

L'évaluation à l'ère des IA génératives : quelles sont les stratégies de détection de contenu généré par l'IA ? Sont-elles la bonne cible ?

Avec l'essor des IA génératives, notamment des modèles comme GPT et d'autres architectures similaires, la question sur la détection de contenu généré par ces IA se pose à différents niveaux. Bien que la mise en place des stratégies pour détecter et prévenir les abus potentiels de ces technologies, en particulier la production de contenu erroné, trompeur, ou nuisible, soit essentielle, la détection de la « tricherie » ne semble pas être la bonne cible. Il est plutôt essentiel de sensibiliser les utilisateurs aux risques associés au contenu généré par des IA et de les éduquer sur l'identification de contenu potentiellement problématique. Nous allons discuter sur ces points, en mettant l'accent sur l'utilisation responsable de ces technologies.