

## EDUCATION

---

<b>Stanford University</b> <i>MSc in Statistics</i>	Stanford, CA <i>June 2022</i>
<b>Stanford University</b> <i>BSc in Mathematical and Computational Science</i>	Stanford, CA <i>June 2021</i>

## EXPERIENCE

---

<b>Graduate Research Assistant</b> <i>Stanford Human Trafficking Data Lab</i>	June 2021 – Present <i>Stanford, CA</i>
<ul style="list-style-type: none"><li>• Developed a containerized, distributed, cloud-native data pipeline to process hundreds of thousands of square kilometers of satellite imagery daily (Docker, Kubernetes, Redis, RabbitMQ, Google Cloud Platform).</li><li>• Developed a computer vision algorithm to identify charcoal production sites in satellite imagery (PyTorch, GDAL).</li><li>• System will be deployed for use by federal prosecutors in Brazil to identify locations where forced labor is being used.</li><li>• Papers forthcoming.</li></ul>	
<b>Research Assistant</b> <i>Stanford Center for Ocean Solutions</i>	June 2020 – June 2021 <i>Stanford, CA</i>
<ul style="list-style-type: none"><li>• Created a deep-learning-based computer vision algorithm to identify small fishing vessels in satellite imagery (PyTorch, GDAL, OpenCV).</li><li>• Analyzed entire near-shore region of the Peruvian EEZ and identified previously unknown locations where illegal, unreported, or unregulated fishing was occurring (Google Cloud Platform, Statsmodels, R).</li><li>• Code available <a href="#">here</a>.</li><li>• Paper forthcoming.</li></ul>	
<b>Tutor</b> <i>Stanford University Mathematical Organization</i>	September 2018 – June 2019 <i>Stanford, CA</i>
<ul style="list-style-type: none"><li>• Tutored Stanford students taking classes in linear algebra, vector calculus, and differential equations.</li></ul>	

## PROJECTS

---

<b>Light-Pipe</b>   <i>Python, Celery, Flask, GDAL, PyTorch, Tensorflow</i>	
<ul style="list-style-type: none"><li>• A Python framework designed to facilitate the deployment of geospatial data pipelines at scale.</li><li>• Provides platform-agnostic API for efficient ETL with cloud-optimized geotiffs.</li><li>• Allows for model training with heterogeneous geospatial data sources and formats.</li><li>• Handles task scheduling and queuing to facilitate distributed processing of geospatial data.</li><li>• Facilitates data ingest from major data providers.</li></ul>	
<b>“Weak Supervision with Incremental Source Accuracy Estimation”</b>	
<ul style="list-style-type: none"><li>• Developed a method to estimate the dependency structure and accuracy of weak supervision sources incrementally using precision matrices and robust principal components analysis.</li><li>• Allows for model training with weakly-supervised training data in on-line settings.</li><li>• Paper available <a href="#">here</a>.</li></ul>	
<b>“Effect of Behavioral Factors on Average Heart Rate During Cardiovascular Exercise”</b>	
<ul style="list-style-type: none"><li>• Designed a fractional factorial experiment to determine whether certain behavioral factors affect average heart rate during cardiovascular exercise.</li><li>• Found that four of the two factors tested were associated with a significant effect on average heart rate during exercise.</li><li>• Paper, data, and code available <a href="#">here</a>.</li></ul>	

## SKILLS

---

<b>Languages:</b> Python, C++, SQL, R, BASH
<b>Tools:</b> Git, Docker, Kubernetes, Apache Airflow, Apache Spark, Argo, PostgreSQL, PostGIS, Rabbit MQ, Redis, Google Cloud Platform, RESTful APIs
<b>Libraries:</b> Celery, Flask, Pytorch, Tensorflow, Scikit-Learn, GDAL, OpenCV, Statsmodels, Rasterio, Numpy, Pandas

## AWARDS

---

<b>National Merit Scholar</b>	April 2017
-------------------------------	------------