Organised by Geotechnical Engineering Technical Division (GETD)

The Institution of Engineers, Malaysia

WEBINAR PRE-AGM TALK ON

"FUTURE OF MACHINE LEARNING IN GEOTECHNICS"

BEM APPROVED CPD HOURS: 2 REF. NO.: IEM24/HQ/199/T (w)

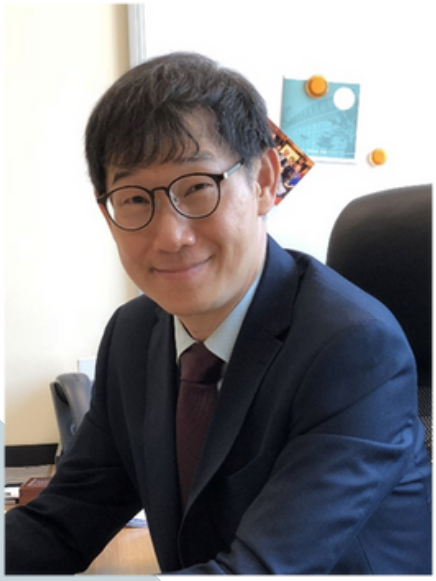
Presented by:

Prof. Phoon Kok Kwang

4 JUNE 2024, TUE
10.00AM - 12.00PM
ONLINE PLATFORM

IEM Students: FOC
IEM Members: RM15 (Online) / RM20 (Offline)
Non-IEM Members: RM70

ABOUT SPEAKER



Kok-Kwang Phoon is Cheng Tsang Man Chair Professor and Provost, Singapore University of Technology and Design (SUTD). He obtained his BEng and MEng from the National University of Singapore (NUS) and his PhD from Cornell University. Prior to joining SUTD, Prof Phoon was Senior Vice-Provost (Academic Affairs) at NUS. He is a Professional Engineer in Singapore, an ASEAN Chartered Professional Engineer, and past President of the Geotechnical Society of Singapore. He is also serving as the Deputy Chief Scientific Advisor (DCSA) to the National Research Foundation, Prime Minister's Office, Singapore and a member of the executive committee of the Academy of Engineering Singapore.

Prof Phoon is particularly interested in developing statistical and other data-driven methods to support decision making in geotechnical engineering. He has edited 4 books and authored 1 book: Model Uncertainties in Foundation Design (CRC Press, 2021). He was bestowed the ASCE Norman Medal twice in 2005 and 2020, the Humboldt Research Award in 2017, the Harry Poulos Award in 2023, and the Alfredo Ang Award in 2024 among many others. He is the Founding Editor of Georisk and past Chair of the ASCE Geo-Institute Risk Assessment and Management Committee and ISSMGE TC304.

SYNOPSIS

Research in data-centric geotechnics is accelerating as a result of tremendous advances in machine learning and AI. ChatGPT from OpenAI and other generative AIs have moved the divide between what a machine can do and what a human can do in a major way. There is near complete consensus that machine learning and AI have the potential to transform the way we work, live, and play in many fundamental ways. It is prudent for geotechnical engineers to understand and to explore the power and the impact of these new tools, particularly their value propositions to practice. An attempt to explore the future of machine learning in geotechnics was made in 2023 (https://doi.org/10.1080/17499518.2022.2087884), but many research papers have been published since then as an indicator of the rate of progress.

Data is now considered to be an asset that is as valuable as our physical infrastructure. This advantage is not well appreciated by most engineers, although it is pivotal to digital transformation. Machine learning and AI universally depend on data for training and validation. ISSMGE TC304 has initiated a database sharing project in called 304dB (http://140.112.12.21/issmge/tc304.htm) in 2017. The largest collection of geotechnical databases would be published in 2024: www.routledge.com/9781032578958 (Site Characterization) and www.routledge.com/9781032579108 (Geotechnical Structures).

Research in data-centric geotechnics is very active. Two companion special issues in Georisk have been published in issue 17, volume 1 (2023) (https://doi.org/10.1080/17499518.2023.2185938) and issue 18, volume 1 (2024) (https://doi.org/10.1080/17499518.2024.2318849). The first workshop on the Future of Machine Learning in Geotechnics (FOMLIG) was successfully held on 5–6 Dec 2023, Okayama, Japan (https://doi.org/10.1080/17499518.2024.2316882). Although research in data-centric geotechnics is in its infancy, its value to practice has been demonstrated in the current body of work. The value of machine learning to practice was the focus of the Fourth Machine Learning in Geotechnics Dialogue (4MLIGD) (https://doi.org/10.1080/17499518.2024.2316879).

Site investigation is the cornerstone of geotechnical engineering. It is not surprising that data-driven site characterization (DDSC) is an important application area in data-centric geotechnics. The challenges in DDSC were first posed in Phoon et al. 2022 (https://doi.org/10.1080/17499518.2021.1896005). They have attracted significant interest from many researchers resulting in the publication of novel solutions. For example, DDSC was the theme for the Third Machine Learning in Geotechnics Dialogue (3MLIGD) (https://doi.org/10.1080/17499518.2022.2105366). The purpose of this lecture is to present the latest research findings in DDSC under the three elements underpinning data-centric geotechnics:

(1) Data centricity. Databases are also called Big Indirect Data (BID) because they are not directly applicable to any specific site. There are at least seven data attributes common to BIDs that are frequently presented as a mnemonic in the literature, MUSIC-3X (Multivariate, Uncertain and Unique, Sparse, Incomplete, and potentially Corrupted with “3X” denoting 3D spatial variability). Data-centricity requires MUSIC-3X to be addressed directly rather than simplifying it to fit the ideal assumptions in classical statistics.

(2) Fit for (and transform) practice. It is not sufficient to demonstrate that a proposed data-driven method is effective in principle based on a “toy” problem. True scale 3D benchmark examples are presented to highlight DDSC challenges. One DDSC method has been implemented in practice in the software Settle3 (Rocscience) (https://www.rocscience.com/help/settle3/documentation/cpt-analysis/cpt-simulation/simulated-cpt-points)

(3) Geotechnical context. Site recognition is related to a fundamental feature in geotechnical practice, namely all sites are different to some extent (site specificity). Hence, a database containing data from multiple sites is BID. The challenge is to quantify “site uniqueness”, directly or indirectly, so that sparse site-specific data can be supplemented by BID to produce a quasi-site-specific model that is more unbiased than a generic model and less uncertain than a site-specific model.