

# Casey Kneale, Ph. D

Grafton, MA | 412-722-4802 | US Citizen | [cpkneale@gmail.com](mailto:cpkneale@gmail.com) | [LinkedIn](#) | [GitHub](#) | [Blog](#)

---

I have 9 years of professional experience transforming real-world needs into software solutions. I've designed and built software products ranging from front-end components to novel analytical methods, including R&D projects at demo scale and production systems handling multiple 100Gb/s workloads. This experience includes developing plant-wide scheduling software for pharma, monitoring solutions for the oil & gas industry (e.g., LNG tankers), and cybersecurity tools analyzing large portions of the internet traffic in the US. I have also had the privilege to work on a DARPA project, consult and give a guest invited lecture.

## PROFESSIONAL EXPERIENCE

---

### **Software Engineer III, Applied Materials (Remote) Apr 2024–Apr 2026**

- Architected and Engineered LDAP based SAML SSO (ComponentSpace C#) to enhance security, reduce tool dependency, and streamline customer onboarding.
- Lead and developed a “flexible layout” configuration product which enabled users to display their data in the form of charts, tables, and other visual components using a simple wizard workflow (Angular, TS, C#, Postgres). This tool reduced custom dashboards legal requirements from GAMP5 to GAMP4 and simplified the creation of multi-component reports/dashboards for end users.
- Lead the creation of a new interactive Gantt chart component and scheduling application (Angular/Typescript). The component allows users to display active tasks, constraints, notes, and included features like dynamic zoom, hierarchical/grouped rows. It is both more efficient and capable than off-the-shelf Gantt charts.
- Created an ETL datasource that ingests alarm data over MQTT which enabled users to be notified of which tasks require immediate attention (Angular, TS, C#, Postgres).
- Co-created dynamic scripting capabilities and alarm management software.
- Performed regular duties, such as: code review, bug fixes, releases, quality assurance, documentation, maintained Jenkins CI/CD pipelines, security updates, design discussions, quality assurance, etc.

### **Software Engineer III & Data Scientist, Redjack LLC (Remote) Jul 2022–Apr 2024**

- Improved the correctness of a monthly run customer-facing data aggregation and reporting tool by optimizing the data structures used and reducing allocations. Reduced the wall-time for a batch of records from 2 days to 2 hrs.
- Wrote an in-memory packet generator (100Gb/S) to bypass our network interface. After profiling and assessing observability metrics, rewrote our network sensor's parallelism model to remove bottle-necks. The sensor emulation module has since been used to test and resolve other issues across engineering.
- Improved the correctness of our internal record differencing tool and optimized the run-time from 66 hours to roughly 5 min on example logs.
- Resolved a correctness issue in our IP address anonymizer algorithm while also reducing the processing time of a batch from more than 2 days to 50 min by code review, reading journal articles and tracking syscalls to PRNG.
- Worked on the data science team to create models (PU learning, PyTorch, etc), as well as tools used to improve the teams' workflows. Roadmapped, prototyped, and implemented features from MITRE standards for cyber resilience.
- Transformed a deeply nested JSON file format into a normalized Apache Parquet format for easy ingestion, smaller files (less cost), faster processing times, and customer requested 3rd party tool accessibility (Rust, DuckDB, Python, etc). Analytic query times dropped from hours to tens of seconds.
- Deployed and statistically evaluated large language AI models (LLMs) for their efficacy and risk in R&D applications.
- Worked to improve the culture and benefit productivity by co-creating ground rules for code reviews. Suggested language and tips to reduce ambiguity and advocated for clear definitions of Accept, Comment, and Reject on pull requests that helped cut the average review time by 50% while maintaining code quality.

### **Software Engineer, RelationalAI (Remote) Apr 2022–Jul 2022**

- Engineered and integrated FIPS compliant encryption-at-rest features for our database product using Rust (OpenSSL v3.0.2), ensuring robust security across Julia and Go code bases with adherence to SLA's.
- Developed Rust modules for secure key management, including hierarchical key handling, rotation, and rekeying.
- Overcame shortcomings in Julia's young ecosystem by creating functionality for performing CRUD operations on blob (Azure blob storage) metadata in atomic operations, and integrating key vault operations.

### **Data Scientist – Quantitative Production Intelligence Tools, DreamView (Remote) Apr 2021–Apr 2022**

- Translated domain expert needs into systems by writing API's (Go) that stored and brought value from a process management graph database. This served as a replacement for a costly and inefficient legacy tool.
- Created and maintained PostgreSQL(GCP), TigerGraph and Neo4j databases along with companion API's for gRPC daemons. The databases were designed for storing/retrieving essential information for our in-house platforms and to improve the tracking of ground truth vs expected time-lines.
- Built new vendor relationships by communicating our needs and contributed to discussions in various open source communities, which allowed DreamView to quickly choose appropriate technologies for the future platform by understanding the technical merit and risks associated with deploying existing solutions.

### **Senior Chemometrician Scientist & Project Manager, Pall Corporation (Westborough, MA) Dec 2020–Apr 2021**

# Casey Kneale, Ph. D

Grafton, MA | 412-722-4802 | US Citizen | [cpkneale@gmail.com](mailto:cpkneale@gmail.com) | [LinkedIn](#) | [GitHub](#) | [Blog](#)

---

- Lead an international cross functional team of 10 people to correct issues related to the acquisition, modeling, and control of Raman spectrometers in bioprocessing.
- Saved the team hundreds of thousands of dollars by reviewing nondisclosure agreements and correcting misunderstandings in statements from 3<sup>rd</sup> parties.
- Constructed dashboards that were used for monitoring real-time multivariate predictions of CQA's. Demonstrated simple but effective multivariate modeling paradigms, and automated control of Raman spectrometers without user intervention via 3<sup>rd</sup> party software SDK's, self created software, and time series modeling strategies.
- Performed design of experiments to validate instrumental performance under various collection conditions.
- Collaborated to build ad hoc data collection tools for business intelligence using technologies like javascript (Puppeteer), Rust, and AWS lambda functions.
- Performed technical interviews to bring in new talent and strengthen Pall's data science and engineering capabilities.

## **Chemometrics Engineer, Kaiser Optical Systems (Ann Arbor, MI)**

**Sep 2019–Dec 2020**

- Developed a semi-automated modeling algorithm for quantitative and qualitative analysis of gaseous Raman spectra. Enabled customers to gain insight from their processes without specializing in analysis or purchasing 3<sup>rd</sup> party tools.
- Delivered Julia and C++ (Armadillo) libraries that brought KOSI's digital signal processing and machine learning modeling capabilities to a 3<sup>rd</sup> party software company.
- Improved and created models for the Optograf product allowing oil and gas customers to quantify their payloads which contained tens of millions of dollars worth of fuel.
- Simplified the quote to sale process in the Oil and Gas industry by creating and implementing tooling that sales personnel were comfortable using. This reduced turn over by more than 10 fold and improved data quality.
- Identified a valuable data stream that was not being used in manufacturing and redirected it for use in R&D. This saved tens of thousands of dollars for future product development.

## **Data Scientist, CCRI (Charlottesville, VA)**

**May 2018–Sep 2019**

- Innovated a solution to a customer requested research topic by creating neural network architectures for cyber threat triage systems via ranking systems and transfer learning from historic data (arXiv/1907.11129).
- Wrote, trained, validated, and delivered a fast character level domain generation detection model (PyTorch) that allowed our team to hone in on suspicious activity on a secured network.
- Created a library (PySpark, Pandas) that performed distributed feature generation operations and built a reliable, ETL pipeline that handled 100's of megabytes of network logs per hour. This effort leveraged Julia, Python, Linux tools, ZMQ, high-performance C++ code which provided a secure data source for ML
- Applied scientific reasoning to image preprocessing and reduced a computer vision models (PyTorch, Darknet) training data size by 100 fold while improving accuracy.

## **GUEST INVITED LECTURE**

*NIR spectroscopy in agricultural sciences*. Fall 2020, UIUC. Dr. Mohammed Kamruzzaman

Presented a lecture with Q&A for senior students in UIUC's agriculture program. The content included a primer on NIR spectroscopy, the different types of NIR measurements, applications to agriculture, and aerial imagery.

Explained several useful figures of merit for NIR imagery, and how statistics and machine learning pushes the envelope.

## **CONSULTING WORK**

VIBEIA

Assessed feasibility of a product and its development for a high value problem as well as the risk for attempting it.

Worked with the tight knit international team to improve computer vision features and models. Transferred knowledge about machine learning and model maintenance.

## **EDUCATION**

[Doctor of Philosophy, Analytical Chemistry \(Chemometrics\)](#) | University of Delaware | Newark, DE | 2015 - 2018

[Bachelors of Science, Chemistry](#) | California University of PA | California, PA | 2011-2015