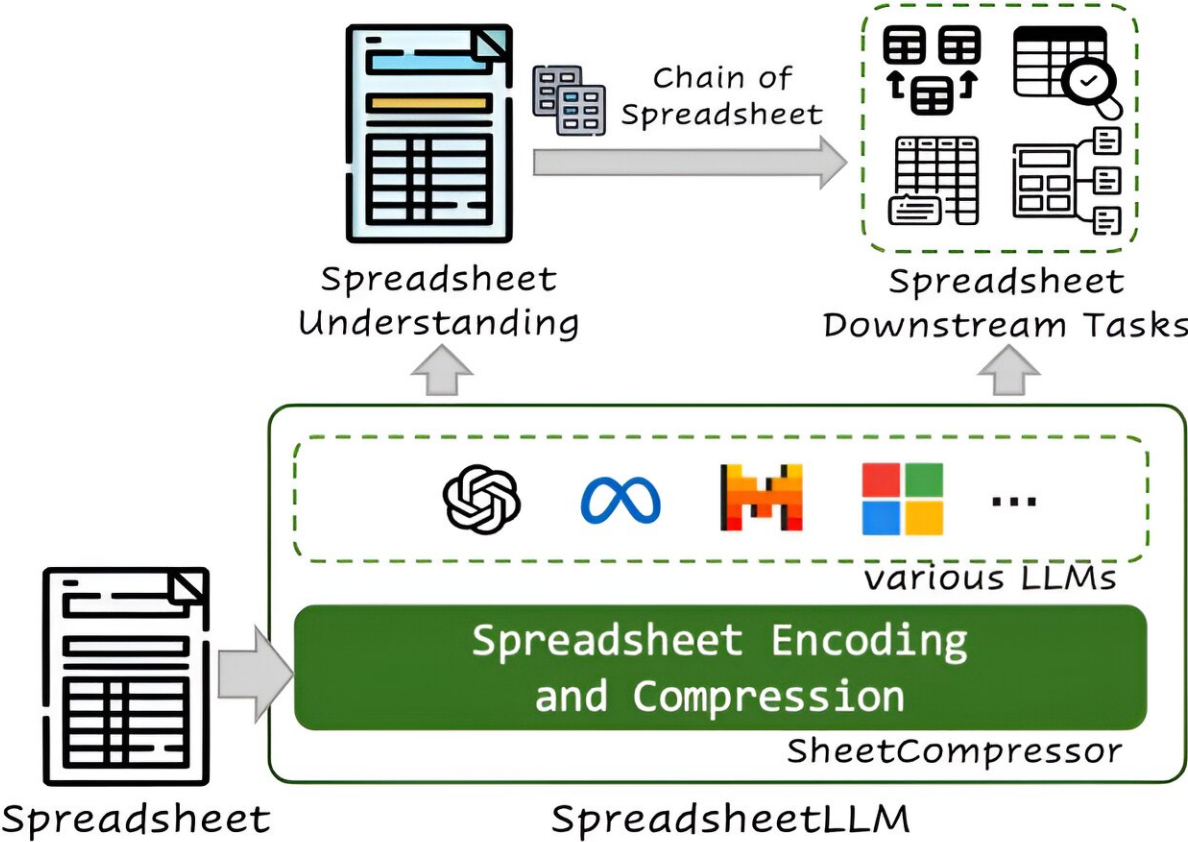# Microsoft unveils software that allows LLMs to work with spreadsheets

July 16 2024, by Bob Yirka

The SPREADSHEETLLM pipeline. Credit: *arXiv* (2024). DOI: 10.48550/arxiv.2407.09025

A team of programmers and AI specialists at Microsoft has developed an

AI tool called SpreadsheetLLM that applies large language model capabilities to spreadsheets. In their study, now [posted](#) on the *arXiv* preprint server, the group developed SheetCompressor, an encoding framework that compresses spreadsheets effectively for use by large language models (LLMs).

LLMs such as ChatGPT are well known, but as more people use them, more of their ability gaps become apparent. One gap is the ability of such models to make sense of spreadsheets. Because of their unique arrangement and capabilities, spreadsheets remain a mystery to LLMs, preventing their use as a [tool](#) for business.

In this new study, the team at Microsoft created a tool that reorganizes a [spreadsheet](#) into a form that LLMs can use as a data source. And, as the team notes, it is based on a concept they call SheetCompressor, a programming tool that allows AI data management and analysis for information in spreadsheets.

To implement SheetCompressor, the researchers split it into three main functions: compression, translation and data format aggregation. The first was implemented by adding what the team describes as anchors throughout a spreadsheet to help an LLM understand what the spreadsheet does.

Once in place, rows and columns are replaced with a skeletonized table. Translation modules are then used to remove empty cells or repeating values. Applying a lossless inverted index [translation](#) in JSON format allows for data [format](#) aggregation.

The team also added other modules to handle unique situations, such as adjacent cells with similar numerical formats. The result is a tool that allows LLMs to use spreadsheets as a data source in a variety of ways.

The research team suggests that SpreadsheetLLM opens the door to using LLM technology to revolutionize the way that spreadsheets are used; from automating data entry, to data analysis, to presentation of complex information in a way that is accessible to people with a variety of backgrounds. And that, they further suggest, will make spreadsheets and the data they hold much more accessible and useful.