

INSTITUTE FOR SOCIAL RESEARCH & DATA INNOVATION



EMPLOYMENT OPPORTUNITY

Be a part of something BIG.

Data Engineer

JOB OPENING ID 324988

IPUMS (<https://www.ipums.org/>) is a leader in the field of quantitative social science research and the largest disseminator of census and demographic data to the world's academic research community. Or, to put it another way - we're on a mission to gather, process, link and publish billions of records spanning hundreds of years and more than 100 countries to demographers, historians, economists, environmental scientists, journalists, policymakers, and others around the globe, who then use the data to do amazing research and make the world a better place.

The IPUMS IT group supports this mission by using leading open source tools to solve complex data and computation challenges and build reliable, scalable web-based data dissemination systems. Your work will be highly visible and will contribute directly to the overall success of our organization. Read more about IPUMS IT at <http://tech.popdata.org/>. IPUMS is part of the Institute for Social Research and Data Innovation (<http://isrdsi.umn.edu>) and closely connected to work at the Minnesota Population Center.

IPUMS and its affiliated units support the work-life balance of our staff with 40 hour work weeks, flexible work hours, and generous vacation and sick leave benefits. The University also offers excellent health insurance, tuition assistance, and retirement benefits. IPUMS IT has a robust professional development fund for staff training and development.

Diversity and inclusion is a core value of our organization. We aspire to create an IT team that represents the diversity of our city, our region, and our world, and to create a space that encourages and embraces inclusiveness, equal opportunity, and respect. We strongly encourage women and members of under-represented groups to apply.

ABOUT THIS ROLE

We are currently seeking a software developer to join our data linking team that is working on large scale data manipulation and creating highly performant distributed software systems. In this role, you will be developing software that links records representing the same individual and their family relationships across every US Census from 1850 - 1940. There are no consistent identifiers across historical censuses, so sophisticated matching algorithms will be required. You will be building on research and IT effort from the past decade to create a dataset that, once released, will enable critical new avenues of research for demographers and historians. This work will also lead to additional data linking opportunities for datasets from around the world and throughout history. You will be working in close collaboration with expert historians, demographers, and data scientists. We use Apache Spark (PySpark and SparkSQL) for a majority of our data processing, with C++ in critical areas. This is an excellent opportunity to grow your career with a cutting -edge software shop at the University of Minnesota and contribute to work that has impact around the globe.

This position will have an annual starting salary of \$78,000+, commensurate with experience.

RESPONSIBILITIES

This position is part of our Microdata Production team and will have responsibility for large-scale datasets and data processing tools surrounding the linking work. This role will work throughout the project life cycle, from architecture and design through implementation and testing to deployment and support. We practice Agile and collaborative software development, which means you will have ample opportunity to work alongside both developers and researchers.

30% Software Architecture and Design. Working in partnership with researchers and data scientists defining business needs, architecting data linking environments, and designing specific data linking systems.

50% Software Implementation and Analysis. Coding, refactoring, testing, executing and analyzing data linking pipelines and linked data outputs in a cross-functional, agile team environment.

10% Deployment and Support. Working with operations team to build out high performance infrastructure to support new data linking pipelines. Developing deployment processes to continuous integration, internal, and production environments. Providing user support to the team and our researchers.

10% Other duties as assigned. Professional development activities, participation in IT working groups, and other tasks as assigned.

QUALIFICATIONS

Required Skills: BA/BS degree required. Two years of work experience in the areas of application/web/systems development with a related BA/BS degree, or four years of work experience in the areas of application/web/systems development with a non-related BA/BS degree.

This role requires technical experience and proficiency with:

- Experience developing software for high performance and/or distributed computing
- Experience with databases (relational or NoSQL)
- Large dataset manipulation
- Linux/UNIX operating systems, including command-line

This role also requires excellent ability to plan work and manage complex tasks/projects. The successful candidate will demonstrate excellent oral and written communication skills for both technical and non-technical academic audiences.

Additional Skills:

Candidates with some of the following skills and/or experience are preferred:

- Familiarity with agile development methods
- Modern data science techniques and approaches
- Developing user-friendly command-line tools
- Machine learning
- Apache Spark or the Hadoop ecosystem
- Helpful language experience: Python, Scala, C/C++
- Prior work experience in a research and/or higher education setting

APPLICATION PROCEDURE

Please apply using the University of Minnesota's online employment system humanresources.umn.edu/jobs and search job opening ID 324988 or use this short url: z.umn.edu/dataengineer2018. Application requirements include a resume, and a cover letter describing your interest and qualifications in the position. Questions concerning the application process may be addressed to Mia Riza, HR Generalist, at mpc-jobs@umn.edu.

Any offer of employment is contingent upon the successful completion of a background check. Our presumption is that prospective employees are eligible to work here. Criminal convictions do not automatically disqualify finalists from employment.

The University of Minnesota shall provide equal access to and opportunity in its programs, facilities, and employment without regard to race, color, creed, religion, national origin, gender, age, marital status, disability, public assistance status, veteran status, sexual orientation, gender identity, or gender expression.