

AWARD NUMBER: W81XWH-15-2-0016

TITLE: An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data

PRINCIPAL INVESTIGATOR: Dr. Jesus Caban

CONTRACTING ORGANIZATION: The Geneva Foundation
Tacoma, WA 98402

REPORT DATE: July 2020

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE July 2020		2. REPORT TYPE Final		3. DATES COVERED 15 Apr 2015 - 14 Apr 2020	
4. TITLE AND SUBTITLE An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-15-2-0016	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Dr. Jesus Caban Email: jesus.j.caban.civ@mail.mil				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The Geneva Foundation 917 Pacific Ave, Suite 600 Tacoma, WA 98402				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Gay Hayden U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012 Phone 301-619-9883				10. SPONSOR/MONITOR'S ACRONYM(S) USAMRAA	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT We propose to design, develop, and validate an interactive visualization framework that physicians assessing TBI/PTSD patients with comorbid symptoms can use to explore and analyze clinical data and that researchers can use to hypothesize new research questions. The primary aims of this project are to (1) extend our interactive visual analytic framework which combines multiple clinical measurements to allow it to be used to explore large collections of clinical data and (2) validate the effectiveness of such visualization systems among clinicians that treat service members.					
15. SUBJECT TERMS Data Visualization, Health Information Technologies, TBI/PTSD, Open Source Tools					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (include area code)

1. **INTRODUCTION:** *Narrative that briefly (one paragraph) describes the subject, purpose and scope of the research.*

An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data

2. **KEYWORDS:** *Provide a brief list of keywords (limit to 20 words).*

Data Visualization, Health Information Technologies, TBI/PTSD, Open Source Tools

3. **ACCOMPLISHMENTS:** The overarching goals of this project are to (1) address the gap between the acquisition of clinical measurements and the diagnosis step by providing an institutive, flexible, and customizable interactive data visualization framework and (2) validate the system among clinicians treating service members diagnosed with TBI / PTSD.

What were the major goals of the project?

List the major goals of the project as stated in the approved SOW. If the application listed milestones/target dates for important activities or phases of the project, identify these dates and show actual completion dates or the percentage of completion.

The first major goal of this project was to design a visual analytic framework that combines multiple clinical measurements and allows the exploration of large collections of clinical data. The second major goal (years 2-4) was to validate the effectiveness and usability of different visualization techniques for exploring large collections of clinical variables with complex associations.

As part of the first major goal we developed an application to visually summarize large, longitudinal datasets by expanding on our previous work and developing a novel visualization method for analyzing large dataset of longitudinal clinical data. During the second and third year we enhanced the applications by adding new visualization techniques, new validation approaches, and new data. The innovative design was possible by leveraging the latest of the informatics and visualization scientific literature, developing a Necklace interface, and leveraging our initial patient time-line framework. The final product was a validated interface for researchers and clinicians to explore clinical data by leveraging the latest interactive visualization techniques. The application has been used by researchers within the TBI community and specific models expanded to other communities.

As part of the second major goal, we validated the effectiveness of our visualization techniques. We validated the applications by performing machine-learning driven validation techniques in conjunction with focus groups, interviews, and tasks-driven experiments. Through evaluation techniques we have addressed many of the functional needs and challenges associated with analyzing the associations of numerous clinical variables.

Finally, we completed the work by sharing the applications with different stakeholders, applying modules of the framework to other clinical domains, and performing pilot projects in other clinical and non-clinical settings.

What was accomplished under these goals?

For this reporting period describe: 1) major activities; 2) specific objectives; 3) significant results or key outcomes, including major findings, developments, or conclusions (both positive and negative); and/or 4) other achievements. Include a discussion of stated goals not met. Description shall include pertinent data and graphs in sufficient detail to explain any significant results achieved. A succinct description of the methodology used shall be provided. As the project progresses to completion, the emphasis in reporting in this section should shift from reporting activities to reporting accomplishments.

During the first year, four different systems were prototyped and developed to perform visualization of tabular, hierarchical, and longitudinal data. First, **VisXplore** was enhanced to become a clinical data visualization system to perform group or single-subject analysis of multivariate tabular, hierarchical, or temporal clinical data. Second, **CoFlow** was developed as an interactive multi-view and exploratory visualization tool designed to analyze longitudinal EHR data. Third, a **graph-based visualization technique** was developed to visually explore the frequency of patients going from one specific clinical diagnosis to other diagnosis. Finally, a **visual summarization approach** was created and tested with thousands of mTBI patients. Each of the tool has a corresponding draft paper describing the design and techniques. See attachments.

During the second year, two different systems were extensively prototyped and developed to effectively summarize the various data elements that are present in Electronic Health Records (EHRs). First, a novel visualization method, **event summary diagrams**, and a corresponding system were built to enable for a large dataset of events to easily be understood through a top-down interactive exploration. This visualization was evaluated with a dataset of thousands of mTBI patients and shown to reduce the visual complexity and analytical capacity required compared to existing techniques. Second, a **timeline-based framework** for aggregating and summarizing EHRs was extensively researched, designed, and developed to overcome the challenges that exist in EHR systems where data integration is lacking and the disparate nature of data creates difficulties for clinicians. Through this framework, a clinician is able to view the entire history of a patient at multiple time scales and develop an understanding of the patient state over time. Each of these tools have a corresponding draft paper describing the design and techniques. See attachments.

During the third we continued to perform a literature review, developed the Necklace interface, continued to enhance the patient time-line framework, developed an effective interface for research clinicians to search at data, and by creating a systematic approach to analyze the patient interaction data.

During the four year we introduced the application of reinforcement learning onto user interactions within a visual analytic system, opening a new avenue of research for building models that can automatically learn and uncover hidden insights. We outlined three research directions that can be addressed using reinforcement learning and used to test the techniques developed under this grant: (i) providing guidance to users towards the completion of a task, (ii) personalizing guidance to match each individual user's behavior and preferences such that the guidance is "polite" and unobtrusive, and (iii) optimizing the visual analytic interface to maximize user performance. Then we present our approach focused on solving these research directions. We show how our approach builds a model that embeds knowledge about a visual analytic system and how it can automatically learn to tailor itself to a user's preferences through the application of providing guidance. Finally, we manually validate the insights embedded within our model to show how it could be further applied to optimizing an interface. Our reinforcement learning-based models can be utilized with user data or in an automatic standalone method, and ultimately allow for analysis that is not possible using standard modelling techniques.

What opportunities for training and professional development has the project provided?

If the project was not intended to provide training and professional development opportunities or there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe opportunities for training and professional development provided to anyone who worked on the project or anyone who was involved in the activities supported by the project. "Training" activities are those in which individuals with advanced professional skills and experience assist others in attaining greater proficiency. Training activities may include, for example, courses or one-on-one work with a mentor. "Professional development" activities result in increased knowledge or skill in one's area of expertise and may include workshops, conferences, seminars, study groups, and individual study. Include participation in conferences, workshops, and seminars not listed under major activities.

While the project was not intended to provide training and professional development, multiple data analysts offsite were hosted to increase the awareness of the use of data and scientific visualization to explore clinical data.

How were the results disseminated to communities of interest?

If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how the results were disseminated to communities of interest. Include any outreach activities that were undertaken to reach members of communities who are not usually aware of these project activities, for the purpose of enhancing public understanding and increasing interest in learning and careers in science, technology, and the humanities.

The results were disseminated through conference talks, training sessions, focus groups, informatics and visualization symposiums, and scientific papers and publications.

What do you plan to do during the next reporting period to accomplish the goals?

If this is the final report, state “Nothing to Report.”

Describe briefly what you plan to do during the next reporting period to accomplish the goals and objectives.

Nothing to Report.

- 4. IMPACT:** *Describe distinctive contributions, major accomplishments, innovations, successes, or any change in practice or behavior that has come about as a result of the project relative to:*

What was the impact on the development of the principal discipline(s) of the project?

If there is nothing significant to report during this reporting period, state “Nothing to Report.”

Describe how findings, results, techniques that were developed or extended, or other products from the project made an impact or are likely to make an impact on the base of knowledge, theory, and research in the principal disciplinary field(s) of the project. Summarize using language that an intelligent lay audience can understand (Scientific American style).

Now the Defense Health Agency (DHA) has a product to visually explore large and complex longitudinal clinical data.

Describe how the findings, results, or techniques that were developed or improved, or other products from the project made an impact or are likely to make an impact on other disciplines.

The impact of our work is touching multiple disciplines and research domains including clinical informatics, health IT, computer science, medicine, and population health.

What was the impact on technology transfer?

If there is nothing significant to report during this reporting period, state “Nothing to Report.”

Describe ways in which the project made an impact, or is likely to make an impact, on commercial technology or public use, including:

- *transfer of results to entities in government or industry;*
- *instances where the research has led to the initiation of a start-up company; or*
- *adoption of new practices.*

Nothing to report

What was the impact on society beyond science and technology?

If there is nothing significant to report during this reporting period, state “Nothing to Report.”

Describe how results from the project made an impact, or are likely to make an impact, beyond the bounds of science, engineering, and the academic world on areas such as:

- *improving public knowledge, attitudes, skills, and abilities;*
- *changing behavior, practices, decision making, policies (including regulatory policies), or social actions; or*
- *improving social, economic, civic, or environmental conditions.*

This project had an impact that goes beyond science and technology. With the tools developed under this grant providers can more effectively review data for mTBI/PTSD patients to better understand the trajectory a patient is following. By being able to visually explore longitudinal data, providers can now quickly review the trajectory of a patient recovery.

- 5. CHANGES/PROBLEMS:** *The PD/PI is reminded that the recipient organization is required to obtain prior written approval from the awarding agency grants official whenever there are significant changes in the project or its direction. If not previously reported in writing, provide the following additional information or state, “Nothing to Report,” if applicable:*

Nothing to Report

Actual or anticipated problems or delays and actions or plans to resolve them

Describe problems or delays encountered during the reporting period and actions or plans to resolve them.

Nothing to Report

Changes that had a significant impact on expenditures

Describe changes during the reporting period that may have had a significant impact on expenditures, for example, delays in hiring staff or favorable developments that enable meeting objectives at less cost than anticipated.

Nothing to Report

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Describe significant deviations, unexpected outcomes, or changes in approved protocols for the use or care of human subjects, vertebrate animals, biohazards, and/or select agents during the reporting period. If required, were these changes approved by the applicable institution committee (or equivalent) and reported to the agency? Also specify the applicable Institutional Review Board/Institutional Animal Care and Use Committee approval dates.

Significant changes in use or care of human subjects

Nothing to Report

Significant changes in use or care of vertebrate animals

Nothing to Report

Significant changes in use of biohazards and/or select agents

Nothing to Report

6. PRODUCTS: *List any products resulting from the project during the reporting period. If there is nothing to report under a particular item, state “Nothing to Report.”*

- **Publications, conference papers, and presentations**

Report only the major publication(s) resulting from the work under this award.

Journal publications. *List peer-reviewed articles or papers appearing in scientific, technical, or professional journals. Identify for each publication: Author(s); title; journal; volume: year; page numbers; status of publication (published; accepted, awaiting publication; submitted, under review; other); acknowledgement of federal support (yes/no).*

- Filip Dabek and Jesus J Caban “VisXplore: Flexible Visualization System for Analyzing Complex Clinical Datasets”, Workshop on Visual Analytics in Healthcare, ACM Digital Library, Oct. 2014
- Filip Dabek, J. Chen, A. Garbarino, and Jesus J. Caban, “Visualization of Longitudinal Clinical Trajectories using a Graph-based Approach”, Workshop on Visual Analytics in Healthcare, ACM Digital Library, Oct. 2015
- Dabek F., Hoover P., Caban J. (2018) Evaluating Mental Health Encounters in mTBI: Identifying Patient Subgroups and Recommending Personalized Treatments. In: Wang S. et al. (eds) Brain Informatics. BI 2018. Lecture Notes in Computer Science, vol 11309. Springer, Cham
- Dabek F., Hoover P., Caban J. (2018) DeepDx: A Deep Learning Approach for Predicting the Likelihood and Severity of Symptoms Post Concussion. In: Wang S. et al. (eds) Brain Informatics. BI 2018. Lecture Notes in Computer Science, vol 11309. Springer, Cham
- Filip Dabek, Jian Chen, and Jesus Caban, “Visual Summarization of a Collection of Temporal Sequences using Adaptive Frequency Mining and Graph-based Event Modeling” [Draft]
- TrajectoryFlow: Visual Summarization of Temporal Sequences [Draft 2019]
- A Reinforcement Learning Approach to Automatically Model and Learn from User Interactions [Draft 2019]

Books or other non-periodical, one-time publications. Report any book, monograph, dissertation, abstract, or the like published as or in a separate publication, rather than a periodical or series. Include any significant publication in the proceedings of a one-time conference or in the report of a one-time study, commission, or the like. Identify for each one-time publication: author(s); title; editor; title of collection, if applicable; bibliographic information; year; type of publication (e.g., book, thesis or dissertation); status of publication (published; accepted, awaiting publication; submitted, under review; other); acknowledgement of federal support (yes/no).

Other publications, conference papers and presentations. Identify any other publications, conference papers and/or presentations not reported above. Specify the status of the publication as noted above. List presentations made during the last year (international, national, local societies, military meetings, etc.). Use an asterisk (*) if presentation produced a manuscript.

- **Website(s) or other Internet site(s)**

List the URL for any Internet site(s) that disseminates the results of the research activities. A short description of each site should be provided. It is not necessary to include the publications already specified above in this section.

- **Technologies or techniques**

Identify technologies or techniques that resulted from the research activities. Describe the technologies or techniques were shared.

- **Inventions, patent applications, and/or licenses**

Identify inventions, patent applications with date, and/or licenses that have resulted from the research. Submission of this information as part of an interim research performance progress report is not a substitute for any other invention reporting required under the terms and conditions of an award.

- **Other Products**

Identify any other reportable outcomes that were developed under this project. Reportable outcomes are defined as a research result that is or relates to a product, scientific advance, or research tool that makes a meaningful contribution toward the understanding, prevention, diagnosis, prognosis, treatment and /or rehabilitation of a disease, injury or condition, or to improve the quality of life. Examples include:

- *data or databases;*
- *physical collections;*
- *audio or video products;*
- *software;*
- *models;*
- *educational aids or curricula;*
- *instruments or equipment;*
- *research material (e.g., Germplasm; cell lines, DNA probes, animal models);*
- *clinical interventions;*
- *new business creation; and*
- *other.*

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Provide the following information for: (1) PDs/PIs; and (2) each person who has worked at least one person month per year on the project during the reporting period, regardless of the source of compensation (a person month equals approximately 160 hours of effort). If information is unchanged from a previous submission, provide the name only and indicate “no change”.

Name: Jesus Caban, PhD

Project Role: PI

Contribution to Project: Dr. Caban tracked progress, organized meetings, evaluated various visualization techniques for exploring large clinical data, and provided overall vision and direction.

Name: Elizabeth Jimenez

Project Role: Developer

Contribution to Project: Ms. Jimenez implemented the visual analytics framework, in addition to developing and evaluating a visualization technique.

Name: Filip Dabek

Project Role: Investigator (Data Scientist for the National Intrepid Center of Excellence)

Contribution to Project: Mr. Dabek implemented the interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

If there is nothing significant to report during this reporting period, state “Nothing to Report.”

If the active support has changed for the PD/PI(s) or senior/key personnel, then describe what the change has been. Changes may occur, for example, if a previously active grant has closed and/or if a previously pending grant is now active. Annotate this information so it is clear what has changed from the previous submission. Submission of other support information is not necessary for pending changes or for changes in the level of effort for active support reported previously. The awarding agency may require prior written approval if a change in active other support significantly impacts the effort on the project that is the subject of the project report.

Nothing to Report

What other organizations were involved as partners?

If there is nothing significant to report during this reporting period, state “Nothing to Report.”

Describe partner organizations – academic institutions, other nonprofits, industrial or commercial firms, state or local governments, schools or school systems, or other organizations (foreign or domestic) – that were involved with the project. Partner organizations may have provided financial or in-kind support, supplied facilities or equipment, collaborated in the research, exchanged personnel, or otherwise contributed.

8. SPECIAL REPORTING REQUIREMENTS

QUAD CHARTS

9. APPENDICES

An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data



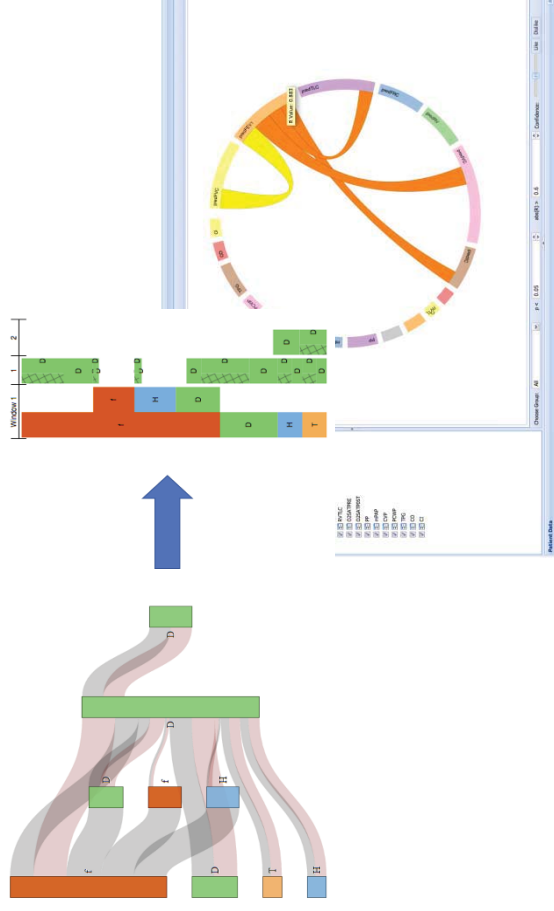
DMRDP

PI: Dr. Jesus J. Caban

Org: National Intrepid Center of Excellence (NICoE) / Geneva

Problem, Hypothesis and Military Relevance

- **Problem:** The large number of evaluation techniques to assess TBI patients poses challenges to clinicians who must integrate disparate measurements to understand the patient's condition.
- **Hypothesis:** In order to successfully analyze a large number of multi-modal clinical variables and integrate different evaluation protocols, new visual analytics techniques and applications are needed.
- **Military Relevance:** The computational models and visual analytical interfaces will use TBI data of military personnel.



Proposed Solution

- **Objective 1:** The design and development of an intuitive framework to visually analyze and explore a large number of clinical variables.
- **Objective 2:** Perform a usability study to validate application with clinical and research staff treating service members diagnosed with TBI/PTSD
- **Objective 3:** Establish a research community. The software application will be shared with researchers and clinicians at different DoD and federal organizations as a validated software application to visually integrate and analyze large number of clinical variables.
- **Budget Expenditure to Date:**
 - Actual Expenditure: \$450,604K **Updated:** 14 July 2020

Timeline and Cost

Activities	FY15	FY16	FY17	FY18	FY19	FY20
R&D visual analytics application						
Compare application with existing techniques						
Perform a usability study						
Validate software application with TBI Data						
Software Integration with other tools and libraries						
Share application						
Estimated Total Budget (\$452k)	\$0	\$75	\$145	\$111	\$87	\$33

Supplemental Material

A word cloud visualization of research trends in TBI research. The words are arranged in a circular pattern, with 'data' and 'models' being the largest and most central. Other prominent words include 'patients', 'techniques', 'predictive', 'measures', 'evaluation', 'support', 'identify', 'different', 'variables', 'used', 'TBI', 'research', 'among', 'patient', 'imaging', 'multi-modal', 'clinicians', 'and/or', 'system', 'analysis', 'complex', 'associations', 'information', 'dependencies', 'willingness', 'trust', 'new'. The colors range from dark blue to light blue.

Award #W81XWH-15-2-0016

Disclosure



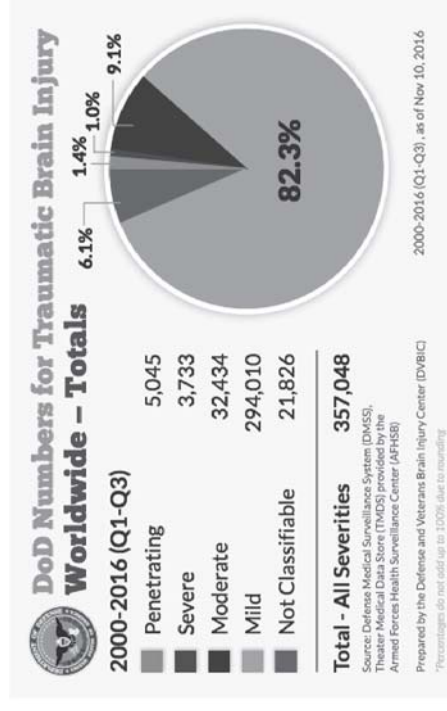
The views expressed in this presentation are those of the authors
and do not reflect the official policy of the Department of
Army/Navy/Air Force, Department of Defense, or U.S. Government.

Project Information



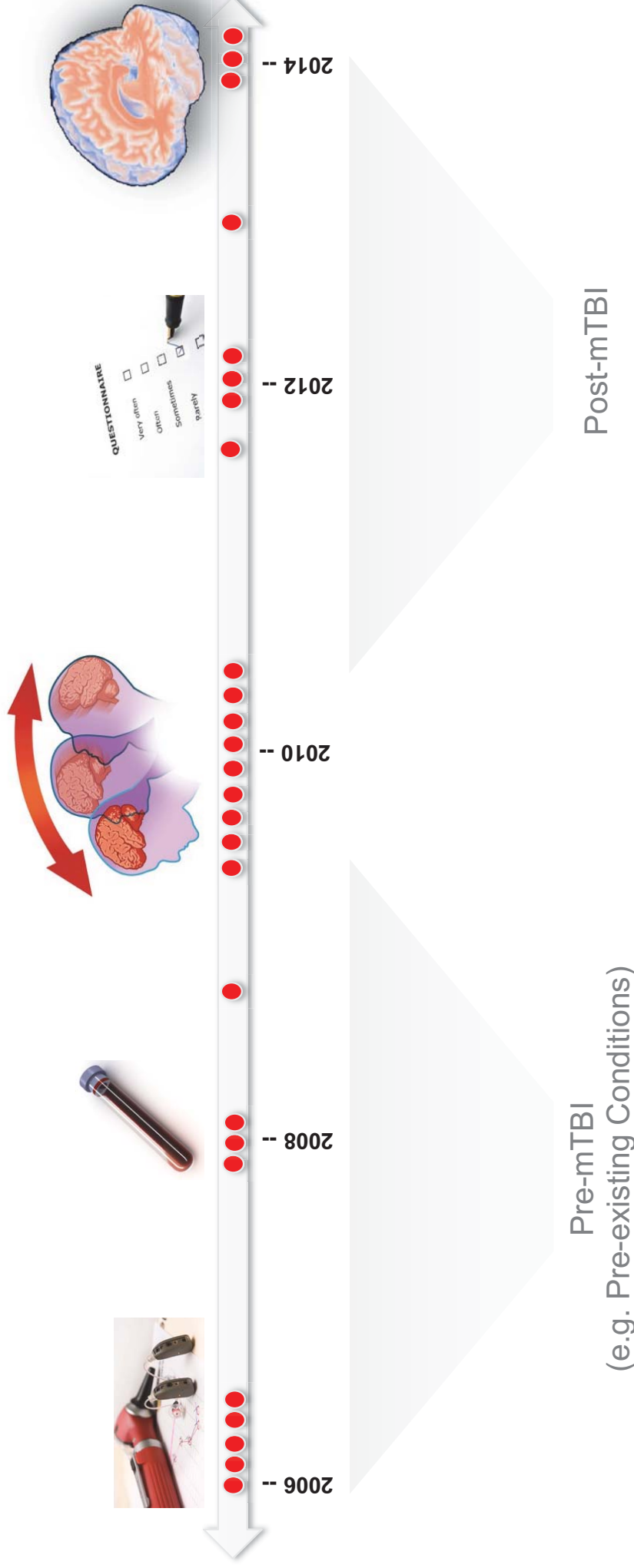
- ❖ **Organization / key partners / institutes:** The Geneva Foundation
- ❖ **Award #:** W81XWH-15-2-0016
- ❖ **Award Mechanism:** Cooperative Agreement
- ❖ **Solicitation:** USAMRMC BAA
- ❖ **Principal Investigator:** Jesus Caban
 - ❖ Key Sub-Awards: (Co-PIs)
- ❖ **Total Cost/Budget:** \$452,332
- ❖ **Period of Performance:** 15 Apr 2015 – 14 Apr 2019
- ❖ **Grants / Contract Officer Representative:** Gay Hayden
- ❖ **Grants / Contract Specialist:** Karen L. Petrore
- ❖ **Related government funding:** N/A

- Traumatic Brain Injury (TBI) continues to be labeled a the signature injury of the wars in Iraq, Afghanistan, and other recent conflicts
- Over 360,000 TBIs occurred in the DoD between 2000 – 20161
 - » ~294,000 are cases of Active Duty mTBI
- mTBI may result in chronic symptomatology
 - » Up to 25% of deployed service members report mTBI symptoms
 - » Following mTBI, most people are expected to recover within a short period of time, however, neurologic, cognitive, and physical symptoms may persist for days to months or longer, requiring ongoing medical treatment



Natural History of TBI

- The National Intrepid Center of Excellence (NICoE) and Intrepid Spirit Centers (ISCs) core mission is to *improve the lives of patients and families impacted by TBI through excellence and innovation*.



NICoE Informatics

Provide services, tools, data, and analysis that can have profound impact in the treatment, management, and overall understanding of TBI.



Solutions Development

- CMT / TBI Portal
- WIIR - TBI
- TBI Document Center
- NICoE Online Referrals
- NICoE Research App



Metrics and Dashboards

- NICoE Dashboards
- NCR/WRNMMC Dashboards
- TBI Dashboards
- Support NICoE Strategic Goals
- Health Outcomes



Data Management

- Data collection
- Data documentation
- Data requests
- Data quality control
- Data Lifecycle control



Informatics Research

- Data Analysis
- Data Modeling (stats & ML)
- Data Mining and Forecasting
- Proposal Development
- USUHS Informatics Partnership



HealthIT

- IT Support
- IT Sustainment
- Data backups
- TeleHealth support

NICoE Informatics

Provide services, tools, data, and analysis that can have profound impact in the treatment, management, and overall understanding of TBI.



Solutions Development

- CMT / TBI Portal
- WIIR - TBI
- TBI Document Center
- NICoE Online Referrals
- NICoE Research App



Metrics and Dashboards

- NICoE Dashboards
- NCR/WRNMMC Dashboards
- TBI Dashboards
- Support NICoE Strategic Goals
- Health Outcomes



Data Management

- Data collection
- Data documentation
- Data requests
- Data quality control
- Data Lifecycle control



Informatics Research

- Data Analysis
- Data Modeling (stats & ML)
- Data Mining and Forecasting
- Proposal Development
- USUHS Informatics Partnership



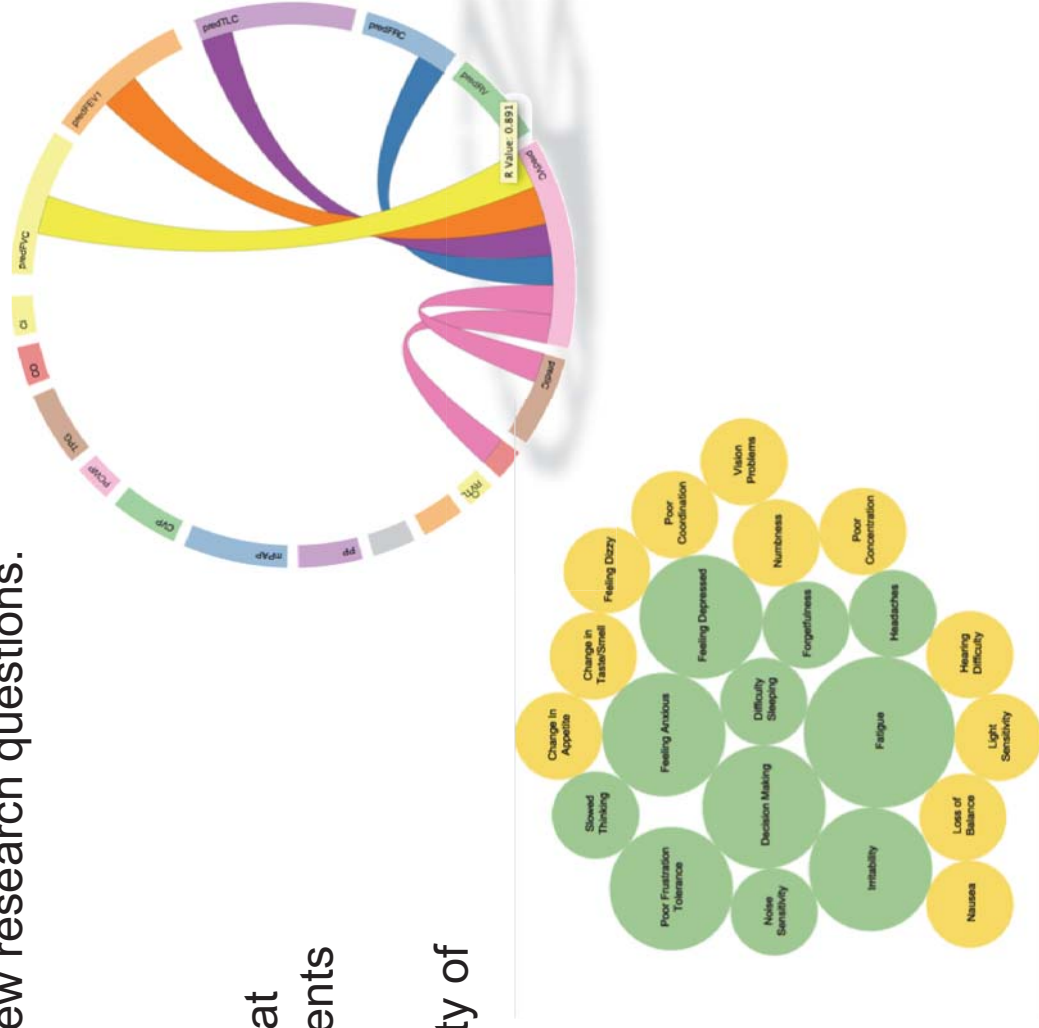
HealthIT

- IT Support
- IT Sustainment
- Data backups
- TeleHealth support



How can we explore clinical data?

- ❖ We propose to design and validate an interactive visual analytics framework that:
 - ❖ clinicians assessing TBI/PTSD patients can use to explore and analyze clinical data
 - ❖ researchers can use to hypothesize new research questions.
- ❖ Primary aims of this project are:
 1. design a visual analytic framework that combines multiple clinical measurements
 2. validate the effectiveness and usability of different visualization techniques



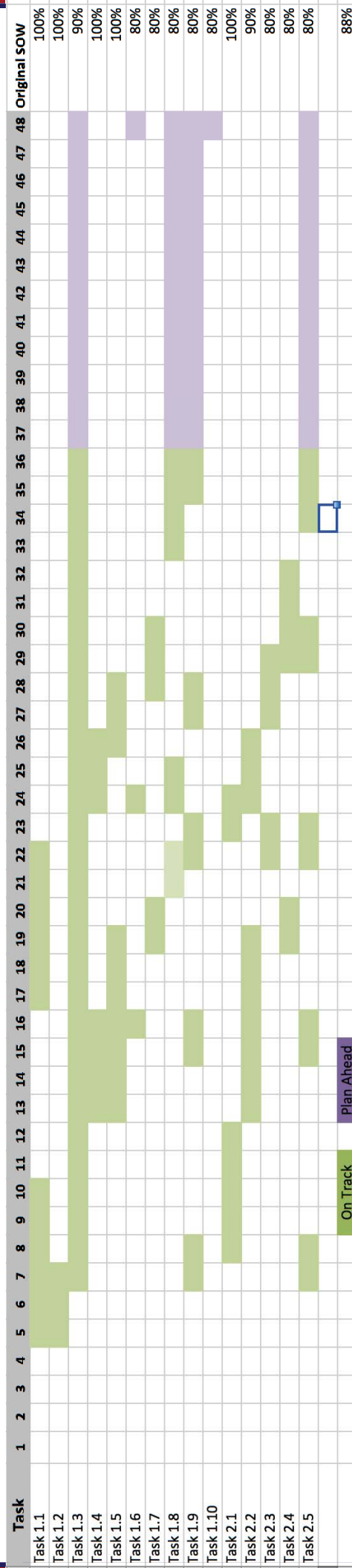
Research Question/Hypothesis & High Level Objectives



- ❖ **Objective 1:** The design and development of an intuitive framework to visually analyze and explore a large number of clinical variables.
- ❖ **Objective 2:** Perform a usability study to validate application with clinical and research staff treating service members diagnosed with TBI/PTSD
- ❖ **Objective 3:** Establish a research community. The software application will be shared with researchers and clinicians at different DoD and federal organizations as a validated software application to visually integrate and analyze large number of clinical variables.



Tasks & Milestone Update



Task	Timeline	Site 1	Site 2
Major Task 1: Design and development of clinical visual analytics framework	Months	NICoE	UMBC
Subtask 1: New software developer and research assistant will be hired and brought up-to speed about our research, the expectations, and deliverables of the project.	1-3	Dr. Caban	Dr. Chen
Subtask 2: Team will compare and evaluate existing data visualization libraries and software tools that could be used to develop a visual analytics framework to explore and analyze clinical data. In particular, Data-Driven Documents, ProtoVis, Flare, and Prefuse will be assessed for their feasibility of extending them to illustrate dependencies among many clinical variables. In addition, software applications such as Tableau, QlikView, and Spotfire will be evaluated.	1-3	Dr. Caban TBD - Software Developer	Dr. Chen TBD - Research Assistant
Subtask 3: Design a modularized visual analytics framework capable of loading, filtering, and illustrating any number of clinical variables.	3-18	Dr. Caban TBD - Software Developer	Dr. Chen TBD - Research Assistant
Subtask 3: Obtain previously collected multi-modal clinical data for 300 mTBI/PTSD subjects and load data into the visualization system. The data will be obtained from the NICoE at no cost to this project.	9-12	Dr. Caban	
Subtask 4: Perform preliminary validation of different visualization techniques	9-15	Dr. Caban	Dr. Chen TBD - Research Assistant
Subtask 5: Prepare and submit research progress report	12	Dr. Caban	Dr. Chen
Subtask 6: Update application based on preliminary validation with clinicians and researchers.	15-16	TBD - Software Developer	TBD - Research Assistant
Subtask 7: Update system based on usability study and suggestions received from users.	20-24	TBD - Software Developer	TBD - Research Assistant
Subtask 8: Share initial framework with different federal organizations including WRNMMC, MHS.	23-24	Dr. Caban	Dr. Chen

Major Task 2: Validation of visualization systems in clinical settings			
Subtask 1: Write and submit expedite IRB protocol to perform usability study	4-8	Dr. Caban	Dr. Chen
Subtask 2: Preliminary validation of the framework with small number of students, clinicians and researchers.	9-15		TBD – Research Assistant
Subtask 3: Perform formal user study of the usability of the system. The system will be tested with over 20 clinicians and 40 students / researchers from Walter Reed and the University of Maryland System.	16-19	Dr. Caban	Dr. Chen TBD – Research
Subtask 4: Analysis of usability study.	20-22	Dr. Caban	Dr. Chen
Subtask 5: Share and/or publish findings of our usability study.	22-24	Dr. Caban TBD - Software Developer	Dr. Chen TBD - Research Assistant

Leveraging existing, approved and non-funded DoD Project for data:

- **Protocol:** DoD IRB #374953
- **PI:** Jesus J Caban, PhD
- **Approval date:** May 30th, 2012 → Dec 2018

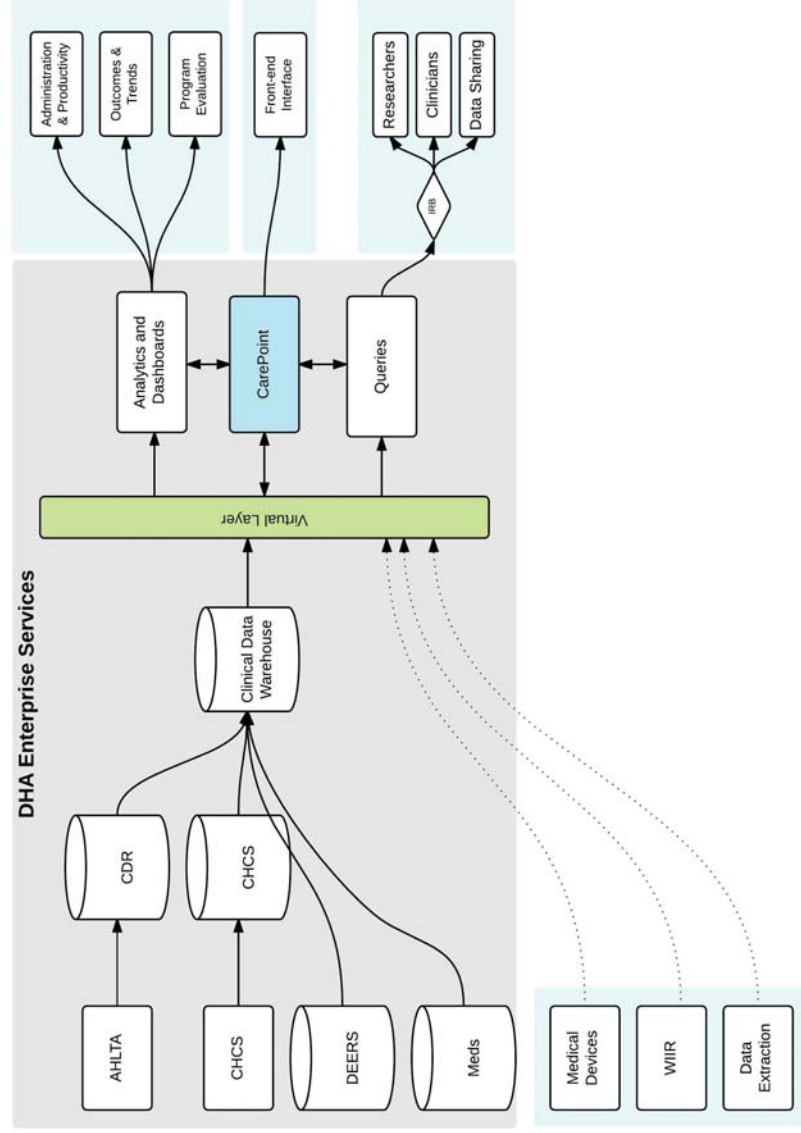
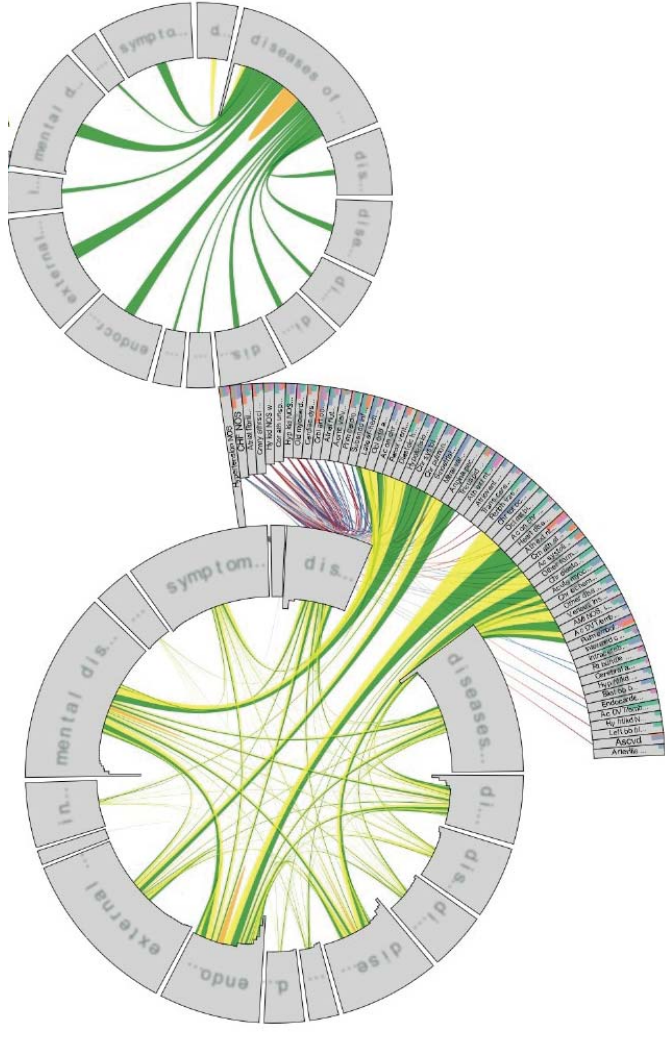


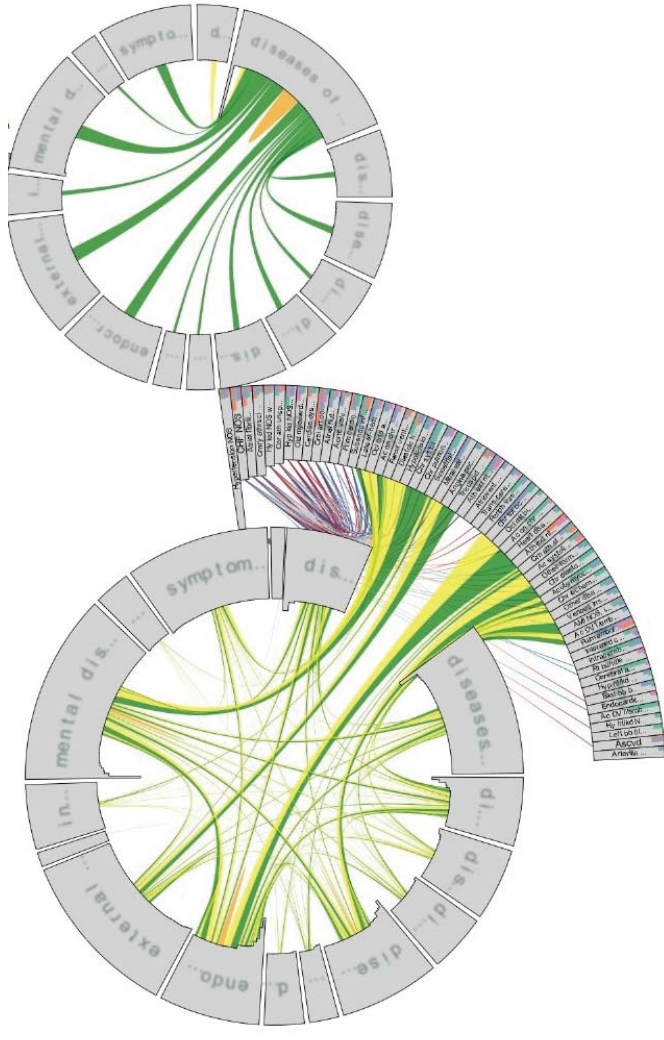
Table #1: Descriptive Statistics of our Dataset

Number of Patients	112,738
Gender	Male 89.6%
	Female 10.4%
Age	18-24 35.8%
	25-34 40.4%
	35-44 18.3%
	45-64 5.4%
	Other 0.1%
Branch of Service	US Army 77,172
	US Marine Corps 15,607
	US Air Force 9,823
	US Navy 9,011
	Other 1,125
SMS with evident War-Related TBIs	44,451
Number of TBI-related Clinical Encounters	17,189,609
Mean number of encounters per patient per year	24
Number of Medication Transactions	9,964,413
Number of Laboratory Results	3,664,461
Number of Radiology Notes	1,123,196

1. 1D/Linear
 - lists of data items, organized by a single feature (e.g., EKG)
2. 2D/Planar (X-Ray, Geospatial)
 - Geographical distribution of a specific disease
3. 3D/Volumetric
 - medical imaging (CT, MRI, etc....)
4. Temporal
 - T=time series
5. nD/Multidimensional
 - category proportions (e.g. Neuroproteomics)
6. Tree/Hierarchical
 - ICD codes, many medical tests
7. Network
 - Clinical research



1. 1D/Linear
 - lists of data items, organized by a single feature (e.g., EKG)
2. 2D/Planar (X-Ray, Geospatial)
 - Geographical distribution of a specific disease
3. 3D/Volumetric
 - medical imaging (CT, MRI, etc....)
4. Temporal
 - T=time series
5. nD/Multidimensional
 - category proportions (e.g. Neurop
6. Tree/Hierarchical
 - ICD codes, many medical tests
7. Network
 - Clinical research



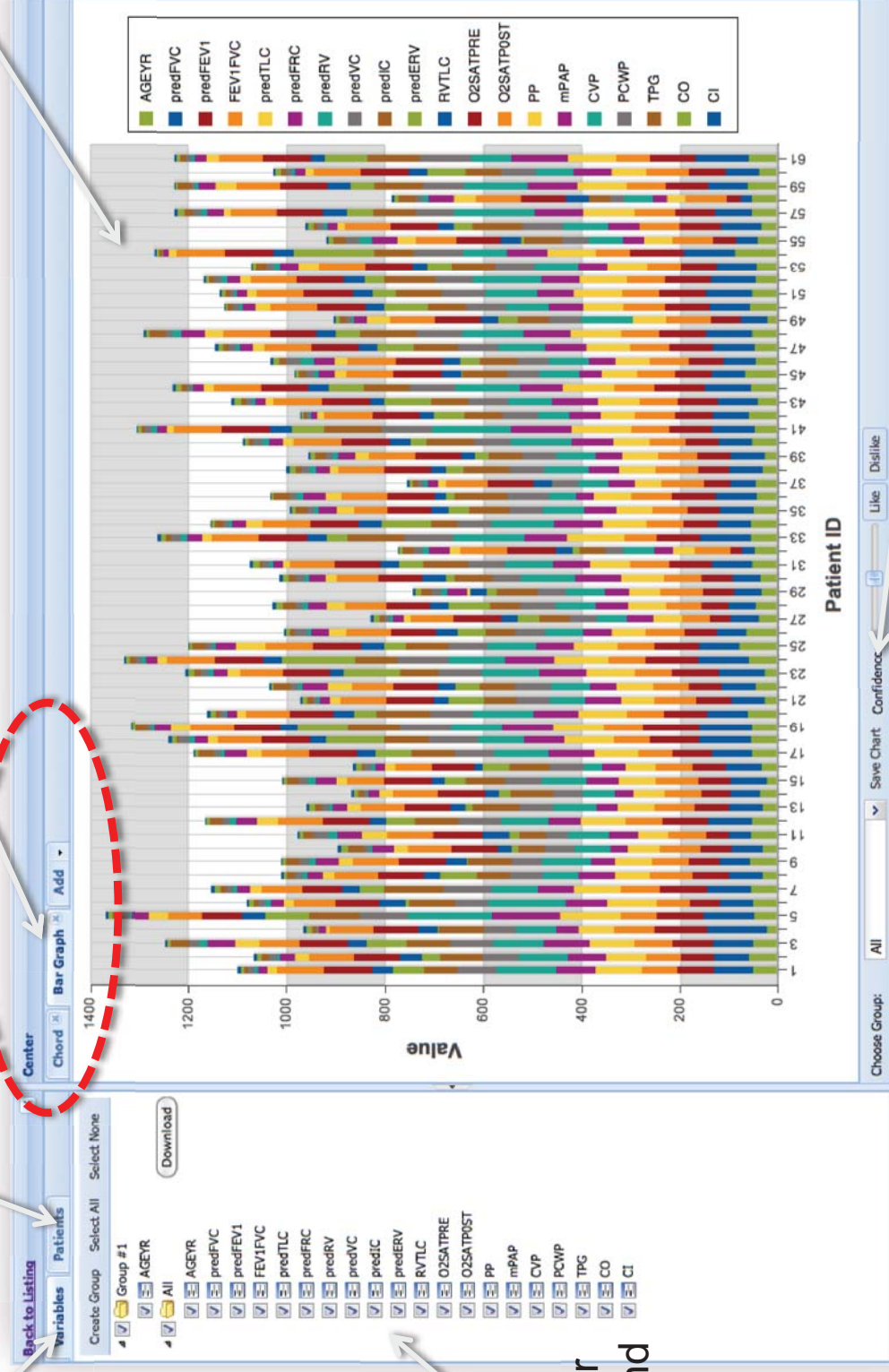
Results - Last Year

List of
Patients

Modules

Clinical
Variables

Chart



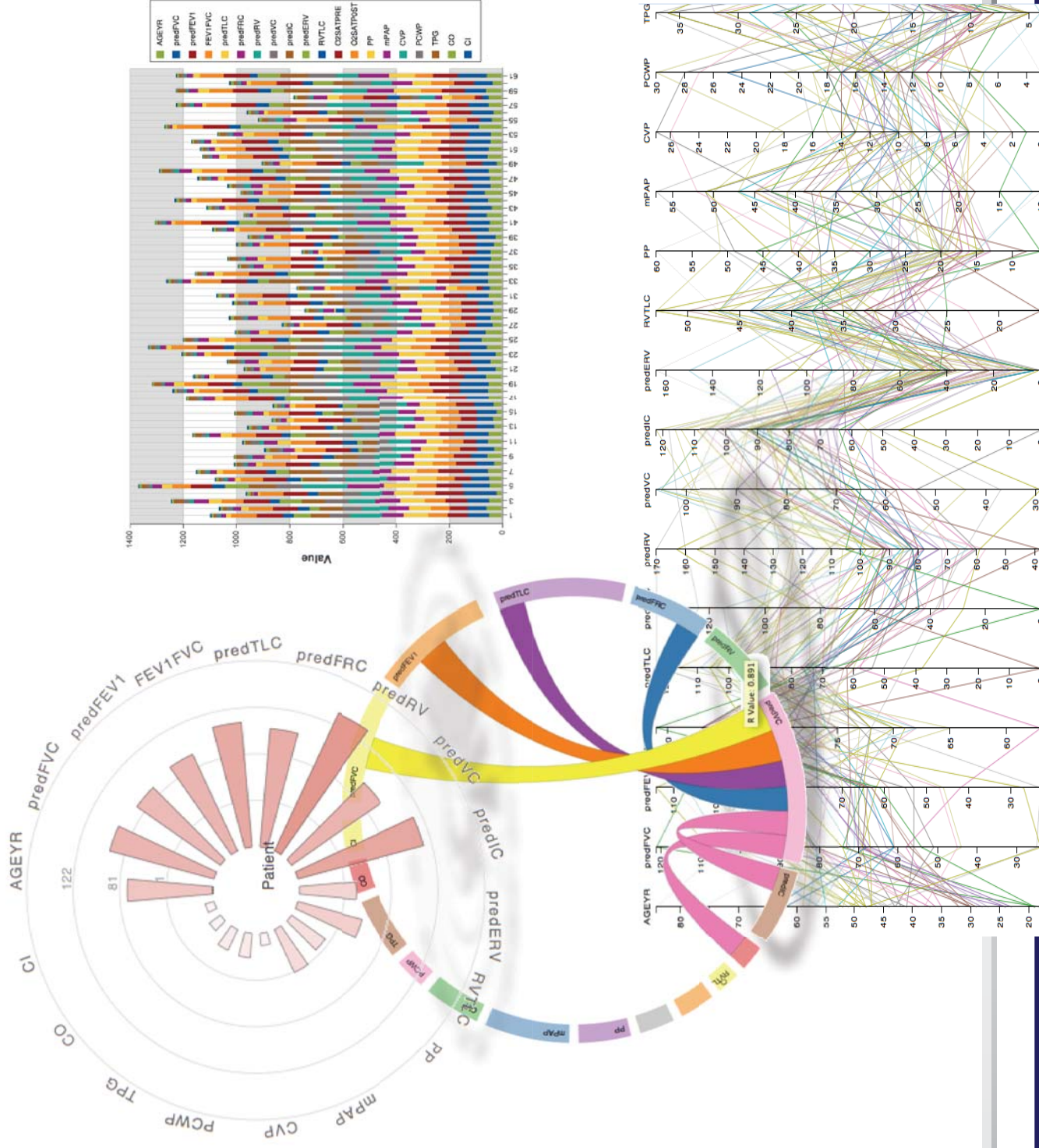
Controls for
Variables and
Patients

Controls for Chart

Results - Last Year



1. Parallel
2. Chord
3. Windrose
4. Lines
5. Statistics
6. Area Chart
7. Histogram
8. Pie Chart
9. Bar Graph
10. Radar Fill



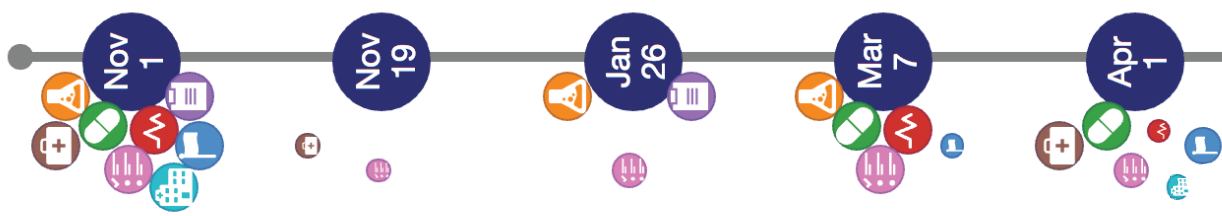
Visually Exploring EHR data for a Single Patient

Physical findings: **Normal**
Vital signs: **Normal**
* See signs: (Reviewed)

ABWHEN IN USER TEST ON 17 JUL 2008 12:24
It was noted that the patient had been
Comments: Schedule retest with PCM and request PCM well cholesterol medication to cover patient's needs.
Laboratory:
LIPID PROFILE (Routine) Result: Result Status
CHOLESTEROL (Total) Result: 160 mg/dL
LIPID PROFILE (Routine) Result: Result Status
HDL CHOLESTEROL (Total) Result: 35 mg/dL

Timeline Construction

- Represent the patient's longitudinal history as a timeline.
- Encodes multiple attributes to a single encounter



Visually encode clinically relevant information into each node.

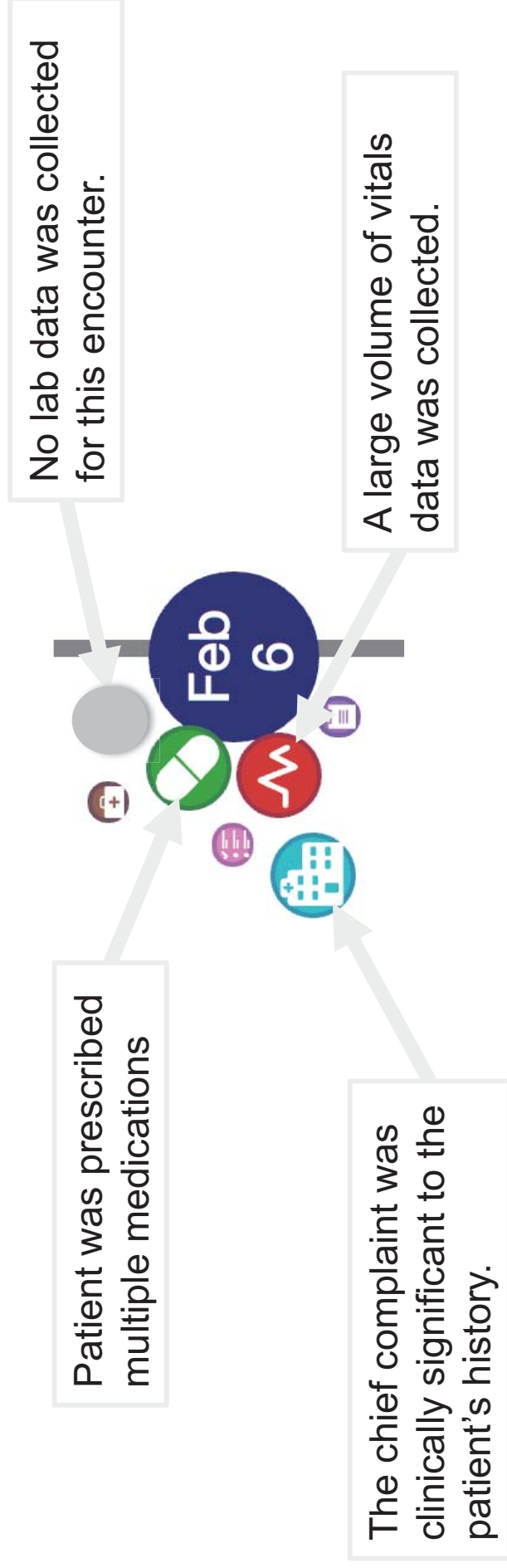
Summary Node

- Encode information about the clinical data elements into attribute nodes:
 - » Lab Test Results
 - » Medications
 - » Vitals
 - » Clinical Note
 - » Diagnosis Codes
 - » Procedure Codes
 - » Radiology Notes



Attribute Nodes

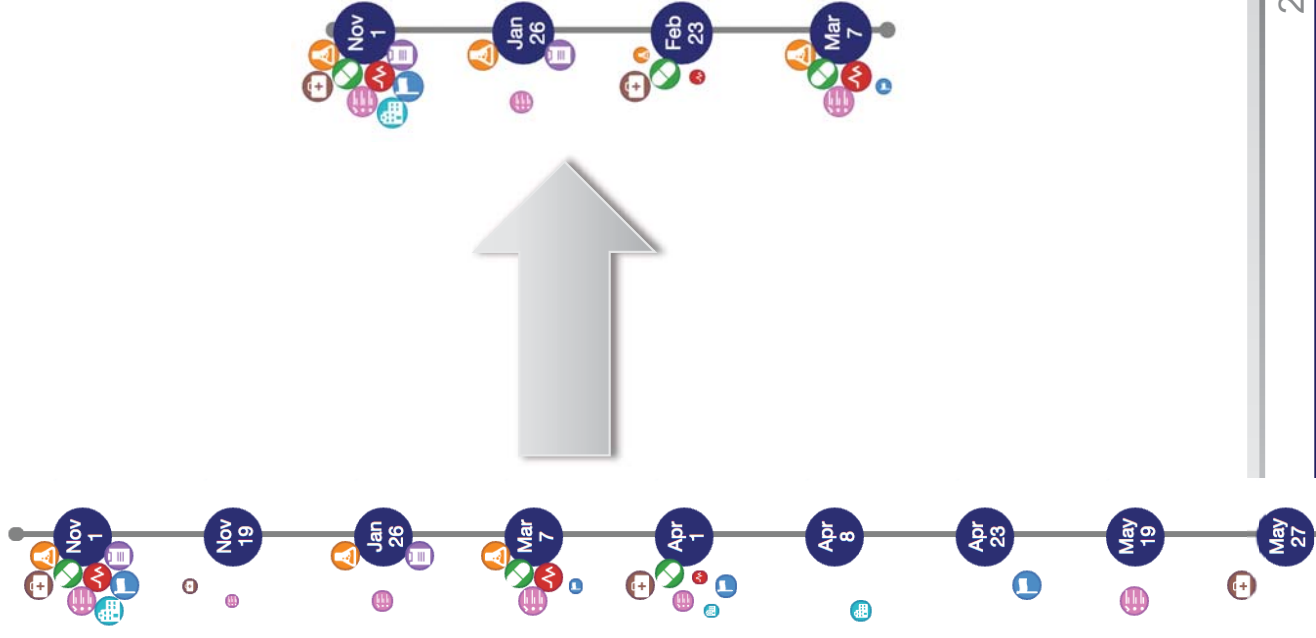
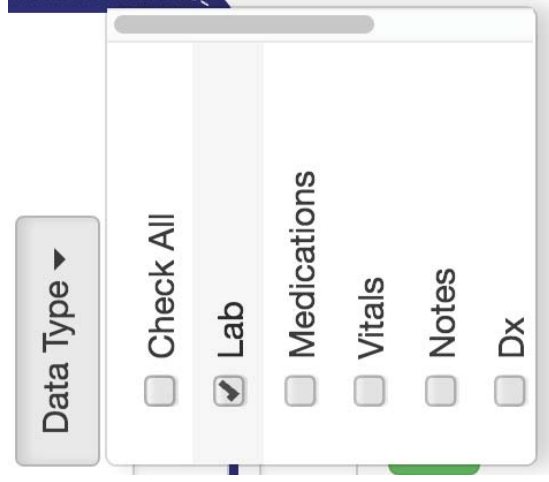
- Size of each attribute node is varied based on:
 - » Volume of data collected for the encounter
 - » Clinical significance of the captured data.



Filtering Example #1



- Only show encounters that contain lab data.



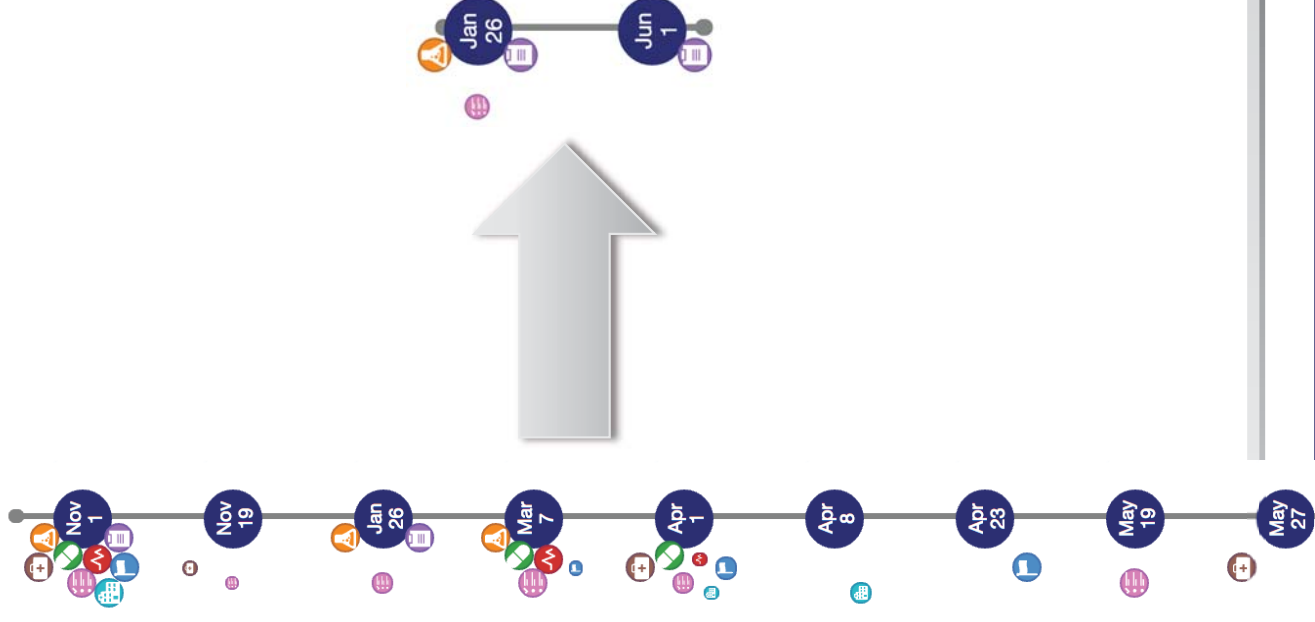
Filtering Example #2



- Only show encounters that occurred with a specific type of provider.

Provider Type ▼

☐ Check All
☐ Audiologist
☐ Neurologist
☐ Physical Therapist
☒ Psychiatry
☐ Speech Pathologist



Timeline Interface

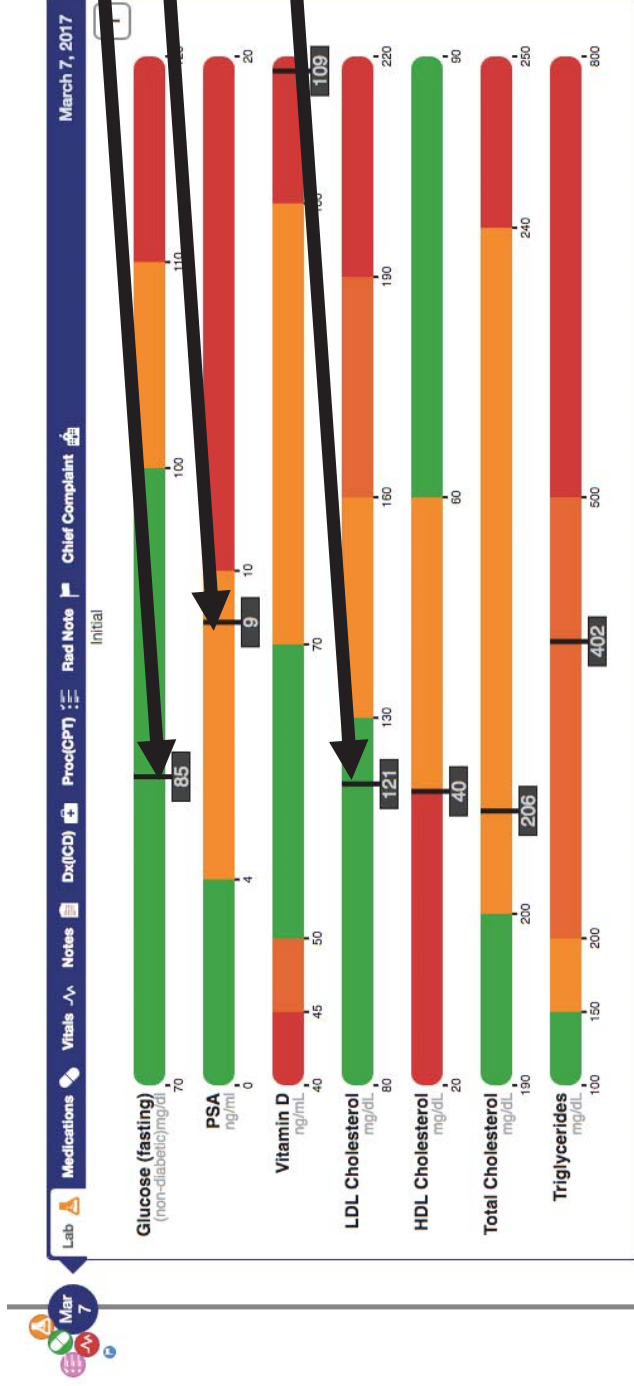


- Allows providers to find the encounter that they want to look at.
- Need to:
 - » Allow to drill down further.
 - » Utilize visualization to communicate each attribute.



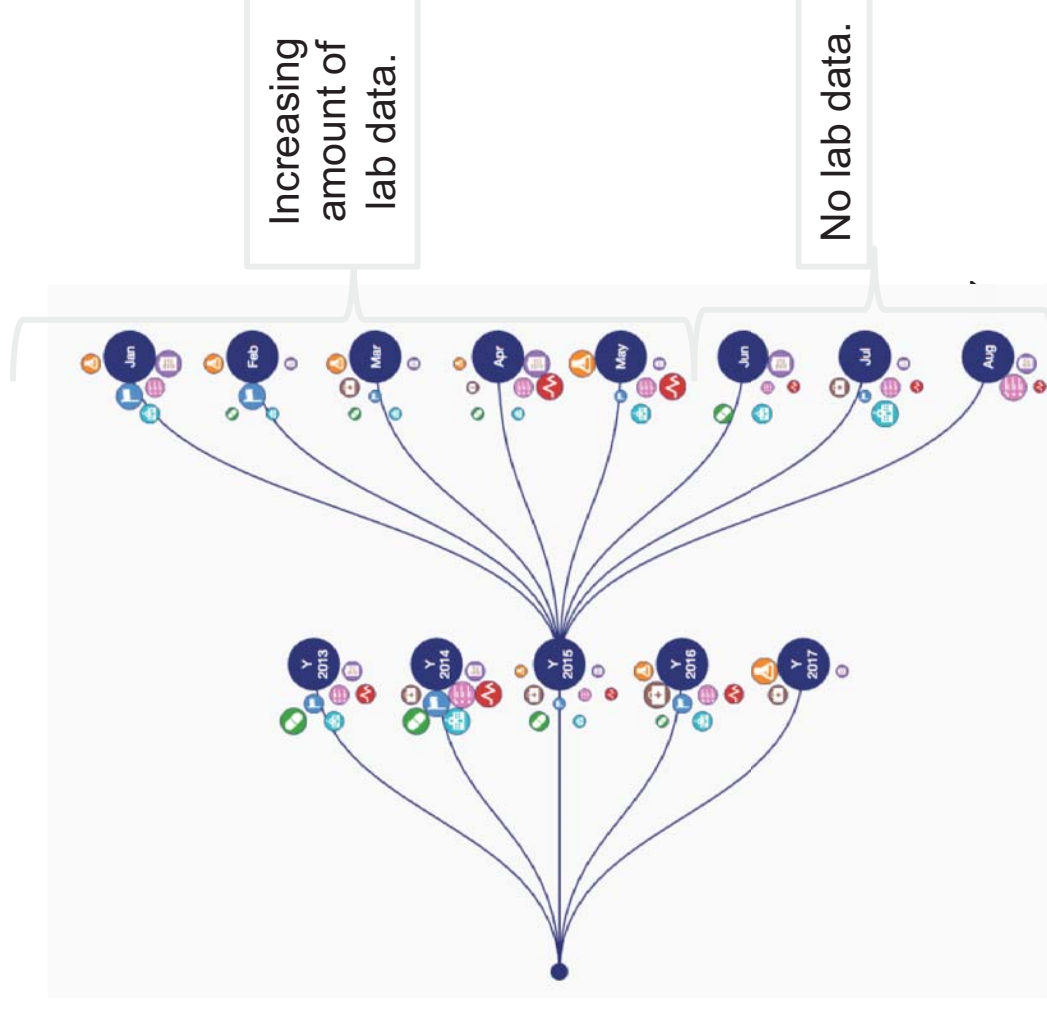
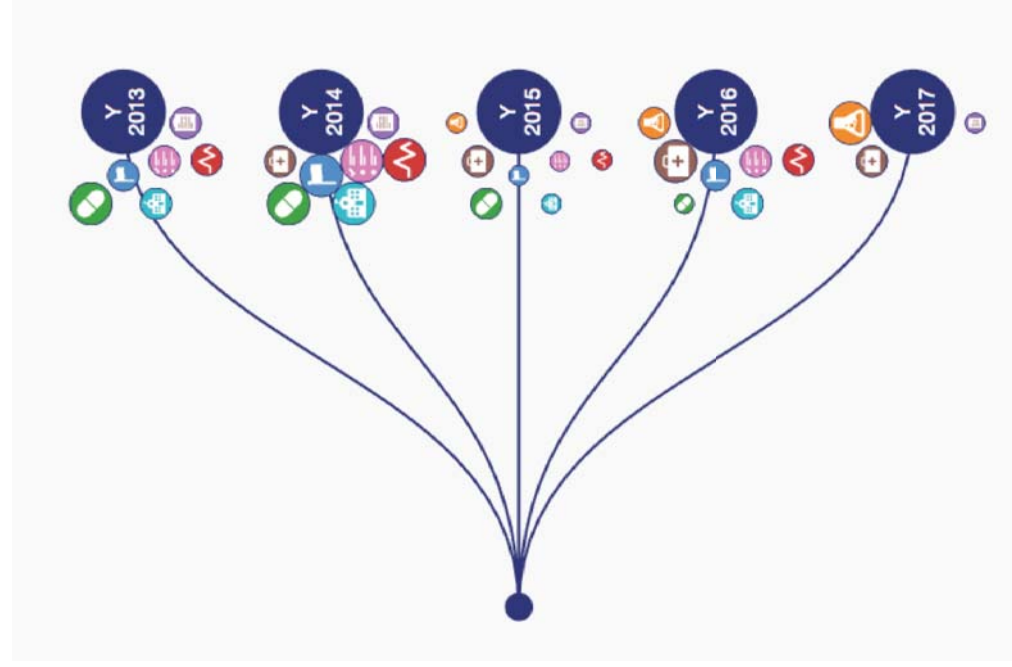
Data View - Labs

- Visualization of lab tests conducted for a given encounter.
- Color scale indicating whether patient is within normal range.

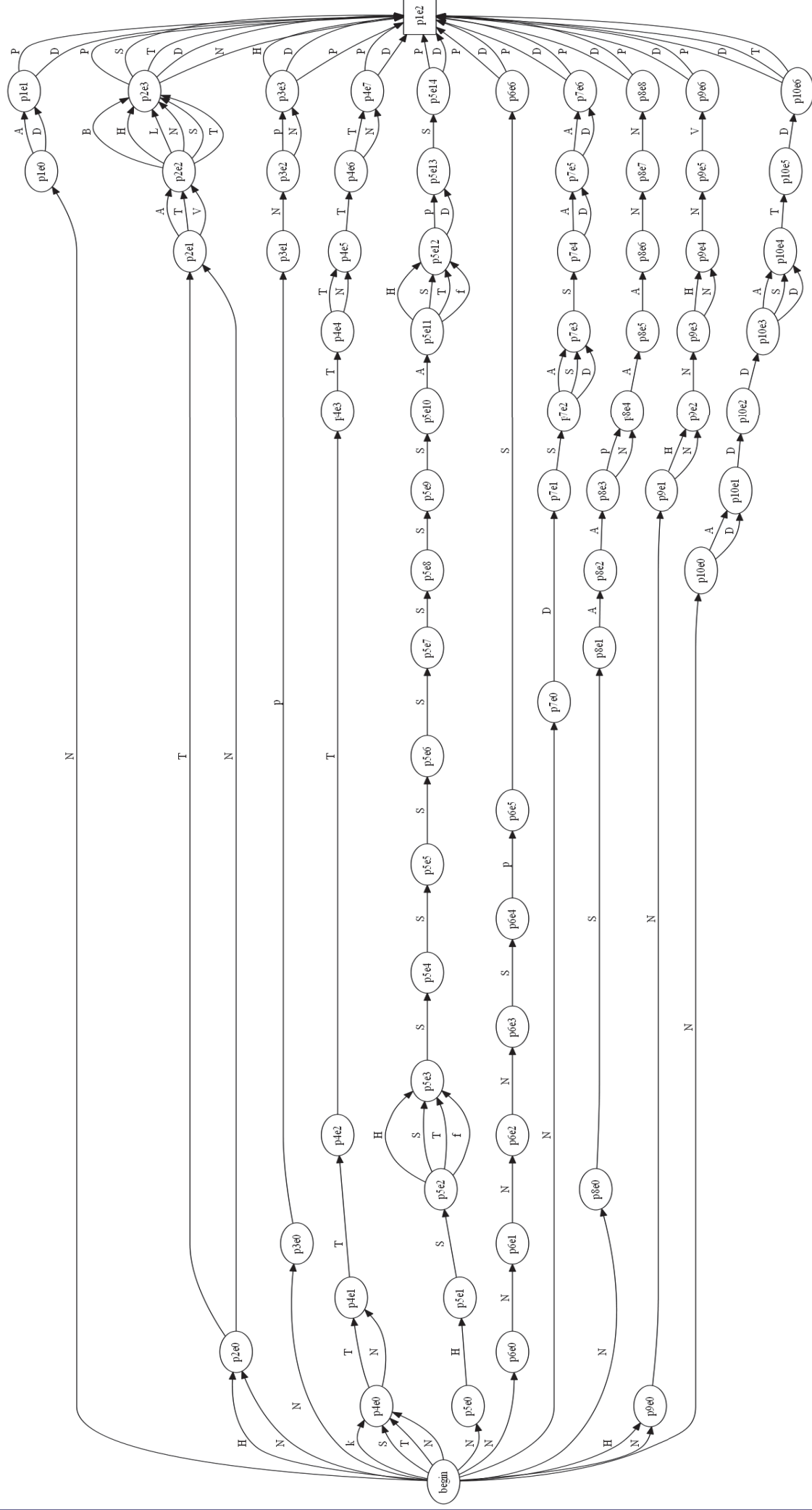


Markers indicating the patient's measured value against the possible range.

Patient Summary

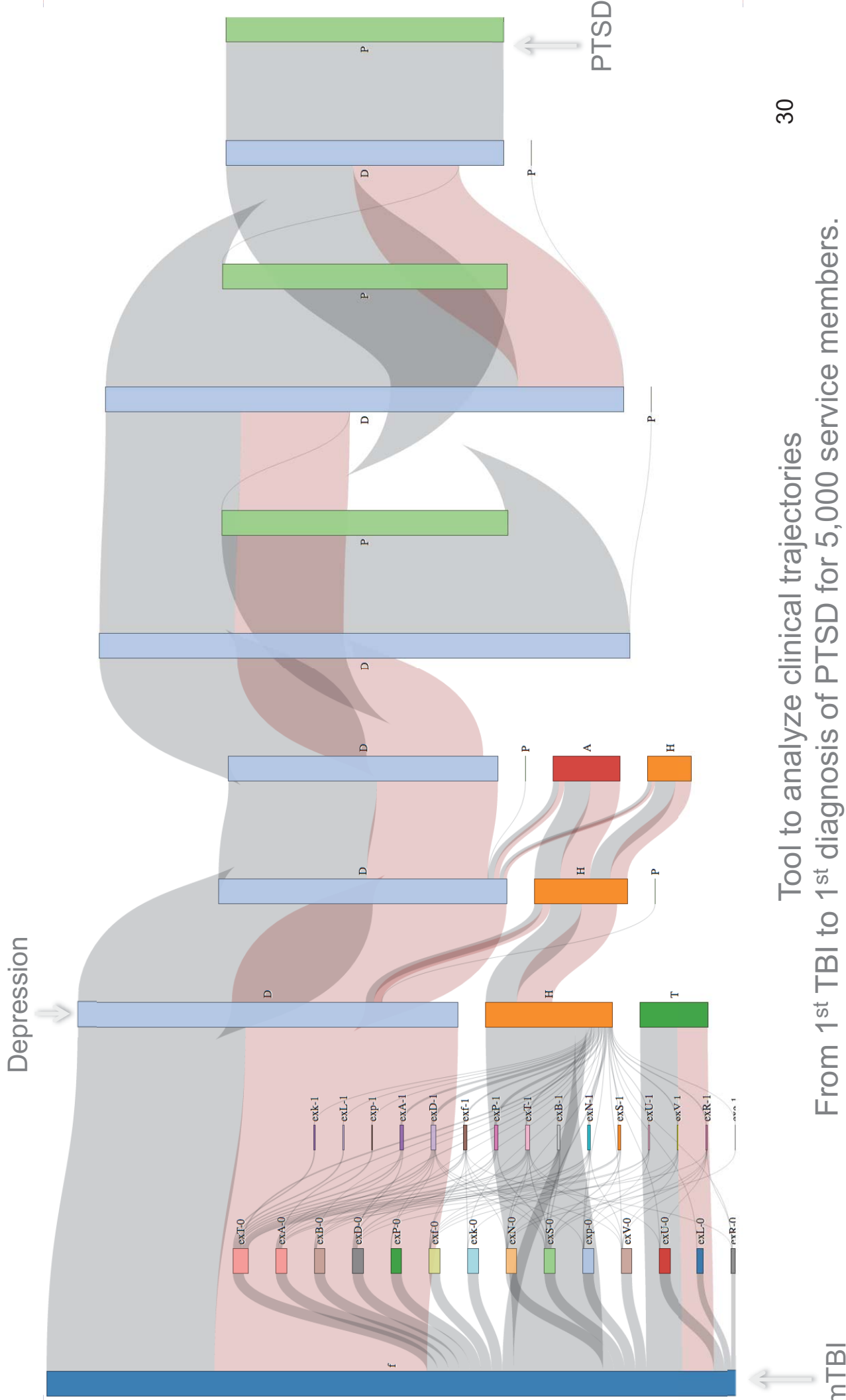


Visually Exploring EHR data for a Population or a Cohort of Patients





Exploring Longitudinal Data for a Cohort of Patients



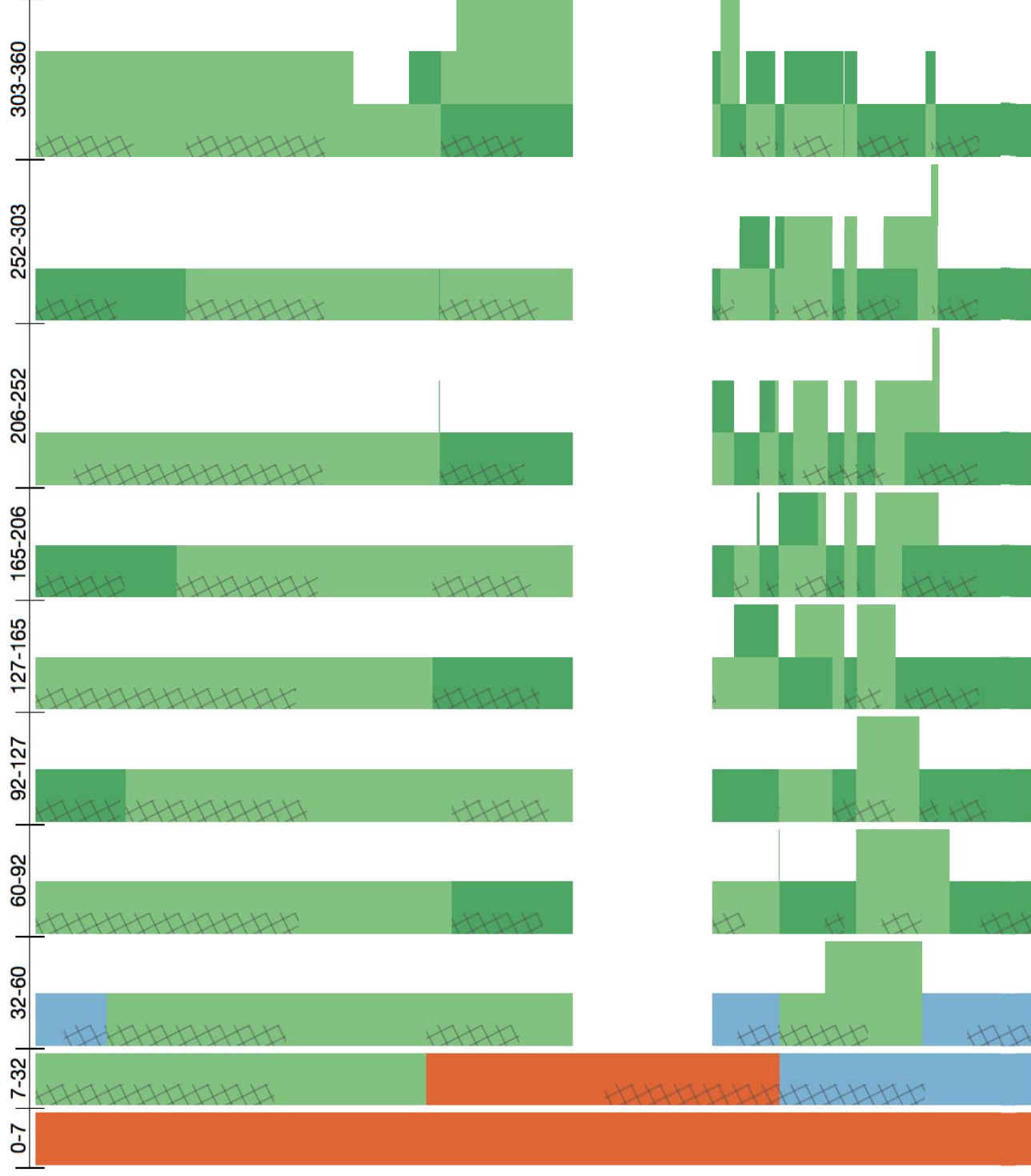
Tool to analyze clinical trajectories

From 1st TBI to 1st diagnosis of PTSD for 5,000 service members.

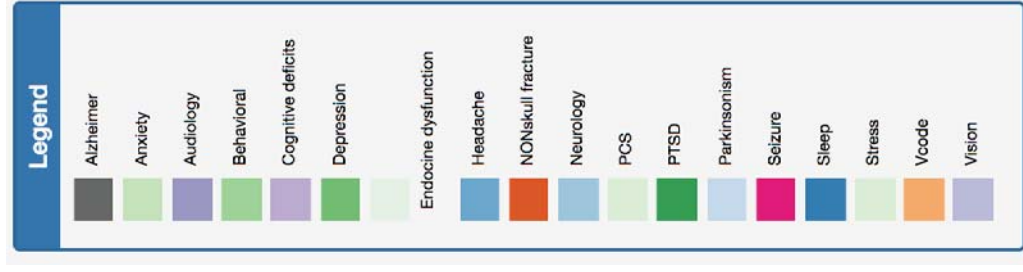
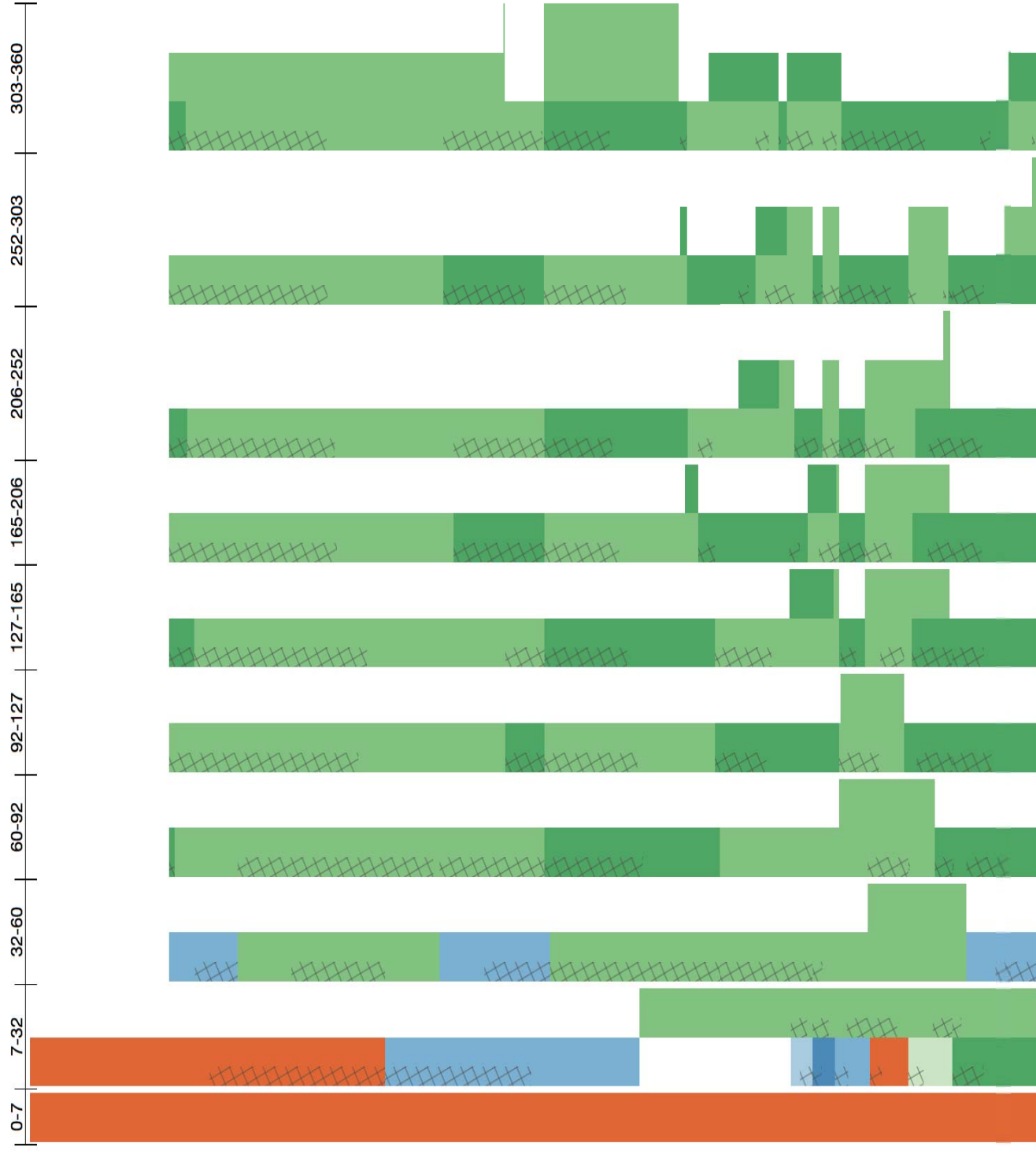
Hierarchical Exploration of Clinical Trajectories



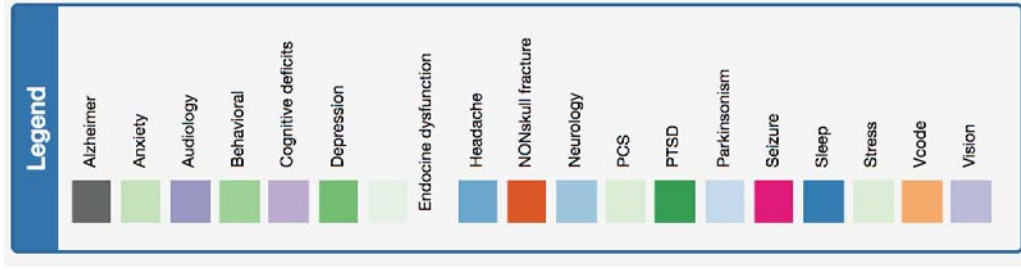
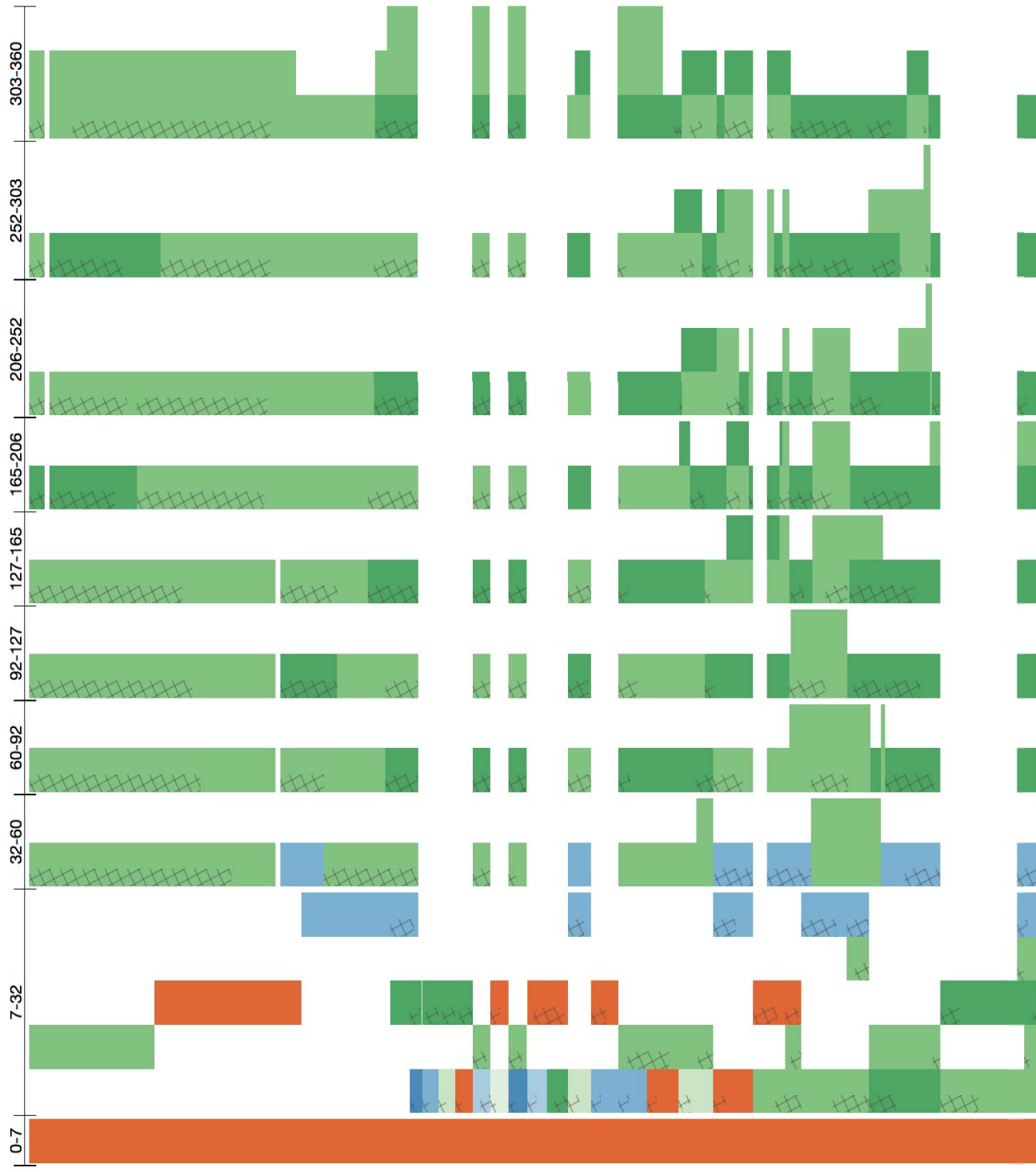
Hierarchical Exploration of Clinical Trajectories

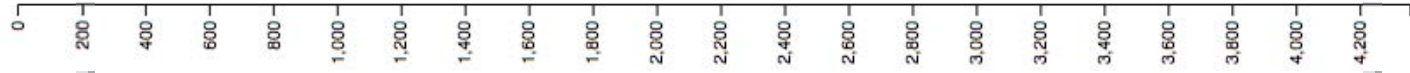


Hierarchical Exploration of Clinical Trajectories



Hierarchical Exploration of Clinical Trajectories





+

+

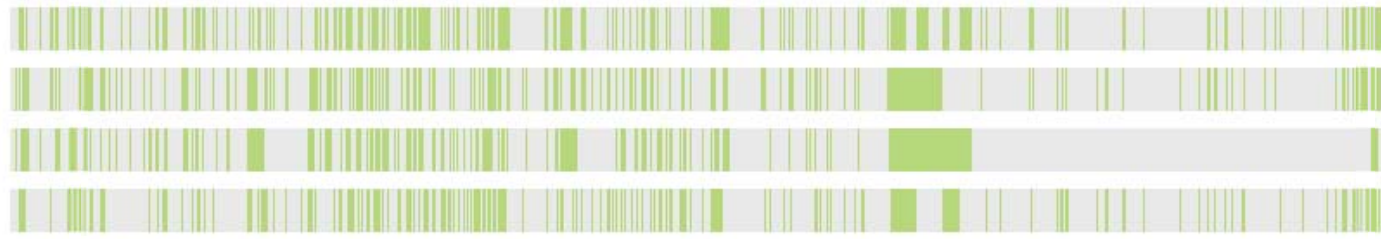
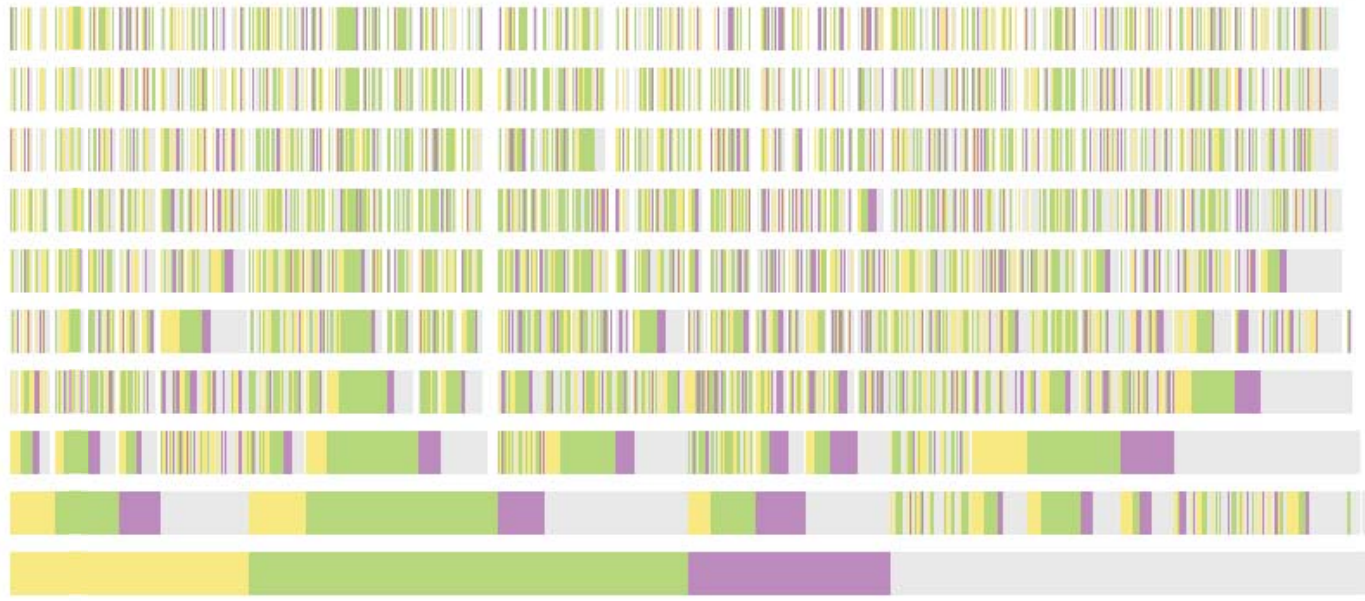
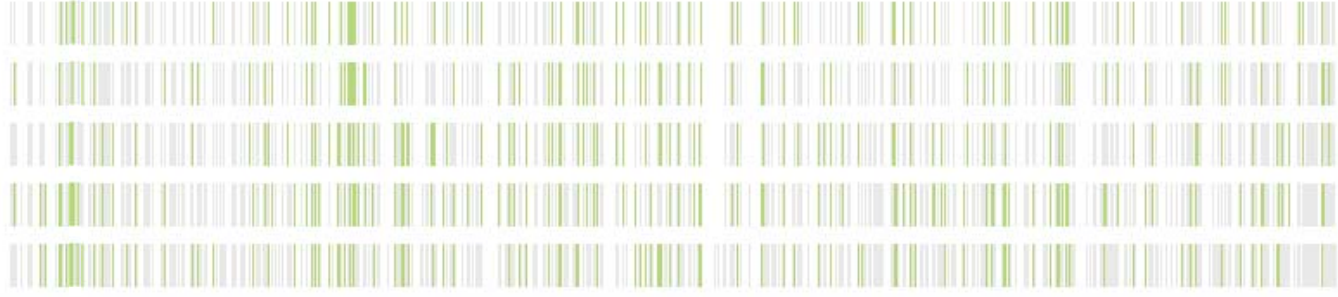
5-15

:

15-20

:

+



- ❖ POP started on April 15th, 2015
 - ❖ Delayed on starting the project. It didn't start until August 2015
 - ❖ Challenges hiring qualified developer to focus on the project due to budget modifications

Risks & Risk Mitigation Plan



- ❖ **Risk:** Delays due to complexity of brining qualified candidates.
- ❖ **Mitigation:** Spent significant amount of time training candidate.

- ❖ Number of research publications, presentations, and abstract
- ❖ Interest and the engagement shown by clinical providers



Visualization of Longitudinal Clinical Trajectories using a Graph-based Approach

Filip Dabek
NICoE
Walter Reed, Bethesda
filip.j.dabek.citr@mail.mil

Jian Chen
Dept. of Computer Science
and Electrical Engineering
University of Maryland, UMBC
jichen@umbc.edu

Alexander Garbarino
Dept. of Computer Science
and Electrical Engineering
University of Maryland, UMBC
garba1@umbc.edu

Jesus J Caban
NICoE
Walter Reed, Bethesda
jesus.j.caban.civ@mail.mil

Understanding the Effects of Concussion on Brain Trajectory

ABSTRACT
The adoption of electronic health records (EHRs) and the increased participation of hospitals and clinics in health information exchange systems have resulted in unique longitudinal data that describes a patient's clinical trajectory. In-depth analysis of that information is important to better understand the general course to recovery or the evolution of a particular disease. Unfortunately, modeling and understanding the progression of a disease is still a challenging task given that most often patients take vastly different paths after being diagnosed with a specific disease or condition. The different trajectories patients follow, the individual temporal events that each patient goes through, the uncertainties associated with clinical diagnoses, and the irregular time intervals between clinical diagnoses present challenges to researchers trying to analyze the common trajectories of a set of N patients. This paper presents a graph-based visualization method to interactively analyze the longitudinal clinical trajectory of a group of patients. The system allows users to select a specific set of events or conditions, filter the data based on different thresholds, and compare different cohorts while using an interactive virtual space that expands as the user continues to analyze and explore the data. The system has been tested with a dataset of over 89,000 patients and 8.7 million clinical events.

Keywords
Data visualization, visual analytics, longitudinal events, healthcare data

1. INTRODUCTION

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation to the source publication. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with permission is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
WACV '15 Chicago, Illinois USA
© 2015 ACM. ISBN 978-1-4503-8271-0
DOI: 10.1145/12235

A Grammar-based Approach to Clinical Trajectory after Injury

Filip Dabek
NICoE
Walter Reed National Military Medical Center
Bethesda, MD

Abstract—Despite the significant amount of longitudinal information that electronic health records (EHRs) include, modeling the progression of a particular disease is still a challenging task. Unfortunately, modeling and understanding the progression of some patients take similar paths (events) to a certain clinical disease, while others take vastly different paths. This paper presents a novel approach to modeling the trajectory of patients by analyzing the significant amount of longitudinal information that EHRs include. The approach is based on the idea of using a grammar to model the progression of a disease. The grammar is a set of rules that define the sequence of events that a patient goes through. The grammar is used to model the progression of a disease by analyzing the significant amount of longitudinal information that EHRs include. The approach is based on the idea of using a grammar to model the progression of a disease. The grammar is a set of rules that define the sequence of events that a patient goes through. The grammar is used to model the progression of a disease by analyzing the significant amount of longitudinal information that EHRs include.

Keywords—Big Data, Machine Learning, Concussion, Informatics, mild Traumatic Brain Injury

Leveraging Big Data to Model the Likelihood of Developing Psychological Conditions After a mild TBI

Filip Dabek and Jesus J. Caban
NICoE
Walter Reed National Military Medical Center
Bethesda, MD

Abstract

A concussion is an invisible and poorly understood mild traumatic brain injury (mTBI) that can alter the way the brain functions. Patients who have been screened positive with mTBI are at increased risk of depression, post-traumatic stress disorder (PTSD), and other psychological conditions. Unfortunately, modeling and understanding the progression of some patients take similar paths (events) to a certain clinical disease, while others take vastly different paths. This paper presents a novel approach to modeling the trajectory of patients by analyzing the significant amount of longitudinal information that EHRs include. The approach is based on the idea of using a grammar to model the progression of a disease. The grammar is a set of rules that define the sequence of events that a patient goes through. The grammar is used to model the progression of a disease by analyzing the significant amount of longitudinal information that EHRs include.

© 2015 Published by Elsevier Ltd.

Keywords—Big Data, Machine Learning, Concussion, Informatics, mild Traumatic Brain Injury

1. Introduction

A concussion is a poorly understood mild traumatic brain injury (mTBI) that can alter the way the brain functions. During the last decade a significant amount of attention has been given to the acquisition of clinical data from patients suffering mTBI and psychological health (PH) problems after a concussion. The increased awareness has been in part driven by the Department of Defense (DoD), the National Football League (NFL), and the National Hockey League (NHL). These organizations have been leading different efforts to raise awareness about the short- and long-term effects of concussions.

A traumatic brain injury (TBI) is defined and indicated by "Any period of loss of or a decreased level of consciousness. Any loss of memory for events immediately before or after the injury. Any alteration in diagnosis or state [4]."

For example, Mackintosh and Shuler studied a dataset of diabetic patients to identify the most common patterns in the progression of a disease. They used a grammar to model the progression of a disease. The grammar is a set of rules that define the sequence of events that a patient goes through. The grammar is used to model the progression of a disease by analyzing the significant amount of longitudinal information that EHRs include. The approach is based on the idea of using a grammar to model the progression of a disease. The grammar is a set of rules that define the sequence of events that a patient goes through. The grammar is used to model the progression of a disease by analyzing the significant amount of longitudinal information that EHRs include.

When PSI and WM data of mTBI patients with comorbid conditions is combined with injury information such as number of concussions or blast exposures, the location analysis can be used to identify patterns in the progression of a disease. To demonstrate this, we collected PSI and WM from 17 service members of similar age (M=36, SD=6, Male=99%).

well as to evaluate whether additional WM is a good indicator to assess a person's ability to remember new information. It held if in short-term memory, concentration, and transfer that

Anticipated Impact as an Outcome of Research

- ❖ Visually exploring the effects of a particular treatment plan and graphically monitoring changes over time will have a significant impact on the way clinicians assess the sequelae of TBI.
- ❖ To guarantee the success of our software application, two clinicians will guide, review, and validate the visualization system so it can be easily used and adapted by other researchers.
- ❖ We believe that the design and development of the software will create a community within the NICoE, Walter Reed, DHA, Intrepid Spirits, Uniformed Services University (USU), and other organizations that will use state-of-the-art visual analytic tools to better understand and explore TBI research data.
- ❖ Components of our visual analytical techniques being incorporated into CarePoint.

Transition Plan



- ❖ Research Community: the software application will be primarily developed to target clinicians and researchers who need to fuse physiological, cognitive, behavioral, and imaging measurements. To accomplish that we plan to:
 - ❖ Include clinicians in regular meetings so they can guide the research and development of specific charts, correlations, and comparisons most clinicians are interested in analyzing.
 - ❖ Validate the framework with other clinicians within the NICoE.
 - ❖ Freely distribute the software application to members of the community.

Transition Plan

- 8 Workshops, 2 tutorials, 1 JAMIA special issue, 1 panel, 1 WG on Visual Analytics in Healthcare
 - » 2010: IEEE Visualization Conference, Salt Lake City, Utah
 - » 2011: IEEE Visualization Conference, Providence, RI
 - » 2012: IEEE Visualization Conference, Seattle, WA
 - » 2013: AMIA, Washington DC
 - » 2014: AMIA, Washington DC
 - » 2015, IEEE Visualization Conference, Chicago, IL
 - 2015: AMIA Tutorial: Introduction to Visual Analytics in Healthcare
 - » 2016, AMIA Chicago, IL
 - 2016: IEEE Visualization Conference, Baltimore, MD
 - » 2017, IEEE Visualization Conference, Phoenix AZ
 - 2017, AMIA Panel, Visualization in Healthcare, Washington DC

In this issue:

Visual analytics in healthcare – opportunities and research challenges—p 260



Proceedings of the IEEE VisWeek Workshop on Visual Analytics in Healthcare: Understanding the Physicians Perspective

October 23rd, 2011
Pr
www.vahc.org



Proceedings of the IEEE VisWeek Workshop on Visual Analytics in Healthcare Open Health Data

October 17th, 2012
Seattle, WA
www.vahc.org



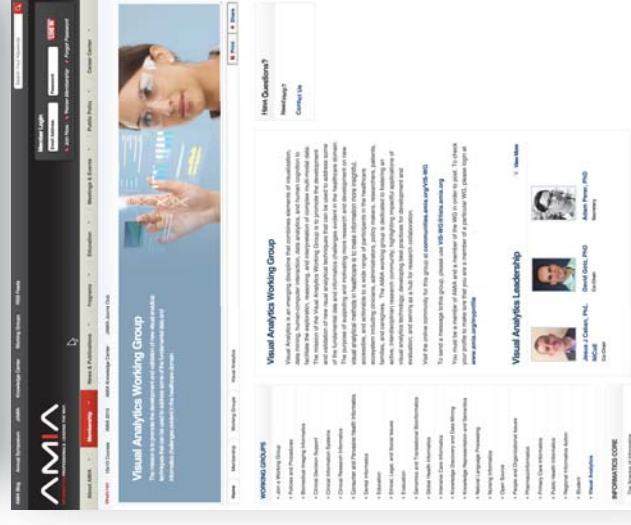
Proceedings of the 2013 Workshop on Visual Analytics in Healthcare:

November 16th, 2013
Washington, DC
www.vahc.org



Proceedings of the 2014 Workshop on Visual Analytics in Healthcare:

November 15th, 2014
Washington, DC
www.vahc.org



www.visualanalyticshealthcare.org/

AMIA Working Group

Conclusions



- ❖ Visualization systems can have a significant impact in how providers look at clinical data
- ❖ This project has the potential to impact different providers that need assistant with the data analysis and exploration
- ❖ Project on track



Contact Info:

Jesus J Caban, PhD
Chief, Clinical & Research Informatics
NICoE, Walter Reed Bethesda
E: jesus.j.caban.civ@mail.mil

The TBI Data Explorer project has continued to be developed. The TBI Data Explorer project will revolutionize how the Data Management team handles the parsing of the data.

NOTE_DATE

Wellness_Assessment

Wellness_Define_Welln_TimeANICoE

Wellness_Experience_Program

Wellness_Goal_Transition_NICoE

Wellness_Injures_Obstac_Challenge

Wellness_Intended_Outcomes

Wellness_Intent_Activ_Hobbi

Wellness_MindBody_Awareness

Wellness_MindBody_Confidence

Wellness_NOTE_DESCRIPTION

Wellness_Need_Add_Resource

Wellness_Personal_Definition

Wellness_Personal_Goal

Wellness_Recommend

Wellness_Take_Away

Reset

wellness *

wellness.json ▼

Load

Parsed In: 00:03:56

Wellness_Intended_Outcomes

```
(
  "array": (
    "on": false,
    "split": "in"
  ),
  "comment": null,
  "date_format": null,
  "dtype": "string",
  "name": "Intended_Outcomes",
  "nan": null,
  "parse": null,
  "prefix": [
    "Wellness_"
  ],
  "restrict": null,
  "suffix": []
)
```

Run

Comment	Computed	Raw
	"Align my chi and reintroduce my soul to my mind."	"Align my chi and reintroduce my soul to my mind."
	"Another way to cope with daily life."	"Another way to cope with daily life."
	"Any and everything that I can."	"Any and everything that I can."
	"Awareness/skillset in how to cope/deal with challenges affecting me."	"Awareness/skillset in how to cope/deal with challenges affecting me."
	"Become a whole person again - and push to a happy rest of my life with my wife and son."	"Become a whole person again - and push to a happy rest of my life with my wife and son."
	"Being able to relax or meditate on cue" On a scale of 1-10, how do you rate your level of mind-body awareness (1 being the lowest level of awareness, and 10 being the highest)?	"Being able to relax or meditate on cue" On a scale of 1-10, how do you rate your level of mind-body awareness (1 being the lowest level of awareness, and 10 being the highest)?
	"Better overall understanding of my injuries and what I can do	"Better overall understanding of my injuries and what I can do

Service Member (SM): Cruise, T.

Date: 3/4/18

Personal Definition of Wellness: "the brown dog"

Interests/Activities/Hobbies: "jumped over the"

Injuries/Obstacles/Challenges: "red cow"

Intended Outcomes: "The orange pumpkin was ready"

On a scale of 1-10, how do you rate your level of mind-body awareness (1 being the lowest level of awareness, and 10 being the highest)? 3

Wellness Assessment: for the big show. Halloween being right around the corner, this was its time to shine. At no other time of the year would such an opportunity present itself. The pumpkin's backstory is

a boring one. It was born underground in the pumpkin patch.

Recommendations: Up down, left right, north south- 3 minutes min.

Follow Up

Date: 7/10/18

How has your definition of wellness developed during your time at the NICoE?

"The sky is blue. The river flows."

What are your wellness goals as you prepare to transition from the NICoE?

"Puppies and kittens are cute."

Are you in need of any wellness resources or information?

"Horses gallop."

What are your most meaningful "take-aways" from the NICoE wellness experience?

"Sunflowers are yellow."

How was your experience of the NICoE's wellness program?

"Monitors are black and paper is white."

On a scale of 1-10, how do you rate your level of mind-body awareness (1 being the lowest level of awareness, and 10 being the highest)? 2

Wellness Assessment: Halloween being right around the corner, this was its time to shine. At no other time of the year would such an opportunity present itself. The pumpkin's backstory is a boring one. It was born underground in the pumpkin patch.

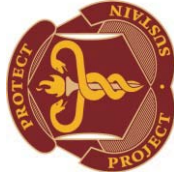
first_name: T.

last_name: Cruise

Referral ID: Referral ID Verified

Warning: Referral ID is usually 4-5 numbers in length

Wellness_Need_Addl_Resource ☐ Yes ☒ No	RAW:	COMPUTED:	COMMENT:
Wellness_Interest_Activ_Hobb ☐ Yes ☒ No no data pulled	RAW:	COMPUTED:	COMMENT:
Wellness_Personal_Definition ☐ Yes ☒ No too long	RAW:	COMPUTED:	COMMENT:
NoteDate ☐ Yes ☒ No	RAW:	COMPUTED:	COMMENT:
Wellness_Injuries_Obstac_Challenge ☐ Yes ☒ No	RAW:	COMPUTED:	COMMENT:
Wellness_Recommend ☐ Yes ☒ No	RAW:	COMPUTED:	COMMENT:
Wellness_Assessment ☐ Yes ☒ No more than field	RAW:	COMPUTED:	COMMENT:



- Updated and expanded on an application that uses Natural Language Processing (NLP) to analyze clinical notes and extract terms. The application allows a provider to easily search and explore notes for a patient. The tool automatically highlights clinical concepts within unstructured provider notes. In addition, the new module can be used to identify commonalities between different EHR notes.
- During the last quarter, incremental improvements were made based on feedback. Additionally, a large amount of time was dedicated to optimizing the interface to handle large amounts of clinical data without sacrificing speed and performance.

The screenshot displays the DHAXELP Analysis interface. At the top, there's a navigation bar with 'Select a Patient', 'History', 'Statistics', and 'Bookmarks'. Below this, a sidebar shows 'Note 331' and 'Note 1039' with their respective last viewed dates. The main content area is divided into several sections: 'Physical Therapy', 'Neurology', 'Audiology', and 'Art Therapy'. Each section contains a 'PATIENT TEST INFORMATION' card with details like 'Indication', 'Height', 'Weight', 'BSA', 'BP', 'HR', 'Status', 'Date/Time', 'Test', 'Doppler', 'Contrast', and 'Technical Quality'. The 'Physical Therapy' section also includes a 'Service' dropdown menu with options like 'CARDIO', 'Allergies', 'Pericard', and 'Abdominal'. The 'Neurology' section includes a 'PATIENT TEST INFORMATION' card with details like 'Indication', 'Height', 'Weight', 'BSA', 'BP', 'HR', 'Status', 'Date/Time', 'Test', 'Doppler', 'Contrast', and 'Technical Quality'. The 'Audiology' section includes a 'PATIENT TEST INFORMATION' card with details like 'Indication', 'Height', 'Weight', 'BSA', 'BP', 'HR', 'Status', 'Date/Time', 'Test', 'Doppler', 'Contrast', and 'Technical Quality'. The 'Art Therapy' section includes a 'PATIENT TEST INFORMATION' card with details like 'Indication', 'Height', 'Weight', 'BSA', 'BP', 'HR', 'Status', 'Date/Time', 'Test', 'Doppler', 'Contrast', and 'Technical Quality'.

This screenshot shows another view of the DHAXELP Analysis interface, focusing on the 'Physical Therapy' and 'Neurology' sections. The 'Physical Therapy' section displays a 'PATIENT TEST INFORMATION' card with details like 'Indication', 'Height', 'Weight', 'BSA', 'BP', 'HR', 'Status', 'Date/Time', 'Test', 'Doppler', 'Contrast', and 'Technical Quality'. The 'Neurology' section displays a 'PATIENT TEST INFORMATION' card with details like 'Indication', 'Height', 'Weight', 'BSA', 'BP', 'HR', 'Status', 'Date/Time', 'Test', 'Doppler', 'Contrast', and 'Technical Quality'. The interface also includes a 'Service' dropdown menu with options like 'CARDIO', 'Allergies', 'Pericard', and 'Abdominal'. The 'Neurology' section includes a 'PATIENT TEST INFORMATION' card with details like 'Indication', 'Height', 'Weight', 'BSA', 'BP', 'HR', 'Status', 'Date/Time', 'Test', 'Doppler', 'Contrast', and 'Technical Quality'.

Radiology

Save as Patient Types * Type to search

Type *

Type to search

Types *

Clear All Medication X

PATIENT/TEST INFORMATION:

Indication: Congestive heart failure.

Height: (m) 67

Weight: (lb): 251

BMI: 35.7 (m²): 2.23 m²

BP (mm Hg): 118/72

HR (bpm): 70

Status: Inpatient

Date/Time: [**2124:1:25**] at 10:51

Test: TTE (Complete)

Doppler: Full Doppler and color Doppler

Comments: None

Technical Quality: Adequate

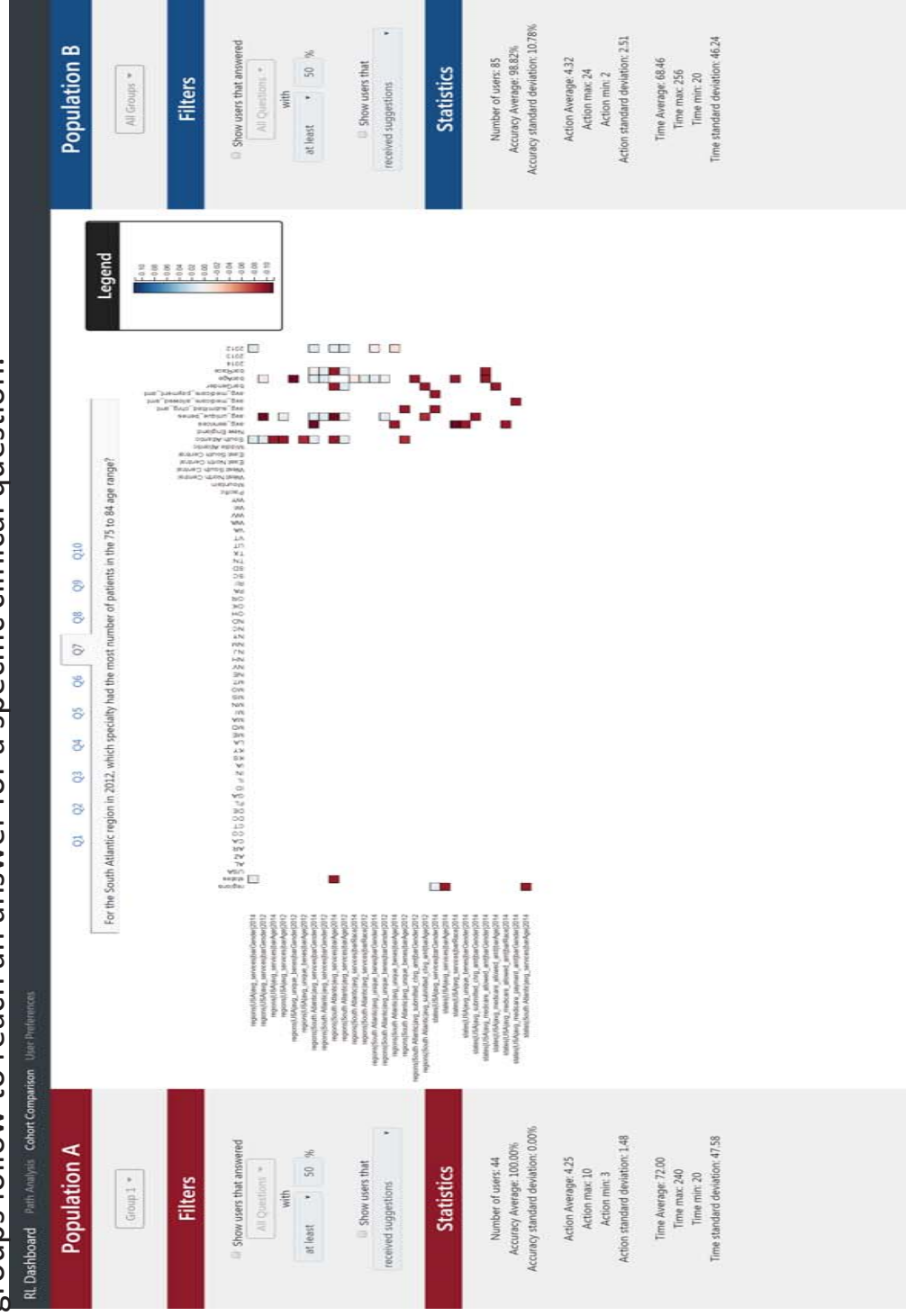
INTERPRETATION:

Findings:

This study was compared to the prior study of [**2122:1:2**].

[illegible]

In order to better understand and validate the effectiveness of different visualization tools, we have developed a method to capture how users interact with the different systems. Given a clinical dataset and a visualization tool or dashboard, we designed an application to compare the correlations, patterns, and flows different groups follow to reach an answer for a specific clinical question.

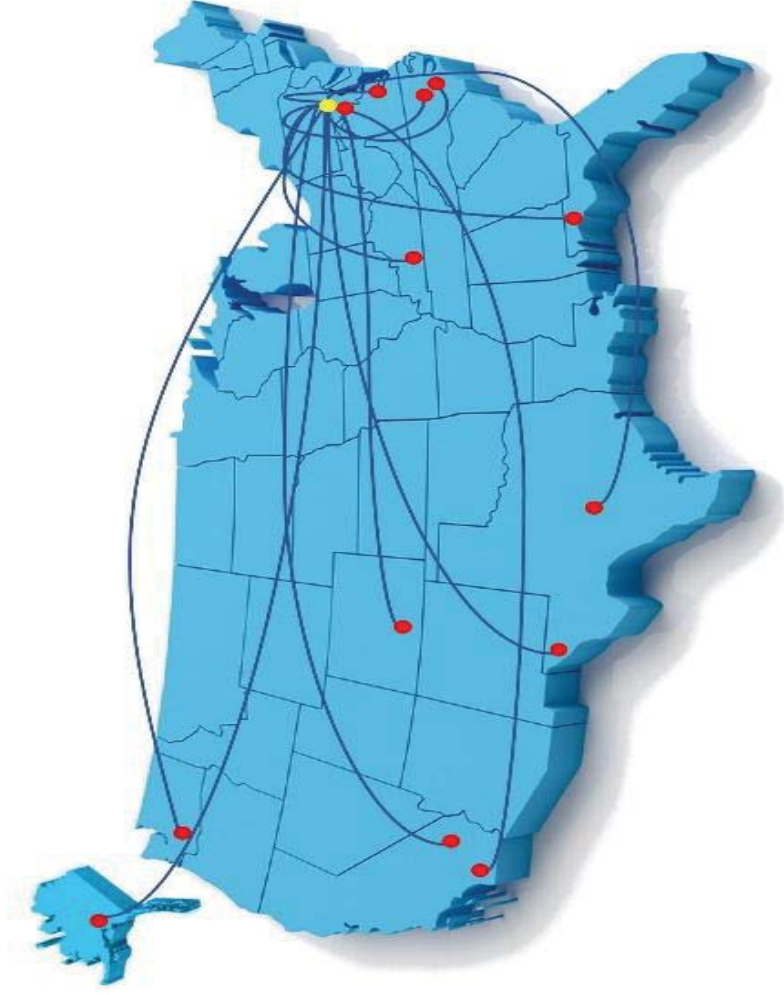


Defense Health Agency

[illegible]

Map of the United States showing the locations of 12 military bases and the flight paths of a traveling salesman problem solution. The bases are: Joint Base Elmendorf-Richardson, Joint Base Lewis-McChord, Fort Carson, Fort Campbell, Fort Bliss, Fort Hood, Eglin AFB, Fort Belvoir, Camp Lejeune, Fort Bragg, Twenty-nine Palms, and Camp Pendleton. The flight path is a series of blue lines connecting the bases in a specific sequence, starting and ending at Joint Base Elmendorf-Richardson.

Design (2)



Description: A poster detailing the features of the TBI Portal.



The TBI Portal is an enterprise registry designed to effectively aggregate, summarize, and visualize disparate data of TBI patients. The portal offers providers a more comprehensive view of the patient's longitudinal data, through embedded dashboards and analytics tools.

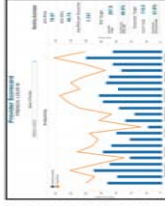
Clinical Registries

As part of the TBI Portal, Clinical Registries have been developed to provide distinct and customizable environments for each Military Treatment Facility (MTF). These environments allow healthcare providers to track clinical referrals, review and measure patient progress, and complete clinical notes. Administrators are also provided the ability to easily measure clinic and provider workloads.



Dashboards

TBI Dashboards are a suite of enterprise metrics and interactive dashboards designed to provide a single analytical framework to help TBI organizations understand organizational trends, quantitative measures, and organizational indicators. Through embedded dashboards and analytics tools, the TBI Dashboard support clinical operations and administrators of TBI clinics with understanding the patient demographics, referral patterns encounter trends, patient health outcomes, measuring and forecasting the incidence of TBI across , and many other metrics.



Research

In an effort to provide clinicians and administrators the tools they require to make informed decisions, the data collected and aggregated within the TBI Portal is analyzed and theories, on the nature of the data, are formulated. Based on those theories, tools are created in hopes of helping users make informed decisions.




TrajectoryFlow: Visual Summarization of Temporal Sequences

Category: Research

Paper Type: Algorithm/Technique

Abstract—Summarizing a collection of temporal sequences is a difficult task given the irregular and variable patterns often found in longitudinal events. Across a wide array of domains, researchers and analysts seek to determine ways to gain an overview of a dataset through the identification of the common temporal paths as well as the trajectories that exist between individual events. While these tasks can be difficult on small and structured datasets, they are increased tenfold on temporal sequences that are noisy, irregular, and voluminous in size. This paper presents an approach that has been designed to visually explore large collections of temporal sequences by combining event mining algorithms with visualization techniques to overcome some of the challenges and complexities of analyzing longitudinal data. The key contributions of this paper are an adaptive window-based frequency sequence mining algorithm designed to identify common subsequences and build overviews of longitudinal trajectories and a novel event summary diagram to display common paths while providing multiple levels of details. The algorithm and visualization techniques have been evaluated through three case studies applied to large-scale longitudinal datasets from different domains.

Index Terms—Visualization of event sequences, sequence mining, event summarization

1 INTRODUCTION

Events exist everywhere in life: online consumers perform a series of comparisons and clicks before purchasing an item [38], patients go to different doctors and therapies during the recovery from a specific condition [11, 24, 26], sports teams make plays and score back-and-forth during a match [24]. Throughout all domains, a temporal event can be described as *a data point denoted by a timestamp and event type*. However, while events can exist in every avenue of life and in a vast array of domains, understanding how events are related to one another and summarizing a large sequence of events remains a difficult task. During the last few years, the increased availability of large-scale temporal datasets that contain complex and long sequences has underscored the importance for designing scalable techniques to easily identify trends and to answer important questions from irregular, noisy, and large longitudinal datasets.

To accomplish the need for exploring event datasets, many techniques have been developed in recent years that provide analysts with the ability to filter, explore, and simplify their data. However, while these approaches have been shown to work on small datasets, they face significant challenges with large-scale longitudinal data as they often (i) present users with the entire dataset which can cause for the visualization to become complicated, noisy, and complex to the point of being discouraging and (ii) struggle to effectively handle large datasets with short, long, and extremely long temporal sequences. Because of these limitations, an approach that is able to handle large scale data, generate a summary, and provide the ability to explore at multiple levels of detail could greatly assist the process of analyzing large scale longitudinal data.

In this paper we present *TrajectoryFlow*, a scalable algorithm and visualization, that has been built to effectively analyze and summarize longitudinal data. The primary contributions of this work include: (1) An adaptive window-based algorithm for extracting common subsequences of a longitudinal dataset. The mining algorithm takes into account the temporal context of the sequences to build an overview of the common paths and trajectories followed by a collection of records, and (2) a novel event summary diagram that reduces the analytical complexity used to visualize the results of event mining approaches. Our diagram also incorporates multiple levels of detail to enable hierarchical and detailed exploration of the data, allowing for a user to analyze from both a summarized view as well as an expanded, detailed view.

The rest of this paper is structured as follows: Section 2 provides a formal definition of longitudinal events and explains the current tasks and challenges of temporal summarization, Section 3 provides a background on the existing approaches for these tasks, Section 4 describes the set of goals that are expected from an effective tool to analyze

temporal events, Section 5 presents our algorithm for event summarization, Section 6 presents our novel event summary diagram used for analysis of the mined information, Section 7 shows our approach through multiple case studies applied to a wide array of domains, and finally Section 8 concludes the paper by summarizing our contributions and the future work that we seek to explore in the area. We also provide an interactive, online tool for demonstration of our approach at: <http://tf.vixsplore.com>

2 TEMPORAL EVENTS

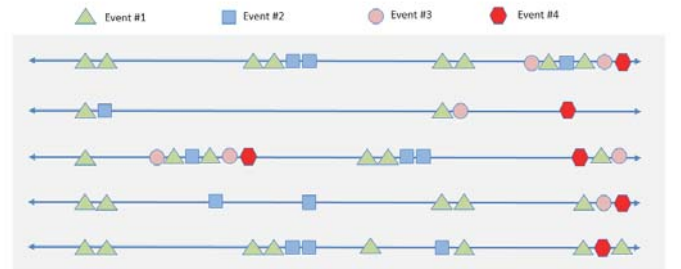


Fig. 1: Illustration showing a set of temporal sequences with the trajectory between event #1 (triangle) and event #4 (hexagon). As the size and complexity of the data increases, there is a need to generate a visual summarization of the sequences to better understand the relationships between events.

2.1 Data Attributes

A longitudinal dataset can be defined as a set of sequences $D = \{S_0, S_1, \dots, S_n\}$ where each sequence $S_i = \langle I_1, I_2, \dots, I_m \rangle$ contains a timestamped series of items I_i arranged in a specific order with respect to the timestamps. The arrangement of the items can be used to derive temporal relationships within the data such as $I_i \prec I_{i+1}$ describing that I_i happened before I_{i+1} . Each itemset, I_i , contains a set of event types such that $I_i = \langle T_1, T_2, \dots, T_p \rangle$ and $T_i \in E$ where $E = \{e_1, e_2, \dots, e_n\}$ is a dictionary of event types. An example of a dataset of sequences is shown in Figure 1, where each row represents a sequence S_i that contains itemsets I_i with one or more event types.

Looking at Figure 1, some of the patterns are easy to notice due to the relatively small nature of the dataset, but at the same time there are more hidden patterns that are difficult to identify and would require

analyzing each individual sequence to be able to pick them out. As the size of this dataset would grow from only 5 rows to hundreds or even millions of rows and sequences from a maximum of 15 data points to millions of data points, these patterns would become extremely difficult to identify and spot from simply looking at the dataset.

With this in mind, Figure 2 shows a summary of the dataset where the summary contains the most common path of events that occur between event #1 (triangle) and event #4 (hexagon). This event summary provides a concise description of the dataset, enables an analyst to easily understand the patterns contained, and represents an example of the summary that researchers seek to obtain so that identifying insights is made easier.



Fig. 2: An example summarization of Figure 1, showing the usefulness of summarizing a dataset of sequences.

2.2 Tasks

Generating an event summary, as shown, and analyzing events present interesting challenges and questions that are applicable to a vast array of domains, ranging from sports matches to business transactions to patient health records. For example, business analysts for online stores are interested in analyzing the comparisons and clicks performed by consumers prior to purchasing an item to be able to properly attribute revenue to the marketing channels that the consumer was directed from. Two common strategies that have been utilized are to only assign revenue to the first or last channel that the user went through, but these strategies severely undervalue the importance of the clicks that the user takes in between and how much they influence the user's final action of making a purchase [38].

In another example, the recent advancements and adoption of Electronic Health Records (EHRs) systems in clinical practice have resulted in rich longitudinal data that can be used to measure the progression of different conditions. The course of most diseases is quite variable and understanding their trajectory can provide key information about the effectiveness of treatments, insight about potential prognoses, and information about the state of the patient. Unfortunately, due to the complexity and challenges associated with the analysis of sparse longitudinal clinical data, EHR data is not often used within clinical settings to study trajectories of similar patients. Some of the contributing factors that have limited the amount of information that is extracted from longitudinal data embedded in EHR systems have been the vastly different paths that similar patients can follow, the irregular intervals at which particular events occur within comparable patients, the comorbidity of diseases many patients experience, the uncertainty associated with clinical diagnosis and coding, and the lack of tools to quickly analyze the information.

For both of these domains the tasks for those that analyze event trajectories, such as business analysts, web optimization analysts, clinicians, and health policy makers, wish to identify are similar and can be summarized as:

- The most common trajectories and how many sequences are associated with each varying path
- An overall trajectory from a specific beginning or ending event, such as the first occurrence of an event
- Which events occur the most in each timeframe of an event trajectory

An approach that is able to successfully target these specific tasks has the potential to support users and make hidden insights much easier to

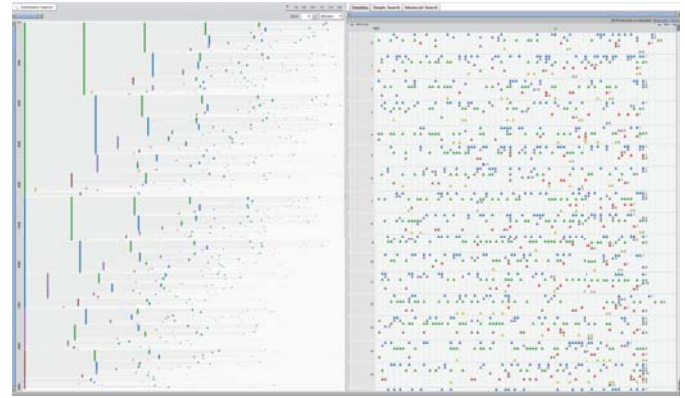


Fig. 3: A screenshot of a popular event mining tool used to visualize and explore the longitudinal patterns within a soccer match dataset of 9,084 games and over 900,000 data points.

uncover in large event datasets. To support these tasks and requirements, we need to ensure that our approach is not limited solely to a particular domain, but rather has the flexibility to tackle the challenges faced by any domains that utilize event datasets.

3 BACKGROUND

In the visualization field, a common method of understanding event sequences has been to visualize a single record to understand its path over time [2, 3, 8, 13, 30]. However, as of late various methods of visualizing event sequences have been explored by stacking sequences vertically on top of one another attempting to merge identical paths to allow for interactive clustering [4, 29], filtering [34], exploration [40], and searching [10, 32–34, 36, 37].

Several systems have been developed for event sequence visualization. EventFlow [24] presents an aggregated display of temporal records and allows users to simplify large, complex datasets into a concise display through filters and merging of events. MatrixWave [41] adapted the concept of a Sankey diagram into a matrix visualization for comparing event sequences. CareFlow [26] used a similarity distance metric to identify a population of similar records and display them in a flow visualization allowing for users to be able to identify the most effective outcomes based on sequence paths. Outflow [35] aligned sequences on a specific event and aggregated them into a directed acyclic graph (DAG) which was used to visualize results in a Sankey-type interactive visualization.

Many of these approaches, while powerful and novel in their applications, choose to display the entirety of a dataset which causes for them to suffer from both performance issues in terms of scalability and high visual complexity in terms of the number of elements present in the visualization. This decision of presenting each individual sequence is reasonable for small datasets, but for large datasets that can contain millions of records and data points such an approach can quickly cause for the visualization to become complex and discouraging for the user to analyze, especially if the aforementioned dataset is very noisy.

To present a use case scenario of employing existing event analysis applications to explore a longitudinal dataset, Figure 3 shows a screenshot of the initial load of using EventFlow [24] to explore the longitudinal patterns within a soccer match dataset of 9,084 games and over 900,000 data points. At first glance we can see that the tool is effective at making it easier for users to identify that the green and blue events are the most common events that occur at the beginning of the sequences, however beyond that initial starting point the visualization becomes complex and patterns become very difficult to identify. At this point, users are expected to take a bottom-up approach and combine events to tune the visualization to identify insights in their data. One can imagine how this problem becomes increasingly more difficult as the size of the dataset grows beyond 9,000 to millions of records and how a top-down approach would assist an analyst in being able to

extract information and highlight trends in datasets of such size.

To help solve this challenge, recent approaches have utilized Frequent Sequence Mining (FSM) algorithms to identify the frequent patterns contained within a dataset and present a concise view to users. Liu et al. utilized an FSM algorithm to present both the most frequent patterns and an icicle plot of all sequences within a dataset [20]. Kwon et al. also utilized an FSM algorithm to display similar information in addition to a histogram and network of the overall dataset [16]. DecisionFlow [11] combined FSM and statistical methods to identify frequent sequences and their correlations to specific outcomes which was presented in a timeline and pattern diagram to understand the statistical results. Frequence [27] modified an FSM algorithm to visualize the most frequent subsequences and their outcome measures in a Sankey diagram. While these approaches assist users in identifying patterns that frequently occur; they lack the ability to construct a summary and simply present the common patterns contained within the entirety of a dataset which can be misleading to the user. FSM algorithms alone are not able to factor in the temporal position of each frequent pattern, thus a pattern that exists at the beginning of a sequence is not indicative of being the same pattern that is contained at the end of a sequence. Therefore, mining an entire dataset as a whole does not guarantee that the identified patterns can be considered the same throughout the entire dataset.

3.1 Limitations

As large-scale longitudinal datasets continue to become widely available for analysis, providing an avenue for exploration through powerful visualization tools is paramount. This paper presents *TrajectoryFlow* to overcome some of the challenges and limitations of existing techniques by addressing several aspects of real world datasets including: (i) real world data contains a lot of uncertainty and noise, making it difficult to simplify the temporal events, thus often resulting in requiring a user to perform a long series of filter and exploration steps rather than being presented an easy and concise view; (ii) real world data tends to be long and often contains patterns early on that can differ from those that occur late, thus making it difficult to summarize longitudinal data; (iii) presenting each individual sequence separately does not allow for scaling with respect to the number of sequences that can be taken into consideration at a given time; and (iv) constructing a trajectory between two events is a common task of interest to researchers, but unfortunately most of the existing interaction techniques are limited at accomplishing such a fundamental task.

4 APPROACH

The goal of this project was to design an algorithm and visualization that enables easy analysis of longitudinal data while addressing some of the fundamental challenges faced by researchers when analyzing temporal events. The specific aims can be summarized as follows:

1. Provide a scalable and efficient algorithm for event summarization such that datasets that contain very long sequences or very many sequences can be quickly analyzed.
2. Develop an effective approach to build a path from a specific event type E_i to another event E_j with the attempt to answer research questions such as “what path do records follow to get to D from A?”.
3. Visually explore the mined event summary while allowing for multiple levels of detail.
4. Reduce the visual complexity and required analytical thought processes of existing event visualizations.

These four goals encompass the premise of summarizing a large dataset and understanding the flow of events within it through an easy to understand visualization that still allows for an analyst to dive into a detailed view of the contained sequences. To accomplish these goals, for our approach, we have built (1) an algorithm, Adaptive-Window-PrefixSpan (AWP-Span), for identifying and visualizing the overall

Table 1: Example Dataset

	Items #1	Items #2	Items #3	Items #4
S_1	{A, C}	{B, D}	{C, D}	
S_2	{A}	{A, C}	{B, D}	{C, D}
S_3	{A}	{B}	{C}	{D}

event sequence flow throughout a dataset, and (2) an event summary diagram that allows for a large collection of temporal events to be visualized and interactively explored. In the next section we will begin by describing the algorithm that we have designed for understanding the flow of events, followed by our novel event summary diagram.

5 MINING

Our first contribution is an adaptive window-based algorithm that introduces a temporal context and an ability to build a trajectory to a popular mining algorithm, PrefixSpan [19, 25]. In order to understand our algorithm, we first provide an explanation and example of FSM algorithms, and most specifically the PrefixSpan algorithm.

5.1 Frequent Sequence Mining

Frequent Sequence Mining (FSM) is a popular technique for finding the most frequently occurring subsequences of events in datasets of sequences [22]. These algorithms define a subsequence as being “frequently occurring” when the number of sequences that it is contained in is more than a certain support threshold. FSM algorithms stem from either Apriori-based or pattern-growth-based approaches [1, 22, 23, 39], but the general structure of an FSM algorithm is to build a frequent set of events that are each of length one and then to continue growing these events by concatenating more events until the defined threshold is met. This results in the algorithm producing a list of frequent subsequences and the number of sequences that each of them appear in.

When computing the frequent subsequences of a dataset, a sequence, $S_i = \langle I_{i1}, I_{i2}, \dots, I_{im} \rangle$, is considered to be a subsequence of another sequence, $S_j = \langle I_{j1}, I_{j2}, \dots, I_{jm} \rangle$ if there are m itemsets in S_j that satisfy $I_{i1} \subseteq I_{j1}, I_{i2} \subseteq I_{j2}, I_{i3} \subseteq I_{j3}, \dots, I_{im} \subseteq I_{jm}$. It can also be said in reverse that S_j contains S_i .

The support, τ , of a subsequence, S_k , is defined as the fraction of all sequences in a dataset that contain the subsequence S_k . In FSM algorithms a support threshold is predefined such that when the “frequently occurring” subsequences are returned, they have a support at least as large as the threshold.

To better explain FSM algorithms, we will provide an example of how to determine if a subsequence is frequently occurring. In this example we will utilize the dataset D shown in Table 1 and a support threshold, τ , of 50% to determine if the subsequence $S_f = \langle A, B, D \rangle$ is frequently occurring. The support threshold of 50% means that at least 50% of all sequences in the dataset, D , must contain the subsequence S_f for it to be considered “frequently occurring”.

Iterating through the sequences we first can see that S_1 does contain S_f as it has A in the first itemset, followed by B in the second itemset, and D in the third itemset. Looking at the next sequence, S_2 , we can see that it contains A in the first itemset, but does not contain B in its second itemset which would initially seem that it does not contain S_f . However, starting from the second itemset we can see that it contains A , followed by a B in the third itemset, and finally a D in the fourth itemset. This means that S_2 does in fact contain S_f and shows how the subsequence can be contained anywhere in the location of the sequence. Moving along to the final sequence, S_3 , we can see that it contains A followed by B in its first two itemsets, but in its third itemset it lacks the D event resulting in it not containing S_f .

Through iteratively looking at the sequences we could see that S_1 and S_2 contained S_f , while S_3 did not, resulting in S_f being present in 66% of the dataset, thus we can conclude that S_f is frequently occurring due to 66% being at least as high as the threshold of 50%. In this example we have shown how FSM algorithms are able to identify all of the frequent subsequences, but instead of taking a brute-force approach they operate on efficient data structures for an optimal search strategy.

5.2 AWP-Span

While FSM algorithms, such as the PrefixSpan algorithm that we have defined, can identify the most frequent subsequences in a dataset, they do not possess the ability to take into account temporal context and build a trajectory between two events. One of the issues with lacking temporal context was highlighted in the example shown in Section 5.1 and Table 1 where the frequent subsequence S_f occurred in a different location within S_2 than it did in S_1 . While the example contained sequences of relatively short length, in real world data the sequences may be very long and frequent patterns at the beginning of a sequence may not necessarily contain the same meaning as a pattern at the end of a sequence. This limitation poses a problem for tasks where an event dataset researcher seeks to identify the most common patterns within a certain stage of all sequences, as was shown in the business analyst example in Section 2.2, because the frequent subsequences obtained from FSM algorithms will potentially be present in a different section of each sequence.

In addition, as event researchers also attempt to understand the various paths that sequences may take, mapping the output of an FSM algorithm to a visualization would result in little information on the trajectories embedded within the dataset. For example, representing the frequent subsequences within a Sankey diagram would simply show the most common patterns within the dataset, but not provide an analyst with information on the most common trajectories and the varying paths followed by sequences.

With these limitations in mind we extend the work of PrefixSpan to include a window technique that is able to take into account the temporal context of a dataset in each window and build an overall pattern followed by the majority of sequences. The pseudocode for our modified PrefixSpan algorithm is shown in Appendix A.

Our algorithm, AWP-Span takes as input a dataset of sequences D , a support threshold τ , and a list of windows W of size ω denoting the number of segments the dataset should be split into. The execution of AWP-Span begins with splitting the dataset D into the windows defined by W and then identifying the most frequently occurring subsequences in each window using PrefixSpan. For example, Figure 4 provides an example of this beginning execution step. In this example the dataset in Figure 4 (left) is split into two windows of size ω such that $W = W_1, W_2$. Then Figure 4 (right) shows a table of the subsequences that occur in W_1 and the subsequences that occur in W_2 . The red lines through certain subsequences indicate that they are not “frequently occurring” as the number of times that they occur is less than the support threshold of three. By splitting a dataset into windows, like we have shown, we are able to incorporate a temporal context into PrefixSpan allowing for it to identify the common patterns at different stages of a sequence rather than ignoring the premise that sequences change over time and patterns in different temporal locations may not be indicative of the same pattern. Being able to evaluate the various stages of a sequence is an important task that event researchers seek to accomplish and this modification of the PrefixSpan algorithm builds this ability directly into the algorithm.

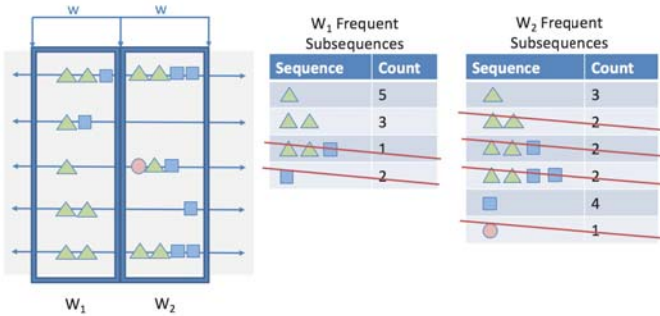


Fig. 4: (Left) An example of a dataset being split into two windows each of size ω and a support threshold of $\tau = 3$. (Right) The frequent subsequences that result from running PrefixSpan on W_1 and W_2 . The red lines indicate the subsequences that were ignored by the algorithm as they occurred less than the support threshold.

Once we have identified the most common subsequences in each window, connecting the windows together such that events flow from one window to the next is the next step in building a common flow and understanding the long term patterns in a dataset.

5.2.1 Window Merge

To connect two adjoining windows, a simple approach would be to concatenate two windows’ frequent subsequences together based on the beginning and ending events. For instance given the frequent subsequences for two example windows $W_1 = [\langle A, B, C \rangle, \langle A, D \rangle]$ and $W_2 = [\langle C, D \rangle, \langle C, F \rangle]$, the subsequence “ABC” in the first window would be connected to “CD” in the second window such that it forms “ABCD”. “ABC” would also be concatenated with “CF” to form “ABCF” and this process would continue for all possible combinations between the two windows. However, connecting windows in this manner is a challenging task as it is not guaranteed that there will be connecting events, which can be seen in the subsequences “AD” in W_1 and “CF” in W_2 where the ending event of ‘D’ in “AD” is not equal to ‘C’ in “CF”.

Another issue with this simple method of concatenation is that there may exist extra information between the frequent subsequences that could be valuable information. In the example of “ABC” and “CD” forming “ABCD”, there may be information in the dataset records that exists between these two subsequences. For example, there may be a record that consists of “ABCEFC” which would indicate that the concatenation “ABCD” is not representative of all records in the dataset. While analyzing an event dataset, researchers are interested in all of the varying paths that sequences can take and this simple concatenation would not represent the smaller, hidden patterns.

Therefore, we recognize that a single event sequence may contain one of two types of information in between its two windows: either noise or no information at all; and to resolve the issues present in the simple concatenation, we develop a method WindowMerge that conjoins windows by using both a direct connection and a wildcard connection, where a direct connection is indicative of no information in between subsequences and a wildcard connection is indicative of extra information. Using this methodology, we can iteratively combine the frequent patterns found in adjoining windows to create an overall flow of events throughout a dataset. We include our WindowMerge in Algorithm 1 of Appendix A.

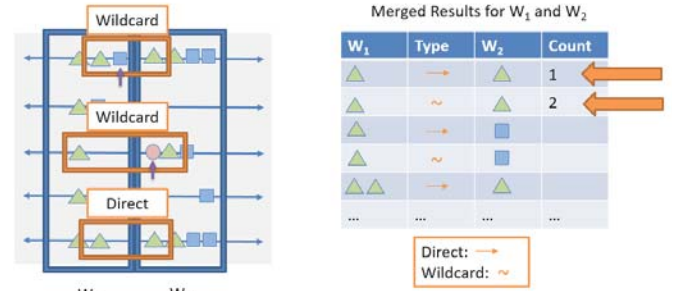


Fig. 5: An example of running WindowMerge on windows W_1 and W_2 . On the right the table is filled in both direct (orange arrow) and wildcard (orange ~) connections for each combination of frequent subsequences found in Figure 4. On the left, the connections for the first two rows are shown and how the counts are determined.

Breaking this method down WindowMerge takes two windows to connect, W_i and W_j , where $W_i < W_j$, and returns the number of direct connections, C_d , and the number of wildcard connections, C_w between the two windows. The method’s execution begins by filtering the dataset of sequences D to only look at the section of the sequences that span the windows such that: $S_{ij} = D \in \{W_i, W_j\}$. Then for each frequent subsequence in each window, $F_{im} \in W_i$ and $F_{jn} \in W_j$ where m is the m th frequent subsequence in W_i and n is the n th frequent subsequence in W_j , it identifies the index locations of these frequent subsequences

in the sequences of the windows such that the index locations are at the closest possible indices to the connection between the windows: $L_{im} = \max(\text{index}(S_{ij}, F_{im}))$ and $L_{jn} = \min(\text{index}(S_{ij}, F_{jn}))$. By choosing the indices closest to the edge where the two windows meet, this ensures that each sequence can only be counted once in the case of repeated events. This concept will be visible in our example that we will show in Figure 5.

Then for each pair of index locations that exist, the method identifies if there exist events between the locations or not. If the locations contain no events in the middle, $L_{jn} - L_{im} = 0$, then the direct connection count for that pair of frequent subsequences is incremented $C_d \leftarrow C_d + 1$, or otherwise if $L_{jn} - L_{im} > 0$ then the wildcard connection is incremented $C_w \leftarrow C_w + 1$.

An example of the WindowMerge method is shown in Figure 5 where the windows W_1 and W_2 from Figure 4 are merged together. In Figure 5, the table on the right is populated with the combination of frequent subsequences that were obtained previously for both windows. Each respective combination is connected via a direct connection and a wildcard connection which can be seen by the arrows and tilde in the table. To compute the number of occurrences for the first combination of frequent subsequences, which corresponds to the first two rows of the table, Figure 5 highlights where the subsequences from the combination occurs in the dataset between the windows. In this example it can be seen that the first sequence contains an extra blue square between the two green triangles indicating that the triangles are joined together by a wildcard connection, then in the third sequence there exists a red circle in between the triangles indicating that it is also a wildcard connection, and finally in the last sequence there exists no extra events between the two green triangles indicating that they have a direct connection between each other. Therefore, the table is filled in with a value of one for the direct connection and a value of two for the wildcard connection. This process would continue for the rest of the combinations of frequent subsequences to obtain the final result between windows.

In this example that we have presented, we can see the issue of repeated events that was mentioned in the definition of the window connections: the first and last sequences in Figure 5 contain multiple triangle events which would alter the counts that we obtain. Had each possible index location been considered for these two sequences then each possible pair of triangles would have been counted resulting in the first sequence containing four wildcard connections and the last sequence containing three wildcard and one direct connection. However, counting multiple connections for a single sequence would present the assumption that the sequence takes multiple varying paths which would not be accurate against the premise that each sequence is a series of events. For example, in the first sequence if we evaluated both *triangle, triangle, square, triangle* and *triangle, square, triangle* then we would be making the assumption that the sequence went through both of these subsequences at different times when in reality the second is just a subset of the first. By misrepresenting the number of sequences in the dataset, we would not be able to correctly evaluate the task of identifying the most common trajectories that sequences followed as a minority trajectory consisting of repeated events would become heavily weighted and lead to inaccurate insights of the dataset.

5.2.2 Algorithm Evaluation

To show an example of how our AWP-Span algorithm is able to provide a reduction and a summary of the overall sequences in a dataset, in Figure 6 Left we should seven sequences contained in a dataset prior to our algorithm being run. This Sankey diagram is limited to only seven sequences due to space limitations, but with these sequences we can already see the complexity that sequences may contain and the difficulty in analyzing these sequences. Furthermore, this Sankey diagram only contains seven sequences and thus the complexity would increase exponentially as the size of the dataset shown would increase.

With this, we show the results of our algorithm in Figure 6 Right where the mined patterns are displayed on the Sankey diagram and the yellow nodes correspond to wildcard connections. We will describe the approach that we take in generating this Sankey diagram in more detail in Section 6, but at first glance we can already see the reduction

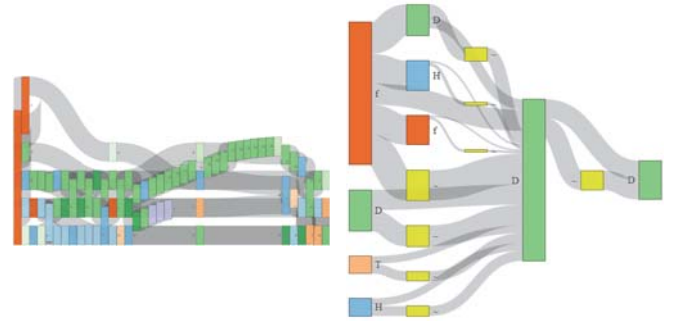


Fig. 6: An example showcasing the reduction that our AWP-Span algorithm performs. (Left) Shows a Sankey diagram of 10 sequences in the dataset, without any reduction or algorithm run. Only 10 sequences are able to be shown due to the space limitations. (Right) Shows the results produced by our AWP-Span algorithm with window size of 30 for the large dataset of sequences. The yellow nodes correspond to wildcard connections.

in complexity and the ease of identifying patterns in our algorithm's results.

5.2.3 Trajectory Analysis

A common challenge faced by researchers when analyzing temporal sequences is exploring the trajectory from one event to another. For example, a researcher commonly asks the question: "What path did sequences follow to get from A to C?". The path that is followed can vary from being short in length, "A -> B -> C", or very long, "A -> B -> E -> F -> G -> C". Being able to understand the trajectory followed, no matter the length of the paths is an interesting area of research and useful for many applications. We embed the ability to build such a trajectory into our algorithm by providing the option to specify beginning, E_b , and/or ending, E_e , events. When either of these events is set, then the algorithm will only consider frequent subsequences that occur between the two events and prune any incoming connections to the beginning event or any outgoing connections from the ending event. We will now define the pruning that our algorithm employs.

Given a sequence, S_1 , our algorithm will identify the index of the first occurrence of E_b such that $I_b = \text{index}(S_1, E_b)$ and the index of the last occurrence of E_e such that $I_e = \text{index}(S_1, E_e)$. Then for each frequent subsequence that is encountered in the algorithm, S_f , it will identify the index of the subsequence in the encompassing sequence such that $I_f = \text{index}(S_1, S_f)$. If the index of the subsequence falls between the first and last occurrence of the beginning events and ending events then it will be considered ($I_b \leq I_f \leq I_e$). However, if it does not fall within the bounds then the algorithm will skip that frequent subsequence.

To remove the connections leading into the beginning event and stemming from the ending event, anytime that the algorithm encounters a beginning event E_b in a frequent subsequence, no incoming connections are made to the event. For example, a subsequence of $\langle A, B, C \rangle$ in window W_i with a beginning event $E_b = 'B'$ will cause the subsequence to be pruned to $\langle B, C \rangle$ as the "A" event occurred before the beginning event. In addition, any frequent subsequences in the prior window, W_{i-1} , will not make a direct or wildcard connection to this subsequence.

Conversely, anytime that the algorithm encounters an ending event E_e in a frequent subsequence, no outgoing connections are made from the event, such that a subsequence of $\langle D, E, F \rangle$ in window W_j with an ending event $E_e = 'E'$ will be pruned to $\langle D, E \rangle$ and no connections to the window W_{j+1} will be made.

By allowing for a trajectory to be built from the first beginning event to the last ending event, there exists the possibility that there may be multiple beginning or ending events contained within these bounds. We do not attempt to remove these repeated events as keeping them allows for researchers to better understand varying pathways. For example, if the beginning event exists in multiple windows then this indicates that the relevant trajectories may have begun at different time points which

can assist in providing researchers the ability to accomplish their task of understanding all of the various paths that the sequences within their dataset can follow.

5.2.4 Adaptive Windows

While the AWP-Span algorithm that we have developed is effective at identifying common subsequences or building a trajectory within a temporal context, we recognized through test runs of our algorithm that identifying the optimal windows to split a dataset on is challenging and can be the deciding factor in obtaining a meaningful result for a specific domain. For example given a sample dataset that contains events on an interval unit of a day and events that are sparsely distributed over the sequences: where a lot of events occur at the beginning of sequences but very few events occur at the end of sequences; setting a fixed window size of $\omega = 30$ days would cause for the number of events in later windows to be significantly less than those in earlier windows. With this difference important patterns may become “chunked” together in earlier windows or not enough events may exist in later windows for patterns to be identified. In addition, identifying the exact number of windows that should exist is another difficult challenge similar to that of identifying the optimal number of clusters in clustering algorithms [7, 12, 28]. Because of these challenges, it would be beneficial to provide an optional functionality to automatically identify the optimal number of windows and for these windows to be varying in size such that areas where data is sparse would be combined and areas of density would be split.

To provide a solution to this challenging problem, we devised a complementary algorithm that returns the optimal window configuration, W , to be used as input to the AWP-Span algorithm. We refer to this algorithm as “AdaptiveWindows” as it adapts the windows such that they fit the dataset and domain that are being analyzed. This adaptation of windows is accomplished by ensuring that the windows are expanded or collapsed such that the number of sequences embedded within each window is similar. We present the algorithm in Algorithm 2 of Appendix A and will provide a brief explanation.

The AdaptiveWindow algorithm takes a dataset of event sequences, D , an application specific window measurement, n , and returns a list of windows with their respective upper and lower bounds. We introduce the idea of an application specific window measurement as it allows for domain knowledge to be embedded into the configuration of the windows. The window measurement, n , is a window size that represents how data data in the domain is usually segmented or it can be a value that the researcher believes would be the best fit for segmentation through windows. For example, in a financial domain a researcher may input a quarter of a year for the window measurement as sales data is typically measured per quarter. In another domain, such as sports, a researcher may choose to split the data based on possession of a ball as changes in possessions may lead to different types of patterns. By providing an avenue to embed domain knowledge and application specific requirements into our AdaptiveWindow algorithm, the windows returned will fit both the dataset and the domain.

Breaking down the algorithm it first creates a temporary list of windows: $W_i = \langle T_0, T_1, \dots, T_k \rangle$ where each temporary window is equally spaced such that they are all of size n : $(T_1 - T_0) == n$. As each sequence in the dataset, D , contains a list of itemsets, I , and each itemset is denoted by a timestamp such that I_i is the timestamp of when the events in I occurred; the first window in the temporary windows begins at the earliest timestamp, t_0 , that exists in the dataset. Building from there the the window boundaries occur at $t_0 + (n * i)$ where i is the window index. Using this domain specific window measurement the algorithm places the itemsets of the data into the respective temporary window in which they occur. Once the temporary windows have been built the number of itemsets in each window is calculated, $|T_k|$, and then averaged together to identify the average across the dataset, $A = \langle |T_0|, |T_1|, \dots, |T_k| \rangle$. This average itemsets, A , represents the average number of events contained in each window when spaced by the domain measurement n .

Once the average has been calculated it is then used to construct the optimal list of windows, W_o , where each window contains approxi-

mately A itemsets. By constructing windows that contain approximately A itemsets each, this means that the final set of windows will be reflective of the domain that was input into the algorithm. To construct the optimal windows, W_o , a famous method of discretization, also known as equal-frequency binning, is utilized [9].

With the optimal windows identified, both sparse and dense areas of events in the dataset will be eliminated allowing for windows of varying size that contain approximately an equal amount of data.

6 VISUALIZATION

Once we obtained the results from our AWP-Span algorithm, we sought to visualize the overall flow of events and understand the various connections between windows. We show the results of our algorithm on two different visualizations: the Sankey diagram and our own event summary design.

6.1 Mapping Common Sequences

Due to the atypical nature of our algorithm results which contain various types of connections (direct and wildcard), we needed to devise a mapping from the results obtained to the Sankey diagram. Next we will describe the mapping that we have defined.

First, we represent each event as a node in the Sankey diagram and place the restriction that each window in our results must contain its own separate set of nodes, such that the windows are initially not connected. By requiring each window to have its own set of nodes, this ensures that the Sankey diagram contains a temporal context as events that happen in different locations of a dataset (in different windows) are not equivalent to one another. To build each window’s set of nodes we convert the frequent subsequences contained in the window to a series of nodes such that similar events that occur in the same index of a subsequence are merged. For example, given a window that contains the frequent subsequences $S_1 = \langle A, C, B \rangle$ and $S_2 = \langle C, D, A \rangle$, C would be represented as a single node due to it occurring in the same itemset index 0, but A would not be merged as it occurs in different indices 0 and 1.

Second, for each pair of nodes in a window that have different index positions, an edge is drawn in between them. Using the previous example of S_1 and S_2 , two edges would be drawn from the node C : an edge from C to B and an edge from C to the A node in the second index. As each edge represents the flow from one event to the next, we encode the number of times that the respective transition occurred between the events into the edges.

Third, we link adjoining windows, W_i and W_j , through two methods: a direction connection and a wildcard node. Given the frequent subsequences of each window, $F_{im} \in W_i$ and $F_{jn} \in W_j$, these two methods are utilized on the last event in each F_{im} such that $E_{im} = F_{im}(-1)$ and the first event in each F_{jn} such that $E_{jn} = F_{jn}(0)$. For the direct connection the two events, E_{im} and E_{jn} , are connected by an edge that is encoded with the direct count between F_{im} and F_{jn} that was identified by our algorithm.

For the wildcard node an extra node is drawn between the two events, E_{im} and E_{jn} , which is labeled as a “wildcard node”. Two edges are then drawn: (1) from the event E_{im} to the wildcard node and (2) from the wildcard node to the event E_{jn} . Both of these edges are encoded with the wildcard count that was returned by the algorithm.

While the wildcard node is powerful at representing extra information that is contained between two events, for some visualization techniques this creates extra nodes and thus noise. Therefore, to reduce this noise and clutter caused by the wildcard nodes we alter the mapping so that instead of creating a wildcard node, a colored edge between E_{im} and E_{jn} is drawn. This change reduces the amount of nodes in the graph and still allows for the difference between direct and wildcard connections to be understood through visual inspection of the graph.

6.2 Sankey Diagram

A popular visualization, the Sankey diagram, has been used to understand the flow of resources in data [31]. We utilize the Sankey diagram by mapping our algorithm’s results to the diagram which allows for us to visualize the flow of events in the various trajectories.

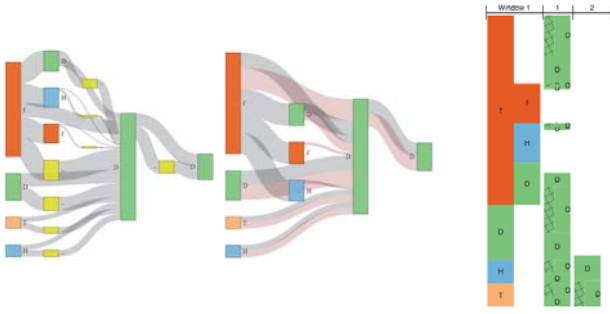


Fig. 7: An example output of our AWP-Span algorithm mapped to a Sankey and our event summary diagrams. (Left) Shows a Sankey diagram containing wildcard nodes causing for extra clutter and noise to exist. (Middle) Shows a Sankey diagram with the wildcard nodes converted into red-colored edges that can be expanded with a click to show the underlying wildcard data which provides multiple levels of detail. (Right) Our event summary diagram where the edges have been removed to decrease the analytical capacity needed to understand the overall flow of events. Wildcard edges are now mapped to a texture rather than a red-colored edge and windows are separated by a vertical gap.

An example of a Sankey diagram produced by mapping our algorithm is shown in Figure 7 Left and Middle where wildcard nodes and edges are used respectively. Through visual inspection we can see that Figure 7 Middle contains less clutter and is easier to understand due to the encoding of the wildcard into the edges.

Analyzing Figure 7 Middle we can see the overall flow of events: first an event of f is the most frequent while D , T , and H also occur at much lower frequencies. Next, the events D , T , and H proceed to another D with the sequences being split about 50% for going to the second D on a direct connection or with extra, infrequent events in between. At the same time, the f event proceeds to a different D , f , and H before ultimately ending up at the same D as the previous events. Once all of the events have reached the large D on the right side of the diagram, only a small portion move further into another repeated event of D . The reason that there exists a small path out of a large event node is that the final D is infrequent in its window relative to the other windows. This means that the D event is a frequent subsequence in its own window, but that the other windows have frequent subsequences of a much higher frequency.

Through this analysis of the Sankey diagram it can be seen how a business analyst could identify the most popular marketing channels, in this case the D channel would be indicative of being very popular, as well as understand the varying paths that users take: most users will end up at the largest D node but in the first column the D , T , and H are all just as likely to end up at D as the f channel. Furthermore, in the healthcare domain a clinician would be interested to see that most patients end up with a diagnosis of D , but that the paths that the patients took differed.

While the Sankey diagram is able to show the flow of sequences between events for all of the windows, it suffers in its visual complexity and ease of use. When analyzing the Sankey diagram, a user is required to follow the edges out from a node to be able to understand the destination events because of the fact that edges can be curved and are overlapped in the diagram. In addition, as the size of the data being represented increases, and in our case the number of windows, the size of the diagram grows significantly causing for it to be difficult to obtain a general overview of the data. These are limitations of the Sankey diagram which directly apply to the data that is attempted to be visualized from our AWP-Span algorithm, and thus we sought to identify a better method of visualizing the data.

6.3 Event Summary Diagram

As the Sankey diagram suffers from difficulties in identifying the flow of events with the spaced-out edges, as was identified by MatrixWave [41],

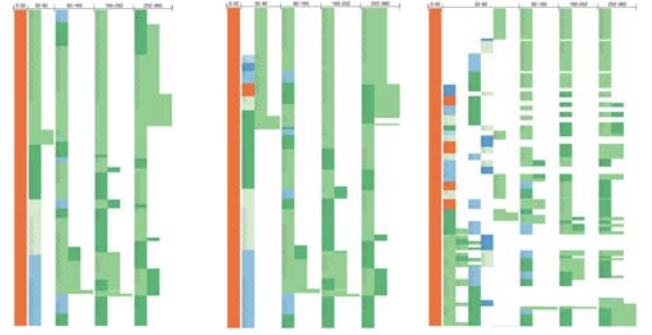


Fig. 8: An example showing the multiple levels of detail available upon clicking on a wildcard edge. In this example, the user clicked on subsequent wildcard edges to show the underlying data.

we developed an edge-free visualization that reduces the complexity of analyzing and understanding the different flow of trajectories. We will outline the steps taken to construct our event summary diagram to overcome these limitations.

First, in the Sankey diagram there is no requirement for event nodes to be aligned as the edge between a set of nodes is flexible to ensure that the connection is fully made. While this ensures flexibility, it also requires immense visual thinking to be able to thoroughly understand the connections. To overcome this issue we removed all edges in the graph and aligned each node next to the node that it originated from. This alignment causes for a clear path to be understood through a quick glance at the graph.

Second, as we removed edges between nodes, and thus removed the encoded wildcard edges, we textured a portion of the destination event node to indicate that the sequence passed through a wildcard prior to reaching this event. Adding a texture to the node adds minimal noise to the diagram while adding key information. We chose to only texture a small portion of the destination node over applying the texture to the entire node, because through a user evaluation we found that texturing the entire node created confusion as to which connection contained the wildcard: whether it was the connection into or out of the node.

Third, as our AWP-Span algorithm is built on the premise of dividing event sequences into windows, we added a vertical spacing between each set of windows so that through visual inspection it is possible to understand the frequent subsequences contained within the window. By ensuring that each window can be individually analyzed, users across many domains can understand a particular section of a dataset. For instance, a clinician could isolate a specific window to evaluate the most common diagnoses and paths followed within a certain time span.

Finally, as each window in our AWP-Span algorithm contains its own frequent subsequences, it is possible that certain windows will contain more sequences than another window. As an example situation: an event A could go to event B with a value of 5, followed by B going to C with a value of 10, where A , B , and C are in separate windows. In this example, the event B would be of size 5 while the event C would be of size 10 and following our first requirement of aligning nodes next to the next that they originated from, this difference in size would not be possible as there is not enough B to flow into C . Therefore, to overcome this limitation we utilize proportions on each window to ensure that the events can be drawn and that the relative frequency within a window is sustained.

We utilize the example data shown in Figure 7 (Left and Middle) to present our new design in Figure 7 Right. Through visual inspection, we can see that our new design contains the same information present in the Sankey diagram, but does so in a simpler and more easily understood manner. This ultimately decreases the analytical capacity required by the user and significantly reduces the virtual screen space occupied by the diagram.

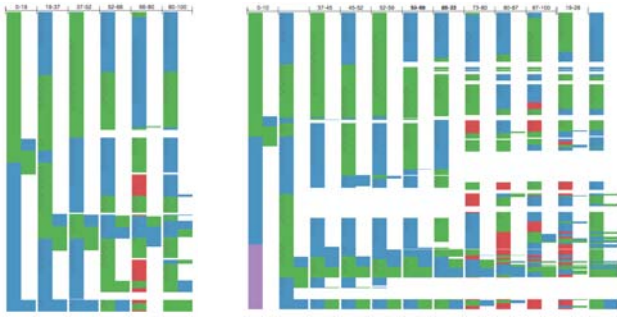


Fig. 9: A dataset of over 9,000 European soccer matches were summarized into event summary diagrams containing (Left) 6 windows and (Right) 12 windows.

6.4 Levels of Detail

As an extension to the wildcard edges produced, we recognize that the information contained in these connections may be interesting and contain useful information, thus we extend both of our visualizations to allow for expanding and collapsing of wildcard connections. Through an interactive action of clicking on a red-colored or textured wildcard edge, the wildcard is exploded such that the events contained within it are drawn as nodes in place of the edge. An example of this is shown in Figure 8, where a wildcard on the blue “H” node was clicked to display the underlying wildcard data. This provides both of our visualizations with multiple levels of detail by helping to identify certain patterns or trends embedded within the wildcards and allowing for further exploration of a dataset of sequences than what the overall summary is able to present.

7 CASE STUDIES

We demonstrate the universal capabilities of our algorithm and visualisation on three datasets from vastly different domains that contain a large amount of records and have been used in real world analysis scenarios. For the purposes of evaluating the insights that can be discovered within these datasets, we interviewed three analysts that each have experience in their own respective domains. An online interactive tool to demonstrate our approach with the different datasets is provided at <http://tf.visxplre.com>

7.1 Soccer Matches

Sports interest analysts and researchers because of the high volume of data that is produced throughout the course of a game: events are produced on and off the field and each player contains an assortment of attributes that can be learned from. For this case study, we utilized a Kaggle dataset that contains 9,084 soccer matches captured from the 2011/2012 season to the 2016/2017 season for the 5 biggest European leagues¹. These matches, in total, produced over 900,000 events which were categorized into 11 different event types. The generated event summary diagrams for two different window counts, n , are shown in Figure 9 where the images correspond to 6 and 12 windows respectively.

We first presented the sports analyst this dataset loaded into an existing event analysis system as we had previously shown in Figure 3. Then we presented our tool with the initial number of windows set to 6, as shown in Figure 9 Left. Immediately upon seeing this visualization, before even attempting to change the number of windows parameter, n , the analyst was intrigued as to the existence of the red events in the 5th window. Upon analyzing the legend, the analyst identified that the red event corresponds to a substitution event and confirmed that this is a typical behavior present in all soccer matches.

Next, the analyst began to change the parameter, n , and saw minimal changes in the visualization until they arrived at a value of 12 windows, shown in Figure 9 Right. At this window configuration, the purple event became present for the first time and through these series of parameter

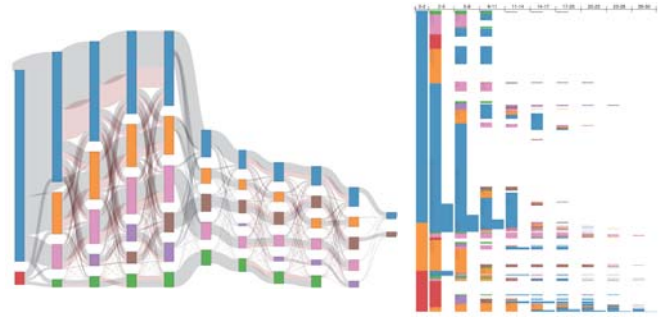


Fig. 10: A web traffic dataset containing 989,818 users was summarized using our AWP-Span to produce a (Left) Sankey diagram and (Right) event summary diagram.

changes the analyst highlighted to us that they noticed two different patterns. The first pattern that they said they were able to uncover is that the purple event, which corresponds to a corner kick, did not appear until they had reached 12 windows. The analyst quickly began to think out loud and realize that as the number of windows increased, the size of each window decreased. Thus they concluded that corner kicks were not frequent enough to appear in the larger size windows, but when confined to only the first 10 minutes of a game corner kicks were frequent enough. This was a very interesting trend to the analyst that they wished to investigate further.

The second pattern that they highlighted and were puzzled by were the large presence of green and blue events, corresponding to free kicks and attempts on goal respectively. The analyst eagerly opened the dataset in a statistical software and produced a histogram of the event types over the dataset where they noted that the green and blue events dominate the entire dataset. Through this they were able to identify a flaw with the dataset before performing any potential future analysis.

The analyst concluded by stating that compared to the existing event analysis system, our visualization was key in performing a top-down analysis of the dataset rather than a bottom-up approach. Most notably, they did not find it possible to arrive at the same insight of substitutions occurring in the later parts of a game in the existing system as its purpose is not catered to capturing a summary of the contained events.

7.2 Web Traffic Logs

We performed the next case study on web traffic logs with an analyst that manages front-end developers and is in charge of ensuring that a system is optimized for the end users. Traffic and access log files are an interesting data source to researchers from many different perspectives ranging from cybersecurity [15, 21], to advertising and marketing purposes [14, 18], to website and usability optimization [6, 17]. However, these datasets commonly exist in large data sizes to the extent of gigabytes and even terabytes of data. At these large scales, it is paramount that analysts are able to fully understand the dataset that they are presented with.

For this case study, we utilized a dataset released by msnbc.com containing the server logs for the entire day of September 28, 1999 [5]. The dataset contains 989,818 tracked users which make an average of 5.7 visits across 17 different categories of pages on the website. Due to extreme outliers present in the data we removed the 10,000 users that made over 50 clicks, leaving us with a still sizable dataset of 979,133 users.

We presented the web traffic analyst with the summary mined from our AWP-Span algorithm mapped to both a Sankey diagram and our event summary diagram as displayed in Figure 10. The analyst immediately was offput by the Sankey diagram, and stated that the only insights that they could safely discover in this “mess” were that users typically visit the blue page first, corresponding to the frontpage of the website, and that over time less users clicked due to the smaller nodes in the Sankey diagram.

However, upon switching to our event summary diagram, the analyst

¹<https://www.kaggle.com/secareanualin/football-events>

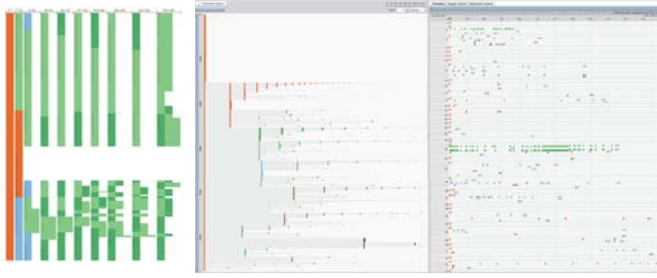


Fig. 11: (Left) An event summary diagram for EHR data. (Right) The same EHR dataset loaded into an existing event analysis system.

began to quickly pick out interesting patterns and trends. They identified that most users stay on the frontpage between clicks, because of the large series of blue in the middle of the diagram, but that some of them move to the orange (news), red (as seen tv), and pink (sports) before coming back to the frontpage, as seen in the upper left section. Next they eagerly pointed to the bottom section of the diagram where the initial red events lead into repeated orange events that continue through all of the windows, which indicates that users that visit the website through the “as seen on tv” (red) section will stay on the website for a long time, clicking through different news (orange) pages. They noted this as being a helpful insight into user retention and pointed out a comparison to the middle section of the diagram where users that entered the website through the frontpage leave after a significantly lower amount of clicks than those that enter through the “as seen on tv” section. Based on this insight, they stated that the frontpage could potentially be optimized to attempt to guide users towards these pages with higher retention rates or that more advertising of the website could be done on TV.

7.3 EHR Data

With the challenge of EHR data in mind, as discussed in Section 2, we performed a final case study with a large, complex, clinical dataset. This dataset consisted of 89,840 patients that have been diagnosed with a mild Traumatic Brain Injury (a.k.a concussion). The longitudinal data included 5.3 million TBI-related clinical encounters over 10 years and over 8.7 million clinical diagnoses (i.e. events), where a TBI-related encounter was defined as a visit to the doctor where the patient was treated with one or more of the conditions that are commonly known to be related to concussions such as behavioral disorder, sleep problems, cognitive deficiencies, and audiology complaints. The patients under consideration had an average of 59.06 encounters.

As a common task in evaluating and understanding TBIs involves analyzing the first year after an injury, we filtered the dataset to contain the first 365 days after a patient’s concussion before running our AWP-Span algorithm. A mined event summary is shown in Figure 11 Left and a subset of the dataset (due to performance limitations) loaded into an existing event analysis system is shown in Figure 11 Right.

The healthcare analyst looked at both types of visualizations displayed in Figure 11 and first noted three similarities: (1) they both show that patients that suffered a concussion (orange) and then had no more events available afterwards, (2) they both indicate that patients have repeated diagnoses, and (3) that Depression (light green) and PTSD (dark green) are very common to occur amongst concussion patients. Beyond these similarities, the analyst liked that the Figure 11 Right showed all of the different types of events, but they stated that they struggle to identify any patterns due to little change in diagnoses/colors for each record. Thus, they moved to the diagram shown in Figure 11 Left and expressed their satisfaction for time periods being present at the top of the diagram to aid in the understanding of the trajectory. With these time periods (window configurations), they found it interesting to see that the size of the windows increased in the later windows, indicating that patients required less healthcare and doctor visits as time went on. Additionally, the analyst noticed that the light blue (headaches) diagnoses only occurred in the first couple of windows after a concussion

which they confirmed follows the medical literature.

The analyst was further interested in understanding what diagnoses the patients had between diagnoses that our system did not find to be common and thus they began to click wildcard edges and arrived at a diagram similar to the one found in Figure 8 Right. Through the expansion of these wildcard edges, the analyst discovered that patients seemed to only have diagnoses of sleep (dark blue) near their concussion, but not in the later windows. They found this interesting and noted that there could be a relation between the diagnoses of sleep and headaches only occurring in similar windows.

8 CONCLUSION

Previously frequent sequence mining (FSM) algorithms had been used to understand the patterns contained within a dataset and did not take into account the location of each pattern. Through the development of our algorithm we have shown how the temporal context of each mined pattern can be utilized by breaking a dataset into temporal windows and constructing a trajectory of the various common paths that sequences may follow. Compared to displaying all possible event paths which suffers from scaling issues, our unique method of utilizing an FSM algorithm provides a condensed overview of an event dataset while still allowing for exploration at multiple levels of detail. In addition, due to the window nature of our algorithm the computation for each window can be run in parallel allowing for results to be obtained quickly.

Through our new novel event summary design we have reduced the visual complexity existent in a typical Sankey diagram. We have demonstrated the effectiveness and usefulness of our algorithm and event summary diagram through three case studies, where analysts from different domains were able to quickly find insights and identify the aid that our approach provides for solving tasks faced each and everyday by event researchers in large scale datasets.

REFERENCES

- [1] R. Agrawal, R. Srikant, et al. Fast algorithms for mining association rules. In *Proc. 20th int. conf. very large data bases, VLDB*, vol. 1215, pp. 487–499, 1994.
- [2] W. Aigner, S. Miksch, B. Thurnher, and S. Biffl. Planninglines: novel glyphs for representing temporal uncertainties and their evaluation. In *Information Visualisation, 2005. Proceedings. Ninth International Conference on*, pp. 457–463. IEEE, 2005.
- [3] R. Bade, S. Schlechtweg, and S. Miksch. Connecting time-oriented data and information to a coherent interactive visualization. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 105–112. ACM, 2004.
- [4] M. Burch, F. Beck, and S. Diehl. Timeline trees: visualizing sequences of transactions in information hierarchies. In *Proceedings of the working conference on Advanced visual interfaces*, pp. 75–82. ACM, 2008.
- [5] I. V. Cadez, D. Heckerman, C. Meek, P. Smyth, and S. White. Visualization of navigation patterns on a web site using model-based clustering. In *KDD*, p. 280, 2000.
- [6] E. H.-h. Chi. Improving web usability through visualization. *IEEE Internet Computing*, 6(2):64–71, 2002.
- [7] M. M.-T. Chiang and B. Mirkin. Intelligent choice of the number of clusters in k-means clustering: an experimental study with different cluster spreads. *Journal of classification*, 27(1):3–40, 2010.
- [8] S. B. Cousins and M. G. Kahn. The visual display of temporal information. *Artificial intelligence in medicine*, 3(6):341–357, 1991.
- [9] J. Dougherty, R. Kohavi, M. Sahami, et al. Supervised and unsupervised discretization of continuous features. In *Machine learning: proceedings of the twelfth international conference*, vol. 12, pp. 194–202, 1995.
- [10] J. A. Fails, A. Karlson, L. Shahamat, and B. Shneiderman. A visual interface for multivariate temporal data: Finding patterns of events across multiple histories. In *Visual Analytics Science And Technology, 2006 IEEE Symposium On*, pp. 167–174. IEEE, 2006.
- [11] D. Gotz and H. Stavropoulos. DecisionFlow: Visual Analytics for High-Dimensional Temporal Event Sequence Data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1783–1792, Dec. 2014. doi: 10.1109/TVCG.2014.2346682
- [12] G. Hamerly and C. Elkan. Learning the k in a_k means. *Advances in neural information processing systems*, 16:281, 2004.

- [13] B. L. Harrison, R. Owen, and R. M. Baecker. Timelines: an interactive system for the collection and visualization of temporal data. In *Graphics Interface*, pp. 141–141. Citeseer, 1994.
- [14] H. Hochheiser and B. Shneiderman. Coordinating overviews and detail views of www log data. In *Workshop on New Paradigms in Information Visualization and Manipulation (NPIM 2000)*, ACM Press, 2000.
- [15] T. Itoh, H. Takakura, A. Sawada, and K. Koyamada. Hierarchical visualization of network intrusion detection data. *IEEE Computer Graphics and Applications*, 26(2):40–47, 2006.
- [16] B. C. Kwon, J. Verma, and A. Perer. Peekquence: Visual analytics for event sequence data. In *ACM SIGKDD Workshop on Interactive Data Exploration and Analytics (IDEA 2016)*, 2016.
- [17] J. Lee and M. Podlaseck. Visualization and analysis of clickstream data of online stores with a parallel coordinate system. In *International Conference on Electronic Commerce and Web Technologies*, pp. 145–154. Springer, 2000.
- [18] J. Lee, M. Podlaseck, E. Schonberg, and R. Hoch. Visualization and analysis of clickstream data of online stores for understanding web merchandising. In *Applications of Data Mining to Electronic Commerce*, pp. 59–84. Springer, 2001.
- [19] T. Li. *Event Mining: Algorithms and Applications*, vol. 38. CRC Press, 2015.
- [20] Z. Liu, Y. Wang, M. Dontcheva, M. Hoffman, S. Walker, and A. Wilson. Patterns and sequences: Interactive exploration of clickstreams to understand common visitor paths. *IEEE Transactions on Visualization and Computer Graphics*, 23(01), 2016.
- [21] K.-L. Ma. Cyber security through visualization. In *Proceedings of the 2006 Asia-Pacific Symposium on Information Visualisation-Volume 60*, pp. 3–7. Australian Computer Society, Inc., 2006.
- [22] F. Massegia, M. Teisseire, and P. Poncelet. Sequential pattern mining. *Encyclopedia of Data Warehousing and Mining*, pp. 1028–1032, 2005.
- [23] T. Mitsa. *Temporal data mining*. CRC Press, 2010.
- [24] M. Monroe, R. Lan, H. Lee, C. Plaisant, and B. Shneiderman. Temporal Event Sequence Simplification. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2227–2236, Dec. 2013. doi: 10.1109/TVCG.2013.200
- [25] J. Pei, J. Han, B. Mortazavi-Asl, J. Wang, H. Pinto, Q. Chen, U. Dayal, and M.-C. Hsu. Mining sequential patterns by pattern-growth: the PrefixSpan approach. *IEEE Transactions on Knowledge and Data Engineering*, 16(11):1424–1440, Nov. 2004. doi: 10.1109/TKDE.2004.77
- [26] A. Perer and D. Gotz. Visualizations to support patient-clinician communication of care. In *ACM CHI Workshop on Patient-Clinician Communication*. Paris, France, 2013.
- [27] A. Perer and F. Wang. Frequence: Interactive Mining and Visualization of Temporal Frequent Event Sequences. In *Proceedings of the 19th International Conference on Intelligent User Interfaces, IUI '14*, pp. 153–162. ACM, New York, NY, USA, 2014. doi: 10.1145/2557500.2557508
- [28] D. T. Pham, S. S. Dimov, and C. Nguyen. Selection of k in k-means clustering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 219(1):103–119, 2005.
- [29] D. Phan, A. Paepcke, and T. Winograd. Progressive multiples for communication-minded visualization. In *Proceedings of Graphics Interface 2007*, pp. 225–232. ACM, 2007.
- [30] 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 *Proceedings of the AMIA Symposium*, p. 76. American Medical Informatics Association, 1998.
- [31] P. Riehmann, M. Hanfler, and B. Froehlich. Interactive Sankey diagrams. In *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005*, pp. 233–240, Oct. 2005. doi: 10.1109/INFVIS.2005.1532152
- [32] K. Vrotsou, K. Ellegård, and M. Cooper. Everyday life discoveries: Mining and visualizing activity patterns in social science diary data. In *Information Visualization, 2007. IV'07. 11th International Conference*, pp. 130–138. IEEE, 2007.
- [33] K. Vrotsou, J. Johansson, and M. Cooper. Activitree: interactive visual exploration of sequences in event-based data using graph similarity. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):945–952, 2009.
- [34] 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 *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 457–466. ACM, 2008.
- [35] K. Wongsuphasawat and D. Gotz. Outflow: Visualizing patient flow by symptoms and outcome. *IEEE VisWeek Workshop on Visual Analytics in Healthcare, Providence, Rhode Island, USA*, pp. 25–28, 2011.
- [36] K. Wongsuphasawat, C. Plaisant, M. Taieb-Maimon, and B. Shneiderman. Querying event sequences by exact match or similarity search: Design and empirical evaluation. *Interacting with computers*, 24(2):55–68, 2012.
- [37] K. Wongsuphasawat and B. Shneiderman. Finding comparable temporal categorical records: A similarity measure with an interactive visualization. In *Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on*, pp. 27–34. IEEE, 2009.
- [38] D. A. Wooff and J. M. Anderson. Time-weighted multi-touch attribution and channel relevance in the customer journey to online purchase. *Journal of Statistical Theory and Practice*, 9(2):227–249, 2015.
- [39] M. J. Zaki. Spade: An efficient algorithm for mining frequent sequences. *Machine learning*, 42(1-2):31–60, 2001.
- [40] E. Zraggen, S. M. Drucker, and D. Fisher. (s—qu)eries: Visual regular expressions for querying and exploring event sequences. *Proceedings of CHI 2015*, 2015.
- [41] J. Zhao, Z. Liu, M. Dontcheva, A. Hertzmann, and A. Wilson. Matrixwave: Visual comparison of event sequence data. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 259–268. ACM, 2015.

A Reinforcement Learning Approach to Automatically Model and Learn from User Interactions

Category: Research

Paper Type: Algorithm/Technique

Abstract—Despite the growth of visual analytics and the widespread use of visualization tools, the mining and modeling of user interactions and behaviors have not kept pace despite the ability to provide critical information about the method in which users interact with the interface and visualization. Existing approaches contain substantial limitations as they rely on expertly crafted rules or for users to examine their own interactions. With this, we introduce the application of reinforcement learning onto user interactions within a visual analytic system, opening a new avenue of research for building models that can automatically learn and uncover hidden insights. In this paper we outline three research directions that can be addressed using reinforcement learning: (i) providing guidance to users towards the completion of a task, (ii) personalizing guidance to match each individual user’s behavior and preferences such that the guidance is “polite” and unobtrusive, and (iii) optimizing the visual analytic interface to maximize user performance. Then we present our approach focused on solving these research directions. We show how our approach builds a model that embeds knowledge about a visual analytic system and how it can automatically learn to tailor itself to a user’s preferences through the application of providing guidance. Finally, we manually validate the insights embedded within our model to show how it could be further applied to optimizing an interface. Our reinforcement learning-based models can be utilized with user data or in an automatic standalone method, and ultimately allow for analysis that is not possible using standard modelling techniques.

Index Terms—Visual Analytics, User Interactions, User Provenance, Machine Learning, Reinforcement Learning

1 INTRODUCTION

User interactions are comprised of the clicks and steps that are taken within a visualization tool or interface. Generally, each interaction requires careful thought and results in a wealth of knowledge embedded within each click. This hidden knowledge has the ability to provide critical information about the method in which users interact with the interface and visualization. However, despite the growth of visual analytics and the widespread use of visualization tools, mining and modelling these user interactions and applying the knowledge embedded within them continues to be a difficult and limited task.

Imagine a typical scenario where a user has a specific task or question to answer: they open a visualization tool, load their dataset, map attributes of the data to visual representations, employ filters to manipulate the data or visual representations, and continue to go through an analytical process of exploring the information until they reach a state where the task under consideration is complete. Throughout this process, no matter the knowledge level of the user, each interaction that they take leads them closer or further away from their goal as well as provides some value to the user: whether that be large or small. There are a wide range of questions that can be asked of these interactions, such as: Does the user seem to be confused by the visualization? Is the user lost and do they need assistance? How important is applying a specific filter onto the data? Can the interface be made better?

Current approaches to analyze user interactions are often not able to answer these questions in an automated fashion because they rely on users analyzing their own process through a historical timeline approach [28], on experts to construct specific rules [37, 48, 56], or on systems to build a graph of a collection of users without objective methods for assigning weights to identify “good” or “bad” paths [9]. To further extract the wealth of knowledge that is contained within the user interaction process, an approach that moves beyond the standard techniques, that can automatically learn by itself, and that can generalize across tasks and domains is needed.

With this, we introduce the application of reinforcement learning onto user interactions within a visual analytic system. Reinforcement learning has the ability to uncover the hidden knowledge and insights embedded within interactions, and most notably can be performed in an automatic manner. Through this introduction, we open up a new avenue of research that we hope will help move the field of analyzing user interactions into more automated approaches that can uncover old insights found in existing approaches as well as new insights that are

only possible with this new technique.

In this paper, we outline three different research directions that can be addressed using reinforcement learning: (i) providing guidance to users towards the completion of a task, (ii) personalizing guidance to match each individual user’s behavior and preferences such that the guidance is “polite” and unobtrusive, and (iii) optimizing the visual analytic interface to maximize user performance. Then we present our approach focused on solving these research directions. We show how our approach builds a model that embeds knowledge about a visual analytic system and how it can automatically learn to tailor itself to a user’s preferences through the application of providing guidance. Finally, we manually validate the insights embedded within our model to show how it could be further applied to optimizing an interface.

2 RELATED WORK

2.1 User Interactions

The capturing, modelling, and analysis of user interaction data for the use of understanding a user’s analytical process in a visualization has been an active research area in the Visualization [11, 13, 19, 29, 39], Human Computer Interaction (HCI) [1, 20, 36], and workflow mining communities [53, 54]. Multiple systems have been developed that provide users with manual methods of understanding their own interactions, which include: presenting a historical timeline of interaction data to the user and allowing for them to analyze their own analytical process [28], visually encoding a user’s interaction history to promote exploration [14], and providing users with the functionality to annotate and share their actions [38]. While these manual techniques allow for users to evaluate their own missteps, the reliance on users to correctly classify their own actions is problem-some for when users are lost and unsure of how to accomplish their task, or even for new tasks that they have never encountered.

To allow for an external evaluation of where users become lost in a visualization, researchers have attempted to understand interactions by capturing and analyzing the user’s analytic provenance by means of eye movement data [2, 26, 30, 40], interaction logs [21, 22], and thinking aloud videos [34]. This work has aided the understanding of a user’s thought process, but relies on an expert to identify the missteps.

With the goal of building systems that automatically learn and adapt to users, Gotz and Zhou defined the different tiers of the visual analytic

process [18] and Ragan et al. developed a framework of user provenance types to assist researchers in evaluating designs and capturing provenance information [42]. Building off of this defined interaction data, Xiao et al. utilized past visualization states and actions to construct an improved and more relevant visualization for future users [57], Hoque et al. proposed a conversational navigation through natural language processing in visual analytics [23], Brown et al. used low-level interactions, specifically mouse activity, to classify fast and slow users as well as personality traits [6], continuing with mouse activity work has utilized it for predicting the next click that a user will take [16, 32], Gotz and Wen built a system that analyzes users' actions using regular expressions to suggest alternate visualizations [17], VisTrails captured user actions in a visualization and analyzed the various branches in activity that users took [3], and Dabek and Caban developed a system that applies a grammar induction algorithm onto a group of user interactions to identify the various paths that users can take. Furthermore, a set of explicit rules were defined on when and how to provide assistance to users, but these rules did not adapt to user's style of analytical thinking [9].

2.2 Reinforcement Learning

While the previous work in modelling and understanding user interactions has helped to uncover some of the knowledge embedded within these interactions, there still exists the need for an approach that can automatically learn from interactions extract new insights and further our understanding. Reinforcement learning is well suited for this problem as it is traditionally flexible to learning a wide variety of tasks and the method in which it learns/trains can be directly translated into a user interface.

Reinforcement learning is a class of machine learning algorithms that learns how to best choose actions to accomplish a task. Recent successful applications of reinforcement learning have been to popular games such as Mario, Atari, and Go [35, 45, 52]. With these applications it has been shown that a simple reinforcement learning algorithm can be applied to a wide range of tasks and that even the same trained model could be utilized across different games, by means of transfer learning [15, 31, 51]. Because of this there exists the potential to utilize reinforcement learning beyond games and venture into other fields [33, 43].

The underlying principle behind reinforcement learning is that the algorithm initially does not know how to get to the end of a task, but is simply notified when it reaches the end of the game and the outcome that it had obtained. The challenging aspect is learning the correct moves to make and how to assign a value to each move that was made in the game. For example, the last move of a game does not necessarily correlate to the result of the game and thus the deciding move may have happened earlier on. In the instance of a game such as Mario, the model would not initially know that collecting the mushroom to grow in size is an important strategy to winning, but over time of playing the game it would learn to assign a larger weight/value to this move.

Therefore, reinforcement learning algorithms have the ability to build a value structure of the possible actions/moves by learning from experiences of playing a game repeatedly or performing simulations of a task.

2.2.1 Chess Example

Chess is a historic game that requires a lot of thought and future planning of moves. It is regarded as a highly intellectual game because while it is a relatively easy game to play, it is difficult to master. Each piece has a different method in which it can be moved and thus some pieces are more valuable than others.

The chess community has performed a lot of analysis into the different strategies of chess as well as identified and coined sets of moves to take. For example, a wide range of key opening moves that maximize a player's chance of winning have been defined [10]. These are moves that required a lot of exploration and analysis to be identified, but reinforcement learning models have been able to identify these types of powerful moves as well as ones that were never known [7, 45].

2.2.2 Definitions

Continuing to use the game of chess as an example, we will define the different elements of a reinforcement learning algorithm:

Agent The actor that takes actions and moves. (Ex. A human chess player or a computer chess player.)

Environment The world in which the agent interacts in. (Ex. The chess board.)

State The current situation of the environment that the agent is in. (Ex. The current situation of the chess board with the pieces in positions that they were last moved to.)

Action Possible moves for the agent to take within the environment. (Ex. Moving a chess piece from one position to another on the board.)

Reward A value assigned to the agent at defined steps of the process. (Ex. Assign +10 for winning and -10 for losing the game of chess.)

3 RESEARCH DIRECTIONS

To help introduce the concept of reinforcement learning on user interactions, we present possible research directions for exploration using these class of algorithms.

R1 Provide Guidance: Typically when confronted with a new task, visual analytic system, or a particularly challenging dataset users may have a difficult time identifying answers to their questions or new insights into their data. Because of this hurdle, users may take missteps or proceed towards a path in which does not lead them towards completing their task and they may ultimately become frustrated and discouraged. Therefore, it is important to ensure that users stay on track by guiding them when they get stuck or require assistance. Through reinforcement learning, it is possible for models that capture the value of interactions, controls, and visualizations within an interface in an automatic manner that does not require user data. This allows for a large variety of paths to be explored and for guidance to be able to be provided to users solving a new task, something that would typically require for a history of data before it would be possible.

R2 Polite Guidance: While determining what to suggest to a user when guiding them is an initial first step, it has been highly documented that providing assistance to users in a polite manner is important to not impede on already established workflows [24]. Therefore, it is crucial to identify when to provide suggestions to users to avoid being intrusive and widely disliked as Clippy in Microsoft Office famously was [55]. As shown in the related work, recent methods for user assistance have relied on manually creating rules to determine when to assist a user, but these methods lack the ability to conform to user preferences and dynamically adapt to the interface that they are being presented with. Therefore, based on this criteria reinforcement learning algorithms can improve how we assist by exploring how the user is interacting, determine when to assist them, and over time adapt to their preferences.

R3 Optimize Interface: The placement of elements within an interface have been shown to be important to the effectiveness of a user accomplishing tasks [25, 44, 47]. This can be extended to visual analytic systems in which the placement of buttons, filters, and visualizations are crucial to assist the user in understanding the data and going through their analytic process.

However, identifying the correct arrangement of an interface is a difficult task which typically requires extensive user studies, use of eye tracking devices, and other methods [12, 27]. These methods rely on the feedback of the users, but can be lackluster as users may not themselves know the best method or arrangement [46]. Additionally, studies rely on the statistics of the users' performance and typically do not place enough emphasis on the interactions performed.

Therefore, utilizing user interactions is a key way to identify which clicks are the most valuable and optimize the interface in

terms of the design and layout in order to maximize the users' performance and analytical understanding.

With these three research directions defined, we will present our approach that is focused on solving the challenge of providing guidance to users in a polite manner. We believe that this is the first approach to automatically learn user's behaviors to tailor the method in which suggestions are presented. Finally, we will show a manual method of learning to optimize an interface in order to show how it could be accomplished and to introduce this avenue of research for future work.

4 APPROACH

For the approach that we will build, we choose to utilize the Q-Learning algorithm [49] due to the interpretability of the underlying matrix that is built. Therefore, an in depth explanation of the algorithm is explored.

4.1 Learning by Playing

Q-Learning is an algorithm that attempts to learn a value for all possible actions within each state. Thus when the algorithm reaches a state, it picks the action with the most value for the given state. For this algorithm, the value is a function such that it is equal to the total reward that would be obtained after picking the associated action if the agent was to act optimally afterwards. The equation for the value can be thought of as:

$$Value(s_t, a_t) \leftarrow \sum_{t=goal} \max_a R(s_t, a) \quad (1)$$

where the sum of the maximum reward at all future states, up to the goal state, is computed.

Comparing this value function to a simple one that is traditionally used in basic intelligence agents it is possible to see how a value is learned. In a basic intelligence agent, the value may be defined as the percent chance of winning the game if the associated action is taken. This simple function ignores the usefulness of the action and has the potential to be improperly learned if the competing agent/player is also unintelligent and almost always loses. However, the value function used in Q-Learning takes into account the impact that taking the action has in a scenario where the agent would be optimal from that point on. It should be noted that while the agent may not initially be acting optimally after the selected action, over time it will learn to be optimal and update all of the actions together.

To obtain the associated value for all actions within each state, Q-Learning builds a matrix/table where the rows correspond to states and the columns correspond to actions. Each time that the agent needs to take an action, it looks at the state (row) that it is currently in and identifies the action (column) with the most value. Then upon taking the action, it receives a reward for taking it. In the most simplest of cases it would receive +10 for a win and -10 for a loss, with all other cases resulting in 0. Using this reward, the algorithm then updates the cell of the action that it had just taken. The value function which is used for this update is defined as:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (2)$$

where $Q(s_t, a_t)$ corresponds to the current cell of the action that was selected; r_{t+1} corresponds to the reward that was obtained upon taking the action; $\max_a Q(s_{t+1}, a)$ is the maximum reward possible at the future state that the agent just moved to (by looking at the entire row of the state that it had just moved to after taking the action); and α and γ are two parameters that are set prior to training the model, such that α is the learning rate, how much the newly acquired reward should override the old information, and γ is the discount factor, how much of the future reward should be taken into account. With this update function and states/actions the Q-Learning algorithm plays the number of games, n , that are set for it to play and builds the associated matrix.

4.1.1 Exploration

While a model is being trained, there may be instances where the algorithm has learned a specific path of moves to take in order to win, but may still not yet have arrived at the best moves to take. To overcome this there exists an instance of randomness present within the algorithm: at each point when the algorithm needs to identify the action with the most value for a given state, it occasionally will choose to select a random action. This ensures that the algorithm is able to continue to explore throughout its learning and potentially identify better actions to take than what it had initially learned.

4.2 Reward Structure

While a simple example of a reward structure is to provide a +10 for winning and a -10 for losing, there are alternate structures that be utilized to achieve different goals. For instance, with the simple reward structure of +10/-10 for a win/loss, it does not matter to the agent how many actions it takes to arrive at a win as long as it wins. This means that for the chess example, it could move around the chess board aimlessly for millions of moves before getting a win and it would be just as valuable of a win as one where it achieved the win in less than 20 moves. This type of example shows the importance of crafting the reward structure such that the desired outcome can be achieved.

To overcome this issue of the number of moves it takes to win, one could add an additional element to the reward structure such that the agent receives -1 for every action that it takes. With this updated structure the agent is still motivated most by a win/loss, but it will also attempt to find the shortest path to arrive at a win because of the constant penalty it receives from making a move. In another example, the reward structure could be modified to motivate the agent to eliminate opponent's pawns by assigning a +1 reward for each time that a pawn is taken by the agent. Through these examples of a movement penalty and pawn reward, it is possible to see how a carefully crafted reward structure, through the magnitude and sign of the rewards, is crucial to arrive at the goal set out.

4.3 Relating back to Viz

A visual analytic interface can be thought of as a game. Similar to chess, the interface can be easy to interact with but difficult to master. Directly relating to the definitions defined in Section 2.2.2 an interface can be thought of as:

Agent User/Robot interacting with the interface

Environment User interface

State Current configuration of the interface and visualizations

Action Clicking a button, filter, visualization, etc.

Reward Assign +10/-10 for completing/not completing a task and -1 for each click

Using these definitions a model could be built and trained for a given user interface. This model would be able to learn about user interactions similarly as it would for a game of chess: the best actions to take at the load of the system or actions which are crucial to ensure success are some examples of aspects that could be learned and identified.

5 VISUALIZATION SYSTEM

To evaluate the effectiveness of applying reinforcement learning algorithms within a visualization user interface, a coordinated multi-view visualization system was built. In this system we sought to create a real-world type scenario similar to those visualization frameworks/tools that are created by Tableau [50] and D3 [4]. In these dashboards, users typically look to either explore the data to uncover insights or have a specific question to answer. For both of these tasks, it is necessary for users to click around the interface to manipulate the view and alter the visualization that is being presented.

The visualization system that we built is shown in Figure 1. The various components of this system include a list of variables on the left-hand side, a clickable map in the top center with controls for switching between partitioning the USA into regions or states, a multiple bar chart in the bottom center with a control for switching between three different graphs and the ability to change the year. Finally, at the top of



Fig. 1: A screenshot of our visualization system built for capturing user interactions in a real-world type scenario.

the screen there is a question along with four multiple answer choices at the top right.

5.1 Dataset

To mimic a real-world scenario, we chose to represent a dataset in the system that would not easily be understood and thus would require exploration of the data for insights to be discovered. Because of this requirement we focused on healthcare data due to its complexity and inherent lack of understanding compared to standard datasets such as weather, sports, etc. The final dataset chosen consists of medicare provider utilization and payment data taken from the Centers for Medicare and Medicaid Services (CMS) website¹. The dataset covers the years 2012, 2013, and 2014 and provides an assortment of variables for each physician: the number of services that they performed, the number of patients that they saw, the average amount that they were reimbursed for services, the average age and race breakdown of patients, etc.

5.2 Questions

For the study, we devised 10 multiple choice questions that were similar to the types of questions asked by media outlets, doctors, patients, and researchers in the healthcare domain^{2 3} and constructed a series of questions. We ensured that each question required interaction for the answer to be visible in the corresponding visualizations. A listing of the 10 questions can be found below:

1. In 2012, how many services did female nurse practitioners perform in the East South Central region?
2. What specialty submitted the highest medicare charges in Nebraska (NE) for 2014?
3. In 2012, which gender submitted the largest payments in New Mexico (NM)?
4. For all physician assistants in Texas (TX) in 2013, which race had the second highest submitted payments?
5. For the Mountain region in 2013, which age group had the highest reimbursements over all specialties?
6. Over all degrees and genders, which specialty had the highest submitted payments for 2014 in Kansas (KS)?
7. For the South Atlantic region in 2012, which specialty had the most number of patients in the 75 to 84 age range?
8. How much were physical therapists reimbursed for white patients in the New England region in 2012?
9. For physician assistants, which gender was able to submit higher amounts in Nevada (NV) for 2012?
10. For family practice physicians in Florida (FL) in 2013, which age group offered the second highest reimbursement?

¹<https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/medicare-provider-charge-data/physician-and-other-supplier.html>

²Wall Street Journal: <http://graphics.wsj.com/medicare-billing/>

³Washington Post: <http://wapo.st/1hAFRFE>

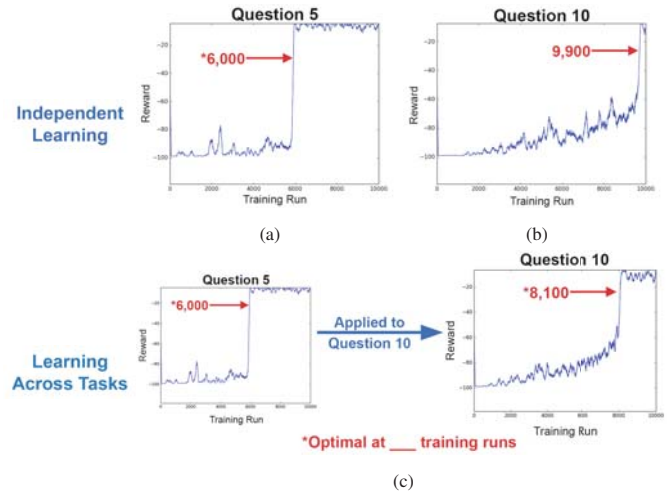


Fig. 2: (a & b) A separate blank model was trained on (a) question 5 and (b) question 10. It can be seen that question 5 became optimal quicker than question 10 and thus can be concluded to be an easier task. (c) The model trained independently on question 5 is used as the starting point for learning question 10 and produces the resulting graph. This shows that the question 5 model had embedded knowledge about the interface and can be applied across tasks.

6 BUILDING MODELS

With our visual analytic system and associated dataset defined, models can be built to solve the tasks presented by the posed questions. As typical applications of reinforcement learning are performed in an automatic way such that a bot undergoes training runs/simulations through the environment in order to build a model, it is similarly possible to do this with a user interface and treat the bot as simulating a user. For a user interface, a bot would pick the different actions to take to manipulate the interface until it is notified of having reached the goal, which is defined as the state at which the solution to the question/task is displayed within the interface.

Because the bot is simply looking for the state at which the solution is visible and the question can be answered, there exists only a single goal state which we define as being the “winning” state. Thus, after attempting multiple reward structures we arrived at one that optimizes for arriving at the goal state quickly: assigning -1 for each action taken and a 0 once arriving at the goal state. With this structure the agent learns to not explore endlessly, but rather to find the goal state to stop being penalized.

To accomplish this training of a bot we utilized a popular reinforcement learning library, OpenAI Gym [5]. A separate model was built for each question and each were trained by running for a series of 10,000 training runs (an attempt at answering the question). The maximum amount of time that it took to train a model was 30 seconds.

To provide an overview of the models we analyzed the training process that each model underwent. For this, we gathered the total amount of reward gained for each run, allowing us to see how well the model performed at each point. By analyzing the reward gained we can identify the moment where the model moves from being unintelligent to the point at which it is able to answer the question and receive a maximal reward the majority of the time. To better see this shift in learning, we applied a mean smoothing of size 100 to the reward data.

Figures 2a & 2b shows the training process for two separate models for questions 5 and 10. The maximal reward is negative for each of these cases because in the optimal scenario it would arrive at the answer in a low number of actions which would be summed to be a negative number. The sign (positive/negative) of the number does not matter as simply arriving at the maximum possible value is the key.

Inspecting Figure 2a we can see that for answering question 5 the bot became optimal at about 6,000 training runs because of the large

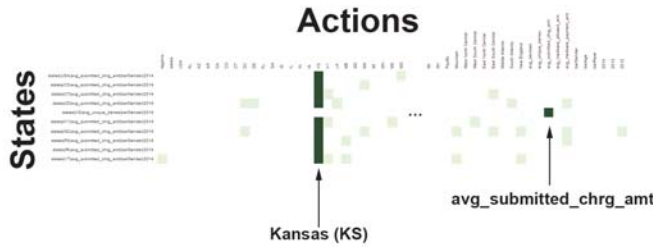


Fig. 3: A heatmap visualization of the trained model for question 6 where rows and columns that had mostly white for the entire span were removed. Inspecting the heatmap, it shows that the model learned that clicking on “Kansas” and the variable “Submitted Amount” were the most important.

vertical spike, while for question 10 in Figure 2b it became optimal in the final few training runs before 10,000. This indicates that the bot found question 10 to be more difficult compared to question 5. This difference in difficulty can reveal a lot about the interface and the way in which users interact with it. For example, a task that is constructed with the intention of being relatively simple for an interface, but is found to be difficult by the model would reveal that the interface may need to be refined to have more specific controls for accomplishing the associated task.

6.1 Learning Across Tasks

While we trained independent models for each question, we can also utilize a trained model as the starting point for a different question to show how knowledge is embedded and the applicability across tasks (referred to as transfer learning). Figure 2c shows this process and the performance on the newly trained task. By analyzing the improved model graph against Figure 2b, it can be seen that this model was able to answer question 10 in almost 2,000 less training runs. This indicates that even though the model was trained for a different question, it had learned and embedded information about the system which was then utilized to identify the path to answering question 10 in a quicker manner. This is directly applicable to a real-world scenario where a never before seen task has arisen and the optimal path needs to be discovered.

6.2 Extracting Optimal Path

Using the built models it is possible to extract the optimal path to complete a task. This can be achieved by starting at the state that corresponds to the beginning of the system, the “start state”, then choosing the action with the most reward, and continuing to move through the states until the goal state has been reached. The sequence of actions that were taken are then considered the optimal path for answering the question and what the agent would look to take.

Looking at question 3 (“In 2012, which gender submitted the largest payments in New Mexico (NM)?”), the extracted path from the model is *[states, NM, submitted amount, 2012]*. This shows that the model learned the correct path to arriving at the goal state because the actions align directly with the keywords of the question. Similar patterns were seen across the different questions.

It is also possible to see the optimal paths by inspecting the model’s matrix. Using question 6 (“Over all degrees and genders, which specialty had the highest submitted payments for 2014 in Kansas (KS)?”) as an example, Figure 3 contains a heatmap representation of the model learned. To construct the heatmap, rows and columns were removed if they contained all white, or nearly white, cells as this indicated almost no reward. This removal constrained the heatmap from thousands of rows down to ten.

This heatmap clearly shows that the model learned that the actions “Kansas” and “Submitted Amount” were the most important and provided the most reward to answer the question. Additionally, very light green cells can be seen which indicate that the model had explored

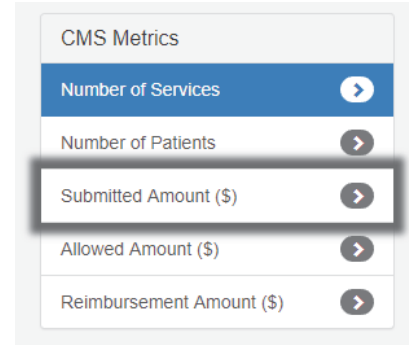


Fig. 4: An example of a suggestion being presented to a user. In this example the user has the “Number of Services” variable selected, but the “Submitted Amount” variable is highlighted to suggest for them to take the associated action.

other pathways that eventually led to the answer, but that it learned that these were not valuable to answering the question.

7 PROVIDING GUIDANCE

With the ultimate goal of providing guidance, we will outline how we provide the suggestions and then how these suggestions are personalized/tailored to the user automatically using reinforcement learning. Figure 4 shows an example of how a suggestion will be presented to a user. In this example the “Number of Services” variable is selected, but the “Submitted Amount” variable is highlighted to suggest for the user to switch to. We chose to keep the suggestion unobtrusive to ensure that we follow the expectation of being polite to the user [55]. Even with this subtle suggestion, our results will show that users did follow the suggestions.

7.1 What to suggest?

As shown in Section 6, we can build a model for each question such that the value of all possible actions in each state of the system can be identified. Using these models, it is possible to provide guidance for users by extracting the action that would provide the most value/lead the user towards answering the question correctly. Therefore, we utilize this knowledge in order to present suggestions to users with the goal of improving their accuracy as well as time and the number of actions taken.

7.2 When to suggest?

While we will be able to utilize our interaction models to provide guidance to users in the form of suggestions, this only solves the challenge of determining what to suggest and does not solve the problem of when to suggest. As it has been highly documented that providing assistance to users in a polite manner is important to not impede on already established workflows [24], it is crucial to identify when to provide suggestions to users to avoid being intrusive and widely disliked as Clippy in Microsoft Office famously was [55]. Recent methods for user assistance have relied on manually creating rules to determine when to assist a user, but these methods lack the ability to conform to user preferences and dynamically adapt to the interface that they are being presented with. Therefore, based on this criteria reinforcement learning algorithms can improve how we assist by exploring how the user is interacting, determine when to assist them, and over time adapt to their preferences.

7.3 Preference Model

Therefore, to personalize this guidance for each user we built a separate model called a “Preference Model” that also utilizes reinforcement learning, but instead focuses on determining the most opportune time to present the user with a suggestion. Some users prefer to receive suggestions frequently, others less so. The model can learn that certain users prefer to always get a suggestion after performing certain

interactions while others may not like a suggestion because they find it too distracting to their analysis process or they already know what they would like to do.

The preference model has a different state, action, and reward structure than the one defined in Section 4, but follows the similar principle of utilizing reinforcement learning and a matrix learned through Q-Learning. To keep this model simple and interpretable we constrained it to two possible actions: “suggest” or “not suggest” which indicate whether or not the user should be presented with a guidance suggestion. Furthermore, when devising the reward structure we sought one that would encourage the model to only present suggestions at the most opportune time and ensure that it remains polite to the user. Thus, we wanted to make sure that the model would choose not present a suggestion if it was unsure of the user’s preference rather than potentially interrupting and agitating the user. With this the reward was structured such that if a suggestion is presented and followed then the model receives a +5, but if the suggestion is not followed then it is penalized with a -5. And if the model chooses not to present a suggestion then it receives a -1. With the high values of +5 and -5 compared to the -1, the model is able to learn that suggestions are very valuable, but risky; and with the constant -1 it is encouraged to identify moments that the user would follow the suggestion and receive a +5.

As we sought for the preference model to be generalize across tasks we defined the states of the model to correspond to the history of actions that a user had taken. Using a history parameter, h , the states are built such that each state corresponds to the last h actions that were taken. For example, given a user creating a chart in Tableau a set of actions within the software could include: *Upload, SetChart, SetColor, FilterVariable* and with $h = 2$, such that the last two actions make up the states, then possible states for this model would include *SetColor – SetColor, SetChart – FilterVariable, SetColor – FilterVariable*, etc. Thus a user at the state *SetColor – SetColor* would have changed the chart type followed by changing the color scheme, resulting in the model learning over all times that the user had performed these two actions whether or not they prefer to have a suggestion displayed to them.

An extra action is also included in the preference model’s possible set of actions called “time”. This action represents a user not having taken an action for the last s seconds and allows the model the ability to evaluate users that may become lost or unsure about where to click next and act upon this.

With this model structure the parameters, h and s , can be configured based on the system it is contained within. For our system we initialized the parameters to $s = 5$, representing a “time” action to be triggered after 5 seconds of inactivity, and $h = 2$, representing a history look-up of two actions.

8 USER STUDY

To apply our models and system, we conducted an IRB approved (Y17TO37149) user study in which participants were presented with our system and asked to answer a series of questions related to the data. The participants were recruited through Amazon Mechanical Turk, presented with a tutorial of the system, and provided a monetary reward of \$1 upon completion of the study.

8.1 User Groups

We divided the participants into two groups. The first group consisted of 50 users that received no guidance, while the second group consisted of 50 users that received guidance, in the form of suggestions, for their last 5 questions. All users received questions in a random order.

We identified 100 users as being a sufficient amount for this user study because after running through the two groups of 50 users, we analyzed the data and were able to find insights immediately. Additionally, the 100 users mimics a real world scenario in which a large cohort of users may not be available. By showing an effective approach on a relatively small cohort we can show that our model is able to learn even with limited data and can extrapolate to what could be learned from a large population that has a vast amount of information and pathways.

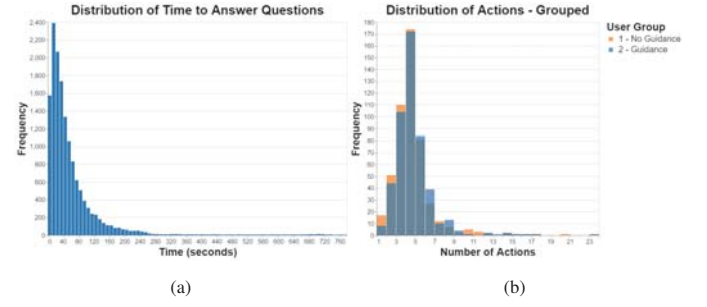


Fig. 5: (a) The distribution of the time spent answering each question by all users. (b) A layered distribution graph showing the distribution of actions for both user groups.

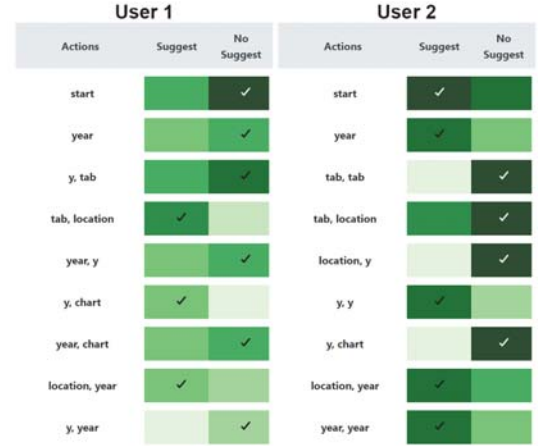


Fig. 6: A heatmap of the preference model built and trained on two different users. The rows correspond to the last two actions that a user took and the block with the darker shade of green (and checkmark) indicates whether or not the user should be provided guidance by means of presenting a suggestion.

8.2 Overall Statistics

On average users scored 78% over all ten questions and took 64.82 (± 62.53) seconds while performing 4.24 (± 2.13) actions per question. Users that correctly answered a question performed 4.32 (± 2.12) actions while users that incorrectly answered a question performed 3.68 (± 2.3) actions. Figure 5a shows the distribution of time to answer a question where users took about 30 seconds to answer a question. Figure 5b shows a layered distribution graph in which the number of actions for each user group is shown.

9 EVALUATION

To evaluate the effectiveness of our approach on providing personalized guidance to users we will perform two different methods of analysis: (i) we will inspect the Preference Models to identify whether individual user preferences were able to be captured and learned by the models and (ii) we will evaluate the statistical impact that the suggestions had on the users.

9.1 Model Inspection

Using the matrix learned by our Preference Model we can examine what was learned about the user. Figure 6 shows a heatmap of the model for two different users. In the first row of the heatmap, which is the initial load of the question before any actions are taken, the model learned a larger value for not providing suggestions to user 1 while the opposite occurred for user 2. This indicates that user 1 preferred not to receive suggestions at the beginning of the task while user 2 did.

Table 1: Pre vs Post Number of Questions Correctly Answered

	<i>Pre</i>	<i>Post</i>
Group 1	193	198
Group 2	180	209

Looking at the entirety of users as a whole, we found that the majority aligned similarly with user 1.

Furthermore, looking at subsequent rows it is easy to identify, based on the checkmarks and different shades of green, moments that the model learned to provide suggestions and moments where it learned to stay dormant. Places at which the shades of green differ significantly indicate that the model learned a definitive decision to make. On the other hand, when the shades of green are close in shade then this indicates that either the model requires more experience to better learn this state or that the user is unsure of which way they prefer the most.

9.2 Guidance Impact

To understand the statistical impact that guidance had on users, we first sought to make sure both groups (first 50 users vs last 50 users) were similar before making the determination of the effect that suggestions had. To accomplish this we ran a t-test on the amount of time spent on all ten questions as a whole and found that the two groups were not significantly different at an $\alpha = 0.05$ (t-statistics=0.903, p-value=0.249). We also ran a t-tests for time taken on each question and once again did not find a significant difference between the two groups (t-statistic=1.15, p-value=0.369). Next, we also ran χ^2 -test for the number of actions that it took a user to complete a question and due to the relatively low number of actions taken and thus low variances, we did not find any significant differences ($\chi^2=0.277$, p-value=0.599). Furthermore, we ran a final test on the accuracy of group 1 versus group 2 and did not also find any significance ($\chi^2=0.0058$, p-value=0.939). Using these tests, we can conclude that the two groups of users behaved similarly and that despite some outliers we can treat them as similar users.

With the two user groups identified as being equal, we analyzed the effect that suggestions had on users and identified two key discoveries when analyzing the first five questions that users saw versus the last five questions that they saw. Because group 2 received suggestions for their last five questions and group 1 did not, we will coin these two sets of questions as “pre” and “post” where “pre” corresponds to the first five questions and “post” corresponds to the last five.

For the first key discovery, we found that the groups did not differ significantly between each other when comparing their pre and post average completion times (t-statistic_{pre}=0.4918, t-statistic_{post}=1.0471). However, comparing all user’s pre against all user’s post did prove to be significant at an $\alpha = 0.05$ (t-statistic=2.31, p-value=0.0229). This is something that we thought could be expected due to the user learning the system, even though we attempted to mitigate it with a tutorial at the beginning of the session. However, through further analysis we found that this was not the case. Looking at only group 1’s pre versus post, we found that their time difference was not significant (t-statistic=1.06, p-value=0.295); and looking at only group 2’s pre versus post, we found that their time differences were significant (t-statistic=2.66, p-value=0.011). Therefore, we can rule out any learning effect in the post section and can conclude that group 2’s that the suggestions made group 2 fast enough that they were able to shift the entire sample population of users as a whole to be significantly different.

For the second key discovery, we found similar results as before: proportionally, users in group 2 had a higher accuracy rate in the post section than in the pre section ($\chi^2=9.08$, p-value=0.0026). Looking at group 1 with the same parameters shows that they proportionally did not exhibit the same behavior ($\chi^2=0.188$, p-value=0.665). This can be seen in Table 1 where a McNemar χ^2 -test of the number of questions answered correctly by both groups shows a significant difference ($\chi^2=69.25$, p-value<0.0001).

With these two discoveries, we can conclude that the guidance provided to users significantly improved their performance, in terms of both accuracy and time. We were not able to identify any significance

in the number of actions that they took due to the small variance within the data. Also, due to a small sample size with randomized question order we were not able to identify significance on a per question basis.

9.3 Suggestion Accuracy

Additionally, we sought to verify if the suggested action chosen by our interaction model was correct. To accomplish this we aggregated each time that a user was presented a suggestion and the associated reward value from the model. Then we split the rewards into the times that the user followed the suggestions versus the times that it did not and ran a t-test on these rewards. With this we found that the rewards were significant at an $\alpha = 0.05$ level such that the suggestions that users followed had a significantly higher reward value than those that users did not (t-statistic=-3.13, p-value=0.002). From this we can gather that when our interaction model was confident in its suggestion (had a high reward value) the user followed it.

We also assessed the effect the number of suggestions a user received to their accuracy and we found that as users received more suggestions their accuracy increased (t-statistic=3.88, p-value=0.0002), indicating that the presence of suggestions was of assistance to them.

10 OPTIMIZE INTERFACE

With **R1** and **R2** accomplished, we will explore how **R3** could similarly be accomplished. First, we will outline the different methods of optimization that could be implemented for a visual analytic system. Then, we will verify that potential avenues of optimization can be identified in our reinforcement learning models using a manual method of inspection. Because the application of these models to automatically optimizing an interface without manually inspection requires substantial research effort we seek for future work in this area to be performed to extend the introduction of reinforcement learning to user interactions that we have shown.

10.1 Methods of Optimization

Interface Layout Users utilize visual analytic interfaces for a wide range of tasks and follow different sets of analytic steps to accomplish the same task. Additionally, tasks across domains may require or highlight different controls. Thus, it is necessary to customize the layout of the tool and its controls for each user and the domain/task that they utilize it for.

Predictive Tasks: A user may commonly perform a series of tasks. For instance, a financial analyst may always apply a green color scheme upon activating a specific chart, while a healthcare analyst may look to always a blue color scheme for the same chart. Rather than requiring the user to make the extra click each time, an interface that can predict common tasks that a specific user performs could automatically change the color scheme to the one that the specific user prefers. Therefore, even though the color scheme cannot be applied as a general rule for all users or cohorts of users, it can be identified for the individual.

Learn Cohorts: While it is typical to personalize to a singular user, it is also possible to use knowledge learned and apply it towards certain cohorts of users. For example, analysts across different domains utilize visual analytic systems differently, focusing on different controls and performing different domain-specific tasks. Because of this, a separate model could be built for each cohort of users to identify the differences. Using these separate models, a new user that has never interacted with the system and is being introduced into a cohort could be provided with a model learned from the overall cohort and through the user’s interactions with the system the model would learn and tune itself to match the unique behaviors of this individual user.

Visualization: Users interpret and understand each type of visualization and interface differently [41]. Some may prefer one over another. For example, in some domains a specific visualization, such as a heatmap, may be widely used causing for experts within the domain to understand a heatmap over other visualizations. Therefore, being able to identify the visualizations that a user prefers or understands the most as well as the variations of the visualization can ensure that users are properly supported by a tool.

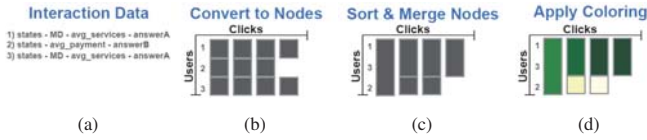


Fig. 7: An example of how we construct our interaction pathway visualization. (a) The interaction data is compiled, (b) a node is placed for each action, (c) the users/rows are sorted and similar actions are merged vertically, (d) a coloring scheme from yellow to green based on reward is applied.

10.2 Model Inspection for Optimization

With different methods of optimization defined, we build the same models as previously built using a bot, but this time we use the user interaction data from the user study. Then we inspect these models built from users to identify knowledge that was gained from the reinforcement learning algorithms. The insights into optimization are broken down into three methods: (1) analysis of interaction pathways, (2) user comparison, and (3) difficulty of the assigned task.

Of note: for each of these insights, due to the size of the visualizations caused by the large matrix learned by the Q-Learning model, we extract and choose to only display parts of the visualizations into the figures. These extractions are insights that we gained from an analysis of our models and their visualizations. We refer the reader to the Appendix for the full visualizations.

10.3 Path Analysis

Evaluating the interaction pathways that users took can reveal differences between users as well as provide information into the types of interactions that are taken within an interface. To understand these pathways, we built a matrix-like visualization where the rows correspond to the users and the columns correspond to the actions that users took, as shown in Figure 7. This allows for being able to follow the path of each user by following along horizontally. Furthermore, we sort the rows of users such that similar users are placed next to each other, allowing for interactions to be grouped together. The sorting is accomplished by adapting the radix sort [8] algorithm to this task. Through the grouping that results from sorting, the amount of users that took each type of interaction is evident by the size of the node.

Figure 8 contains a visualization of the final pathway visualization for the sequences of actions that users took to answer question 9. The color scheme was taken from the reward values learned from the models. In addition, red vertical bars were placed on the rows of users that had answered the question incorrectly.

Analyzing this visualization there are two dominant paths to highlight: the large section of users that correctly answered the question at the top, denoted as “Group A”, and the large section of users that answered the question incorrectly towards the bottom, denoted as “Group B”. These two sections immediately jump out because of the similarity that exists within a large block of users and the fact that the majority of Group A users answered the question correctly whereas the majority of Group B answered incorrectly.

As indicated in black text on top of each node, users in Group A took the path [“states”, “NV”, “2012”, “Submitted Amount”, “Answered Correctly”] whereas the majority of the users in Group B took the path [“2012”, “states”, “NV”, “Answered Incorrectly #1”]. The interesting realization between these two paths is that there exists an intersection in the actions, such that Group B had taken a subset of the moves made by Group A. However, Group B had not taken all of the actions that Group A did, presumably forgetting to switch to the “Submitted Amount” variable. Additionally, the model had identified, without any user data, that the path taken by Group A was more valuable than the path taken by Group B. This can be seen in the third action of the path by comparing the darker shade of green for “2012” against the light green for “NV” within the same column.

Based on this knowledge, it can be learned that it is advantageous for users to follow the path of clicking on the “states” button first rather than

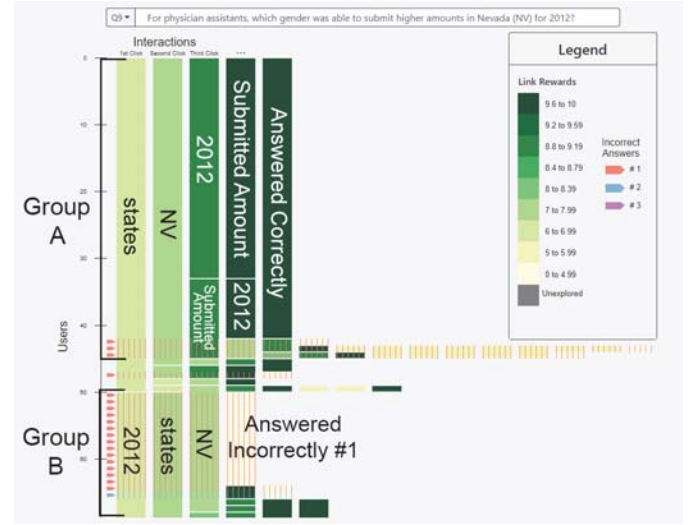


Fig. 8: A visualization of the interaction pathways for question 9, as built by Figure 7. Users in Group B took a subset of the actions taken by users in Group A, presumably forgetting to change to the “Submitted Amount” variable. Additionally, the trained model identified that the Group A actions were more valuable, as evidenced by the darker shades of green for the third column. With this, the data suggest to optimize the interface to ensure that users take the path that leads them towards answering correctly.

first clicking on the year “2012”. Therefore, the data suggests that the interface could be optimized to ensure that this aspect is presented first. A potential layout alteration that could be made would be to place each set of controls vertically such that the controls could be treated as steps, thus forcing users to follow this path for a specific question. While this layout alteration may be most applicable to Amazon Mechanical Turk users, it still shows that our model is able to identify these kinds of insights.

10.4 User Comparison

Another important aspect of understanding user interactions is to understand how users differ. Identifying where users took missteps and the reasons behind the different actions that they took is paramount in having a deep and rich understanding of where the interface could be optimized. Common techniques for performing user comparisons consist of using statistical methods in which the amount of time, number of actions, and correctness are taken into account. However, these techniques are limiting in that they are not able to identify the interactions at which users differed. For example, if two cohorts of users both took about the same amount of time and clicks, how did they differ? With this, we show our interaction model applied to this problem on two different comparisons.

10.4.1 Correct vs Incorrect

The first type of common analysis that is performed is to compare users that answered a question correctly against those that did not. For this, we utilize question 6 (“Over all degrees and genders, which specialty had the highest submitted payments for 2014 in Kansas (KS)”) and split the users into two groups: those that answered the question correctly (blue) and those that did not (red). A model is trained for each group and then the two models are subtracted to produce a heatmap included within the Appendix. We extracted from the heatmap a single state into Figure 9a and a single action over many states into Figure 9b. For both of these representations, a more blue cell indicates that the blue group took those actions more frequently and a more red cell indicates the same but for the red group. Additionally, the values of the cells correspond to the reward learned by the model.

In Figure 9b it can be learned that both sets of users had taken the

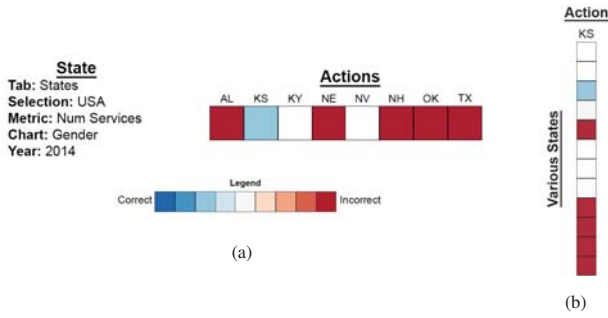


Fig. 9: A comparison of users that answered a question correctly versus those that did not. (a) For a single state it can be seen that the incorrect users struggled to identify the state Kansas. (b) For the action Kansas (KS), it can be seen that correct users took the action at a singular state while the incorrect users took the action from an assortment of states.

correct action, but that correct users seemed to have always been in the same state of the system when they had taken it compared to incorrect users that had taken it from a variety of system configurations. To understand why incorrect users had clicked on Kansas from different states, Figure 9a shows that for a single state incorrect users had clicked on a wide variety of US states while correct users always clicked on the correct one. Additionally, the light shade of blue within the figure indicates some of the incorrect users had also clicked on “KS” and that it wasn’t limited to only correct users, which would correspond to a dark shade of blue.

The knowledge extracted from this figure provides two helpful insights: (1) incorrect users struggled to identify the state of Kansas and thus the interface should be updated to assist these users that require assistance and (2) some incorrect users had taken the correct action of Kansas and thus further inspection into other states within the model is required to understand where those users had deviated. Both of these insights could be utilized to help optimize the interface, such as displaying the names of the states directly onto the map to make them easier to find for users that are lost.

10.4.2 How do correct users differ?

Another interesting comparison that is difficult and sometimes even impossible to perform is to understand how users that correctly answered a question differ. The reason that this task is typically very difficult is because even when a divergence of paths is able to be found, identifying the reason for the divergence requires immense research: did the user diverge because they made a mistake or did they diverge because they had extra knowledge? In the latter case, it is near impossible for standard techniques to make a concrete conclusion as two users may have taken a similar amount of actions, but one of the users may have identified a more intelligent manner to arrive at the answer. For our interaction model, this analysis is simple.

To accomplish this we built a model utilizing only users that correctly answered a question. We then normalized each row (state) of the model such that its reward values would be between 0 and 1, and plotted it as a heatmap. An extraction of a singular row/state within the heatmap for question 4 is shown in Figure 10.

Because this heatmap only contains correct users and is not a subtraction of two cohorts as we saw in the correct versus incorrect users, the analysis on this heatmap is different. To evaluate this heatmap, we must look for rows (states) where the variance between actions is high. These variances can reveal the extra knowledge that expert users may possess and thus would provide information into alternate use cases or benefits of a system that were not intended or identified upon its construction. While these types of insights would be more possible and easier to identify within a more complex system where alternate pathways have the potential to be present, we provide a simple example from our user study on a question where the insight was able to be found.



Fig. 10: A comparison of correct users for question 9 in which for a single state there was variance in the action reward between the users. While the “submitted amount” correctly led users to the answer, it was discovered that the other actions would present a visualization that would also provide users with the correct answer.

Inspecting Figure 10 and the different rewards for each action for the given state, we can see that correct users had clicked on three different variables: “Number of Patients”, “Submitted Amount”, and “Allowed Amount”. The dominant and most reward path consisted of “Submitted Amount” which is what the question had asked for. However, other correct users had also taken alternate paths and upon inspection of these variables we found that the user would still arrive at the same answer even if they had taken these wrong paths. This insight reveals to us that the dataset had some commonalities which for this question had allowed for users to take a multitude of paths.

11 DISCUSSION

With this work, we have laid the foundation for a more intelligent and in depth learning of user interactions. The model and framework that we have outlined utilizes reinforcement learning to automatically model and learn, thereby removing the previous limitation of experts constructing rules or users analyzing their own interactions via historical timelines. This newly applied technique on interaction data allows for analyses and applications that would not be possible with standard techniques, thus opening a new window into future work to build upon this work.

12 LIMITATIONS

Due to the new introduction of reinforcement learning to visualization, we acknowledge that our approach contains limitations but we recognize this as being a necessary first step. First, the task provided to users was simple in nature compared to many visualization tasks that users perform on a daily basis. This constraint in the difficulty was necessary in order to evaluate the effectiveness of our modelling without needing to account for the noise and challenges associated with complex tasks. However, because the theory and concept of reinforcement learning have been proven to be successful for this task, scaling to more complex tasks can easily be accomplished.

Second, while one of the exciting applications of such automatic models is towards exploration tasks where a user enters a visualization system without a predefined question or task, our user study did not tackle such tasks. We believe that this type of research is interesting and should be a long-term goal, but there exists work that is still required to be completed before we can reach this goal and accomplish such tasks.

REFERENCES

- [1] J. Allen, C. I. Guinn, and E. Horvitz. Mixed-initiative interaction. *IEEE Intelligent Systems and their Applications*, 14(5):14–23, 1999.
- [2] G. Andrienko, N. Andrienko, P. Bak, D. Keim, S. Kisilevich, and S. Wrobel. A conceptual framework and taxonomy of techniques for analyzing movement. *Journal of Visual Languages & Computing*, 22(3):213–232, 2011.
- [3] L. Bavoil, S. Callahan, P. Crossno, J. Freire, C. Scheidegger, C. Silva, and H. Vo. VisTrails: enabling interactive multiple-view visualizations. In *IEEE Visualization, 2005. VIS 05*, pp. 135–142, Oct. 2005. doi: 10.1109/VISUAL.2005.1532788

- [4] M. Bostock, V. Ogievetsky, and J. Heer. D³ data-driven documents. *IEEE transactions on visualization and computer graphics*, 17(12):2301–2309, 2011.
- [5] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba. Openai gym. *arXiv preprint arXiv:1606.01540*, 2016.
- [6] E. Brown, A. Ottley, H. Zhao, Q. Lin, R. Souvenir, A. Endert, and R. Chang. Finding waldo: Learning about users from their interactions. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1663–1672, Dec. 2014. doi: 10.1109/TVCG.2014.2346575
- [7] M. Campbell. Knowledge discovery in deep blue. *Communications of the ACM*, 42(11):65–67, 1999.
- [8] T. H. Cormen. *Introduction to algorithms*. MIT press, 2009.
- [9] F. Dabek and J. J. Caban. A grammar-based approach for modeling user interactions and generating suggestions during the data exploration process. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):41–50, 2017.
- [10] N. De Firmian. *Modern Chess Openings: MCO-15*. Random House Incorporated, 2008.
- [11] W. Dou, D. H. Jeong, F. Stukes, W. Ribarsky, H. Lipford, and R. Chang. Recovering Reasoning Processes from User Interactions. *IEEE Computer Graphics and Applications*, 29(3):52–61, May 2009. doi: 10.1109/MCG.2009.49
- [12] C. Ehmke and S. Wilson. Identifying web usability problems from eye-tracking data. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI... but not as we know it-Volume 1*, pp. 119–128. British Computer Society, 2007.
- [13] M. Eirinaki and M. Vazirgiannis. Web Mining for Web Personalization. *ACM Trans. Internet Technol.*, 3(1):1–27, Feb. 2003. doi: 10.1145/643477.643478
- [14] M. Feng, C. Deng, E. M. Peck, and L. Harrison. Hindsight: Encouraging exploration through direct encoding of personal interaction history. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):351–360, 2017.
- [15] C. Fernando, D. Banarse, C. Blundell, Y. Zwols, D. Ha, A. A. Rusu, A. Pritzel, and D. Wierstra. Pathnet: Evolution channels gradient descent in super neural networks. *arXiv preprint arXiv:1701.08734*, 2017.
- [16] E. Y. Fu, T. C. Kwok, E. Y. Wu, H. V. Leong, G. Ngai, and S. C. Chan. Your mouse reveals your next activity: Towards predicting user intention from mouse interaction. In *Computer Software and Applications Conference (COMPSAC), 2017 IEEE 41st Annual*, vol. 1, pp. 869–874. IEEE, 2017.
- [17] D. Gotz and Z. Wen. Behavior-driven visualization recommendation. In *Proceedings of the 14th International Conference on Intelligent User Interfaces, IUI '09*, pp. 315–324. ACM, New York, NY, USA, 2009. doi: 10.1145/1502650.1502695
- [18] D. Gotz and M. Zhou. Characterizing users’ visual analytic activity for insight provenance. In *IEEE Symposium on Visual Analytics Science and Technology, VAST '08*, pp. 123–130, Oct. 2008. doi: 10.1109/VAST.2008.4677365
- [19] T. Green, R. Maciejewski, and S. DiPaola. ALIDA: Using machine learning for intent discernment in visual analytics interfaces. In *2010 IEEE Symposium on Visual Analytics Science and Technology (VAST)*, pp. 223–224, Oct. 2010. doi: 10.1109/VAST.2010.5650854
- [20] I. Grudin. History and focus. *interaction*, 1994.
- [21] H. Guo, S. R. Gomez, C. Ziemkiewicz, and D. H. Laidlaw. A case study using visualization interaction logs and insight metrics to understand how analysts arrive at insights. *IEEE transactions on visualization and computer graphics*, 22(1):51–60, 2016.
- [22] J. Heer, J. Mackinlay, C. Stolte, and M. Agrawala. Graphical histories for visualization: Supporting analysis, communication, and evaluation. *IEEE transactions on visualization and computer graphics*, 14(6), 2008.
- [23] E. Hoque, V. Setlur, M. Tory, and I. Dykeman. Applying pragmatics principles for interaction with visual analytics. *IEEE transactions on visualization and computer graphics*, 24(1):309–318, 2018.
- [24] E. Horvitz, A. Jacobs, and D. Hovel. Attention-sensitive Alerting. In *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence, UAI'99*, pp. 305–313. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999.
- [25] J. Hurtienne and L. Blessing. Design for intuitive use-testing image schema theory for user interface design. In *16th International Conference on Engineering Design*. Citeseer, 2007.
- [26] R. Jacob and K. S. Karn. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind*, 2(3):4, 2003.
- [27] R. J. Jacob. Eye tracking in advanced interface design. *Virtual environments and advanced interface design*, pp. 258–288, 1995.
- [28] N. Kadivar, V. Chen, D. Dunsmuir, E. Lee, C. Qian, J. Dill, C. Shaw, and R. Woodbury. Capturing and supporting the analysis process. In *IEEE Symposium on Visual Analytics Science and Technology, 2009. VAST 2009*, pp. 131–138, Oct. 2009. doi: 10.1109/VAST.2009.5333020
- [29] S. Kandel, A. Paepcke, J. Hellerstein, and J. Heer. Enterprise Data Analysis and Visualization: An Interview Study. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2917–2926, Dec. 2012. doi: 10.1109/TVCG.2012.219
- [30] S.-H. Kim, Z. Dong, H. Xian, B. Upatising, and J. S. Yi. Does an eye tracker tell the truth about visualizations?: findings while investigating visualizations for decision making. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2421–2430, 2012.
- [31] J. Kirkpatrick, R. Pascanu, N. Rabinowitz, J. Veness, G. Desjardins, A. A. Rusu, K. Milan, J. Quan, T. Ramalho, A. Grabska-Barwinska, et al. Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13):3521–3526, 2017.
- [32] T. C. Kwok, E. Y. Fu, E. Y. Wu, M. X. Huang, G. Ngai, and H.-V. Leong. Every little movement has a meaning of its own: Using past mouse movements to predict the next interaction. In *23rd International Conference on Intelligent User Interfaces*, pp. 397–401. ACM, 2018.
- [33] Y. Li. Deep reinforcement learning: An overview. *arXiv preprint arXiv:1701.07274*, 2017.
- [34] Z. Liu and J. Heer. The effects of interactive latency on exploratory visual analysis. *IEEE transactions on visualization and computer graphics*, 20(12):2122–2131, 2014.
- [35] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*, 2013.
- [36] A. J. Munro, K. Höök, and D. Benyon. *Social navigation of information space*. Springer Science & Business Media, 2012.
- [37] G. M. Olson, J. D. Herbsleb, and H. H. Rueter. Characterizing the sequential structure of interactive behaviors through statistical and grammatical techniques. *Human-Computer Interaction*, 9(4):427–472, 1994.
- [38] A. Perer and B. Shneiderman. Systematic yet flexible discovery: guiding domain experts through exploratory data analysis. In *Proceedings of the 13th international conference on Intelligent user interfaces*, pp. 109–118. ACM, 2008.
- [39] P. Pirolli and S. Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, vol. 5, pp. 2–4, 2005.
- [40] M. Pohl, M. Schmitt, and S. Diehl. Comparing the readability of graph layouts using eyetracking and task-oriented analysis. In *Computational Aesthetics*, pp. 49–56, 2009.
- [41] R. O. Prates, C. S. de Souza, and S. D. Barbosa. Methods and tools: a method for evaluating the communicability of user interfaces. *interactions*, 7(1):31–38, 2000.
- [42] E. Ragan, A. Endert, J. Sanyal, and J. Chen. Characterizing Provenance in Visualization and Data Analysis: An Organizational Framework of Provenance Types and Purposes. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):31–40, Jan. 2016. doi: 10.1109/TVCG.2015.2467551
- [43] A. Raghu, M. Komorowski, L. A. Celi, P. Szolovits, and M. Ghassemi. Continuous state-space models for optimal sepsis treatment—a deep reinforcement learning approach. *arXiv preprint arXiv:1705.08422*, 2017.
- [44] D. Roldán-Álvarez, E. Martín, M. García-Herranz, and P. A. Haya. Mind the gap: impact on learnability of user interface design of authoring tools for teachers. *International Journal of Human-Computer Studies*, 94:18–34, 2016.
- [45] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.
- [46] A. Sonderegger and J. Sauer. The influence of design aesthetics in usability testing: Effects on user performance and perceived usability. *Applied ergonomics*, 41(3):403–410, 2010.
- [47] C. Speier and M. G. Morris. The influence of query interface design on decision-making performance. *MIS quarterly*, pp. 397–423, 2003.
- [48] M. Streit, H.-J. Schulz, A. Lex, D. Schmalstieg, and H. Schumann. Model-driven design for the visual analysis of heterogeneous data. *IEEE Transactions on Visualization and Computer Graphics*, 18(6):998–1010, 2012.
- [49] R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*,

- vol. 1. MIT press Cambridge, 1998.
- [50] Tableau. <http://www.tableau.com/>.
- [51] M. E. Taylor and P. Stone. Transfer learning for reinforcement learning domains: A survey. *Journal of Machine Learning Research*, 10(Jul):1633–1685, 2009.
- [52] J. Togelius, S. Karakovskiy, J. Koutník, and J. Schmidhuber. Super mario evolution. In *Computational Intelligence and Games, 2009. CIG 2009. IEEE Symposium on*, pp. 156–161. IEEE, 2009.
- [53] W. Van der Aalst, T. Weijters, and L. Maruster. Workflow mining: Discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering*, 16(9):1128–1142, 2004.
- [54] W. M. Van der Aalst, B. F. van Dongen, J. Herbst, L. Maruster, G. Schimm, and A. J. Weijters. Workflow mining: a survey of issues and approaches. *Data & knowledge engineering*, 47(2):237–267, 2003.
- [55] B. Whitworth. Polite computing. *Behaviour & Information Technology*, 24(5):353–363, Sept. 2005. doi: 10.1080/01449290512331333700
- [56] W. Willett, J. Heer, and M. Agrawala. Scented widgets: Improving navigation cues with embedded visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1129–1136, 2007.
- [57] L. Xiao, J. Gerth, and P. Hanrahan. Enhancing Visual Analysis of Network Traffic Using a Knowledge Representation. In *Visual Analytics Science And Technology, 2006 IEEE Symposium On*, pp. 107–114, Oct. 2006. doi: 10.1109/VAST.2006.261436

A APPENDIX

Below are the full visualizations that were utilized to create the figures within the paper.

A.1 Dashboard

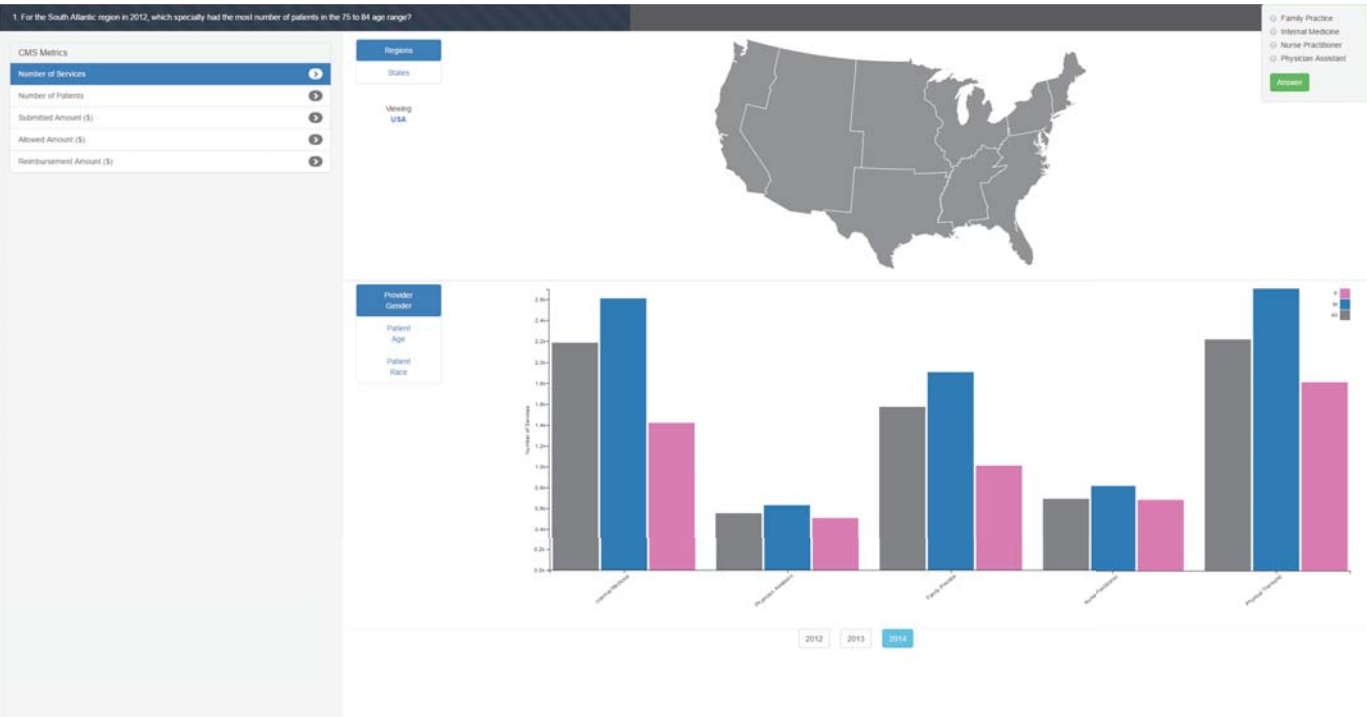


Fig. 11: A large screenshot of our visualization system built for capturing user interactions in a real-world type scenario.

A.2 Path Analysis

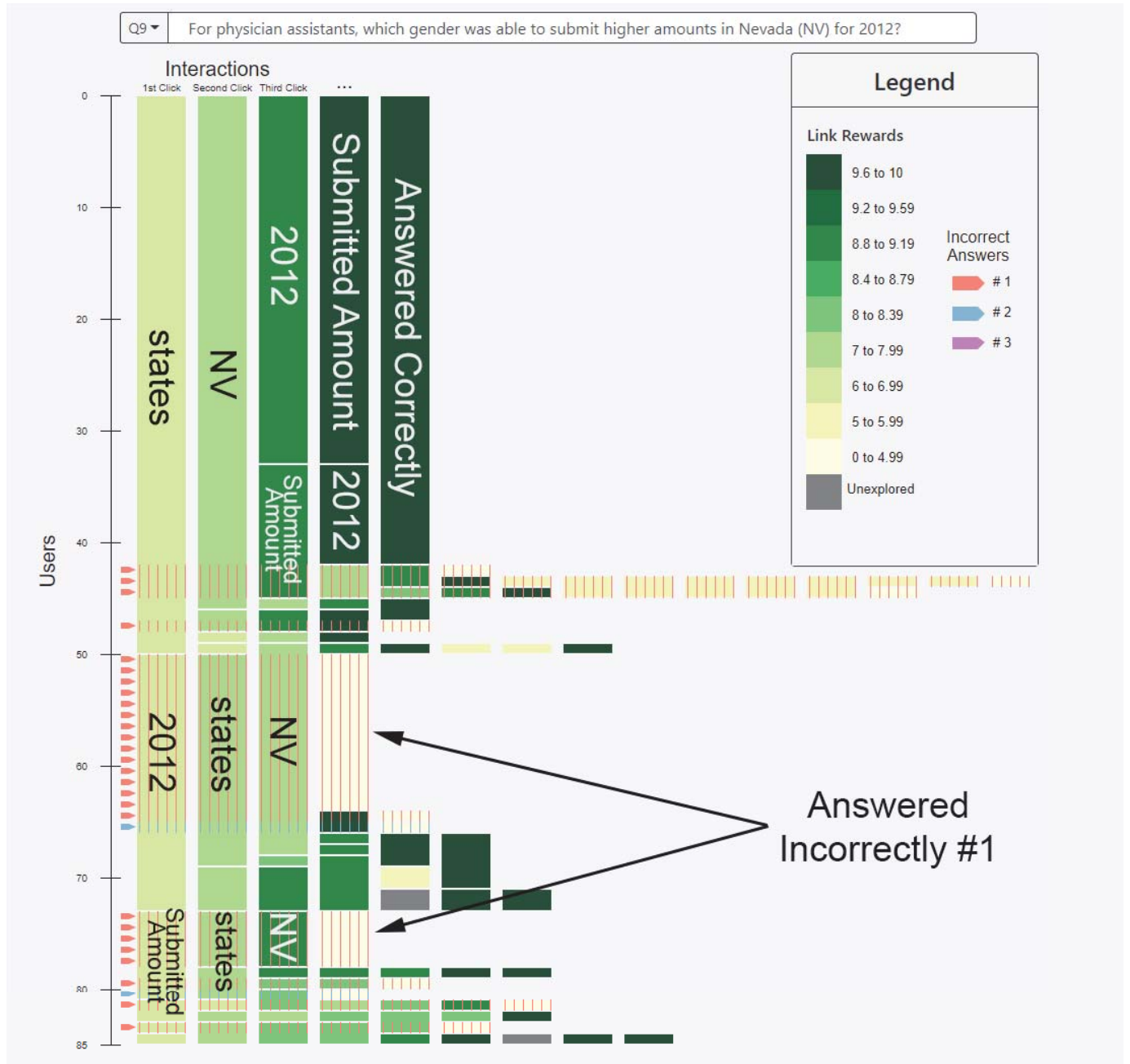


Fig. 12: A large screenshot of our visualization system built for capturing user interactions in a real-world type scenario.

A.3 Correct vs Incorrect

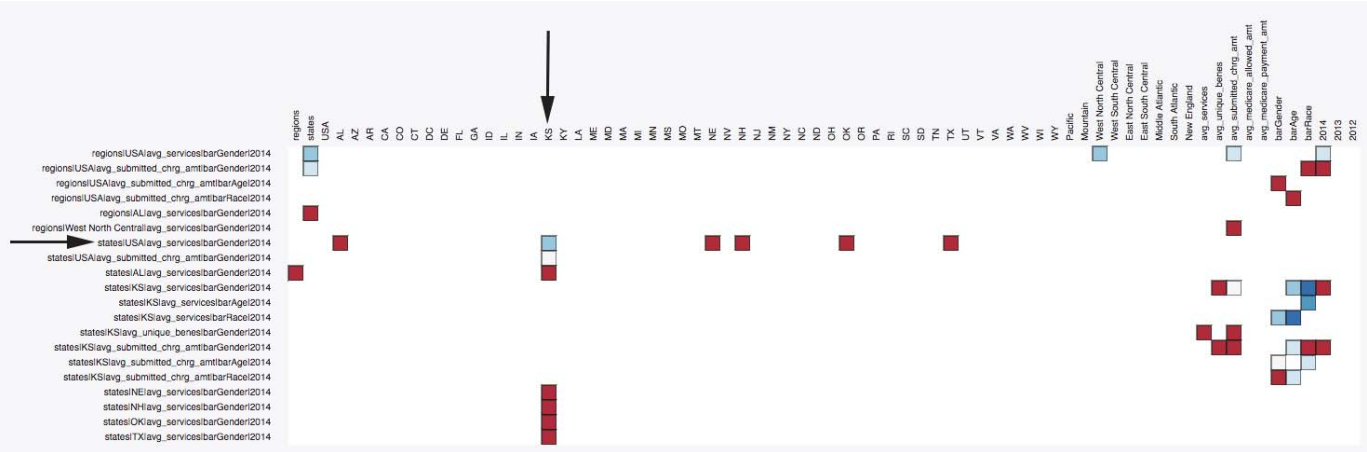


Fig. 13: A large screenshot of our visualization system built for capturing user interactions in a real-world type scenario.

A.4 How do correct users compare?

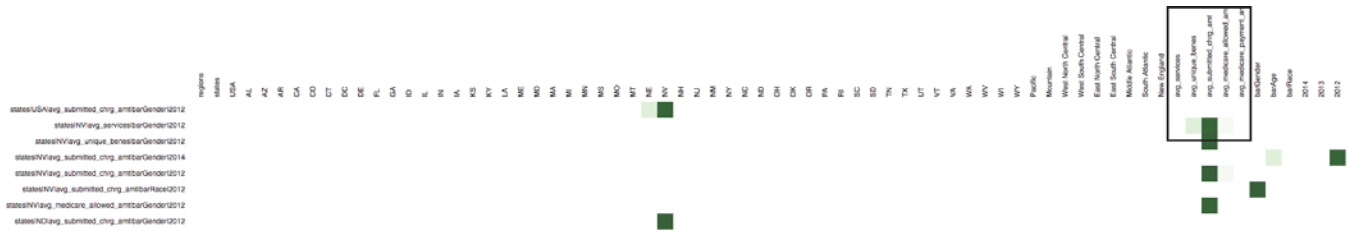


Fig. 14: A large screenshot of our visualization system built for capturing user interactions in a real-world type scenario.

A.5 Preference Model

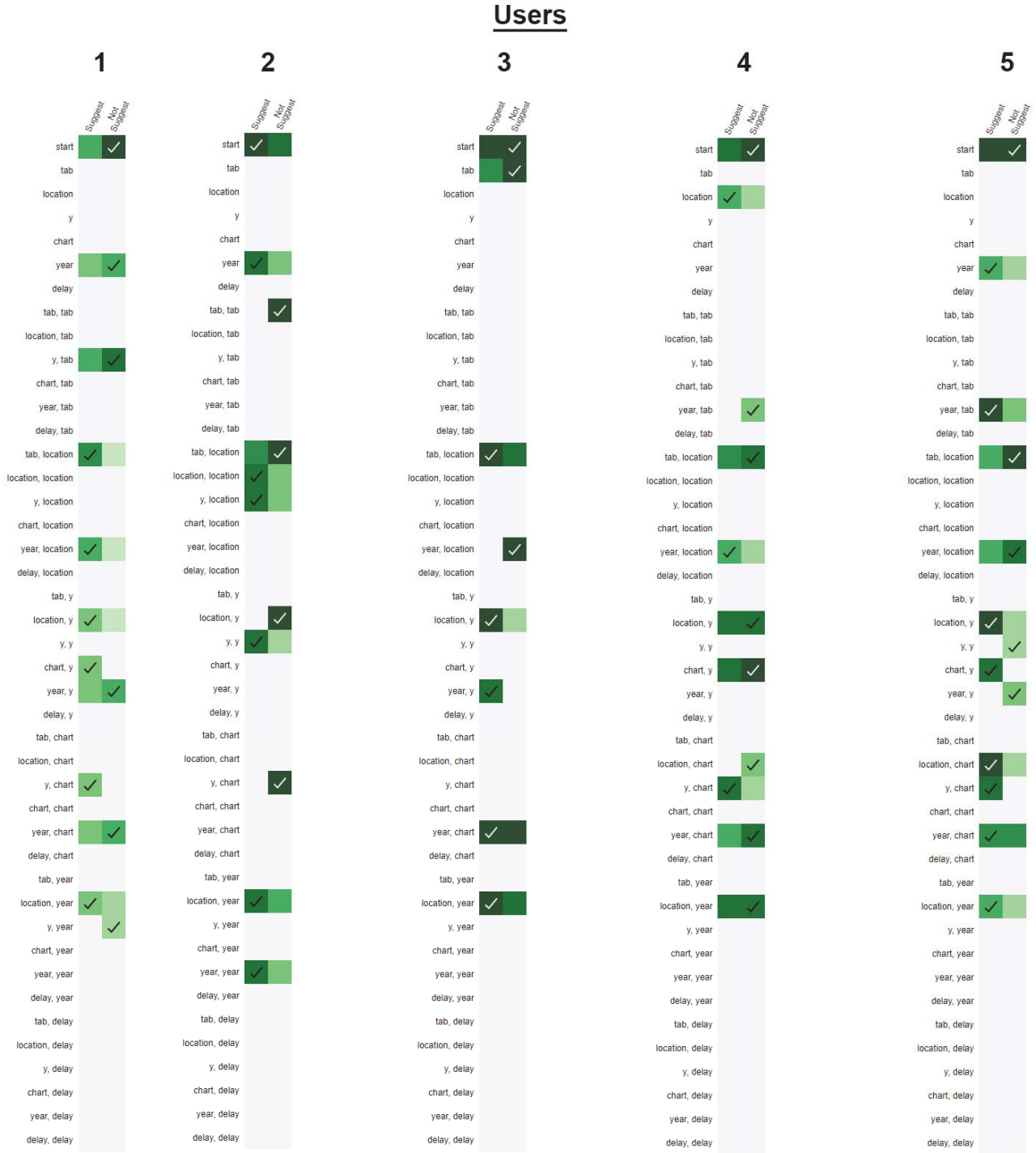


Fig. 15: Heatmaps of the preference model built and trained on different users. The rows correspond to the last two actions that a user took and the block with the darker shade of green (and checkmark) indicates whether or not the user should be provided guidance by means of presenting a suggestion.

TrajectoryFlow: Visual Summarization of Temporal Sequences

Filip Dabek*

National Intrepid Center of Excellence
Walter Reed National Military Medical
Center, Bethesda, MD

Jian Chen†

The Ohio State University

Jesus J Caban‡

National Intrepid Center of Excellence
Walter Reed National Military Medical
Center, Bethesda, MD

ABSTRACT

Summarizing a collection of temporal sequences is a difficult task given the irregular and variable patterns often found in longitudinal events. Across a wide array of domains, researchers and analysts seek to determine ways to gain an overview of a dataset through the identification of the common temporal paths as well as the trajectories that exist between individual events. While these tasks can be difficult on small and structured datasets, they are increased tenfold on temporal sequences that are noisy, irregular, and voluminous in size. This poster presents an approach that has been designed to visually explore large collections of temporal sequences by combining event mining algorithms with visualization techniques to overcome the challenges and complexities of analyzing longitudinal data.

1 INTRODUCTION

Events exist everywhere in life: online consumers perform a series of clicks before purchasing an item, patients go to different doctors and therapies during recovery, sports teams make plays and score back-and-forth during a match. Throughout all domains, a temporal event can be described as a *data point denoted by a timestamp and event type*. However, while events can exist in every avenue of life and in a vast array of domains, understanding how events are related to one another and summarizing a large sequence of events remains a difficult task. During the last few years, the increased availability of large-scale temporal datasets that contain complex and long sequences has underscored the importance for designing scalable techniques to easily identify trends and to answer important questions from irregular, noisy, and large longitudinal datasets.

In this poster we present *TrajectoryFlow*, a scalable algorithm and visualization, that has been built to effectively analyze and summarize longitudinal data. The primary contributions of this work include: (1) An adaptive window-based algorithm for extracting common subsequences of a longitudinal dataset. The mining algorithm takes into account the temporal context of the sequences to build an overview of the common paths and trajectories followed by a collection of records, and (2) a novel event summary diagram that reduces the analytical complexity used to visualize the results of event mining approaches. We provide an interactive, online tool for demonstration of our approach at: <http://tf.visxplorer.com>

2 BACKGROUND

In the visualization field, a common method of understanding event sequences has been to visualize a single record to understand its path over time [5] and to stack sequences vertically on top of one another attempting to merge identical paths [2,4]. Many of these approaches, while powerful and novel in their applications, choose to display the entirety of a dataset which causes for them to suffer performance issues in terms of scalability, high visual complexity caused by the large number of elements present in the visualization, and thus through these two issues the visualization becomes complex and

discouraging for the user to analyze, especially if the aforementioned dataset is very noisy.

3 AWP-SPAN

Frequent Sequence Mining (FSM) is a popular technique for finding the most frequently occurring subsequences of events in datasets of sequences [3], which produces a list of frequent subsequences and the number of sequences that each of them appear in. While FSM algorithms can identify the most frequent subsequences in a dataset, they do not possess the ability to take into account temporal context and build a trajectory between two events.

Our algorithm, AWP-Span, first splits a dataset of sequences into windows as denoted by W . Then an FSM algorithm, PrefixSpan, is utilized to identify the frequent patterns in each window that meet the frequent threshold of τ . Figure 1 provides an example of where two windows are utilized such that $W = W_1, W_2$ and where subsequences that are not “frequently occurring” ($\tau = 3$) are crossed out with a By splitting a dataset into windows, like we have shown, we are able to incorporate a temporal context into PrefixSpan allowing for it to identify the common patterns at different stages of a sequence rather than ignoring the premise that sequences change over time and patterns in different temporal locations may not be indicative of the same pattern. Being able to evaluate the various stages of a sequence is an important task that event researchers seek to accomplish and this modification builds this ability directly into the algorithm.

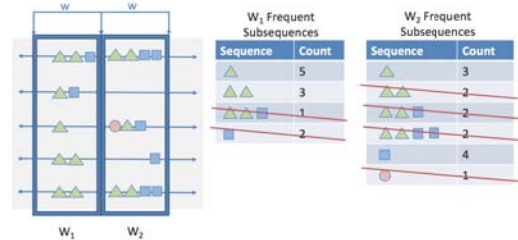


Figure 1: An example of our AWP-Span algorithm

4 SANKEY DIAGRAM

Due to the atypical nature of our algorithm results, we devised a mapping from the results obtained to the Sankey diagram [6], where each window's nodes were constructed from the frequent subsequences contained in the window and edges were drawn between nodes in different indices. To link adjoining windows we recognized that a single event sequence may contain one of two types of information in between its two windows: either noise or no information at all. Therefore we utilize both direct connections and wildcard nodes. A direct connection was drawn when no events existed in between the two frequent subsequences of a window, while a wildcard node was inserted when there was extra, infrequent information contained.

While the wildcard node is powerful at representing extra information that is contained between two events, for some visualization techniques this creates extra nodes and thus noise. To reduce this noise and clutter caused we altered the mapping so that instead of creating a wildcard node, a colored edge between events is drawn.

5 EVENT SUMMARY DIAGRAM

As the Sankey diagram suffers from visual complexity, ease of use, and difficulties in identifying the flow of events, we developed an edge-free visualization that reduces the complexity of analyzing

*e-mail: filip.j.dabek.ctr@mail.mil

†e-mail: chen.8028@osu.edu

‡e-mail: jesus.j.caban.civ@mail.mil



Figure 2: An example output of our AWP-Span algorithm. (Left) Shows a Sankey diagram containing wildcard nodes causing for extra clutter and noise to exist. (Middle) Shows a Sankey diagram with the wildcard nodes converted into red-colored edges. (Right) Our event summary diagram where the edges have been removed to decrease the analytical capacity needed to understand the overall flow of events. Wildcard edges are now mapped to a texture rather than a red-colored edge and windows are separated by a vertical gap.

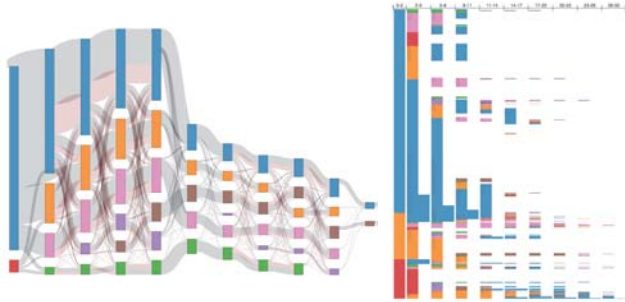


Figure 3: A web traffic dataset containing 989,818 users was summarized using our AWP-Span to produce a (Left) Sankey diagram and (Right) event summary diagram.

and understanding the different flow of trajectories. For our event summary diagram we first enforce the requirement that all event nodes must be aligned next to the node that they originated from. Next we removed the edges between nodes and encoded the wildcard edges using a texture. Finally we added a vertical spacing between each set of windows so that it is possible to understand the frequent subsequences that occurred within each window.

We utilize the example data shown in Figure 2 (Left and Middle) to present our new design in Figure 2 Right. Through visual inspection, we can see that our new design contains the same information present in the Sankey diagram, but does so in a simpler and more easily understood manner, thereby decreasing the analytical capacity required by the user and significantly reducing the virtual screen space utilized.

6 CASE STUDY - WEB TRAFFIC LOGS

We performed the first case study on web traffic logs based on an msnbc.com dataset containing the server logs for the entire day of September 28, 1999 [1], which contains 989,818 tracked users that make an average of 5.7 visits across 17 different categories of pages on the website. We presented a web traffic analyst with Figure 3. The analyst immediately was offput by the Sankey diagram, and stated that the only insights that they could safely discover in this “mess” were that users typically visit the blue page first, corresponding to the frontpage of the website.

With our event summary diagram, the analyst quickly picked out interesting patterns and trends. They identified that most users stay on the frontpage between clicks, because of the large series of blue in the middle of the diagram, but that some of them move to the orange (news), red (as seen tv), and pink (sports) before coming back to the frontpage, as seen in the upper left section. Next they eagerly pointed to the bottom section of the diagram where the initial red events lead into repeated orange events that continue through all of the windows, which indicates that users that visit the website through the “as seen on tv” (red) section will stay on the website for a long

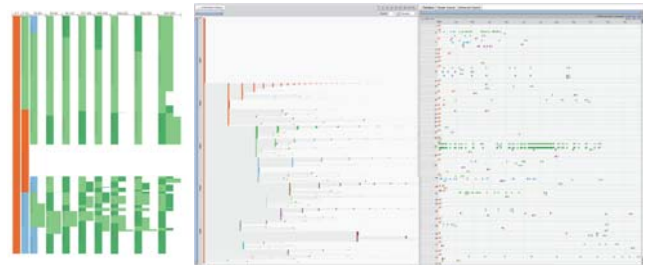


Figure 4: (Left) An event summary diagram for EHR data. (Right) The same EHR dataset loaded into an existing event analysis system.

time, clicking through different news (orange) pages. They noted this as being a helpful insight into user retention and pointed out a comparison to the middle section of the diagram where users that entered the website through the frontpage leave after a significantly lower amount of clicks than those that enter through the “as seen on tv” section. Based on this insight, they stated that the frontpage could potentially be optimized to attempt to guide users towards these pages with higher retention rates or that more advertising of the website could be done on TV.

7 CASE STUDY - EHR DATA

We performed a second case study with a large, complex, clinical dataset consisting of the first year after a mild Traumatic Brain Injury (a.k.a concussion) for 89,840 patients. The longitudinal data included 5.3 million TBI-related clinical encounters over 10 years and over 8.7 million clinical diagnoses (i.e. events). A mined event summary is shown in Figure 4 Left and a subset of the dataset (due to performance limitations) loaded into an existing event analysis system is shown in Figure 4 Right.

The healthcare analyst looked at both visualizations displayed in Figure 4 and first noted three similarities: (1) they both show that some patients suffered a concussion (orange) and then had no more events available afterwards and (2) they both indicate that patients have repeated diagnoses. Beyond these similarities, the analyst liked that EventFlow showed the raw sequences on the right, but that they struggled to identify any patterns besides the repeated events. Thus, they moved to the event summary diagram and expressed their satisfaction for time periods being present at the top of the diagram to aid in the understanding of the trajectory. With these time periods, they found it interesting to see that the size of the windows increased in the later windows, indicating that patients required less healthcare as time went on. Additionally, the analyst noticed that the light blue (headaches) diagnoses only occurred in the first couple of windows after a concussion which they confirmed follows the medical literature.

REFERENCES

- [1] I. V. Cadez, D. Heckerman, C. Meek, P. Smyth, and S. White. Visualization of navigation patterns on a web site using model-based clustering. In *KDD*, p. 280, 2000.
- [2] D. Gotz and H. Stavropoulos. DecisionFlow: Visual Analytics for High-Dimensional Temporal Event Sequence Data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1783–1792, Dec. 2014. doi: 10.1109/TVCG.2014.2346682
- [3] F. Masegla, M. Teisseire, and P. Poncelet. Sequential pattern mining. *Encyclopedia of Data Warehousing and Mining*, pp. 1028–1032, 2005.
- [4] M. Monroe, R. Lan, H. Lee, C. Plaisant, and B. Shneiderman. Temporal Event Sequence Simplification. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2227–2236, Dec. 2013. doi: 10.1109/TVCG.2013.200
- [5] 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 *Proceedings of the AMIA Symposium*, p. 76. American Medical Informatics Association, 1998.
- [6] P. Riehmann, M. Hanfler, and B. Froehlich. Interactive Sankey diagrams. In *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005*, pp. 233–240, Oct. 2005. doi: 10.1109/INFVIS.2005.1532152

Necklace: A Guided Interactive Multilayer Network Comparison Interface for Medical Claim Analysis

Gaoqi He^{‡,∇}, Xiaohui Bian^{†,∇}, Alexander Garbarino[†], Hadi Kharrazi[‡], Jesus Caban[△], Ping Zhang^{*}, and Jian Chen^{*}

[‡]East China Normal University, [†]University of Maryland Baltimore County, [∇]East China University of Science and Technology,

[‡]Johns Hopkins University, [△]NICoE, Walter Reed National Military Medical Center, ^{*}The Ohio State University

ABSTRACT

We present Necklace, a multivariate comparative data exploration interface for medical claims: billing information submitted by hospitals for payment by commercial and government health plans. Necklace uses a series of interconnected rings to represent event co-occurrences: nodes on rings represent events, and novel glyph small multiples to reveal relationships suitable for multilayer network visualizations. Our tool has two novel methods: a guided interaction for visual exploration of complex relationships and a spatial-saving *glyph small multiples* (GSM) to compare features on multilayer networks.

1 INTRODUCTION

Assessing healthcare costs in electronic health records (EHRs) is exceptionally challenging, especially for tasks related to analyzing medical claims. Some exemplar public-health researchers' claim-analysis tasks are *What attributes contribute the most to the increase in health-care cost (overview)? How does the socioeconomic-ethnic status influence claims over-time among populations and associated diagnoses (complex comparison)?* Addressing these tasks demands exploratory analysis of multivariate heterogeneous attributes among factors such as individual demographic characteristics (e.g., gender, race) and attributes (e.g., insurance type), as well as the disease and prescription history. Implementing these queries using database languages is not easy either because programming the complex queries is not well-supported or because the process of forming complex queries diminishes the fluidity of analytical workflow [1].

Current creative EHR visualization solutions to focusing on event trajectories for medical records (e.g., Perer et al.'s MatrixFlow [2], and Gotz et al.'s DecisionFlow [3]), though effective in showing diagnosis and disease progression, have limited use for claims for several reasons. First, trajectories in claims are less meaningful unless associated relationships are explained. These trajectories summarize a population by abstracting so much details away that causes and effects are obscured. Second, temporal ordering is least relevant, perhaps because claims are filed after the diagnostic events; Instead, we find co-occurrences that form multilayer networks of relationships.

For these reasons, we have designed and implemented Necklace (Figure 1), an interface for multivariate comparisons of complex multivariate relationships to reveal data distributions to guide researchers' explorations, that contributes to: (1) A guided interaction using forward-backward event cooccurrence to help experts explore large multivariate temporal health data; and (2) a multivariate comparative glyph design called *glyph small multiples* that saves screen space while showing patterns among cooccurred attributes.

2 NECKLACE INTERFACE

Necklace is a web-browser based event co-occurrence multivariate visual comparison tool on an unlimited panning and zoomable canvas, with a series of rings (like pearls in a necklace): a node on a ring represents a diagnosis and directional curved links between nodes

show events co-occurring before (in yellow) or after (in blue) the current diagnosis. Associated claims for multivariate comparison are displayed using novel glyphs design called *glyph small multiples* (GSM) for comparative analyses (Figure 2).

2.1 Needs Analysis

Data Characterization. We use a subset of curated medical claim database from the Centers for Medicare & Medicaid Services (CMS) of 691 patients and associated 12649 associated claims for post-traumatic stress disorder (PTSD) filed between January 1st 2008 and December 31 2010. PTSD is a common mental disorder that affects as much as 20% of veterans, 8% of the US population, and up to 86% among refugees with gross cost occurring years after the condition is controlled. The claim database contains patient ID, age, sex, race, state, claim history such as prescription, doctors/providers, medicine refills, and payments from insurance and related diagnosis codes in standard ICD-9 format. Essentially, the claim database is a multilayer network of multiple connected edges of various attributes, such as co-occurrences characterized by patients' demographics (e.g., sex and race, categorical), disease type (categorical), and insurance (e.g., providers and state).

Task Characterization. More than 12 hours' extensive discussions with five professional public health researchers and a year spent in designing Necklace highlight the importance of synopsis tasks of multivariate and multipatient data queries. These discussions and design do not set focused on PTSD but instead on claim analysis in general and how visualization could impact the public health domain. Important tasks that can be benefited by visualizations are those relevant to managing multi-patient and multiple variables in an interactive system to let them slice and dice and drill down variables and relationships to reveal unknown patterns in data. A participant said "If we select all patients with similar conditions, we can see related conditions associated with this patient group in terms of *what* these conditions are, *when* the highest cost occurs with a series of co-current events? and *how* the claim progressed." Participants also suggest that during interactive query, the interaction could start with a specific query of a population and go up a population to see a trend, e.g., "let's see not just females but the entire population distribution or vice versa".

2.2 Guided Interaction

Since the main challenge is to leverage co-occurring diagnoses among multivariate relationships and these relationships are often complex, we use *guided interaction*, i.e., interaction supported by visual display, to show *contextual information* to help users formulate more effective actions or knowledge with minimal visual uncertainty. In a sense, the visualizations on the interface themselves will steer viewer-controlled data exploration.

We provide action-guidance by effective visualization of variables of current interest enabled by multiple comparisons and layered network visualizations of these rings (pearls). Rings of are two kinds: root and branch rings. The root ring contains an overview of diagnosis associated with PTSD. Clicking a node (diagnosis) on the ring brings a branch-ring showing all co-occurring diagnoses. After viewers click on a node, Necklace shows the events related to the

*e-mail: chen.8028@osu.edu

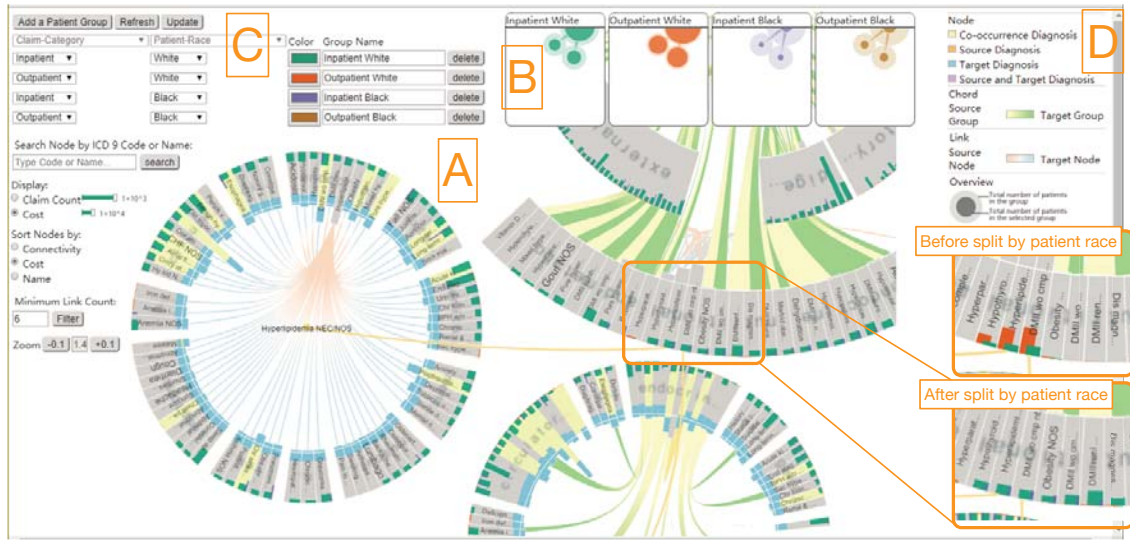


Figure 1: The Necklace User Interface for PTSD Claim Data. The interface has four parts: (A) the main interactive visualizations; (B) overview subgraphs; (C) patient group exploration, and (D) legend of visual mappings in these pearls. The callouts in the orange boxes show the multivariate comparison interface by patient group. This “Before split by patient race” comparing inpatient and outpatient costs (inpatient in green, outpatient in orange) shows that outpatient leads to higher average costs than inpatient for “endocrine, nutritional and metabolic diseases and immunity disorders.” When we group the data by race, however, it shows that race is not a deterministic factor in the inpatient population for “hyperlipidemia” (elevated cholesterol), revealing that high cholesterol introduces similar costs in inpatient blacks as whites.

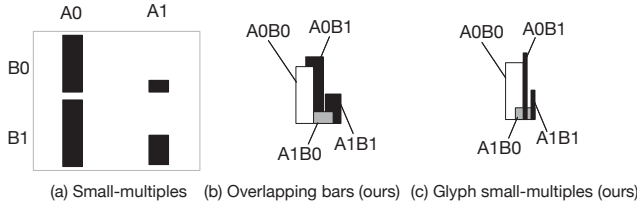


Figure 2: Design Alternatives for Multivariate Category Comparisons.

selected diagnosis in a new branch ring. Viewer can further divide the data by categories and all pearls automatically update themselves in response to the viewer’s input. The incidence rate of diagnoses is the average of all patient costs associated with that node, encoded into the height of a bar overlaid on the corresponding node. The costs of selected types are shown as cost bars in nodes, and thus users can find diagnosis events which lead to high cost intuitively (Figure 1, “after split by patient race”). Figures in the supplementary material show a workflow of the guided exploration process.

2.3 Multivariate Comparison Design

Glyphs are presented for comparisons in several ways. First, the overlapping bars in a single glyph condition reduce the space over side-by-side bars; these bars are ordered with higher values at the bottom and the bar height shows the average cost per patient cohort category. This design permits seeing trends at a glance. For example, the “Before race split by patient” (Figure 1 callout) shows that outpatient claims are progressively more expensive in the three categories, *hypothyroid*, *hyperlipidemia (high cholesterol)*, and *diabetes mellitus without mention of complication... or DMII w/o cmp...* in the *endocrine, nutritional and metabolic diseases and immunity disorders* group and the average claims for the outpatient (perhaps incurred afterwards) is much higher than those submitted by inpatient. The overlapping method in Figure 2(b) lets us reduce the space used by a fourth compared to Figure 2(a). It also accelerates

trend-detection based on our qualitative analysis.

To answer questions such as “which patient cohort costs most when data are categorized by inpatient or outpatient type and by race (black or white)?”, we design categorical glyph small multiples (GSM), with each glyph displaying multiple levels in two categories (A and B). Here categories (A and B) are encoded as bar width (wide and slim) and levels within the dominant category in different colors (white and gray for the two levels in A) and nondominant category in black (two quantitative information in B). These bars are overlaid so that both trend and relations within and between categories can be found quickly. Compared to small-multiples, GSM lets us further reduce the space used by half.

3 CONCLUSION

Identifying groups of individuals or entities relationships in health-care claims can benefit from carefully crafted comparative data visualizations to guide viewers complex multivariate analyses. This design can be reused for other multiple variables, time series and multilayer network visualizations. The supplementary material shows the entire exploration workflow.

ACKNOWLEDGMENTS

This work is supported in part by DoD USAMRAA-13318046, USA, and the Open Project Program of the State Key Lab of CAD&CG (No. A1913), Zhejiang University, China. The work is carried out when X. Bian visits the Interactive Visual Computing Lab at UMBC in the US.

REFERENCES

- [1] J. Walny, S. Carpendale, N. Henry Riche, G. Venolia, and P. Fawcett. Visual thinking in action: Visualizations as used on whiteboards. *IEEE TVCG*, 17(12):2508–2517, 2011.
- [2] Adam Perer, Fei Wang, and Jianying Hu. Mining and exploring care pathways from electronic medical records with visual analytics. *Journal of Biomedical Informatics*, 56:369 – 378, 2015.
- [3] D. Gotz and H. Stavropoulos. DecisionFlow: Visual analytics for high-dimensional temporal event sequence data. *IEEE TVCG*, 20(12):1783–1792, 2014.

Summary of Quarterly Reports

(a) Award Number: W81XWH-15-2-0016

(b) Log Number: GRANT11367272

(c) Project Title: An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data

(d) Principal Investigator Name: Dr. Jesus Caban

(e) Principal Investigator Organization and Address:

National Intrepid Center of Excellence

Walter Reed National Military Medical Center

8901 Wisconsin Ave

Bethesda, Maryland 20889

(f) Principal Investigator Phone and Email:

jesus.j.caban.civ@mail.mil

(301) 319-3759

(g) Report Date: 30 July 2018

(h) Report Period: 15 April 2018 - 15 July 2018

SECTION 1 -- Accomplishments:

What were the major goals of the project?

The first major goal of this project is to design a visual analytic framework that combines multiple clinical measurements and allows the exploration of large collections of clinical data. In addition, the second major goal is to validate the effectiveness and usability of different visualization techniques for exploring large collections of clinical variables with complex associations.

In this quarter, the PI focused on:

1. [Aims 1.3, 1.7, 1.8] Started to optimize, enhance, and increase the flexibility of the *Patient Timeline* tool based on comments and feedback received from different clinical staff members and the latest technology. Specifically, started to convert the initial prototype to Vue.js, a JavaScript framework capable of powering advanced web applications. The initial prototype of the *Patient Timeline tool* was developed to assist physicians explore the longitudinal medical data of a patient. The software application continues to be enhanced so it can be better suitable for a potential technology transfer. In addition, the front and back-end of the system have been upgraded to be more scalable to real-world clinical data that is often noisy and scatter. See figures #1 and #2 for more screenshots of the application.

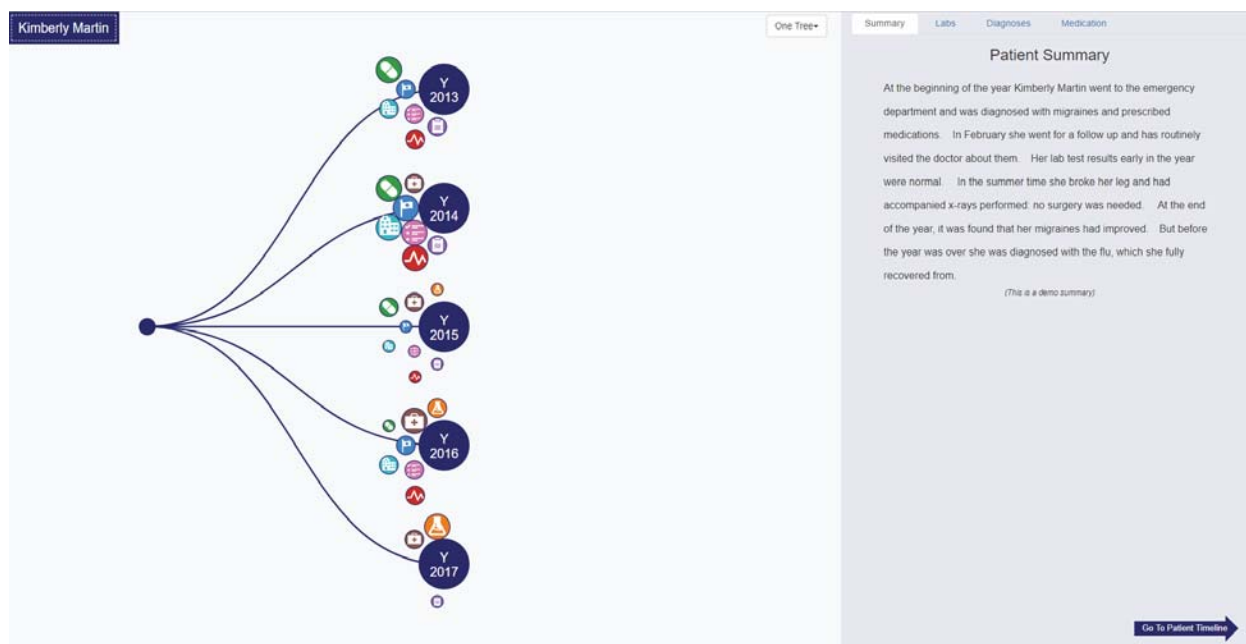


Figure 1: The Patient Timeline starts by displaying a tree that provides a quick glimpse of the data available for a patient.

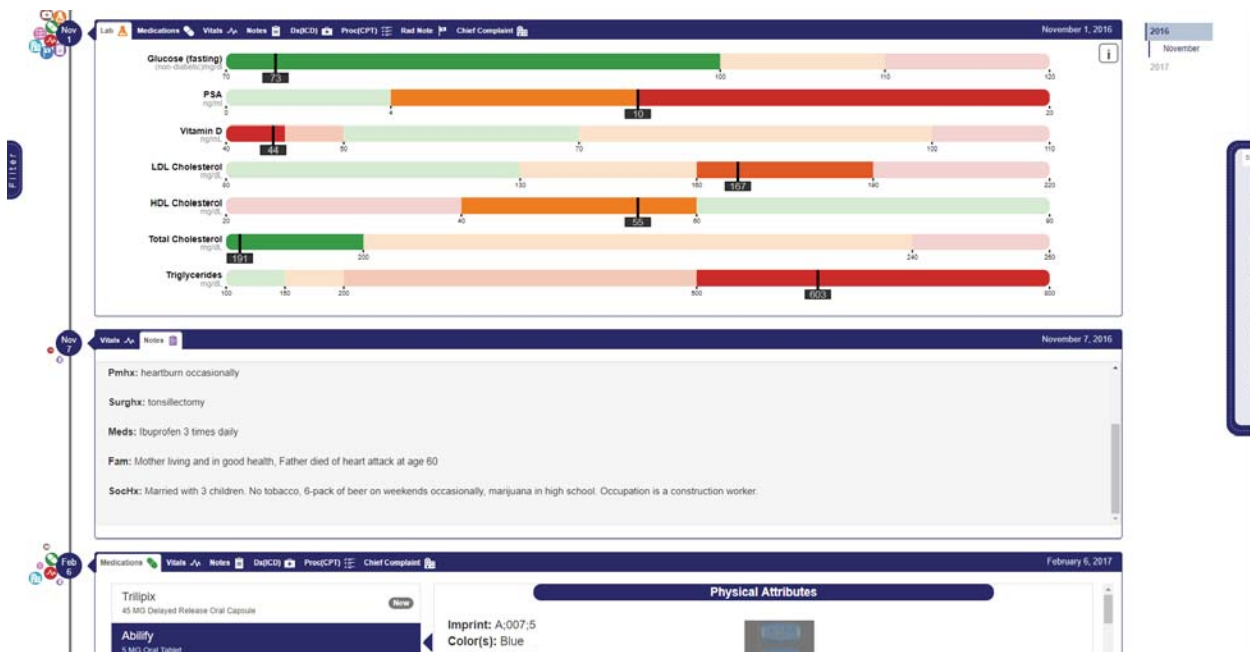


Figure 2: The nodes are presented on a timeline and individual days of the patient's data can be explored in panels.

2. [Aims 1.3, 1.7, 1.8, 1.9] During the last quarter, we completely redesigned and updated an application that uses Natural Language Processing (NLP) to analyze clinical notes and extract terms. The application allows a provider to easily search and explore notes for a patient. The tool automatically highlights clinical concepts within unstructured provider notes. The new module can be used to compare clinical notes and highlight changes over time. In addition, the new module can be used to identify commonalities between different EHR notes. Figure 3 shows an example of the application. Figure 4-5 show the latest interface to visually review clinica notes.

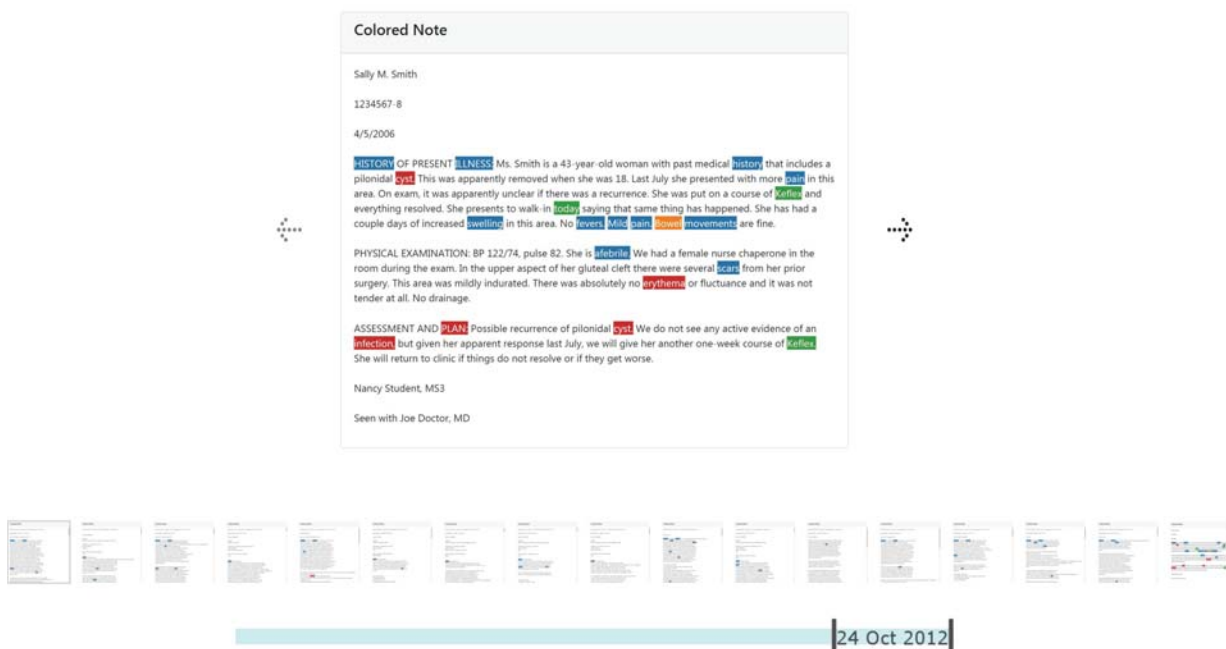


Figure 3: The nodes are presented on a timeline and individual days of the patient's data can be explored in panels.

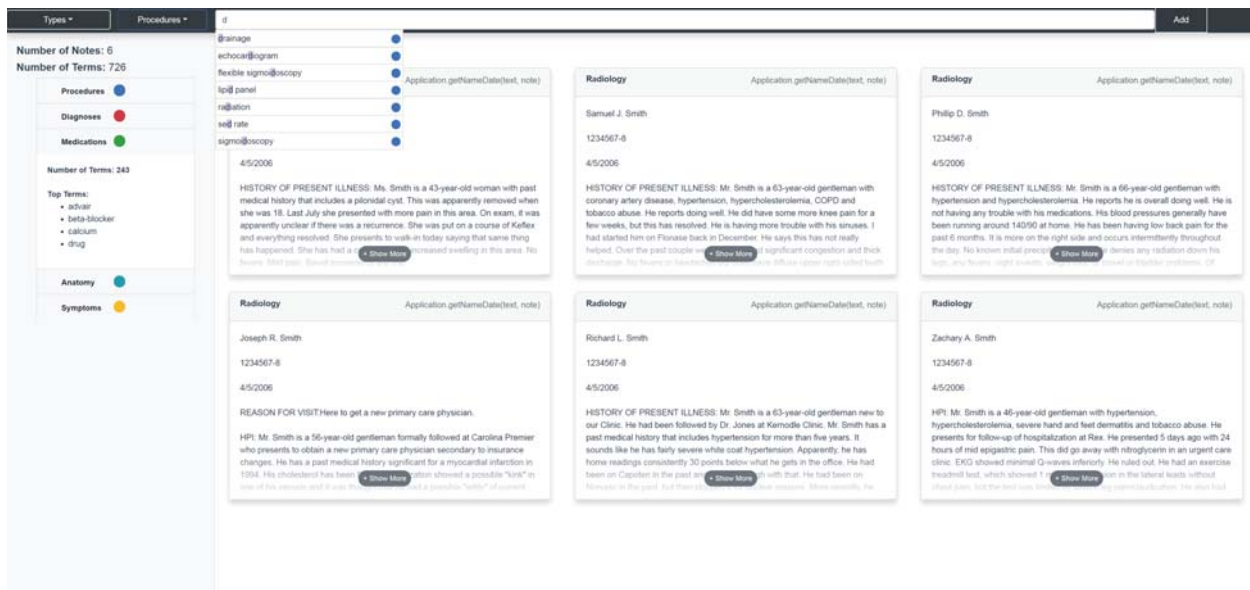


Figure 4: New application that allows providers to search keywords and clinical concepts within clinical notes.

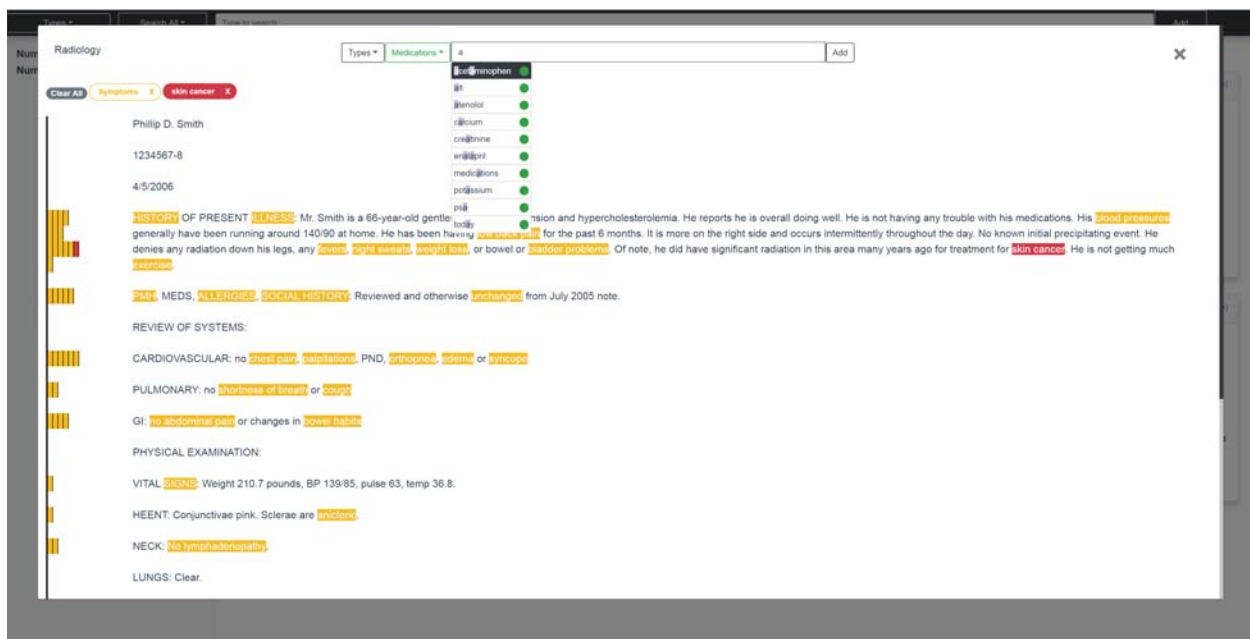


Figure 5: Screenshots of the module under development to analyze / compare clinical notes.

3. [Aims 1.3, 1.7, 1.8, 1.9] In order to better understand and validate the effectiveness of different visualization tools, we have developed a method to capture how users interact with the different systems. Given a clinical dataset and a visualization tool or dashboard, we designed an application to compare the correlations, patterns, and flows different groups follow to reach an answer for a specific clinical question.

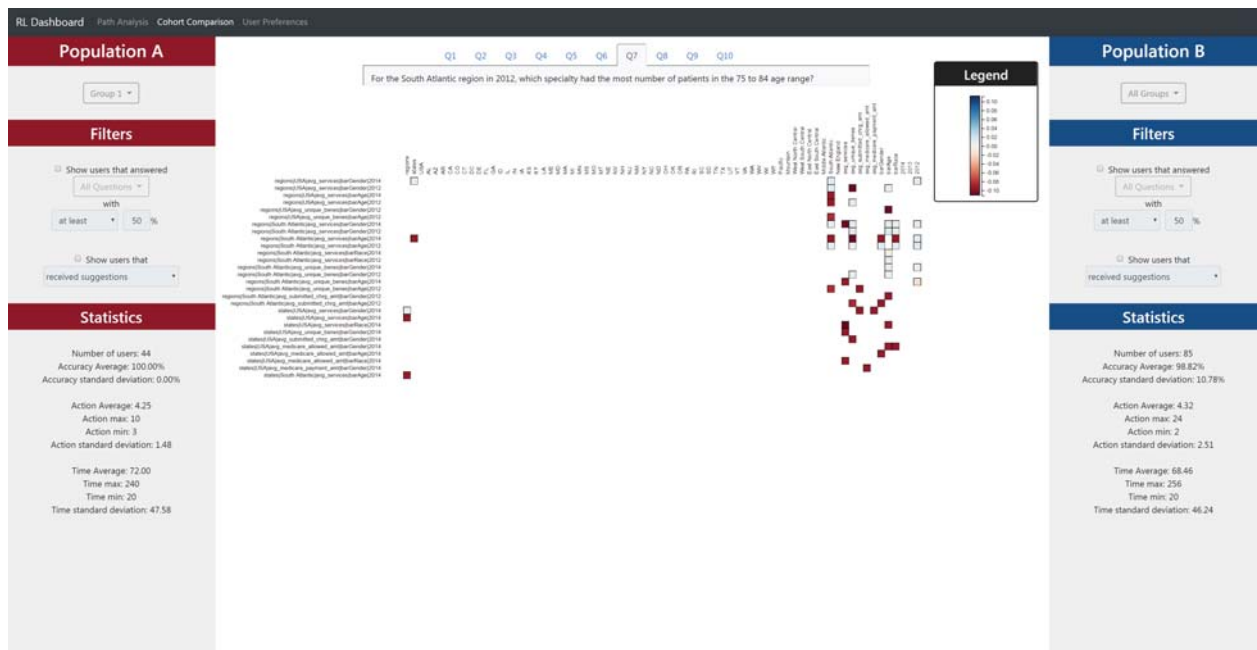


Figure 4: Screenshot of a tool designed to explore and understand how users interact with different visualization tools.

What was accomplished under these goals?

Major Task 1: Design and development of clinical visual analytics framework		
Task	Comments	% Completed
Subtask 1: New software developer and research assistant will be hired.	After 6 months recruiting for a software developer capable of working within the DoD and willing to accept the low salary that was approved under this award, we were able to hire Elizabeth Jimenez to support the project.	100%
Subtask 2: Team will compare and evaluate existing data visualization libraries and software tools.	This task has been completed and the PI (as an international leader in data visualization) continues to assess new software tools and programming languages to make sure	100%
Subtask 3: Design a modularized visual analytics framework.	The PI continues to develop a state-of-the-art visualization framework to support the analysis, integration, and visualization of clinical variables of patients diagnosed with mTBI. During the last quarter new functionalities have been implemented including: (1) a patient timeline to allow providers to explore the clinical variables of a patient over time (2) machine learning algorithm to learn and model how users interact with the visualization systems (3) algorithm to summarize and visually represent clinical trajectories (4) visualization technique that uses a Sankey	80%

diagram and a Sankey-like diagram to visualize clinical trajectories.		
Subtask 4: Obtain previously collected multi-modal clinical data for 300 mTBI/PTSD subjects and load data into the visualization system. The data will be obtained from the NICoE at no cost to this project.	Data has been obtained. IRB protocol continues to be active. Submitted continuing review and received approval until Dec 31 2017.	100%
Subtask 5: Perform preliminary validation of different visualization techniques	This task has been completed and the PI (as an international leader in data visualization) continues to validate, incorporate, and assess different visualization techniques that can be included within the framework.	100%
Subtask 6: Prepare and submit research progress report	All quarterly progress reports and annual reports have been submitted.	100%
Subtask 7: Update application based on preliminary validation with clinicians and researchers.	The team continues to work closely with several providers that are giving feedback, comments, and suggestions about the usability of the system, the effectiveness of different visualization techniques, and the specific tasks that are more relevant for the different clinical disciplines.	75%
Subtask 8: Update system based on usability study and suggestions received from users.	The system continues to improve as the team receives comments and suggestions from users. We are still collecting data from a usability study to be able to update the system based on the feedback of many users within a controlled environment.	75%
Subtask 9: Share initial framework with different federal organizations including WRNMMC, MHS, DCoE, and NIH. Share final software application with the US Army Medical Research and Material Command and other federal organizations.	The framework, novel visualization design, and ideas have been shared with a number of researchers including WRNMMC, NICoE, USU, San Antonio, Madigan, DHA, and other DoD organizations. As more features, enhancements, and modules are added to the tool, we will continue to disseminate the tool following appropriate channels.	60%
Subtask 10: Prepare final report	All quarterly progress reports and annual reports have been submitted.	60%

Describe the Regulatory Protocol and Activity Status (if applicable).

The primary objective of this project is to design and develop a visual analytic framework that combines multiple clinical measurements and allows the exploration of large collections of clinical data. To develop the software tool no clinical data is needed. Only when validating the effectiveness of the application is when de-identified clinical data will be needed. To obtain data, an existing retrospective IRB protocol is being used.

(a) Human Use Regulatory Protocols

TOTAL PROTOCOL(S): Integration of Image and Clinical Measurements of TBI Patients using Machine Learning Techniques

PROTOCOL(S): WRNMMC IRB #374953

Submitted to and Approved by:

- Submitted and reviewed by Walter Reed National Military Medical Center, Department of Research Programs
- Submitted and reviewed the USAMRMC regulatory office
- Active since 2012

Status: an existing retrospective IRB protocol (that has been reviewed by USAMRMC multiple times) will be used. During the fall of 2015 an IRB amendment will be submitted to explicitly include visualization techniques within the list of techniques that will be used to analyze and explore the data. Submitted continuing review and approved until Dec 2018.

(b) Use of Human Cadavers for Research Development Test & Evaluation (RDT&E), Education or Training

Not applicable.

(c) Animal Use Regulatory Protocol

Not applicable

What do you plan to do during the next reporting period to accomplish the goals and objectives?

During the next reporting period, further progress will be made on the visual analytics framework by continuing to research new and novel techniques that can be used to illustrate data. In addition, we hope to finalize the questions that will be used to validate the different systems. Finally, some of the tools developed by the group might be tested within the DoD enterprise system.

SECTION 2 – Products:

Nothing to report

SECTION 3 - Participants & Other Collaborating Organizations

What individuals have worked on the project?

Name: Jesus Caban, PhD

Project Role: PI

Nearest person month worked: 1

Contribution to Project: Dr. Caban has organized meetings, tracked progress of the project, and evaluated various visualization techniques for exploring large clinical data.

Name: Filip Dabek

Project Role: N/A (Data Scientist for the National Intrepid Center of Excellence)

Nearest person month worked: 1

Contribution to Project: Mr. Dabek has begun implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

Name: Elizabeth Jimenez

Project Role: Jr. Software Developer

Nearest person month worked: 3

Contribution to Project: Ms. Jimenez worked on implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

SECTION 4 –Changes/Problems:

1. Actual Problems or delays and actions to resolve them

Nothing to Report

2. Anticipated Problems/Issues

Nothing to Report

SECTION 5 – Special Reporting Requirements: Quad Chart

(a) Award Number: W81XWH-15-2-0016

(b) Log Number: GRANT11367272

(c) Project Title: An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data

(d) Principal Investigator Name: Dr. Jesus Caban

(e) Principal Investigator Organization and Address:

National Intrepid Center of Excellence

Walter Reed National Military Medical Center

8901 Wisconsin Ave

Bethesda, Maryland 20889

(f) Principal Investigator Phone and Email:

jesus.j.caban.civ@mail.mil

(301) 319-3759

(g) Report Date: 30 Oct 2018

(h) Report Period: 16 July 2018 - 15 Oct 2018

SECTION 1 -- Accomplishments:

What were the major goals of the project?

The first major goal of this project is to design a visual analytic framework that combines multiple clinical measurements and allows the exploration of large collections of clinical data. In addition, the second major goal is to validate the effectiveness and usability of different visualization techniques for exploring large collections of clinical variables with complex associations.

In this quarter, the PI focused on:

1. [[Aims 1.3, 1.7, 1.8, 1.9] The TBI Data Explorer and TBI Parsers projects are being developed in tandem, with a goal of transforming the way researchers collect, store, parse, and check data. The new TBI Parser process consists of a custom system, developed in Python, that receives a JSON template file containing the structure of the clinical document to be parsed. To assist in the process of creating and debugging template files, the TBI Data Explorer Interface was developed with two main sections: Parser and QC. The interface was developed using VueJS and Bootstrap.

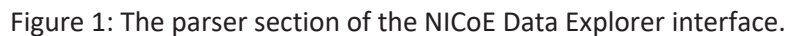
The TBI Data Explorer Parser tab will help with the development of final version of these templates. The first version/base template is built using only a few files of sample data, ensuring that the template file captures the overall structure of the clinical document. After the initial template has been built, the Parser tab can be used to quickly view how changes will affect the data and to optimize the various regular expressions necessary. Therefore, upon loading the Parser tab the user chooses the template that they are working on as well as any additional information and then clicks “Load” for the system to generate the interface for the specific template. In Figure 1, the left section shows the user the fields that have been identified in the template; it also keeps track of which fields have been visited by fading their color and which field is currently being viewed by changing its font color. The middle section contains the JSON that belongs to the currently selected field on the left-most section. The right section shows all of the possible raw data values that were extracted from the clinical documents, as well as the resulting “computed” and “comment” values that are extracted based on their field configuration. On click of Run, or when the user clicks outside of the orange-bordered editable section, the right-most section updates in real-time.

The second section, titled QC for Quality and Control, assists in the process of verifying that the data extracted using a given template is correct. The current QC process consisted of manually opening clinical documents and comparing the values within the clinical document against a large and difficult to scroll through Excel spreadsheet of values. Therefore, for this interface the goal was to build an easy to use interface that allows for the user that is QCing the data to be able to: focus on their task at hand, speed up the process, maintain a high level of quality control or to even exceed the current level of quality control.

The resulting interface is shown in Figure 1. In this interface the clinical document is shown on the left, removing the need for the user to have to locate and open the document as in the existing process, and the extracted values are shown on the right hand side. At the top of the screen the user is shown a progress bar which is updated as they complete each clinical document, allowing for them to track their progress.

The idea of the right hand side of the screen is for the user to indicate whether or not the values extracted match what is shown within the clinical document. In the existing process, the user would simply mark a cell within an Excel spreadsheet as green or red. However, in this new interface and process the user selects “Yes” or “No” and then an assortment of options are available to them.

All of these steps, including the alerting of missing information, is updated in real time for the user and is a tremendous improvement over existing research QC processes.



Invalid Sections

WJ-III

Rapid_Pix

SLP_Woodcock_Johnson_RapidPic

RBMT

Scaled Scores

SLP_RBMT3_SM_Route_Immediate_Recall

- WJ-III ✓ -
 - Rapid Pix ✓ -
 - SLP Woodcock Johnson RapidPic ✓ -

Raw	Computed	Comment
10/27	10	/27

Value Correct: Yes No

Reason Value Incorrect: Partially Parsed

Computed True Value:

Comment True Value:
 - SLP Woodcock Johnson RapidPic Timing ✓ -

Raw	Computed	Comment
10/27	10	/27

Value Correct: Yes No
- RBMT ✓ -
 - Scaled Scores ✓ -
 - SLP RBMT3 PM Belongings Delayed Recall ✓ +
 - SLP RBMT3 SM Route Immediate Recall ✓ -

Raw	Computed	Comment
10/27	10	/27

Value Correct: Yes No

Submit

Figure 3: An example of a user indicating whether or not the data extracted matches the data in the original clinical document, and the new process of entering information pertaining to match. When a submission is attempted, if it fails, the top populates with the sections that have failed the validation process.

[Aims 1.3, 1.7, 1.8, 1.9] Following the success and feedback received from the Patient Timeline that was developed in previous quarters, a version 2 of the interface/system was discussed, mocked up, and designed which is shown in Figure 4. The design is currently in progress and has continued to undergo multiple iterations, but through the design of this mockup more productive discussions have been able to be had. In the next quarter, the development of this timeline is expected to begin such that an interface for users and clinicians to utilize will be built.

Select a Patient

Types

Type to search

Statistics

Bookmarks

History

Note: 331
Last Viewed:
10 - 31-2018 16:37:23

Note: 1039
Last Viewed:
10 - 31-2018 16:37:31

Clear History

Clear All

Medications

skin

Physical Therapy

Admission Date: ["*2124-1-24*"] Discharge Date: ["*2124-3-8*"]

Service: CARDIOTHORACIC

Allergies:
Percocet

Attending ["*First Name3 (LF) 165*"]
Chief Complaint:
Increasing SOB

Major Surgical or Invasive Procedures

Show More

Audiology

PATIENT/TEST INFORMATION:
Indication: Congestive heart failure.
Height (n): 67
Weight (b): 251
BSA (m²): 2.23 m²
BP (mm Hg): 118/72
HR (bpm): 70
Status: Inpatient
Date/Time: ["*2124-1-25*"] at 10:51
Test: TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Intraop redo redo sternotomy, MVR, TVR. Evaluate valvular function, ventricular function, Aortic Atheroma.
Weight (b): 199
BP (mm Hg): 102/63
HR (bpm): 85
Status: Inpatient
Date/Time: ["*2124-2-3*"] at 10:32
Test: TEE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Post-op with significant wound drainage, hypotension, evaluate for pericardial effusion
Height (n): 67
Weight (b): 200
BSA (m²): 2.02 m²
BP (mm Hg): 95/53
HR (bpm): 75
Status: Inpatient
Date/Time: ["*2124-2-9*"] at 21:19
Test: Portable TTE (Focused views)
Doppler: Limited Doppler
Contrast: None

Show More

Art Therapy

PATIENT/TEST INFORMATION:
Indication: ? Congestive heart failure.
Weight (b): 210
BP (mm Hg): 125/65
HR (bpm): 78
Status: Inpatient
Date/Time: ["*2124-2-16*"] at 10:55
Test: Potable TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Art Therapy

Baseline artifact precludes assessment of atrial mechanism, which could be atrial fibrillation, etc. Ventricular pacing at a rate of 70 beats per minute.
Compared to the previous tracing of ["*2124-2-1*"], allowing of lead placement variation, no diagnostic change. Clinical correlation is suggested.
TRACING #2

Show More

Select a Patient

Types

head

I

Statistics

Bookmarks

History

Note: 331
Last Viewed:
10 - 31-2018 16:37:23

Note: 1039
Last Viewed:
10 - 31-2018 16:37:31

Clear History

Clear All

Medications

head

heart

heart block

heart failure

ct head

ct head w/o contrast

head ct

heart surgery

healing

Physical Therapy

Admission Date: ["*2124-3-8*"]

Service: CARDIOTHORACIC

Allergies:
Percocet

Attending ["*First Name3 (LF) 165*"]
Chief Complaint:
Increasing SOB

Major Surgical or Invasive Procedures

Show More

Audiology

PATIENT/TEST INFORMATION:
Indication: Congestive heart failure.
Height (n): 67
Weight (b): 251
BSA (m²): 2.23 m²
BP (mm Hg): 118/72
HR (bpm): 70
Status: Inpatient
Date/Time: ["*2124-1-25*"] at 10:51
Test: TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Intraop redo redo sternotomy, MVR, TVR. Evaluate valvular function, ventricular function, Aortic Atheroma.
Weight (b): 199
BP (mm Hg): 102/63
HR (bpm): 85
Status: Inpatient
Date/Time: ["*2124-2-3*"] at 10:32
Test: TEE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Post-op with significant wound drainage, hypotension, evaluate for pericardial effusion
Height (n): 67
Weight (b): 200
BSA (m²): 2.02 m²
BP (mm Hg): 95/53
HR (bpm): 75
Status: Inpatient
Date/Time: ["*2124-2-9*"] at 21:19
Test: Portable TTE (Focused views)
Doppler: Limited Doppler
Contrast: None

Show More

Art Therapy

PATIENT/TEST INFORMATION:
Indication: ? Congestive heart failure.
Weight (b): 210
BP (mm Hg): 125/65
HR (bpm): 78
Status: Inpatient
Date/Time: ["*2124-2-16*"] at 10:55
Test: Potable TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Art Therapy

Baseline artifact precludes assessment of atrial mechanism, which could be atrial fibrillation, etc. Ventricular pacing at a rate of 70 beats per minute.
Compared to the previous tracing of ["*2124-2-1*"], allowing of lead placement variation, no diagnostic change. Clinical correlation is suggested.
TRACING #2

Show More

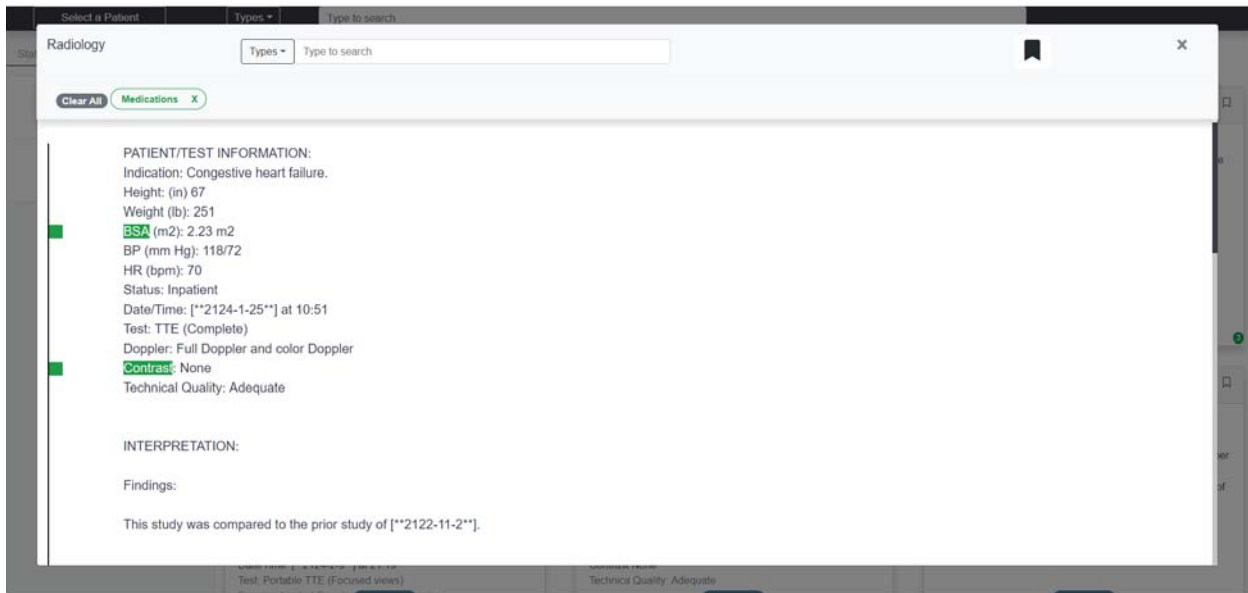


Figure 5 (a, b, & c): Screenshots of the module under development to analyze / compare clinical notes. This application that allows providers to search keywords and clinical concepts within clinical notes.

What was accomplished under these goals?

Major Task 1: Design and development of clinical visual analytics framework		
Task	Comments	% Completed
Subtask 1: New software developer and research assistant will be hired.	After 6 months recruiting for a software developer capable of working within the DoD and willing to accept the low salary that was approved under this award, we were able to hire Elizabeth Jimenez to support the project.	100%
Subtask 2: Team will compare and evaluate existing data visualization libraries and software tools.	This task has been completed and the PI (as an international leader in data visualization) continues to assess new software tools and programming languages to make sure	100%
Subtask 3: Design a modularized visual analytics framework.	The PI continues to develop a state-of-the-art visualization framework to support the analysis, integration, and visualization of clinical variables of patients diagnosed with mTBI. During the last quarter new functionalities have been implemented including: (1) a patient timeline to allow providers to explore the clinical variables of a patient over time (2) machine learning algorithm to learn and model how users interact with the visualization systems (3) algorithm to summarize and visually represent clinical trajectories (4) visualization technique that uses a Sankey diagram and a Sankey-like diagram to	85%

visualize clinical trajectories.

Subtask 4: Obtain previously collected multi-modal clinical data for 300 mTBI/PTSD subjects and load data into the visualization system. The data will be obtained from the NICoE at no cost to this project.	Data has been obtained. IRB protocol continues to be active. Submitted continuing review and received approval until Dec 31 2017.	100%
Subtask 5: Perform preliminary validation of different visualization techniques	This task has been completed and the PI (as an international leader in data visualization) continues to validate, incorporate, and assess different visualization techniques that can be included within the framework.	100%
Subtask 6: Prepare and submit research progress report	All quarterly progress reports and annual reports have been submitted.	100%
Subtask 7: Update application based on preliminary validation with clinicians and researchers.	The team continues to work closely with several providers that are giving feedback, comments, and suggestions about the usability of the system, the effectiveness of different visualization techniques, and the specific tasks that are more relevant for the different clinical disciplines.	80%
Subtask 8: Update system based on usability study and suggestions received from users.	The system continues to improve as the team receives comments and suggestions from users. We are still collecting data from a usability study to be able to update the system based on the feedback of many users within a controlled environment.	80%
Subtask 9: Share initial framework with different federal organizations including WRNMMC, MHS, DCoE, and NIH. Share final software application with the US Army Medical Research and Material Command and other federal organizations.	The framework, novel visualization design, and ideas have been shared with a number of researchers including WRNMMC, NICoE, USU, San Antonio, Madigan, DHA, and other DoD organizations. As more features, enhancements, and modules are added to the tool, we will continue to disseminate the tool following appropriate channels.	70%
Subtask 10: Prepare final report	All quarterly progress reports and annual reports have been submitted.	70%

Describe the Regulatory Protocol and Activity Status (if applicable).

The primary objective of this project is to design and develop a visual analytic framework that combines multiple clinical measurements and allows the exploration of large collections of clinical data. To develop the software tool no clinical data is needed. Only when validating the effectiveness of the application is when de-identified clinical data will be needed. To obtain data, an existing retrospective IRB protocol is being used.

(a) Human Use Regulatory Protocols

TOTAL PROTOCOL(S): Integration of Image and Clinical Measurements of TBI Patients using Machine Learning Techniques

PROTOCOL(S): WRNMMC IRB #374953

Submitted to and Approved by:

- Submitted and reviewed by Walter Reed National Military Medical Center, Department of Research Programs
- Submitted and reviewed the USAMRMC regulatory office
- Active since 2012

Status: an existing retrospective IRB protocol (that has been reviewed by USAMRMC multiple times) will be used. During the fall of 2015 an IRB amendment will be submitted to explicitly include visualization techniques within the list of techniques that will be used to analyze and explore the data. Submitted continuing review and approved until Dec 2018. A continuing review has been submitted to extend our research effort until December 2019.

(b) Use of Human Cadavers for Research Development Test & Evaluation (RDT&E), Education or Training

Not applicable.

(c) Animal Use Regulatory Protocol

Not applicable

What do you plan to do during the next reporting period to accomplish the goals and objectives?

During the next reporting period, further progress will be made on the visual analytics framework by continuing to research new and novel techniques that can be used to illustrate data. In addition, the new data explorer and QC system will be tested with a lot of data. Finally, some of the tools developed by the group might be tested within the DoD enterprise system.

SECTION 2 – Products:

Nothing to report

SECTION 3 - Participants & Other Collaborating Organizations**What individuals have worked on the project?**

Name: Jesus Caban, PhD

Project Role: PI

Nearest person month worked: 1

Contribution to Project: Dr. Caban has organized meetings, tracked progress of the project, and evaluated various visualization techniques for exploring large clinical data.

Name: Filip Dabek

Project Role: N/A (Data Scientist for the National Intrepid Center of Excellence)

Nearest person month worked: 1

Contribution to Project: Mr. Dabek has begun implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

Name: Elizabeth Jimenez

Project Role: Jr. Software Developer

Nearest person month worked: 3

Contribution to Project: Ms. Jimenez worked on implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

SECTION 4 –Changes/Problems:

1. Actual Problems or delays and actions to resolve them

Nothing to Report

2. Anticipated Problems/Issues

Nothing to Report

SECTION 5 – Special Reporting Requirements: Quad Chart

(a) Award Number: W81XWH-15-2-0016

(b) Log Number: GRANT11367272

(c) Project Title: An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data

(d) Principal Investigator Name: Dr. Jesus Caban

(e) Principal Investigator Organization and Address:

National Intrepid Center of Excellence

Walter Reed National Military Medical Center

8901 Wisconsin Ave

Bethesda, Maryland 20889

(f) Principal Investigator Phone and Email:

jesus.j.caban.civ@mail.mil

(301) 319-3759

(g) Report Date: 30 Jan 2019

(h) Report Period: 16 Oct 2018 - 15 Jan 2019

SECTION 1 -- Accomplishments:

What were the major goals of the project?

The first major goal of this project is to design a visual analytic framework that combines multiple clinical measurements and allows the exploration of large collections of clinical data. In addition, the second major goal is to validate the effectiveness and usability of different visualization techniques for exploring large collections of clinical variables with complex associations.

In this quarter, the PI focused on:

1. [[Aims 1.3, 1.7, 1.8, 1.9] The TBI Data Explorer and TBI Parsers projects are being developed in tandem, with a goal of transforming the way researchers collect, store, parse, and check data. The new TBI Parser process consists of a custom system, developed in Python, that receives a JSON template file containing the structure of the clinical document to be parsed. To assist in the process of creating and debugging template files, the TBI Data Explorer Interface was developed with two main sections: Parser and QC. The interface was developed using VueJS and Bootstrap.

The TBI Data Explorer Parser tab will help with the development of final version of these templates. The first version/base template is built using only a few files of sample data, ensuring that the template file captures the overall structure of the clinical document. After the initial template has been built, the Parser tab can be used to quickly view how changes will affect the data and to optimize the various regular expressions necessary. Therefore, upon loading the Parser tab the user chooses the template that they are working on as well as any additional information and then clicks “Load” for the system to generate the interface for the specific template. In Figure 1, the left section shows the user the fields that have been identified in the template; it also keeps track of which fields have been visited by fading their color and which field is currently being viewed by changing its font color. The middle section contains the JSON that belongs to the currently selected field on the left-most section. The right section shows all of the possible raw data values that were extracted from the clinical documents, as well as the resulting “computed” and “comment” values that are extracted based on their field configuration. On click of Run, or when the user clicks outside of the orange-bordered editable section, the right-most section updates in real-time.

The second section, titled QC for Quality and Control, assists in the process of verifying that the data extracted using a given template is correct. The current QC process consisted of manually opening clinical documents and comparing the values within the clinical document against a large and difficult to scroll through Excel spreadsheet of values. Therefore, for this interface the goal was to build an easy to use interface that allows for the user that is QCing the data to be able to: focus on their task at hand, speed up the process, maintain a high level of quality control or to even exceed the current level of quality control.

The resulting interface is shown in Figure 1. In this interface the clinical document is shown on the left, removing the need for the user to have to locate and open the document as in the existing process, and the extracted values are shown on the right hand side. At the top of the screen the user is shown a progress bar which is updated as they complete each clinical document, allowing for them to track their progress.

The idea of the right hand side of the screen is for the user to indicate whether or not the values extracted match what is shown within the clinical document. In the existing process, the user would simply mark a cell within an Excel spreadsheet as green or red. However, in this new interface and process the user selects “Yes” or “No” and then an assortment of options are available to them.

Figure 2 shows an example of a user marking a data value as being incorrect, another as correct, and finally being alerted that they had not marked anything for the final data value. For the first data value in Figure 3 it can be seen how the interface reacts to the user and allows them to select a reason for why the data did not match, as well as fields to enter the correct values of the data elements.

All of these steps, including the alerting of missing information, is updated in real time for the user and is a tremendous improvement over existing research QC processes.

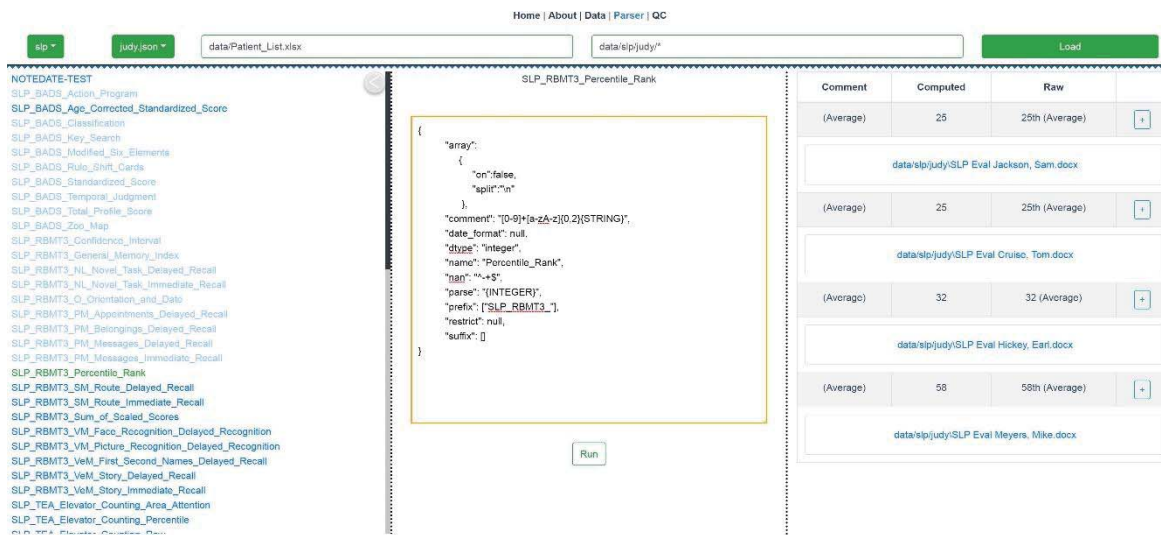


Figure 1: The parser section of the NICOE Data Explorer interface.

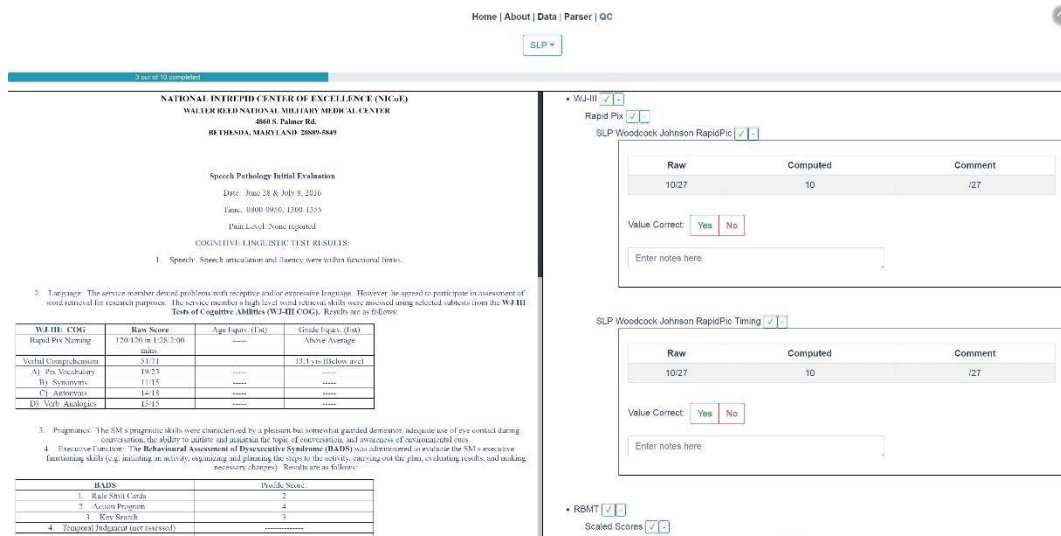


Figure 2: The QC (quality control) section of the TBI Data Explorer interface.

Invalid Sections

WJ-III

Rapid_Pix

SLP_Woodcock_Johnson_RapidPic

RBMT

Scaled Scores

SLP_RBMT3_SM_Route_Immediate_Recall

- WJ-III ✓ -
 - Rapid Pix ✓ -
 - SLP Woodcock Johnson RapidPic ✓ -

Raw	Computed	Comment
10/27	10	/27

Value Correct: Yes No

Reason Value Incorrect: Partially Parsed

Computed True Value: Enter what computed value should be

Comment True Value: Enter what comment value should be

Enter notes here
 - SLP Woodcock Johnson RapidPic Timing ✓ -

Raw	Computed	Comment
10/27	10	/27

Value Correct: Yes No

Enter notes here
- RBMT ✓ -
 - Scaled Scores ✓ -
 - SLP RBMT3 PM Belongings Delayed Recall ✓ +
 - SLP RBMT3 SM Route Immediate Recall ✓ -

Raw	Computed	Comment
10/27	10	/27

Value Correct: Yes No

Enter notes here

Submit

Figure 3: An example of a user indicating whether or not the data extracted matches the data in the original clinical document, and the new process of entering information pertaining to match. When a submission is attempted, if it fails, the top populates with the sections that have failed the validation process.

[Aims 1.3, 1.7, 1.8, 1.9] Following the success and feedback received from the Patient Timeline that was developed in previous quarters, a version 2 of the interface/system was discussed, mocked up, and designed which is shown in Figure 4. The design is currently in progress and has continued to undergo multiple iterations, but through the design of this mockup more productive discussions have been able to be had. In the next quarter, the development of this timeline is expected to begin such that an interface for users and clinicians to utilize will be built.

Select a Patient

Types

Type to search

Statistics

Bookmarks

History

Note: 331
Last Viewed:
10 - 31-2018 16:37:23

Note: 1039
Last Viewed:
10 - 31-2018 16:37:31

Clear History

Clear All

Medications

skin

Physical Therapy

Admission Date: ["*2124-1-24*"] Discharge Date: ["*2124-3-8*"]

Service: CARDIOTHORACIC

Allergies:
Percocet

Attending ["*First Name3 (LF) 165*"]
Chief Complaint:
Increasing SOB

Major Surgical or Invasive Procedures

Show More

Audiology

PATIENT/TEST INFORMATION:
Indication: Congestive heart failure.
Height (n): 67
Weight (b): 251
BSA (m²): 2.23 m²
BP (mm Hg): 118/72
HR (bpm): 70
Status: Inpatient
Date/Time: ["*2124-1-25*"] at 10:51
Test: TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Intraop redo redo sternotomy, MVR, TVR. Evaluate valvular function, ventricular function, Aortic Atheroma.
Weight (b): 199
BP (mm Hg): 102/63
HR (bpm): 85
Status: Inpatient
Date/Time: ["*2124-2-3*"] at 10:32
Test: TEE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Post-op with significant wound drainage, hypotension, evaluate for pericardial effusion
Height (n): 67
Weight (b): 200
BSA (m²): 2.02 m²
BP (mm Hg): 95/53
HR (bpm): 75
Status: Inpatient
Date/Time: ["*2124-2-9*"] at 21:19
Test: Portable TTE (Focused views)
Doppler: Limited Doppler
Contrast: None

Show More

Art Therapy

PATIENT/TEST INFORMATION:
Indication: ? Congestive heart failure.
Weight (b): 210
BP (mm Hg): 125/65
HR (bpm): 78
Status: Inpatient
Date/Time: ["*2124-2-16*"] at 10:55
Test: Potable TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Art Therapy

Baseline artifact precludes assessment of atrial mechanism, which could be atrial fibrillation, etc. Ventricular pacing at a rate of 70 beats per minute.
Compared to the previous tracing of ["*2124-2-1*"], allowing of lead placement variation, no diagnostic change. Clinical correlation is suggested.
TRACING #2

Show More

Select a Patient

Types

head

I

Statistics

Bookmarks

History

Note: 331
Last Viewed:
10 - 31-2018 16:37:23

Note: 1039
Last Viewed:
10 - 31-2018 16:37:31

Clear History

Clear All

Medications

head

heart

heart block

heart failure

ct head

ct head w/o contrast

head ct

heart surgery

healing

Physical Therapy

Admission Date: ["*2124-3-8*"]

Service: CARDIOTHORACIC

Allergies:
Percocet

Attending ["*First Name3 (LF) 165*"]
Chief Complaint:
Increasing SOB

Major Surgical or Invasive Procedures

Show More

Audiology

PATIENT/TEST INFORMATION:
Indication: Congestive heart failure.
Height (n): 67
Weight (b): 251
BSA (m²): 2.23 m²
BP (mm Hg): 118/72
HR (bpm): 70
Status: Inpatient
Date/Time: ["*2124-1-25*"] at 10:51
Test: TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Intraop redo redo sternotomy, MVR, TVR. Evaluate valvular function, ventricular function, Aortic Atheroma.
Weight (b): 199
BP (mm Hg): 102/63
HR (bpm): 85
Status: Inpatient
Date/Time: ["*2124-2-3*"] at 10:32
Test: TEE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Post-op with significant wound drainage, hypotension, evaluate for pericardial effusion
Height (n): 67
Weight (b): 200
BSA (m²): 2.02 m²
BP (mm Hg): 95/53
HR (bpm): 75
Status: Inpatient
Date/Time: ["*2124-2-9*"] at 21:19
Test: Portable TTE (Focused views)
Doppler: Limited Doppler
Contrast: None

Show More

Art Therapy

PATIENT/TEST INFORMATION:
Indication: ? Congestive heart failure.
Weight (b): 210
BP (mm Hg): 125/65
HR (bpm): 78
Status: Inpatient
Date/Time: ["*2124-2-16*"] at 10:55
Test: Potable TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Art Therapy

Baseline artifact precludes assessment of atrial mechanism, which could be atrial fibrillation, etc. Ventricular pacing at a rate of 70 beats per minute.
Compared to the previous tracing of ["*2124-2-1*"], allowing of lead placement variation, no diagnostic change. Clinical correlation is suggested.
TRACING #2

Show More

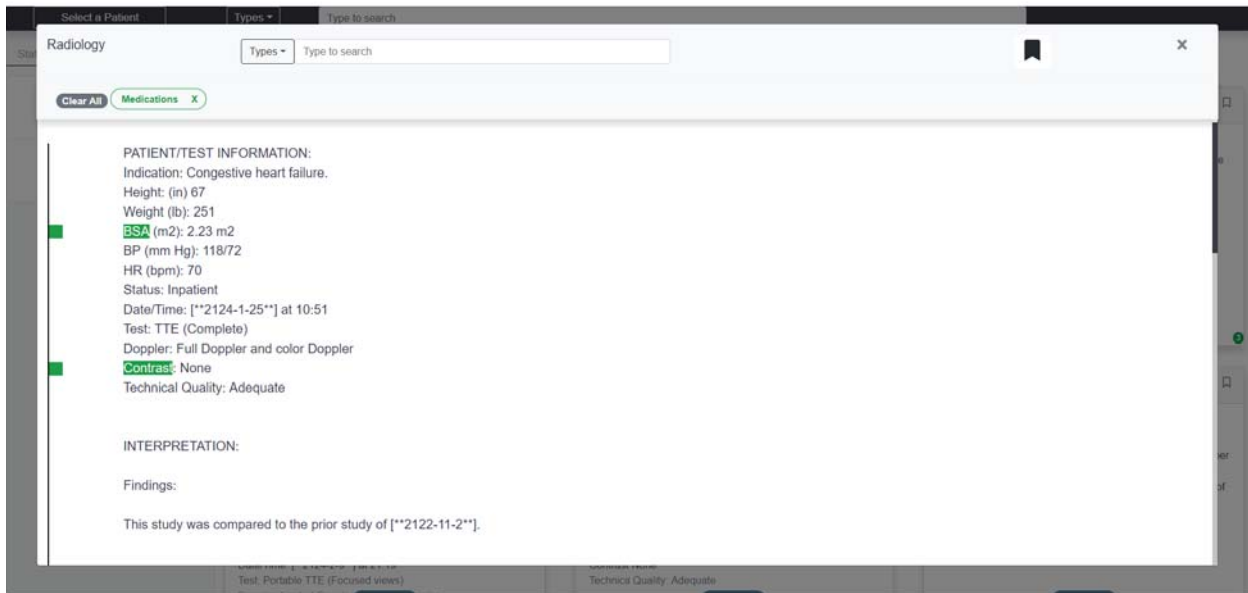


Figure 5 (a, b, & c): Screenshots of the module under development to analyze / compare clinical notes. This application that allows providers to search keywords and clinical concepts within clinical notes.

What was accomplished under these goals?

Major Task 1: Design and development of clinical visual analytics framework		
Task	Comments	% Completed
Subtask 1: New software developer and research assistant will be hired.	After 6 months recruiting for a software developer capable of working within the DoD and willing to accept the low salary that was approved under this award, we were able to hire Elizabeth Jimenez to support the project.	100%
Subtask 2: Team will compare and evaluate existing data visualization libraries and software tools.	This task has been completed and the PI (as an international leader in data visualization) continues to assess new software tools and programming languages to make sure	100%
Subtask 3: Design a modularized visual analytics framework.	The PI continues to develop a state-of-the-art visualization framework to support the analysis, integration, and visualization of clinical variables of patients diagnosed with mTBI. During the last quarter new functionalities have been implemented including: (1) a patient timeline to allow providers to explore the clinical variables of a patient over time (2) machine learning algorithm to learn and model how users interact with the visualization systems (3) algorithm to summarize and visually represent clinical trajectories (4) visualization technique that uses a Sankey diagram and a Sankey-like diagram to	85%

visualize clinical trajectories.

Subtask 4: Obtain previously collected multi-modal clinical data for 300 mTBI/PTSD subjects and load data into the visualization system. The data will be obtained from the NICOE at no cost to this project.	Data has been obtained. IRB protocol continues to be active. Submitted continuing review and received approval until Dec 31 2017.	100%
Subtask 5: Perform preliminary validation of different visualization techniques	This task has been completed and the PI (as an international leader in data visualization) continues to validate, incorporate, and assess different visualization techniques that can be included within the framework.	100%
Subtask 6: Prepare and submit research progress report	All quarterly progress reports and annual reports have been submitted.	100%
Subtask 7: Update application based on preliminary validation with clinicians and researchers.	The team continues to work closely with several providers that are giving feedback, comments, and suggestions about the usability of the system, the effectiveness of different visualization techniques, and the specific tasks that are more relevant for the different clinical disciplines.	80%
Subtask 8: Update system based on usability study and suggestions received from users.	The system continues to improve as the team receives comments and suggestions from users. We are still collecting data from a usability study to be able to update the system based on the feedback of many users within a controlled environment.	80%
Subtask 9: Share initial framework with different federal organizations including WRNMMC, MHS, DICOE, and NIH. Share final software application with the US Army Medical Research and Material Command and other federal organizations.	The framework, novel visualization design, and ideas have been shared with a number of researchers including WRNMMC, NICOE, USU, San Antonio, Madigan, DHA, and other DoD organizations. As more features, enhancements, and modules are added to the tool, we will continue to disseminate the tool following appropriate channels.	70%
Subtask 10: Prepare final report	All quarterly progress reports and annual reports have been submitted.	70%

Describe the Regulatory Protocol and Activity Status (if applicable).

The primary objective of this project is to design and develop a visual analytic framework that combines multiple clinical measurements and allows the exploration of large collections of clinical data. To develop the software tool no clinical data is needed. Only when validating the effectiveness of the application is when de-identified clinical data will be needed. To obtain data, an existing retrospective IRB protocol is being used.

(a) Human Use Regulatory Protocols

TOTAL PROTOCOL(S): Integration of Image and Clinical Measurements of TBI Patients using Machine Learning Techniques

PROTOCOL(S): WRNMMC IRB #374953

Submitted to and Approved by:

- Submitted and reviewed by Walter Reed National Military Medical Center, Department of Research Programs
- Submitted and reviewed the USAMRMC regulatory office
- Active since 2012

Status: an existing retrospective IRB protocol (that has been reviewed by USAMRMC multiple times) will be used. During the fall of 2015 an IRB amendment will be submitted to explicitly include visualization techniques within the list of techniques that will be used to analyze and explore the data. Submitted continuing review and approved until Dec 2018. A continuing review has been submitted to extend our research effort until December 2019.

(b) Use of Human Cadavers for Research Development Test & Evaluation (RDT&E), Education or Training

Not applicable.

(c) Animal Use Regulatory Protocol

Not applicable

What do you plan to do during the next reporting period to accomplish the goals and objectives?

During the next reporting period, further progress will be made on the visual analytics framework by continuing to research new and novel techniques that can be used to illustrate data. In addition, the new data explorer and QC system will be tested with a lot of data. Finally, some of the tools developed by the group might be tested within the DoD enterprise system.

SECTION 2 – Products:

Nothing to report

SECTION 3 - Participants & Other Collaborating Organizations**What individuals have worked on the project?**

Name: Jesus Caban, PhD

Project Role: PI

Nearest person month worked: 1

Contribution to Project: Dr. Caban has organized meetings, tracked progress of the project, and evaluated various visualization techniques for exploring large clinical data.

Name: Filip Dabek

Project Role: N/A (Data Scientist for the National Intrepid Center of Excellence)

Nearest person month worked: 1

Contribution to Project: Mr. Dabek has begun implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

Name: Elizabeth Jimenez

Project Role: Jr. Software Developer

Nearest person month worked: 3

Contribution to Project: Ms. Jimenez worked on implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

SECTION 4 –Changes/Problems:

1. Actual Problems or delays and actions to resolve them

Nothing to Report

2. Anticipated Problems/Issues

Nothing to Report

SECTION 5 – Special Reporting Requirements: Quad Chart

(a) Award Number: W81XWH-15-2-0016

(b) Log Number: GRANT11367272

(c) Project Title: An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data

(d) Principal Investigator Name: Dr. Jesus Caban

(e) Principal Investigator Organization and Address:

National Intrepid Center of Excellence

Walter Reed National Military Medical Center

8901 Wisconsin Ave

Bethesda, Maryland 20889

(f) Principal Investigator Phone and Email:

jesus.j.caban.civ@mail.mil

(301) 319-3759

(g) Report Date: 30 Oct 2018

(h) Report Period: 16 Oct 2018 - 15 Jan 2019

SECTION 1 -- Accomplishments:

What were the major goals of the project?

The first major goal of this project is to design a visual analytic framework that combines multiple clinical measurements and allows the exploration of large collections of clinical data. In addition, the second major goal is to validate the effectiveness and usability of different visualization techniques for exploring large collections of clinical variables with complex associations.

In this quarter, the PI focused on:

1. [[Aims 1.3, 1.7, 1.8, 1.9] The TBI Data Explorer and TBI Parsers projects are being developed in tandem, with a goal of transforming the way researchers collect, store, parse, and check data. The new TBI Parser process consists of a custom system, developed in Python, that receives a JSON template file containing the structure of the clinical document to be parsed. To assist in the process of creating and debugging template files, the TBI Data Explorer Interface was developed with two main sections: Parser and QC. The interface was developed using VueJS and Bootstrap.

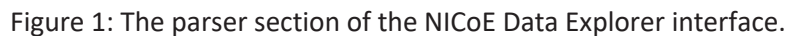
The TBI Data Explorer Parser tab will help with the development of final version of these templates. The first version/base template is built using only a few files of sample data, ensuring that the template file captures the overall structure of the clinical document. After the initial template has been built, the Parser tab can be used to quickly view how changes will affect the data and to optimize the various regular expressions necessary. Therefore, upon loading the Parser tab the user chooses the template that they are working on as well as any additional information and then clicks “Load” for the system to generate the interface for the specific template. In Figure 1, the left section shows the user the fields that have been identified in the template; it also keeps track of which fields have been visited by fading their color and which field is currently being viewed by changing its font color. The middle section contains the JSON that belongs to the currently selected field on the left-most section. The right section shows all of the possible raw data values that were extracted from the clinical documents, as well as the resulting “computed” and “comment” values that are extracted based on their field configuration. On click of Run, or when the user clicks outside of the orange-bordered editable section, the right-most section updates in real-time.

The second section, titled QC for Quality and Control, assists in the process of verifying that the data extracted using a given template is correct. The current QC process consisted of manually opening clinical documents and comparing the values within the clinical document against a large and difficult to scroll through Excel spreadsheet of values. Therefore, for this interface the goal was to build an easy to use interface that allows for the user that is QCing the data to be able to: focus on their task at hand, speed up the process, maintain a high level of quality control or to even exceed the current level of quality control.

The resulting interface is shown in Figure 1. In this interface the clinical document is shown on the left, removing the need for the user to have to locate and open the document as in the existing process, and the extracted values are shown on the right hand side. At the top of the screen the user is shown a progress bar which is updated as they complete each clinical document, allowing for them to track their progress.

The idea of the right hand side of the screen is for the user to indicate whether or not the values extracted match what is shown within the clinical document. In the existing process, the user would simply mark a cell within an Excel spreadsheet as green or red. However, in this new interface and process the user selects “Yes” or “No” and then an assortment of options are available to them.

All of these steps, including the alerting of missing information, is updated in real time for the user and is a tremendous improvement over existing research QC processes.



Invalid Sections

WJ-III

Rapid_Pix

SLP_Woodcock_Johnson_RapidPic

RBMT

Scaled Scores

SLP_RBMT3_SM_Route_Immediate_Recall

- WJ-III ✓ -
 - Rapid Pix ✓ -
 - SLP Woodcock Johnson RapidPic ✓ -

Raw	Computed	Comment
10/27	10	/27

Value Correct: Yes No

Reason Value Incorrect: Partially Parsed

Computed True Value:

Comment True Value:
 - SLP Woodcock Johnson RapidPic Timing ✓ -

Raw	Computed	Comment
10/27	10	/27

Value Correct: Yes No
- RBMT ✓ -
 - Scaled Scores ✓ -
 - SLP RBMT3 PM Belongings Delayed Recall ✓ +
 - SLP RBMT3 SM Route Immediate Recall ✓ -

Raw	Computed	Comment
10/27	10	/27

Value Correct: Yes No

Submit

Figure 3: An example of a user indicating whether or not the data extracted matches the data in the original clinical document, and the new process of entering information pertaining to match. When a submission is attempted, if it fails, the top populates with the sections that have failed the validation process.

[Aims 1.3, 1.7, 1.8, 1.9] Following the success and feedback received from the Patient Timeline that was developed in previous quarters, a version 2 of the interface/system was discussed, mocked up, and designed which is shown in Figure 4. The design is currently in progress and has continued to undergo multiple iterations, but through the design of this mockup more productive discussions have been able to be had. In the next quarter, the development of this timeline is expected to begin such that an interface for users and clinicians to utilize will be built.



Figure 4: A design mockup of the next iteration of the Patient Timeline.

[Aims 1.3, 1.7, 1.8, 1.9] Updated and expanded on an application that uses Natural Language Processing (NLP) to analyze clinical notes and extract terms. The application allows a provider to easily search and explore notes for a patient. The tool automatically highlights clinical concepts within unstructured provider notes. In addition, the new module can be used to identify commonalities between different EHR notes.

During the last quarter, incremental improvements were made based on feedback. Additionally, a large amount of time was developed to optimizing the interface to handle large amounts of clinical data without sacrificing speed and performance. Figure #5 shows the latest interface.

Select a Patient

Types

Type to search

Statistics

Bookmarks

History

Note: 331
Last Viewed:
10 - 31-2018 16:37:23

Note: 1039
Last Viewed:
10 - 31-2018 16:37:31

Clear History

Clear All

Medications

skin

Physical Therapy

Admission Date: ["*2124-1-24*"] Discharge Date: ["*2124-3-8*"]

Service: CARDIOTHORACIC

Allergies:
Percocet

Attending ["*First Name3 (LF) 165*"]
Chief Complaint:
Increasing SOB

Major Surgical or Invasive Procedures

Show More

Audiology

PATIENT/TEST INFORMATION:
Indication: Congestive heart failure.
Height (n): 67
Weight (b): 251
BSA (m²): 2.23 m²
BP (mm Hg): 118/72
HR (bpm): 70
Status: Inpatient
Date/Time: ["*2124-1-25*"] at 10:51
Test: TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Intraop redo redo sternotomy, MVR, TVR. Evaluate valvular function, ventricular function, Aortic Atheroma.
Weight (b): 199
BP (mm Hg): 102/63
HR (bpm): 85
Status: Inpatient
Date/Time: ["*2124-2-3*"] at 10:32
Test: TEE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Post-op with significant wound drainage, hypotension, evaluate for pericardial effusion
Height (n): 67
Weight (b): 200
BSA (m²): 2.02 m²
BP (mm Hg): 95/53
HR (bpm): 75
Status: Inpatient
Date/Time: ["*2124-2-9*"] at 21:19
Test: Portable TTE (Focused views)
Doppler: Limited Doppler
Contrast: None

Show More

Art Therapy

PATIENT/TEST INFORMATION:
Indication: ? Congestive heart failure.
Weight (b): 210
BP (mm Hg): 125/65
HR (bpm): 78
Status: Inpatient
Date/Time: ["*2124-2-16*"] at 10:55
Test: Potable TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Art Therapy

Baseline artifact precludes assessment of atrial mechanism, which could be atrial fibrillation, etc. Ventricular pacing at a rate of 70 beats per minute.
Compared to the previous tracing of ["*2124-2-1*"], allowing of lead placement variation, no diagnostic change. Clinical correlation is suggested.
TRACING #2

Show More

Select a Patient

Types

head

I

Statistics

Bookmarks

History

Note: 331
Last Viewed:
10 - 31-2018 16:37:23

Note: 1039
Last Viewed:
10 - 31-2018 16:37:31

Clear History

Clear All

Medications

head

heart

heart block

heart failure

ct head

ct head w/o contrast

head ct

heart surgery

healing

Physical Therapy

Admission Date: ["*2124-3-8*"]

Service: CARDIOTHORACIC

Allergies:
Percocet

Attending ["*First Name3 (LF) 165*"]
Chief Complaint:
Increasing SOB

Major Surgical or Invasive Procedures

Show More

Audiology

PATIENT/TEST INFORMATION:
Indication: Congestive heart failure.
Height (n): 67
Weight (b): 251
BSA (m²): 2.23 m²
BP (mm Hg): 118/72
HR (bpm): 70
Status: Inpatient
Date/Time: ["*2124-1-25*"] at 10:51
Test: TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Intraop redo redo sternotomy, MVR, TVR. Evaluate valvular function, ventricular function, Aortic Atheroma.
Weight (b): 199
BP (mm Hg): 102/63
HR (bpm): 85
Status: Inpatient
Date/Time: ["*2124-2-3*"] at 10:32
Test: TEE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Post-op with significant wound drainage, hypotension, evaluate for pericardial effusion
Height (n): 67
Weight (b): 200
BSA (m²): 2.02 m²
BP (mm Hg): 95/53
HR (bpm): 75
Status: Inpatient
Date/Time: ["*2124-2-9*"] at 21:19
Test: Portable TTE (Focused views)
Doppler: Limited Doppler
Contrast: None

Show More

Art Therapy

PATIENT/TEST INFORMATION:
Indication: ? Congestive heart failure.
Weight (b): 210
BP (mm Hg): 125/65
HR (bpm): 78
Status: Inpatient
Date/Time: ["*2124-2-16*"] at 10:55
Test: Potable TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Art Therapy

Baseline artifact precludes assessment of atrial mechanism, which could be atrial fibrillation, etc. Ventricular pacing at a rate of 70 beats per minute.
Compared to the previous tracing of ["*2124-2-1*"], allowing of lead placement variation, no diagnostic change. Clinical correlation is suggested.
TRACING #2

Show More

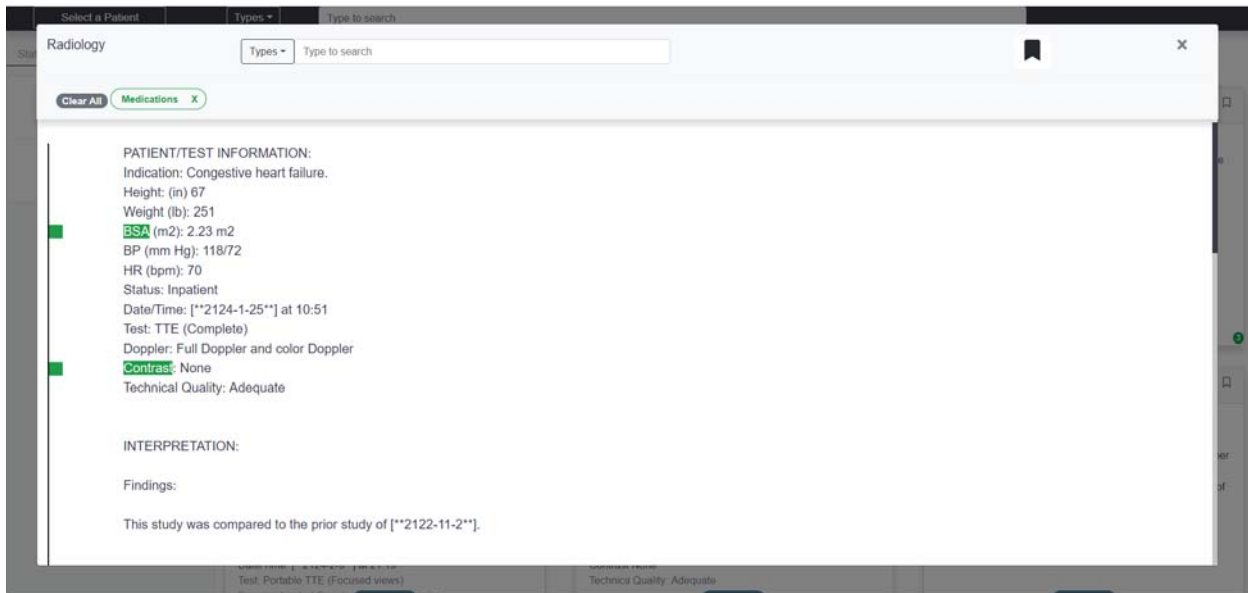


Figure 5 (a, b, & c): Screenshots of the module under development to analyze / compare clinical notes. This application that allows providers to search keywords and clinical concepts within clinical notes.

What was accomplished under these goals?

Major Task 1: Design and development of clinical visual analytics framework		
Task	Comments	% Completed
Subtask 1: New software developer and research assistant will be hired.	After 6 months recruiting for a software developer capable of working within the DoD and willing to accept the low salary that was approved under this award, we were able to hire Elizabeth Jimenez to support the project.	100%
Subtask 2: Team will compare and evaluate existing data visualization libraries and software tools.	This task has been completed and the PI (as an international leader in data visualization) continues to assess new software tools and programming languages to make sure	100%
Subtask 3: Design a modularized visual analytics framework.	The PI continues to develop a state-of-the-art visualization framework to support the analysis, integration, and visualization of clinical variables of patients diagnosed with mTBI. During the last quarter new functionalities have been implemented including: (1) a patient timeline to allow providers to explore the clinical variables of a patient over time (2) machine learning algorithm to learn and model how users interact with the visualization systems (3) algorithm to summarize and visually represent clinical trajectories (4) visualization technique that uses a Sankey diagram and a Sankey-like diagram to	85%

visualize clinical trajectories.

Subtask 4: Obtain previously collected multi-modal clinical data for 300 mTBI/PTSD subjects and load data into the visualization system. The data will be obtained from the NICOE at no cost to this project.	Data has been obtained. IRB protocol continues to be active. Submitted continuing review and received approval until Dec 31 2017.	100%
Subtask 5: Perform preliminary validation of different visualization techniques	This task has been completed and the PI (as an international leader in data visualization) continues to validate, incorporate, and assess different visualization techniques that can be included within the framework.	100%
Subtask 6: Prepare and submit research progress report	All quarterly progress reports and annual reports have been submitted.	100%
Subtask 7: Update application based on preliminary validation with clinicians and researchers.	The team continues to work closely with several providers that are giving feedback, comments, and suggestions about the usability of the system, the effectiveness of different visualization techniques, and the specific tasks that are more relevant for the different clinical disciplines.	80%
Subtask 8: Update system based on usability study and suggestions received from users.	The system continues to improve as the team receives comments and suggestions from users. We are still collecting data from a usability study to be able to update the system based on the feedback of many users within a controlled environment.	80%
Subtask 9: Share initial framework with different federal organizations including WRNMMC, MHS, DICOE, and NIH. Share final software application with the US Army Medical Research and Material Command and other federal organizations.	The framework, novel visualization design, and ideas have been shared with a number of researchers including WRNMMC, NICOE, USU, San Antonio, Madigan, DHA, and other DoD organizations. As more features, enhancements, and modules are added to the tool, we will continue to disseminate the tool following appropriate channels.	70%
Subtask 10: Prepare final report	All quarterly progress reports and annual reports have been submitted.	70%

Describe the Regulatory Protocol and Activity Status (if applicable).

The primary objective of this project is to design and develop a visual analytic framework that combines multiple clinical measurements and allows the exploration of large collections of clinical data. To develop the software tool no clinical data is needed. Only when validating the effectiveness of the application is when de-identified clinical data will be needed. To obtain data, an existing retrospective IRB protocol is being used.

(a) Human Use Regulatory Protocols

TOTAL PROTOCOL(S): Integration of Image and Clinical Measurements of TBI Patients using Machine Learning Techniques

PROTOCOL(S): WRNMMC IRB #374953

Submitted to and Approved by:

- Submitted and reviewed by Walter Reed National Military Medical Center, Department of Research Programs
- Submitted and reviewed the USAMRMC regulatory office
- Active since 2012

Status: an existing retrospective IRB protocol (that has been reviewed by USAMRMC multiple times) will be used. During the fall of 2015 an IRB amendment will be submitted to explicitly include visualization techniques within the list of techniques that will be used to analyze and explore the data. Submitted continuing review and approved until Dec 2018. A continuing review has been submitted to extend our research effort until December 2019.

(b) Use of Human Cadavers for Research Development Test & Evaluation (RDT&E), Education or Training

Not applicable.

(c) Animal Use Regulatory Protocol

Not applicable

What do you plan to do during the next reporting period to accomplish the goals and objectives?

During the next reporting period, further progress will be made on the visual analytics framework by continuing to research new and novel techniques that can be used to illustrate data. In addition, the new data explorer and QC system will be tested with a lot of data. Finally, some of the tools developed by the group might be tested within the DoD enterprise system.

SECTION 2 – Products:

Nothing to report

SECTION 3 - Participants & Other Collaborating Organizations

What individuals have worked on the project?

Name: Jesus Caban, PhD

Project Role: PI

Nearest person month worked: 1

Contribution to Project: Dr. Caban has organized meetings, tracked progress of the project, and evaluated various visualization techniques for exploring large clinical data.

Name: Filip Dabek

Project Role: N/A (Data Scientist for the National Intrepid Center of Excellence)

Nearest person month worked: 1

Contribution to Project: Mr. Dabek has begun implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

Name: Elizabeth Jimenez

Project Role: Jr. Software Developer

Nearest person month worked: 3

Contribution to Project: Ms. Jimenez worked on implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

SECTION 4 –Changes/Problems:

1. Actual Problems or delays and actions to resolve them

Nothing to Report

2. Anticipated Problems/Issues

Nothing to Report

SECTION 5 – Special Reporting Requirements: Quad Chart

(a) Award Number: W81XWH-15-2-0016

(b) Log Number: GRANT11367272

(c) Project Title: An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data

(d) Principal Investigator Name: Dr. Jesus Caban

(e) Principal Investigator Organization and Address:

National Intrepid Center of Excellence

Walter Reed National Military Medical Center

8901 Wisconsin Ave

Bethesda, Maryland 20889

(f) Principal Investigator Phone and Email:

jesus.j.caban.civ@mail.mil

(301) 319-3759

(g) Report Date: 30 Oct 2018

(h) Report Period: 16 July 2018 - 15 Oct 2018

SECTION 1 -- Accomplishments:

What were the major goals of the project?

The first major goal of this project is to design a visual analytic framework that combines multiple clinical measurements and allows the exploration of large collections of clinical data. In addition, the second major goal is to validate the effectiveness and usability of different visualization techniques for exploring large collections of clinical variables with complex associations.

In this quarter, the PI focused on:

1. [[Aims 1.3, 1.7, 1.8, 1.9] The TBI Data Explorer and TBI Parsers projects are being developed in tandem, with a goal of transforming the way researchers collect, store, parse, and check data. The new TBI Parser process consists of a custom system, developed in Python, that receives a JSON template file containing the structure of the clinical document to be parsed. To assist in the process of creating and debugging template files, the TBI Data Explorer Interface was developed with two main sections: Parser and QC. The interface was developed using VueJS and Bootstrap.

The TBI Data Explorer Parser tab will help with the development of final version of these templates. The first version/base template is built using only a few files of sample data, ensuring that the template file captures the overall structure of the clinical document. After the initial template has been built, the Parser tab can be used to quickly view how changes will affect the data and to optimize the various regular expressions necessary. Therefore, upon loading the Parser tab the user chooses the template that they are working on as well as any additional information and then clicks “Load” for the system to generate the interface for the specific template. In Figure 1, the left section shows the user the fields that have been identified in the template; it also keeps track of which fields have been visited by fading their color and which field is currently being viewed by changing its font color. The middle section contains the JSON that belongs to the currently selected field on the left-most section. The right section shows all of the possible raw data values that were extracted from the clinical documents, as well as the resulting “computed” and “comment” values that are extracted based on their field configuration. On click of Run, or when the user clicks outside of the orange-bordered editable section, the right-most section updates in real-time.

The second section, titled QC for Quality and Control, assists in the process of verifying that the data extracted using a given template is correct. The current QC process consisted of manually opening clinical documents and comparing the values within the clinical document against a large and difficult to scroll through Excel spreadsheet of values. Therefore, for this interface the goal was to build an easy to use interface that allows for the user that is QCing the data to be able to: focus on their task at hand, speed up the process, maintain a high level of quality control or to even exceed the current level of quality control.

The resulting interface is shown in Figure 1. In this interface the clinical document is shown on the left, removing the need for the user to have to locate and open the document as in the existing process, and the extracted values are shown on the right hand side. At the top of the screen the user is shown a progress bar which is updated as they complete each clinical document, allowing for them to track their progress.

The idea of the right hand side of the screen is for the user to indicate whether or not the values extracted match what is shown within the clinical document. In the existing process, the user would simply mark a cell within an Excel spreadsheet as green or red. However, in this new interface and process the user selects “Yes” or “No” and then an assortment of options are available to them.

Figure 2 shows an example of a user marking a data value as being incorrect, another as correct, and finally being alerted that they had not marked anything for the final data value. For the first data value in Figure 3 it can be seen how the interface reacts to the user and allows them to select a reason for why the data did not match, as well as fields to enter the correct values of the data elements.

All of these steps, including the alerting of missing information, is updated in real time for the user and is a tremendous improvement over existing research QC processes.

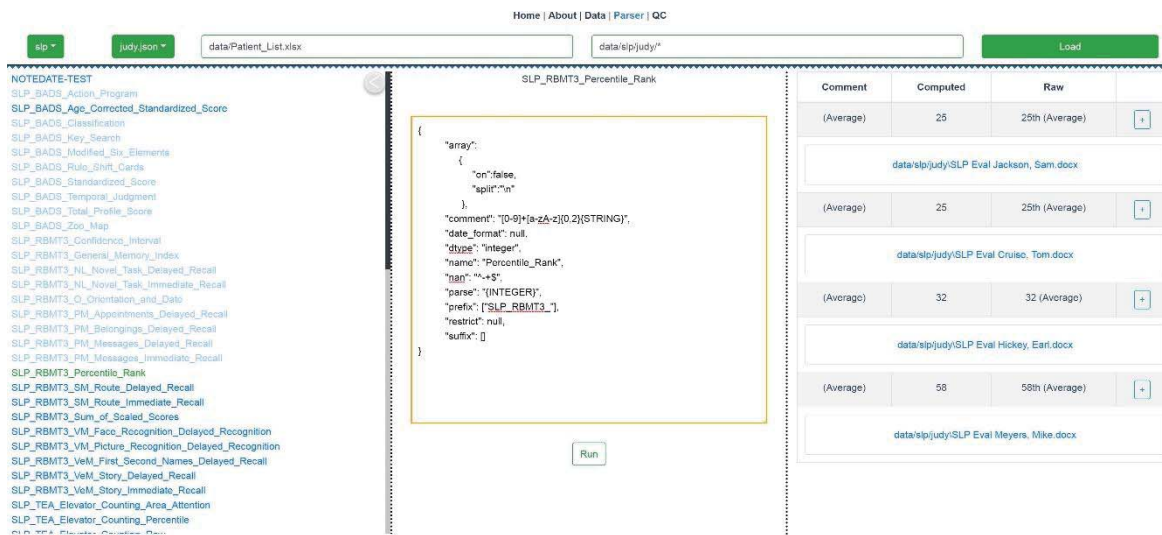


Figure 1: The parser section of the NICOE Data Explorer interface.

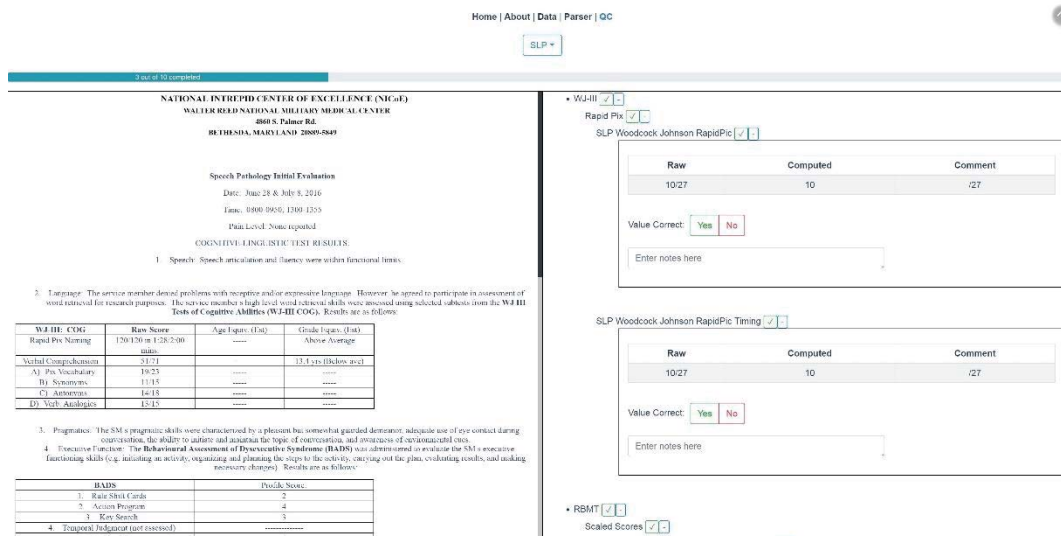


Figure 2: The QC (quality control) section of the TBI Data Explorer interface.

Invalid Sections

WJ-III

Rapid_Pix

SLP_Woodcock_Johnson_RapidPic

RBMT

Scaled Scores

SLP_RBMT3_SM_Route_Immediate_Recall

- WJ-III ✓ -
 - Rapid Pix ✓ -
 - SLP Woodcock Johnson RapidPic ✓ -

Raw	Computed	Comment
10/27	10	/27

Value Correct: Yes No

Reason Value Incorrect: Partially Parsed

Computed True Value: Enter what computed value should be

Comment True Value: Enter what comment value should be

Enter notes here
 - SLP Woodcock Johnson RapidPic Timing ✓ -

Raw	Computed	Comment
10/27	10	/27

Value Correct: Yes No

Enter notes here
- RBMT ✓ -
 - Scaled Scores ✓ -
 - SLP RBMT3 PM Belongings Delayed Recall ✓ +
 - SLP RBMT3 SM Route Immediate Recall ✓ -

Raw	Computed	Comment
10/27	10	/27

Value Correct: Yes No

Enter notes here

Submit

Figure 3: An example of a user indicating whether or not the data extracted matches the data in the original clinical document, and the new process of entering information pertaining to match. When a submission is attempted, if it fails, the top populates with the sections that have failed the validation process.

[Aims 1.3, 1.7, 1.8, 1.9] Following the success and feedback received from the Patient Timeline that was developed in previous quarters, a version 2 of the interface/system was discussed, mocked up, and designed which is shown in Figure 4. The design is currently in progress and has continued to undergo multiple iterations, but through the design of this mockup more productive discussions have been able to be had. In the next quarter, the development of this timeline is expected to begin such that an interface for users and clinicians to utilize will be built.

Select a Patient

Types

Type to search

Statistics

Bookmarks

History

Note: 331
Last Viewed:
10 - 31-2018 16:37:23

Note: 1039
Last Viewed:
10 - 31-2018 16:37:31

Clear History

Clear All

Medications

skin

Physical Therapy

Admission Date: ["*2124-1-24*"] Discharge Date: ["*2124-3-8*"]

Service: CARDIOTHORACIC

Allergies:
Percocet

Attending ["*First Name3 (LF) 165*"]

Chief Complaint:
Increasing SOB

Show More

Major Surgical or Invasive Procedures

Audiology

PATIENT/TEST INFORMATION:
Indication: Congestive heart failure.
Height (n): 67
Weight (b): 251
BSA (m²): 2.23 m²
BP (mm Hg): 118/72
HR (bpm): 70
Status: Inpatient
Date/Time: ["*2124-1-25*"] at 10:51
Test: TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Intraop redo redo sternotomy, MVR, TVR. Evaluate valvular function, ventricular function, Aortic Atheroma.
Weight (b): 199
BP (mm Hg): 102/63
HR (bpm): 85
Status: Inpatient
Date/Time: ["*2124-2-3*"] at 10:32
Test: TEE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Post-op with significant wound drainage, hypotension, evaluate for pericardial effusion
Height (n): 67
Weight (b): 200
BSA (m²): 2.02 m²
BP (mm Hg): 95/53
HR (bpm): 75
Status: Inpatient
Date/Time: ["*2124-2-9*"] at 21:19
Test: Portable TTE (Focused views)
Doppler: Limited Doppler
Contrast: None

Show More

Art Therapy

PATIENT/TEST INFORMATION:
Indication: ? Congestive heart failure.
Weight (b): 210
BP (mm Hg): 125/65
HR (bpm): 78
Status: Inpatient
Date/Time: ["*2124-2-16*"] at 10:55
Test: Potable TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Art Therapy

Baseline artifact precludes assessment of atrial mechanism, which could be atrial fibrillation, etc. Ventricular pacing at a rate of 70 beats per minute.
Compared to the previous tracing of ["*2124-2-1*"], allowing of lead placement variation, no diagnostic change. Clinical correlation is suggested.
TRACING #2

Show More

Select a Patient

Types

head

I

Statistics

Bookmarks

History

Note: 331
Last Viewed:
10 - 31-2018 16:37:23

Note: 1039
Last Viewed:
10 - 31-2018 16:37:31

Clear History

Clear All

Medications

head

heart

heart block

heart failure

ct head

ct head w/o contrast

head ct

heart surgery

healing

Physical Therapy

Admission Date: ["*2124-3-8*"]

Service: CARDIOTHORACIC

Allergies:
Percocet

Attending ["*First Name3 (LF) 165*"]

Chief Complaint:
Increasing SOB

Show More

Major Surgical or Invasive Procedures

Audiology

PATIENT/TEST INFORMATION:
Indication: Congestive heart failure.
Height (n): 67
Weight (b): 251
BSA (m²): 2.23 m²
BP (mm Hg): 118/72
HR (bpm): 70
Status: Inpatient
Date/Time: ["*2124-1-25*"] at 10:51
Test: TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Intraop redo redo sternotomy, MVR, TVR. Evaluate valvular function, ventricular function, Aortic Atheroma.
Weight (b): 199
BP (mm Hg): 102/63
HR (bpm): 85
Status: Inpatient
Date/Time: ["*2124-2-3*"] at 10:32
Test: TEE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Neurology

PATIENT/TEST INFORMATION:
Indication: Post-op with significant wound drainage, hypotension, evaluate for pericardial effusion
Height (n): 67
Weight (b): 200
BSA (m²): 2.02 m²
BP (mm Hg): 95/53
HR (bpm): 75
Status: Inpatient
Date/Time: ["*2124-2-9*"] at 21:19
Test: Portable TTE (Focused views)
Doppler: Limited Doppler
Contrast: None

Show More

Art Therapy

PATIENT/TEST INFORMATION:
Indication: ? Congestive heart failure.
Weight (b): 210
BP (mm Hg): 125/65
HR (bpm): 78
Status: Inpatient
Date/Time: ["*2124-2-16*"] at 10:55
Test: Potable TTE (Complete)
Doppler: Full Doppler and color Doppler
Contrast: None
Technical Quality: Adequate

Show More

Art Therapy

Baseline artifact precludes assessment of atrial mechanism, which could be atrial fibrillation, etc. Ventricular pacing at a rate of 70 beats per minute.
Compared to the previous tracing of ["*2124-2-1*"], allowing of lead placement variation, no diagnostic change. Clinical correlation is suggested.
TRACING #2

Show More

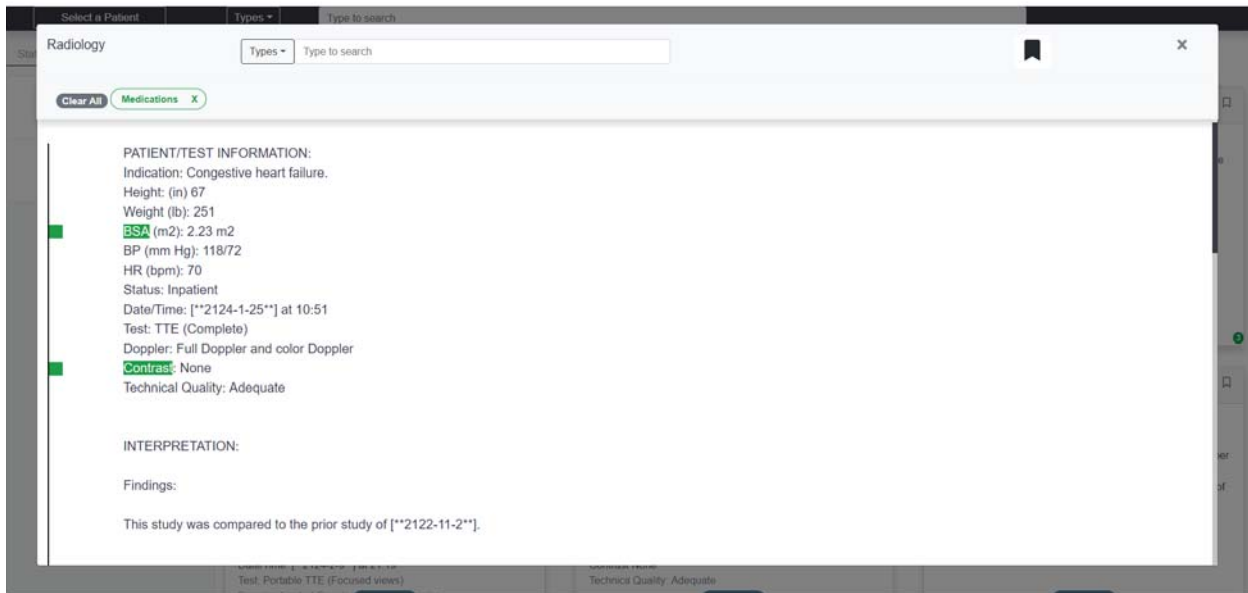


Figure 5 (a, b, & c): Screenshots of the module under development to analyze / compare clinical notes. This application that allows providers to search keywords and clinical concepts within clinical notes.

What was accomplished under these goals?

Major Task 1: Design and development of clinical visual analytics framework		
Task	Comments	% Completed
Subtask 1: New software developer and research assistant will be hired.	After 6 months recruiting for a software developer capable of working within the DoD and willing to accept the low salary that was approved under this award, we were able to hire Elizabeth Jimenez to support the project.	100%
Subtask 2: Team will compare and evaluate existing data visualization libraries and software tools.	This task has been completed and the PI (as an international leader in data visualization) continues to assess new software tools and programming languages to make sure	100%
Subtask 3: Design a modularized visual analytics framework.	The PI continues to develop a state-of-the-art visualization framework to support the analysis, integration, and visualization of clinical variables of patients diagnosed with mTBI. During the last quarter new functionalities have been implemented including: (1) a patient timeline to allow providers to explore the clinical variables of a patient over time (2) machine learning algorithm to learn and model how users interact with the visualization systems (3) algorithm to summarize and visually represent clinical trajectories (4) visualization technique that uses a Sankey diagram and a Sankey-like diagram to	85%

visualize clinical trajectories.

Subtask 4: Obtain previously collected multi-modal clinical data for 300 mTBI/PTSD subjects and load data into the visualization system. The data will be obtained from the NICOE at no cost to this project.	Data has been obtained. IRB protocol continues to be active. Submitted continuing review and received approval until Dec 31 2017.	100%
Subtask 5: Perform preliminary validation of different visualization techniques	This task has been completed and the PI (as an international leader in data visualization) continues to validate, incorporate, and assess different visualization techniques that can be included within the framework.	100%
Subtask 6: Prepare and submit research progress report	All quarterly progress reports and annual reports have been submitted.	100%
Subtask 7: Update application based on preliminary validation with clinicians and researchers.	The team continues to work closely with several providers that are giving feedback, comments, and suggestions about the usability of the system, the effectiveness of different visualization techniques, and the specific tasks that are more relevant for the different clinical disciplines.	80%
Subtask 8: Update system based on usability study and suggestions received from users.	The system continues to improve as the team receives comments and suggestions from users. We are still collecting data from a usability study to be able to update the system based on the feedback of many users within a controlled environment.	80%
Subtask 9: Share initial framework with different federal organizations including WRNMMC, MHS, DICOE, and NIH. Share final software application with the US Army Medical Research and Material Command and other federal organizations.	The framework, novel visualization design, and ideas have been shared with a number of researchers including WRNMMC, NICOE, USU, San Antonio, Madigan, DHA, and other DoD organizations. As more features, enhancements, and modules are added to the tool, we will continue to disseminate the tool following appropriate channels.	70%
Subtask 10: Prepare final report	All quarterly progress reports and annual reports have been submitted.	70%

Describe the Regulatory Protocol and Activity Status (if applicable).

The primary objective of this project is to design and develop a visual analytic framework that combines multiple clinical measurements and allows the exploration of large collections of clinical data. To develop the software tool no clinical data is needed. Only when validating the effectiveness of the application is when de-identified clinical data will be needed. To obtain data, an existing retrospective IRB protocol is being used.

(a) Human Use Regulatory Protocols

TOTAL PROTOCOL(S): Integration of Image and Clinical Measurements of TBI Patients using Machine Learning Techniques

PROTOCOL(S): WRNMMC IRB #374953

Submitted to and Approved by:

- Submitted and reviewed by Walter Reed National Military Medical Center, Department of Research Programs
- Submitted and reviewed the USAMRMC regulatory office
- Active since 2012

Status: an existing retrospective IRB protocol (that has been reviewed by USAMRMC multiple times) will be used. During the fall of 2015 an IRB amendment will be submitted to explicitly include visualization techniques within the list of techniques that will be used to analyze and explore the data. Submitted continuing review and approved until Dec 2018. A continuing review has been submitted to extend our research effort until December 2019.

(b) Use of Human Cadavers for Research Development Test & Evaluation (RDT&E), Education or Training

Not applicable.

(c) Animal Use Regulatory Protocol

Not applicable

What do you plan to do during the next reporting period to accomplish the goals and objectives?

During the next reporting period, further progress will be made on the visual analytics framework by continuing to research new and novel techniques that can be used to illustrate data. In addition, the new data explorer and QC system will be tested with a lot of data. Finally, some of the tools developed by the group might be tested within the DoD enterprise system.

SECTION 2 – Products:

Nothing to report

SECTION 3 - Participants & Other Collaborating Organizations

What individuals have worked on the project?

Name: Jesus Caban, PhD

Project Role: PI

Nearest person month worked: 1

Contribution to Project: Dr. Caban has organized meetings, tracked progress of the project, and evaluated various visualization techniques for exploring large clinical data.

Name: Filip Dabek

Project Role: N/A (Data Scientist for the National Intrepid Center of Excellence)

Nearest person month worked: 1

Contribution to Project: Mr. Dabek has begun implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

Name: Elizabeth Jimenez

Project Role: Jr. Software Developer

Nearest person month worked: 3

Contribution to Project: Ms. Jimenez worked on implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

SECTION 4 –Changes/Problems:

1. Actual Problems or delays and actions to resolve them

Nothing to Report

2. Anticipated Problems/Issues

Nothing to Report

SECTION 5 – Special Reporting Requirements: Quad Chart

Quad Chart