

## Information Science and Technology Seminar Speaker Series



Cliff A. Joslyn

Pacific Northwest National Laboratory

### Semantic Graph Database Analytics: Complex is the New Big Data

Wednesday, December 18, 2013

3:00 - 4:00 PM

TA-3, Bldg. 1690, Room 102 (CNLS Conference Room)

**Abstract:** Two of "big data"'s canonical "Four V" values are Volume and Variety (the others being Velocity and Veracity). Addressing high volume information which is simple and homogeneous is itself enough of a problem; but when combined with complexity and heterogeneity, computational challenges can be exponentiated. PNNL is advancing the concepts and technologies around Semantic Graph Databases (SGDBs) as a NoSQL database paradigm for handling big, complex, open-world data. Where traditional, relational databases (RDBs) are structured as rows and columns, typed by schemata, and queried through SQL, SGDBs are structured as labeled nodes and edges, typed by ontologies, and queried through graph-like languages including SPARQL. While not all graph data are best addressed in SGDBs, SGDB concepts and structures arise naturally when graph data have directionality and a variety or complexity of labels and attributes carrying the semantic and logical content of the information being represented. In this talk I will introduce SGDB concepts and survey recent PNNL work, including PNNL's SGEM engine for high performance SGDBs; our descriptive semantic statistical approaches to characterizing large SGDBs; the role and value of hybrid graph/relational data models; the search for benchmark test data sets; and the mathematical issues involved in extending network science approaches to the labeled, directed graph structures typical of SGDBs.

**Biography:** Cliff Joslyn is the Chief Scientist for Knowledge Sciences in the National Security Directorate of the Pacific Northwest National Laboratory in Seattle, Washington. Previously he was with LANL from 1996-2007, and was a Team Leader for Knowledge and Information Systems Sciences in the Computer Science Division; and an NRC Research Associate at NASA's Goddard Space Flight Center from 1994-1996. Dr. Joslyn holds a BA in Cognitive Science and Mathematics with High Honors in Cybernetics and Science of Mind from Oberlin College, and an MS and PhD in Systems Science from SUNY Binghamton. A computer professional for over 30 years, his research interests include discrete mathematics and applied order theory, knowledge discovery in databases, theoretical cybernetics, ontology management and knowledge systems, computational semiotics, qualitative modeling and simulation, and generalized information theory.