Director of Data Science

BLOODPAC Consortium | Blood Profiling Atlas in Cancer

Organization: BLOODPAC Consortium (Blood Profiling Atlas in Cancer)
Location: Remote (with occasional travel to consortium events)
Type: Full-Time | Non-Profit Consortium
Reports To: Executive Director
Posting Date: April 2026

ABOUT BLOODPAC

The Blood Profiling Atlas in Cancer (BLOODPAC) is a pre-competitive public-private consortium of 88+ industry, academic, and government partners united by a single mission: to accelerate the development, validation, and clinical use of liquid biopsy assays. Through collaborative working groups and prospective studies, BLOODPAC generates the evidence needed to advance liquid biopsy technologies — from circulating tumor DNA to multi-analyte blood-based biomarker profiling — toward regulatory acceptance and clinical utility.

BLOODPAC is currently re-envisioning its data strategy, shifting focus toward measuring and amplifying the scientific impact of its consortium activities. Central to this new direction is a publication portal — a modern web platform that tracks, curates, and showcases the research output stemming from BLOODPAC studies and member collaborations. Data science is evolving alongside this strategy: from infrastructure-heavy data curation to driving analytical insights, developing intelligent applications, and enabling member-facing tools that make the consortium's scientific contributions visible and accessible.

ROLE OVERVIEW

BLOODPAC is seeking an exceptional Director of Data Science to lead the next chapter of the consortium's data strategy. This is a hands-on technical leadership role with real ownership: you will be the primary architect and developer of BLOODPAC's data infrastructures and will define the analytical approach for prospective studies. Equally important is the ability to think strategically about where data fits into a maturing consortium — including the long-term path toward making study datasets accessible to the research community.

The ideal candidate is a builder first: someone who is energized by designing systems and writing production-grade code, conversant with AI-augmented development workflows, and capable of bringing scientific rigor to both data analysis and software design. Familiarity with the liquid biopsy or broader oncology data landscape is a significant advantage.

KEY RESPONSIBILITIES

Consortium Engagement & Scientific Leadership

- Serve as BLOODPAC's primary data science voice in working groups, steering committees, and external-facing forums including regulatory discussions with the FDA.
- Contribute to peer-reviewed publications writing, and regulatory submissions as a data science expert.

- Guide and mentor data contributors across consortium member institutions on best practices for data collection, documentation, and analysis.

Data Analysis Strategy for Prospective Studies

- Define and lead the analytical strategy for BLOODPAC-organized prospective studies, from study design through publication-ready results.
- Collaborate with working groups and consortium members to ensure data collection protocols are designed with downstream analysis in mind.
- Develop reproducible analysis pipelines for multi-omic liquid biopsy data (ctDNA, cfRNA, proteomics, epigenomics) and communicate findings to scientific and regulatory stakeholders.
- Translate analytical outputs into evidence that supports BLOODPAC's regulatory engagement and member communications.

Long-Term Data Strategy

- Develop and steward a roadmap for how BLOODPAC study datasets can be made accessible to the broader research community, balancing openness with appropriate governance, data security, and member interests.
- Evaluate and recommend frameworks, infrastructure, and partnerships that could support future data sharing — whether through federated models, controlled-access repositories, or platform integrations.
- Ensure data handling practices meet de-identification standards, and applicable regulatory requirements.

Publication Portal & Application Development

- Take ownership of the BLOODPAC Publication Portal — a platform for tracking, curating, and showcasing the consortium's scientific output — and continue its development from a working MVP toward a production-grade member-facing tool.
- Design and build new features end-to-end: AI-assisted metadata extraction, controlled vocabulary tagging, publication submission workflows, download gating, and impact analytics.
- Architect and code additional data applications, internal tools, and analytical dashboards as BLOODPAC's data strategy evolves.
- Leverage AI-assisted development tools — such as Claude Code, GitHub Copilot, or equivalent agentic coding frameworks — to accelerate delivery without sacrificing quality.
- Manage the full software development lifecycle: requirements, design, implementation, testing, deployment, and iteration, across a cloud-native stack.

REQUIRED QUALIFICATIONS

- Advanced degree (PhD or MS) in Bioinformatics, Computational Biology, Data Science, Computer Science, or a closely related field.
- Track record of first- or co-authored peer-reviewed publications in a relevant domain.
- Solid foundation in clinical research data, biospecimen management, or oncology research.
- Experience with genomic or multi-omic data: bioinformatics pipelines, sequencing data (WGS/WES/ctDNA-targeted panels), or similar workflows.

- Excellent written and verbal communication skills, with the ability to engage both technical and non-technical audiences.
- Ability to build and sustain relationships across large, multi-stakeholder networks — spanning academic institutions, federal agencies (NIH, FDA, DOD), and industry partners
- Knowledge of data management best practices and strategies for metadata harmonization including data access governance models, FAIR principles, or controlled-access data sharing infrastructure,
- 5+ years of hands-on software development experience in Python and/or TypeScript/JavaScript, with a track record of shipping production applications — not just analysis scripts.
- Demonstrated ability to design and build data-intensive applications, APIs, or platforms from the ground up.

PREFERRED QUALIFICATIONS

- Background in liquid biopsy technologies: ctDNA, cfDNA, circulating tumor cells, proteomics, or other blood-based biomarker modalities.
- Experience working in or with pre-competitive public-private consortia, multi-site clinical studies, or regulatory (FDA) engagement.
- Background in machine learning or statistical modeling for biomarker discovery, controlled vocabulary tagging, or clinical outcome prediction.
- Hands-on experience with AI-assisted coding tools — Claude Code, GitHub Copilot, Cursor, or similar agentic development frameworks — is a meaningful differentiator for this role.
- Familiarity with modern web application stacks (React, TypeScript, Supabase, Vercel, or comparable platforms) and/or cloud environments (AWS, GCP, Azure.)
- Demonstrated interest in policy initiatives, standards development, best practice or governance efforts related to data science, analytics, or healthcare research data.

WHAT WE OFFER

- A rare opportunity to shape the data and technology strategy of a consortium that sits at the intersection of cancer research, diagnostics, and regulatory science.
- The autonomy to build systems and tools that will define how BLOODPAC measures and communicates its scientific impact for years to come.
- Direct collaboration with leaders from pharma, diagnostics, academic medical centers, and government agencies including the FDA.
- A mission-driven, collaborative culture with flexible remote work and a genuine commitment to work-life balance.

HOW TO APPLY

Please submit a CV/resume and a brief cover letter describing your relevant experience and interest in BLOODPAC's mission. Links to public code repositories or relevant publications are welcomed. Applications may be sent to info@bloodpac.org.