Senior Design Instructors for 2020-2021

Derek Aguiar, Ph.D.   Ion Mandoiu, Ph.D.
Hanna Aknouche-Martinsson, Ph.D.  Fei Miao, Ph.D.
Mukul Bansal, Ph.D.   Jacob Scoggin, Ph.D.
Jinbo Bi, Ph.D.   Dong-Guk Shin, Ph.D.
Phillip Bradford, Ph.D.   Wei Wei, Ph.D.
Benjamin Fuller, Ph.D.   Yufeng Wu, Ph.D.
Seung-Hyun Hong, Ph.D.

Contact Information

Joseph Johnson, Ph.D., Associate Professor in-Residence,  
Associate Director of Undergraduate Programs in Computing and  
CSE Industrial Liaison  

joseph.2.johnson@uconn.edu

Sanguthevar Rajasekaran, Ph.D., Professor and Department Head  
sanguthevar.rajasekaran@uconn.edu

University of Connecticut  
Computer Science and Engineering  
371 Fairfield Way, Unit 4155  
Storrs, CT 06269-4155  
Telephone (860) 486-3719  
Web: http://www.cse.uconn.edu/
HDF5 Interpretation and Visualization

The sponsor of this project, the Carrier Corporation, uses state-of-the-art technologies for model-based systems engineering (MBSE) of complex systems. One of the model-based design toolkits employed by Carrier is a software called Sandia Dakota. The Dakota software uses a file format known as HDF5 to contain the outputted information from the analysis results. Dakota’s HDF5 output files have a difficult structure to follow, in turn producing results which are difficult to interpret. Carrier is looking to develop a method by which they can automatically process such HDF5 files and make more efficient use of the information contained within them. Leading into the project which the team has been presented by the company: develop an application which can seamlessly parse an HDF5 file produced via the Dakota application, and use the information contained within the file to provide the user with an interface in which they can better analyze and visualize the information contained within. The application is a desktop application which is capable of reading and processing HDF5 files and utilizes web-based technologies in order to create a pleasant and intuitive interface allowing users to extract significant information from HDF5 files through producing versatile plots and visualizations of the data contained in various ways. The application will allow the Carrier team to easily visualize the results of methods, compare and contrast variables, find optimal values, and more. The application grants users the ability to create a variety of custom visualizations from the datasets contained within a wide-array of user provided HDF5 files. This application was developed using the Electron software framework, React and Bootstrap components, as well as the H5PY and Plotly libraries for data processing and visualization generation.
SmartEMR Using Clinical Data Integration

SmartEMR is a web-based application designed to empower doctors, nurses, and other healthcare professionals to gain better insights into stored electronic medical records, or EMRs. SmartEMR leverages powerful machine learning tools to provide analysis on key free-form data types, including clinical notes and medical images. The system is implemented through a microservice architecture and exposed to the user through a secure web interface. The frontend for SmartEMR is built from the ground up for usability, security, and scalability. The site uses a simple tab-based navigation scheme to access various data analysis services, minimizing the learning curve for users to begin analyzing their patient data. The first of the system’s microservices is a natural language processing service. The service is mainly accessed via the SmartQuery engine, which prompts the user for plain-English input queries and converts them to formalized SQL statements that are then executed and displayed. In addition, the service can tokenize, or parse, key prescription information from free text. The key function of the analytics web service is to create insightful data visualizations from plain-English queries. This service works in tandem with the natural language processing service to create the user-experience present in the SmartQuery engine. The imaging microservice is responsible for storing, updating, classifying, and returning relevant medical images. The site uses several established endpoints to allow for the uploading of medical images to new or existing patient profiles, to query images using natural language, and to classify medical images and return their relevant descriptors. The final microservice in the SmartEMR architecture is a clinical web service designed to generate sophisticated sample patient data. The service uses an algorithmic data creation scheme designed to produce extremely realistic patient histories that use genetics to predict physical attributes.
Unary Cipher

The unary cipher shown in this application is a variation of the very popular Vernam Cipher or One-Time Pad. This ciphertext is absolutely secure, no matter the strength of the computer trying to crack it. This is because the ciphertext does not commit to the plaintext. As an example, if a message of "Hello World" is sent, the Unary Cipher may create a ciphertext that when deconstructed with the wrong key, will result in the message "I'm hungry". Due to the number of different keys that can be used to decrypt to meaningful plaintexts, even if a computer were to find every possible key, there is no guarantee that the one chosen will even be the correct plaintext. The best part of Unary, is that the same key can be used over and over again. This cipher is convenient, useful, and most of all, very secure.

To learn more about Vernam: https://en.wikipedia.org/wiki/One-time_pad

To learn more about the cipher, you can read the full paper here: https://eprint.iacr.org/2020/389/20200524:211603
Covid-19 Contact Tracing and Stochastic Risk Analysis App

Our group sought to develop a Contact Tracing and Risk Analysis application that was aimed at providing a community like a university with a tool to combat the spread of Covid-19. To achieve this objective and lower the barriers to adoption, the application was designed to work on both Androids and Apple iPhones. Rather than writing two separate sets of code, our group used an open-source framework called Nativescript that is used to develop mobile apps on iOS and Android simultaneously. The application was written in Typescript, JavaScript, XML, and SCSS. The group also used the Google Firebase platform for user authentication, and as a NoSQL Realtime database. User authentication pertains to verifying the identity of a user through account creation and login. In terms of what was accomplished by the team this semester, users are able to submit a variety of data forms that ask about their Covid-19 and flu vaccination status, Covid-19 symptoms as outlined by the CDC, and the results of any Covid-19 tests taken. Users are able to query these data forms by specifying a maximum quantity of data forms to retrieve, the type of data form to retrieve, and a date and time range between which to query the database. Users are then able to review the results of these queries. Accounts are associated with a unique ID generated by Firebase. All data in the database is stored with the ID as part of the path. This was done to make it efficient to perform a query. The entire database is not searched when a query is performed. The query is localized to a region of the database using the unique user id. A logical next step for the project would be to use Bluetooth to log when two authenticated users are in close proximity for a duration of time above a threshold such as 5 or 10 minutes. Then users can be informed at some point in the day that they have been in close proximity to someone who has been exhibiting symptoms of Covid-19 or who has self-reported a positive Covid-19 test.
Our project, Lab Training Simulation, is a virtual reality lab environment where users can perform either of two labs. We implemented two different experiments. The first was a Biology lab: Transpiration & Photosynthesis that allows users to learn the process of photosynthesis. We also implemented a Chemistry lab: Distillation of Cyclohexane-Toluene from which users can learn how a mixture of two different solutions can be separated by heat.

To implement this project, we mainly used Unity, an established, free platform for XR development and Blender, a free and open-source program for building 3D models from scratch. In Unity we were able to create a lab environment, add our assets (lab equipment) and write scripts to take the user through the steps of the lab. We used blender to make specific lab equipment that was not already available to us in Unity’s free asset store. We also used Windows Standalone as our platform to export our project. Users are able to download the executable from a website and then run the program locally on their device.

We wanted to make this project the best possible learning experience it could be for the user, so we focused on making it interactive and adding different learning components. These learning components included two different modes: a tutorial mode and a quiz mode. In either mode the user would be able to interact with the lab equipment and complete the steps themselves. The tutorial mode outlines the steps for the user to follow. The quiz mode allows the user to complete it themselves and tells the user if they’ve made an error and need to start over.

Apart from just making our project a quality learning experience, we found a couple of key benefits to the virtual reality lab setting including. These benefits include visualization of safety protocols and lab procedures, speedup of long experiments or slowdown of fast experiment, and studying how modifying the procedure impacts the experiment without risk of causing harm.
Machine Learning for Cybersecurity Applications

In this project we created a machine learning algorithm that, when given the log files from a server, can tell if a server is under attack from certain cyber security threats. The cyber threats our model can detect are SQL injections and Denial of Service attacks. Our project has potential to link two important and growing fields within computer science.

For our project we set up a Tomcat Apache server running a simple pet clinic website. This website allowed users to enter data in a database, such as the name and address of clients, and search that same database. We then attacked this website using Kali Linux tools and had python scripts running to simulate normal usage. During both attacks and simulated normal usage, JavaMelody was used to capture the server logs to feed to our machine learning algorithm.

We preprocessed our data to only contain what we found to be the necessary attributes from the server logs. Doing this helped the model not be distracted by attributes that had no correlation with being in an attacked state or normal state.

Once data was generated and preprocessed, we trained a Long Short Term Memory (LSTM) model to identify whether the server was in an attacked state or normal state. A LSTM model is a recurrent neural network which will hold onto a data point for an arbitrary amount of time depending on the specific cell in the network. This is great for our project since it is important for the model to remember what attacks which took place a long time ago look like and what recent attacks look like.

Finally, after data is run through the model, we scored how the model performed using a variety of methods, such as the receiving operator characteristic (ROC) curve and F1 Score, and visualized it using principal component analysis (PCA) so that it could be viewed in two dimensions despite having over 20 dimensions. When our project is running in real time, it can deliver a notification of when a server is under attack based upon our pre-trained model.
Enhanced Campus Tour

Touring a college campus can make or break a prospective student's decision to attend a university. With the COVID-19 pandemic, it is harder for those students to go out and visit the university they dream of attending. With the Enhanced Campus Tour, it is easier for students to view the UConn campus from the safety of their home. Using virtual reality technology, students will receive a similar experience as if they were participating in person. The user will be able to tour campus and view popular buildings through a web application developed in Unity. Throughout the tour, when you look at a building or structure that is commonly used, an informational display will appear where you can learn more about that building. This application allows the user to be able to access the tour on any device and anywhere with an internet connection. The Enhanced Campus Tour will allow students to experience life at UConn while remaining socially distant and keeping our fellow Huskies safe.
DevOps Pipeline for Kuali Financial Systems

Our project for UConn Information Technology Service (ITS) focuses on their Development and IT Operations (DevOps) pipeline for Kuali Financial Systems (KFS). KFS is a suite of financial software built on open-source platforms for higher education. The ITS DevOps team manages and modifies the KFS environment by tracking issues, programming features, and then the final step of deploying these features using a pipeline. The pipeline begins with code in a central repository and an automated job that then compiles, builds, and deploys the updated repository to their application server. The goal of this project will be to enhance this pipeline to include unit and integration testing, issue tracking, artifact management, and automation. These changes will deploy code versions in a more stable manner as well as automate the entire process to faster deploy and/or troubleshoot upcoming changes. The DevOps Team’s motivation behind this project was to merge technologies within the Atlassian Suite, though some other open-source technologies will also be seen integrated within this pipeline. Merging functionality across Atlassian products allows the DevOps team to more comfortably connect and manage their code pipeline.
Cigna Healthcare Literacy Enhancement

The goal for this project, as presented by Cigna, was to create a digital solution to increase healthcare literacy and/or digital engagement, especially among underserved, low income populations. Our team developed the Prescription Interpreter, an Android mobile application which scans prescription labels utilizing the mobile camera, translates the image to text using a custom made machine learning algorithm, and displays visual medical adherence instructions. The Prescription Interpreter was designed to emphasize clarity, ease of use, accessibility, personalization, and diversity. Complex medical terminology is avoided in favor of visual icon-based directions. To promote personalization and inclusion, the visual directions feature an avatar which can be customized for skin tone, hair color, and hair style, so that users will feel represented regardless of their background. Lastly, our application’s most innovative feature separates it from other applications on the market with an ability to automatically interpret how and when a user should take their prescriptions via machine learning solely based on a picture of the prescription’s written directions, even when the directions contain some degree of spelling errors.
Connecticut National Guard Cyber Range

The Connecticut National Guard (CTNG) Cyber Range is a cloud-hosted cyber range to facilitate training for its personnel and partners. In this environment, the CTNG must be able to create a scenario to faithfully replicate either a critical infrastructure or local municipality’s network and provide either hosted or remote-connected adversaries. (i.e. the CTNG provides the threat, Red Team, or facilitates bringing a Red Team via remotely connected systems). The range must provide a realistic set of threats, vulnerabilities and exploits that mimic real-world attacks and provides a means for “blue team” systems to be added to the network, post exploitation, simulating a cyber incident response. The end state is a well-orchestrated storyline that ties into a realistic set of attacks that mimic threats from script kiddy, Low Level hackers through advanced persistent threat and nation-state actors.
Sonalyists Interface Crowdsourcing Environment

The Sonalyists Human-Autonomy Interaction Laboratory (HAIL) conducts research, development, test, and evaluation of Human-Machine Interfaces (HMIs) for a variety of systems for the Navy, Space Force, Air Force, Army, and more. The UConn student team developed the Interface Crowdsourcing Environment (ICE) to digitize the crowdsourcing of HMIs and enable more rapid processing and analysis of design data from diverse groups of participants. The result is better HMIs for applications such as mission-critical Department of Defense systems.

A multitude of methods are used to design HMIs, including interviews, focus groups, surveys, and job observations. One additional method is to 1) provide end users a blank paper outline of the displays and a list of HMI components that must be included and 2) have users draw how they would ideally like the HMI to look. Analysts then manually recreated each paper drawing in PowerPoint and manually assigned colors to different components. Finally, they adjusted transparencies until heatmaps are generated showing where each HMI component should go. While this manual approach yielded valuable insights, it was labor intensive, error prone, and could not explore differences across demographic groups.

ICE is a web-based platform that consolidates and automates the planning, collection, analysis, visualization, and exploration of this information crowdsourced from a large and diverse group of users. ICE enables simpler and more cost-effective data collection at greater scale, drastically increasing the quality of analytics and design of HMIs. ICE fully removes the need for paper/pen data collection, automatically processes and analyzes data based on common methods, and gleanes insights through interactive visualizations, where analysts can rapidly toggle data on or off based on demographic features of users or other parameters. This application will greatly improve analyst workflows and empower analysts to design better HMIs.
Wildfires are a destructive force that can be created by humans and nature. They are challenging to predict, although they are most likely to occur in areas that are very dry. For this project, we are going to focus on wildfires in California because they are common in that area, and due to climate change, they are occurring more frequently and are spreading over larger areas.

This project, Trackit!, is a web-based wildfire predictor application. It allows users to interact with a machine learning model that predicts the likelihood of wildfires in California. Users will click on a location marker pin on our interactive map, and a pop up will alert the user of the risk level of a wildfire in the selected area in the next two weeks. They can use this to help ensure their safety when traveling to or within California. We trained a deep neural network using historical wildfire data its associated weather data to predict the probability of a wildfire occurring. The weather data was web scraped from the Farmer’s Almanac website, which provides accurate historical data. The wildfire data was exported from the National Oceanic and Atmospheric Administration (NOAA) website.
Mental Health & Wellness Application

The Mental Health & Wellness Application was created to provide easier access to SHaW’s Mental Health Services for UConn students. The application primarily features a chat service, allowing UConn students to connect with SHaW professionals through messaging at any time. This service is targeted towards those looking to seek advice not warranting a full session with a therapist. Due to COVID-19, it has become more difficult to seek help in person. The alternatives (over the phone and Zoom calls) can also be extremely daunting. This can result in less individuals seeking help when they need it the most.

The project consists of two parts: a patient website and a therapist program. The patient website will be utilized by UConn students and was created using the Django framework and Python for its backend programming. In order to chat with therapists, users of the site have to be registered to ensure that they are UConn students. Once logged in, the user will gain access to links to the chat rooms of all available therapists. When not utilizing the chat rooms, users can find articles, programs, and other useful services in the resources section of the site.

The therapist program will be utilized by SHaW mental health professionals. The Graphical User Interface (GUI) for this program was created through Python’s Tkinter module. When the program is opened, the therapist user is prompted to either log in or register a new account. Once logged in, the therapist can indicate their ready status to the site through a toggle button at the top of the screen. Once they indicate they are ready, the therapist can open up their chat room using the messaging button. When they are not chatting with patient users, the therapist can utilize the calendar feature to input their availability and view times they have previously indicated to be working. During sessions, the therapist can also utilize the notes functionality to write down important information they may need in the future.
Tragedy App

The Tragedy App is an interactive web game accessible through internet browsers that is meant to allow individuals to learn and experience an economic problem known as The Tragedy of the Commons. In this game, users can become either a host or a player. Hosts are in charge of administering games, while players join the games created by the host and play the game. During each round of the game, each player is given a set number of resources that they are supposed to allocate between “Farming”, and “Pasturing”, with any unallocated resources sent to the “Reserve”. At the end of each round, each player is scored based on their choices. However, in the event that the cumulative allocations to “Reserve” pass a certain amount, they could lose points instead.
The team was tasked by General Dynamics Electric Boat to develop a communication program that would be used to achieve underwater wireless data transfer. The team elected to leverage the Transmission Control Protocol, also known as TCP, to transmit data. TCP allows the program to form a connection between two hosts. Once this initial connection is made, data can continue to be exchanged over this connection. The team knew security would be a top priority and to ensure that the data being exchanged between the two hosts couldn’t be accessed and read by a third party we implemented TLS or Transport Layer Security on top of TCP. TLS uses two personalized keys and certificates which are used to encrypt the data being sent. This encryption is the key to the security of the program. With a secure connection, the team then focused our attention on potential data errors that would occur from transmitting data underwater. To combat this, we implemented Reed-Solomon error correction into the connection. This detects and corrects errors in the data so that the receiving party gets the correct data even in error prone environments.
Smart EMR

An Electronic Medical Records (EMR) is a collection of digital medical information corresponding to a patient from various clinics. The goal of this project is to integrate clinical data from Epic and dental data from axiUm into an open source EMR management program called Smart EMR. The clients are clinical researchers who will have access to Smart EMR with a key for both axiUm and Epic in order to retrieve information corresponding to different patients. These researchers will be able to access the data within Smart EMR through a user-friendly web User Interface (UI). Additionally, researchers can query the data through regular sentences which are parsed with a natural language parser. The resulting data is examined by an AI model that creates suggestions and different visualizations corresponding to the type of data returned. This way, researchers have an easier time creating inferences and finding associations between diseases and genes that may have been missed otherwise.

The data for this project has been managed to meet various constraints and to handle various types. Privacy has to be maintained due to HIPPA compliance rules within the medical field, thus all of the test data is fake. Current biological data sets are based around dental and skeletal features, but our system is extensible and implementation of other biological sets is still possible. Unstructured and structured data, such as a doctor’s narrative or lab data respectively, have been considered and handled too through the usage of MySQL and MongoDB. All structured patient information such as lab data and gene mutations are stored within MySQL. All unstructured data types, such as DICOM images or doctor notes, are stored within MongoDB. Finally, all forms on the web pages check for validity in order to remain secure and prevent cyber-attacks.
StockU - Algorithmic Trading Made Easy

Team 18's Senior Design Project is a functional implementation of an educational algorithmic stock trading platform called StockU. This platform brings the benefits of algorithmic trading to the average person, without a need for coding. StockU features a no-code algorithm builder, live stock data feeds, and an easy to navigate UI for a quality user experience. The brokerage service and API used for live stock data comes from Alpaca LLC, which is a FINRA licensed brokerage service and data provider. Through OAuth verification, we are able to have users sign in using their Alpaca account and maintain secure user data retrieval.

In the algorithm maker, users are able to drag and drop blocks which allow them to construct conditional statements. An example of a conditional statement would be “buy stock in AAPL if the closing price is 5% higher than the opening price, and sell if the closing price is down 10%.” Users can make up to five algorithms like this, save them to their profile and receive email notifications if an algorithm is triggered and makes purchases. As this is just a prototype, all purchases and sales are done through a “paper trading account”, meaning that they will be treated like a real account, with balances adjusting and transactions going through, but no real money will be spent. However, Alpaca does give us the capability of accepting real transactions in the future if this prototype is taken into production.

This application was developed using a MERN stack (MongoDB Atlas, Express, React.js, and Node.js). The database itself is accessed using a Python API developed using Flask while the OAuth system is built on NodeJS. Using Netlify and Google Cloud, these APIs and the front-end are publicly hosted for a fully functional application.

If you have any questions feel free to email us:

Rohan Parikh – rohan.2.parikh@uconn.edu
Ben Levy – ben.levy@uconn.edu
Michael Allen – michael.g.allen@uconn.edu
Henry Nguyen – henry.d.nguyen@uconn.edu
Samantha Grubb – samantha.grubb@uconn.edu
The purpose of our project was to create a piece of software that displayed voting record information in such a way that had high confidence, easy readability, and was user-friendly. Before the completion of our software, the UConn VoTeR center used a large CSV file to audit ballot results which was very difficult to read, but our software was designed to fix this. As a group, we designed a product that allowed the user to read in a ballot, parse the data, and display the information in both a graphical format, as well as a table format. We employed the use of drop-down menus to allow the ballot information to be split up into different batches which are pre-defined groups of ballots, as well as individual ballots to allow easy readability of confirmed and questionable votes in each race or even for each candidate. Our product will be delivered to The UConn Voting Center at the end of the semester and will be incorporated this summer to be used by staff and interns.
Securely Streaming PMU Data To The Cloud

We worked with ISO-NE for our senior design project. ISO-NE is a non-for-profit organization responsible for managing the power grid in New England. To give a little background, ISO-NE has on-premise systems that their control centers rely on for real-time operation. These systems are currently configured for failover, however they may still be subject to system failures, thus a cloud based backup solution was needed. The data that control centers rely on is called PMU data, which are sourced from Phasor Measurement Unit (PMU) devices located on the electric grid throughout New England. They provide time-stamped measurements of voltage, current, and frequency and are important because they reflect real system states.

Our group was tasked with securely moving this PMU data to the cloud from the on-premises infrastructure. For security we utilized the TLS protocol and for our cloud service we used Amazon Web Services (AWS). Our project can be divided into 4 components, the PMU Adapter, the Kinesis-DynamoDB Adapter, the Database, and Kinesis Data Analytics. The PMU adapter receives data from the PMU devices, parses the data accordingly and pushes it to a Kinesis stream. The Kinesis-DynamoDB Adapter processes the data from the Kinesis stream and writes the data to DynamoDB, which is where the PMU data ends up being stored. This is our main workflow. For deployment, the PMU Adapter and Kinesis-DynamoDB Adapter are written in C++ / NodeJS and Java respectively, both have a Docker image, and both will be deployed using the ECS container service on AWS. Kinesis Data Analytics is integrated with the Kinesis stream, performs analytics, and sends them to another DynamoDB table. This allows us to monitor system status and detect issues in real-time.

The featured image is a diagram of our whole workflow, figure 1 expands on the PMU Adapter workflow, and figure 2 expands on the Kinesis-DynamoDB Adapter workflow.
The current COVID-19 pandemic has displaced us from our classrooms to our homes while relying on our smartphones to stay connected. Transitioning to distance learning has interrupted our ability to make new connections that are essential to our college experience.

Our UStudy smartphone application facilitates social interaction and connections between students on the UConn campus. The UStudy platform allows each student to enter course-specific chat rooms to ask questions, make study groups, and remain united as a classroom again from virtually anywhere.
Herzberg Denial of Service De-Amplifier

The goal of this senior design project is to create an interactive simulation of a SYN-ACK flood denial of service attack on a network topology where a defense mechanism being built by Doctor Amir Herzberg and doctorate student Anna Mendonca has been deployed. Their defense mechanism, consisting of a DDD filter, control mechanism, and port changer, relies on changing the client’s source port as it appears on outgoing packets to hidden ports. The goal of the simulation is to provide a proof of concept to demonstrate that such a design would successfully be able to mitigate the attack, provide data about the attack as it progresses in the simulation, and make it accessible to a broader audience, such as students and other interested researchers, by providing both an animation of the flow of data in the simulation and a few graphs showing the behavior of the simulation over time. Users can specify almost all attributes of the topology, such as the bandwidths between all of the network devices, the sizes of their queues, data transmission and retransmission rates, and whether the defense mechanism is active or not to allow for comparison between the attack being execute with a defense and without a defense. The goal is for this simulation to provide both an educational opportunity to help individuals learn about the defense mechanism, as well as to provide useful data for when the system is complete.
Synchrony POST

The Synchrony POST project entails a proof-of-concept application designed to streamline the functionality of Synchrony’s existing CareCredit mobile application by simplifying patient-provider interactions. For context, CareCredit is a credit card that allows patients to finance expensive medical expenses. The accompanying CareCredit mobile application provides customers with a digital credit card for making secure payments to their health care providers. This project aims to prototype and develop potential enhancements to this application. One such enhancement includes adding the ability for users to check in to their providers using Near-Field Communication (NFC) and geolocation technology to allow for easy and secure hands-free billing. Currently, the CareCredit app is designed to primarily be used by end-users, so to facilitate the enhanced check-in and payment process, an accompanying interface for CareCredit health care providers was developed. Additionally, to help promote usage of CareCredit, a machine learning capable recommendation algorithm was developed to predict which CareCredit providers a user would be most likely to be interested in. To accompany all of these proposed features to the CareCredit mobile application, a full set of backend services were architected and developed. Considering the sensitive nature of payment data, a fully fleshed-out security implementation was also developed as a core feature rather than an afterthought. In addition to being secure, all backend services follow a microservice architecture pattern to ensure the scalability of all proposed functionality.
Predicting Drug Side Effects from Chemical Signatures

Drug development is an expensive process that requires major investments of time and resources in drugs that may turn out to exhibit severe side effects during clinical trials. To accelerate drug discovery, many machine learning methods have been developed to predict side effects of candidate drugs. However, available data processing pipelines are difficult to execute, evaluations of methods are opaque, and comparisons between them are sparse.

In our project, we implemented eight side effect prediction methods and wrapped them into a user-friendly API. To increase the generality of our analyses, we processed two independent datasets. One is the FDA’s FAERS database containing clinical reports of adverse reactions to prescribed drugs. Since these data are input by different doctors, the data is unorganized, inconsistent, and incomplete. We created a pipeline to identify duplicate reports, correct invalid entries using string alignment, and summarize the strength of association between drugs and side effects using a statistic that adjusts for biases in case reporting.

We then established a baseline for comparison – a classifier that ignores chemistry and just predicts the empirical frequency of each side effect. Surprisingly, we found that this method is similarly accurate as chemistry-based models as measured by the two most popular metrics in this field. In fact, we developed statistical theory and simulations that explain these results and evaluated our models on other metrics along several different axes.

These results led us to the conclusion that a finer-grained approach must be taken to understand the predictions of side effect models. To aid in this, we created a visualization tool that drug developers can use to interpret chemical spaces and uncover trends in model behavior normally hidden by aggregate scores. In summary, our work increases the quality, interpretability, and accessibility of side effect modeling in the pharmaceutical industry.
The Courtesy Messaging Signaling System (CMSS) is a mobile messaging application designed to facilitate communication between people who share the road, such as drivers and pedestrians. The CMSS is intended to make the road a safer place by giving advanced warning about any potential road hazards and allowing drivers to communicate their intentions, among other situations. This project is an extension upon an ECE Senior Design project from last year, where the team created a system where users could send messages to a Raspberry Pi via Bluetooth, which were then routed to an LED panel. The CMSS was transitioned to a system only dependent on a smartphone due to the desire not to have any specific hardware limitations. The app allows users to enter a picture and description of their vehicle which will be used for identification on the road. When active, the app uses Bluetooth to scan for other nearby users and notifies the server when the nearby users change. Upon tapping a button, a user says a phrase they would like to communicate to other users of the app. This phrase is then matched to one of several predetermined phrases to avoid offensive or confusing messages. The message is then sent to a server where the server forwards it to all phones which have been recorded as being nearby. Through use of the CMSS communication on the road will be enhanced, providing a safer environment for drivers and pedestrians.
Model-Based Systems Engineering Switch Configuration

Sonalysts is a company based in Connecticut that has integrated advanced technology capabilities of a defense research, development, and engineering firm while also providing solutions to services involving graphics design, sound design and set construction or exhibitory business. Sonalysts’ goal in sponsoring this system engineering project is to introduce the team upon the subject of model-based systems engineering of complex systems that are built from the synergy of computational and physical components, in this instance, an Arista switch is the main focal point. This exposure increases the awareness of systems engineering while also illustrating its concepts through projects. Another worthy note for the importance of this project is that major changes have been initiated by the Department of Defense which prioritizes the practice of Agile software development process as a means to improve the software quality and speed of delivery.

The goal of the project is to generate Arista switch configurations, which will then be loaded the new configuration onto the switches. Our task was to model the switch on a systems modeler called Cameo along with the necessary components and possible security features, if necessary, of a nominal target network. Additionally, the model exports all the essential configuration of the switch. The team also developed Python based code to convert the output of the model into configuration files that could be loaded onto the switch. The ultimate goal of this software development is to eliminate any manual updates that occur when configuration files are generated before being loaded onto the switch.
Startup Scorecard® is a web application designed to provide startups and investors with useful insights derived from data collected in the Pre app. Pre is an application created by Funding Founding that uses gamification to promote engagement between startups and potential investors during pitch events by having audience members make mock investments in participating startups. Startups compete with each other to earn the most investments while investors compete to make smart investments.

Our task was to create a web application that takes advantage of the large amount of data collected by Pre to help startups gain a greater understanding of their strengths, weaknesses, and potential investors while also helping investors make informed investment decisions. We built the web application using Flutter written in Dart, and we used Firebase for features such as Cloud Firestore, Cloud Functions, and hosting.

The data collected by Pre is stored in Firestore. A cloud function triggered by the web app pulls data from Firestore and serves the documents needed to create the various charts and widgets on our dashboard. Using the cloud function along with Firebase hosting allows us to save the static documents in the cache to limit the number of read calls to our database. For features that require real-time synchronization, such as chat, we directly accessed Firestore.

Another purpose of Startup Scorecard is to facilitate engagement between startups and potential investors following the pitch event. This was accomplished by adding features such as messaging and the ability to follow and track the growth of a startup.
Covid-19 Classification System

The purpose of this project was to abstract away some of the more technical aspects of configuring a neural network for the purposes of image classification. Specifically, our intentions were to create a platform where users could create and manage image datasets, create models from those datasets, and classify images against them. Our final product did not fully realize those goals, but we were able to prove that a scalable, distributed, classification system is possible.

This was accomplished through the development of three pieces of software.

1. The Worker – A thin python program which connects to the data service, to be discussed, in order to fulfill any user submitted classifications. There may be any number of these instances enabled at any moment. No request will be assigned more than one worker assuming the currently assigned worker produces a result within 72 hours. 

Github

2. The Data Service. – The core component of the platform. This is the centralized storage location for all of the image classification submissions. There are a collection of operations made available by this service to go about submitting, classifying, and viewing the results of a classification against a specific image.

Github

3. The React Webapp – The user facing web application where images may be uploaded and results obtained.

Github
Deep Reinforcement Learning - Humanoid Robot Simulations

“Deep Reinforcement Learning – Humanoid Robot Simulations” is a project that explores the realm of Reinforcement Learning (RL) algorithms and policies for simulating human-like actions. The contributing team utilizes tools including Open AI Gym, Multi-Joint with Contacts (MuJoCo) physics engine, Tianshou RL framework, Tensorflow, Deepmind Control, Pybullet, and Robosumo. Open AI Gym was used as a playground to create environments for XML models designed to perform specific tasks, while serving as the bridge between MuJoCo and Tianshou. The environments define action spaces, steps, rewards used by the agent to control the humanoid model's movement in an attempt to perform tasks such as standing up or walking. Additionally, custom ant, humanoid, and cheetah based models were developed and trained to understand the capabilities of both MuJoCo and OpenAI Gym.

The most successful work was built off of Tianshou, which is a fast-speed modularized framework based on PyTorch, supporting a wide variety of RL policies. The three primary policies used for training with MuJoCo’s “Humanoid-v3” model were Proximal Policy Optimization (PPO), Deep Deterministic Policy Gradient (DDPG), and Twin Layered Deep Deterministic Policy Gradient (TD3). Models trained using PPO were incapable of reaching a high enough reward threshold such that the humanoid could sustain a walking motion, while TD3 produced models that stood stably, but moved minimally. The best results have shown to come from DDPG, where the humanoid was able to find a stable “tip-toeing” position by locking the hips and either extending or locking the arms while in movement.

Feel free to send questions to the team!

Eric Wang – ericwang@uconn.edu
Gurman Singh – gurman.singh@uconn.edu
Tianze Ran – tianze.ran@uconn.edu
Maggie Cheung – maggie.cheung@uconn.edu
Ayman Braik – ayman.braik@uconn.edu
Sikorsky 720 Customer Hub Web Portal

We developed a feature rich customer portal for users to interact with the products and services provided at Sikorsky. We worked to create new features, along with basic features, that would help in the purchasing process of products and services. More specifically, we added a live chat feature for customer support, a vehicle status widget, and a calendar widget, on top of the usual login, signup, dashboard, and checkout features. Using Java, Spring Boot, Angular, AWS, and Tawk, we were able to create a prototype portal for Sikorsky.
Noise Life

Noise Life is an Android application created for auditory researchers, with the primary goal of being used with research participants to track their daily activities and ambient dB level of their surroundings over a long period of time. Our application prioritizes collecting accurate, meaningful data for the sponsor, providing the functionality to remind the user to update their activity not only periodically, but also when high variances in ambient dB are detected. Our app also offers a lot of customizability for researchers, including a secure settings page where they can start and stop the data recording process, as well as set various parameters pertaining to how often dB data will be collected, how often the user will be reminded to update their current activity, and under what conditions such notifications should be triggered. Our app is built on the Android operating system and will be published as an open-source project on GitHub. This allows other researchers to freely use, improve, or adapt the application to their own needs with relative ease, due to the plethora of resources available on Android development.

Our app is built on 5 core components, sensor data collection, user activity tracking, a notification system, UI and settings, and lastly, data storage and security. Sensor data collection focuses on the recording, averaging, and analyzing of noise data from a microphone connected to the phone. Activity tracking focuses on how we track the participant’s activities without being too invasive, while still providing interesting and usable data for the researchers. Another major component is the notification system, which reminds the user to update their activity both periodically, and during periods of high ambient dB, to collect more relevant and interesting activity data for the researchers. Overall, the notification system will allow us to narrow the gap between the sensor data and activity data, resulting in more useful data for the sponsor. For data storage, we had to design our app to never store raw audio data, only the average dB value, to prevent any privacy concerns. Another focus of this component was making sure the data was stored in a clear, readable format, and was easily exportable from the phone for the researchers to analyze.

Our application consists of two main tabs, one which is used by the research participant to input and update their current activity, and the other, securely locked behind a password, containing the settings panel. The first tab allows the user to select between a list of predefined “general” activities that are universally shared across most participants, such as Sleeping, Eating, Studying, Exercising, etc. For more personalized data collection, we have also included some custom activity fields, which the participant can fill in with their own specific activities that do not neatly fit into the other, predefined activities. The settings page, on the other hand, will be exclusively for the researchers, allowing them to stop, start, and tweak many parameters of the recording process. Some of the important options are the dB calibration offset to calibrate the noise data on the fly for different external microphones, as well as the ability to set the interval over which noise data will be recorded, then averaged and stored into one value. Other options include the ability to set the dB threshold at which our application should start notifying the user to update their activity, allowing the researchers to get more recent activity updates when the ambient dB increases beyond that threshold, as well as an option to set a recurring notification interval, periodically reminding the user to update their activity.
Predicting Outcomes of MLB Games

The sports industry has always relied on statistics and past data to make inferences about future trends and performance. This sector has become so lucrative that the sports analytics market is projected to reach nearly $5 billion by 2021. Major League Baseball is one of the largest consumers and industry drivers of these analytics. Statistical analysis is nothing new in baseball, and has been used to give teams an edge against their competition since its inception. Our team set out to create a predictive model that allows users to select matchups between two different baseball teams. In particular, users can select the away team, the home team, the away starting pitcher, and the home starting pitcher as input to the model and will be presented with the win probability for both of the selected teams. To accomplish this, the group leveraged the power of machine learning and used comprehensive data from Retrosheet to train various models. For the matchup predictions, Random Forest Classifiers gave the best results. The team also used machine learning order to make predictions for the 2021 standings for the end of the season using Ridge Regression. An accompanying webapp was made and can be accessed by anyone by visiting the website: sdpteam32.herokuapp.com. On the site, users can interact with the predictor and find links to see the source code. The data processing and machine learning was done in Python in the form of Jupyter Notebooks, while the webapp used a combination of HTML, CSS, and Python in conjunction with the Flask microframework.
Gamification of Training ("Adventrain")

Adventrain is a mobile application written using React Native that is designed to enhance the employee training experience through the use of gamification techniques. Instead of the typical training format used in many corporate environments, Adventrain uses AR along with a story-based experience to administer entertaining matching exercises to the user. These training modules are assigned through an infrastructure of user managed groups and tests which are stored in a PostgreSQL database backend system. Once a user has been assigned a module, they can access it through a comprehensive, personally tailored user dashboard before they take its associated test. These trainings divide each test’s content into stages that the user can complete to gain practice with the material. Users select a specific stage to begin practicing with by finding various virtual objects in a digital or AR space. Once an object has been found, the user is presented with the corresponding training data for that object’s category and a matching game that reinforces this information. When the user passes a stage by matching pairs of question and answer cards, they are rewarded with the digital object they selected. Once this process has been completed for all stages of the module, the user is free to begin a traditional testing process in order to meet the training standards of their employer or department head.
Rapid Cyber Exploit Reporting via Mobile App

Our team has designed and created a mobile-accessible web application used to report cyber attacks and exploits. The development of this application was sponsored by CT GMIS with the Town of Manchester to facilitate the current cyber exploitation reporting process. Time is of the essence when an attack occurs, yet often the news of cyber exploits doesn’t reach technicians and defender agencies until weeks after the attack has been detected. With this application, participating Connecticut municipalities can report cyber exploits in a timely and secure manner. Additionally, technicians can receive notification of cyber incidents in real time, removing unnecessary delays in reporting.

Our application makes use of various frameworks and services to come together and form a complete prototype. We use MongoDB as a database to store records in conjunction with Django to add, archive, or retrieve records as needed for our backend. With Angular, we provide a reactive, accessible frontend interface to display detected incidents, each with running comment chains to keep track of updates or edits. Administrators can also manage organizations and users. Using AWS, our application sends out notifications to all affected users and tagged agencies. This cloud service is also expected to be used by the sponsor for hosting and deploying the final application. Although the original goals of this project were to create both Android and Apple mobile applications, a slight shift in primary sponsorship to CT GMIS during the Spring semester changed our priority to making a single mobile-friendly web application.
Secure Embedded Architecture

Our team first researched the vulnerabilities that the board may be susceptible to such as Cache Attacks, side channel attacks, electromagnetic attacks, etc. Based on the research, we would come with attacks to attempt on the developmental board. Then we were tasked to gather metrics based on the cryptographic algorithms, AES and RSA, by using a prototype secure boot process by a software only approach and only hardware-assistance to gather the metrics. The metrics gathered were execution times, memory footprint, and level of protection of the board. Levels of protection being how the board would respond to things such as the wrong key given with the data to be encrypted, the key being reversed, or a one-bit difference in the key.
Path Planning with Deep Neural Nets

The modern world relies on road vehicles for almost every aspect of life. However, the speed at which these vehicles operate can make them deadly in human hands. Improvements in computing hardware, sensors, and machine learning algorithms have the potential to reduce the deaths, injuries, and damage to property caused by automotive accidents as well as increasing the economic efficiency of road vehicles. Moving into the future, new methods must be developed to minimize the number of accidents on the road. New control systems may be able to outperform human drivers in safety and could replace them entirely.

Mitsubishi Electric Research Laboratories has contacted the University of Connecticut Senior Design program and has asked them to design a vehicle testbed and implement a deep neural network for self-driving vehicles. CSE Senior Design Team 36 focused on the development of these methods in collaboration with a team of Electrical Engineering students who assembled the vehicle testbed that would evaluate the validity of the path planning algorithm. The neural network is designed to plan the path of the vehicle and appropriately adjust its course to avoid collisions and stay within the lane. The specific neural network structure that was used by the senior design team was inspired by a proposal from Dr. Karl Berntorp, published as US Patent 9989964B2. The vehicle testbed includes a sensor suite that will allow the vehicle to make real-time measurements of its surroundings and make decisions accordingly to maintain safe operation.
Path Planning with Deep Neural Nets

Our project is to create a new and revamped website for the Connecticut Crash Data Repository. The Connecticut Crash Data Repository is a website that compiles all the available data for vehicle crashes in the state of Connecticut. For this project we will be working with Director Dr. Eric Jackson and his team at the Connecticut Transportation Safety Research Center. The main goal of the website is to provide users with analytics on motor vehicle accidents and the factors that contribute to those accidents. These factors include but are not limited to drug use, road condition, distracted driving, and weather conditions. We have been tasked with updating the graphics and layout of their current existing site, as well as creating new and exciting data visualizations that are easier for viewers to understand. Furthermore, we have been tasked with performing analysis on toxicology data of drivers involved in accidents, as well as performing analysis of the effects of the COVID-19 pandemics on vehicle accidents in the state. Finally, due to complications with the foundation that the original site was built on, we have been tasked with increasing performance on the website, as well as decreasing the load time across the site. These are the tasks outlined to us by Dr. Jackson that our team will be focusing on with our website. To handle the data set displayed by the website, we will be using a MSSQL Database and an Apache server. For the data visualizations and website formatting we will be using Plotly Dash and Python.
Remote Firearm Detection

In recent years, the US has experienced many tragic shooting events. Businesses and public places need a security system design that can identify, deter, and eliminate threats with concealed firearms before tragedies occur. Our project uses modern AI and object detection to identify both concealed and openly visible firearms. The system has the flexibility to forward the detection results to a web-app interface or log the detections for future reference. We designed our project to find a good balance between size, speed, and cost – all aspects of the system are self-contained on a single small form factor computer – NVIDIA’s Jetson Nano – along with either a standard or infrared camera (for detecting concealed weapons). The end result is a model that can detect open carry rifles and pistols at ~84 % MaP and concealed at ~70% MaP.
Cut Room Scheduling

In the sewn goods industry, a cutting room is the industrial environment in which materials (usually fabrics) are spread and cut for apparel manufacturing. The processes of spreading and cutting are often complex and challenging to manage. Cutting rooms contain multiple tables on which the material is spread, often in multiple layers (called plys). Often, a manual spreader system is used in conjunction with an automatic cutter. Unfortunately, the automatic cutter systems are usually quite cost prohibitive and as a result a cutting room will often only have one that must be shared. Currently, cutting room managers must manually decide how to schedule the spreading and cutting of different orders in order to utilize these minimal machines and tables. This manual scheduling is suboptimal as it is simply not feasible for a manager to accurately determine the most efficient schedule that satisfies all of the various parameters. To solve this issue, our project team designed and developed an algorithm that, when given a list of cut orders and the needed information about them, will automatically determine what the best scheduling order is based on optimizing a series of factors (order due dates, order priorities, maximum cutter usage, and fastest schedule).
Human Position and Posture Detection System for Human Robotic Interaction Control

The main goal of this project was to design a posture detecting and replicating system that could be incorporated into a Kebbi robot. The system extracts body position, posture, and gestures using the Kebbi robot's built-in camera, then those features are processed using computer vision. After processing the robot's motors are driven to replicate the user's gestures and posture and the data frames are logged. The end goal was so users can educate and assist children using the Kebbi robot.
Encaptiv Marketing Website and Serverless API

The company encaptiv strives for perfection in online presentation, webinar, and event hosting. In order to provide useful functionality and a great user experience, encaptiv sought to improve their public marketing website and build an ARM (audience relationship management) product. To put it simply, encaptiv’s goal is to maximize user value, and our team was thrilled at the opportunity to assist in this goal.

For the website portion of our project (Figure 1), our approach was to use industry standard Javascript framework (Vue) with an industry standard language (Typescript). The benefits of using a modern Framework include improved source code readability/maintainability, dynamically generated content, and support for adaptive designs. The result is a website that was designed to be adaptive inside and out, giving its users and its developers a seamless experience.

The umbrella project of an ARM (audience relationship management) product contains several complex components. An important component is an API (Application Programming Interface). An API is typically any software that allows for software-to-software interaction and data sharing. The API we created (Figure 2) allows for recording and retrieving audience data while providing an abstraction layer between the application and its data, enabling simple, reusable access to data. With serverless technologies (AWS Lambda), it’s possible to create small functional pieces of code accessible to an API without needing a traditional server set up, which can be costly and difficult to maintain.

We employed the Agile development method throughout to improve organization, get continuous feedback, and ensure sponsor requested changes were completed before strict deadlines. Writing user stories helped our team prepare demonstrations of the user value gained on a week-by-week basis. This constant communication and continuous feedback helped our team produce finished projects that were to our sponsors exact specifications.
Oracle Creation with the Zap Protocol

In the context of blockchains, the oracle problem is loosely defined as the security, authenticity, and trust conflict between third-party oracles and the trustless execution of smart contracts. The Zap Protocol, a project sponsored by the New York Blockchain Center and the Synapse Foundation, aims to solve this by democratizing the oracle creation process. This encourages a robust marketplace, wherein data providers and other decentralized services compete with each other. Our project was split into multiple phases. The first was a learning phase, where we learned the terminology of the blockchain space and became familiar with some of the technologies used by developers. The second phase entailed academic research on the oracle problem. Through our research, we helped develop novel ideas for possible solutions to the oracle problem. This in turn informed the Zap team of potential design changes to their protocol. The third phase consisted of helping to drive open-source development in several programming languages. By expanding their codebase from JavaScript to other popular languages such as Python and Java, more developers from a wider range of disciplines could participate in the Zap ecosystem. The final phase involved implementing our own oracle that serves to end users the latest basketball scores and relevant statistics from all major American basketball leagues, both at the college and professional level. We also built an accompanying web application, and thoroughly documented the process to serve as a model for future users of the Zap platform.
HDF5 Interpretation and Visualization

The sponsor of this project, the Carrier Corporation, uses state-of-the-art technologies for model-based systems engineering (MBSE) of complex systems. One of the model-based design toolkits employed by Carrier is a software called Sandia Dakota. The Dakota software uses a file format known as HDF5 to contain the outputted information from the analysis results. Dakota’s HDF5 output files have a difficult structure to follow, in turn producing results which are difficult to interpret. Carrier is looking to develop a method by which they can automatically process such HDF5 files and make more efficient use of the information contained within them. Leading into the project which the team has been presented by the company: develop an application which can seamlessly parse an HDF5 file produced via the Dakota application, and use the information contained within the file to provide the user with an interface in which they can better analyze and visualize the information contained within. The application is a desktop application which is capable of reading and processing HDF5 files and utilizes web-based technologies in order to create a pleasant and intuitive interface allowing users to extract significant information from HDF5 files through producing versatile plots and visualizations of the data contained in various ways. The application will allow the Carrier team to easily visualize the results of methods, compare and contrast variables, find optimal values, and more. The application grants users the ability to create a variety of custom visualizations from the datasets contained within a wide-array of user provided HDF5 files. This application was developed using the Electron software framework, React and Bootstrap components, as well as the H5PY and Plotly libraries for data processing and visualization generation.
Model-Based Systems Engineering
Switch Configuration

Sonalysts is a company based in Connecticut that has integrated advanced technology capabilities of a defense research, development, and engineering firm while also providing solutions to services involving graphics design, sound design and set construction or exhibitory business. Sonalysts’ goal in sponsoring this system engineering project is to introduce the team upon the subject of model-based systems engineering of complex systems that are built from the synergy of computational and physical components, in this instance, an Arista switch is the main focal point. This exposure increases the awareness of systems engineering while also illustrating its concepts through projects. Another worthy note for the importance of this project is that major changes have been initiated by the Department of Defense which prioritizes the practice of Agile software development process as a means to improve the software quality and speed of delivery.

The goal of the project is to generate Arista switch configurations, which will then be loaded the new configuration onto the switches. Our was tasked with modeling the switch on a systems modeler called Cameo along with the necessary components and possible security features, if necessary, of a nominal target network. Additionally, the model exports all the essential configuration of the switch. The team also developed Python based code to convert the output of the model into configuration files that could be loaded onto the switch. The ultimate goal of this software development is to eliminate any manual updates that occur when configuration files are generated before being loaded onto the switch.