

Tyler Bikaun presents: MaintIE: A Fine-Grained Annotation Schema and Benchmark for Information Extraction from Maintenance Short Texts

Maintenance short texts (MST) are value resources for asset management as they offer insights into the condition and maintenance of machines and infrastructure. However, extracting and using this information efficiently on a large scale can be challenging.

In this presentation, [Tyler](#) will introduce "MaintIE," a novel multi-level annotation scheme designed for entity recognition and relation extraction within MST. This scheme categorises information into five primary classes - PhysicalObject, State, Process, Activity, and Property - and further breaks them into 224 specific entities. Additionally, it includes six relations tailored to the context of MST.

Tyler will discuss the creation of two corpora using MaintIE. The first is a meticulously annotated corpus of 1,076 texts characterised by its high quality and fine-grained detail. The second is a larger, coarse-grained corpus comprising 7,000 texts, which has been instrumental in enhancing the capabilities of fine-grained information extraction.

The presentation will explore the performance of various deep-learning models using these corpora, setting benchmarks for automated entity recognition and relation extraction in MST. Tyler will demonstrate how the MaintIE scheme, corpus, and models can be adopted by industry through their public release made available under the MIT license to foster further research and innovation in this domain.

When - 1.00pm - 15 December 2023

Where - Microsoft Teams meeting

Join on your computer, mobile app or room device

[Click here to join the meeting](#)