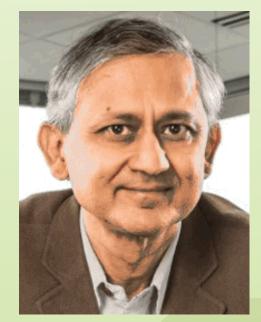## An IBM Research Distinguished Speaker Series

Tuesday, July 1, 10:00 - 11:30 A.M.
Accelerated Discovery Lab (H2-200)

Transforming Big Data Into Smart Data

Speaker: Amit P. Sheth
Ohio Center of Excellence in Knowledge-enabled Computing,
Kno.e.sis Wright State University, Dayton, Ohio

Big Data has captured a lot of interest in industry, with anticipation of better decisions, efficient organizations, and many new jobs. Much of the emphasis is on the challenges of the four Vs of Big Data: Volume, Variety, Velocity, and Veracity, and technologies that handle volume, including storage and computational techniques to support analysis (Hadoop, NoSQL, MapReduce, etc.) However, the most important feature of Big Data, the raison d'etre, is none of these 4 Vs -- but value. In this talk, I will forward the concept of Smart Data that is realized by extracting value from a variety of data, and how Smart Data for growing variety (e.g., social, sensor/IoT, health care) of Big Data enable much larger class of applications that can benefit not just large companies but each individual. This requires organized ways to harness and overcome the four V-challenges. In particular, we will need to utilize metadata, employ semantics and intelligent processing, and go beyond traditional reliance on ML and NLP.

For harnessing volume, I will discuss the concept of Semantic Perception, that is, how to convert massive amounts of data into information, meaning, and insight useful for human decision-making. For dealing with Variety, I will discuss experience in using agreement represented in the form of ontologies, domain models, or vocabularies, to support semantic interoperability and integration. Lastly, for Velocity, I will discuss somewhat more recent work on Continuous Semantics, which seeks to use dynamically created models of new objects, concepts, and relationships and uses them to better understand new cues in the data that capture rapidly evolving events and situations.

Smart Data applications in development at Kno.e.sis come from the domains of personalized health, energy, disaster response and smart city. I will present examples from a couple of these.

Amit P. Sheth is an educator, researcher, and entrepreneur. He is the LexisNexis Eminent Scholar and founder/executive director of the Ohio Center of Excellence in Knowledge-enabled Computing (Kno.e.sis). Kno.e.sis conducts research in social/sensor/semantic data and Web 3.0 with real-world applications and multidisciplinary solutions for translational research, healthcare and life sciences, cognitive science, material sciences, etc. Kno.e.sis' activities have resulted in Wright State University being recognized as a top organization in the world on World Wide Web in research impact. Prof. Sheth is one of top authors in Computer Science, World Wide Web and databases (cf: Microsoft Academic Search; Google H-index=85). His research has led to several commercial products, many real-world applications, and two earlier companies with two more in early stages. One of these was Taalee/Voquette/Semagix, which was likely the first company (founded in 1999) that developed Semantic Web enabled search and analysis, and semantic application development platforms. He is founding EIC of IJSWIS and co-EIC of IJSWIS and DAPD.