Within the architecture, engineering, and construction (AEC) domain, simulation modeling is mainly used to facilitate decision-making by enabling the assessment of different operational plans and resource arrangements, that are otherwise difficult (if not impossible), expensive, or time consuming to be evaluated in real world settings. The accuracy of such models directly affects their reliability to serve as a basis for important decisions such as project completion time estimation and/or resource allocation. Compared to other industries, this is particularly important in construction and infrastructure projects due to the high resource costs and the societal impacts of these projects. Discrete event simulation (DES) is a decision making tool that can benefit the process of design, control, and management of construction operations. Despite recent advancements, most DES models used in construction are created during the early planning and design stage when the lack of factual information from the project prohibits the use of realistic data in simulation modeling. The resulting models, therefore, are often built using rigid (subjective) assumptions and design parameters (e.g. precedence logic, activity durations). In all such cases and in the absence of an inclusive methodology to incorporate real field data as the real project evolves, modelers rely on information from previous projects (a.k.a. secondary data), expert judgments, and subjective assumptions to generate simulations that can reasonably predict future performance. These and similar shortcomings have to a large extent limited the use of traditional DES tools to preliminary studies and long-term planning of construction projects.

The goal of this doctoral research is to investigate the design and implementation of a framework capable of (1) capturing multi-modal process data from field agents (i.e. construction resources) using ubiquitous mobile sensors, and (2) extracting computer-interpretable knowledge through vigorous data mining and machine learning techniques to update DES models corresponding to the real system operations. To achieve this goal, a preliminary study was conducted to demonstrate the feasibility and applicability of data-driven knowledge-based simulation modeling with focus on data collection using a wireless sensor network (WSN) and a rule-based taxonomy of activities. The resulting knowledge-based simulation models performed very well in properly predicting key performance measures of real construction systems. Next, a pervasive mobile data collection and mining technique was adopted and an activity recognition framework for construction equipment and worker tasks was developed. Data was collected using smartphone accelerometers and gyroscopes from construction entities to generate significant statistical time and frequency domain features. The extracted features served as the input of different types of machine learning algorithms that were applied to various construction activities. The trained predictive algorithms were then used to extract activity durations and calculate probability distributions to be fused into corresponding DES models. Results indicate that the generated data-driven knowledge-based simulation models outperform the static models created based upon engineering assumptions and estimations with regard to compatibility of performance measure outputs to reality.