Introduction to the HICSS-49 Data Science for Collaboration Minitrack

Lakshmi Iyer
Information Systems and Supply Chain Management (ISSCM)
The University of North Carolina Greensboro
435 Bryan Building
Greensboro, NC 27402-6165
Lsiyer@uncg.edu

Souren Paul
Graduate School of Computer & Information Sciences
Nova Southeastern University
3301 College Avenue
Fort Lauderdale, FL 33314
Souren.paul@gmail.com

Lina Zhou
Information Systems Department
University of Maryland Baltimore County
Baltimore, MD 21250
zhoul@umbc.edu

Driven by the ever-rising phenomenon of big data, data science has emerged to fulfill the critical need of analyzing the data at great breadth and depth and communicating the results of data analysis to various stakeholders so as to improve their decision making and enhance their competitive advantages. There is exploding interest in organizations looking for ways to increase value from data science and using it to address business challenges. However, there is a dearth of discussion of its role in the context of collaboration. This minitrack fills the void that focuses on data science for collaboration.

Data science for collaboration is the study of the generalizable extraction of knowledge from data to support human collaboration within and across groups and organizations. The new knowledge gained is expected to be actionable for achieving collaborative goals such as generating, choosing, negotiating, and executing. Data science for collaboration couples a systematic study of collection, aggregation, organization, processing, and analysis of data. In addition, it requires deep understanding of formulating problems valuable for collaboration, engineer effective solutions to the collaboration problems, and ways to effectively communicate findings across roles ranging from business managers to data analysts. The emerging heterogeneous, voluminous, and unverified data present both opportunities and new challenges for addressing collaboration problems.

The inaugural session of the minitrack will include one paper session, consisting of three papers. These papers cover the following topics of interest to the minitrack: data science for internal collaboration in groups and organizations, data science for inter-organizational collaboration, crowdsourcing for collaborative tasks, security and privacy in collaborative data science, and data science in collaborative creation.

The first paper, "*Collaborative Information Service: The Security Question*" by Asim Kumar Pal, Subrata Bose, has investigated the problem of collaborative computing for an information service. They study the role of a flexible privacy model that can simplify the job of security in a collaborative environment.

In the second paper, "*An Architecture to Enhance Collaboration in Scientific Software Product Line*", Anrafel F. Pereira, José Maria N. David, Regina Braga, Fernanda Campos propose and evaluate a new framework named Collaborative PL-science by introducing collaboration features for improving awareness and communication among scientists on top of the existing PL-science core. The collaboration features address a long-time challenge for scientists who work together from distributed geographical locations.

To handle the big data in text, particularly scientific literature, Diogo Nolasco Ferreira Sousa, Jonice de Oliveira Sampaio propose a method based on a popular text mining technique for automatic labeling topics of scholar data in the third paper, "Detecting Knowledge Innovation through Automatic Topic Labeling on Scholar Data". The method can contribute to knowledge innovation through research collaboration and topic discovery.

These papers will help to stimulate thinking about future research agenda in data science for collaboration. Many thanks to Jay Nunamaker and Bob Riggs for their guide in setting up the minitrack.

1530-1605/16 $31.00 © 2016 IEEE
DOI 10.1109/HICSS.2016.48