# Table of Contents

- **Dean’s Message** .................................................. 3
- **Big Data** .............................................................. 4
- **Data Science** ......................................................... 8
- **Data and Software Security** ..................................... 12
- **Image and Video Processing** ..................................... 14
- **Big Knowledge, Bioinformatics, and Medical Informatics** .................................................. 16
- **Distributed Systems, Mobile Computing, and Networking** .................................................. 18
- **Cloud, and High Performance Computing** .................................................. 20
- **Augmented Reality, Virtual Reality Gaming and Graphics** .................................................. 24
- **Advanced Algorithms** ............................................. 30
- **Social/Mobile/Education** .......................................... 34
The Ying Wu College of Computing (YWCC) was established in 2001 and has grown rapidly to become the second-largest college at New Jersey Institute of Technology, consisting of the Computer Science and Informatics departments. Enrolling more than 3,200 students at all levels and graduating more than 800 computing professionals every year, YWCC is the largest generator of computing tech talent in the greater New York Metro area.

The tenure-track faculty and Ph.D. student population of YWCC has also grown significantly in the last decade. These talented individuals are responsible for the high-quality academic research described in this report, research that is financially supported by a variety of government funding agencies and corporate entities, published in top venues, and in some cases patented and commercialized.

Spanning a wide spectrum of topics, from human-computer interaction to cybersecurity to sophisticated data science algorithms, the research conducted by the YWCC researchers ranges from deep mathematical theory to very practical applications. Committed to sharing beyond publication, much of the software developed in these projects is made available to the general scientific community through open-source repositories.

I invite you to read through the abstracts of the research projects described in this report and encourage you to reach out to the individual researchers for more details, if needed. We welcome new ideas, collaborations and any form of research partnership imaginable.

Sincerely,

Craig Gotsman
Distinguished Professor
Dean, Ying Wu College of Computing
New Jersey Institute of Technology
Faculty at NJIT work on a variety of Big Data topics including dimension reduction, complex workflows and energy optimization, as well as data mining research with emphasis on keyword search within graphs and trees. The research is a combination of algorithms and techniques to improve performance of large-scale data systems and extract relevant information.

Our researchers have presented their findings at top-tier conferences including VLDB and SIGMOD, and have received various awards including the SIGMOD Test of Time Award in 2015. Researchers have also received best paper awards at conferences for high-performance computing, parallel and distributed computing, multimedia and other relevant themes.
Dimensionality and Scalability Issues in High Dimensional Spaces
For many, fundamental operations in the areas of search and retrieval, data mining, machine learning, multimedia, recommendation systems and bioinformatics, the efficiency and effectiveness of implementations depends crucially on the interplay between measures of data similarity and the features by which data objects are represented. When the number of features (the data dimensionality) is high, the discriminative ability of similarity measures diminishes to the point where methods that depend on them lose their effectiveness. Our research proposes techniques for local feature selection with application to search and clustering and is currently looking for ways of finding lower dimensions that increase the data discriminability using machine learning techniques.

Multi-Instrument Database of Solar Flares
Solar flares are the most prominent manifestation of the sun’s magnetic activity. They emit radiation that could potentially damage power systems, interfere with civilian and military radio frequencies and disrupt spacecraft operations. To improve analysis, in collaboration with the department of Physics, we aim to integrate and enrich solar data captured by various solar flare observing instruments around the world.

Research Area: Multi-media databases, spatio-temporal databases, recommender systems.

Visualization-Driven Transparency in Human-Data Interaction
To communicate actionable insights from data-driven predictive models, we are developing interactive interfaces that can summarize complex information in a transparent and user-interpretable way and can dynamically update the presented information based on user feedback and preferences. These interfaces integrate the outcome of the predictive models with novel visualization techniques for enabling human judgment and reasoning about data-driven patterns. This ensures we leverage the best of both worlds: computational power for fast extraction of patterns and perceptual and cognitive human faculties for enabling experts (e.g., domain scientists, social media content moderators) to transparently reason about the context, significance and implications of the patterns.

An example of one such interface is MyriadCues (myriadcues.njitvis.com) with which scientists can distinguish between “good” and “bad” simulation models and incorporate their expert feedback for understanding how model rankings change with respect to different parameters and simulation scenarios. By validating the efficacy of this interface with scientists working at national labs, we have demonstrated how visualization can play a critical role in automating the workflow for more trustworthy and effective simulation models.

Research Area: Data visualization, visualization techniques.
Building Faster, Energy-Efficient Analytics Pipelines for Decision-Making

Big data analysis is used to analyze problems related to massive datasets. Today, these datasets are loaded from storage into memory, manipulated and analyzed using high performance computing (HPC) algorithms and then returned in a useful format. This end-to-end workflow provides an excellent platform for forensic analysis, however there is a critical need for systems that support decision-making with a continuous workflow. HPC systems must focus on ingesting data streams; incorporating new microprocessors and custom data science accelerators that assist with loading and transforming data; and accelerating performance by moving key data science tasks and solutions from software to hardware. These workflows must be energy-efficient and easy to program, while reducing transaction times by orders of magnitude. Analysts and data scientists must be able to ask queries in their subject domain and receive rapid solutions that execute efficiently, rather than requiring sophisticated programming expertise.

Scalable Graph Algorithms

Our research is supported in part by an NVIDIA AI Lab (NVAIL) award. NVIDIA is a leading technology company that makes GPU accelerators such as the DGX Deep Learning server. We contribute to RAPIDS.ai, an open GPU data science framework for accelerating end-to-end data science and analytics pipelines entirely on GPUs. These new analytics pipelines are more energy-efficient and run significantly faster, which is critical for making swift, data-driven decisions.

Scalable Graph Learning Algorithms

This project, with the support of a Facebook Research, AI System Hardware/Software Co-Design research award, focuses on developing scalable graph learning algorithms and implementations that open the door for learned graph models on massive graphs. Our lab investigates 1) scalable high performance graph learning systems based on graph convolution network algorithms by improving the workflow on shared memory NUMA machines, balancing computation between threads, optimizing data movement and improving memory locality; and 2) graph learning algorithm-specific decompositions and new strategies for graph learning that can inherently scale well while maintaining high accuracy.

Data Science Leadership

Bader is the founding Director of NJIT’s Institute for Data Science, which includes the Center for Big Data, Cybersecurity Research Center and Structural Analysis of Biomedical Ontologies Center. He also serves on the leadership team for the Northeast Big Data Innovation Hub, supported by the National Science Foundation.

Research Area: Data science, high-performance computing.

Revolutionizing Processing of Big Datasets

To reduce processing time of information in complex scientific computing infrastructures, we develop practical computing and networking toolkits to improve efficiency of complex workflows in big data and high-performance environments. Visual inspection can often detect complex models or discover new patterns in big data environments. We develop visualization algorithms for 3D-volume data generated by scientific computations on supercomputers. Visual feedback is critical to the understanding and validation of physics models used for simulations in computational sciences.

Reducing Energy Consumption in Big Data Computation

The transfer of big data across high-performance networks consumes a significant amount of energy. Employing two widely adopted power models — power-down and speed scaling — we have made inroads into green computing in big data environments. Our approach allows network providers to reduce operational costs and reduce carbon dioxide emissions.

Uncovering Low-Level, Hazardous Radiation

Radioactive substances and biological agents present a serious threat to public health and safety, particularly in densely populated areas. We develop reliable tools to detect and contain radioactive materials to protect the populace and reduce the risk of radiological dispersal devices, such as “dirty bombs.”

Research Area: Big data, green computing and networking, parallel and distributed computing.
The amazing abundance of data available today has created opportunities for corporations, educators and governments to gain previously unavailable insights through deep analysis of datasets. This knowledge discovery has created real-world, actionable intelligence.

Our faculty work on all aspects of data science including natural language parsing, machine learning and deep learning. These techniques provide tools to address issues ranging from security, to fake news detection, to healthcare and genomic data analysis. The work of our researchers has been published in a wide variety of venues including Nature, Science, Nature Medicine, Cancer Discovery and Nature Communications, as well as top-tier data mining, machine learning, biostatistics and bioinformatics conferences and journals.
Deep Reinforcement Learning Network for Traffic Light Cycle Control

Existing inefficient traffic light cycle control causes numerous problems, such as long delays and waste of energy. To improve efficiency, taking real-time traffic information as an input and dynamically adjusting the traffic light duration accordingly is a must. In this project, we propose a deep reinforcement learning model to control the traffic light cycle. In the model, a convolutional neural network is employed to map states to rewards. The model incorporates multiple optimization elements to improve the performance, such as dueling network, target network, double Q-learning network and prioritized experience replay.

Deep Neural Network Modeling to Predict Vehicle Path Flow

Predicting full spectrum vehicle path flow is very challenging since it is determined by both spatial and temporal factors, in addition to many other random factors, such as human behavior and weather. We are developing a Spatio-Temporal Fuzzy Neural Network to accurately predict vehicle path flow in the near term. The deep neural network is composed of stacked Convolutional Long Short-Term Memory, fuzzy neural networks, convolutional layers and fully connected layers. Extensive experiments on big datasets are being conducted to evaluate the efficacy and efficiency of the model.

Research Area: Deep Learning, blockchain technologies, IoT.

Advanced Analytics and Learning for Genomics

We are developing statistical modeling and machine learning techniques, with applications to biology, medicine and healthcare fields. We are especially concerned with how model-driven approaches, theories or empirical applications can be used to address various challenges arising from exploiting large and complex data. Current research topics of interest include, but are not limited to, deep learning, probabilistic and statistical models and theories, machine learning and data mining theories, models and systems. A focus area is analysis of genomic and genetic data, with application for cancer and genetic disorders. This includes extensive collaborations with biologists, geneticists and physicians.

KPI-Driven Content Understanding and Generation with Applications to Digital Marketing

In digital marketing it is crucial to distribute valuable, relevant and consistent content to attract/acquire a clearly defined audience and to optimize various Key Performance Indexes (KPIs), such as conversion rate and click-through rate. This project focuses on developing machine learning techniques for content understanding with application to digital marketing. “Understanding” content is not interpreting its literal meaning (such as the detection of an object, e.g., a dog, in an image), but instead refers to elucidating its implication in driving KPIs. This project in collaboration with Adobe will facilitate content personalization using artificial intelligence-driven technologies.

Research Area: Advanced data analytics, data mining, statistical learning.

Big Data Analytics Laboratory - Data Science with Humans in the-loop

The Big Data Analytics Lab (BDaL) is an interdisciplinary research laboratory that focuses on large-scale data analytics problems that arise in different application domains and disciplines. One primary focus of our lab is to investigate an alternative computational paradigm that involves “humans-in-the-loop” for large-scale analytics problems. These problems arise at different stages in a traditional data science pipeline (e.g., data cleaning, query answering, ad-hoc data exploration or predictive modeling), as well as from emerging applications. We study optimization opportunities that arise because of this unique man-machine collaboration and address data management and computational challenges to enable large-scale analytics with humans-in-the-loop. Our focus domains are social networks, healthcare, climate science, retail and business and spatial data training.

Research Area: Human-in-the-loop large scale data analytics, optimization algorithms.
Preserving Sensitive Healthcare Data
Networks that share Electronic Health Records (EHRs) containing a patient’s personal and clinical data promise improved continuity of care and better health outcomes. However, these networks put highly sensitive patient information at risk and expose healthcare providers to legal jeopardy. We have created “DeepPrivate,” a system that uses machine learning techniques to protect personal health information against cyber attacks.

Enable Real-Time Drug Abuse Risk Behavior Detection
Some Twitter users are surprisingly willing to divulge private details of their lives, including their own drug abuse activities. However, such tweets are hidden within hundreds of millions of other posts, literally needles in a very large haystack. We apply machine-learning methods to recognize drug-abuse-related tweets in a large collection of Twitter postings. Many tweets are geographically tagged, and all are temporally labeled. This allows us to recognize hot spots of drug activity and peak days and times to provide near-real time information to public health officials to plan appropriate tactical and strategic responses.

Ontology-Based Interpretable Deep Learning
Machine learning models are trained with large amounts of data and achieve a certain level of competency in interpreting and classifying new input data. However, even if they work very well, it is nearly impossible to say “why they work well” and lingering doubt persists that in some situations that the classification output of the model might be wrong. In applications such as self-driving cars, this could have spectacularly negative consequences. In our research, we tie predictions of the model to a set of keywords taken from a predefined vocabulary of relevant terms. In other words, the number of words “hardcoded into the model” that influence the outcome produced by a machine learning model for a new input is reduced, and those words are taken from a limited and relevant collection of terms (an ontology). This makes the output of the model easier to interpret as it becomes independent from terms that are irrelevant to the application domain.

Research Area: Social network analysis, machine learning, spatio-temporal data mining.

Mining Concise and Useful Tips for Programming From Heterogeneous Data Sources
A tremendous amount of data containing coding tips or suggestions exists on a wide range of data sources, such as YouTube, Q&A websites, reviews, etc. Mining such tips can help software practitioners, such as developers and students, to learn and use a programming language. We are developing deep learning techniques to accurately classify and extract small, useful and concise coding suggestions from crowdsourcing websites.

Next Generation Question and Answer System for Software Developers
Typically, a software developer’s question contains long text, code examples, “software semantics” and diagrams. This data diversity hinders many state-of-the-art techniques, and systems perform very poorly when automatically answering such questions. We are developing AI techniques that analyze developers’ questions and documentation using deep learning techniques.

Research Area: Software engineering, artificial intelligence, machine learning.

Mining Big Data through Deep Learning
We are designing and implementing new deep learning algorithms and architectures for mining big data. We have developed a 3D-atrous convolutional neural network, used it as a deep visual feature extractor, and stacked convolutional long short-term memory networks on top of the feature extractor. This allows us to capture not only deep spatial information but also long-term temporal information in the data. In addition, we use stacked denoising autoencoders to learn latent representations of the data, to construct feature vectors suitable for classification. We also develop new recurrent neural networks to mine big time series data for stock market forecasting and space weather prediction. Currently, we are building a deep learning framework with generative adversarial networks. This framework will be used for stochastic video prediction, image synthesis and image-to-image translation. The framework can handle model uncertainty as well as data uncertainty and sparsity. Thus, our deep learning models are suited for big data applications that have few, incomplete, imperfect, missing, noisy or uncertain training data.

Research Area: Data mining, machine learning, deep learning, data science.
Data and Software Security

The security research group has done extensive work in homomorphic encryption, supply chain security, blockchain and device/application security. It also makes significant contributions to open source software.

Kurt Rohloff received the first DARPA Young Faculty Award at NJIT. Initial results of a project in cliptography led to a conference award (top 3 rated) paper at Asiacrypt.
Improving Smartphone Reliability and Security
We have developed a wide range of approaches to improve smartphone reliability and security, including static and dynamic program analyses, record-and-replay systems, runtime systems for Moving Target Defense, app self-healing and automatic test generators. These approaches have been released as open source code and have found security and reliability issues in many popular apps, such as NPR News, AirBnB, Waze and Facebook.

Protecting Army Networks
As part of the Cybersecurity Collaborative Research Alliance, a joint effort between the U.S. Army Research Laboratory, Applied Communication Services and six universities, our research aims to advance the foundations of cybersecurity in the context of military networks. We are using a two-pronged approach: theoretical foundations of, and practical approaches for, agile defense of Army devices and networks, as well as measuring and predicting socio-cognitive factors involved in human decision-making during cybersecurity engagements.

Research Area: Programming languages, software engineering and the smartphone side of systems/security.

Combating Data Leaks: PALISADE
We developed a widely used open-source lattice encryption library and software engineering tools for a new family of encryption technologies. This software library provides encrypted computing capabilities such as homomorphic encryption, allowing organizations to outsource computation to cloud computing environments without risking privacy and leaking sensitive information to potential adversaries.

Improving Usability of Open-Source Software
Funded by the first DARPA Young Faculty Award at NJIT and the IARPA HECTOR project, our MARSHAL and Verona projects focus on making it easier to rapidly optimize open-source software to run on commodity hardware. This research focuses on the deployment of the PALISADE open source lattice encryption library on embedded systems.

Research Area: Encrypted computing, lattice encryption implementation, homomorphic encryption, cryptographic program optimization.

Improving Smartphone Reliability and Security
We have developed a wide range of approaches to improve smartphone reliability and security, including static and dynamic program analyses, record-and-replay systems, runtime systems for Moving Target Defense, app self-healing and automatic test generators. These approaches have been released as open source code and have found security and reliability issues in many popular apps, such as NPR News, AirBnB, Waze and Facebook.

Protecting Army Networks
As part of the Cybersecurity Collaborative Research Alliance, a joint effort between the U.S. Army Research Laboratory, Applied Communication Services and six universities, our research aims to advance the foundations of cybersecurity in the context of military networks. We are using a two-pronged approach: theoretical foundations of, and practical approaches for, agile defense of Army devices and networks, as well as measuring and predicting socio-cognitive factors involved in human decision-making during cybersecurity engagements.

Research Area: Programming languages, software engineering and the smartphone side of systems/security.

Blockchain Technology
Our research focuses on advancing three core layers of blockchain: decentralized applications, blockchain itself (consensus) and supporting cryptographic primitives. We have been pushing asynchronous consensus towards both theoretically and practically optimal performance, and we are also working with JD.com to deploy asynchronous consensus in real-world large systems. Application-wise, we have designed a decentralized cybersecurity management system, and showcased possible consumer cost reduction via decentralizing Amazon MTurk tasks on Ethereum. We also designed new cryptographic tools, such as proof-of-storage-time, to enable clients to verify continuous availability of the storage service in either the cloud or the decentralized storage market.

Cliptography: Cryptography Against Subversion Attacks
Software or hardware implementations (including open-source ones) of cryptographic functionalities may deviate from formal specifications. We are designing a new generation of cryptographic specifications such that either the subverted implementation can be easily detected via simple blackbox testing, or we can preserve security even when a malicious implementation contains backdoors.

Research Area: Big data analytics, cyber security, parallel and distributed computing.

Defending Software Supply Chains Against Hackers
Funded by DARPA and the NSF, we have developed “in-toto”, an open-source framework that promises to safeguard software for developers and end users. in-toto provides organizations with insights into the software development and distribution chain, such as having a provable assurance that proper software development practices have been followed. With in-toto in place, it will be more difficult for malicious code to be slipped into software products, thus raising the bar significantly for attackers. Through integrations, in-toto is currently used by thousands of companies and has improved the security of millions of users.

Research Area: Big data analytics, cyber security, parallel and distributed computing.
Our faculty focused on image and video processing continue to advance the science and technology and work with experts in healthcare, state and local government, security and the private sector to examine obstacles and leverage advances in image and video processing to address them.

From the diagnosis of pneumonia and brain tumors, to automated systems that can monitor traffic incidents, roadway congestion and public safety risks, the image processing group is making advances with significant work in image forensics and watermarking, pattern and face recognition, medical image analysis, video analytics, deep neural networks and smart transportation systems.
Deep Learning for Pneumonia Classification and Segmentation on Medical Images

Automatic identification of pneumonia on medical images has attracted intensive study. In this project, we develop a novel joint-task architecture that can learn pneumonia classification and segmentation simultaneously. Two modules, including an image preprocessing module and an attention module, are developed to improve both classification and segmentation accuracies. Experimental results performed on the massive dataset of the Radiology Society of North America have confirmed its superiority over other existing methods.

Deep Morphological Neural Networks and Applications

Given a target image, determining suitable morphological operations and structuring elements is a cumbersome and time-consuming task. In this project, we propose new morphological neural networks, which includes a nonlinear feature extraction layer, to learn the structuring element correctly and an adaptive layer to automatically select appropriate morphological operations. We also use them for their classification applications, including hand-written digits, geometric shapes, traffic signs and brain tumors. Experimental results show higher computational efficiency and higher accuracy when compared with existing convolutional neural network models.


Making Intelligent Transportation Systems Smarter

The New Jersey Department of Transportation (NJDOT) has designated more than 400 CCTV video cameras distributed throughout the state to perform incident monitoring, traffic congestion control and public safety operations. Video streams from these cameras feed to a back-end system. There, video analytics software is used to perform target detection and incident monitoring applications. We are actively working on incorporating wireless sensor networks, hierarchical edge-computing and advanced computer vision to mitigate the challenging problems in various illumination and weather conditions in order to achieve fast and automated video-based traffic monitoring.

Video Analytics Pilot Studies and Testing of Technologies

We propose a new modular approach for statistical modeling of traffic incidents and model selection in order to improve state-of-the-art traffic incident detection and monitoring. We have investigated and developed fully-automated video analytics systems to replace human operators for automated traffic incident detection and monitor e-cameras installed along the major New Jersey highways. We test the proposed technologies and benchmark their performance through online testing using NJDOT test cameras.

Research Area: Computer vision, pattern and face recognition, video processing.
Faculty research has been applied to a number of important applications. Yehoshua Perl and James Geller are Fellows of the American College of Medical Informatics who have worked on improving medical ontologies. An ontology is a mechanism used to encode diagnoses in electronic health records.

Results have been published in AMIA, the top U.S. conference in Medical Informatics, and in JBI, one of the top journals.

We are currently working to develop a new interface ontology specifically for Coronavirus and COVID concepts. Using such terms for annotating Electronic Health Records will make it easier to discover unexpected new symptoms of a disease. Such symptoms (e.g., toe lesions) were discovered with an unnecessary delay in the past three months. This technology will contribute to future pandemic preparedness. We are also applying a similar method to cardiology.
Summarizing, Visualizing and Correcting Large Collections of Medical Terms

Biomedical professions have collected large ontologies (repositories) of medical terms covering diseases, diagnosis, drugs, anatomy, bacteria, genes, chemicals, medical procedures and more. We have developed a theoretical framework to create concise summaries of large ontologies and software tools to visualize these summaries. Using this software, medical experts can easily browse large medical ontologies making it easier to recognize errors in structure and content. Recently, we have used advanced machine learning methods for update and curation of biomedical ontologies. We are also developing methods to interpret large collections of formal medical knowledge in ontologies, which will aid in the prediction of dangerous drug interactions.

James’ Research Area: Medical informatics, semantic web technology, object-oriented database modeling, knowledge representation.

Yehoshua’s Research Area: Medical informatics, semantic web, ontologies.

Michael’s Research Area: Controlled terminologies, ontologies, object-oriented databases, conceptual modeling, open data models.

Open Knowledge Network for Spatial Decision Support

This project focuses on research critical to the development of Open Knowledge Networks (OKN) through the combination and testing of participatory and automated ontology development processes. Three domain-specific case studies (wildland fire, water quality and biodiversity conservation) will build on participatory Geographic Information System (GIS) and ontology development work through engagement of problem-focused stakeholder networks. At the same time, the utility of automated tools for resource discovery, ontology development and social network analysis will be tested in these real-world problem environments. Through integration and comparison of these techniques, we will deliver insights into efficient and effective methods for OKN development.

Capturing, Managing and Utilizing GeoSpatial Multimedia Data for Collaborative Field Research

This project focuses on developing spatial social network analytics to integrate multimedia data in space, time and within a network. The software captures and transfers geospatial multimedia data in a variety of different field settings ranging from developed countries with advanced IT infrastructures to countries that still lack a reliable access to the internet; manage and explore the collected data; apply qualitative, quantitative and spatial data analysis with mining algorithms, visual representations and interactions.

Research Area: GIS, urban informatics.

Deep Learning for Medical AI Problems

We work on several medical AI problems involving images and DNA. In collaboration with neuroscientists, we are creating new models for tumor identification in brain MRI images and are studying them across different data sources. With vascular surgeons we are proposing novel models for identifying vessel and plaque in vascular ultrasound images from real patients. We are studying simple random networks for the classification of histopathology slide images and find them to be highly accurate there.

Adversarial Robust Machine Learning with 01 Loss:

Machine learning models today are highly accurate but not very robust. They can be fooled to misclassify data with minor perturbations known as adversarial attacks. Adversaries targeting a given convex model are known to affect other convex models. We find this transferability phenomenon to be less effective between 01 loss and convex losses such as hinge and logistic, both of which are approximations to 01 loss and known to be affected by outliers. Consequently, it is harder to attack 01 loss models with convex substitute model black-box attacks, and when the black-box attacker is 01 loss, the attack is highly ineffective on all models. Based upon these observations, we are researching novel algorithms and design implementations for scalable and faster 01 loss models.

Research Area: Machine learning, medical informatics.
Distributed Systems, Mobile Computing and Networking

Our researchers are investigating a wide range of topics related to distributed systems, mobile computing and networking. An area of particular interest and research focus is the development of blockchain systems. Faculty research in that area is exploring and testing the potential for increasingly secure, independent organizations that help promote an individual user's personal agency in the blockchain space, while improving a user's data privacy.

Other areas the group is studying include such diverse issues as how to address the online advertising challenge and balancing revenue objectives and maintaining a high-quality user experience, and improving access to available parking in urban environments.

The results of our research have been published in premier and competitive venues, such as ICDCS, PODC and WWW.
**Blockchain and the Sharing Economy**

Thanks to the emergence of blockchain technology, we are now moving quickly toward a new world of decentralized and secure organizations. Given the blockchain at hand, a large number of individuals can coordinate, interact and trade directly, thus governing themselves without the help of centralized platforms. This project studies how individual people can interact through the blockchain to crowdsource human intelligence securely and efficiently. The design is secure and robust, as it not only eliminates the vulnerable reliance on centralized third-parties, but also mitigates the inherent transparency issues of the blockchain to attain data privacy as well as user anonymity. In addition, the design is efficient and practical, and even financially cheaper than the existing centralized systems such as Amazon’s Mechanical Turk.

**A Generic Superlight Client for Permissionless Blockchains**

This project conducts a systematic study on the light client protocol of permissionless blockchains, in the setting where full nodes and light clients are rational. In the game-theoretic model, we design a super light client protocol to enable a light client to employ some relaying full nodes (e.g., two or one) to read the blockchain. The protocol is generic, i.e., it can be deployed ignoring underlying consensuses, and it is also superlight, i.e., the computational cost of the light client to predicate the (non)existence of a transaction in the blockchain becomes a small constant. Since our protocol resolves a fundamental challenge of broadening the usage of blockchain technology, it captures a wide variety of important use-cases such as multi-chain wallets, DApp browsers and more.

**Research Area:** Deep learning, blockchain technologies, IoT.

---

**Predictive Online Advertisement**

Online advertising is a multi-billion dollar industry. We aim to improve online ad efficacy and reduce ad-annoyance of users through predictive machine learning algorithms. We propose probabilistic latent class models that predict the viewability (in the technical sense of online advertising) of any given scroll depth for a user-page pair and deep learning models to predict the viewability of any page depth for any given user dwell time. Currently, we are studying techniques to find a balance between publisher’s revenue and user experience in the context of the ad-blocking battles between users and web publishers.

**Finding Free Parking in Cities**

Finding an available free parking space in a city during peak hours is challenging. We have created a system for assigning free curbside parking spaces to drivers in cities that reduces driving time to the parking spot and walking time from the parking spot to the destination. Currently, we are designing a distributed system for free parking assignment that is scalable and protects the driver’s privacy.

**Research Area:** Mobile computing and sensing, ad hoc and vehicular networks, cloud and disturbed systems.
Researchers are investigating multiple aspects of cloud computing. Reducing latency is one of the key aspects of the research and it includes use of parallelism and algorithms for scheduling.

A key result of this research is enabling applications to fully leverage the computing capability of hardware and devices to achieve optimal data processing speeds. Research results such as APPLES, a solution for reducing spin-lock overhead, have attracted interest from leading companies such as Tencent and Didi.
Enabling High-Performance Cloud Computing

The persistent uploading, downloading and processing of images, videos and files to and from the cloud can lead to inefficiency and delayed response times due to irregular computing demands. Our research focuses on live migration of virtual machines, as well as containers that will help alleviate the problem and improve cloud servers, such as Amazon’s Elastic Compute Cloud and Microsoft Azure. It will also help meet the power and computing requirements of mobile and enterprise cloud applications.

Jing Li
Assistant Professor
jing.li.2@njit.edu

Parallel Real-Time Systems

Real-time systems need to provide timing guarantees for latency-critical applications in cyber-physical systems that interact with humans or the physical environment. Examples span autonomous vehicles, drones, avionic systems and robotics, to structural health monitoring systems and hybrid simulation systems in earthquake engineering. However, as parallel machines become ubiquitous, we face challenges in designing real-time systems that can fully utilize the efficiencies of parallel computing platforms. We are developing parallel real-time systems by exploiting the untapped efficiencies in the parallel platforms, drastically improving the system performance of a cyber-physical system.

Research Area: Real-time systems, parallel computing, cyber-physical systems, interactive online cloud services architecture.

Andrew Sohn
Associate Professor
andrew.sohn@njit.edu

Scheduling for Interactive Cloud Services

Delivering consistent interactive latencies (i.e., response delays) is the key performance metric of interactive cloud services that significantly impacts user experience. The need to guarantee low-service latency, while supporting increasing computational demands due to complex functionalities of the services, requires parallel scheduling infrastructure to effectively harness parallelism in the computation and efficiently utilize system resources. Our research, for the first time, designs, analyzes and implements scheduling strategies that are measurably good and practically efficient to provide various quality-of-service guarantees on cloud service latency.

Alex Gerbessiotios
Associate Professor
alexg@njit.edu

Enabling High-Performance Cloud Computing

The persistent uploading, downloading and processing of images, videos and files to and from the cloud can lead to inefficiency and delayed response times due to irregular computing demands. Our research focuses on live migration of virtual machines, as well as containers that will help alleviate the problem and improve cloud servers, such as Amazon’s Elastic Compute Cloud and Microsoft Azure. It will also help meet the power and computing requirements of mobile and enterprise cloud applications.

Research Area: Adaptive computing infrastructure, autonomous migration of Linux virtual machines, maximizing cluster utilization.

Multi-Core and Many-Core Algorithm Design, Analysis and Implementation

We study models of computation that abstract and capture parallelism in the presence of multiple memory hierarchies and cores. New approaches are needed to make multi-core architectures accessible to software designers in domains such as machine learning and big data. Abstracting the programming requirements of such architectures in a useful and usable manner is necessary to increase processing speed and improve memory performance.

Parallel Computing Techniques in Sequential (AKA Serial) Computing

The norm in computing is to port sequential algorithms that work on one processor into multi-core or parallel algorithms intended for multiple cores and processors. Amdahl’s Law highlights the limitations of using multiple cores in programs with an inherently sequential component that is not amenable to parallelization. We address this by exploring the utilization of parallel computing techniques to speed up a sequential program by exploiting the multiple memory hierarchies present in contemporary microprocessors, even if its multi-core capabilities are left unexploited.

Research Area: Architecture-independent parallel algorithm design and implementation.
Virtualization of Heterogeneous and Non-Uniform Memory Hierarchies

Accessing data saved in heterogeneous and non-uniform memory hierarchies is an increasingly important factor for the performance of many applications. However, in the cloud, system software, particularly virtualization software, causes data-intensive applications to suffer a significant performance penalty. We improve memory virtualization technology to build virtual memory hierarchies. Virtual memory hierarchies have similar architectures and features as those of new memory systems. Thus, they can effectively serve as the portal for applications in virtual machines to efficiently access the data in new memory systems.

System Software for Scalable Computation in the Cloud

As the resources available for computation keep increasing in today’s computers (e.g., multi-core/many-core processors and accelerators), we need ways to scale the performance of these computers by taking advantage of the extra resources. The objective is to guarantee that applications in the cloud can achieve higher performance when presented with more resources.

Research Area: Cloud computing infrastructures, parallel and distributed systems, operating systems, database systems, computer architecture.
Augmented Reality, Virtual Reality, Gaming and Graphics

The research being done by faculty in our augmented reality, virtual reality, gaming and graphics groups is as diverse and innovative as one would expect from researchers engaged in these rapidly evolving fields. This group has faculty focused on the relationship between people, human cognition and augmented realities, and developing tools that improve user experience. Our research on the potential applications of this technology range from access to live-saving healthcare services, advances in crowd simulation that have implications for urban and space planning, to the use of virtual environments and simulation to improve learning outcomes, and ways to use AR/VR tools to increase efficiencies and support innovation in manufacturing settings.
Design of Virtual Environments and Interfaces to Support Information, Perception and Cognition

Our research examines the design of virtual and augmented reality hardware, software interaction techniques and applications to augment or change user thinking and cognitive performance. This research is done with teams within the distributed M.I.N.D. Labs working within the context of medical, scientific and defensive military applications. For example, in an international project collaborating with teams in Spain and Korea, we examine how a brain-computer augmented reality environment can support and trace complex consumer decision making, political decision making and psycho-therapeutic applications. Users of the AR environment are also instrumented (i.e., wear) untethered brain sensors (FNIR), other psychophysiological sensors (e.g., EEG, EMG) and behavioral sensors (e.g., facial emotion-sensing) to detect and measure how features of the virtual environment influence “thinking”. These “brain-sensing” measures (i.e., neurocognitive indicators) are used to adapt the virtual environment to the user in real-time, for example, a virtual character’s behavior or the relative amount/complexity of information within medical, decision-making or military applications. Other projects in this research stream involve virtual environments that change the perception of the user’s body or how social cognition is affected by immersive, hyperrealistic avatars during negotiation, training and decision-making. The design of virtual environments to augment “thinking” (cognition) also applies to interactive scientific visualizations. These AR environments make physical forces or microscopic phenomena experienceable by the senses and allow the user to perceive relations in data. For example, in a recent astrophysics project, we created an augmented reality environment that uses streaming data from head-worn, magnetic sensors to directly see and experience the earth’s local magnetic fields.

Research Area: Virtual and augmented reality systems, components for brain-computer interfaces, real-time public opinion measurement.

Virtual Human Teleportation

Virtual reality (VR) and 360° video are emerging technologies that can enable virtual human teleportation to any remote corner of the globe. This requires ultra-low latency Gigabit-per-second (Gbps) wireless speeds and hyper data-intensive computing. Our research investigates novel synergies at the intersection of 6DOF 360° video representation methods, edge computing, UAV-IoT and millimeter wave and free space optics wireless technologies that transmit data using much higher electromagnetic wave frequencies to enable the ultra-high data rates and ultra-low latencies required by next generation societal VR applications.

Real-Time Structure-Aware Reinforcement Learning

Reinforcement Learning (RL) provides a natural paradigm for decision-making in diverse emerging applications that operate in unknown environments and with limited data of unknown stochastic characteristics. Paramount to the effective operation of these ultra-low latency applications, such as IoT sensing, autonomous navigation and mobile virtual and augmented reality is the ability to learn the optimal operation actions online and as quickly as possible. Existing state-of-the-art RL methods either take too long to converge or are too complex to deploy. Our research examines novel structure-aware RL methods that integrate basic system knowledge to compute learning actions updates across multiple states or even the entire state-space of the problem of interest, in parallel. To address the challenge of computational complexity that is introduced, our methods integrate analysis that help effectively trade-off learning acceleration and computing complexity.

Societal Applications

Our research focuses on interdisciplinary synergies to enable next generation applications. For instance, an NIH project at the intersection of networked VR, artificial intelligence and low-vision rehabilitation that aims to enable novel, previously inaccessible and unaffordable healthcare services to be delivered broadly and affordably. Other projects include the integration of VR, real-time RL and soft-exoskeletons for future physical therapy and the synergy of UAV-IoT and VR towards next generation forest fire monitoring.

Research Area: Immersive communication, augmented/virtual reality.
**Computational Design of Virtual and Real Worlds**

Virtual worlds are growing in terms of complexity and interactivity and are challenging to construct efficiently and realistically. Interior scenes are an intrinsic part of virtual worlds. They are also needed for non-interactive applications, such as furniture marketing, residential showcasing and many other similar commercial applications. Unfortunately, most of the aesthetic and creative process of designing interiors, real or virtual, is mostly manual. To accelerate this process, we utilize machine learning and computer vision to develop several computational interior design tools. Our research assists users by providing suggestions on which furniture to pick, how well it matches the room's current style and where to place it. Hence, we can facilitate content creation needs for multiple industries and, most importantly, accommodate our common human needs to create our own optimal interior environment.

**Simulating the Motion of Multi-Agent Crowds**

Crowd simulation is an essential part of multiple domains, from interactive media systems, pedestrian analysis in evacuation scenarios and urban planning, to autonomous multi-agent path planning and coordination. A crowd is a collection of independent, self-actuated agents, where each agent has individual navigational goals in this shared environment. Agent movement is controlled by a navigation algorithm, which needs to ensure that an agent progresses towards its goal while avoiding collisions. Computing collision-free agent motion is difficult, due to the complexity of such dynamic interactions. Despite 30+ years in crowd simulation research, simulation methods have subspecialized and are computationally effective for either sparse or dense crowds but not for both. Our work is the first method that allows real-time simulation of both dense and sparse crowds for up to 100K agents. This result is made possible by reframing agent motion as a constrained mathematical optimization problem that can be solved effectively on modern GPU computer hardware. Aside from the immediate practical implications for the gaming and visual effects industry, these results inspire future research on large-scale pedestrian dynamics which can affect urban and space planning.

**Research Area:** Machine learning for visual computing, crowd simulation and dynamics modeling, computer graphics, computer vision.

---

**Creating Immersive Classroom Environments**

As augmented and virtual reality (AR/VR) technologies become more mainstream, educational technologies will increase use of AR/VR capabilities. Examining AR/VR’s impact on STEM education through the fields of learning science and human computer interaction allow for unique understanding of how these emerging technologies will impact our knowledge-based society. Identifying STEM educational use cases of AR/VR may help scope research efforts, and improve technology efficacy. As rich multimedia experiences become everyday facets of student lives, the delivery mechanisms of their educational content need to keep pace and supplement the traditional methods of instruction. Demonstrating that a game-structured virtual learning space can deliver comparable learning outcomes as a traditional, in-person learning environment has been a focus of the research, which was proven in recent work involving students taught binary counting in both settings. The learning outcomes from that study were the same, showing that students are capable of learning in a virtual, game-based environment as well as they do in an in-person setting.

**Research Area:** Interactive computer graphics for training and educational applications.

---

**Human Computer Collaboration**

We explore how humans and computers can work together to solve problems by conceptualizing creativity as a search in a structured space of computational artifacts. Our work has been applied to understanding game design, assessing learning in educational games, creating game content and behavior, as well as facilitating human creativity.

**Research Area:** Real-time systems, parallel computing, cyber-physical systems, interactive online cloud services, architecture.
Mixed Reality Simulations to Improve Training

Our research focuses on creating simulations using computer graphics, multimedia and mixed reality to improve training in different domains including healthcare simulation. We are especially interested in research involving virtual humans and multisensory experiences. We have developed a new class of augmented reality patient simulators called Physical-Virtual Patients (PVP) that allow healthcare educators to interact with a life-size simulated patient by providing real-time physical tactile cues (e.g. temperature, pulse), auditory cues (e.g. speech, heart sounds), rich dynamic visual cues, such as facial expressions (e.g. pain, emotions) and changes in appearance (e.g. skin color, wounds).

Research Area: Augmented, virtual, and mixed reality, 3D graphics, virtual humans and synthetic reality, modeling simulation and training.

Increasing Diversity in STEM through Mentorship

This project explores how to increase participation and diversity in computing, especially for underserved and underrepresented minorities in STEM. We partner with local nonprofits and schools to provide programming experience to K-12 students, specifically using near-peer mentors (college students) to teach and engage younger middle school and high school students in a 9-Saturday programming camp called Newark Kids Code. We also provide an introductory computing course for in-service high school teachers during the summer.

Research Area: Human-computer interaction (HCI), computing education research (CER).

Advancing STEM Education with Gaming and Virtual Reality

We explore unique approaches to teaching STEM (Science, Technology, Engineering and Math) topics in formal and informal learning environments. Our work on “Gidget” (helpgidget.org)—an online game to teach students introductory programming concepts—is effective in attracting and engaging a broad audience, including women and underrepresented minority groups in computing. Our work with CSpresso uses an interactive VR environment to teach middle school and high school students introductory computing science concepts such as binary counting and sorting algorithms.

Research Area: Geometric Modeling and Computational Design

The traditional separation of design and manufacturing is currently one of the bottlenecks in the product development process leading to multiple expensive feedback loops. Our research aims at developing advanced computational tools that integrate expert knowledge in 3D geometric modeling, physical simulation and smart algorithms in order to facilitate this process, which will become an important factor for success in increasingly competitive digital markets.

Computational Fabrication

We perform research at the intersection of applied mathematics, computer graphics and computational design. We investigate problems in geometric modeling and geometry processing, shape analysis and synthesis, as well as physical modeling and simulation with the goal of developing novel algorithmic solutions for the rapidly-growing domain of digital fabrication. Our research aims to provide novel solutions for the future design and manufacturing market.

Research Area: Computer graphics, geometric modeling, geometry processing, computational fabrication.
**Immersive Cross-Reality Applications**

Our research examines general areas of interactive Cross-Reality (XR), which encompasses virtual, augmented and mixed reality applications and serious game development. We specialize in eye and body tracking, as well as multi-sensory augmentations. Specific topics include the design, development and evaluation of novel XR and cross-model (visual, audio and/or haptic) user experiences through simulation (walking in European cities and driving in NJ and NY), as well as building augmented-reality collaborative applications.

**Multi-User Gaming and Collaborative Platforms**

Virtual collaboration has received a lot of attention recently as many people are forced to work away from their usual workspaces. Providing a realistic environment where people can reliably and efficiently collaborate on tangible objects and models will help many businesses. Primarily, this is relevant to city planners, military and law enforcement, as well as educational settings. We are also interested in a multi-calibration platform between various augmented and mixed reality devices such as mobile phones, Hololens and Magic Leap.

**Visualization of Large Datasets in VR and AR**

Large datasets such as ontology trees or visibility graphs, when loaded into virtual or augmented reality devices, can pose many challenges. For example, continuous loading of data into a mobile device. Similarly, there are no established methods for the most user-friendly way to visualize large data clouds. Hence, we combine various computer science algorithms with user-studies to develop the most efficient ways to visualize large datasets.

**Research Area:** Immersive and collaborative cross reality, navigation, gaze/body tracking.

---

**Margarita Vinnikov**

Assistant Professor

margarita.vinnikov@njit.edu
Advanced Algorithms

There is a wide variety of formal methods and algorithmic research at NJIT. The work includes optimization, risk assessment, numerical algorithms and program repair.

Researchers publish in premiere conferences and journals and have received numerous prestigious honors and awards, including Best Paper prizes at top venues and NSF CAREER grants.
Dynamic Graph Algorithms

Graphs are used extensively to model various kinds of networks, like transportation networks or social networks. In most of real life applications these networks change over time so their characteristics are changing as well. The goal of dynamic graph algorithms is to compute these characteristics over time as efficiently as possible. Clearly, the required output can be recomputed from scratch at each time point. However, in many cases, the slow pace of change relative to the size of the network enables much faster computation. One such characteristic is the Maximal Independent Set (MIS) of the graph, which has extensive connections to a plethora of fundamental combinatorial optimization problems, such as maximum matching, minimum vertex cover, and graph coloring. We have developed several dynamic algorithms for computing MIS including the first sublinear amortized update time algorithm for maintaining an MIS in dynamic graphs.

Location Problems on Euclidean Metrics

Location problems are an important class of combinatorial optimization problems that arise in several applications (e.g., choosing facility sites in a supply chain, placing servers in a telecommunication network and clustering data). The underlying distance function in many cases is Euclidean. It is natural to ask whether the Euclidean metrics can be leveraged to obtain more efficient algorithms than the ones known on a general metric space. We considered one such location problem, the classical k-Supplier Problem, and showed that, indeed, there exists an algorithm for this problem on Euclidean metrics that beats the lower bound on the time required for any such algorithm on a general metric space.

Research Area: Algorithms, mathematics of artificial intelligence (AI) optimization.

Spectral Network Analysis

Networks, also known as graphs, are objects of central interest in Data Science. A network can be mapped to linear operators whose spectral properties encode valuable connectivity information, enabling the design of novel algorithms for network analysis. These are, in turn, based on efficient algorithms for fundamental computational problems, such as systems of linear equations and other generalized regression problems. We have designed and implemented very fast and mathematically robust numerical algorithms in the context of network analysis. We have also contributed to spectral graph theory by developing a better understanding of spectral properties of graphs with a prescribed cut structure. These novel methods have potential for applicability. Our research includes concrete applications in data mining and machine learning problems on very large graphs.

Exact Algorithms

A multitude of well-studied computational problems are conjectured to require exponential time for their solution. Current research aims to develop a “fine-grained” understanding of the computational complexity of these problems beyond the classical NP-completeness theory. The design of faster “exact” algorithms for such problems and for their parameterized versions is of key importance in the area. In this context, we pioneered the general method of “algebraic fingerprints” that reduces various combinatorial problems to monomial detection problems that are, in turn, solved via algebraic algorithms. This research has led to breakthrough results for classical algorithmic problems, such as the Hamiltonian Cycle problem and single exponential time algorithms for problems parameterized by treewidth.

Research Area: Fast Linear system solvers, spectral graph algorithms, machine learning, data mining.
Science and Engineering in Program Repair
For the past decade, researchers in software engineering have been working on developing automated tools for program repair. In our research, we are interested in evolving theoretical foundations for this discipline and analyzing the impact of these on the state-of-the-art and the state-of-the-practice in program repair.

Redundancy: The Mutants’ Elixir of Immortality
Mutants are syntactic variants of a base program used in software testing to assess the quality of test data. Test data is good if, whenever we modify a program, the test data exposes the difference. Sometimes a mutant may be semantically equivalent to a base program despite being syntactically distinct. In this work, we analyze the probability of equivalence between a base program and a mutant, or between two mutants, and we use this probability to analyze the mutation properties of the base program.

Research Area: Software engineering, high assurance systems engineering, discrete mathematics, software metrics.

Global Optimization
To solve the pervasive optimization problems in engineering, science and commerce, we are developing “global optimization” algorithms where the objective is to solve optimization problems without getting stuck in local minima. This has applications in the design of fuel-efficient aircraft, the error rate of classification algorithms and financial investing.

Research Area: Global optimization, probabilistic analysis of algorithms.

Efficient Computational Methods for Risk Assessment
Many disparate fields suffer from uncertainties with detrimental consequences, such as large losses in financial portfolios or failures of critical infrastructure due to natural disasters. Modern society crucially depends on gaining a better understanding of the likelihood and impacts of such calamitous rare events. Our work devises novel approaches for substantially reducing statistical errors in Monte Carlo simulation, a computational technique that can be employed to study risks in decision-making and analytics. One project focuses on designing improved methods for probabilistic safety assessments of nuclear power plants. The work also applies to evaluating the reliability of complex systems (e.g., navigation computers of aircrafts or package-tracking systems for overnight-delivery companies) and the dependability of supply chains.

Research Area: Monte Carlo simulation, risk analysis, applied probability, statistics.

Algorithm Design to Combat Discriminations in Rideshare Services
Rideshare platforms, such as Uber and Lyft, have gained increasing popularity in recent years. One of the central tasks facing Uber and Lyft is the matching policy pairing drivers and riders. Recently, it has been reported that drivers cancel riders based on their demographic attributes such as gender, race and disability, either intentionally or unintentionally. In our project, we try to leverage the power of algorithm design to curb discriminative cancellations from drivers to riders and improve the social welfare overall.

Research Area: Artificial intelligence, approximation and randomized algorithm, rideshare and crowdsourcing markets.
Social, Mobile and Education

Social media and mobile communications have transformed how people interact with one another and how they access information. The social, mobile and education group conduct research that looks at how users engage with these platforms and technologies and examine key issues and opportunities that our increasing reliance and use of these platforms can create. For instance, a grant from the Mozilla Foundation examines how female and LGBTQ live streamers cope with online harassment, while other research focuses on the dissemination of disinformation via social platforms.

Our researchers focused on education are working to improve learning outcomes, automate instructor feedback and examine the effect of mobile technology on student learning and education delivery.

The crisis informatics work was one of three papers that received a best-paper nomination at this year’s ISCRAM conference.
Exploring the Link Between Technology and Well-Being

We are exploring the connection between health and technology usage with an emphasis on the role social media plays in facilitating social support and psychological well-being, such as understanding the impact of mental health disclosures on live streams and how people deal with online harassment. We also look at how enclosed spaces, such as windowless offices, influence well-being and are designing technologies to augment physical spaces.

New Digital Economies

Games, esports and digital patronage systems have novel digital economies with unique digital currency. On the surface, they appear as recreational activities but represent a complex financial ecosystem that could have implications in the future of the digital economy. Our research focused on spending behaviors and exchange patterns of virtual goods in games and other alternative financial platforms and how these activities are tied with creative content generation.

Social Media Content Moderation

Supported by the National Science Foundation, this work focuses on the work of volunteer moderators on various social media platforms and understanding how safe spaces develop online.

Research Area: Content moderation practices, teamwork in e-Sports, social media usage and well-being.

Brook Wu
Associate Professor
yi-fang.wu@njit.edu

Early Detection of Fake News on Social Media

A major challenge of effective and early detection of fake news is fully utilizing the limited data observed at the early stage of news propagation. We propose a novel deep neural network to detect fake news early mainly by combining user- and post-based features into status-sensitive crowd responses. Experimental results show that our proposed model can detect fake news with greater than 90% accuracy within 5 minutes after it starts to spread and before it is retweeted 50 times. Most importantly, our approach requires only 10% labeled fake news samples to achieve this effectiveness under PU-Learning settings. We plan to extend this work by incorporating additional social context data extracted from user interactions to further enhance user representations and prediction accuracy.

Automating Feedback for Meaningful Learning

Writing-to-Learn (WTL) is a pedagogical strategy which uses writing to improve students’ understanding of course content. However, most existing feedback systems for writing are mainly focused on improving students’ writing skills rather than their conceptual development. We have developed an automatic approach to generate timely, actionable and individualized feedback to facilitate meaningful learning by comparing knowledge representations extracted from teaching materials and individual students’ writing assignments. Our results show that students include more domain concepts and relationships from the course materials in their assignments when using our system. We plan to extend this work to generate feedback for instructors to reinforce the discussion of core concepts in class.

Research Area: Text mining, information extraction, information retrieval.
Open-Source Book on IT Concepts and Applications in Three-Tier Architectures

The research presented in our book aims to contribute to the open-source educational literature on three-tier environments, specifically from the viewpoint of undergraduates specializing in information technology and computer science. While the book considers various applications and technologies that arise in a full-stack environment, the underlying focus and objective is not on web programming but on exploring the technologies, concepts and development challenges that occur in these very complex environments.

Research Area: IT education.

Detecting Multi-Platform Online Disinformation Campaigns

This project develops multi-platform, multi-modal analytics tools for comparing behaviors of potentially malicious online actors with an eye toward detecting inauthentic coordinated action and foreign influence campaigns in online social platforms. Current results suggest such malicious campaigns are consistent with prior work on propaganda but exhibit a distinct signal in how dissimilar the general content they share appears, suggesting an effort to artificially inflate the perceived diversity of their ideas.

Research Area: IT education.

Retrieving Critical Information in Social Media Data During Crises

Social media, blogs and other online information sources contain large volumes of data, especially during, and in the aftermath of, crises. This research project is part of an ongoing effort with the Incident Streams track at the National Institute of Standards and Technology (NIST) and works to develop standardized datasets, tools and machine learning models for classifying social media data by information types and priorities. We use these systems to inform emergency response officers and support their efforts to find and respond to the non-trivial amount of critical information people share on social media during times of crisis.

Research Area: Social media mining, crisis informatics, political engagement, disinformation.

Redefining Student Engagement

Participatory learning deepens knowledge through active involvement in the entire lifecycle of assignments, exams and other course activities, including peer grading and self-assessment. We are developing a web application that functions as an educational dashboard, integrating commenting and mentoring capabilities and automatically assigning tasks to students.

Wearable Technology in Classrooms

Classroom dynamics continue to evolve as more students will use their personal wearable technology (watches, wristbands and contact lenses) in ways that are undetectable by instructors and peers, which would create an atmosphere of distrust. We are the first to systematically investigate and determine how best to design learning, teaching and assessment when personal wearable technology is used undetectably in classroom settings.

Research Area: Web engineering, link-based services, hypermedia, community informatics.

Open-Source Book on IT Concepts and Applications in Three-Tier Architectures

The research presented in our book aims to contribute to the open-source educational literature on three-tier environments, specifically from the viewpoint of undergraduates specializing in information technology and computer science. While the book considers various applications and technologies that arise in a full-stack environment, the underlying focus and objective is not on web programming but on exploring the technologies, concepts and development challenges that occur in these very complex environments.

Research Area: IT education.

Detecting Multi-Platform Online Disinformation Campaigns

This project develops multi-platform, multi-modal analytics tools for comparing behaviors of potentially malicious online actors with an eye toward detecting inauthentic coordinated action and foreign influence campaigns in online social platforms. Current results suggest such malicious campaigns are consistent with prior work on propaganda but exhibit a distinct signal in how dissimilar the general content they share appears, suggesting an effort to artificially inflate the perceived diversity of their ideas.

Research Area: IT education.

Retrieving Critical Information in Social Media Data During Crises

Social media, blogs and other online information sources contain large volumes of data, especially during, and in the aftermath of, crises. This research project is part of an ongoing effort with the Incident Streams track at the National Institute of Standards and Technology (NIST) and works to develop standardized datasets, tools and machine learning models for classifying social media data by information types and priorities. We use these systems to inform emergency response officers and support their efforts to find and respond to the non-trivial amount of critical information people share on social media during times of crisis.

Research Area: Social media mining, crisis informatics, political engagement, disinformation.
COLLEGE OF COMPUTING FACTS AND FIGURES

DEGREE PROGRAMS

Undergraduate
B.S. Business and Information Systems
B.S. Computer Science
B.S. Computing and Business
B.S. Data Science*
B.S. Human-Computer Interaction
B.A./B.S. Information Systems
B.S. Information Technology
B.S. Web and Information Systems

Graduate
M.S. Computer Science
M.S. Information Systems
M.S. Computing and Business
M.S. Business and Information Systems
M.S. Software Engineering
M.S. IT Administration and Security
M.S. Cyber Security and Privacy
M.S. Data Science

Doctoral Degrees
Ph.D. in Computer Science
Ph.D. in Information Systems

*pending approval

FACULTY

Tenure-Track
42

Instructional Lecturers
27

DIVERSITY

20% of Students are Women

21% of Students are Underrepresented Minorities

ALUMNI

Number of Annual Graduates
800

Average Starting Salary
$73,748

TOTAL STUDENTS 3,119
BS 2,256
MS 753
PhD 110