Facebook owner Meta seeks to train AI model on European data as it faces privacy concerns

June 10 2024, by Kelvin Chan

The Meta logo is seen at the Vivatech show in Paris, France, Wednesday, June 14, 2023. Meta said Monday, June 10, 2024, it wants to use data from users in privacy-conscious Europe to train its artificial intelligence models. It's facing concerns about data protection while battling to keep up with rivals like OpenAI and Google. Credit: AP Photo/Thibault Camus, File

Meta wants to use data from users in privacy-conscious Europe to train its artificial intelligence models, the social media giant said Monday as it faces concerns about data protection while battling to keep up with rivals like OpenAI and Google.

The company, which owns Facebook, Instagram and WhatsApp, said that in order to better reflect the "languages, geography and cultural references" of its users in Europe, it needs to use public data from those users to teach its Llama AI large language model.

Meta's AI training efforts are hampered by stringent European Union data privacy laws, which give people control over how their personal information is used. Vienna-based group NOYB, led by activist Max Schrems, complained last week to 11 national privacy watchdogs about Meta's AI training plans and urged them to stop the company before it starts training Llama's next generation.

AI language models are trained on vast pools of data that help them predict the most plausible next word in a sentence, with newer versions typically smarter and more capable than their predecessors. Meta's AI assistant feature has been [baked into](#) Facebook, Instagram and WhatsApp for users in the U.S. and 13 other countries, but [notably not Europe](#).

"If we don't train our models on the public content that Europeans share on our services and others, such as public posts or comments, then models and the AI features they power won't accurately understand important regional languages, cultures or trending topics on social media," Stefano Fratta, global engagement director of Meta's privacy policy, said in [blog post](#).

"We believe that Europeans will be ill-served by AI models that are not informed by Europe's rich cultural, social and historical contributions."

Fratta said other companies including Google and OpenAI have already trained on European data. Meta won't use private messages to friends and family nor content from European users who are under 18, he said.

Since May 22, the company has sent 2 billion notifications and emails to European users explaining its plans and linking to an online form to opt out, Fratta said.

The latest version of Meta's privacy policy is set to take effect on June 26, indicating that training for the next model will start soon after.