Kno.e.sis Invited Guest Speaker Series

Title: Scalable Querying of Semantic Web Data Models: Challenges and Opportunities
Speaker: Kemafor Anyanwu
Date: March 29, 2013
Time: 11 am
Place: 292 Joshi (Brandeberry)

Abstract: Recent advancements in Semantic Web publishing technologies are ushering in the era of "big Semantic Web data". We now have a rapidly growing number of publicly available Semantic Web datasets spanning a variety of domains. Attempting to harness the collective knowledge represented in combinations of these datasets, quickly gives rise to a "big and heterogeneous data" problem.

Although, there are numerous research activities now focused on "big data", many of them focus on data that is fairly homogeneous and either completely structured or completely unstructured. However, Semantic Web data models impose very few constraints on structure compared to relational models, leading to graph-like or semi(structured) nature. Further, Semantic Web data collections often contain data expressed using multiple vocabularies or schemas requiring processing that involves reasoning about equivalences that are not explicitly stated. These requirements lead to workloads that do not match assumptions made by traditional relational query optimization techniques. Unfortunately, many existing Semantic Web data query processing techniques still rely on relational-like query optimization techniques and, the impact of this mismatch is much more palpable at large scales. The use of cloud computing platforms to support elastic scaling requirements adds yet another dimension of challenges because several key assumptions underlying traditional query optimization techniques do not transfer to popular platforms like Google's MapReduce. In this talk, I will present an overview of the efforts being undertaken by the Semantic Computing Research Lab at North Carolina State University to address some of these challenges and highlight some open research opportunities.

Bio: Kemafor Anyanwu is an Assistant Professor of Computer Science and director of the Semantic Computing Research Lab at North Carolina State University. She received a Ph.D. in Computer Science from the University of Georgia in 2007. Her research interests include Semantic Web data management, data analytics and mining, and their applications. The two themes of her research activities revolve around developing optimization techniques for large scale Semantic Web data processing and developing query primitives and languages for supporting more sophisticated querying on the Semantic Web. She has served on program committees of different tracks of conferences such as ISWC, ICDE, ICSC and was on the organizational committee for WWW2010 held in Raleigh. She reviews for journals such as TKDE, IJSWIS and has been a guest editor for IJSWIS. Her work with her student received the best paper award in JIST 2012. Her work is funded by grants from the NSF and industry awards like the IBM Faculty awards.
http://research.csc.ncsu.edu/coul/