

Artificial intelligence (AI) technology company Upstage (CEO Kim Sung-hoon) announced on the 9th that it has signed a memorandum of understanding (MOU) with AI language data company Flitto (CEO Lee Jung-soo) to build AI language data.

Currently, the development of giant language models (LLMs) led by big tech companies is biased towards English, and relatively, languages in Asia have insufficient learning data. In response, the two companies will work together to build multilingual language data, such as Japanese and Thai, to lay the foundation for developing high-quality LLMs even for low-resource languages.

Specifically, the two companies will strengthen cooperation in various fields, including ? building a benchmark dataset for the Korean language model evaluation platform 'Open-Ko LLM Leaderboard' ? operating a multilingual LLM leaderboard ? building low-resource language data and localizing LLMs using low-resource languages ? strengthening data supply partnerships for enterprise LLM construction.

Upstage plans to secure high-quality low-resource language data through this and further expand the supported languages of its pre-trained LLM Solara to develop customized models specialized in various regional languages, such as Southeast Asia. Solara currently supports Korean and English, and plans to expand its supported languages to include Japanese and Thai by the end of the year.

Flitto plans to contribute actively to the advancement of language models through synergy between language data and AI technology based on this agreement, and to enhance its language collection technology competitiveness through multilingual parallel corpus construction know-how and text, image, and voice datasets without copyright issues.

Upstage CEO Kim Sung-hoon said, "Securing high-quality language data is an essential task in the context of the generative AI craze triggered by language models." He added, "Upstage will work on data

enhancement to ensure that more people around the world can experience generative AI innovation through this cooperation with Flitto."

Flitto CEO Lee Jung-soo said, "The field of low-resource language learning has emerged as a key factor in the performance of giant language models." He added, "Through the cooperation between the two companies, we will show how much positive contribution the combination of high-quality data and advanced technology can make to the domestic generative AI ecosystem."