A fresh look at the Future of Technical Language Processing Research in the eyes of ChatGPT



A/Prof Wei Liu

Theme Lead

Theme 1
2023-04-21

Virtual - Researchers Catch-up host online from Curtin University

"Unprecedented" still feels like an under-statement to describe the "astronomical" growth of the user-base of ChatGPT, which attracted over 1 million users in less than a week and 100 million users in less than two months. There are claims that the only other application that enjoy the same "vertical" adoption is Pokémon GO.

For AI researchers, in particular language researchers, this is so disruptive that research methodologies (if not research directions) must be re-defined and re-designed. It sets new benchmarks or at least demands for comparisons/integrations across almost all language-based tasks. Technical language processing (TLP) is not spared.

- What are the opportunities and challenges lie ahead for us? Automatic triple extraction without/with minimal labelling effort?
- Will "prompt engineering" - how to ask LLM complex queries become a research direction?
- ChatGPT and Large Language Models (LLM) are capturing/memorising both common-sense and domain-specific knowledge by "experiential" learning, very much like how our cognitive system develops. Do we still need to build explicit knowledge repositories using ontologies? Would a "know-everything" LLM be adaptive and conversant enough to provide the translational role of ontologies for the interoperability of heterogenous systems?
- Do we even need to build domain-specific technical language models in house, or one day that LLM vendors (e.g. OpenAI) will have good-enough data governance to allow companies to trust them with proprietary data?

We have more questions than answers. Let's take this opportunity to start a "chat" about ChatGPT and LLMs, with a focus on information extraction and knowledge graph construction from maintenance work orders.