Neural Language Model based Intelligent Semantic Information Retrieval on NCDOT Projects for Knowledge Management

Siddharth Banerjee, S.M.ASCE[1], Colin M. Potts[2], Arnav H. Jhala, Ph.D.[3], Edward J. Jaselskis, Ph.D., P.E., A.M.ASCE[4]

[1]Graduate Research Assistant, Department of Civil, Construction and Environmental Engineering, North Carolina State University, Raleigh, NC, 27695; email: sbaner22@ncsu.edu
[2]Graduate Research Assistant, Department of Computer Science North Carolina State University, Raleigh, NC, 27606; email: cmpotts@ncsu.edu
[3]Associate Professor, Department of Computer Science, North Carolina State University, Raleigh, NC 27606; email: ahjhala@ncsu.edu
[4]E.I. Clancy Distinguished Professor, Department of Civil, Construction and Environmental Engineering, North Carolina State University, Raleigh, NC 27695; email: ejjasels@ncsu.edu

ABSTRACT

The North Carolina Department of Transportation (NCDOT) created a new knowledge repository called Communicate Lessons, Exchange Advice, Record (CLEAR) as an official platform for end-users to store and retrieve knowledge. Through the CLEAR program, end-users can enter lessons learned and best practices gained in their workplace in addition to soliciting solutions to any ongoing issue. This paper briefly reviews the development of CLEAR and proposes an intelligent knowledge transference process of information on NCDOT projects using natural language processing and knowledge graphs based on neural language models developed by the CLEAR project team. The CLEAR project includes a collection of documented lessons learned and best practices. The AI model learns an inference model of the domain vocabulary from various sources such as contract documents, textbooks, and specifications. This model allows the system to make meaningful connections between lessons learned and best practices within CLEAR and the project-specific domain knowledge. The model output will initially be shown to NCDOT team members belonging to various project lifecycle phases such as design, construction, and maintenance to certify the usefulness of the generated keywords and thereby the AI model in an iterative manner until the model has been appropriately fine-tuned. Necessary modifications will be made to the model based on the feedback obtained from project personnel to ensure high-quality output. In the long run, this automation in information retrieval will encourage NCDOT personnel to use the CLEAR program as a part of their routine work to improve project workflow processes.