Date: 18.08.2025

Fr-app

Online-Ausgabe

frapp.ch/fr
1752 Villars-sur-Glâne
https://frapp.ch/fr

Genre de média: En ligne
Type de média: Plateformes d'informations
Page Visits: 1'012'700

Vue Web

Hes·so

Ordre: 1073023
N° de thème: 375009

Référence: 93efab2f-9480-439d-92e4-dd03a7b1d87d
Coupure Page: 1/3

Cybersécurité

Agents IA: une automatisation à double tranchant

18.08.2025, Frapp - Mattia Pillonel

Les agents d'intelligence artificielle, capables d'automatiser certaines tâches en ligne, facilitent aussi la désinformation et l'escroquerie. Explications.

On les appelle des "agents". Ce sont des programmes d'intelligence artificielle capables d'agir de manière autonome. Là où les IA conversationnelles se contentent de répondre à nos questions, les agents passent à l'action, en allant par exemple chercher des informations sur nos réseaux sociaux ou dans nos mails; ils peuvent aussi réserver des vacances, commander des habits pour les utilisateurs, ou même créer des sites web en seulement quelques minutes.

Depuis quelque temps, ils ont de plus en plus la cote. OpenAI, par exemple, a lancé ChatGPT Agent en juillet, et certains utilisateurs se sont rapidement amusés à envoyer l'IA faire des parties d'échecs en ligne contre des joueurs humains, pour tester les limites de la technologie.

Mais derrière les promesses d'automatisation et de facilitation du quotidien, se cachent aussi des risques en matière de cybersécurité et de désinformation. "Avec ces agents, on va contextualiser l'IA à fonctionner d'une certaine façon", résume JeanHennebert, professeur en informatique à la Haute école d'ingénieurs et d'architecture de Fribourg et responsable de l'Institut d'IA et Systèmes Complexes. Pour lui, l'ingénierie sociale, pratique de manipulation psychologique à des fins d'escroquerie, est l'un des risques majeurs. En ayant accès aux réseaux sociaux, l'IA peut "comprendre qui vous êtes, comment vous réagissez et même, s'il y a des contenus vocaux, imiter votre voix", avertit le professeur.

Les agents ont la capacité d'automatiser la reconnaissance de cibles, optimiser les campagnes de phishing et personnaliser les attaques. Si la police cantonale n'a pas de chiffres précis sur les cyberescroqueries boostées à l'IA, elle confirme que celle-ci est utilisée pour l'automatisation d'échanges de discussion avec les personnes lésées, la modification d'images et de vidéos — par exemple faire croire qu'une célébrité fait la promotion d'une plateforme crypto — ou encore la réalisation de sites web et de fausses annonces. "L'IA permet avant tout aux escrocs de gagner du temps, de crédibiliser les scénarios et de défier les instruments de sécurité mis en place par les banques, les plateformes ou les marketplaces", nous explique-t-elle.

Désinformation en quelques clics

"L'IA est effectivement un amplificateur de plein de choses qui se passaient déjà auparavant", observe Jean-Hennebert. Mais au-delà des risques d'attaques spécifiques sur les utilisateurs, le spécialiste identifie une autre menace insidieuse: la création et la diffusion massive de désinformation.

La facilité déconcertante avec laquelle il est désormais possible de créer des sites web en est une bonne illustration. "J'ai fait l'expérience en utilisant Manus, une IA chinoise un peu plus ouverte sur ce qu'on peut lui demander", explique Jean Hennebert. "Je lui ai demandé de raisonner sur les arguments climato-sceptiques les plus convaincants, confrontés aux arguments scientifiques. Elle produit une réponse élaborée qui est d'ailleurs très bien faite. Je lui demande ensuite de me créer un site web sur cette base. Quelques dizaines de secondes plus tard, on se retrouve effectivement avec une page, qui expose cette argumentation climato-sceptique de façon convaincante."

Et la cerise sur le gâteau? Le site peut être indexé par les moteurs de recherche comme Google, amplifiant sa visibilité. "On peut ainsi imaginer créer des centaines de pages qui diffusent de l'information critique", alerte Jean Hennebert. "Et ce qui serait encore plus inquiétant, c'est que ces mêmes sites entrent ensuite dans les données d'entraînement des IA."

Garder un esprit critique

Date: 18.08.2025

Fr-app

Online-Ausgabe

frapp.ch/fr
1752 Villars-sur-Glâne
https://frapp.ch/fr

Genre de média: En ligne
Type de média: Plateformes d'informations
Page Visits: 1'012'700

Vue Web

Hes·so

Ordre: 1073023
N° de thème: 375009

Référence: 93efab2f-9480-439d-92e4-dd03a7b1d87d
Coupure Page: 2/3

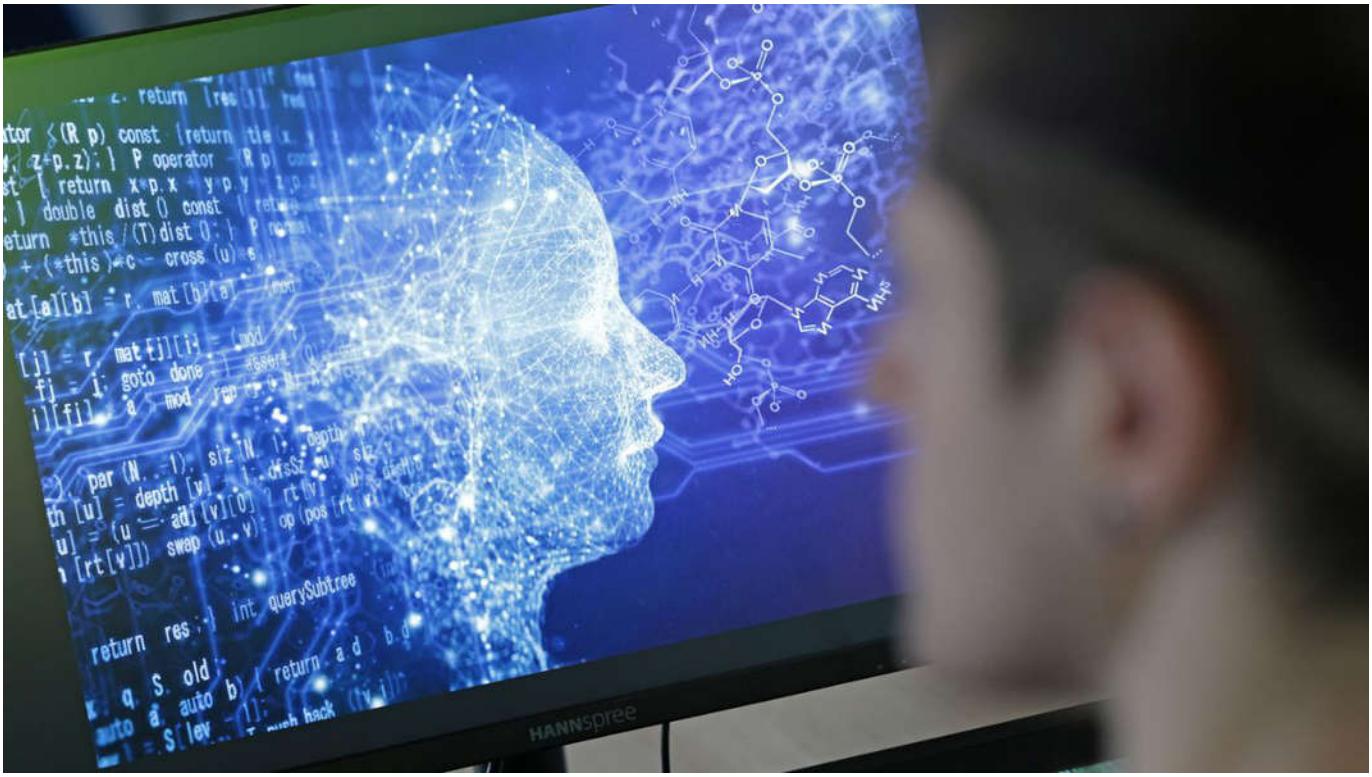
Alors comment se protéger contre cette montée des menaces automatisées? Un retour aux fondamentaux de la sécurité numérique s'impose pour Luca Haab, professeur à l'HEIA-FR spécialisé dans la cybersécurité.

Face aux attaques de phishing personnalisées, il faut porter une attention particulière aux détails comportementaux. "Si vous savez qu'un ami vous parle toujours de son chien quand vous faites du small talk par exemple, mais que lors d'une discussion en ligne qui semble venir de lui, il n'y a aucune mention de chien, il faut se méfier", explique l'expert. En cas de doute, il ne faut surtout pas hésiter à prendre un autre canal, "que ce soit par téléphone, un email ou autre, et poser la question si c'est bien avec cette personne qu'on parle."

Luca Haab explique que les évolutions technologiques récentes ont déjà forcé des adaptations des systèmes de sécurité. "Certaines banques avaient mis en place une reconnaissance vocale pour permettre aux clients s'authentifier avant de pouvoir discuter avec leur propre banquier, mais ces systèmes ne sont tout simplement plus fonctionnels."

Mais face à la désinformation, là, c'est un peu plus compliqué. "Quand on cherche une information, la première chose qu'on va faire, c'est aller sur Google et regarder les dix premiers résultats. Dans l'exemple du professeur Hennebert, cela va devenir de plus en plus compliqué parce que l'information correcte sera noyée par des fake news. Actuellement, nous n'avons pas de réponse technique pour trier facilement cela", regrette Luca Haab. La responsabilité retombe finalement sur l'utilisateur, qui doit toujours garder un esprit critique face aux informations trouvées sur internet et valider ses recherches avec des sources de confiance.

Un esprit critique qui est de mise aussi dans l'utilisation des outils IA, rappellent finalement les deux professeurs. "Tout citoyen et citoyenne doit utiliser sciemment ces outils, se poser des questions et ne pas juste demander à ChatGPT ou autre IA telle que le Chat de Mistral, mais plutôt réfléchir par soi-même", insiste Luca Haab.



L'IA permet aux escrocs de gagner du temps, de crédibiliser les scénarios et de défier les instruments de sécurité mis en place (image prétexte). © KEYSTONE