FRIDAY!!

Visualization Design Sprints for Online and On-Campus Courses 🡪 Talking about how to teach viz to students.

Dispersion vs Disparity: Hiding Variability Can Encourage Stereotyping When Visualizing Social Outcomes 🡪 These results demonstrate that data visualizations about social inequity can be misinterpreted in harmful ways and lead to stereotyping. Design choices can influence these biases: Hiding variability tends to increase stereotyping while emphasizing variability reduces it.

D-BIAS: A Causality-Based Human-in-the-Loop System for Tackling Algorithmic Bias 🡪 We found that D-BIAS helps reduce bias significantly compared to the baseline debiasing approach across different fairness metrics while incurring little data distortion and a small loss in utility. Moreover, our human-in-the-loop based approach significantly outperforms an automated approach on trust, interpretability and accountability. Basically an algorithm that creates a viz model that seeks to reduce bias.

MONDAY!!

A Predictive Visual Analytics System for Studying Neurodegenerative Disease Based on DTI Fiber Tracts 🡪 Diffusion tensor imaging (DTI) has been used to study the effects of neurodegenerative diseases on neural pathways, which may lead to more reliable and early diagnosis of these diseases as well as a better understanding of how they affect the brain. We introduce a predictive visual analytics system for studying patient groups based on their labeled DTI fiber tract data and corresponding statistics.

Reflections on Visualization Research Projects in the Manufacturing Industry 🡪 The rise of Industry 4.0 and cyber-physical systems has led to an abundance of large amounts of data, particularly in the manufacturing industry. Visualization and visual analytics play essential roles in harnessing this data. They have already been acknowledged as being among the key enabling technologies in the fourth industrial revolution.

**[most relevant 2nd]**Visual Comparison of Language Model Adaptation 🡪 Our comparative visualizations show the differences in the adapted embedding vectors and prediction outcomes for diverse human-interpretable concepts (e.g., person names, human qualities).

**[relatively easier to read]**Sporthesia: Augmenting Sports Videos Using Natural Language 🡪 Sporthesia takes raw video footage and commentary text of racket-based sports as input, and outputs an augmented video. To achieve this, three key steps are taken: 1) detecting the visualizable entities in the text, 2) mapping the entities to visualizations, and 3) scheduling the visualizations to play with the raw video

**[Kinda relevant use case]**ConceptExplainer: Interactive Explanation for Deep Neural Networks from a Concept Perspective: To this end, we design, develop, and validate CONCEPTEXPLAINER, a visual analytics system that enables non-expert users to interactively probe and explore the concept space to explain model behavior at the instance/class/global level.

Seeing What You Believe or Believing What You See? Belief Biases Correlation Estimation: hrough two crowdsourced experiments, we demonstrate that supposedly objective assessments of the strength of a correlational relationship can be influenced by how strongly a viewer believes in the existence of that relationship. While data visualizations are typically thought to present objective truths to the viewer, these results suggest that existing personal beliefs can bias even objective statistical values people extract from data.

**[Most relevant so far]**Cartolabe: A Web-Based Scalable Visualization of Large Document Collections: This visualization module aims at enabling users with no expertise in visualization and data analysis to get an overview of their corpus, and to interact with it: exploring, querying, filtering, panning and zooming on regions of semantic interest.

**[Boring]**FlowNL: Asking the Flow Data in Natural Languages:

**[Interesting, consult later]**Scientometric Analysis of Interdisciplinary Collaboration and Gender Trends in 30 Years of IEEE VIS Publications: Our primary goal in this work is to call the visualization community’s attention to the critical topics of collaboration, diversity, and gender. Our research offers critical insights through the lens of diversity and gender to help accelerate progress towards a more diverse and representative research community.

TUESDAY-THURSDAY!!

**[Consult this primarily on how to solve a problem]**ClinicalPath: a Visualization tool to Improve the Evaluation of Electronic Health Records in Clinical Decision-Making: In this paper, we propose ClinicalPath, a visualization tool for users to track a patient’s clinical path through a series of tests and data, which can aid in treatments and diagnoses. Our proposal is focused on patient’s data analysis, presenting the test results and clinical history longitudinally. Both the visualization design and the system functionality were developed in close collaboration with experts in the medical domain to ensure a right fit of the technical solutions and the real needs of the professionals. We validated the proposed visualization based on case studies and user assessments through tasks based on the physician’s daily activities. Our results show that our proposed system improves the physicians’ experience in decision-making tasks, made with more confidence and better usage of the physicians’ time, allowing them to take other needed care for the patients.

**[Kinda relevant but too long]**SD^2: Slicing and Dicing Scholarly Data for Interactive Evaluation of Academic Performance: In this paper, we present an interactive visualization framework, SD2 , to enable flexible data partition and composition to support various analysis requirements within a single system. SD2 features the hierarchical histogram, a novel visual representation for flexibly slicing and dicing the data, allowing different aspects of scholarly performance to be studied and compared. We also leverage the state-of-the-art set visualization technique to select individual researchers or combine multiple scholars for comprehensive visual comparison. We conduct multiple rounds of expert evaluation to study the effectiveness and usability of SD 2 and revise the design and system implementation accordingly. The effectiveness of SD2 is demonstrated via multiple usage scenarios with each aiming to answer a specific, commonly raised question.

polyphony: an Interactive Transfer Learning Framework for Single-Cell Data Analysis

**[Good paper, easy to understand, fun to read]**PMU Tracker: A Visualization Platform for Epicentric Event Propagation Analysis in the Power Grid: As a result, we have developed PMU Tracker, an event localization tool that supports power grid operators in visually analyzing and identifying power grid events and tracking their propagation through the power grid’s network.

**[Not very fun to read]**CosmoVis: An Interactive Visual Analysis Tool for Exploring Hydrodynamic Cosmological Simulations: We introduce CosmoVis, an open source web-based visualization tool for the interactive analysis of massive hydrodynamic cosmological simulation data. CosmoVis was designed in close collaboration with astrophysicists to enable researchers and citizen scientists to share and explore these datasets, and to use them to investigate a range of scientific questions.

**[Ironically, not very funny JK not very enjoyable to read]**DeHumor: Visual Analytics for Decomposing Humor: DeHumor , a visual analytical system for analyzing humorous behaviors in public speaking. To intuitively reveal the building blocks of each concrete example, DeHumor decomposes each humorous video into multimodal features and provides inline annotations of them on the video script.

**[Meta Analysis, we could do something similar?]**Thirty-Two Years of IEEE VIS: Authors, Fields of Study and Citations: Driven by this movement, we investigated (1) who contributed to VIS, and (2) where VIS stands in the scientific world. We examined the authors and fields of study of 3,240 VIS publications in the past 32 years based on data collected from OpenAlex and IEEE Xplore, among other sources.

**[Boring to read]**GenoREC: A Recommendation System for Interactive Genomics Data Visualization: GenoREC enables genomics analysts to select effective visualizations based on a description of their data and analysis tasks. Here, we present the recommendation model that uses a knowledge-based method for choosing appropriate visualizations and a web application that enables analysts to input their requirements, explore recommended visualizations, and export them for their usage.

**[Good paper to read]**COVID-19 EnsembleVis: Visual Analysis of County-level Ensemble Forecast Models: In this paper, we built COVID-19 EnsembleVis, a web-based interactive visual interface that allows the assessment of the errors of ensembles and individual models by enabling users to effortlessly navigate through and compare the outputs of models considering their space and time dimensions.

**[Interesting]**Real-Time Visual Analysis of High-Volume Social Media Post: In this work, we present an interactive system that enables the visual analysis of streaming social media data on a large scale in real-time. We propose an efficient and explainable dynamic clustering algorithm that powers a continuously updated visualization of the current thematic landscape as well as detailed visual summaries of specific topics of interest.

**FRIDAY!!!**

**[SSH paper in viz but not very relevant]**Topic modeling of SSH logs using latent dirichlet allocation for the application in cyber security: We discovered that brute-force attacks and legitimate users could easily be separated through the vocabulary found in their logs. We also discovered that due to the overlapping topic features of the LDA model, cyber-attacks can be further classified and differentiated according to their topical weights.

**[Come back to this and check the papers it is referencing]**NViZ: An Interactive Visualization of Network Security Systems Logs: We describe NViZ a graphical visualization tool developed to visualize log data generated by network security devices and services like firewall, intrusion detection system and domain name system. It can generate interactive visualization graphs identifying popular websites visited, active users, DNS query patterns, IDS alert types and their distribution and also peer network connections of infected and non-infected machines.

**[Definitely related, keep coming back to this, note Kippo-graph tool mentioned in the paper, note the background and related work of this paper]***Analysis and Visualization of SSH Attacks Using Honeypots*: In this paper we present the results of a research honeypot’s operation, which undertook the role of a web trap for attackers who target the SSH service in order to gain illegal server access. The fake system has remained online and fully operational during a course of several consequent months, capturing attacks and logging all malicious activity. Lastly we present a visualization tool aimed at helping security researchers during the analysis and conclusions drawing phases, for use with the same SSH honeypot implementation software as outlined in this work.

Above paper cites:

* **[]** *Large-scale Network Monitoring for Visual Analysis of Attacks*:Within the system, the monitored network is mapped to a TreeMap visualization, the attackers are arranged at the borders and linked using splines parameterized with prefix information. In a series of case studies, we demonstrate how the tool can be used to judge the relevance of alerts, to reveal massive distributed attacks, and to analyze service usage within a network.
* ***[MOST RELEVANT, WILL NEED TO CITE THIS]****VIAssist: Visual Analytics for Cyber Defense:* Analysis of voluminous computer network data has become a common practice for cyber defense, but few tools provide adequate support for cyber-infrastructure defenders’ workflow, visual exploration, IP geo-location, scalability, collaboration, or reporting. The state-of-the-art in visual analysis tools for cyber defense is typically no more than spreadsheets and primitive charting. While familiar to users, this approach ignores the human perceptual ability to identify novel patterns and anomalies when data is presented graphically. This paper reports on a visual analytics systems, VIAssist, being developed for cyber-infrastructure protection that helps cyber defenders better understand the massive, multi-dimensional datasets to protect our nation’s critical infrastructure.

Above paper is cited by:

* **[Get back to this later]**Tracking smart grid hackers

**[Definitely related, keep coming back to this**]Visualization of System Log Files for Post-incident Analysis and Response: Post-incident analysis of a security event is a complex task due to the volume of data that must be assessed, often within tight temporal constraints. System software, such as operating systems and applications, provide a range of opportunities to record data in log files about interactions with the computer that may provide evidence during an investigation. Data visualization can be used to aid data set interpretation and improve the ability of the analyst to make sense of information. This paper posits a novel methodology that visualizes data from a range of log files to aid the investigation process. In order to demonstrate the applicability of the approach, a case study of identification and analysis of attacks is presented.

Visualizations are accompanied by case studies and evaluations by domain experts.

I have an idea for another paper:

What if we make a visualization that shows what it takes to get into every top rated US conference?

NViZ: An Interactive Visualization of Network Security Systems Logs(*cite and cited by are not as important*)

Analysis and Visualization of SSH Attacks Using Honeypots**[DONE!!!!!!!](possible cite)**

* Cites:
  + Large-scale Network Monitoring for Visual Analysis of Attacks**(possible cite)**
    - Cites:
      * Closing-the-Loop in NVisionIP: Integrating Discovery and Search in Security Visualizations(**THIS IS USEFUL IF I DECIDE INCORPORATING SYMBOLIC PATTERN MATCHING TO MY DATA**)
    - Cited by:
      * [**CURRENTLY HERE**]
  + VIAssist: Visual Analytics for Cyber Defense(*definite cite*)
    - Cites:
      * NVisionIP: NetFlow visualizations of system state for security situational awareness
    - Cited by:
      * Situ: Identifying and explaining suspicious behavior in networks(**VERY IMPORTANT**)(*definite cite*)
        + Cites:

Focusing on context in network traffic analysis(*may be useful*)

* + - * + Cited by:

VITALflow: Visual Interactive Traffic Analysis with NetFlow(*may be useful*)

VKE: a Visual Analytics Tool for CyberSecurity Data(**NEED TO ACCESS**)

Network security data visualization(**NEED TO ACCESS**)

* + - * Bubblenet: A cyber security dashboard for visualizing patterns(**VERY IMPORTANT**)(*definite cite*)
        + Cites:

Visualizing Real-Time Network Resource Usage(*may be useful*)

7 Key Challenges for Visualization in Cyber Network Defense(*may be useful to respond to, we could set up a framework*)

Fast detection and visualization of network attacks on parallel coordinates(*may be useful*)

IMap: Visualizing network activity over internet maps.(*may be useful*)

Visualizing Cyber Attacks using IP Matrix(*may be useful*)

* + - * + Cited by:

NetCapVis: Web-based Progressive Visual Analytics for Network Packet Captures(*may be useful)*

Data Visualization in Cybersecurity(*may be useful)*

* + - * A visual analytics approach to detecting server redirections and data exfiltration(may be useful later)
* Cited by:
  + Tracking smart grid hackers
    - Cites:
    - Cited by:

Visualization of System Log Files for Post-incident Analysis and Response

* Cites
* Cited by

NetVis: A Network Security Management Visualization Tool Based On Treemap(this could come in handy when looking at what to write about in the related section of my paper)

PortVis: A Tool for Port-Based Detection of Security Events

* Cites
  + Visualizing Network Data(**seems to be the paper that everyone has cited, published March 1995**)
  + The Spinning Cube of Potential Doom(can be referred to, need to check with Prof whether this is too old)
* Cited by
  + <https://dl.acm.org/doi/10.1145/1029208.1029220>
  + Cyberattack Graph Modeling for Visual Analytics(may refer to this)
  + Visual Analytics for Cyber Security Domain: State-of-the-Art and Challenges(**CITE THIS!!!**)
  + Cite [1] and [2] and [6] from this paper: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7883232> (don’t cite this paper though!!)
  + “A survey of security visualization for computer network logs”): (**Fun paper to read, not sure if I will cite**)

Look at what this paper cites: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4126216>

Possible cite due to its recency and closeness to what we are doing:

<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9931797>

Check from citation 1 next time: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1532063>

TASKS FOR MONDAY:

Cites and Cited by **NViZ: An Interactive Visualization of Network Security Systems Logs**

Cites and Cited by **VIAssist: Visual Analytics for Cyber Defense**

Cites and Cited by **Large-scale Network Monitoring for Visual Analysis of Attacks**

Cites and Cited by **Tracking smart grid hackers**

Cites and Cited by **Visualization of System Log Files for Post-incident Analysis and Response**