

Multiple PhD Openings in Big Data Tools and Infrastructure R&D at UNC Charlotte (Starting Immediately–Fall '16, Spring '17, Summer '17)

Contact Email: atalukde@uncc.edu

Web: <http://cs.uncc.edu/directory/talukder-ashit>

Prof. Ashit Talukder, Professor and Director, Data Visualization Center, UNC Charlotte
Director, Data Science Initiative, UNC Charlotte-CCI

Please email resumes, research experience/interests in word or PDF format, GRE/TOEFL scores, and scan of transcripts with subject line: PhD openings in Big Data Tools and Infrastructure R&D

Summary PhD Research Description

Data Scientist, Big Data/Data Science System Middleware and Tools Research and Development, Data Science and Analytics System R & D, Data Mining, NLP , Semantic and Entity Mining.

Desired qualifications: MS in Computer Science, EE, ECE, Information Science, or related area with interests and background in Data Science, HPC, Data Management, Software Engineering.

Candidates will work on a large scale Big Data research and development program led by Professor Ashit Talukder (Director of Data Science at CCI and Director of the UNC Charlotte Data Visualization Center), in collaboration with multiple universities and research labs across the US. The PhD students will be an integral part of the project and member of the Charlotte Data Visualization Center, comprised of many core faculty working on research in Data Science, Data Analytics, Visual analytics, and Visualization. The candidates will work with other PhD students and closely interact with faculty and post docs at the Center and at UNC. The focus of the PhD's research will be on research and development of Big Data and Data Analytics software building blocks, tools and algorithms to enable distributed analytics across distributed repositories with distributed search, information retrieval, and semantic mining capabilities.

All candidates should have an inherent ability and passion for programming and implementation of concepts and algorithms. It is also desirable if candidates have experience/knowledge with some of the following areas.

Desirable (but Not Necessary) Knowledge and Experience with the Following Topics:

- Experience with OS and platform middleware, multiple OS, multiple programming languages
- Experience with or knowledge of VMs and containerization (dockers etc)
- Knowledge of cloud programming and development tools and environments
- Knowledge of HPC systems and administration of IT systems
- Knowledge or awareness of open source tools that are useful to build custom data-access, storage, processing environments including tools such as: Cloudera secured Hadoop platform, the Apache KNOX network security service, the Cloudera Navigator data management service, the Integrated Rule-Oriented Data System (iRODS) authorization and metadata management system, the Sentry role-based data authorization service, and the Dataverse Networks data catalog system
- Expertise in, or knowledge of, distributed data analytics algorithms and implementation of analytics tools and environments
- Knowledge of tools for end to end data lifecycle management and visualization, including data storage, management, curation, indexing, search and information retrieval, query and APIs
- Knowledge of SOA and REST web services for scalable systems