KAJIMA-FOMLIG AI CHALLENGE

(August 26-28 2026, 4FOMLIG, Seoul)

Purpose

The aim of this challenge is to foster interdisciplinary collaboration, innovation, and practical problem-solving skills among students through a project-based learning framework. By bringing together students from different countries and engaging with Kajima Corporation Technical Research Institute Singapore (KaTRIS) as an industry partner, the challenge seeks to bridge academic knowledge and real-world applications.

Challenges

Two types of challenges are proposed at 4FOMLIG:

Challenge 1: Automatic extraction of geotechnical foundation data (graduate/PhD level), See Pages 2 ~ 4 for details

Challenge 2: Developing a method for identifying expansive soil (undergraduate level), See Pages 5 ~ 6 for details

Awards

Challenge 1: The winner will receive a cash prize of approximately USD 2,500.

Challenge 2: The winner will receive a cash prize of approximately USD 1,900.

*Some teams will also receive approximately USD 1,000 as an Encouragement Award. In some cases, no awards may be given.

*The winner is required to attend the award ceremony and give a presentation in person at 4FOMLIG.

How to participate?

If you are keen to participate in KAJIMA-FOMLIG AI challenge, please contact the following organizers:

Challenge 1 → Stephen Wu (Institute of Statistical Mathematics), stewu@ism.ac.jp

Challenge 2 → Takayuki Shuku (Tokyo City University), tshuku@tcu.ac.jp

*A single team may participate in two challenges.

Tentative Timeline

Call Opens	26th May 2026
Registration Deadline	30th June 2026
Submission Deadline	27th July 2026
Winner Announcement	Before 3rd August 2026
Workshop Presentation	26th ~ 28th August 2026

Organizers

Takayuki Shuku, Tokyo City University, Japan

Stephen Wu, Institute of Statistical Mathematics, Japan

Kensuke Date, Kajima Technological Research Institute Singapore (KaTRIS), Singapore

Yasuhiro Yokota, Kajima Technological Research Institute Singapore (KaTRIS), Singapore

A Benchmark for Automatic Extraction of Structured Data from Pile Test Papers

1. Benchmark Objective and Scope

Pile load-test papers contain a large amount of reusable foundation engineering data, including pile geometry, construction information, soil stratigraphy, in-situ test records, and load-movement responses. These data are essential for building pile databases, calibrating design methods, comparing interpretation criteria, and developing data-driven geotechnical models. However, most of these data remain embedded in heterogeneous literature sources rather than being provided in a directly machine-readable database format. As a result, constructing a high-quality pile test database still relies heavily on manual review, engineering judgement, and repeated cross-checking across text, tables, figures, captions, and appendices.

This benchmark is designed to evaluate automatic data extraction from pile test papers. The task is not limited to locating isolated values in PDF documents. Instead, the expected output is a set of structured, database-ready records that can be mapped to the target fields used in the DUT/CFA/836 pile database. These records cover pile-related metadata, soil profile information, CPT/SPT data, and pile load-movement curves. Therefore, the benchmark focuses on whether an extraction system can recover both the numerical values and the engineering context required to organize them correctly.

A central motivation of this benchmark is that pile test papers often contain target information in forms that are difficult for generic document extraction systems. Relevant evidence may be distributed across different parts of a paper, the correspondence between a curve and a pile record may be implicit, key data may appear only in figures, and the same engineering quantity may be expressed using inconsistent units, abbreviations, or local terminology. To reflect these challenges, the benchmark is constructed around a difficulty-aware design: target data are first defined according to the database requirements, then each paper is annotated using five extraction difficulty categories, and finally the training/test split is determined with reference to the resulting difficulty patterns.

2. Target Data to Be Extracted

The benchmark focuses on extracting structured geotechnical records from pile test papers. The target data are not defined as isolated numerical values, but as linked records that describe the tested pile, the ground condition, the available in-situ test evidence, and the measured load-movement response. This design follows the practical structure of pile databases, where a load-movement curve is meaningful only when it can be associated with the corresponding pile geometry, pile type, loading condition, and soil profile.

Four main groups of target data are considered in the benchmark: pile metadata, soil profiles, CPT/SPT data, and static load-test response data. These groups cover the core information required to reconstruct a pile test case from the literature. The design also reflects a balance between geotechnical completeness and extraction feasibility: the target schema should be rich enough to support downstream database construction, but not so broad that the benchmark becomes dominated by rarely reported or highly specialized information.

Table 1. Target data groups in the pile-test data extraction benchmark

Target data group	Representative fields to be extracted	Required record linkage
Pile metadata	Pile ID, pile type, installation method, pile material, pile diameter, pile length, toe depth, and shaft length	Each pile metadata record should be linked to the corresponding test pile and paper/source case.
Soil profiles	Soil layer depth, layer thickness, soil description, soil classification, groundwater level, and basic soil properties when available	Soil profile records should be linked to the corresponding site, borehole, or pile location.
CPT/SPT data	CPT depth, cone resistance, sleeve friction or friction ratio when available, SPT depth, SPT N-value, and test location	CPT/SPT records should be linked to the relevant site investigation record and, where possible, to the corresponding pile or site area.
Static load-movement data	Loading direction, loading procedure, test type and complete load-movement curve	Each load-movement record should be linked to the corresponding test pile and static load test.

This target schema defines the expected output of the benchmark. A successful extraction system should not only identify values such as pile diameter, SPT N-value, or applied load, but also organize these values into coherent pile test records. For example, a digitized load-movement curve must be assigned to the correct pile, and a soil profile or CPT/SPT record must be associated with the relevant site or pile location. Therefore, the benchmark evaluates both field-level extraction and record-level organization.

3. Difficulty Taxonomy for Manual and Automatic Extraction

The target fields defined above cannot be extracted reliably by treating pile test papers as simple containers of text and tables. In many cases, the required information is distributed across different parts of the paper, linked through implicit engineering context, embedded in figures, or expressed using non-standard terminology, units, and abbreviations. Therefore, this benchmark introduces a difficulty taxonomy to describe the main obstacles encountered during manual review and to provide a structured basis for evaluating automatic extraction systems.

Five difficulty categories are defined in the benchmark. These categories are not intended to measure the overall quality of a paper. Instead, they describe the specific extraction challenges that affect whether target data can be located, interpreted, normalized, and linked to the correct database record. A single paper may contain more than one difficulty category, because different target fields may be presented in different forms within the same source document.

This taxonomy converts qualitative manual-review observations into a structured representation of extraction difficulty. For each paper, the presence or absence of the five categories can be encoded as a binary difficulty vector. This representation provides a bridge between data extraction analysis and benchmark construction: it allows papers to be compared according to their difficulty patterns rather than only by publication source, pile type, or geographic origin. The resulting paper-level difficulty vectors are then used in the subsequent clustering and training/test split design.

Table 2. Difficulty taxonomy for pile-test data extraction

Difficulty	Difficulty category	Definition
D1	Distributed and incompletely signposted target information	Target data are scattered across sections, tables, figures, captions, or appendices, and the paper does not clearly indicate where all required fields can be found.
D2	Implicit correspondence between evidence and target records	The evidence is present, but its correspondence to a specific pile, soil profile, CPT/SPT record, or load-test curve is not explicitly stated and must be inferred from context.
D3	Figure-based target data that are difficult to parse	Important target data are embedded in figures that are dense, low-contrast, multi-panel, weakly labeled, or otherwise difficult to digitize and interpret reliably.
D4	Non-standard expression of the same target quantity	The same target quantity is reported using inconsistent units, ranges, summary values, nominal descriptions, or criterion-dependent expressions, requiring normalization before database entry.
D5	Terminology, abbreviation, and multilingual normalization difficulty	Target fields are expressed using domain-specific shorthand, inconsistent terminology, mixed notation, abbreviations, or non-English language, requiring semantic normalization.

4. Performance evaluation

A total of 31 manually processed papers are included in this benchmark, where 22 papers were selected as “training set” to be used by the participants to develop and validate their data extraction algorithms and the remaining 9 papers are hidden to the participants for final performance test purpose.

The designed split provides a balance between representativeness, difficulty coverage, and training adequacy. It covers the main types of difficulties and avoids overfitting the evaluation design to a single type of difficulty. As a result, the test set is better suited for assessing whether an extraction system can generalize across different difficulty patterns in pile test papers.

Developing A Practical Method For Identifying Expansive Soil

OUTLINE

Expansive soils can cause a wide range of engineering problems, and KAJIMA Technical Research Institute Singapore (KaTRIS) has faced such issues for many years in construction projects across Southeast Asia. To minimize the risks associated with expansive soils, KaTRIS has worked on developing a method for identifying them. However, proposing a universal identification method is challenging because soil behavior is highly variable, and existing testing methods and evaluation criteria differ significantly among countries and organizations. To address this issue, KaTRIS conducted extensive literature reviews and numerous laboratory tests to develop a practical identification method. This research required substantial time and effort from experienced engineers and researchers. This challenge raises the following question:

> *"Can undergraduate students with no prior knowledge of expansive soils develop a practical identification method within a short period of time using advanced AI tools?"*

The objective of this challenge is to develop a practical method for identifying expansive soils. Although the challenge is primarily intended for undergraduate students, master's and PhD students are also welcome to participate.

Key Rules

- No datasets, research papers, or technical reports specific to this challenge will be provided by the organizers. Participants are encouraged to use publicly available information and datasets.
- The submitted explanatory materials (slides, videos, or extended abstracts) must include:
  - i) an explanation of how the method was developed, and which AI tools were used;
  - ii) the total time spent developing the method; and
  - iii) the results of model validation. Any reasonable validation approach may be used.
- The submitted materials should be understandable to judges who were not involved in the development of the method.
- Participants may also develop an agentic AI tool that determines whether the soil at a given location is expansive soil based on the input latitude and longitude.
- It is desirable to make effective use of figures, tables, photographs, maps, and other relevant information as well.
- The proposed method should emphasize practicality, simplicity, and interpretability for real-world engineering applications.
- Participants are encouraged to explain the assumptions and limitations of their proposed method.

Evaluation Criteria

The submitted methods will be evaluated by the evaluation panel based on their practicality, validation quality, effective use of AI tools, development efficiency, novelty of the method, and the clarity of the submitted materials. The evaluation panel may also compare the proposed methods with the method developed by KaTRIS.

Award

The winner will receive a cash prize of approximately USD 1,900. Some teams will also receive approximately USD 1,000 as an Encouragement Award. In some cases, no awards may be given to any teams. The winner is required to attend the award ceremony and give a presentation (5 min ~ 10 min) in person at 4FOMLIG.

References

Participants are encouraged to review introductory materials and previous studies on expansive soils, including their engineering behavior, identification methods, and evaluation criteria. Since there is no universally accepted method for identifying expansive soils, participants are expected to investigate and compare different approaches used in previous studies and engineering practice. Examples of useful references include:

a) Dakshanamurthy, V., and Raman, V. (1973). “A Simple Method of Identifying an Expansive Soil.” Soils and Foundations, 13(1), 97–104.

b) Snethen, D. R. et al. (1977). An Evaluation of Expedient Methodology for Identification of Potentially Expansive Soils. Federal Highway Administration (FHWA).

Timeline

Call Opens	27th May 2026
Registration Deadline	30th June 2026
Submission Deadline	23rd July 2026
Evaluation Period	24th July 2026 ~ 31st July 2026
Winner Announcement	1st August 2026
Workshop Presentation	26th ~ 28th August 2026

*Participants need to submit your explanatory materials to tshuku@tcu.ac.jp by the submission deadline.

Organizers

Takayuki Shuku, Tokyo City University, Japan

Tomonori Mikami, Kajima Technological Research Institute (KaTRI), Japan

Kensuke Date, Kajima Technological Research Institute Singapore (KaTRIS), Singapore

Yasuhiro Yokota, Kajima Technological Research Institute Singapore (KaTRIS), Singapore

If you have any questions, please contact Takayuki Shuku (tshuku@tcu.ac.jp).