

Senior Software Engineer – Genomics

To apply, visit the [UC Berkeley Jobs website](#) and keyword search for 17901, the Job ID for this position.

Note that the salaries listed on the [UC Berkeley Jobs site](#) are ranges for the Berkeley campus in general. We are able to make higher salary offers for qualified candidates. Berkeley also provides [a generous benefit package](#).

Departmental Overview

The AMPLab has been awarded a five-year \$10M, NSF “Expeditions in Computing” program grant, which was announced by the White House as part of their “Big Data” research initiative, a 4.5 year, \$5M Darpa XData contract and has raised \$6M to date from Industry sources. The director is Michael Franklin, co-principal investigators are Michael Jordan, Scott Shenker and Ion Stoica. The founding group included eight faculty members and dozens of Ph.D. students who were backed by 18 industry titans, including Google, Amazon Web Services and SAP. AMP researchers are developing powerful new tools to help extract key information from Big Data, a term coined for the dizzying array of measurements, images, audio, video, tweets, texts and more that has grown ever larger, faster and more diverse. A key part of AMPLab’s mission as promised to these sponsors is to deliver scalable, robust and reliable software that can be used by researchers inside and outside of Berkeley as well as by government, industrial and other organizations that need to make sense of massive amounts of data. The purpose of this position is to ensure that the software developed by the lab is of sufficient functionality, robustness and quality to meet these high expectations.

Next generation sequencing has created a deluge of genomic data that geneticists are wrestling to analyze with current software. The AMPLab is looking for 2 Senior Developers to build scalable systems to enable timely insights from these ever-increasing data sets. The AMPLab is the birthplace of the Berkeley Data Analytics Stack (BDAS) comprised of Apache Spark, Shark, GraphX, MLlib and more. As a member of the Global Alliance for Genomics and Health and with collaborators from Mt. Sinai, Oregon Health Sciences University, UC Santa Cruz, GenomeBridge, The Broad Institute and others, the AMPLab is in a unique position to apply “Big Data” technology to genomics to help find the cause and inform treatment of genetic diseases like Cancer. This is a great opportunity to apply your software engineering skills to make a huge impact on people’s lives. All software is released as Apache-licensed open source to benefit the larger community.

Responsibilities

70% Open-source software development on SNAP, ADAM and Avocado projects for alignment, pre-processing and variant calling respectively. This requires an understanding of Apache Spark, Hadoop and Avro as well as Parquet and the Scala programming language.

10% Coordinate with our clinical partners at Mt. Sinai, OHSU and others to ensure the systems we’re building have the features and performance characteristics that are needed.

10% Answer questions and provide technical support on the open-source mailing.

10% Work with the infrastructure and testing teams to ensure that end-to-end performance and concordance tests are working well and providing the metrics needed to monitor software development progress.

Required Qualifications

This is a two-year contract position with the possibility of being extended a third year

We are searching for two talented Senior Developer’s for the job.

Contract positions may be extended based on operational demand. Contract positions are eligible to participate in the health and welfare programs offered by UC Berkeley.

You should enjoy working in a collaborative environment with a diverse group of researchers, clinicians and engineers to solve hard problems in genomics, have experience in open-source systems and understand how to develop robust and scalable distributed systems.

- Strong programmer in Java and preferably Scala.
- Experience with Apache Hadoop, NoSQL systems and preferably Apache Spark.
- Working knowledge of existing software in genomics, e.g. samtools, Picard and GATK.
- Experience with cloud services like Amazon Web Services, Google Compute Engine or Microsoft Azure.
- A desire to learn more about Biology or, preferably, a working knowledge of genomic algorithms and analysis best practices.

Criminal Background Check

This position has been designated as sensitive and may require a Criminal Background Check. We reserve the right to make employment contingent upon successful completion of a Criminal Background Check.

Equal Employment Opportunity

The University of California is an Equal Opportunity/Affirmative Action Employer. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, disability, or protected veteran status. For the complete University of California nondiscrimination and affirmative action policy see: <http://policy.ucop.edu/doc/4000376/NondiscrimAffirmAct>

If you’d like more information about your EEO rights as an applicant under the law, please see: http://www1.eeoc.gov/employers/upload/eeoc_self_print_poster.pdf

Work Environment

As of January 1, 2014 the University of California, Berkeley is a smoke and tobacco free workplace. Information about our Smoke and Tobacco Free policy is available at:

<http://www.tobaccofree.berkeley.edu/>