Using Jaccard Similarity to Identify New Issues from AEC Project Team Meeting Minutes

**Hasan Gokberk Bayhan,[1] Yao Ma,[2] Joseph Thekinen, Ph.D.,[3]
Jiliang Tang, Ph.D.,[4] and Sinem Mollaoglu, Ph.D., A.M.ASCE[5]**

[1]Ph.D. Student, Construction Management Program, School of Planning, Design, and Construction, Michigan State Univ., East Lansing, MI 48824; e-mail: bayhanha@msu.edu
[2]Ph.D. Candidate, Department of Computer Science and Engineering, Michigan State Univ., East Lansing, MI 48824; e-mail: mayao4@msu.edu
[3]Research Associate, School of Planning, Design, and Construction, Michigan State Univ., East Lansing, MI 48824; e-mail: thekinen@msu.edu
[4]Assistant Professor, Department of Computer Science and Engineering, Michigan State Univ., East Lansing, MI 48824; e-mail: tangjili@msu.edu
[5]Professor, Construction Management Program, School of Planning, Design, and Construction, Michigan State Univ., East Lansing, MI 48824; e-mail: sinemm@msu.edu

ABSTRACT

Keeping track of issues and their documentation in Architecture, Engineering, and Construction (AEC) projects demand significant amounts of time, budget, and effort. While various types of documents and software aid coordination in AEC projects, project team meeting minutes, developed as a follow-up to periodic project team meetings, continue to be the most common and prominent type of documentation across project types for recording team communications, tasks, and assignments. Presently, due to its dynamic nature, identifying unique project issues and tracking their progress from meeting minutes is a manual process that is time-consuming and susceptible to error. This study aims to automate the identification of project issues and track resolution timelines using project team meeting minute documents via the Jaccard Similarity method. In this study, over 50 AEC project team meeting minutes documents of varying formats from three different projects of various sizes were collected, automatically converted, and coded to train the Jaccard Similarity model for detecting new and continuing issues. Accuracy, precision, recall, and F1 parameters were tested, and the accuracy rates of 81.86% to 94.18% were obtained. The study provides the groundwork to automate the analysis of issue complexity, detection of bottlenecks, and analysis of expertise assignments for issue resolution.