relational information, is a timely topic within the information retrieval community. A large body of research exists to automatically extract temporal cues from natural language texts [1]. However, the process to accurately combine temporal cues into a single coherent temporal ordering of described events remains an open research challenge. Temporal Information Extraction (TIE) is used in many application domains, such as information retrieval, multidocument summarization, and question answering. TIE also provides opportunities in the clinical domain, such as patient timeline visualizations [6] or patient selection for clinical trials [13]. However, due to suboptimal accuracy, output from TIE models contains uncertainty. Because of the sensitive nature, clinical data should be verified before it can be imported in electronic health records (EHR).

We argue that healthcare professionals should be provided with tools to help them process information stored in clinical reports. They should be supported by natural language processing (NLP) models that lower their workload and provide them with actionable information. However, experts should remain in control [15]. Temporal information is critical in clinical areas [3], yet it is hard for NLP models to *accurately* extract this information from clinical reports. In this paper, we present a visual analytics dashboard [7] that could support healthcare professionals to act as validators. We propose Temporal Information Extraction Visualization (TIEVis) to help healthcare professionals, such as general practitioners, review, process and annotate extracted temporal information from *clinical reports* in a timeline of a patient's condition.

2 TEMPORAL INFORMATION EXTRACTION VISUALIZATION (TIEVIS)

2.1 Uncertainty and NLP Models

Representation and visualization of uncertainty about the timing of events is crucial, especially for automatically extracted information, as in many cases it is impossible to determine an exact event timing based purely on the text (e.g., from a common phrase like *a CT-scan was performed last week* exact timing of the CT-scan is impossible, most-likely the CT-scan was performed on a working day during that week, but it could in principle have been any day of the week). Recently, a set of neural regression models was developed [9] to extract clinical event timelines together with their associated temporal uncertainty: predicting both the most-likely timing of the events and their temporal bounds, based on (1) the document-creation time, (2) temporal expressions (phrases like *yesterday*, *04-08*, or *in two weeks*) that lie close to the events of interest, and (3) the context of the events in the text (e.g., words like *since*, *before*, or *after*). In these models, the clinical text is internally represented by the word-embedding techniques GloVE [11], and ELMo [12], pre-trained for the clinical domain on MIMIC-III, a large dataset of clinical reports of around 40k patients [5]. The full details of the data and models behind TIEVis are described in [9].

2.2 Visual Analytics Dashboard

The ability to get a visual overview of NLP model output is crucial for successful patient data handling. Users need tools to verify output, which can increase users trust [8, 15]. After selecting a clinical report (Fig. 1A), the user is immediately presented with a visual overview - a timeline (Fig. 1B) - of all temporal events detected in the clinical report. Each event is visualized with a lower bound date, a most likely date, and an upper bound date. Color codes are used to create a distinction between different types of events. The report is shown in its entirety for reference (Fig. 1A).

TIEVis supports users to zoom into certain time ranges to check the period before or after a certain event or date. The user can zoom into a certain time range of interest by selecting a period in the timeline. Second, the user can filter types of events and can decide, for example, to only show problems and treatments. When a user hovers over an event - either in the report or in the timeline - the corresponding event is highlighted in the timeline or report.

Timely and relevant information is imperative for data-driven healthcare such as computerized medication monitoring, patient profiling and stratification. For example, for computerized medication monitoring it is important that medication is correctly registered in the health records. Healthcare professionals can visually detect anomalies or mistakes in the timeline.

3 EVALUATION

3.1 Study design

An online evaluation protocol has been setup to evaluate TIEVis with healthcare professionals. We developed a set of seven tasks to solve using TIEVis, while thinking aloud to stimulate the exploration of all functionalities: (1) list main treatments, (2) only show treatments and problems, (3) explain why patient received medication, (4) zoom in on events that happened before admission, (5) manually add treatment, (6) list tests and correct time ranges, (7) explore other texts. Afterwards, they were asked to fill in a familiarity questionnaire of the different concepts; a trust and transparency questionnaire; and a System Usability Score (SUS) questionnaire [2]. The duration of each session was set at 30 minutes + 10 minutes questionnaires. During the evaluation sessions, annotated data from the i2b2 temporal challenge [16] is used [9]. We recruited five healthcare professionals with the assistance of the Academic Center for General Practice and the University Hospital.

3.2 Results

The quantitative results (Fig. 2A) have been augmented with feedback from the think aloud sessions. Transcripts were coded deriving 42 initial codes. After three iterations, five common themes have been derived (see Fig. 2B).

The temporal, visual overview was perceived useful as **context is key**. Participants appreciated that both text and timeline were visualized side-by-side to show events in their context. However, a clinical report does not contain all patient information. An **integration with the electronic health record (EHR)** is therefore needed to visualize the complete patient profile, after all "*a healthcare professional will not use additional programs.*" (P2) They wanted to import the events from TIEVis directly into their EHR as "*those problems are very important to include in the medical file!*" (P3).

Clinical reports are often incomplete. E.g., one text mentioned that a procedure was performed in March 2004, "*so the system cannot know the exact date*." (P1) Since not even a human annotator can reach optimal accuracy, output from TIE models contains **uncertainty**. However, as illustrated in Fig. 2A, healthcare professionals indicated they trusted the visualizations and felt confident using them. They were not suspicious of the output (1 exception - a person who always distrust computer generated output). The basis of this trust could be explained by the highly perceived transparency. They found TIEVis usable (SUS: 79), transparent and do not think it is harmful, nor deceptive. All users understood why the visualization was showing the events like it did, and that it was clear where it was based on. The ability to **manually update and correct events** was important to establish trust: "*Adding additional events is something you should be able to do as a physician.*" (P5). On the other hand, knowing the exact dates is not always needed. "*We often don't know either. [...] It is more important that the event is registered than it is on an exact date.*" (P2) Since "*problem-oriented registration is important*" (P3), the main improvement suggested by three out of five participants is to visualize **relationships**, such as a connection between treatments and associated problems.

4 DISCUSSION

Different methods have been developed for visual representation of electronic medical and patient records [14, 17], but it remains a significant challenge to gain insight in structures hidden in patient data, which severely hinders data-driven patient profiling and stratification. Our proof-of-concept dashboard TIEVis successfully visualized the output of a state-of-the-art NLP model [9] that extracts temporal patient information in time-context.

By putting a human, e.g., healthcare professional, back into the loop to guide the analysis, interactive data visualization has an important role to play [7]. The visual analytics approach relies on interactive and integrated visualizations for exploratory data analysis in order to identify unexpected trends, outliers or patterns. Integrating visual analytics into textual analysis poses its own challenges due to the high information content and the data not fitting the classic types of data encoding. As a result the state of the art of combining integration of user feedback with textual data is very limited. In this study, however, we showed that visual analytics can indeed be successful in increasing the trust of healthcare professionals in NLP models. Our results show that the dashboard is perceived transparent and participants highlighted the importance of remaining in control (all data can be correctly manually from the dashboard).

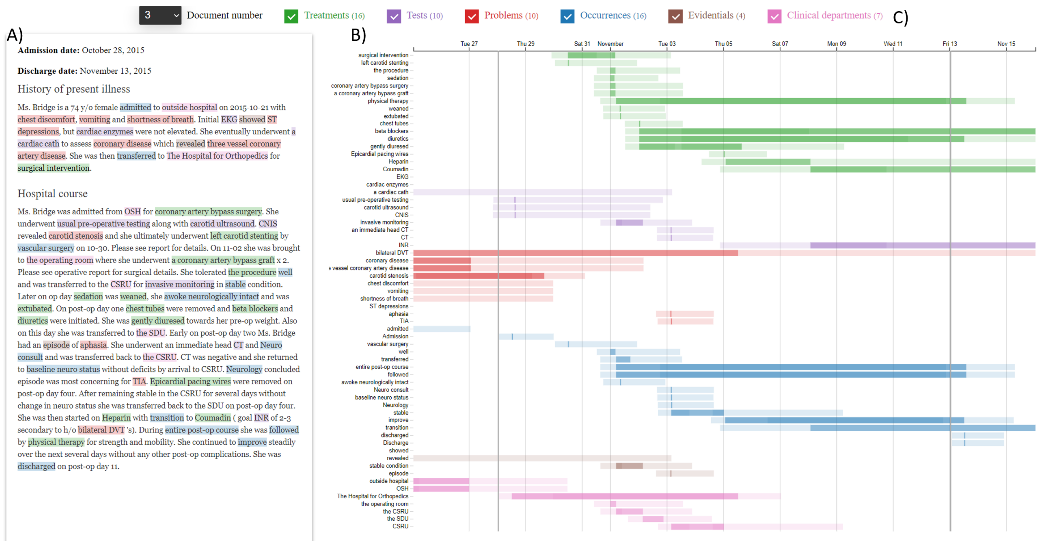

Figure 1: Overview of TIEVis. Clinical events (marked words) are extracted from a clinical report (A) and time periods are estimated for each event. Users can hover over events to bi-directionally highlight them in the report and the visualization (B). An interactive demo is available at: https://augment.cs.kuleuven.be/tievis/

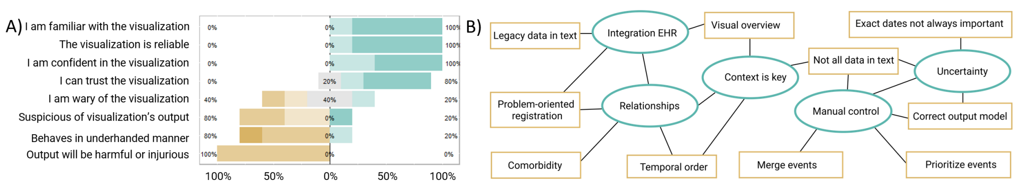

Figure 2: A) Diverging bar chart with the quantitative results on the trust questionnaire. B) Thematic map illustrating the five themes (green) resulting from the simplified codes (orange). EHR = Electronic Health Record.

ACKNOWLEDGMENTS

This work is part of ACCUMULATE, with project number IWT-SBO 150056 financed by Flanders Innovation & Entrepreneurship, and IMPERIUM supported by the Research Foundation Flanders (FWO) under Grant G0A3319N.

REFERENCES

[1] Ghada Alfattni, Niels Peek, and Goran Nenadic. 2020. Extraction of temporal relations from clinical free text: A systematic review of current approaches. *J. Biomed. Inform* 108 (2020), 103488.

[2] John Brooke. 1986. System usability scale (SUS): a quick-and-dirty method of system eval. user info. *Reading, UK: Dig. Equipment Co Ltd* 43 (1986), 1–7.

[3] C Combi and Y Shahar. 1997. Temporal reasoning and temporal data maintenance in medicine: issues and challenges. *Comput. Biol. Med.* 27, 5 (Sept. 1997), 353–368.

[4] R. De Croon, J. Klerkx, and E. Duval. 2015. Design and Evaluation of an Interactive Proof-of-Concept Dashboard for General Practitioners. In *2015 ICHI*. IEEE, Dallas, TX, USA, 150–159.

[5] Alistair EW Johnson, Tom J Pollard, Lu Shen, H Lehman Li-Wei, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi, and Roger G Mark. 2016. MIMIC-III, a freely accessible critical care database. *Scientific data* 3, 1 (2016), 1–9.

[6] Hyuckchul Jung, James Allen, Nate Blaylock, William de Beaumont, Lucian Galescu, and Mary Swift. 2011. Building timelines from narrative clinical records: initial results based-on deep natural language understanding. In *Proceedings of BioNLP 2011 workshop*. ACL, Portland, Oregon, USA, 146–154.

[7] Daniel A Keim, Jörn Kohlhammer, Geoffrey Ellis, and Florian Mansmann. 2010. *Mastering the Information Age - Solving Problems with Visual Analytics*. Eurographics Association, Germany.

[8] B. C. Kwon, M. Choi, J. T. Kim, E. Choi, Y. B. Kim, S. Kwon, J. Sun, and J. Choo. 2019. RetainVis: Visual Analytics with Interpretable and Interactive Recurrent Neural Networks on EMRs. *IEEE Trans Vis Comput Graph* 25, 1 (2019), 299–309.

[9] A. Leeuwenberg and M. F. Moens. 2020. Towards Extracting Absolute Event Timelines From English Clinical Reports. *IEEE/ACM Trans. Audio, Speech, Language Process* 28 (2020), 2710–2719.

[10] Aurélie Névéol, Hercules Dalianis, Sumithra Velupillai, Guergana Savova, and Pierre Zweigenbaum. 2018. Clinical NLP in languages other than english: opportunities and challenges. *J. of biomed. sem.* 9, 1 (2018), 12.

[11] Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. Glove: Global vectors for word repres.. In *2014 EMNLP*. ACL, Doha, Qatar, 1532–1543.

[12] Matthew Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018. Deep Contextualized Word Representations. In *Proc. 2018 North American Chap. of the ACL: Human Language Technologies, Volume 1 (Long Papers)*. ACL, New Orleans, Louisiana, 2227–2237.

[13] Preethi Raghavan, James L Chen, Eric Fosler-Lussier, and Albert M Lai. 2014. How essential are unstructured clin. narratives and information fusion to clin. trial recruitment? *AMIA Jt Summits Transl Sci Proc* 2014 (April 2014), 218–223.

[14] Alexander Rind, Taowei David Wang, Wolfgang Aigner, Silvia Miksch, Krist Wongsuphasawat, Catherine Plaisant, and Ben Shneiderman. 2013. Interactive Information Visualization to Explore and Query Electronic Health Records. *Foundations and Trends® in Human–Computer Interaction* 5, 3 (2013), 207–298.

[15] D. Sacha, H. Senaratne, B. C. Kwon, G. Ellis, and D. A. Keim. 2016. The Role of Uncertainty, Awareness, and Trust in Visual Analytics. *IEEE Trans Vis Comput Graph* 22, 1 (2016), 240–249.

[16] Weiyi Sun, Anna Rumshisky, and Ozlem Uzuner. 2013. Evaluating temporal relations in clinical text: 2012 i2b2 Challenge. *J. Am. Med. Inform. Assoc.* 20, 5 (Sept. 2013), 806–813.

[17] Vivian L West, David Borland, and W Ed Hammond. 2015. Innovative information visualization of electronic health record data: a systematic review. *J. Am. Med. Inform. Assoc.* 22, 2 (March 2015), 330–339.