CybersÃ©curitÃ© : Encadrer lâ□□utilisation de lâ□□IA gÃ©nÃ©rative
SÃ©curitÃ©
PostÃ© par : JerryG
PubliÃ©e le : 6/11/2023 13:00:00

LancÃ© par la PremiÃ¨re ministre en septembre dernier, le comitÃ© de lâ□□intelligence artificielle gÃ©nÃ©rative a pour objectif de prÃ©senter des propositions concrÃ¨tes d'ici Ã six mois afin d'adapter la stratÃ©gie du Gouvernement au sujet de l'IA, en plein essor.

Selon la derniÃ¨re enquÃªte annuelle de McKinsey, un tiers des entreprises utilisent rÃ©guliÃ¨rement lâ□□IA gÃ©nÃ©rative dans au moins une fonction, et 60 % des organisations ayant dÃ©clarÃ© avoir adoptÃ© lâ□□IA utilisent lâ□□IA gÃ©nÃ©rative.

Si les avantages de son utilisation sont considÃ©rables, elle nâ□□est pas sans risque. En juin dernier, plus de 100 000 comptes ChatGPT ont Ã©tÃ© piratÃ©s, puis mis en vente sur le Dark Web. Il faut donc avoir conscience que ces outils ne garantissent pas la sÃ©curitÃ© ou la confidentialitÃ© des donnÃ©es.

Selon Ray Canzanese, Threat Research Director du Threat Labs de Netskope, si les applications d'IA gÃ©nÃ©rative ont le potentiel de faciliter le quotidien des collaborateurs, elles peuvent Ã©galement exposer de maniÃ¨re significative les donnÃ©es sensibles et les organisations Ã des risques de compromissions de donnÃ©es et de non-conformitÃ© :

Â« Selon nos recherches, pour 10â□□000 utilisateurs, les entreprises subissent chaque mois environ 183 incidents de partage de donnÃ©es sensibles avec lâ□□application ChatGPT. Les codes source sont les plus partagÃ©es dans ChatGPT, avec un taux mensuel de 158 incidents pour 10â□□000 utilisateurs.

On dÃ©nombre Ã©galement des donnÃ©es rÃ©glementÃ©es, notamment financiÃ¨res ou mÃ©dicales, et des informations personnellement identifiables (IPI). Les organisations doivent donc prendre des contre-mesures sÃ©rieuses pour prÃ©server la confidentialitÃ© et la sÃ©curitÃ© de leurs donnÃ©es sensibles.

Cela implique de surveiller l'utilisation et l'abus potentiel d'applications dâ□□Intelligence Artificielle gÃ©nÃ©rative, en limitant l'exposition d'informations sensibles par leur intermÃ©diaire et en sÃ©curisant les documents confidentiels contre la perte accidentelle et le vol.

Pour relever le dÃ©fi, la visibilitÃ© sera clÃ©. Les Ã©quipes de sÃ©curitÃ© doivent en effet s'appuyer sur des outils automatisÃ©s capables de surveiller en permanence les applications telles que ChatGPT auxquelles les utilisateurs de l'entreprise accÃ¨dent ; et plus particuliÃ¨rement comment, quand, Ã partir d'oÃ¹ et Ã quelle frÃ©quence.

Il est essentiel de comprendre Ã©galement les diffÃ©rents niveaux de risque que chaque outil reprÃ©sente pour l'organisation et d'avoir la capacitÃ© de dÃ©finir des politiques de contrÃ´le d'accÃ¨s granulaires, en temps rÃ©el sur la base de cette catÃ©gorisation ainsi que des conditions de sÃ©curitÃ© qui peuvent changer au fil du temps.

Alors que les applications plus explicitement malveillantes devraient Ãªtre bloquÃ©es, la responsabilitÃ© de l'utilisation dâ□□applications dâ□□IA gÃ©nÃ©rative devrait Ãªtre confiÃ©e aux utilisateurs. Il nâ□□y aura en effet pas de futur sans ces technologies, lâ□□objectif est donc de permettre, mais de restreindre leur utilisation Ã des activitÃ©s dÃ©finies et pertinentes.

En parallÃ¨le, les Ã©quipes de sÃ©curitÃ© ont la responsabilitÃ© de sensibiliser les collaborateurs aux applications et aux activitÃ©s jugÃ©es risquÃ©es. Cela peut se faire via des alertes en temps rÃ©el et des workloads automatisÃ©s de formation, en impliquant l'utilisateur dans les dÃ©cisions d'accÃ¨s aprÃ¨s avoir reconnu le risque.

Les utilisateurs commettent des erreurs et peuvent, par nÃ©gligence, mettre en danger des donnÃ©es sensibles. Par consÃ©quent, il est primordial de limiter le tÃ©lÃ©chargement et la publication de donnÃ©es hautement sensibles via ChatGPT et autres outils similaires.

Seules des techniques modernes de prÃ©vention des pertes de donnÃ©es (DLP) permettent d'y parvenir, car elles sont capables d'identifier automatiquement les flux de donnÃ©es critiques et de catÃ©goriser les messages sensibles avec un niveau de prÃ©cision trÃ¨s Ã©levÃ©. Â»