Project Summary

Title: TUES: EAGER: Scaffolding Big Data for Authentic Learning of Computing

We propose to create curriculum and technologies that leverage "Big Data" to create authentic and engaging learning experiences in computational thinking and computer science courses. We ask for support through the EAGER process because we have a unique, time-sensitive opportunity at Virginia Tech to profoundly impact the general education curriculum through a new non-major's Computational Thinking course, whose first offering will be in the Fall 2014 semester. The materials developed for Computational Thinking also will be used to improve introductory courses in Computer Science, extending our prior work that used real-time data sources. Further, collaborating with Virginia Tech's Department of Engineering Education, we will explore how our work could support authentic and motivating introductory computing experiences for engineering students.

We address the critical challenge of weaving together curriculum, pedagogy, and tools to engage groups of learners that have starkly different interests and expectations about computation. Our solution uses authentic project-based experiences that engage multi-disciplinary groups of learners with computation. We leverage the widespread availability of "Big Data" sources from the Internet. But to provide Big Data resources to students without a programming background requires careful scaffolding of the technology to manipulate Big Data streams. With the proper scaffolding, meaningful and motivating projects can easily be deployed by instructors to quickly and seamlessly incorporate this compelling dimension into new learning experiences. This proposal builds on the success of the RealTimeWeb project, our framework for rapidly building real-time, web-data centered assignments in introductory CS courses.

The proposed work will (1) develop novel curriculum resources including interactive learning materials integrating graphical programming, visualization, and in-line execution; (2) expand and enhance the RealTimeWeb with new Big Data streams and interfaces to visualization and other critical services; and (3) develop, apply, and analyze an extensive set of assessment measures including assessments related to achievement of learning objectives, student motivation, and the dynamics of student cohorts.

Intellectual Merit: By introducing authentic, massively-sized datasets of relevance into the early undergraduate curriculum, we can create a more engaging experience for students where they develop a foundational knowledge of Big Data concepts. This project will provide an excellent opportunity to improve the theoretical understanding of the problems and best practices for teaching Big Data, even as it offers a practical method for individual instructors to begin using these resources. This project will also provide insights about the challenge of designing engaging and relevant contexts for projects that positively impact CS learning outcomes, while simultaneously enabling students to work with one of the most important technologies of the modern era.

Broader Impact: This project will improve recruitment, retention, and engagement of students within STEM disciplines. Prior research indicates that women are more likely to study and excel in Computer Science when content is contextualized and proven useful for solving real-world problems. Non-major students can be given realistic assignments that can more directly relate to their intended line of work, further increasing their motivation to succeed during their time in a CS course. Appropriately redesigning programming projects to involve interesting, contextualized Big Datasets is likely to improve the relevance and attractiveness of Computer Science to a wider community.