

Job Offer: Data Engineer – R&D Multi-Omics & AI Infrastructure

About Brenus Pharma

Brenus Pharma is a French clinical-stage biotech company based in Lyon, France, developing next-generation in vivo immunotherapies targeting solid tumors, which are responsible for the majority of therapeutic failures in oncology. Our mission is to provide an effective treatment option for patients who have no alternative to chemotherapy.

Our proprietary Stimulated Ghost Cell (SGC) technology is designed to mimic cancer biology and educate the patient's immune system to recognize and eliminate malignant cells.

Our lead candidate, STC-1010, targets microsatellite-stable (MSS) colorectal cancer, a form that responds poorly to current immunotherapies, and entered a Phase I/IIa clinical trial in 2025. Supported by a €23M Series A round raised from leading European investors and public stakeholders (Bpifrance, France 2030, etc.), Brenus aims to become a key player in immuno-oncology by combining breakthrough innovation with tangible impact for patients.

The Role

We are looking for a driven Data Engineer to act as a core builder of our biological data infrastructure. Reporting to the scientific and bioinformatics leadership, your role will be highly cross-functional, sitting at the intersection of cloud infrastructure, bioinformatics, and upcoming AI initiatives. You will be responsible for operationalizing our data flows—translating complex wet-lab and multi-omics data into clean, scalable, and machine-readable assets to power our immunology models.

Your Key Missions

Mission 1 – Build & Scale Brenus' Sovereign Cloud Data Backbone

- Contribute to the design, structuring, and deployment of Brenus' cloud-based data environment to centralize internal scientific and bioinformatics data.
- Establish robust data organization, storage, documentation, and access principles following FAIR guidelines to improve traceability, reproducibility, and long-term maintainability.
- Ensure the day-to-day maintenance, reliability, and evolution of the data infrastructure in close collaboration with the bioinformatics team. Ensure data security and traceability align with European regulations (GDPR and HDS compliance).

Mission 2 – Connect & Automate Multi-Omics Data Pipelines

- Integrate internal datasets with relevant public biological datasets to support bioinformatics analyses and broader scientific exploration.
- Contribute to the structuring and operational use of workflows supporting RNA-seq / transcriptomics, proteomics, and other evolving omics datasets.
- Develop and automate robust ETL/ELT pipelines and data processes that improve data accessibility, quality control, reuse, and long-term robustness across projects.

Mission 3 – Enable the Future of AI/ML Application

- Help lay the foundations for AI-agent-enabled workflows and future AI/ML-driven data capabilities, in interaction with management / strategy and external AI or digital partners.
- Collaborate cross-functionally to translate complex scientific needs into scalable software solutions.

What You Bring (Profile & Skills)

Education & Experience:

- Engineer degree, Master's degree, or Ph.D. in bioinformatics, data engineering, computational biology, computer science, or a related field with exposure to life sciences.
- 3–5 years of experience in a data engineering, bioinformatics engineering, or related role, ideally in an industry environment (A relevant Ph.D. counts towards this experience).

Technical Stack:

- **Programming:** Strong command of Python, with familiarity with R to support bioinformatics integrations, Linux, and Git, along with experience with CI/CD pipelines (e.g., GitHub Actions, GitLab CI). Solid expertise in dataset design, querying, and managing both relational (SQL, Oracle) and non-relational (NoSQL) databases, including familiarity with S3-compatible object storage or columnar data formats (e.g., Apache Parquet), for high-performance data processing and automation.
- **Cloud Infrastructure & IaC:** Hands-on experience building and managing scalable cloud environments for sensitive data. Strong proficiency in automating deployments using Infrastructure as Code (IaC) tools like Terraform or Ansible. Direct experience with OVHcloud or similar sovereign European clouds is a major advantage.
- **Bioinformatics Pipelines & Reproducibility:** Hands-on experience containerizing scientific applications (**Docker / Singularity**) and building scalable, reproducible multi-omics workflows. Strong proficiency with industry-standard orchestrators like **Nextflow (nf-core)** or **Snakemake** to automate complex biological analyses.
- **Biological Data Fluency:** You possess a solid practical understanding of high-dimensional biological datasets and omics environments. You are comfortable structuring and processing RNA-seq / transcriptomic datasets, and you know how to navigate public biological databases (e.g., TCGA, GEO). Direct experience handling proteomics data or working within immuno-oncology is a significant advantage.

Soft Skills:

- Ability to work cross-functionally with bioinformatics, management / strategy, and external partners.
- You are a builder at heart, comfortable in a fast-paced, agile startup environment.
- Professional English required.

Culture, Values

By joining the team, you become part of a company where science, people and purpose move forward together. At Brenus, our culture is built on curiosity, learning, and the desire to grow, both individually and collectively. We value creativity, open-mindedness, teamwork, and the joy of sharing ideas, challenges, and new experiences. Collaboration, mutual support, and finding solutions together are at the heart of how we work every day. Beyond our projects, we aim to have a positive impact in healthcare: advancing scientific innovation by improving patients' lives.

- Location: Lyon, France (hybrid working model).
- Start: ASAP.
- Type of contract: CDI (Status Cadre).

To Apply

Send your CV to George Alzeeb at galzeeb@brenus-pharma.com, along with a brief explanation of your motivation for the role and why your background is relevant.