Fiche de synthèse veille technologique

Thèmes de veille technologique :

- Cybersécurité
- IA

Définition des thèmes :

Cybersécurité : La cybersécurité regroupe l'ensemble des pratiques, technologies et processus visant à protéger les systèmes informatiques, les réseaux, les programmes et les données contre les cyberattaques. Ces menaces peuvent inclure le vol d'informations, les logiciels malveillants, les ransomwares, les attaques par déni de service (DDoS) et le piratage des infrastructures critiques.

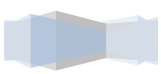
IA : L'intelligence artificielle (IA) désigne l'ensemble des technologies permettant à des machines d'imiter certaines capacités humaines, comme l'apprentissage, la compréhension du langage, la reconnaissance d'images et la prise de décision. L'IA repose sur des algorithmes avancés et des modèles d'apprentissage automatique (machine learning) pour analyser de grandes quantités de données et s'améliorer en continu.

Définition de la stratégie de collecte :

La stratégie de collecte désigne l'ensemble des méthodes, outils et processus mis en place pour recueillir, analyser et exploiter des informations pertinentes dans un domaine donné. Elle est essentielle dans des contextes variés tels que la veille technologique, la veille concurrentielle, la gestion des données ou encore la cybersécurité.

Exploitation des données, dans le respect des bonnes pratiques :

Pour l'exploitation des données dans le respect des bonnes pratiques, j'utilise les outils FlipBoard et google alert.
J'utilise ces deux outils car ce sont des outils graphiques, rapides et simples d'utilisation.
D'autres outils peuvent être utilisés pour la veille technologique comme Feedly, Meltwater ou encore Next.
J'ai créé des répertoires qui rassemblent tous les articles autours de mes thèmes de veille.



La veille technologique sur l'intelligence artificielle et la cybersécurité nécessite l'identification de sources fiables et diversifiées. Les publications académiques et

scientifiques constituent une base essentielle pour comprendre les avancées théoriques et technologiques dans ces deux domaines. Les recherches publiées dans des revues spécialisées ou accessibles via des bases de données scientifiques permettent d'accéder à des analyses rigoureuses sur des sujets tels que l'apprentissage automatique appliqué à la cybersécurité ou encore les nouvelles méthodes de détection des cyberattaques.

Exemples :

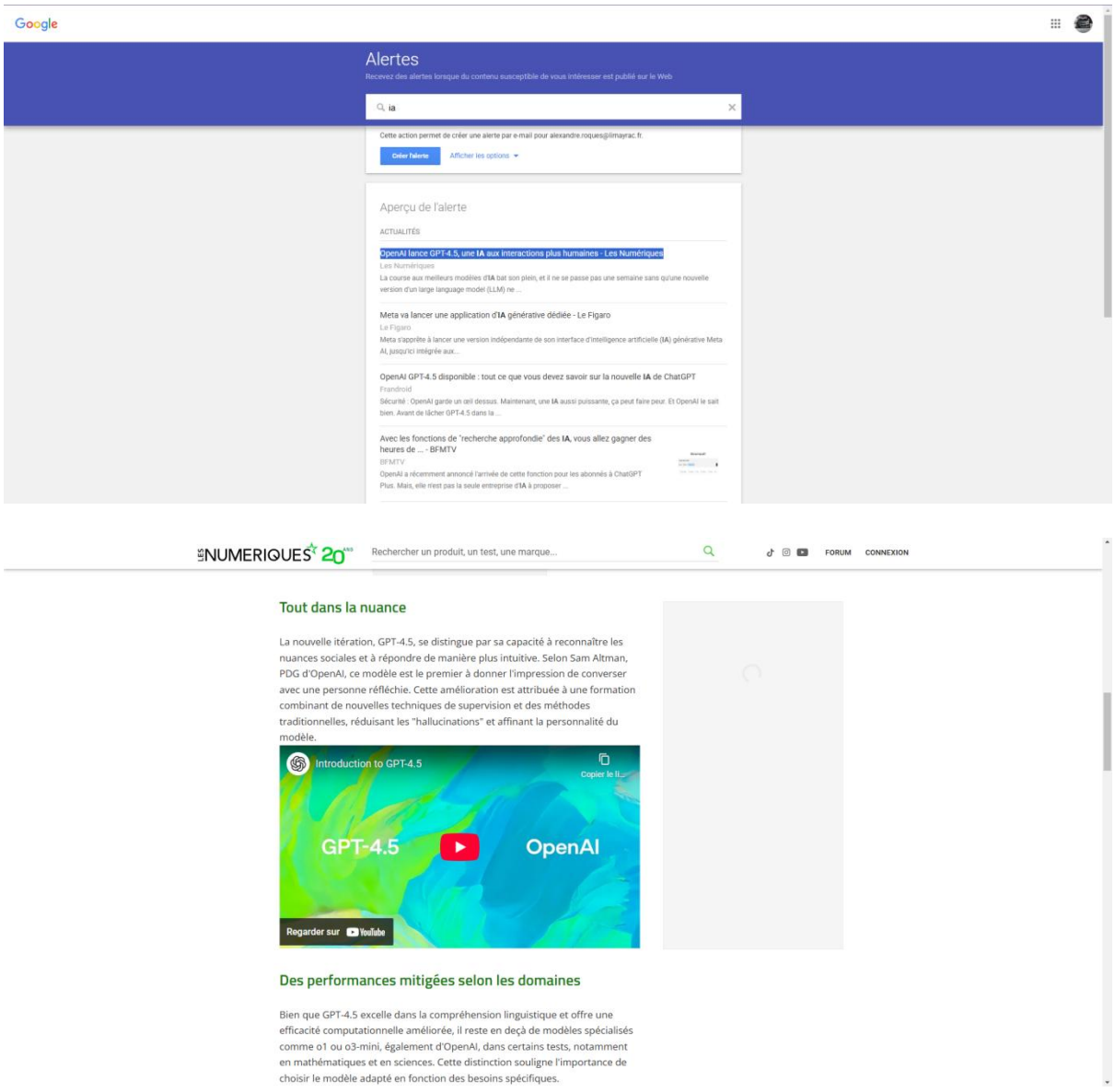
Pour illustrer ces propos on va illustrer ces propos avec deux exemples sur les différents thèmes.

Cybersécurité :

Article récupéré sur l'outil flipboard traitant d'une cyberattaque sur Disney à l'aide de l'intelligence artificielle :

Ce média traite de tout le secteur du numérique, au-delà de l'industrie. Il peut traiter de divers domaines, notamment des aspects matériels (électronique), services (Saas), réglementation (RGPD) ainsi que l'intelligence artificielle (IA).

IA :



Le site propose des actualités sur les objets et loisirs numériques, des essais comparatifs de matériels réalisés par son laboratoire pour guider les choix des lecteurs.

3

Synthèse des Informations

Une fois les données collectées, il est important de les structurer afin d'en extraire les éléments clés et d'en faciliter l'exploitation. La synthèse analytique consiste à résumer les

principales avancées en IA et en cybersécurité, en mettant en lumière les innovations technologiques, les nouvelles menaces identifiées et les évolutions réglementaires.

Ce travail permet de mieux comprendre les interactions entre l'IA et la cybersécurité, notamment en ce qui concerne l'utilisation de l'apprentissage automatique pour la détection proactive des cyberattaques.

Synthèse article de cybersécurité :

Synthèse – Fuite de Données de Disney suite à une Cyberattaque Alimentée par l'IA

Un pirate informatique a mené une cyberattaque contre un ingénieur de Disney en obtenant l'accès à des informations internes sensibles. Alors qu'il était en communication avec l'équipe de cybersécurité de l'entreprise, l'ingénieur a reçu un message sur Discord, dans lequel le hacker affirmait détenir des renseignements confidentiels sur sa vie personnelle et professionnelle. Pour prouver ses dires, il a révélé un détail concernant un déjeuner dont seuls les membres d'un canal Slack privé de Disney avaient connaissance.

Refusant de céder au chantage, l'ingénieur a immédiatement alerté la police. En représailles, le pirate a publié 44 millions de messages internes de Disney, divulguant des informations clients, des numéros de passeport d'employés ainsi que des données stratégiques liées aux revenus des parcs et des services de streaming. Se faisant appeler Nullbulge, il prétend appartenir à un groupe d'hacktivistes russes opposés à l'usage de l'intelligence artificielle, mais les experts en cybersécurité estiment qu'il s'agit plutôt d'un individu isolé, motivé par des gains financiers.

L'attaque a eu de lourdes conséquences pour l'ingénieur, dont les comptes bancaires et les numéros de sécurité sociale ont été compromis. Le pirate a également pris le contrôle de ses réseaux sociaux et des comptes Roblox de ses enfants, exposant sa vie privée et celle de sa famille. Quelques semaines plus tard, Disney l'a licencié après la découverte de contenus pornographiques sur son ordinateur professionnel, une accusation qu'il dément formellement. Ce renvoi lui a fait perdre son assurance santé et d'importantes primes financières, estimées à 200 000 dollars.

Sa famille, convaincue qu'il est victime d'un piratage, a témoigné en sa faveur. Sa sœur a déclaré qu'un outil téléchargé "juste pour s'amuser avec les enfants" aurait pu être infecté sans qu'il en ait conscience. Face à cette situation dramatique, ses proches ont lancé une campagne GoFundMe afin de l'aider financièrement, soulignant que cette cyberattaque affecte non seulement l'ingénieur, mais aussi ses deux jeunes enfants.

Synthèse article sur L'IA :

Synthèse – OpenAI Dévoile GPT-4.5 : Une IA Plus Naturelle et Intuitive

OpenAI a annoncé officiellement le lancement de GPT-4.5, une mise à jour de son modèle d'intelligence artificielle, le 27 février 2025. Conçu pour les utilisateurs de ChatGPT Pro, ce modèle offre la promesse d'une interaction plus authentique, d'une meilleure compréhension des émotions et d'une diminution des fautes de raisonnement.

Face à une compétition acharnée avec des modèles tels que DeepSeek R1, Claude 3.7 et Microsoft Phi 4, OpenAI devait impérativement apporter une amélioration notable. GPT-4.5 se démarque par son aptitude à saisir les subtilités sociales et situationnelles, fournissant de ce fait des réactions plus instinctives et appropriées. Selon Sam Altman, le directeur général d'OpenAI, ce modèle est le premier à donner l'impression d'avoir une conversation avec un individu réfléchi.

En dépit de ces progrès, GPT-4.5 ne surpasse pas encore les modèles spécialisés, en particulier dans les domaines des mathématiques et des sciences, où d'autres alternatives comme o1 et o3-mini d'OpenAI continuent à montrer de meilleures performances. En outre, ce modèle demande une formation et une exploitation coûteuses, ce qui restreint pour l'instant son accès à un auditoire limité.

L'introduction de GPT-4.5 a généré des réactions partagées, certains anticipant déjà l'arrivée de GPT-5, prévu pour mai 2025. Ce modèle à venir comportera un moteur de raisonnement innovant (O3) et aura pour but d'harmoniser les différentes variantes d'IA, ce qui rapprochera OpenAI de son but d'atteindre une intelligence artificielle générale (AGI).

Diffusion des Résultats :

Pour qu'une veille technologique soit performante, il ne faut pas qu'elle se résume à la collecte et l'examen des données, elle doit aussi être communiquée aux acteurs concernés. Les comptes rendus internes sont le format de choix pour communiquer les résultats de la surveillance aux équipes techniques, aux décideurs et aux responsables en cybersécurité.

Il est également conseillé de publier les résultats pour participer à la prise de conscience des problématiques liées à l'IA et à la cybersécurité. La rédaction d'articles de blog ou de publications sur les réseaux professionnels contribue à rendre accessibles certaines innovations technologiques et à éveiller un large public aux dangers associés aux cyberattaques, ainsi qu'aux possibilités offertes par l'intelligence artificielle.

Pour conclure, s'impliquer dans des conférences, webinaires et tables rondes, constitue une bonne occasion de diffuser les résultats des analyses issues de la veille et d'interagir avec d'autres spécialistes du secteur. Ces manifestations contribuent à approfondir la connaissance des enjeux liés à l'intelligence artificielle et à la cybersécurité, tout en stimulant la coopération entre chercheurs, entreprises et organismes publics.