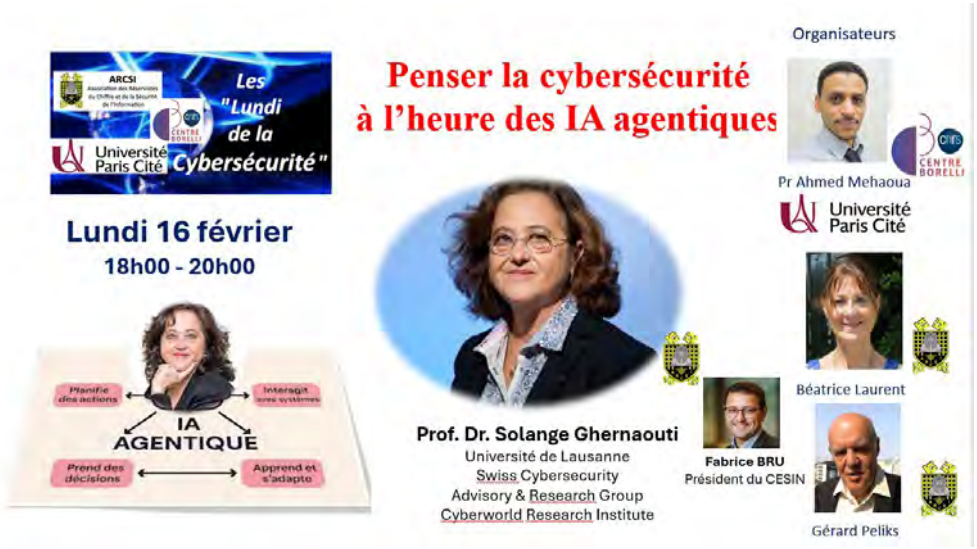
Lettre des "Lundi de la cybersécurité" n° 89

Penser la cybersécurité à l'heure des IA agentiques



L'IA Agentique, une nouvelle révolution qui va bouleverser l'utilisation de l'Internet

Le web a donné le premier grand engouement pour l'utilisation de l'Internet. Puis **les moteurs de recherches** ont grandement participé à l'utilisation, par tous, de ce réseau mondial. Quelques mots clés entrés et le web propose de nombreux liens contenant ces mots. Plus récemment, **les moteurs de réponses** ont encore agrandi cet attrait avec l'IA Générative. On pose une question et le web, après consultation de tas de serveurs et analyse statistique des réponses à travers des « réseaux de neurones » artificiels propose un texte généralement très intéressant pour fournir une réponse (voir par exemple à la fin de ce message, ce que répond Gemini à ma question posée.

Une nouvelle révolution aujourd'hui s'annonce : Des agents « intelligents » capables d'initiatives, d'adaptations et de décisions stratégiques dans des environnements complexes. Et ces agents accomplissent l'objectif qui leur est donné et qui peut évoluer avec le temps.) avec une autonomie encadrée et évolutive.

C'est l'**IA** dite **Agentique.** Cette nouvelle application va rendre d'immenses services pour une collaboration augmentée entre l'humain et la machine. Mais, comme toute technologie nouvelle, elle sera aussi utilisée par les cybercriminels, et les cyber-combattants du camp adverse**.** Cela pose bien sûr de nouveaux problèmes. L'humain peut-il faire confiance à des systèmes informatiques autonomes ? Les mécanismes et les actions de ces systèmes sont-ils toujours explicables ? Pour ces évoquer ces points fondamentaux, dans notre prochain

« Lundi de la Cybersécurité », nous avons choisi une experte de l'IA Agentique, très connue pour nous expliquer en quoi l'agenticité change la donne de la cybersécurité.

Je donne la plume à Solange GHERNAOUTI

Penser la cybersécurité à l'heure des IA agentiques

Inscrite dans un cadre de réflexions de nature transdisciplinaire et selon une approche globale, notre réflexion traitera de l'intelligence artificielle sous l'angle de la génération de risques complexes pour nous intéresser à ce que fait l'IA agentique à la cybersécurité en termes d'enjeux, de défis et d'éléments de solutions.

En proposant une grille de lecture des risques et des apports de l'IA agentique à la sécurité des organisations nous questionnerons les pratiques de cybersécurité.

Qui est Solange GHERNAOUTI ?

Docteure en informatique de l'université Paris Sorbonne, Solange Ghernaouti est spécialiste de renommée internationale des questions de sécurité liées aux technosciences. Membre de l'Académie suisse des sciences, Chevalier de la Légion d'honneur, présidente de la Fondation SGH – Institut de recherche Cybermonde, elle est professeure honoraire de l'université de Lausanne.

Auteure de nombreux livres scientifiques et de vulgarisation dont certains sont traduits en plusieurs langues. Dunod vient de publier la 8ème édition revue et augmenté de son livre de cours avec exercices corrigés « Cybersécurité. Analyser les risques. Mettre en œuvre les solution ».

Son roman « Off » co-écrit avec Philippe Monnin *(Slatkine 2023) est lauréat du Prix Littéraire Environnement 2025.*

Puisqu'on ne change pas une équipe qui gagne, notre prochain roman consacré à l'intelligence sortira en mars 2026 toujours aux éditions Slatkine.

Demande d'inscription au lundi 16 février, par visio-conférence à partir de 18 h 00.

Nos « Lundi de la cybersécurité » sont gratuits et veulent vous offrir une fête technologique. Demandez votre inscription, par courriel, nous vous enverrons, un peu avant le jour de l'événement, un hyperlien pour entrer dans la visioconférence Zoom.

Les demandes d'inscriptions sont à adresser à **Béatrice Laurent** co-organisatrice de nos évènements :

beatricelaurent.CDE@gmail.com

Les prénoms, noms et adresses mails des inscrits seront connus des organisateurs et communiqués aux intervenants. Si vous voulez être ajoutés à ma liste de distribution des lettres des « *Lundi de la cybersécurité* » mensuels, demandez-le-moi par mail (gerard.peliks@noos.fr).

Si vous vous inscrivez pour assister à notre évènement, soyez connectés le **lundi 16 février, dès 18 h 00 ou mieux, un peu avant**. La salle d'attente Zoom sera ouverte aux alentours de **17 h 00** pour des conversations informelles entre intervenants et participants. Ce sont toujours d'agréables échanges qui se font avant 18 h 00.

Quelques minutes avec une association : Le CESIN

Suivant la tradition de nos « Lundi de la cybersécurité », entre l'exposé des intervenants et la session questions / réponses, qui commence autour de 19 h 30, nous donnons pour une quinzaine de minutes la parole à une organisation ou à une personne qui opère dans l'écosystème du numérique et dans la sécurité de l'information.

Cette séquence des **quelques minutes avec une association** sera animée par **Fabrice BRU, président du CESIN.**

Le coin des « Lundi de la cybersécurité » précédents.

Vous trouverez, sur le web de l'ARCSI en
https://www.arcsi.fr/evt_passes.php
tous les détails, slides, parfois aussi les enregistrements et des comptes rendus.
La retransmission du « Lundi de la Cybersécurité » précédent est disponible ici

Le coin-coin de l'IA Générative

Allons-nous obtenir une « hallucination » ou une réponse qui décrit une situation bien réelle ?

Prompt posé à Gemini :

L'IA Agentic peut-elle aider à diminuer le risque cyber ?

L'IA agentique (ou IA par agents) est l'une des évolutions les plus prometteuses pour la cybersécurité en 2025. Contrairement à l'IA traditionnelle qui se contente d'analyser ou de suggérer, l'IA agentique est capable de raisonner, de planifier et surtout d'agir de manière autonome pour atteindre un objectif de sécurité.

Voici comment elle transforme concrètement la défense numérique pour diminuer les risques :

1. **De la détection passive à la réponse autonome**

L'IA classique (comme dans les outils EDR ou SIEM) alerte souvent un humain d'une anomalie. L'IA agentique, elle, prend l'initiative :

- Investigation intelligente : L'agent IA peut recevoir une alerte, interroger de lui-même les bases de données de menaces, vérifier l'identité de l'utilisateur sur Okta et corréler les logs réseau sans attendre qu'un analyste le fasse.
- Remédiation immédiate : Si l'attaque est confirmée, l'agent peut décider de couper une connexion réseau, d'isoler une machine infectée ou de réinitialiser un mot de passe en quelques secondes, là où un humain mettrait des minutes (ou des heures).

2. Défense contre les attaques "à la vitesse de la lumière"

Les attaquants utilisent désormais eux-mêmes des agents IA pour automatiser des attaques par phishing massif ou pour trouver des vulnérabilités dans le code.

- Contre-offensive automatisée : Seule une IA agentique peut répondre avec la même rapidité pour bloquer des milliers de tentatives d'intrusion simultanées qui s'adaptent en temps réel.
- Pentesting continu : Des agents peuvent être configurés pour tester en permanence vos propres défenses, en simulant des comportements d'attaquants pour trouver des failles avant qu'elles ne soient exploitées.

3. Gestion de la fatigue des alertes

L'un des plus gros problèmes des centres d'opérations de sécurité (SOC) est le volume écrasant d'alertes.

- Réduction du bruit : Les agents IA peuvent regrouper des dizaines d'alertes liées à un même incident et ne présenter à l'humain qu'un rapport complet plutôt que 50 notifications éparses.
- Focus stratégique : En automatisant les tâches répétitives de niveau 1, elle permet aux experts de se concentrer sur les menaces complexes et la stratégie.

Les nouveaux risques à surveiller

L'introduction d'agents autonomes crée aussi de nouveaux défis :

- L'agent compromis : Si un attaquant prend le contrôle d'un agent de sécurité qui a des droits d'administration, les dégâts peuvent être massifs.
- Hallucinations d'action : Un agent pourrait interpréter un comportement légitime comme une menace et couper l'accès à un serveur critique par erreur.
- Gouvernance : Il devient crucial de définir des "garde-fous" (*guardrails*) pour limiter ce qu'un agent a le droit de faire sans validation humaine.

Nous écouterons **Solange GHERNAOUTI** nous dire ce qu'elle en pense.

JE M'INSCRIS Inscrivez-vous (mail à beatricelaurent.cde@gmail.com) ou cliquez sur le rectangle rouge :

Gérard