Research Statement

My research focuses on building intelligent mobile systems that leverage creative computational solutions to practically improve the lives of millions of people. To increase the potential impact and reach of my work, I publish in high-profile interdisciplinary journals and flagship computer science conferences, and present my work at specialty venues related to the project (e.g. audiology, hematology, cardiology, and underwater networking). This approach has allowed me to get valuable feedback from experts in multiple domains, which has been extremely helpful towards achieving adoption. My projects span both hardware and software and fall under three broad themes:

1) Reimagining healthcare delivery on a global-scale. Health inequity is seen most vividly in low- and middle-income countries where many patients lack access even to basic medical resources such as hearing tests. However, health inequities manifest even in wealthy countries like the US, where individuals in neighboring zip codes can experience 10-20 year differences in life expectancy. I developed mobile systems that can significantly bridge these equity gaps, both in the US and globally. I worked on the first wireless earbuds to screen for newborn hearing loss (Nature Biomedical Engineering '22), detected middle ear fluid using active sonar on smartphones and a paper cone (Science Translational Medicine '19), and showed how to perform blood clot testing at home using a smartphone (Nature Communications '22). I also built systems to address pressing issues in public health; specifically, I developed algorithms that run on smart speakers to detect cardiac arrest (Nature Digital Medicine '19) as well as a wearable injector system to reverse the fatal effects of opioid overdose (Scientific Reports '21). With the goal of deploying these technologies to millions of users, my work is now being commercialized by two startups, Wavely Diagnostics, where I am a co-founder, and which recently obtained FDA-listing for their first product, and Sound Life Sciences (acquired by Google). I was also a lead contributor to CovidSafe (now WA Notify), a COVID-19 contact tracing app built with Microsoft, which became part of official efforts by the WA Department of Health to manage the pandemic.

2) Systems for planetary-scale environmental monitoring. Though it is well known that human actions have had exceedingly negative impacts on underwater ecosystems, it has long been challenging for researchers to track the state of these rapidly evolving environments. I developed computing tools that allow researchers and citizen scientists to understand the effects of human actions on underwater environments at scale by lowering the prohibitive costs of field equipment. Specifically, I built open-source, field-tested communication systems that use speakers and microphones on unmodified smart devices to enable underwater acoustic communication at a range of 100m (SIGCOMM '22) and 3D underwater tracking (under submission). By taking a software-only approach, these systems can scale to millions of devices, which would not otherwise be possible had we relied on specialized hardware. The primitives I built lay the foundation for a broader set of research applications I am now pursuing, including tiny sensor motes for microplastic sensing that can help to reverse the effects of habitat destruction and ultimately support the growth of biodiverse ocean ecosystems. With the introduction of the latest Apple Watch, which can operate at underwater depths of 100m, I am excited about having my technology integrated into these devices for potential use by millions of people.

3) Real-time intelligence for mobile applications. Deep learning systems have shown a remarkable ability to excel at intelligence-based tasks, but these neural networks are often compute-intensive, which can make it challenging to create real-time AI systems that run on resource-constrained mobile devices. Furthermore, there are significant real-time constraints that require these networks to operate on small blocks (≤ 10 ms). I have been working on deep learning methods to enable these systems for tasks like semantic noise cancellation which filter out undesired sound classes (e.g. traffic) and allow sounds of interest (e.g. alarm clock). To overcome these challenges, we design an encoder-decoder architecture that combines computationally efficient dilated convolution layers for processing large receptive fields of size $R$ with complexity $O(\log R)$, with transformers as decoders to achieve state of the art performance (under review). Our system can generalize to real-world environments, and can operate with previously unseen sound classes. Furthermore, our system can maintain the binaural and spatial properties of the desired sound classes in the output audio stream.
The rest of this statement provides a detailed description of my work in these domains and proposals for future directions.

1) Reimagining healthcare delivery on a global-scale

Low-cost Newborn Hearing Screening Using Earbuds (Nature Biomedical Engineering ‘22). Hearing loss in newborns is particularly deleterious, often delaying neurological and social development unless detected shortly after birth. Tests for hearing loss typically require patients to respond to auditory stimuli of different sound levels, which is not possible in newborns. I developed two systems to perform hearing screening by detecting otoacoustic emissions, which are very soft sounds produced only by a healthy cochlea (inner ear), in response to auditory stimulus. The first system is a pair of low-cost wireless earbuds that use wireless sensing algorithms to reliably separate otoacoustic emissions from in-ear reflections and echoes. The second system is a $10 ear probe that leverages the speakers from everyday earphones, and attaches to a smartphone to test for otoacoustic emissions. The tool has been evaluated in a clinical study at Seattle Children’s Hospital where it correctly flagged all newborns with a hearing loss, and achieved performance comparable to a $5000 medical device designed to perform the test. This work has the potential to democratize access to hearing screening and reduce global health inequity in this area. We presented this work to government officials and clinicians at Kenya’s Ministry of Health and University of Nairobi, and are starting deployments at local clinics in Nairobi.

Detecting Middle Ear Fluid Using Smartphones (Science Translational Medicine ’19). Ear infections are by far the leading cause of pediatric healthcare visits. They are typically caused by a buildup of fluid behind the ear. Currently, the initial screen for middle ear fluid is via visual interpretation with an otoscope, which has an accuracy as low as 51%. Although more accurate methods like tympanometry and pneumatic otoscopy exist, they require significant expertise and referral to a specialist. We developed an accurate and accessible smartphone-based tool to detect middle ear fluid in children. Our smartphone-based tool assesses eardrum mobility by (i) sending a soft acoustic chirp into the ear canal using the smartphone speaker, (ii) detecting reflected sound from the eardrum using the smartphone microphone, and (iii) employing a machine learning model to classify these reflections and predict middle ear fluid status. The technique does not require any additional attachments to the smartphone beyond a paper funnel. The system was tested in a clinical study at Seattle Children’s Hospital on 98 pediatric ears, and obtained a sensitivity of 85% and a specificity of 82%, comparable to specialist tools. This work was selected as an annual highlight by the American Association for the Advancement of Science (AAAS). Based on it, I had the opportunity to author a column on the landscape of digital medicine in The Lancet (2019). This work is commercialized by UW startup Wavelly Diagnostics where I am a co-founder. Wavelly’s first product using this technology is now FDA-listed and available to select early access participants and health systems.

Blood Clot Testing Using Smartphones (Nature Communications ’21). Millions of people have heart conditions such as atrial fibrillation that require blood thinners, and an increased risk of death from blood clotting. Frequent blood clot testing is a reality for these individuals, who need to routinely monitor the effect of blood thinners on their clotting times. We developed a frugal blood clot testing system that uses a smartphone’s vibration motor to vibrate a small cup containing just a drop of blood obtained from a fingerstick, and the camera that tracks the micro-mechanical movements of a copper particle that moves and rotates in the blood sample. By tracking precisely when the copper particle stops moving as a result of the increased viscosity of blood from coagulation, we can accurately calculate clotting time. We ran a clinical study on plasma and whole blood samples from healthy patients and those with blood disorders and demonstrated a high degree of correlation with gold-standard laboratory equipment. Given the ubiquity of smartphones in the global setting, this technology may also provide affordable and effective blood clot testing in low-resource
environments. This work won three awards in the Hematology/Coagulation and Clinical Translation Science division at the American Association for Clinical Chemistry conference (2022).

**Contactless Cardiac Arrest Detection Using Smart Devices (npj Digital Medicine ’19)** Out-of-hospital cardiac arrest is a leading cause of death worldwide. Rapid diagnosis and initiation of cardiopulmonary resuscitation (CPR) is the cornerstone of therapy for victims of cardiac arrest. Yet a significant fraction of cardiac arrest victims have no chance of survival because they experience an unwitnessed event, often in the privacy of their own homes. An under-appreciated diagnostic element of cardiac arrest is the presence of agonal breathing, an audible biomarker and brainstem reflex that arises in the setting of severe hypoxia. Using real-world labeled 9-1-1 audio of cardiac arrests, we trained a support vector machine to accurately classify agonal breathing instances in real-time in a bedroom environment. We prototyped our contactless system using commodity smart devices (Amazon Echo and Apple iPhone) and demonstrated its effectiveness at identifying cardiac arrest-associated agonal breathing instances played over the air. This paper is ranked #3 on Altmetric out of all publications from npj Digital Medicine, and is licensed to the University of Washington startup Sound Life Sciences (acquired by Google).

**Closed-loop Wearable Naloxone Injector System (Scientific Reports ‘21).** Overdoses from non-medical use of opioids can lead to respiratory failure, cardiac arrest, and death when left untreated. Opioid toxicity is readily reversed with naloxone, a competitive antagonist that can restore respiration. However, there remains a critical need for technologies that can administer naloxone in the event of unwitnessed overdose events. We built a closed-loop wearable injector system that measures respiration and apneic motion associated with an opioid overdose event using a pair of on-body accelerometers and administers naloxone subcutaneously upon detection of an apnea. Our system was evaluated in: (i) an approved supervised injection facility where people self-inject opioids under medical supervision and (ii) a hospital environment where we simulate opioid-induced apneas in healthy participants. This work led to collaborations with West pharma and demonstrated a technical pathway for smart auto-injectors which the company is now pursuing.

**COVID-19 Apps and Symptom Tracking (IEEE Bulletin on Data Engineering ‘20).** During the onset of the pandemic lockdowns, I worked with researchers, volunteers, and engineers at UW and Microsoft round the clock to design and build a contact tracing solution from the ground up that could be downloaded by millions of people. Unlike other projects I worked on, this work was significantly more time-sensitive since even small project delays would have tangible effects on public health. In under a month, we designed privacy sensitive protocols (PACT) for Bluetooth and GPS-based contact tracing and developed an open-source system that was presented to Microsoft leadership and other universities. I learned valuable lessons about the need for leadership and decisive action during such times as a lead system contributor. It was very gratifying to see that these efforts were adopted by the WA Department of Health as part of official efforts to combat COVID-19. The app continues to be maintained by the Brotman Baty institute for Precision Medicine.

2) Systems for planetary-scale environmental monitoring

**Bringing Underwater Networking to the 21st Century (SIGCOMM ’22).** Since its inception, underwater digital acoustic communication has required custom hardware that has neither economies of scale nor pervasiveness. We presented the first acoustic system that brings underwater messaging capabilities to mobile devices like smartwatches and smartphones. Our software-only solution leverages microphones and speakers, ubiquitous in today’s devices to enable acoustic underwater communication between mobile devices. To achieve this, we designed a communication system that in
real-time adapts to differences in frequency responses across mobile devices, changes in multipath and noise levels at different locations and dynamic channel changes due to mobility. We evaluated our system in six different real-world underwater environments with depths of 2-15 m in the presence of boats, ships and people fishing and kayaking, and achieved a range of up to 100 m. With Apple’s latest smartwatch able to go to underwater depths of 100m, our software-based approach has the potential to make underwater messaging capabilities widely available to anyone with a mobile device. *All software and data for this system has been made open source.*

**Bringing Underwater GPS to Smartphones and Watches (under preparation).** We present a novel underwater system that uses acoustics to bring 3D tracking to commodity smartphones and watches. To achieve this, we designed a real-time underwater tracking protocol that computes the time-of-flight between a smart device and a floating buoy. To address the severe effects of multipath in underwater environments, we present a dual-microphone optimization algorithm that more reliably identifies the direct path of signals sent between devices and buoys. Further, our system works across different smartphones and watches even in the presence of clock drifts. Though existing underwater tracking research is targeted for custom hydrophone hardware, we believe that our work breaks new ground by demonstrating a path to bringing 3D underwater tracking capabilities to billions of existing smartphones and watches.

**3) Real-time intelligence for mobile applications**

**Real-time Target Sound Extraction (under review).** In this project, we present a novel neural network architecture, *Waveformer*, that can enable target sound extraction in real-time, where the goal is to extract sound signals of interest from a mixture of various overlapping sounds, given clues about the target sound class. The challenge here is that real-time applications impose significant algorithmic and computational constraints, requiring networks to operate on small blocks (≤ 10 ms) with a limited number of lookahead samples for each block, which can significantly degrade performance. We present the first deep learning method to perform target sound extraction in real-time, while achieving state-of-the-art performance for this task. To achieve this, we design an encoder-decoder model that consists of a stack of dilated causal convolution layers as the encoder, and the decoder layer of a transformer architecture as the model’s decoder. The advantage of this hybrid architecture is that the dilated causal convolutions can process large receptive fields of size $R$ efficiently with a computational complexity of $O(\log R)$, while transformer-based architectures can provide state of the art performance gains. This is in contrast to prior transformer-only architectures used for speech separation which had a complexity of $O(R)$. In our system, the encoder’s output which is conditioned with the target sound class is considered as the transformer decoder’s ‘target’ and the encoder’s unconditioned output is considered as the transformer decoder’s ‘memory’. Using self-attention on the ‘target’ followed by cross-attention between the ‘target’ and ‘memory,’ the decoder can then generate the mask for extracting the specified target sound and produce the output signal.

**Enabling Semantic Noise Cancellation for Intelligent Earables (under preparation).** Imagine being able to precisely program the acoustic environment around you by semantically specifying exactly what sound classes you want filtered out (e.g. traffic), and what you want to allow in (e.g. alarm clock). The goal of this project is to enable such a semantic noise cancellation system that is capable of running on resource-constrained mobile devices such as hearing aids and earphones in real-time using an optimized *Waveformer*-based architecture. Further, to ensure that our system can generalize to in-the-wild scenarios, we train it on a mixture of both synthetically rendered audio data which can simulate reverb and multipath effects, as well as data collected over the air in an anechoic chamber on our hardware setup, which best represents ground truth target recordings. Additionally, our system is designed to cancel sounds classes that were not part of the original training set by conditioning the network with audio embeddings of the new unseen sounds. Our system is also designed to preserve the binaural and spatial properties in the canceled audio stream.

**Future research directions**

I believe that mobile systems will be a key enabler for technological change in the coming decades. My goal is to lead the community in this direction by pursuing a mix of both curiosity-driven and socially impactful projects. Below I outline
several directions for future research centered around developing new techniques that can enable novel classes of mobile systems, and exploring application areas that have the potential to impact millions of people.

**Small-data intelligent mobile systems.** State of the art machine learning systems for tasks like speech recognition (e.g. OpenAI’s Whisper) are typically built on big datasets that cater to the vast majority of use cases. However, it is challenging for these systems to generalize to test data not present in the original training set. Recent efforts by Google to adapt their speech recognition technology to the speech patterns of individuals with speech impediments required them to record hundreds to thousands of phrases over the course of several hour-long recording sessions, which is a challenging task to undertake. This is in marked contrast to the short two-minute speech exercises conducted by speech pathologists to measure the full range of speech production mechanisms. I am interested in the development of real-time machine learning models that run on mobile devices and are fine-tuned using orders of magnitude less data than current systems, using few-shot or zero-shot based learning methods; thus, enabling practical deployments of personalized AI systems.

**Real-time ML mobile applications.** Current techniques that enable large machine learning models to run on the resource constraints of mobile devices typically rely on techniques like quantization and pruning which may not be sufficient for supporting low-latency applications such as speech separation, or for obtaining state-of-the-art performance. I am interested in building on top of my recent Waveformer work that took an application-specific approach to this optimization problem by developing a hybrid encoder-decoder model which combined the speed of stacked dilated convolutional layers and the performance gains of transformer-based models, which have increasingly been used to obtain state of the art performance while achieving real-time performance on mobile devices. There are also many future opportunities for hybrid architectures such as systems that leverage application-specific signal processing algorithms as a preprocessing step to generate embeddings for custom, light-weight neural networks that run in real-time.

**Mobile systems for health.** Going forward, I plan to focus on developing new techniques and tools for clinical monitoring and diagnostics for which there currently do not exist any solutions. I am interested in expanding on my work that was able to detect cardiac arrests on smart speakers by developing ambient monitoring systems and methods that can be used to contactlessly track different health conditions over extended periods of time. For example, can we develop a system that leverages active sonar technology and depth cameras to detect the occurrence of rare seizure events in a privacy-preserving way? Can we develop a continuous-learning machine learning system that adapts to the unique physiological patterns of each individual? I have also been interested in the development of mobile and non-invasive diagnostic tools that can detect the presence of disease from the gaseous chemical compounds in an individual's breath. The challenge here is in developing accurate systems that can detect even small chemical traces from mixtures of multiple gaseous compounds, so that diseases can be detected even during early stages.

**CS for the environment.** There has been renewed urgency in monitoring the effects of pollution and climate change on ecosystem biodiversity. Understanding these effects requires the development of new computing and sensing tools that can enable people to keep track of the state of the environment at scale. I plan to do this by building on top of my work to bring new capabilities to smart devices underwater. For example, I am interested in developing new imaging methods using acoustic and LiDAR sensors on smart devices that can be used even in turbid, low-visibility waters where smartphone cameras would not work. I am also interested in designing tiny sensor motes that can be dispersed into open waters to track the level of contaminants like microplastics in the ocean and communicate its findings to nearby base stations, while being powered by environmental forces like the sun or the waves.

My vision illustrating the breadth of mobile systems research as a field, including work I have completed, and work I am interested in conducting in the future.