The Data Scientist will partner with researchers utilizing high performance computing to solve "big data" problems in their research including: interacting with novel big data architectures; the evaluation and implementation of new technologies and algorithms; and the development of materials for documentation and training. This individual will utilize experience, creativity and innovation to solve conceptual computing problems raised by cutting edge research.

As Data Scientist you will diagnose a wide range of data problems; design, code, test, debug, install, document and maintain complex programs. You will partner with Purdue researchers to understand what value and insights are possible and may be derived from their data. The Data Scientist will author and submit work for scientific publications as well as provide technical direction and leadership to ITaP staff and ensure the quality of RCAC services.

Purdue Research Computing is a research computing center providing advanced computing resources and support services including access to leading-edge computational and data storage systems, as well as expertise in a broad range of high performance computing activities, to support the computationally-intensive research of Purdue faculty, staff and research partners nationwide. For more information, visit our website: https://www.rcac.purdue.edu/.

Required:
- Bachelor's degree.
- Five years of experience in analyzing scientific application requirements in big data including resolving problems with the interaction of applications, applications and hardware, and data management.
- In lieu of degree, consideration will be given to an equivalent combination of related education and required work experience.
- Experience using high-performance computing resources to produce scientific results, or in support of high-performance computing resources.
- Experience with software design and implementation including the development of programs and software tools necessary to create big data and high-performance computing applications and scientific workflow.
- Experience with distributed computing, parallel programming, and big data software; specifically MPI, OpenMP, MapReduce, Spark or Hadoop.
- Experience working with large volumes of data, developing custom analysis methods, and optimizing performance of high-performance computing applications.
- Experience with higher-level languages such as R, Matlab and Python.
- Knowledge related to parallel and distributed computing.
- Knowledge of and skill using a wide range of hardware, scientific programming languages, scripting languages, commercial and public domain software packages, operating systems, and tools.
- Ability to work effectively both individually and as part of a team.
• Strong verbal and written communication skills including the ability to consult with and provide support for researchers, make presentations, develop and present training material, and develop documentation.

Preferred:
• Bachelor's degree in engineering, math, physical or life sciences, statistics, computer science or related field.
• Graduate degree.

Additional Information:
• A background check is required for employment in this position.
• FLSA: Exempt (Not Eligible for Overtime)
• Retirement Eligibility: Defined Contribution Waiting Period.
• Purdue University is an EEO/AA employer. All individuals, including minorities, women, individuals with disabilities, and protected veterans are encouraged to apply.

Apply online: http://purdue.taleo.net/careersection/wl/jobdetail.ftl?lang=en&job=1701407