



CAPWIC 2024 Submissions

[4] Kiymet Akdemir ([Virginia Tech](#)) and Pinar Yanardag ([Virginia Tech](#)). *Mitigating Intersectional Bias in Text-to-Image Diffusion Models*.

Abstract. Diffusion-based text-to-image models have rapidly gained popularity for their ability to create detailed and realistic images from textual descriptions. However, these models often mirror the biases found in their training datasets, particularly affecting marginalized groups. While previous efforts in debiasing language models have focused on specific types of bias, such as racial or gender biases, initiatives to mitigate intersectional bias have been minimal. Intersectional bias refers to the specific kind of bias experienced by individuals belonging to two or more marginalized groups, such as a black woman. Addressing intersectional bias is particularly critical because it combines multiple forms of discrimination, such as those based on race, gender, leading to compounded negative impacts. In this paper, we propose a method that tackles intersectional bias in diffusion based text-to-image models. Our method utilizes a pre-trained model, applying disentangled modifications to diminish bias while preserving the original quality of unaltered concepts. Extensive experiments show that our method not only matches but surpasses existing work in mitigating intersectional bias across various attributes.

[5] Tatiana Chakravorti ([Pennsylvania State University](#)) and Sarah Rajtmajer ([Pennsylvania State University](#)). *A survey study of research reproducibility and transparency in India and USA*.

Abstract. The increasing focus on reproducibility, replicability, and transparency in scientific research reflects growing concerns about the reliability of published findings across various disciplines. Open science and the science of science communities have initiated innovative efforts to address these issues. However, this survey with 430 scientists reveals that many researchers remain unaware of open science practices and the replication crisis in India as well as in the USA. In the U.S., research institutions are progressively adopting open science practices, including data sharing and open-access publishing, but there is a pressing need for similar initiatives in India. This survey, reflecting the challenges faced by researchers in 2023, highlights the lack of incentives for practices ensuring reproducibility and replication. The publication model has been identified as a significant factor contributing to the reproducibility and replication crisis. Additionally, incorporating different cultural perspectives can enhance the validity and reliability of findings, making them more relevant and applicable across various contexts.

[6] Marina Lin (Thomas Jefferson High School for Science and Technology). *Machine Learning and Density Functional Theory for the Accelerated Discovery of Novel Iron-Doped Al₂B₂ as a Highly Efficient Hydrogen Evolution Catalyst*.

Abstract. The unsustainable reliance on fossil fuels has fueled the search for innovative Hydrogen Evolution Reaction (HER) electrocatalysts. The discovery process has two main problems: 1) Investigating new materials is time-consuming and costly, and 2) Although precious metals (Pt) have been proven to be high-efficiency HER catalysts, they are scarce and expensive, limiting their large-scale industrial usage. This study research combines state-of-the-art machine learning (ML) models with density functional theory (DFT) computations to discover novel, highly efficient HER catalysts. To train the predictive model, I compiled electronic features and DFT descriptors for 318 distinct metal borides in M₂B₂ and M₂B₁ stoichiometry with undoped and single transition metal (TM) doped configurations (TM/M₂B₂ and TM/M₂B₁, respectively). To predict the Gibbs free energy of hydrogen adsorption (ΔG_H), I employed eight ML algorithms and achieved a correlation coefficient (R²) of 91.7% using extremely randomized trees. The trained ML models were applied to 1092 new thermodynamically stable TM/M₂B₂, achieving R² = 90.32%. Furthermore, the ML models recommended Fe/Al₂B₂ with a near-zero ΔG_H of -0.086 eV. The DFT calculations on Fe/Al₂B₂ confirmed its exceptional HER catalyst potential, outperforming Pt across various hydrogen coverages (θ from 1/9 to 5/9). Our generalizable DFT-based ML model revolutionizes the computational material development process, and the discovery of Fe/Al₂B₂ is a breakthrough in HER catalysis. Our research aligns with CAPWIC's focus on computing research by showcasing how advanced computing techniques can lead to significant breakthroughs in other fields. It emphasizes the importance of diversity in computing, specifically in terms of interdisciplinary collaboration.

[7] Judy Wang (Georgetown University) and Hongmei Chi (Florida A&M University). *Explore Phishing Email Generation and Detection via ChatGPT and Gemini*.

Abstract. With the rapid advancements in natural language processing and artificial intelligence, language models like ChatGPT and Gemini have emerged as powerful tools capable of generating human-like text. However, this raises concerns about their potential misuse in generating phishing emails that are difficult to distinguish from legitimate messages. This paper investigates the use of ChatGPT for both phishing email generation and detection. Firstly, we explore the capabilities of ChatGPT/Gemini in crafting phishing emails that imitate genuine communications, allowing cybercriminals to scale their attacks rapidly and target specific individuals or organizations. By examining the characteristics and patterns of these generated emails, we identify key indicators that differentiate them from legitimate messages. We assess the effectiveness and robustness of the proposed detection methods through empirical evaluations using a diverse dataset of phishing emails, genuine communications, and LLM-generated messages. The results shed light on the challenges and opportunities in leveraging ChatGPT's and Gemini's capabilities for cybersecurity purposes.

[9] Xinyu Wang (Pennsylvania State University) and Sarah Rajtmajer (Pennsylvania State University). *Understanding the disparity between Public and Private Opinion in Social Media*.

Abstract. The proliferation of social media has given rise to complex communicative behaviors, one of which is the 'spiral of silence'—a social phenomenon suggesting that people tend to remain silent when they believe their views are opposed by the majority, to avoid social isolation. This study expands on this concept by investigating how the perceived level of support from one's social network could influence the likelihood of expressing or withholding personal opinions. Through an analysis of 390 survey responses and employing a mixed-methods approach, this research aims to address questions regarding the existence of the disparity between public and private opinions, the contextual factors shaping these perceptions, the role of topic polarity on the willingness of self-expression, and how the perceived support or lack thereof from online connections associates with the level of disparity in expressed opinions. The findings aim to dissect the reasons underlying the divergence between public and private opinions and the underpinnings of self-censorship, particularly how such dynamics inhibit the proactive challenge and correction of misinformation. The study enhances our understanding of the mechanisms of the spiral of silence and examines the potential consequences of this reticence in the digital communication landscape. It offers valuable insights into the potential ramifications for public discourse, which have become increasingly pertinent in an era where online platforms are pivotal to societal and political narratives.

[10] Yelyzaveta Satynska (Roanoke College) and Rayford Johnson (Roanoke College). *A Maximal Independent Set Of Vertices On a Cellular Grid.*

Abstract. The use of wireless networks generated by cell towers has been growing. If two cell towers are located close to each other and have a frequency similar in value, interferences occur causing unstable network connection. To avoid interferences, network providers are given minimum reuse distance at which two towers can be assigned the same frequency.

To minimize interferences between towers and create room for future assignments in the surrounding area, it is best to use the least number of frequencies and assign each tower a frequency as far in value as possible from towers around it. As it would be difficult to consider all the different distances between cell towers in the real world, we idealize the problem by using a cellular grid C representing physical space and graph theory notations - a vertex is a potential tower location, an edge is a connection between two towers that can interfere. Assigning frequencies to vertices is referred to as coloring.

It was proven that for a reuse distance k , the minimum number of frequencies required to color an infinite grid C is $3k^2+3k+1$. However, it was not proven that each pair of vertices is assigned a frequency as far apart as possible in value in the suggested assignment.

The goal of our research is to find out whether it is possible to color an infinite grid C using $3k^2+3k+1$ frequencies for a reuse distance k such that each pair of vertices is assigned colors as far apart as possible in value.

[11] Van Ha Nguyen (undergraduate student at Virginia Tech), Bharathi Ganesan (undergraduate student at Virginia Tech) and Mary Zeno (undergraduate student at Virginia Tech). *Defense in Augmented Reality.*

Abstract. Augmented Reality (AR) technology integrates digital elements like images, videos, and interactive content into the real-world view through digital devices. With the growing adoption of AR, there is an escalating risk of data breaches due to mismanaged permissions among various simultaneous applications. In this case, the world demands a quick, efficient, and secure application to boost user satisfaction while enhancing privacy and data protection in AR. This study focuses on developing a system where the user's gaze directs which application is in the foreground, thereby offering a targeted solution to ensure data security and user privacy in AR environments.

[13] Yanfu Yan (William & Mary), Kevin Moran (University of Central Florida) and Denys Poshyvannyk (William & Mary). *Enhancing Code Understanding for Impact Analysis by Combining Transformers and Program Dependence Graphs.*

Abstract. Impact analysis (IA) is a critical software maintenance task that identifies the effects of a given set of code changes on a larger software project with the intention of avoiding potential adverse effects. IA is a cognitively challenging task that involves reasoning about the abstract relationships between various code constructs. Given its difficulty, researchers have worked to automate IA with approaches that primarily use coupling metrics as a measure of the "connectedness" of different parts of a software project. Many of these coupling metrics rely on static, dynamic, or evolutionary information and are based on heuristics that tend to be brittle, require expensive execution analysis, or large histories of co-changes to accurately estimate impact sets.

In this work, we introduce a novel IA approach, called Athena, that combines a software system's dependence graph information with a conceptual coupling approach that uses advances in deep representation learning for code without the need for change histories and execution information. Previous IA benchmarks are small and suffer from tangled commits, making it difficult to measure accurate results. Therefore, we constructed a large-scale IA benchmark, from 25 open-source software projects, that utilizes fine-grained commit information from bug fixes. On this new benchmark, our best performing approach configuration achieves an mRR, mAP, and Hit@10 score of 60.32%, 35.19%, and 81.48%, respectively. Through various ablations and qualitative analyses, we show that Athena's novel combination of program dependence graphs and conceptual coupling leads it to outperform the simpler baseline by 10.34%, 9.55%, and 11.68% with statistical significance.

[14] Julia Larson (James Madison University). *Design of a Wearable Coaching System for Shoulder Rehabilitation*.

Abstract. According to UNC Health, stroke recovery is a top reason patients attend physical therapy sessions. Strokes can be problematic for a patient's shoulders because damage to the rotator cuff muscles can result in chronic shoulder dislocations due to the weight of the patient's arm. Continuous at-home physical therapy is crucial for recovery, but many patients fail to adhere to their exercises.

The goal of the proposed research is to explore a wearable computing coaching system to monitor and guide patients through rehabilitation movement execution. The system provides feedback and guidance to the patient to increase adherence to at-home exercises by ensuring accurate rehabilitation movement to prevent injury to the patient. This research is ongoing, and this flash-talk will discuss the design decisions made regarding a shoulder rehabilitation system using haptic feedback to instruct and coach a patient through their prescribed at-home exercises.

[16] Cassandra Williams (Christopher Newport University) and Samuel Henry (Christopher Newport University). *Biomedical Question Answering Using Large Language Models*.

Abstract. With developments in Natural Language Processing (NLP) and Large Language Models (LLM), we can now sift through a vast amount of information with ease. However, these models are not perfect and frequently generate incorrect information. While this is a harmless annoyance for most questions, it can pose a serious danger for medical questions. This research aims to improve biomedical question answering systems using large language models. Specifically, we evaluate several different LLMs, prompt engineering techniques, and methods of incorporating context (snippets from Biomedical publications containing relevant information) into the prompts. We evaluate these methods on the BioASQ dataset, a benchmark dataset consisting of 5049 question-context-answer triplets. Our current techniques are on-par with state-of-the-art methods with the benefit of being much smaller models that run locally.

[17] Prianka Mandal (William & Mary). *"Belt and suspenders" or "just red tape"?: Investigating Early Artifacts and User Perceptions of IoT App Security Certification*.

Abstract. As IoT security regulations and standards emerge, the industry has begun adopting the traditional enforcement model for software compliance to the IoT domain, wherein Commercially Licensed Evaluation Facilities (CLEFs) certify vendor products on behalf of regulators (and in turn consumers). My talk will focus on a simple but timely question: does the traditional model work for IoT security, and more importantly, does it work as well as consumers expect it to? I will describe our investigation (recently accepted to USENIX Security 2024) of the initial artifacts resulting from IoT compliance certification, and user perceptions of compliance, in the context of certified mobile-IoT apps, i.e., critical companion and automation apps that expose an important IoT attack surface. I will explore three key questions: (1) are certified IoT products vulnerable?, (2) are vulnerable-but-certified products non-compliant?, and finally, (3) how do consumers perceive compliance enforcement? I will describe our systematic analysis of 11 mobile-IoT apps certified by IOXT, along with an analysis of 5 popular compliance standards, and a user study with 173 users, that together yielded 17 key findings. Particularly, I found significant vulnerabilities that indicate gaps in certification, but which do not violate the standards due to ambiguity and discretionary language. I will conclude with a discussion on future directions towards a "belt and suspenders" scenario of effective assurance that most users desire, from the status quo of "just red tape", through objective checks and balances that empower the regulators and consumers to reform compliance enforcement for IoT.

[18] Afsara Benazir (University of Virginia), Zhiming Xu (University of Virginia) and Felix Xiaozhu Lin (University of Virginia). *Leveraging cache features to enable SLU on tiny devices.*

Abstract. This paper addresses spoken language understanding (SLU) on microcontroller-like embedded devices, integrating on-device execution with cloud offloading in a novel fashion. We leverage temporal locality in the speech inputs to a device and reuse recent SLU inferences accordingly. Our idea is simple: let the device match incoming inputs against cached results, and only offload inputs not matched to any cached ones to the cloud for full inference. Realization of this idea, however, is non-trivial: the device needs to compare acoustic features in a robust yet low-cost way. To this end, we present XYZ, a speech cache for tiny devices. It matches speech inputs at two levels of representations: first by sequences of clustered raw sound units, then as sequences of phonemes. Working in tandem, the two representations offer complementary tradeoffs between cost and efficiency. To boost accuracy even further, our cache learns to personalize: with the mismatched and then offloaded inputs, it continuously finetunes the device's feature extractors with the assistance of the cloud. We implement XYZ on an off-the-shelf STM32 microcontroller. The complete implementation has a small memory

footprint of 2 MB. Evaluated on challenging speech benchmarks, our system resolves 45%–90% of inputs on device, reducing the average latency by up to 80% compared to offloading to popular cloud speech recognition services. The benefit brought by our proposed XYZ is notable even in adversarial settings – noisy environments, cold cache, or one device shared by a number of users.

[19] Prianka Mandal (William & Mary), Amit Seal Ami (William & Mary), Victor Olaiya (William & Mary), Sayyed Hadi Razmjo (William & Mary) and Adwait Nadkarni (William and Mary). *Poster: "Belt and suspenders" or "just red tape"?: Investigating Early Artifacts and User Perceptions of IoT App Security Certification.*

Abstract. As IoT security regulations and standards emerge, the industry has begun adopting the traditional enforcement model for software compliance to the IoT domain, wherein Commercially Licensed Evaluation Facilities (CLEFs) certify vendor products on behalf of regulators (and in turn consumers). Since IoT standards are in their formative stages, we investigate a simple but timely question: does the traditional model work for IoT security, and more importantly, does it work as well as consumers expect it to? Our work, recently accepted to USENIX Security 2024, investigates the initial artifacts resulting from IoT compliance certification, and user perceptions of compliance, in the context of certified mobile-IoT apps, i.e., critical companion and automation apps that expose an important IoT attack surface. We focus on three key questions: (1) are certified IoT products vulnerable?, (2) are vulnerable-but-certified products non-compliant?, and finally, (3) how do consumers perceive compliance enforcement? Our systematic analysis of 11 mobile-IoT apps certified by IOXT, along with an analysis of 5 popular compliance standards, and a user study with 173 users, together yield 17 key findings. We find significant vulnerabilities that indicate gaps in certification, but which do not violate the standards due to ambiguity and discretionary language. Further, these vulnerabilities contrast with the overwhelming trust that users place in compliance certification and certified apps.

[20] Ying Li (William & Mary), Yuhui Bao (Northeastern University), Pranav Vaid (Stanford University), Gongyu Wang (Lightmatter), Adwait Jog (University of Virginia), Darius Bunandar (Lightmatter), Ajay Joshi (Lightmatter) and Yifan Sun (William & Mary). *TraceSim: a Lightweight Simulator for Large-Scale DNN Workloads on Multi-GPU Systems.*

Abstract. Deep Neural Networks (DNNs) have become increasingly capable of tasks ranging from image recognition to human-language generation, predominantly powered by GPUs. With the growing DNNs' complexity and the datasets' size, training DNNs with a large number of GPUs is turning into a prevalent strategy. However, when designing and deploying such systems, designers usually rely on testing configurations on hardware, which is both costly and inflexible. While an alternative solution is to test on GPU simulators, they are often too time-consuming. Addressing these challenges, we present TraceSim. TraceSim relies on kernel-level execution traces collected from PyTorch execution and can simulate new hardware configurations with extremely high performance and reasonable accuracy. TraceSim is expected to serve as an essential tool in evaluating the system-level design of large-scale DNN training, especially for massive DNN models.

[21] Dian Chen (Virginia Tech), Paul Yang (Virginia Tech), Ing-Ray Chen (Virginia Tech), Dong Sam Ha (Virginia Tech) and Jin-Hee Cho (Virginia Tech). *SusFL: Energy-Aware Federated Learning-based Monitoring for Sustainable Smart Farms*.

Abstract. We propose a novel energy-aware federated learning (FL)-based system, namely SusFL, for sustainable smart farming to address the challenge of inconsistent health monitoring due to fluctuating energy levels of solar sensors. This system equips animals, such as

cattle, with solar sensors with computational capabilities, including Raspberry Pis, to train a local deep-learning model on health data. These sensors periodically update Long Range (LoRa) gateways, forming a wireless sensor network (WSN) to detect diseases like mastitis. Our proposed SusFL system incorporates mechanism design, a game theory concept, for intelligent client selection to optimize monitoring quality while minimizing energy use. This strategy ensures the system's sustainability and resilience against adversarial attacks, including data poisoning and privacy threats, that could disrupt FL operations. Through extensive comparative analysis using real-time datasets, we demonstrate that our FL-based monitoring system significantly outperforms existing methods in prediction accuracy, operational efficiency, system reliability (i.e., mean time between failures or MTBF), and social welfare maximization by the mechanism designer. Our findings validate the superiority of our system for effective and sustainable animal health monitoring in smart farms. The experimental results show that

SusFL significantly improves system performance, including a 10% reduction in energy consumption, a 15% increase in social welfare, and a 34% rise in Mean Time Between Failures (MTBF), alongside a marginal increase in the global model's prediction accuracy.

[24] Zheyue Caroline Yao (CS Department, Swarthmore College) and Yue Zhang (CS Department, Swarthmore College). *3D Terrain Generation with Procedural Noise and Pseudo-Delaunay Triangulation*.

Abstract. Terrain generation has been an essential field in game development as well as in film industries with the increasing use of special effects. This project is dedicated to demonstrating 3 different procedural noise terrain generation algorithms: Perlin Noise, Diamond-Square, and Worley Noise in implementation and observing their "skeleton" by performing 3D mesh using Pseudo-Delaunay Triangulation. With a special focus on analyzing the respective strength and weakness of the three algorithms, a few artworks featuring various kinds of terrains and surfaces are created to illustrate the actual visual effect and diverse characteristics of the algorithms. These artworks include a variety of styles inspired by real world and masterpieces in art history, such as traditional Chinese-Style Shan-Shui, volcano, human cell-like tissue, and more. Besides demonstrating high adaptability of the algorithms by creating different artworks, this project also highlights their extensive flexibility through direct modifications and adjustments on certain parameters and inputs. As a public repository on Github, this project gives people that are interested in terrain generation and procedural noise an opportunity to visualize and experiment with their own ideas.

[25] Caroline Gihlstorf (University of Virginia) and Yangfeng Ji (University of Virginia). *Studying the impact of finetuning language models with part of speech tags on their syntactic abilities and levels of gender bias.*

Abstract. Despite the ability of language models to produce impressively human-like text, the extent to which models are learning generalizable rules about language as opposed to spurious correlations is not fully known. Models reproduce biases and prejudice from learned correlations in their training data, and the extent to which they are using similar syntactic and semantic mechanisms as humans do is unclear. Understanding what kind of information helps and hinders models' abilities to understand linguistic information and avoid reproducing biases and prejudice is important in better characterizing how they learn from the text they train on. We hypothesize that models that have a more robust understanding of syntactic information may be less prone to rely on spurious correlations between tokens that reproduce biases and prejudice. We are investigating whether or not fine-tuning English language models on part-of-speech (POS) tags, which we consider to provide syntactic information, a) improves their syntactic generalization abilities and b) in so doing, reduces their gender bias. We are testing pretrained language models on a variety of test suites to establish baselines for their performance. We then plan to fine-tune the models with explicit POS information and reevaluate the fine-tuned models' performance on the baselines. Our goal is to determine whether this fine-tuning procedure aids model performance and, if so, what this may reveal about what information models are using to learn language.

[26] Elizabeth Palmieri (University of Virginia) and Yangfeng Ji (University of Virginia). *An Empirical Study on Aspect Based Text Summarization.*

Abstract. Aspect based text summarization is becoming an increasingly pertinent focus of summarization based tasks for large language models. The inclusion of an aspect is particularly helpful for users who want a text summarized based on a particular subject—this allows for the summarization task to center around a focused topic rather than the text itself. For instance, a user is able to guide the model to generate a summary of a famous person regarding their personal life as opposed to their career, allowing for the summary to be more tailored to the user's needs.

The goal of this work is to improve the quality of the text generated by large language models for an aspect based text summarization task. In order to achieve this, we perform an empirical study utilizing various models performed on a variety of benchmark datasets. We evaluate the performance of these models and contribute results which surpass the current benchmarks for aspect based text summarization.

[27] Aarushi Sachdeva (Swarthmore College) and Kenneth Barkdoll (Swarthmore College). *Applying Genetic Algorithms for Colour Based Image Clustering.*

Abstract. This study explores an innovative method for creating visually appealing image collages through genetic algorithms, focusing on color-based clustering. It utilises a novel crossover function that operates on both the horizontal and vertical planes of a 2D grid, enabling a nuanced recombination of genetic material to better preserve color clusters. Simultaneously, the tailored fitness function rewards localized color uniformity and penalizes deviations from the original color distribution. These modifications ensures collages not only cluster similar colors effectively but also faithfully represent the original image dataset's color palette. This approach demonstrates the genetic algorithm's efficacy in handling spatially localized information and its potential for creative organization of visual content.

[28] Soneya Binta Hossain (University of Virginia) and Matthew Dwyer (University of Virginia). *Ensuring Critical Properties of Test Oracles for Effective Bug Detection*.

Abstract. With software becoming essential in all aspects of our lives, especially in critical areas like medical and avionic systems, the need for robust and reliable software is more critical than ever. Even seemingly insignificant software bugs can compromise system stability and security, as evidenced by a simple copy-paste error in Apple devices accepting invalid SSL certificates and a date formatting issue causing a widespread Twitter outage. These realities underscore the need for effective testing and bug detection mechanisms to ensure software reliability. At the heart of this challenge are test oracles, a fundamental component of testing, which play a crucial role in detecting software bugs.

Recognizing the pivotal role of test oracles, my research conducts large-scale studies to understand their impact on bug detection effectiveness, identify limitations in existing test adequacy metrics and automated oracle generation methods. Based on the findings, my research identifies three key properties of test oracles essential for effective bug detection, referred to as CCS (check, correct, strong). These properties ensure that test oracles thoroughly check codes, are correct based on the specification, and strong for bug detection. To enforce the CCS properties, my research introduces a set of methods, leading to the development of OracleGuru framework that significantly enhances the quality of test oracles.

[29] Ananya Prakash (Virginia Polytechnic Institute and State University). *Human-in-the-loop classifier of human-written and AI-written text*.

Abstract. Since OpenAI launched CHATGPT, several Large Language Models (LLMs) are becoming widely used to mimic human conversation. As human interaction becomes increasingly online through dating apps, classroom interactions and electronic consultations, we need a tool that can help us differentiate if we are communicating with another human or an AI system. In 2023, AI21, a start-up, conducted a large-scale "imitation game" with 1.5 million participants to evaluate if they could identify if they were chatting online with an AI bot. 68% of the participants were able to successfully identify a bot, particularly by asking moral questions. My research aims to leverage the inability of LLMs to master human moral reasoning and the human intuition displayed in the AI21's experiment, to develop a human-in-the-loop(HITL) Machine Learning model that can classify if a given text snippet containing a moral question and corresponding response is human-written or AI-written. The dataset will consist of moral questions derived from popular datasets such as ETHICS and corresponding responses from human participants as well as GPT-4. Classification models including Naive Bayes, Linear SVM and Logistic Regression will be experimented with to identify the best performing model on the data. Human crowd workers will then label the same data based on their intuition as "human-written" or "AI-written" and this will be fed to the ML model. The hypothesis is that the HITL ML model will achieve a higher classification accuracy, resulting in a system that can be further developed to be integrated to online chatting applications.

[30] Shikha Shrestha (Swarthmore College) and Ben Mitchell (Swarthmore College). *Assessing Model Performance Estimation in the Presence of Missing Data*.

Abstract. In machine learning, the partitioning of data into training and testing sets is standard practice to construct and evaluate models. However, the presence of missing data poses challenges to model performance assessment. This study investigates the impact of missing values on model performance across various imputation methods. We investigated the differences in the performance estimation of random forest classifiers when using the most common imputation methods to replace missing values. Missing Values were artificially introduced into the dataset at varying proportions. Depending on whether the missing data was present in just the train, or both train and test set, we observed a clear pattern of overestimation or underestimation in model performance. Specifically, we observed a pronounced tendency for model performance to be overestimated when missing values were present in both the training and testing sets, although the extent of overestimation varied among the imputation methods. Notably, this tendency for overestimation was more prominent when utilizing imputation techniques such as Random Forest compared to others. Consistent with previous findings, our results also indicated that model performance deteriorates as the proportion of missing values in the training set increases, irrespective of imputation technique. This finding underscores the need for caution when working with data featuring missing values in a machine learning context.

[31] Desire Omotoye (George Mason University) and Baseerat Chaudhry (George Mason University). *Enhancing Post-Acute Cardiovascular Care: A Cross-Dataset Analysis of Healthcare and Wearable Technology Access in the Washington Metropolitan Area*.

Abstract. Over the years, we have seen massive surges in hospitalizations and mortality rates due to issues with chronic conditions, such as cardiovascular heart disease (CVD). This has become more challenging with the emergence of viruses like COVID-19, introducing additional and potentially unknown risk factors that could affect symptoms and care of heart conditions. These challenges disproportionately impact historically marginalized communities, where barriers to access may hinder their ability to obtain and sustain adequate post-acute care following diagnosis and treatment. In recent years, we have seen technical solutions emerge, namely telehealth and wearable devices, that aim to address gaps in healthcare access and provide support for monitoring conditions. However, limited research has explicitly investigated how access to healthcare, both physical and digital, may affect the ability of marginalized communities to effectively monitor chronic conditions. In this paper, we establish a foundation for understanding this problem and potential solutions by conducting an exploratory, cross-dataset analysis of healthcare access and engagement in the Washington Metropolitan area. Given the prevalence of cardiovascular disease in historically marginalized communities, we focus our initial analyses on access and engagement with facilities and technologies for monitoring cardiovascular health. Our analysis suggests that the utilization of wearables among individuals at risk of CVD may be strongly associated with demographic factors such as age, race, education, income, and health insurance coverage. Moreover, the fraction of individuals that use wearables are more likely to monitor and share data from their devices with their healthcare providers.

[32] Sarah Wilsey (Randolph-Macon College). *How Many Cards to the Center of the Universe: The Connectedness of Randolph Macon College.*

Abstract. Graph theory allows us to navigate the connections within a graph, which often represent more than just a collection of edges and vertices. One of the biggest applications of graph theory is exploring and visualizing a social network. A social network graph takes a population and represents each person with a vertex and each direct relationship with an edge connecting the two vertices. This presentation will explore the social network of Randolph Macon College through the context of graph theory by examining the relationships between students the probability of being connected to one another. The presentation combines the experimental probability and theoretical probability of one student being connected with another student, producing an overview of the social network of Randolph Macon College.

[33] Natalie Andrus (Virginia Tech) and Wei Lu Wang (Virginia Tech). *Implications for Design in Outdoor Asynchronous Video Sharing .*

Abstract. Life circumstances often necessitate adoption of computer-mediated communication, specifically the leveraging of asynchronous video sharing [AVS]. Popular use cases for AVS include people talking over long distances with loved ones and uploading content onto social apps. This interest in AVS calls for an examination of user needs and an assessment of available technologies. We are currently performing two diary studies that will inform the design of AVS platforms. Seventeen participants were selected, on a volunteer basis, from both undergraduate and graduate computer science HCI courses at Virginia Tech. These students represent adults aged 18-30, with routine levels of technological familiarity. Over three weeks, participants will share outdoor experiences via Marco Polo, an ASV app. At the end of each week, participants report on their experience using the provided diary study templates. Our preliminary results see participants reporting a desire for time-monitoring tools that mediate video length. Another theme is uncertainty of video-content reception due to lack of immediate audience feedback. As the study continues, we expect more AVS user needs and challenges to emerge. Specifically, we expect a preference for front-facing camera recordings, a disinterest in voice effects, and an increased willingness to share. To conclude our study, we will review the completed reports, conduct focus-group interviews, perform qualitative analysis, and reflect on design implications for emotionally impactful AVS tools.

[34] Jordan Horrall (Virginia Tech), Jennifer Chandran (Virginia Tech), Pallavi Chavan (Virginia Tech), Will Gaudian (Virginia Tech), Yaw Owusu Snr (Virginia Tech), Phillip Jordan (Virginia Tech), Aditya Rana (Virginia Tech) and John Zhao (Virginia Tech). *Building Tools to Support Diary Studies.*

Abstract. Diary studies offer a way to explore the long-term use of technologies. Unlike many other experimental approaches, diary studies collect diary entries from users over many days or weeks, providing insights that cannot be gleaned from lab-based studies. Diary studies are a good match for academia as well, providing a low-effort way to explore novel technologies. However, currently lacking are the tools that allow diary study teachers and administrators to craft diary experiences and aggregate results, and tools to help students understand their results and engage and dialogue about them.

This project seeks to design and implement a new interactive web platform for diary-based learning to improve education, particularly focusing on topics like human-computer interaction, computer ethics, and web design. This emerging platform features various functionalities, including software matching, diary submission, diary analysis, and intelligent template generation. Highlighted in the project is the design of interactive systems using human-computer interaction techniques, systems implementation through full-stack development in front-end development (JavaScript/React/HTML and CSS), back-end development (MongoDB), and AI Topic Modelling Techniques (Gensim).

Our poster highlights the current prototype system, including screenshots showing key features and envisioned results gleaned from the system. Highlighted in the poster are scenarios that help professors craft diary study experiences for their students and scenarios that help students glean insights from diaries that they and their fellow classmates created.

[35] Abigail Ensogna (James Madison University). *Labelling Data in the Field of Wearable Computing*.

Abstract. Recording and identifying human motion can be done largely through the use of inertial sensors. The data from these devices provides insight into human movement such as step counts, mobility, and exercise performance. This information is helpful to humans, but only when paired with context. Often times, this context is gained through the labeling of data after motion is captured; this process is usually time-consuming and burdensome. To address this challenge, we are examining new systems to efficiently label data in the moment during capture. By doing this, we enable ourselves to train machine learning algorithms to recognize these exercises in real-time with precision and accuracy.

Our current approach is to allow for the user to record the start times of exercises as they are performed by the participant, and later merge these labels with the sensor data according to Unix time. Python has been the language of choice during this research because of its libraries (such as pandas) that help to synthesize multiple time series with ease.

Moving forward, this research aims to implement the system and collect data through a user study to determine the effectiveness of our strategy.

[37] Sindhuja Madabushi (Virginia Tech), Haider Ali (Virginia Tech), Ahmad Khan (Virginia Tech), Mengmeng Gu (Virginia Tech) and Jin-Hee Cho (Virginia Tech). *Privacy Preserving and Feature Importance Based Incentive Mechanism in Vertical Federated Learning*.

Abstract. Federated Learning (FL) is a collaborative approach that allows multiple participants to train a machine-learning model without sharing private data. Despite research on incentivizing participation in horizontal FL, vertical FL - participants contribute different features for the same set of samples - has seen less focus on incentive mechanisms. The reluctance to join such federations often stems from concerns about privacy, computational costs, and resource demands. To address this, we propose PERFACY-FL, an innovative incentive mechanism encouraging participation by offering privacy assurances and rewarding clients based on their data's impact on model accuracy. Clients contributing to high-quality data significantly to improve the model's performance receive greater rewards. Additionally, to safeguard client privacy and prevent data reconstruction by others in the federation, the server implements differential privacy techniques. This approach aims to balance privacy preservation with the motivation for increased participation, thereby enhancing the global model's performance in a secure and equitable manner.

[38] Qi Zhang (Virginia Tech), Lance M Kaplan (US DEVCOM Army Research Laboratory), Feng Chen (University of Texas, Dallas), Dong Jeong (University of the District of Columbia) and Jin-Hee Cho (Virginia Tech). *Winning the Social Media Influence Battle: A New Approach to Understanding and Spreading True Information*.

Abstract. In the world of social media, where we share and receive lots of information, there is a challenge called the "Competitive Influence Maximization" (CIM) problem. This is when groups like political parties or companies try to influence people's choices, especially when those choices are binary (agree or disagree). This study particularly focuses on preventing false information from spreading which may cause harm to someone's reputation, cause money loss, or mislead public opinion. To solve this problem on social media, people oversimplify things by viewing opinions as just yes or no, without considering how people form their opinions in a more complicated and changing way. Our research uses AI algorithms, called "deep reinforcement learning," where human users' opinions are modeled by a belief model that can represent uncertain opinions to understand how sure or unsure people are about their beliefs, aiming to spread truthful information more effectively. Our study aims to find the best "seed users," who are key people who can help spread true information effectively because of their influence and the way their opinions evolve. The key contribution of this study is to propose an uncertainty-based opinion model that can significantly combat false information spreads. In addition, we conduct the competitive influence analysis by two parties spreading true and false information respectively where seed nodes are selected based on different criteria when a given online social network's users and the connectedness among them are not fully known.

[39] Kiymet Akdemir (Virginia Tech) and Pinar Yanardag (Virginia Tech). *MIST: Mitigating Intersectional Bias in Text-to-Image Diffusion Models*.

Abstract. Diffusion-based text-to-image models have rapidly gained popularity for their ability to create detailed and realistic images from textual descriptions. However, these models often mirror the biases found in their training datasets, particularly affecting marginalized groups. While previous efforts in debiasing language models have focused on specific types of bias, such as racial or gender biases, initiatives to mitigate intersectional bias have been minimal. Intersectional bias refers to the specific kind of bias experienced by individuals belonging to two or more marginalized groups, such as a black woman. Addressing intersectional bias is particularly critical because it combines multiple forms of discrimination, such as those based on race, gender, leading to compounded negative impacts. In this paper, we propose a method that tackles intersectional bias in diffusion based text-to-image models. Our method utilizes a pre-trained model, applying disentangled modifications to diminish bias while preserving the original quality of unaltered concepts. Extensive experiments show that our method not only matches but surpasses existing work in mitigating intersectional bias across various attributes.

[40] Emily Chang (University of Virginia), Caroline Gihlstorff (University of Virginia) and Jade Gregoire (University of Virginia). *Endangered Languages—How much data do we need to model them well?*

Abstract. Saving and revitalizing endangered languages is crucial to maintaining cultural diversity. Because endangered languages lack sufficient data to pre-train language models, fine-tuning serves as a potential method for language models to specialize in endangered languages. We seek to identify a lower limit to the amount of data required to fine-tune a model around a “low-resource” language. We adopt a transfer learning approach and establish linguistic proxies. English serves as a high-resource language and French serves as a low-resource language. We finetune a pre-trained English language model on a separate target language: French. The monoglottism of the language model ensures that the model has no prior knowledge about the target language and allows us to simulate fine-tuning a language model. We will evaluate model performance several times during the process of fine-tuning with increasing amounts of data, simulating a low-resource scenario. In estimating how much data is needed for this kind of transfer learning to be successful, our findings aid linguists and developers in designing language models that preserve low-resource languages and cultural diversity.

[41] Emily Chang (University of Virginia). *Considering Anthropogenic Factors in PM 2.5 Machine Learning Models*.

Abstract. Current machine learning models fail to account for anthropogenic factors when predicting air pollution levels and can only forecast a few days into the future. By gathering temporal, environmental (e.g. weather), and anthropogenic data (e.g. unemployment) from Albemarle County, Virginia, I constructed two datasets: (1) a dataset comprised of temporal and environmental data and (2) a dataset comprised of temporal, environmental, and anthropogenic data. These datasets are fed into classical machine learning models (e.g. linear regression) as well as deep learning models. To test how generalizable these models are, I will evaluate models on data from Arlington and Charles County, Virginia, communities that are respectively more urban and rural than Albemarle County.

I found that a combination of environmental and anthropogenic seasonal data can improve model performance. For example, the Long Short-Term Memory Model (LSTM) could forecast a year into the future with relatively low error. Model performance decreased when evaluated on data that came from outside of the training distribution (e.g. data from Arlington and Charles County). This finding suggests that a combination of environmental and anthropogenic data cannot generalize to data originating from outside of the distribution.

These findings can aid communities in developing early warning systems that notify local leaders when air pollution might spike, giving them time to develop emergency response plans that protect populations vulnerable to pollution (e.g. the homeless, children, elderly).

[42] Emily Chang ([University of Virginia](#)). *A Methodology for Controlling Smart HVAC Systems in Planetary Environments*.

Abstract. To explore the Moon and Mars, astronauts need smart habitats that support life in harsh environments and remain operational when vacant. Smart HVAC systems are instrumental in smart habitats because HVAC systems maintain healthy indoor air quality levels and thermal comfort. The current research approach is to share Earth-collected data with robotic systems that then process and act on the physical environment. However, planetary environments are radically different from Earth's. The current research approach raises the following questions: (1) how might an algorithm driven by data collected on Earth handle the constraints of a lunar environment and (2) how might a smart habitat, trained on data that reflects Earth-like conditions, behave in a lunar environment?

To answer these questions, the study seeks to (1) develop a supervised machine learning model (MLM) framework that allows HVAC systems developed on Earth to operate within a space environment and (2) develop a smart control approach for HVAC systems that maximizes thermal comfort under limited energy resources. Three MLMs will be developed, forecasting lunar outdoor temperatures, solar energy capture, and indoor temperatures respectively. Because energy is a limited resource, the framework seeks to "budget" energy use. Energy is used only when necessary (i.e. periods of extremely low or high temperatures). As a study that integrates humanity's knowledge of technology made on Earth with the constraints of a space environment, this project informs how life-support systems are developed in future deep space habitats.

[43] Meriel von Stein ([University of Virginia](#)) and Sebastian Elbaum ([University of Virginia](#)). *Distribution-Aware Feature Mitigation for Testing and Evolving Autonomous Robots*.

Abstract. My recent research seeks to learn transformations for inputs that are not strictly within the input distribution for the system under test due to sensor versioning, such as switching to a more cost-effective camera sensor that ultimately degrades image quality. These types of inputs can cause undefined and potentially catastrophic system behavior. I use an encoder to capture features important to the predictions of the system's navigation DNN in order to learn safe transformations of image inputs. This approach leverages the normal distribution of the navigation predictions to fulfill the encoder's assumptions. Furthermore, it can handle common image transformations with as little as a 5K sample dataset.

Next, I plan to enhance current saliency and attribution techniques to uncover globally important features in DNNs for regression tasks, rather than features pertinent to only single images and only for classification tasks as is typical of the current state of the art techniques. Leveraging this new technique to uncover globally important features, I can then identify spurious inputs with greater accuracy, enhancing the explainability of safety-critical systems. Detecting these spurious correlations ahead of time can help developers anticipate features for which the network does not generalize well, identifying features that may lead to system failures before they happen. This technique can be used to identify adversarial patch attacks or naturally-occurring features of the driving environment that have simply been encoded incorrectly by the neural network, as well as to provide an explainability artifact to improve end-user trust.

[44] Ebtesam Al-Haque ([George Mason University](#)) and Brittany Johnson ([George Mason University](#)). *Exploring Developer Interactions with AI-Assisted Tools*.

Abstract. AI-assisted tools such as ChatGPT have drastically changed the landscape of software development over the past year. Existing studies explore how developers could benefit from this through controlled studies and benchmarks on simple use cases. However, not much is known about how developers interact with these tools in practice. In this project, we provide a comprehensive overview of how developers leverage ChatGPT in open source codebases by analyzing a collection of ChatGPT conversations mined from various artifacts on GitHub. Our aim is to map these interactions to various stages of the software development life cycle, identify the range of development activities developers seek support for and which of them are adequately supported by ChatGPT. We also evaluate the adoption of AI-generated code within repositories. Preliminary findings suggest that not only do developers use ChatGPT for implementation and debugging, but they also use it for activities such as planning and design. Our work gleans insights into the practical applications and limitations of ChatGPT in a real-world setting and has implications for designing the next generation of AI-assisted tools.

[45] Eileen Krepkovich (Charlottesville Women in Tech), Kim Wilkens (Charlottesville Women in Tech) and Elaine Cheng (Charlottesville Women in Tech). *Beyond Graduation: Finding Your Community as a Professional Woman in Technology*.

Abstract. As college students, women who are in technology fields often find themselves in the minority. Luckily, most students are able to access supportive student groups such as SWE, WIC, women in data science, and others. But what happens when you graduate and land your first job working as a women in tech? Many women find themselves at companies that are even more gender- imbalanced than their college programs, often in new cities where they have to build new connections from scratch. Many networking groups for women in technology are hard to find or only offer conferences or meetups once or twice a year. It can be challenging to find a supportive community that offers guidance from women's perspective on professional situations such as salary negotiation, getting recognition for your work, and professional development. And as women progress in their careers, additional challenges emerge, including navigating management tracks, maternity leave policies, and juggling family life. Charlottesville Women in Tech (CWIT) is a non-profit that was created to address this need for connection and support among women who work in technology careers. This panel will feature CWIT board members who will share their stories about feeling alone in new jobs, seeking connections, and creating a strong, thriving group of local women who can network and guide each other through career (and life) challenges. The panel will discuss ways to find women in similar situations, how to create networking opportunities, and how to continue fostering connections throughout your entire career.

[46] Shawal Khalid (Virginia Tech) and Chris Brown (Virginia Tech). *Smart Contract Generation: A Comparison Between Generative AI Tools*.

Abstract. Smart contracts, self-executing agreements on the blockchain, have revolutionized the execution and management of agreements by enhancing transparency, security, and efficiency. However, their development poses challenges such as coding complexity, vulnerabilities, and auditing difficulties, particularly in the manual coding process. This complexity and manual approach increase the risk of vulnerabilities and security threats, while also complicating auditing and ensuring security, leading to potential oversights. This research aims to address these challenges by exploring generative AI tools for efficient smart contract development.

The motivation behind this research stems from the desire to boost efficiency, accessibility, and innovation in smart contract creation. By automating coding tasks, the development process can be accelerated.

The research objectives is to assess the efficiency improvement brought about by generative AI, investigate its impact on development timelines, and ensure consistency and standardization in smart contract development. The methodology involves gathering a diverse dataset of smart contracts, selecting generative AI tools (such as ChainGPT, BardAI, ChatGPT), and designing experiments to evaluate their effectiveness based on metrics such as code correctness, efficiency, and adherence to best practices.

It is important to note that once deployed, smart contracts are immutable and cannot be changed, highlighting the importance of thorough development and testing processes. Ultimately, this research aims to uncover the full potential of generative AI as an ally in creating smarter, faster, and more reliable smart contracts, thus contributing to the advancement of blockchain technology and its applications in various industries.

[47] Scott McCrickard (Virginia Tech), Megan Olsen (Loyola (MD)), Mohammed Seyam (Virginia Tech), Chandani Shrestha (James Madison University) and Jixiang Fan (Virginia Tech). *Making voices heard: Classroom tools and techniques for augmenting student opinions* .

Abstract. Research suggests that thoughts, questions, and opinions from underrepresented student groups risk going unheard or underrepresented. This panel presents tools and techniques from both research and classroom efforts that have been helpful in augmenting students from groups that are typically underrepresented in computer science. Areas to be discussed include machine learning, human-computer interaction, ethics and professionalism, and games and gaming. Topics to be explored include examples of helpful prompts and cues to inspire discussion, roles of qualitative and quantitative feedback, and the advantages and risks of anonymity in sharing opinions. Additional readings and approaches, as well as extended opportunities to engage with the panelists, will be made available.

Each panelist will talk briefly about a tool or technique that has been particularly helpful to them in increasing engagement , highlighted by connections to teaching and learning settings. Significant time will be left for questions and comments from the attendees. The panel should prove useful and interesting to both faculty and students.

[48] Stephen Watson (James Madison University), Hannah Lam (James Madison University) and Alex Milanese (James Madison University). *Application of Traveling Salesman Problem*.

Abstract. This project aims to develop a web application that offers personalized travel path recommendations for city visitors, with Harrisonburg as our focal point. However, the application can be tailored to any city worldwide with Google Maps accessibility. The project will be making use of the Google Maps API as our main tool in finding distance and timing data relative to the users provided destinations.

In the application, users, upon entering the application, can select specific destinations they wish to visit within the city. Subsequently, the application generates an optimal driving circle or path with the shortest driving time as consideration, encompassing all selected sites. Leveraging the back end, we translate this challenge into the classical Traveling Salesman Problem (TSP). We present two solutions: an optimal solution for scenarios with a low number of selected sites, focusing on minimizing driving time, and an approximate solution for cases with a higher number of sites, ensuring shorter driving distances. The finalized recommended circle or path is seamlessly displayed on Google Maps for users' convenience.

[49] Genevieve Purcell (University of Virginia), Ananya Ananda (University of Virginia) and Aaron Bloomfield (University of Virginia). *Poster: A Computing Club for Middle School Girls*.

Abstract. We present a poster on the implementation and execution of a computing club for middle school girls. Volunteer students from the University of Virginia travel to The Village School (villageschool.us), a middle school for girls in grades 5-8, once a week to host a one hour long computing club session. Our goals are four-fold: to increase their interest in and awareness of the field, for the girls to connect with UVA students that they can relate to through their shared interest in computing, for them to learn more about both the field overall and programming in particular, and for them to feel more comfortable when they later enter other situations in CS that are more male dominated. The instructional content started with the Hour of Code, and is following general CS1 introductory curricula using Scratch and Python. The pedagogical style focuses on hands-on exercises with active learning concepts throughout. Although the results are too early for proper objective analysis, the subjective results are that the girls love the club, connect with the UVA students, and are excited to learn more. With over one fifth of the school participating in the club, it is one of the their most popular activities. Future plans include bringing in female professors, introducing more advanced programming as their interest dictates, and even having them compete in a hack-a-thon. A deliverable at the end will be a guide for how others can replicate our success.

[50] Prajakta Belsare (James Madison University). *Leveraging Wearable Technology to Understand Glycemic Trends in Diabetes: Unraveling Seasonal and Behavioral Influences.*

Abstract. Diabetes, a prevalent chronic condition affecting millions globally, poses significant challenges in achieving optimal glycemic control. Despite advancements in wearable technology, patient-generated data from devices like continuous glucose monitors (CGMs) remain underutilized. This underutilization limits our understanding of glycemic trends, particularly concerning seasonal variations and changes in behavior.

This talk presents findings from a study leveraging CGM data from people with primarily type 1 diabetes. This work explores how glycemic trends fluctuate across holidays, seasons, days/weeks, and population subgroups. By analyzing the interplay between lifestyle behaviors, food intake, and diabetes outcomes, we aim to bridge the gap in understanding the influence of seasonal and behavioral factors on glycemic control.

This research underscores the importance of utilizing wearable technology to capture real-world data and highlights the need for tailored interventions targeting vulnerable periods for poor glycemic control. Insights gained from this study have the potential to inform personalized approaches to diabetes management, ultimately improving both short- and long-term outcomes for individuals living with diabetes.

[51] Mary Kate McGranahan (Virginia Tech). *Using Technology to Unleash the Power of Positive Self-Talk in Endurance Sports.*

Abstract. Positive self-talk has the potential to improve performance in endurance sports, but how can athletes remind themselves to engage in positive self-talk during the exact moments when their brain is beginning to give into negative thoughts? I propose that the wearable technology many endurance athletes already carry with them during activity could be key to encouraging positive self-talk in long endurance events such as the marathon, therefore increasing performance.

Self-affirmation activities can provide long-term benefits for people including decreasing stress and anxiety which helps to increase goal attainment. In addition, research in endurance sports has shown that the brain is often a limiting factor in a human's endurance capacity. The brain sets off alarm bells when the body starts to feel pain, long before it is at its limit. I will conduct a systematic literature review of positive affirmation practices in sports. I will combine this information with data from a diary study that I am currently running on positive affirmation writing apps to inform the design of an app that pushes notifications to a user's phone or watch during a marathon reminding them of positive affirmations that they have been using to help them throughout their training. I will then scrape various resources such as twitter for written race recaps from athletes who have run marathons and use sentiment analysis to pinpoint moments in the race where descriptions commonly turn negative and design the app to send notifications at those moments.

[52] Susan Zehra (Old Dominion University), Syed Rizvi (Old Dominion University) and Na Na (N/A). *Enhancing Security in Vehicular Networks: A Novel Algorithmic Approach to Safeguarding Inter-vehicular Communications and Privacy.*

Abstract. Vehicular networks have become a crucial technology for implementing various safety applications for both drivers and passengers. These networks are currently receiving significant attention due to their ability to facilitate access to a diverse range of ubiquitous services. However, the growing popularity of vehicular networks has also led to an increase in security vulnerabilities within their inter-vehicular services and communications, resulting in a rise in security attacks and threats.

Ensuring the security of vehicular networks is paramount, as their deployment should not compromise the safety and privacy of stakeholders. Effectively defending against a broad spectrum of attacks necessitates the development of efficient and reliable security solutions and services. This scientific research adopts a structured and innovative approach to enhance security in vehicular networks. The study introduces a novel algorithm designed to secure vehicular networks and outlines its response to potential security threats, vulnerabilities, and privacy issues within the vehicular environment.

[53] Zahra Rizvi (Grafton High School), Sophia Rizvi (Grafton High School) and Yousuf Mohammed (Virginia Peninsula Community College). *MedAI: Pioneering Rheumatoid Arthritis Detection with Python-Driven EHR Diagnosis Extraction*.

Abstract. This project introduces an innovative diagnostic tool designed to detect Rheumatoid Arthritis using a Python-based pipeline. The pipeline focuses on constructing machine learning classifiers for extracting diagnoses from Electronic Health Record (EHR) data, specifically Natural Language and Narrative data. After conducting a comprehensive market survey of such tools, it was determined that the approach is universally applicable. The pipeline successfully identifies Rheumatoid Arthritis patients within EHR data from different medical centers. The source code is planned to be made open source, accompanied by extensive documentation for installation and usage on both Windows and Linux/Windows (Dev) systems. Tool dependencies are managed through custom kernels or pip. An interactive demo is scheduled to be accessible as a Kaggle session, and the pipeline's workflow will combine automated processes with manual evaluations, simplifying the diagnostic extraction process for physicians. The study's findings and practical applications will be detailed in the poster.

[54] Grace Chamberlain (James Madison University), Kendra Letchworth-Weaver (James Madison University) and Lily Jade Joyce (James Madison University, Rensselaer Polytechnic Institute). *Computational Investigation of High Entropy Oxides*.

Abstract. High entropy oxides have the potential to improve lithium-ion batteries, thermoelectric materials, and corrosion shielding, but synthesizing novel compositions of high entropy oxides can be time-consuming and expensive due to their high formation temperatures. Using high-performance computing (HPC) resources at James Madison University and the University of Virginia, we can run JDFTx, a density-functional theory software, to model theoretical high entropy oxides, predict their formation temperatures, and collect statistical data on, for example, the bond lengths between the atoms in the material. However, since we do not have experimental data with which to compare our results when investigating theoretical materials, we tested the reliability of our predictions through altering different structural characteristics of a proposed composition. We then analyzed the effect of changing a given structural characteristic on the predicted formation temperature of the material and found no major differences in the formation temperatures, showing that we can continue to use HPC and our computational methods to investigate theoretical high entropy oxides and produce significant results.

[55] Xiaoxiao Gan (Virginia Tech). *Exploring Challenges of Integrating GUI Testing Framework in CI/CD Workflows on Open Source Repositories*.

Abstract. This study aims to conduct an empirical study on Graphical User Interface (GUI) testing in Continuous Integration/Continuous Delivery (CI/CD) processes, exploring its potential challenges when integrating GUI testing into popular CI/CD platforms, such as GitHub Action, in the open-source repositories. GUI testing, which is used to evaluate the user interface of applications, plays an important part in software quality assurance. However, little is known about how exactly popular GUI testing frameworks and techniques, such as Selenium-based technologies, are adopted in CI/CD workflows and their impact on open-source software development processes. Therefore, this work focuses on investigating developers' perception on challenges, practices and impacts when introducing GUI testing into CI/CD workflows. To gather data, purposive sampling will be employed to select participants for an opinion survey distributed via email, followed by in-depth interviews to garner more comprehensive insights.

[56] Huayu Liang (Virginia Tech). *Understanding End User and Software Engineer Perceptions of Dark Patterns*.

Abstract. Dark patterns refer to user interface design choices crafted to manipulate, deceive, or trick users into taking actions that they might not otherwise choose to do. These design patterns exploit cognitive biases or draw on the user's susceptibility to persuasive techniques to execute certain behaviors that benefit the designers or stakeholders, often at the expense of the user's interests. Dark patterns can be found everywhere in digital interfaces, such as websites and mobile apps. There are many studies by researchers about dark patterns to bring more awareness about it. Studies include but are not limited to generating a taxonomy of dark patterns, understanding deceptive practices of dark patterns on online platforms, and exploring the dark patterns from an end-user perspective. Existing research studies highlight the impacts of dark patterns and their potential harm on user decision-making, therefore helping facilitate recommendations for stakeholders including researchers, practitioners, and regulators to study and develop new tools or policies to help mitigate and reduce the use of dark patterns. However, very little academic literature has explored software developers/engineers' perceptions of dark patterns which are the ones who design and implement the user interface designs. This study aims to understand the perceptions of dark patterns not only from end-users but also from software engineers, gaining more insights into dark patterns from stakeholders such as staff from software engineering companies.

[57] Khyati Kiyawat (University of Virginia) and Kevin Skadron (University of Virginia). *Performance modeling of memory-bound kernels for Processing-In-Memory(PIM) architectures*.

Abstract. With the advent of data-intensive applications like bio-informatics, AI/ML, etc., which are primarily memory-bound, innovative solutions like bringing compute closer to the data are being explored. While academia and industry have proposed various near/in-memory architectures, evaluating their performance remains challenging. An ideal approach to evaluating the performance of different PIM architectures would be to have a standard benchmark suite comprising applications from various workloads. However, a standard infrastructure to compile the benchmarks and execute them on real PIM hardware or simulators is missing. Although industry-based PIM prototypes undergo real hardware testing, these are rarely available to academia. Consequently, academia mostly relies on functional simulators and analytical models that lack consideration of timing parameters like refresh intervals. Until efforts to build a generic compilation and simulation framework are realized, a step closer to achieving a fair performance evaluation could involve autogenerating Read and Write(R/W) address traces for various kernels and running them on DRAMsim3, a commonly used DRAM simulator to obtain the total cycle count for any given DRAM configuration.

We propose to design a benchmark suite for PIM and evaluate the inline kernels by autogenerating R/W traces for different input dimensions and data partition schemes. Since manual calculation is time-consuming and error-prone, the proposed approach will offer more accurate insights into various data-parallel kernels. This ongoing-work aims to present a step-by-step approach for analyzing various PIM-amenable kernels across different categories of PIM architectures.