

Senior Data Scientist



The Opportunity

As a Senior Data Scientist on the Omics Team at Pheno.AI, you will play a pivotal role in managing and analyzing large-scale omics datasets, significantly contributing to our cutting-edge research and product development. You will utilize clinical and biological data from the Human Phenotype Project (HPP), one of the world's most comprehensive collections of deep-phenotype multi-omics datasets—over 10,000 individuals across several dozens of measured phenotypes—to uncover insights that enhance human health. Collaborating closely with a multidisciplinary team comprising data scientists, life scientists, clinicians, and software engineers, you will design, develop, and implement advanced algorithms and machine learning models, such as predictive models, deep learning architectures, and integrative multimodal analyses. Your expertise will be instrumental in navigating the complexities of data processing, ensuring high-quality data handling, and adhering to rigorous scientific standards in line with best practices in the field.

Key Responsibilities

- **Data Ownership and Technical Expertise:** Lead the management of omics data, employing state-of-the-art methodologies to ensure accuracy, reproducibility, and integrity. Serve as the principal expert on omics datasets, guiding internal stakeholders and external partners on best practices and novel analytical approaches.
- **Data Quality Assurance:** Manage the end-to-end data lifecycle—from initial collection and processing in the lab or clinic to integration into analytical workflows—ensuring quality, consistency, and relevance for research applications.
- **Research Leadership:** Drive innovative research efforts involving omics datasets to generate impactful scientific discoveries.
- **Collaboration and Project Leadership:** Build and maintain strong relationships with academic institutions and industry partners, leading collaborative projects to accelerate scientific research and the application of omics-based insights to improve patient outcomes.

Requirements

- **Educational Background:**
 - Ph.D. in Computer Science, Bioinformatics, Computational Biology, or a related discipline, with at least 2 years of industry experience as a Data Scientist focused on healthcare or omics data analysis.
 - Alternatively, an M.Sc. in a computational discipline with 5 years of industry experience, including at least 2 years as a Data Scientist focused on healthcare or omics data analysis.
- **Experience in Analyzing Omics Data:** Demonstrated expertise working with diverse biological datasets at large-scale, including experience in one or more of the following areas:
 - Transcriptomics: e.g., bulk RNA-Seq, scRNA-Seq

- Genomics: sequencing data analysis (e.g., whole-genome sequencing, whole-exome sequencing), large-scale genome-wide association studies (GWAS), variant calling pipelines, or related NGS datasets and analytical workflows
- Microbiome: metagenomic sequencing, microbial functional profiling, etc.
- Metabolomics: e.g., mass spectrometry (LC-MS/MS)
- **Machine Learning Expertise:** Experience with applying advanced machine learning techniques to high-dimensional biological datasets, including designing, training, and evaluating supervised and unsupervised models.
- **Programming Proficiency:** Strong programming skills in Python, with deep familiarity in data science libraries and frameworks such as Pandas, NumPy, PyTorch, and scikit-learn.
- **Analytical and Communication Skills:** Ability to analyze complex data problems, develop actionable insights, and clearly communicate findings to multidisciplinary audiences.
- **Passion for Impact:** Deep enthusiasm and commitment to driving healthcare innovation through data-driven research.

Preferred Qualifications

- **Cloud Computing Experience:** Practical experience using cloud platforms (especially AWS) to manage, scale, and deploy data analysis workflows.
- **Deep Learning Experience:** Experience applying deep learning techniques (e.g., neural networks, transformers) to biological or clinical datasets.
- **Project Management:** Track record of successfully leading complex research initiatives from conception through to impactful outcomes.
- **Leadership and Mentorship:** Experience mentoring, coaching, and guiding junior scientists and fostering collaborative research environments.

About Pheno.AI

Pheno.AI is a pioneering biotech company based in Tel Aviv, Israel, focused on revolutionizing healthcare through the power of data. As the driving force behind the Human Phenotype Project (HPP), we are committed to collecting, organizing, and making accessible one of the world's most extensive longitudinal cohort and biobank initiative, designed to map the full spectrum of human health and disease. Our work enables researchers and clinicians worldwide to unlock new insights into human health, driving advancements in precision medicine, disease prediction, and drug development. By leveraging cutting-edge technology and a collaborative research environment, Pheno.AI is at the forefront of transforming how we understand and treat human health.

How to apply

Send your resume to hr@pheno.ai