**HARRY “JAKE” REED**

**706-836-5595** **hreed.ga@gmail.com** **** [**https://tinyurl.com/jreedportfolio**](https://tinyurl.com/jreedportfolio)

**DATA SCIENTIST | BIOINFORMATICS | MACHINE LEARNING**

Driven and results-oriented Data Scientist with a PhD in Bioinformatics and a comprehensive background in machine learning, high-dimensional data analysis, and security management. Demonstrated expertise in developing innovative analytical tools and leading cross-functional teams to drive impactful research and operational efficiency. Skilled in data-driven decision-making, project management, and fostering collaborative environments. Adept at leveraging advanced data analysis techniques to enhance security protocols and optimize research methodologies.

* Demonstrated expertise targeting opportunities to expand relationships by continuously proposing solutions that meet objectives and add strategic value. Adept in delivering organizational change management to obtain the best results for programs focusing on improvement, technological advancement, and end-user success.
* Interpersonal, organizational, and critical thinking skills, including an ability to work effectively within an interagency environment on complex issues requiring negotiation and consensus-building.
* Excellent interpersonal, verbal, and written communication skills. Stellar ability to communicate.
* Maintains an active awareness of industry best practices, changes, and trends utilized to improve processes and policies.
* Fluent in applying research and data analysis to make decisions, craft strategy, and track progress towards organizational goals.
* **Software:** Advanced proficiency in Microsoft Office Suite (Excel, PowerPoint, Word), SharePoint, Programming Languages: R, Python, BASH, Developer Tools: Git, Docker, SVN, Singularity, Libraries: ggplot2, Seurat, Pandas, NumPy, Seaborn, Matplotlib

**AREAS OF EXPERTISE**

Bioinformatics Analysis, Genomic Data Analysis, Next-Generation Sequencing (NGS), RNA-Seq, R Programming, Python, Machine Learning, Data Visualization, Statistical Analysis, Cancer Genomics, Transcriptomics, Differential Expression Analysis, Bioinformatics Pipelines, High-Performance Computing (HPC), Data Mining, Biostatistics, Omics Data Integration, Computational Biology, Gene Expression Profiling, Pathway Analysis, CRISPR Data Analysis, Single-Cell RNA-Seq, Unix/Linux, Git, Data Management, Scripting, Bioinformatics Tools, Multi-Omics Analysis, Software Development, Sequence Alignment, Big Data Analysis., High-Performance Computing (HPC), Statistical Analysis, Package Development, Professional Writing, Science Communication, Security Management

**PROFESSIONAL EXPERIENCE**

**Postdoctoral Fellow, Bioinformatics**

**Georgia Cancer Center**  **07/2022 – Present**

* **Research Leadership**: Spearheaded and managed multiple high-profile bioinformatics research projects focusing on Topological Data Analysis and telomere length dynamics, resulting in publications in leading scientific journals.
* **Tool Development and Innovation**: Innovated and developed a suite of advanced data science tools for analyzing telomere dynamics in human cells and chronic lymphocytic leukemia. These tools significantly improved the accuracy and efficiency of biological research and data interpretation.
* **Team Leadership and People Management**: Led a team of two graduate students, providing mentorship on advanced data analysis methodologies, statistical modeling, and best practices in bioinformatics. Fostered a culture of continuous learning and professional development.
* **Collaborative Research Initiatives**: Collaborated closely with interdisciplinary faculty researchers, providing expert data analysis support and driving forward novel scientific discoveries through cross-functional collaboration.
* **Data Security and Management**: Implemented robust data security measures to protect sensitive research data, ensuring compliance with institutional and federal guidelines for data handling and management.

**Project Manager, 82nd Airborne Infantry**

**United States Army**  **07/2004 – 07/2008**

* **Project Management and Leadership**: Led multiple high-stakes projects focused on enhancing security measures. Successfully managed cross-functional teams to implement secure and efficient data management solutions.
* **Security Management**: Developed and implemented comprehensive security management protocols to safeguard sensitive data and systems against threats. Conducted regular security audits and risk assessments to identify vulnerabilities and improve overall security posture.
* **Data Analysis and Reporting**: Utilized advanced data analysis techniques to evaluate system performance and security metrics. Developed detailed reports and dashboards to monitor key performance indicators (KPIs) and inform strategic decision-making.
* **People Management**: Oversaw a team of specialists, providing guidance on best practices in security and project management. Conducted training sessions to enhance team skills in security protocols and data analysis techniques.
* **Stakeholder Engagement**: Engaged with internal and external stakeholders, including government agencies and private sector partners, to align project goals and deliverables with broader organizational objectives. Facilitated regular meetings to ensure clear communication and alignment of priorities.

**EDUCATION**

**Published Data Science Tools**

* **TECAT (Patent Pending)**: A cutting-edge data science tool designed to accurately determine size distributions of biological macromolecules, with significant applications in precision medicine and biopharmaceuticals. This tool integrates data from the NASA Twin Study and includes a dynamic shiny dashboard for comprehensive data visualization.
	+ **Impact**: Enhanced the capability of researchers to analyze complex biological data with higher precision, contributing to advancements in the field of bioinformatics.
* **RPointCloud (Available on CRAN)**: An innovative tool utilizing persistent homology for the topological data analysis of large numeric matrices. This package enables researchers to visualize high-dimensional data in a 3D space, providing deeper insights into data structure and relationships.
	+ **Impact**: Facilitated the exploration and understanding of complex data sets, supporting groundbreaking research in various scientific domains.
* **TLD (R and Python Package)**: A versatile package for analyzing telomere dynamics, leveraging Python libraries for visualization and R for advanced data analysis. This tool demonstrates cross-platform compatibility and flexibility in bioinformatics applications.
	+ **Impact**: Improved the efficiency of telomere analysis, allowing for more accurate assessments in research and clinical applications.
* **MinimapR (Available on CRAN)**: A comprehensive wrapper for minimap2, enhancing DNA sequence mapping capabilities. This tool underscores expertise in bioinformatics tool development and genomic analysis.
	+ **Impact**: Streamlined DNA sequencing workflows, increasing the speed and accuracy of genomic data analysis.

**EDUCATION**

**PhD in Bioinformatics**
Weill Cornell Medicine, New York, NY | Aug 2021

* + Specialized in machine learning models, high-dimensional data sets, and topological data analysis.

**MD**
Weill Cornell Medicine, New York, NY | Aug 2013 – Aug 2015

* + Completed Years 1 and 2; Passed USMLE Step 1

**BS in Chemistry and Biology, Magna Cum Laude with Honors**
Augusta University, Augusta, GA | May 2013

**PROFESSIONAL PUBLICATIONS**

McGee RL 2nd, Reed J, Coombes CE, Herling CD, Keating MJ, Abruzzo LV, Coombes KR. Topological Structures in the Space of Treatment-Na¨ıve Patients with Chronic Lymphocytic Leukemia. Cancers (Basel). 2024 Jul 26;16(15). doi: 10.3390/cancers16152662. PubMed PMID: 39123390; PubMed Central PMCID: PMC11311631.

Grigorev K, Foox J, Bezdan D, Butler D, Luxton JJ, Reed J, McKenna MJ, Taylor L, George KA, Meydan C, Bailey SM, Mason CE. Haplotype diversity and sequence heterogeneity of human telomeres. Genome Res. 2021 Jul;31(7):1269-1279. doi: 10.1101/gr.274639.120. Epub 2021 Jun 23. PubMed PMID: 34162698; PubMed Central PMCID: PMC8256856.

Reed J, Kirkman LA, Kafsack BF, Mason CE, Deitsch KW. Telomere length dynamics in response to DNA damage in malaria parasites. iScience. 2021 Feb 19;24(2):102082. doi: 10.1016/j.isci.2021.102082. eCollection 2021 Feb 19. PubMed PMID: 33644714; PubMed Central PMCID: PMC7887396.

O’Hara NB, Reed HJ, Afshinnekoo E, Harvin D, Caplan N, Rosen G, Frye B, Woloszynek S, Ounit R, Levy S, Butler E, Mason CE. Metagenomic characterization of ambulances across the USA. Microbiome. 2017 Sep 22;5(1):125. doi: 10.1186/s40168-017-0339-6. PubMed PMID: 28938903; PubMed Central PMCID: PMC5610413.

Calhoun SF, Reed J, Alexander N, Mason CE, Deitsch KW, Kirkman LA. Chromosome End Repair and Genome Stability in Plasmodium falciparum. mBio. 2017 Aug 8;8(4). doi: 10.1128/mBio.00547-17. PubMed PMID: 28790200; PubMed Central PMCID: PMC5550746