

# MICHAEL CROCKETT

[linkedin.com/in/rniice](https://linkedin.com/in/rniice) • [blueskyengineering.net](http://blueskyengineering.net) • [michael.crockett@gmail.com](mailto:michael.crockett@gmail.com) • (925) 961-7238

## PROFILE

Engineer specializing in the development of full lifecycle data product oriented solutions. 5+ Years Data Science, Machine Learning, & Analytics; 10+ Years Software Engineering; 10+ Years Product Development.

## EDUCATION & CERTIFICATIONS

- Harvey Mudd College, Claremont, CA — B.S. Engineering
- Board for Professional Engineers — PE License Mechanical Systems; [M35629](#)

## EXPERIENCE

### Data Scientist, PostPilot; Remote, WA — January 2024 - June 2025

- Innovated data pipeline on GCP for brand-specific new customer prospect targeting ranking engine including core application integration using Python, FastAPI, Docker, Cloud Run, Cloud Storage, BigQuery, Terraform, and Artifact Registry.
- Responsibilities included translation of core business objectives into programmatic standardized components used by other data scientists, setup and maintenance of source code using best practices, data cleansing routines, feature selection and encoding processes, and the reporting of metrics used to monitor reliability, performance, and quality of output from automated systems.
- Created prospecting model as a binary classifier by using XGBoost to predict new consumers similar to existing brand customer profiles by leveraging data gathered from existing consumer Shopify order history combined with consumer data (Acxiom, Epsilon).
- First year initial data product release driving considerable demand, resulting in over \$10 Million new revenue stream for PostPilot.
- Second year data product designed using Tensorflow Functional API builds on prior campaign learnings by harnessing conversion feedback from prior treatment groups as well as introducing a more sizable class imbalance to improve specificity on track for over \$20 Million new revenue FY2025.

### Staff Data Scientist, Maxar; Remote, WA (merged with Aurora Insight Jan 2023) — 2023 - 2024

- Refined core signal detection data model in support of enabling automated selection and delivery of RF signals matching customer specified characteristics without compromising on security by using AWS GovCloud delivery endpoints on S3.
- Developed a novel approach to RF signal search by applying VGG16 CNN and PCA to transform spectrum-isolated RF signal FFT waveforms into most differentiating extracted features from the total population of RF signal detections, enabling the rapid ranking and selection of lookalike signals from the population of all detected signals using Postgres Vector RDS on AWS.
- Harnessed RF similarity search to curate training datasets used for building more robust specialized signal classifiers.
- Applied RF similarity search to report global geospatial-spectral-temporal trends for signal characteristics of interest to DOD and commercial partners.
- Maxar sold RF Solutions to [HawkEye 360](#) end of FY2023, resulting in divestment of my division and thus role due to redundancy.

### Data Scientist, Aurora Insight; Remote, WA — 2021 - 2023

- Created production data processing pipeline, signal conditioning methods, and novel core model design for detecting, segmenting and classifying signals from broad spectrum RF measurements collected by a constellation of orbiting satellites on AWS by using Python, Numpy, Docker, Batch, RDS, S3, and ECR.
- Applied clustering and anomaly detection techniques to discover, rank, visualize, and deliver geo-spatial spectral temporal signal detection patterns of interest to DOD and commercial customers leading to [\\$2 Million NRO Contract](#).
- Hosted a week-long on-site data science hackathon including topic scoping, solution path ideation, responsibility division, and component consolidation into the delivery of a specialized company-wide set of API tools and Label Studio instance used for cross-team curation and communication of ground truth datasets for signal detection and classification purposes.
- Produced an interactive RF signal detection geospatial analysis and report creation web application using Streamlit and KeplerGL to streamline the capacity of a 4-person business development team to create deliverables for customers without engineering support.

### Data Scientist, Hewlett Packard; Remote, WA — 2020 - 2021

- Adapted legacy software practices for identifying, communicating, and tracking global printer organization quality control issues by incorporating AI and automation techniques into existing global quality assurance business operations.
- Created a specialized multi-class classification model using BERT in order to predict customer issue root cause from globally aggregated and translated call center transcriptions on a weekly basis.
- Developed and deployed an issue type anomaly detection method (PoT) and interactive dashboard using Streamlit in order to highlight outliers in aggregated quality control issue spikes by type, product line, and affected country used to proactively respond to emerging firmware, hardware, and quality issues at scale enabling precision root cause isolation in near real time.

### Software Engineer, Hewlett Packard; Vancouver, WA — 2016 - 2020

- Inspired firmware development team to adopt a new approach to solve a big data challenge by creating a proof of concept demonstrating multi-system onboard 3D Printer refined diagnostics platform using Python, Numpy, and Scikit-Image.

- Extended proof of concept into production by creating an onboard 3D Printer Analytics SDK using the facade software pattern to tightly coupled the firmware data model to a consistent and intuitive API access methods so that analytics developers only need to write a diagnostic routine once, ensuring critical diagnostic routines gather, process, and report subsystem specific metrics in standardized format without interruption across numerous printer subsystems.
- Edge diagnostics proved mission critical value to overall 3D printer fleet health monitoring, and tactical value to customer support during the COVID pandemic, by making it possible to proactively identify issue root cause remotely and send replacement parts necessary without requiring multiple costly site visits by maintenance technicians.
- Refined and augment core thermal camera algorithms used in calibration and diagnostic routines with dynamic part monitoring and reporting to enhance detectability of fusing process part temperature anomalies during layer by layer fusing process.
- Developed realtime remote fleet monitoring software using React, Plotly, and MQTT for visualizing live streaming data from R&D Lab 3D Printers in order to streamline operations between engineering and remote lab technicians.

### **Software Engineer, Autodesk; San Francisco, CA — 2014 - 2016**

- Created USB and Serial communication libraries for consumer 3D Printer Spark Platform using NodeJS and C++.
- Developed software to translate between various print profile settings into a proprietary Spark Platform format using NodeJS.
- Curated and integrated a database of consumer-grade FDM material feedstock supplies to deliver print profile ready configurations to users of Print Studio, and subsequently Fusion 360.
- Innovated R&D software algorithms and hardware interfaces used to extend typical FDM 3D printer capabilities resulting in several patents for large format FDM printing processes including part retention and multi-printer single part system coordination.

### **SKILLS**

- **Languages:** Python, SQL, DBT, NodeJS, JavaScript, JSX, React, C, C++, HTML, CSS
- **Cloud:** Terraform, AWS, S3, RDS, SQS, Lambda, Batch, EC2, CodeArtifact, GCP, Cloud Storage, BigQuery, Cloud Run, IAM, Artifact Registry
- **Techniques:** Data Modeling, Machine Learning, Deep Learning, Analytics, Algorithms, Data Visualization, Statistics, Jupyter, Web Scraping, Data Cleaning, Scripting, Statistical Modeling, Quantitative Analysis, Data Analysis, Business Analysis, Optimization, Collaborative Influence on Strategy
- **Libraries:** Pandas, Dask, Numpy, Scikit-Learn, Scikit-Image, SciPy, TensorFlow Keras, PyTorch, FastAPI, Streamlit, KeplerGL, XGBoost, YOLO, OpenCV, TSLearn, MQTT, Matplotlib, Plotly
- **AI Orchestration:** LangGraph, LlamaIndex, Fast MCP, OpenAI, n8n, Ollama, LLMs, RAG, Hugging Face, browser-use
- **Workflow:** Linux, macOS, Docker, Git, JIRA, VSCode, Label Studio, WindSurf, Cursor

### **PRESENTATIONS**

- (1) Rapid & TCT 2017 - Speaker
- (2) 3D Printed Violins - Producer

### **PATENTS**

- (1) Active Fire-Blocking Wind Deflector
- (2) Controllable Release Build Plate For 3D Printer
- (3) Vacuum-Assisted Incidental Build Material Collection
- (4) Thermal Characteristic Control in a Build Material
- (5) Build Material Spreaders
- (6) Adjustments Based on the Age of Fusing Agents
- (7) Build Material Spreading Wall
- (8) Automated Conveyance of Articles in Vapor Processing