Dr. Mohith Surya Kiran Kasula

(317) 756-7183 mohithkasula@gmail.com
linkedin.com/in/mohithksl/ github.com/mskkasula plhi/mohith-kasula/

Professional Summary

Health informatics professional with clinical training as a dentist and hands-on experience in U.S. public health, academic research, and healthcare data projects. I help teams turn fragmented clinical data into reliable, analysis-ready information and practical dashboards that support care delivery, research, and operations. Colleagues value that I can speak the language of both clinicians and data teams, ask the right questions about workflow and data quality, and follow through with clean, well-documented solutions using Python, SQL, and modern EHR tools.

Experience

Indiana University Indianapolis — Aug 2024 - Sep 2025

Research Assistant at Purkayastha Lab for Health Innovation — *Indianapolis, IN*

- Designed, tested, and validated the SuperAssist mobile app to support secure supervisor-provider communication, embedding Hive and Firebase databases for persistent local storage and selective cloud syncing.
- Developed Python- and SQL-based analytics workflows to monitor provider burnout and workload trends, enabling real-time dashboards and data validation checks for research reporting.
- Improved data reliability and user experience by optimizing synchronization logic across devices and sessions, ensuring consistent access to client updates while safeguarding provider self-reported data.

Virginia Department of Health — Dec 2024 - Mar 2025

Data Analyst — *Richmond, VA*

- Extracted, cleaned, and analyzed public health data from REDCap and other internal databases using Python, SQL, and Tableau to generate dashboards and reports for statewide health equity initiatives.
- Collaborated with policy analysts, epidemiologists, and external partners to design data collection instruments, reporting structures, and visualizations aligned with Virginia's public health priorities.
- Performed QA checks, data validation, and discrepancy resolution to improve data completeness, consistency, and readiness for regulatory reporting and program evaluation.
- Automated recurring reports and created Mail Merge workflows from REDCap exports, streamlining communication and documentation for grant reporting and partner updates.

Indiana University Purdue University Indianapolis — Aug 2022 - May 2024

Graduate Research Assistant at Purkayastha Lab for Health Innovation — *Indianapolis, IN*

- Led the migration of 200,000+ patient records from the MIMIC-IV database into OpenEMR, designing and executing ETL pipelines across 20+ clinical tables while maintaining data integrity and HIPAA compliance.
- Mapped MIMIC-IV fields to OpenEMR and OMOP CDM, automating RxNorm and drug code mapping with Python, and implementing SQL-based validation rules to ensure accurate population of EHR fields.
- Built robust batch loading and dependency handling scripts, achieving 98%+ accuracy in EHR display and enabling faculty to use OpenEMR as a live training environment for 200+ health informatics students.
- Evaluated 75 AI/ML algorithms across domains (vision, audio, NLP) for reusability and reproducibility, identifying environment-level issues (CUDA, dependencies) and documenting barriers to reuse.
- Implemented Docker-based containerization workflows in cloud environments (Jetstream2) and contributed to AI-readiness recommendations grounded in FAIR principles and metadata standardization.

GTechnologies Pty Ltd — Aug 2023 - Dec 2023

Data Science Intern (Remote) — *Sydney, Australia*

- Automated extraction of renal parameters from unstructured pathology reports using Python and NLP techniques, significantly reducing manual chart review and data entry time.
- Developed predictive models to analyze renal health trends and forecast risk patterns, supporting improved clinical decision-making and earlier identification of high-risk patients.
- Designed visualizations of renal health indicators using Matplotlib and Seaborn, enabling clinicians and stakeholders to interpret patient cohorts and temporal trends quickly.

Narayana Dental College and Hospital — Apr 2021 - Mar 2022

Dental Informatics Intern — *Nellore, AP, India*

- Supported implementation and optimization of EHR systems across dental departments, improving patient data capture, chart retrieval, and documentation workflows.
- Developed and delivered training modules for students and staff on EHR use, improving data entry accuracy and adherence to health information standards.
- Collaborated with clinicians and administrative staff to streamline patient information retrieval and reporting, reducing manual errors and enhancing care coordination.

Skills & Expertise

Programming & Query Languages: Python, SQL, R, SAS, XML.
Analytics & Data Science: Data Cleaning and Preparation, Exploratory and Statistical Analysis, Predictive Modeling, Machine Learning, Deep learning, Natural Language Processing (NLP), Data Reporting and Visualization.
Frameworks & Libraries: Pandas, Numpy, Scikit-Learn, TensorFlow, PyTorch, Matplotlib, Seaborn, Plotly, ggplot2.
Tools & Platforms: Tableau, Power BI, REDCap, Qualtrics, Talend Open Studio, ArcGIS, RStudio, SPSS, Unix, MS Office, GitHub, Azure, Flutter.
Healthcare Standards & Concepts: ICD-10, SNOMED CT, RxNorm, HL7, HIPAA, HCPCS, NDC, CPT, LOINC, OMOP CDM, EHR workflows.
Project & Process: Health IT Lifecycle Management, Requirements Gathering, Cross-functional Team Coordination, Agile, User Training and Adoption.
EHR & Health IT Systems Systems: OpenEMR, OpenMRS, Cerner (CST Cerner/PowerChart tutorials).

Projects

MIMIC-IV to OMOP/OpenEMR: Healthcare Data Migration | *Python, SQL, Talend* **Aug 2023 - May 2024**

- Extracted and transformed records for 40,000+ patients from MIMIC-IV into OpenEMR and OMOP CDM, standardizing data across 33+ source tables for educational and research use.
- Authored schema mappings and Python-based transformation pipelines, integrating RxNorm for medications and generating synthetic identifiers to support HIPAA-compliant de-identification.
- Implemented SQL-based data validation rules, batch loading with referential integrity checks, and reverse-loading verification, achieving 98%+ accuracy of clinical views in the EHR.

Reproducible AI Frameworks for Health Research| *Python, Docker, GitHub, Jetstream2* **Jan 2023 - May 2024**

- Assessed reusability and reproducibility of 75 open-source ML algorithms relevant to digital health, identifying environment and dependency barriers to clinical research reuse.
- Developed containerized execution workflows using Docker/Singularity to standardize environments and improve reproducibility of computational experiments.
- Contributed to infrastructure recommendations for FAIR-compliant AI pipelines in healthcare, emphasizing metadata capture, environment documentation, and open science practices.

Statistical Analysis of Location-Based Trends in Accommodation Pricing | *R, RStudio* **Jan 2023 - May 2023**

- Analyzed 48,000+ lodging records to study price variation by room type and distance to a central hub, mirroring geospatial analysis approaches used in public health and access-to-care studies.
- Ensured data quality via outlier handling (IQR capping), normality checks (Q–Q plots), and variance tests (Bartlett's), and applied non-parametric hypothesis tests (Kruskal–Wallis, Mann–Whitney U).

Predictive Modeling for Depression Classification in U.S. Adults | *Python, SQL* **Aug 2022 - Dec 2022**

- Integrated and engineered features from national datasets (NHANES, NAMCS) to evaluate gaps between patient self-reported depression and provider-documented diagnoses.
- Applied data cleaning, encoding, and chi-square tests to identify key predictors (comorbidities, age, race, insurance), informing risk stratification strategies.
- Built ML models (XGBoost, Random Forest) with SMOTE class balancing, achieving 85% accuracy and highlighting underdiagnosed at-risk subpopulations.

Education

Indiana University Purdue University Indianapolis **Aug 2022 - May 2024**
Master of Science in Health Informatics *Indiana, USA*

Dr. NTR University of Health Sciences **Sep 2016 - Mar 2022**
Bachelor of Dental Surgery *Andhra Pradesh, India*

Licenses & Certifications

- Licensed Dentist by the Dental Council of India.
- Data Analytics Essentials
- Python Data Analysis
- Project Management: Healthcare Projects
- GCP – Social and Behavioral Research Best Practices for Clinical Research
- Biomedical Researcher
- Biomedical Responsible Conduct of Research
- Introduction to CST Cerner
- PowerChart Basics
- Privacy and Confidentiality in CST Cerner

Publications

- M. S. K. Kasula, S. H. Sudalagunta, K. Sunchu, and S. Purkayastha, "Utilizing ETL Processes to Enhance Healthcare Education: Migrating MIMIC-IV to OpenEMR," *2024 IEEE 37th International Symposium on Computer-Based Medical Systems (CBMS)*, Guadalajara, Mexico, 2024, pp. 186-191, doi: 10.1109/CBMS61543.2024.00039.
- R. Quick and M. S. K. Kasula, "AI Readiness: A Reusability Study of Popular AI Algorithms," *2025 58th Hawaii International Conference on System Sciences (HICSS)*, The Big Island, HI, USA, 2025, pp. 7351-7358, url: https://hdl.handle.net/10125/109731.