

Introduction to the HICSS-51 Data Science and Digital Collaborations Minitrack

Lakshmi S. Iyer
Appalachian State University
Boone, NC 28608
iyerLs@appstate.edu

Souren Paul
Nova Southeastern University
Fort Lauderdale, FL 33314
Souren.paul@gmail.com

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

Data science and digital collaboration is a systematic study of collection, aggregation, organization, processing, and analysis of data. In addition, it requires a deep understanding of formulating problems valuable for collaboration, and engineering effective solutions to the collaboration problems and ways to effectively communicate findings across roles ranging from business managers to data analysts. Organizations have an explosive interest in looking for ways to increase value from data science and using it to address business challenges. One promising way for businesses and organizations to enhance their performance or competitiveness is by investigating how data science and analytics can facilitate collaboration both internally and externally. For example, businesses are trying to understand how data science and analytics can help engage customers and improve operation efficiency and how it can use social media to support corporate knowledge management. Another area is collaborative generation and creation of ideas and solutions through crowdsourcing and online communities (such as dominodatalab.com). Access to heterogeneous, voluminous, and unverified data presents both new opportunities and challenges for addressing collaboration problems.

The third offering of the minitrack includes one paper session, consisting of three papers covering the following areas of interest: use of enterprise crowdsourcing system (ECS) for enhancing knowledge sharing for collaboration, social collaboration for visual analytics design, and a communication model for data miners and domain users. The first paper, “The Effect of Enterprise Crowdsourcing Systems on Employees' Innovative Behavior and Job Performance”, aims to understand how factors of organization hosted ECS can impact employees in sharing innovative ideas to improve job performance. This is an important and relevant context for digital collaboration as gathering innovative ideas from employees continues to be a challenge in organizations. A survey of 183 employees and their managers reveals that creative knowledge application (CKA) with ECS increase

employees' satisfaction on ECS and their innovative behavior. In addition, the effect of CKA on job performance is mediated mainly by ECS satisfaction. The preliminary findings of this study indicate what organizations can do to refine their ECSs and innovation initiatives.

The second paper, “Topic Analysis through Streamgraph via Shiny Application: a Social Collaborative Approach”, addresses the challenges of integrating humanistic components into text mining and visualization by introducing a reactive web framework that allows for interactive data analysis and visualization. The development of the framework drew on the philosophical notion of rhizomatic model that follows principles of multiplicity and heterogeneity. The framework organizes ideas around concepts and tasks, while preserving a fluidity of conversation among team collaborators. By recognizing social interaction as a part of visualization system design, the framework supports interactive user-driven customization. In addition, the research develops workflow to improve collaboration and development at all stages of a project. The advantages of the proposed framework are demonstrated by developing applications via a direct collaboration with researchers.

The third paper entitled “A Communication Model that Bridges Knowledge Delivery between Data Miners and Domain Users”, proposes a communication model to facilitate knowledge delivery between data miners and domain users. When data miners and domain users come from different fields of knowledge, effective communication of the results of data mining may not be achieved. The author(s) suggest the role of a data liaison specialist to foster collaboration between data miner and domain user. They propose a multi-stage communication model between data miner and domain users and explain how data liaison specialist can facilitate the communication at each stage. These papers will help foster new ideas and future research agenda in data science and digital collaboration.