

## BIOINFORMATICS SCIENTIST

**Location:** Remote

**Type:** Full-time

**Organization:** The Align Foundation

### About Us

At the Align Foundation, we imagine a world with abundant scientific data. We believe that automation is key to scaling collaboration in life sciences – bringing people together to execute data-intensive, collaborative projects that result in predictive models of biology.

We are building scalable, cloud-native infrastructure for large open datasets and AI-ready pipelines, designed to accelerate discovery across the scientific community.

To help realize this vision, we are seeking a **Bioinformatics Scientist** who excels at connecting biology, statistics, machine learning, and cloud infrastructure. In this role, you'll collaborate closely with experimental scientists to design studies with high-quality, shareable data in mind from the outset. You'll develop robust, scalable analysis pipelines that make this data accessible and reusable by the broader scientific community, accelerating discovery and enabling reproducibility at scale.

### What You'll Do

- **Experimental Design & Statistics:** Collaborate closely with experimentalists to design robust biological experiments and apply statistical methods to generate high-quality, AI-ready datasets.
- **Pipeline Development:** Build and automate reproducible bioinformatics workflows using cloud-native infrastructure and scalable data processing tools.
- **Data Schema Design:** Develop extensible, structured data schemas to enable efficient storage, access, and analysis of biological datasets.
- **Programming & Tooling:** Write clean, production-quality code in Python.
- **Cloud Infrastructure:** Deploy and manage workflows and data in AWS cloud environments.
- **Data Management:** Implement and optimize SQL databases and structured data repositories.
- **Data Analysis:** Work with diverse biological datasets, including microbial, gene expression, and protein function data, and rapidly adapt to new data types as needed.
- **Visualization:** Create informative and high quality data visualizations for projects, partnerships and publications.

## What We're Looking For

- **Experience:** 3–5 years in bioinformatics, computational biology, biostatistics, or a related field.
- **Programming Skills:** Proficiency in Python and familiarity with version control using Github.
- **Cloud-Native Development:** Hands-on experience with cloud platforms and bioinformatics workflow orchestration (e.g., Nextflow, and Snakemake).
- **Database Knowledge:** Strong understanding of relational databases (SQL) and structured data models.
- **Scientific & Statistical Rigor:** Experience in experimental design, biological data analysis, and application of statistical methods.
- **Machine Learning Exposure:** Familiarity with fundamental machine learning or AI-related statistics, and an understanding of how structured biological datasets support AI model development.
- **Collaborative Mindset:** Ability to work cross-functionally with engineers, scientists, and domain experts in a mission-driven environment.
- **Bioinformatics Experience:** Familiarity extracting data from existing biological databases (e.g., PDB, UniProt and NCBI), working with protein structures (e.g., AlphaFold, PyMOL, and FoldX) and numerous bioinformatics softwares (e.g., Samtools, DESeq2, bedtools, jackhmmer, BLAST, HHblitz, and MMseqs).

## Nice to Have

- Experience analyzing large microbial, gene expression, or protein expression datasets.
- Experience designing or contributing to large public datasets or open science initiatives.
- Knowledge of containerization (Docker) and cloud orchestration tools.
- Practical experience supporting machine learning or AI model training pipelines.

## Why Join Us?

- **Mission-Driven Work:** Help make scientific data abundant, structured, and actionable — accelerating predictive biology.
- **Impact at Scale:** Shape datasets and pipelines that will power community-wide AI models in life sciences.
- **Collaborative Environment:** Join a technical, interdisciplinary team passionate about open data, automation, and scientific reproducibility.
- **Flexible, Remote-Friendly Workplace:** Work remotely with a focus on outcomes, innovation, and collaboration.

## Salary and Benefits

The Align Foundation offers competitive, market-based compensation, based on experience, skills, and location. For U.S.-based employees, we provide:

- Health Insurance (medical, dental, vision)



- 401(k) matching
  - Paid family leave (up to 13 weeks)
  - Paid time off, sick leave, and holidays (plus your birthday)
  - Monthly wellness stipend
  - Education reimbursement and childcare support
- (\*For non-U.S. employees, benefits are tailored to your location.\*)

### How to Apply

Submit your CV, a brief cover letter, and, if available, a sample of your work via this application form: <https://forms.gle/SgK5Y9nLYjxZQVng5>

We aim to fill this position by mid-Summer 2025.

### The Align Foundation welcomes everyone

We believe diversity enriches our team and so strive to hire people with a wide range of identities, backgrounds, and experiences.

We are an equal opportunity employer. That means we don't discriminate on the basis of race, religion, color, national origin, gender, sexual orientation, age, marital status, veteran status, or disability status.