ICANN82 | Forum de la communauté – Table ronde de NARALO sur l'IA et l'utilisation malveillante du DNS
Lundi 10 mars 2025 – 15h00 à 16h00 PST

MICHELLE DESMYTER

Veuillez prendre place. Nous allons commencer dans quelques instants. Merci.

Cette séance va commencer. Veuillez lancer l'enregistrement. Merci.

Bonjour et bienvenue à tout le monde à la « Table ronde de NARALO : utilisation malveillante du DNS et intelligence artificielle ». Je m'appelle Michelle DeSmyter. Je suis responsable de la participation à distance.

Cette séance est enregistrée et elle est régie par les normes de comportement de l'ICANN ainsi que la règle anti-harcèlement de l'ICANN.

Les questions et les commentaires seront lus à haute voix s'ils sont soumis dans la fenêtre questions-réponses. L'interprétariat est disponible en anglais, français et espagnol. Si vous souhaitez prendre la parole, veuillez lever la main sur Zoom. Les participants à distance auront la permission d'activer leur micro pour prendre la parole. Les participants sur place utiliseront un micro. Veuillez

indiquer votre nom, et la langue dans laquelle vous parlerez, si elle est autre que l'anglais. Veuillez parler à un rythme raisonnable.

Sur ce, je donne la parole au président de NARALO, Greg Shatan.

GREG SHATAN

Bonjour à toutes et à tous. Bienvenue à notre table ronde sur l'utilisation malveillante du DNS et l'intelligence artificielle. Ce sont des sujets qui viennent à point nommé.

Nous avons des sommités qui sont avec nous. Tout d'abord, je vais vous présenter à ma droite Jeff Bedser. C'est un entrepreneur. C'est un leader dans tout ce qui a trait à l'atténuation de l'utilisation malveillante du DNS, et c'est le PDG de CleanDNS. Et donc il s'intéresse à la cessation des utilisations malveillantes. Il a été détective privé, et donc il s'appuie sur son expérience sur les thématiques qui nous intéressent ici, à At-Large. Il siège à différents conseils d'administration. Il a 25 ans d'expérience dans le secteur de la cybercriminalité. Donc il est à SSAC depuis 2007. Et donc, auparavant, il a été au Conseil d'administration PIR. Jeff, merci d'être avec nous.

Tim Maurer, à sa droite. Il est actuellement le directeur senior des politiques de cybersécurité chez Microsoft. Il collabore avec les parties prenantes internes et externes de Microsoft. Auparavant, il travaillait au sein du gouvernement américain, dans les domaines de la cybersécurité et de la technologie. Et il était également donc dans le conseil éponyme de la Maison-Blanche. Il a passé dix ans à

travailler dans la cybersécurité et il est l'auteur donc d'un ouvrage sur la cybersécurité.

Et à sa droite, nous avons Ana Neves. Vous la connaissez probablement comme étant la représentante du GAC, donc du Portugal au GAC. Elle est membre donc de l'IGF MAG. Donc c'est un conseil d'administration qui permet à l'IGF de continuer. Elle est donc responsable du conseil d'administration de .PT notamment. Donc elle est vice-présidente de la Commission des Nations Unies sur les technologies, les sciences et le développement. Elle a une longue expérience. Et en 2018, elle a été donc reconnue comme étant chef de file dans ce domaine. Donc cela donne l'impression que, moi, je ne suis qu'un tout petit acteur dans ce domaine.

Sans plus attendre, je voudrais lancer la discussion, si vous le voulez bien. Passons à notre première diapositive. Avant de poursuivre, je voudrais dire que toute cette présentation a été rédigée par le Copilot Microsoft. Alors, Dieu n'est pas mon copilote. C'est Microsoft qui est mon copilote. Donc Copilot. Nous dit donc que l'usage malveillant du DNS fait référence à un abus, un usage abusif du DNS à des fins malveillantes tel que l'hameçonnage, la diffusion de logiciels malveillants. Donc, il y a malware, botnet, donc l'hameçonnage, le dévoiement, le spam.

Bon, à proprement parler, le spam n'est pas considéré comme étant donc un usage abusif du DNS. Nous en avons parlé il y a quelques années déjà sur les limites et les définitions de l'usage

malveillant du DNS. Eh bien, nous allons tout simplement nous contenter de cette définition dans notre débat.

Comme vous le savez, comme chacun le sait, comme Copilot le sait, donc, l'intelligence artificielle a un rôle à jouer dans le combat que nous menons. Je vous laisserai le soin de lire tout cela en temps utile. Je préfère écouter mes collègues, plutôt que ce que le Copilot de Microsoft a à dire. Donc les PowerPoint sont aussi ceux qui ont été choisis par Microsoft, ainsi que les images. D'habitude, je suis très créatif, mais, aujourd'hui, j'ai décidé de laisser Copilot faire le travail.

Alors tout d'abord, je voudrais donner la parole à Tim. Mais avant de lui donner la parole, l'utilisation malveillante du DNS est aussi manifestée par l'IA, et l'IA -- des exemples de cela dans ces diapositives, l'IA crée de nouvelles formes et des mutations dans l'utilisation malveillante du DNS. Donc c'est pour le meilleur et pour le pire que nous avons l'IA. Donc il y a donc cette interface générale de la cybersécurité et de l'IA. Comment est-ce que cela touche le DNS ? Tim est le seul -- donc la seule personne ICANN qui a des renseignements éclairés.

TIM MAURER Merci beaucoup, Greg. C'est un plaisir d'être avec vous. Je discutais avec Ana. Et la dernière fois que nous nous étions rencontrés, c'était avant la pandémie. Et c'est un plaisir tout particulier d'être

avec vous aujourd'hui, car c'est une communauté qui nous intéresse beaucoup.

Chez Microsoft, nous avons beaucoup de choses à faire en commun pour, donc, la défense de la cybersécurité et faire en sorte que les progrès pour l'humanité puissent avancer. J'ai rejoint Microsoft il y a un an à peu près. Ça a été une expérience fascinante d'être donc dans une entreprise qui est tête de file de l'innovation en ce qui concerne l'utilisation malveillante du DNS.

Nous, nous savons que l'intelligence artificielle existe déjà depuis longtemps. Et elle est utilisée pour de nombreuses choses, y compris l'atténuation de l'utilisation malveillante du DNS. Et avec Copilot, il y a donc l'intelligence artificielle générative qui nous permettra d'abattre les cloisons, les barrières et permettre à un grand nombre de personnes d'utiliser l'IA générative et d'utiliser les outils pour atténuer cette utilisation malveillante du DNS. J'espère pouvoir avoir un échange interactif.

Alors notre entreprise envisage l'IA dans le contexte de la sécurité et de l'utilisation malveillante du DNS à travers trois angles. Comment maximiser les capacités de défense ? Comment mobiliser les clients et les communautés telles que la vôtre, les inciter à adopter cette technologie IA le plus rapidement possible ? Et comment utiliser les nouveaux outils pour faire progresser leur activité, comment utiliser l'IA à des fins de défense, d'atténuation ? Et plus cela sera fait rapidement, mieux cela nous permettra donc

non seulement de rattraper, mais de devancer ceux qui sont nos ennemis.

Deuxièmement, je souhaiterais souligner que notre entreprise est depuis longtemps déjà concentrée sur la prise de mesures en amont. Donc, nous suivons beaucoup de pistes de personnes qui mènent des activités non autorisées. Et nous avons constaté que les acteurs malveillants utilisent les produits d'intelligence artificielle. Donc, de toute façon, nous ne souhaitons pas que les criminels soient plus productifs, donc nous prenons des mesures en amont. En tout cas, c'est l'engagement de nos équipes dirigeantes. Nous souhaitons donc habiliter les défenseurs de la cybersécurité.

Et enfin, notre entreprise se concentre sur ce qui doit être fait de notre côté pour faire en sorte que la technologie soit utilisée à bon escient et comment protéger nos produits et nos services qui, comme vous le savez, donc, sont sous la menace de nombreux acteurs malveillants. Donc voilà le cadre que je voulais donner pour notre échange.

GREG SHATAN

Merci beaucoup Tim. Vous avez mentionné la défense, donc la défense contre les utilisations malveillantes du DNS. Et donc c'est une transition vers Jeff Bedser, dont la vie a été définie par ces notions de défense et de nettoyage du DNS. Je ne dirais pas de lui que c'est le shérif de notre ville, mais en quelque sorte il a un badge. Mais en fait, nous n'en avons pas besoin. Alors je voudrais que vous

creusiez ce sujet. Vous avez des diapositives qui ne sont pas rédigées par le Copilot de Microsoft, mais par vos soins. Alors vous pouvez nous les présenter.

JEFF BEDSER

Merci Greg. Je ne nierai pas que j'ai utilisé l'intelligence artificielle avec succès. Alors je vais dire que j'ai échangé avec ce groupe. Hier, nous avons parlé donc de l'utilisation malveillante du DNS de façon générale. Alors ici, nous allons parler du DNS. À quel point est-il le propre, net, en quelque sorte ? Et nous avons dit l'autre jour que, si vous y pensez, donc, il faut savoir que nos ennemis sont déjà très loin. Ils nous ont devancés peut-être. Et il va falloir avancer à grands pas.

Ce que je vais dire, c'est que nous avons déployé des modèles. Parce qu'il y a donc des utilisations malveillantes qui existent, mais nous ne les connaissons pas toutes, car elles ne sont pas toutes signalées. Donc, il y a quatre ou cinq sources où nous pouvons trouver des cas signalés. La plupart du temps, on ne signale pas l'hameçonnage par SMS, les différents types d'hameçonnage. La meilleure solution, c'est de faire de la détection rapide.

Et c'est-ce que fait CleanDNS. Nous sommes entre les entreprises qui détectent les abus, entre, donc, les opérateurs de registre, les bureaux d'enregistrement notamment. Et nous essayons d'accélérer non seulement le signalement, mais la détection, afin

d'accélérer les processus. Donc on m'a souvent appelé donc comme étant le shérif.

Nous avons découvert dans une étude qu'il y a 96 heures entre le signalement et l'atténuation. Et donc combien de personnes peuvent cliquer sur un lien d'hameçonnage pendant ce temps ? Et donc combien d'entre vous pourraient se faire avoir pendant ce laps de temps ?

J'en arrive à mon point suivant. SAC115 était le modèle qu'on a décidé d'utiliser pour CleanDNS, parce que beaucoup de personnes dans l'industrie nous ont indiqué le fait que c'était un très bon document académique, mais que les solutions dont on parlait dans ce document n'allaient pas émerger. Et donc nous avons créé un cercle vertueux. Et je suis reconnaissant envers les personnes qui utilisent ce terme.

Nous tentons de résoudre le problème, mais pas seulement. Nous tentons de faire en sorte que les signalements soient de meilleure qualité et que notre secteur soit meilleur. Parce qu'une fois qu'on a trouvé un problème, il s'agit de trouver une solution et de tenter de l'éliminer. Donc il s'agit d'un modèle d'agrégation de données pour pouvoir ajuster les différents processus.

Je vais donc maintenant passer à l'IA. Diapo suivante, s'il vous plaît !

Alors tout d'abord l'IA est meilleure que les êtres humains. Si vous êtes ici à cette conférence et que vous n'avez pas un smartphone,

expliquez-moi pourquoi. Certaines entreprises vous font croire quelque chose qui n'existe pas, qu'il s'agisse de banques de télécommunications ou de votre médecin. Il y a toujours un faire croire, une image qui vous fait croire que la page sur laquelle vous êtes est bien celle de votre médecin ou de votre banque, par exemple. Mais les acteurs de l'hameçonnage ne connaissent pas toutes les entreprises du monde, tandis que les systèmes de l'IA peuvent tirer les données du système qui permettent de comparer les images et d'analyser les images, les logos, par exemple, des entreprises, pour savoir s'il y a une tentative d'attraper et de capturer des références, des informations.

Et quand je parle d'analyse d'image, je veux parler notamment de capture d'écran pour voir, pour détecter un hameçonnage par SMS ou un hameçonnage simple. Donc on peut faire une capture d'écran, l'analyser, analyser son contenu. Et l'IA est très douée pour dire que oui, cette image est structurée de façon qui me fait dire qu'il s'agit là d'un hameçonnage. Il s'agit aussi de se pencher sur le texte et les autres éléments de cette image.

Et voilà, je ne suis pas très doué pour faire défiler les diapos de ma présentation, et c'est écrit en très petit surtout, donc ce n'est pas très lisible.

Donc, au cours des 20 dernières années, notre secteur a été basé sur un système de signalement par e-mail basé sur des données structurées. Et l'analyse d'image vous permet de tenter de savoir ce qui se passe. Lorsque quelque chose nous arrive, eh bien, on

veut souvent raconter l'histoire, expliquer ce qui s'est passé et expliquer, par exemple, ces histoires à tout le monde. Ces histoires sont très fascinantes, mais elles prennent beaucoup de temps à lire. Et donc il s'agit de savoir quel est le nom de domaine, l'URL impliquée, et de quoi on l'accuse. D'hameçonnage ? Bien, c'est la seule chose qui m'intéresse.

Le reste de l'histoire ne m'intéresse pas pour tous les efforts d'atténuation des risques. Et je vais encore une fois lire le texte écrit en tout petit caractère. Donc on parle d'intégration de vecteurs et de similarités par bloc de texte. Par exemple dans le cadre des courriers indésirables, est-ce qu'une entreprise utilise la même structure, le même format pour toute une série de mails ? Et on peut utiliser les relations contextuelles pour comprendre qu'il existe 1000 exemplaires d'un même mail, par exemple. Et on peut se baser aussi sur l'heure, par exemple, à laquelle l'e-mail a été envoyé pour savoir, pour déterminer ou pas s'il s'agit d'une tentative d'hameçonnage.

Donc la détection des logos des entités, la détection des données personnelles et de beaucoup de domaines qui vont être utilisés pour l'utilisation malveillante du DNS ou qui ont été utilisés à cette fin sont inclus dans des pages de parking de domaine. Et lorsque les gens consultent cette page, eh bien, ils ne voient aucun problème. Et puis ils cliquent sur un lien et arrivent sur une page d'hameçonnage. Et donc c'est aussi intéressant de pouvoir utiliser l'analyse d'image, car on n'a pas besoin de se servir d'êtres

humains pour ce faire. Et cela réduit le temps nécessaire pour détecter tout usage abusif du DNS.

Alors nous souhaitons, bien sûr, peaufiner, ajuster ce processus, car il est très bon pour détecter les tentatives d'hameçonnage, mais tout le monde ici sait qu'il y a bien d'autres types de formes de l'utilisation malveillante du DNS. Et la plupart des parties représentées dans la ccNSO et la Chambre des parties contractantes le savent. Comment donc ajuster ce processus pour pouvoir détecter qu'une telle ou telle forme de l'utilisation malveillante du DNS est effectivement une utilisation malveillante ? C'est difficile, car s'il faut traiter le code et déterminer s'il s'agit d'un code malveillant, eh bien, les systèmes sont exposés à des risques. Et donc, si on utilise l'IA, on peut déterminer les schémas, la façon dont tel ou tel utilisateur a été ciblé.

Et enfin l'aspect IA générative. On l'utilise pour analyser notamment les communications que reçoivent les destinataires. Et on utilise cela pour créer des modèles. Et tout le monde sait, bien sûr, comme Greg l'a expliqué, qu'il faut toujours être prudent avec l'IA générative, car parfois cela peut aller loin. Il y a des documents qui existent sur l'utilisation malveillante du DNS et qui comportent énormément de notes de bas de page qui menaient à des sources, en fait, qui n'existaient pas. Donc il faut faire très attention quand on utilise l'IA générative.

Diapo suivante, s'il vous plaît. Qu'en est-il des failles ? La transparence. Alors, toutes les personnes qui travaillent avec moi ne cherchent pas à prétendre que l'on comprend tout de cette technologie. Il faut aussi faire attention aux biais liés aux données. Et c'est une inquiétude tout à fait légitime, notamment lorsqu'on se penche sur les différents jeux de données qu'on utilise. Par exemple, on a souvent des faux positifs, mais on ne sait pas qu'il s'agit d'un faux positif. Si, par exemple, j'ai toute une série d'images qui ne proviennent pas d'une banque et qui sont incluses dans un jeu de données, eh bien, cela peut parfois mener à des biais liés à ces données et donner des faux positifs, des résultats faux positifs ou des faux négatifs.

Et il y a aussi la question de la gestion des cas sensibles et des données sensibles. La question du caractère privé des données. Et je pense que c'est ma dernière diapo. Et donc je vais vous rendre la parole, Greg.

GREG SHATAN Merci beaucoup Jeff. C'était très intéressant et tout à fait opportun de parler de cela. Je vais maintenant me tourner vers Ana Neves. Ana, je vous en prie, allez-y. Dites ce que vous voulez.

ANA NEVES Merci beaucoup. Bon, je dis ce que je veux. Alors je vais commencer par dire que c'est un plaisir d'être ici, à Seattle, et notamment ici, au sein de l'At-Large. C'est toujours un honneur de participer aux

réunions de l'At-Large, car, vraiment, depuis le début de mes activités au sein du GAC, j'ai toujours défendu le fait que le GAC et l'At-Large devraient travailler de concert, car nous devrions accroître nos efforts dans le secteur public, et ce pour les utilisateurs finaux.

Alors que puis-je dire après ces interventions de la part de mes collègues Jeff et Tim concernant les utilisations malveillantes du DNS. Eh bien, ce que je peux dire concernant Microsoft, c'est qu'il est très bon de savoir que vous disposez d'une politique cherchant à atténuer les utilisations malveillantes du DNS. Mais qu'en est-il des autres entreprises qui ne disposent pas de ce genre de politiques, qui envisagent de créer et de mettre en œuvre ce genre de politiques ? Pouvons-nous nous fier aux entreprises ? Pouvons-nous compter sur elles pour qu'elles se dotent de ce genre de politiques ?

Aussi, s'agissant de la perspective du secteur public que je voudrais vous présenter, raison pour laquelle aussi j'ai été invitée à participer à cette table ronde, eh bien, moi, je voudrais insister sur la coopération entre le secteur public et privé et sur l'importance des partenariats, la coordination avec les autorités de réglementation.

Et il faut, bien sûr, toujours garder la cybersécurité à l'esprit, et, notamment, au sein des gouvernements lorsqu'ils envisagent de mettre en place ou d'élaborer des politiques. Il faut toujours qu'ils gardent cela à l'esprit. Et tous ces efforts sont importants pour

garantir une utilisation de l'IA permettant de garantir la sécurité du DNS.

Autre point important que j'aimerais soulever, eh bien, c'est le développement de cadres de gouvernance de l'IA. Je pense qu'il est judicieux d'en parler ici à l'ICANN et avec l'At-Large. Nous devrions envisager des politiques qui permettent une sécurité alimentée par l'IA, qui soit alignée sur les principes de la gouvernance de l'Internet. Et la révision SMSI+20 a pour objectif d'intégrer ces processus en parallèle du Pacte numérique mondial. Et on devrait vraiment garder tout ceci à l'esprit dans le cadre de ces processus et de ces négociations.

Autre point important, il s'agit d'habiliter les utilisateurs finaux. Il nous faut davantage de transparence pour permettre aux individus de faire des choix informés s'agissant de leurs paramètres de sécurité. Il est nécessaire d'informer les utilisateurs finaux. Aussi, le renforcement des capacités est toujours important pour renforcer nos efforts. Aussi, je dirai que le renforcement de la coopération, l'accroissement des partenariats public-privé, le développement de cadre de gouvernance de l'IA et l'habilitation des utilisateurs finaux sont les points sur lesquels je souhaiterais insister, notamment dans le cadre de ma première intervention.

GREG SHATAN Merci beaucoup Ana. Pourrions-nous en revenir à la diapo sur laquelle je m'étais arrêté, s'il vous plaît ? Alors, passons à la diapo

suivante, ce qui va nous permettre de mentionner certains des points que Copilot a soulevés.

Alors on peut déjà observer des attaques d'hameçonnage alimentées par l'IA, générées par l'IA, qui imitent la confiance, qui contournent les filtres de sécurité, qui analysent la syntaxe, la grammaire qui est utilisée.

L'IA peut être aussi utilisée pour préjuger les références. Elle peut analyser des dizaines de milliers de références en une seconde, ce qui permet donc des attaques de force brute, et ce, de manière répétée. Cela permet un grand volume d'attaques.

Ensuite, il existe aussi des logiciels malveillants qui peuvent être alimentés par l'IA, c'est-à-dire par exemple un code qui s'adapte de manière dynamique pour éviter qu'il ne soit détecté, pour l'aider à le dissimuler, des logiciels antivirus. Donc tout cela peut se faire aussi grâce à l'IA.

Diapo suivante, s'il vous plaît. Nous pouvons aussi observer des reconnaissances et des outils de scan qui agissent à un rythme très rapide pour automatiser les attaques de pénétration, en identifiant notamment les points faibles des infrastructures de sécurité. Et il y a aussi des attaques d'ingénierie sociale automatisées qui permettent de contourner les filtres traditionnels. Et cela indique clairement la raison pour laquelle nous devrions utiliser l'IA de manière sophistiquée et continuer à améliorer nos fonctionnalités de défense.

Diapo suivante, s'il vous plaît ! Certaines de ces solutions, bien sûr, sont rendues possibles par les outils IA. Nous fixons une confiance 0 à 1 — niveau de confiance 0. Vous avez constaté qu'il y a beaucoup de méthodes d'authentification multifactorielle notamment. Donc nous avons rehaussé les mesures.

Diapositive suivante, s'il vous plaît. Donc, les plateformes doivent donc rehausser leur niveau d'intelligence. Il y a bien évidemment le facteur humain. Il faut de toute façon former des personnes réelles, des employés formés, des utilisateurs finaux pour leur permettre de reconnaître l'hameçonnage, les attaques. Il faut une sensibilisation constante.

Au sein de mon entreprise, j'ai réalisé la formation annuelle que nous faisons sur la cybersécurité. Diapositive suivante. En tant qu'avocat, je connais le besoin de continuer, donc, dans le domaine de création de lois. Et il faut que l'aspect juridique et éthique avance aussi rapidement. Il faut que l'on puisse régir l'IA de façon plus responsable. Et il y a des entreprises qui travaillent avec l'IA. Il faut qu'elles soient conscientes des utilisations abusives du DNS. Ce n'est pas forcément la première chose qui vient à l'esprit. Alors il faut savoir si les entreprises qui utilisent l'IA collaborent suffisamment avec le reste de l'écosystème. Ensuite, il faut continuer à prendre des mesures proactives. Il faut donc parfois même avoir un certain niveau de paranoïa.

Ici, c'est la biographie de Jeff Bedser. C'est Copilot qui m'a proposé cette biographie. Elle semble assez exacte, sans trop d'erreurs.

Diapositive suivante, s'il vous plaît. Donc j'ai tout d'abord demandé Tim à --revenons en arrière. Alors, regardez. Dès l'année indiquée à l'avant-dernier paragraphe, à l'année 2071. En fait, c'est parce qu'il y avait une note en bas de page que je n'avais pas retirée. C'est une erreur humaine en réalité.

Diapositive suivante, s'il vous plaît. Voilà, ça, c'est la première tentative d'utilisation de l'IA. Donc c'est plus ou moins précis, mais pas forcément à jour. Diapositive suivante. Donc j'ai ajouté Microsoft dans l'invite. Et on voit là le résultat. Donc Microsoft ne connaît pas vraiment bien Microsoft -- le Copilot ne connaît pas encore bien Tim. Diapositive suivante. J'ai voulu en savoir plus sur ses hobbies, ce qu'il souhaite faire, ce qu'il aime faire. Dans ses loisirs, il aime faire du vélo, des randonnées, les activités en plein air, la photo.

Diapositive suivante. Ici, c'est la biographie d'Ana. Et ça n'est pas à la hauteur de la biographie qu'elle nous a proposée. Mais ce n'est pas mal. Ensuite, nous avons la diapositive suivante. C'est la première tentative. Donc, ici, il y a son deuxième prénom que, moi, je n'avais pas utilisé. Elle est donc directrice de la santé numérique, donc, à l'Imperial College de Londres.

Diapositive suivante. Ici, c'est moi. Si vous voulez avoir des renseignements sur ma personne. Alors, ici donc, on nous donne des dates inexactes, un nombre d'années inexact en ce qui concerne ma participation à l'ICANN. Diapositive suivante. Et voilà, ici c'est une photo de Seattle sous le brouillard -- dans le brouillard,

telle que choisie par le Copilot de Microsoft. C'est une jolie photo, mais c'est un petit peu une image d'apocalypse.

Alors maintenant, entamons la discussion sur les questions de l'IA défensive et offensive. Je vais commencer par vous Jeff. Qu'est-ce que vous anticipez dans les étapes suivantes, quand vous voyez les changements, les évolutions dans l'IA ? Comment est-ce que cela peut avoir un impact sur l'utilisation malveillante du DNS ?

JEFF BEDSER

C'est une question intéressante. Je pense que l'espace que nous allons voir, c'est donc la gestion des identités, les fausses identités, la capacité de l'IA à créer des cartes d'identité locales avec l'identité de personnes réelles, avec des photos, de sorte qu'il sera impossible de déceler les fraudes.

Cela est une véritable possibilité. Donc ce n'est pas son IA à lui, sans sarcasme. Donc par exemple, quand on a un numéro de carte de crédit qui n'a pas encore été utilisée. On peut la faire correspondre avec une personne qui est domiciliée en Pennsylvanie. On a un permis de conduire issu donc de Pennsylvanie. Alors on prend une photo et donc on a une image qui est générée. Et donc quand nous allons dans un domaine avec des numéros de carte de crédit volée, il n'y a pas de signalement. Et au bout du compte, tout cela est généré en 4 à 5 secondes.

Et là, on arrive à la cybercriminalité, qui est donc constituée de numéros de carte de crédit volée, d'identité volée. Et vous avez ces identifications, ces usurpations d'identité. C'est ce qui émerge.

GREG SHATAN

Merci Jeff. Tim, je vais vous poser plus ou moins la même question. Je vais vous poser donc la question suivante : les développeurs d'IA sont-ils conscients de notre espace d'abus du DNS comparé aux autres espaces malveillants ?

TIM MAURER

Alors la réponse est oui. Tout simplement parce que c'est dans l'intérêt de Microsoft d'être actif dans ce secteur et de faire en sorte que le DNS ne fasse pas l'objet d'une utilisation malveillante par des acteurs malveillants. Nous essayons d'anticiper la façon dont les personnes malveillantes l'utiliseront.

Il y a eu donc des prévisions de scénarios beaucoup plus apocalyptiques. Nous n'avons pas vu donc l'apparition de ces logiciels malveillants. Donc nous utilisons l'IA de façon progressive pour contrer les menaces que nous connaissons déjà. Je pense qu'il est intéressant de remonter à il y a quelques mois pour voir comment les défenseurs utilisaient l'IA. Et on voit qu'il est important de savoir comment la technologie est utilisée, comment les invites sont utilisées et comment traduire cela en technique défensive.

Par exemple, si vous utilisez Copilot pour des mesures de sécurité, pour produire des rapports, pour détecter des anomalies, et donc, quand vous voyez qu'il y a une anomalie qui est détectée, il faut la production d'un résumé. Et il faut plutôt inclure davantage d'anomalies potentielles plutôt que moins, afin que nous puissions tenir compte des faux positifs. Nous préférons détecter de faux positifs, plutôt que de laisser passer des anomalies non détectées.

Donc il faut exploiter l'IA pour pouvoir utiliser les ressources humaines de façon optimale. Donc, tout à l'heure, vous avez parlé donc de l'usurpation d'identité. Alors, du point de vue de la défense, il y a un manque de compétences, un manque de formation. Alors, nous allons essayer d'utiliser l'IA pour réduire le nombre d'entrées et permettre aux personnes qui n'ont pas forcément de fortes compétences pour renforcer le travail de défense.

Pour revenir à ce que disait Ana, il y a donc l'importance des interactions entre le secteur public et privé. Au cours des dernières années, il a été très intéressant de voir les effets de la publication de ChatGPT et la rapidité avec laquelle les gouvernements ont mis en œuvre des règles, des politiques. Et évidemment, les entreprises ont dû avancer rapidement pour s'y conformer. Et nous avons dû faire en sorte que nous ayons des mesures d'atténuation. Et il a été intéressant de voir à quel point, avec quelle rapidité, les régulateurs et les gouvernements ont mis en place des règles que

les entreprises, telles que Microsoft, par exemple, ont dû se conformer.

GREG SHATAN Merci. Ana. Alors de votre point de vue, du point de vue européen, je voudrais savoir ce que vous en pensez. Nous, nous sommes aux États-Unis. Les choses changent considérablement aussi, mais je serais intéressé à connaître davantage d'autres points de vue.

ANA NEVES Oui, absolument. En Europe, je pense que nous comprenons les menaces. Nous savons que ces menaces d'utilisation malveillante du DNS augmentent rapidement. Nous essayons d'élaborer des directives pour la prévention de l'utilisation malveillante, tout en tenant compte de la confidentialité des données.

Nous avons un rapport élaboré au sein de l'Union européenne. Donc, il y a cette étude sur l'usage malveillant, donc, qui propose des recommandations pour les opérateurs de registre et les bureaux d'enregistrement. Il est intéressant de voir que les domaines tels que .DE et .UE sont des domaines qui sont moins utilisés, qui font moins l'objet d'usages malveillants. Alors là, l'utilisation malveillante n'est pas aussi prononcée.

Alors, la question de la mise en œuvre. Le travail de collaboration entre les secteurs public et privé, cela est très important. Il est important de travailler ensemble aussi, de travailler avec les chercheurs dans le domaine de l'IA. Ce sont les personnes qui

peuvent nous aider à mettre en place des mesures de prévention. Il faut prévenir. Il y a donc ces mesures, ces actions multipartites, l'importance d'inclure les chercheurs, les milieux universitaires, les développeurs d'IA dans tous ces processus, cela est extrêmement important. Et quand cette collaboration, cette implication est possible, quand on peut être à l'écoute des différentes parties prenantes, des différents acteurs, eh bien, nous parvenons à des politiques publiques, notamment pour répondre à ces questions d'utilisation malveillante du DNS.

GREG SHATAN Merci Ana.

Il y a beaucoup de personnes dans cette salle ainsi qu'en ligne avec nous aussi. Maintenant, je voudrais me tourner vers la boîte de questions-réponses sur Zoom. Si vous avez des questions à poser, n'hésitez pas à les y poser, et Michelle DeSmyter les recevra. Et si vous êtes sur Zoom, n'hésitez pas à lever votre main dans Zoom. Je vois que la main de Joanna est levée. Alors si vous n'êtes pas assis autour de la table, nous avons un micro volant que nous vous apporterons si vous avez une question à poser. Donc, restez bien assis. Nous vous amènerons un micro si vous avez une question à poser.

Nous allons commencer avec Joanna.

JOANNA KULESZA

Cindy, Joanna Kulesza au micro. Merci beaucoup pour toutes ces présentations. C'est incroyable d'en apprendre autant sur l'IA et le DNS.

J'ai une question par rapport pour rebondir à l'intervention vis-à-vis de l'intervention d'Ana. Et merci aussi à Tim pour cette présentation assez complète. Alors je ne sais pas si vous pouvez répondre à cette question en tant qu'employé de Microsoft, donc n'hésitez pas à parler de politiques mondiales de cybersécurité si vous le désirez.

Donc je vais commencer par quelques informations de contexte pour cette question. D'un côté, Microsoft a été essentiel pour appuyer l'Ukraine après l'invasion de la Russie. Les analyses de menaces à la sécurité, de Microsoft, ont été essentielles pour l'Ukraine. Et bien sûr, il semblerait maintenant que le partage d'informations a été ralenti. Maintenant, il semblerait qu'il pourrait reprendre. Cela démontre bien comment les réglementations nationales peuvent cibler les tentatives de garantir la cybersécurité, que ce soit au niveau de l'utilisation malveillante du DNS ou du partage d'informations.

Mais d'un autre côté, Ana a mentionné qu'il existe des améliorations -- des réglementations améliorées sur l'IA, notamment la législation sur l'IA. Et Senator Cruz, en décembre de l'année dernière, a dit que cette législation empêchait l'innovation et pourrait être considérée comme une intervention étrangère dans le développement de l'IA aux États-Unis.

Et cela m'amène à ma question. Notre bon ami, Marietje Schaake, a publié un livre intitulé IT Coup. Et je suis certaine que vous comprendrez le concept derrière ce titre. Nous en avons parlé récemment. Et il mentionnait qu'il était peut-être le bon moment pour les chercheurs qui n'étaient pas à l'aise à l'idée de mener leurs recherches aux États-Unis, il était peut-être temps pour eux de déménager en Allemagne, en Union européenne.

Et donc voilà, je ne sais pas si c'est une question à laquelle Microsoft souhaitera répondre, mais que pensez-vous de ceci ? Quel est le rôle de la réglementation s'agissant du développement de l'IA pour la cybersécurité ? Et avant que Jonathan me remette à ma place, je voudrais savoir comment les utilisateurs finaux peuvent peut-être travailler avec les gouvernements et avec le GAC pour garantir que l'IA permette de protéger les individus des cybercriminels. Je pense que c'est une façon assez diplomatique de poser une question, et je sais que vous pouvez, bien sûr, identifier les questions géopolitiques que je tente de traiter au travers de cette question.

Voilà, si vous avez une information à nous partager, n'hésitez pas.

TIM MAURER

Merci beaucoup Joanna. Je suis très heureux de vous revoir. Alors oui, bien sûr, je peux répondre aussi en tant qu'employé de Microsoft, car notre entreprise a été très claire s'agissant des défis géopolitiques auxquels nous sommes confrontés. Et notre entreprise est engagée pour la cybersécurité. Pour notre

entreprise, nos clients, pour les initiatives de sécurité, vraiment, la sécurité est un élément essentiel pour notre entreprise.

Alors que peut faire la communauté ? Eh bien, je pense que nous en sommes à un tournant, s'agissant des innovations technologiques. Les progrès que nous engrangeons grâce à l'IA et grâce à Microsoft nous ont permis de déployer de nouveaux produits et services régulièrement. Et nous allons continuer de le faire. Et la communauté peut nous aider à mieux utiliser la technologie pour le bien, notamment dans le contexte de la défense et de la sécurité. Et cela, vous pouvez être sûre que l'idée aussi, bien sûr, est pour certaines personnes de monétiser ces technologies, de savoir comment cela peut servir leurs objectifs.

Donc, nous, nous allons devoir adopter ces technologies. Si vous considérez qu'il y a des questions de recherche sur lesquelles Microsoft ou d'autres entreprises pourraient se concentrer, n'hésitez pas à nous le faire savoir, car nous sommes ici dans le même bateau. Et nous voulons vraiment tenter de maximiser le potentiel de la technologie pour le bien. Et donc, que ce soit au sein du GAC ou au travers d'autres mécanismes, nous avons besoin des contributions de la communauté.

Concernant votre premier point concernant l'UE et les réglementations, notre entreprise a été très claire. Nous pensons qu'il y a certains domaines pour lesquels nous avons besoin de réglementation. Microsoft s'inquiète peut-être plus du fait que la réglementation puisse devenir trop prescriptive et qu'elle ne

tienne pas compte du rythme accéléré du développement de la technologie. Il s'agit aussi de prendre en compte le fait que, dans certains pays, avec des différentes législations, eh bien, justement, ces différences de législations peuvent parfois créer des problèmes. Et donc actuellement, nous pensons que nous sommes à un tournant. Nous allons faire de grandes innovations technologiques. Peut-être est-il temps maintenant d'inverser la tendance au niveau systémique en faveur de la défense. Et nous, c'est vraiment un de nos objectifs et nous espérons pouvoir vraiment tirer le plein potentiel -- exploiter le plein potentiel de ces technologies pour la défense.

GREG SHATAN Merci beaucoup pour cette question. Alors la boîte questions-réponses, et puis, ensuite, nous donnerons la parole à Hadia.

MICHELLE DESMYTER Nous avons une question de la part de [Saïd Najib]. Quelles méthodes sont-elles utilisées au sein de CleanDNS par rapport à l'IA ?

JEFF BEDSER Eh bien, les méthodes que nous utilisons se retrouvent dans la présentation que j'ai donnée en interne. Nous utilisons Copilot. Mais sur notre plateforme, nous utilisons ce qui est disponible. Mais je ne peux malheureusement pas divulguer, en fait, les IA que nous utilisons bien sûr.

MICHELLE DESMYTER Question suivante, de Sivasubramanian : Pouvons-nous utiliser l'IA -- l'IA pourrait-elle être utilisée pour concevoir des sites d'hameçonnage qui seraient plus authentiques ?

JEFF BEDSER Alors, je vais répondre à cette question.

Alors, un site d'hameçonnage, ce n'est pas qu'une page sur laquelle on atterrit. Il y a des détails liés à l'infrastructure qui gère ce site. Donc pour la plupart des sites d'hameçonnage, eh bien, en fait, ils ont été conçus à des fins d'hameçonnage. La plupart du temps, l'enregistrement du nom de domaine est malveillant, donc l'âge du nom de domaine est un facteur clé, ainsi que les IP, la réputation. Très souvent, on se rend compte que l'adresse IP provient d'un hôte, d'un site d'hébergement bulletproof, comme on les appelle. Et voilà, si vous hébergez ce genre de site, cela veut probablement dire que vous vous livrez à des activités criminelles. Donc ce n'est pas qu'une page sur laquelle on atterrit. Mais effectivement, ces pages deviennent de plus en plus authentiques, semblent de plus en plus authentiques grâce à l'IA.

GREG SHATAN Nous allons maintenant donner la parole à Hadia Elminiawi.

HADIA ELMINIAWI Merci Greg. Ma question s'adresse à Ana.

Vous avez parlé de transparence s'agissant de l'IA. Aussi, ma question est la suivante. En pratique, que signifie la transparence ? Est-ce que cela sous-tend une IA open source ? Est-ce qu'on sait que l'IA open source ne permet pas de nous donner des éclaircissements sur la façon dont l'IA prend des décisions ? Il y a aussi des systèmes de boîtes noires, même parfois lorsqu'on utilise des systèmes open source.

Donc en termes pratiques, la transparence, qu'est-ce que ça veut dire ?

ANA NEVES

Merci pour cette question qui est très intéressante. Quand j'ai parlé de transparence et d'IA, je voulais parler notamment de transparence et de responsabilité. Les modèles d'intelligence artificielle doivent pouvoir être expliqués. Et donc il faut permettre aux opérateurs du DNS et aux agents de la cybersécurité de comprendre les processus de prise de décision de l'IA. Si on ne comprend pas ce qu'on contrôle, ce contre quoi on lutte, cela nous rend la tâche très difficile. Donc il faut que les modèles d'intelligence artificielle puissent être expliqués et compréhensibles, car cela nous permet de savoir peut-être plus facilement où est le problème.

GREG SHATAN

Merci. Jonathan, est-ce que vous aviez une question à poser ?

JONATHAN ZUCK

Oui, effectivement. Alors mon IA, bon, je l'appellerai du renseignement un peu antique pour détecter les tentatives d'hameçonnage, mais il semblerait que les courriels d'hameçonnage sont d'une façon ou d'une autre liés à des entités reconnaissables, des marques reconnaissables. Et très souvent dans ces e-mails, il y a un lien qui vous amène vers un site de dévoiement.

Est-ce qu'on peut savoir d'où vient cet e-mail si on se penche sur les liens qui sont intégrés dans cet e-mail, afin de savoir si ces liens proviennent bien de l'entité qui, semblerait-il, nous a envoyé ce courriel, ce courrier électronique ? Je me souviens un jour, j'ai été très surpris lorsque j'ai voulu envoyer un e-mail sur Gmail et Google m'a averti du fait que j'avais prévu de joindre une pièce annexée à cet e-mail. J'étais très enthousiasmé par cela. Mais en même temps, peut-être qu'il faut faire preuve de prudence par rapport à tout ce genre de fonctionnalités.

TIM MAURER

Eh bien, effectivement, il existe toute une série de systèmes en place qui sont basés sur l'apprentissage machine et sur les algorithmes. Et comparés à il y a 15 ou 20 ans, ces processus nous permettent de nous garantir qu'on ne reçoive pas trop d'e-mails. L'idée est de pouvoir assurer donc [d'e-mails] menant à des formes d'utilisation malveillante du DNS. Si un site est enregistré d'une façon ou d'une autre et qu'on s'inquiète de l'identité derrière ce site, peut-être à cause de la longueur de la page, eh bien, il y a un

potentiel à exploiter ici. Que doit savoir l'être humain pour pouvoir agir par rapport à ce genre d'e-mails ?

GREG SHATAN Nous avons déjà dépassé le délai imparti, mais je vais accepter encore une question. Amrita.

AMRITA CHOUDHURY Merci. J'avais deux questions. Je voudrais savoir de votre part, Tim et Jeff, ceci. Nous avons parlé de l'utilisation abusive de l'anglais et de la façon dont on peut utiliser l'IA pour aggraver cette utilisation malveillante. Mais qu'en est-il des non-anglophones ? On constate que ce genre de forme d'utilisation malveillante du DNS augmente. Dans les pays majoritaires, il y a des utilisateurs finaux qui ne sont pas si habiles s'agissant des outils numériques aussi. Que faites-vous pour ces utilisateurs finaux qui sont confrontés à des utilisations malveillantes du DNS ? Est-ce que vous avez des tutoriels, des vidéos ? Je pense que, autrefois, vous aviez des formations à l'habileté numérique. Donc voilà, que faites-vous pour ces utilisateurs finaux qui s'y connaissent moins ?

TIM MAURER Alors Microsoft est basé partout dans le monde et propose des services, des produits, dans différentes langues. Et nos efforts en matière de sécurité se font en anglais, mais pas seulement. Donc, nous avons vraiment pour objectif de faire en sorte que nos technologies soient accessibles à tout le monde, partout dans le

monde, et de faire en sorte que les mesures de sécurité et d'atténuation reflètent aussi le profil des utilisateurs qui les utilisent.

S'agissant du renforcement des capacités et de la formation, l'année dernière, nous avons annoncé, notamment concernant le Kenya et le Sud, l'hémisphère sud de notre monde, eh bien, que nous souhaitions que cette partie du monde soit plus habile en matière d'utilisation du numérique. Et bien sûr, nous voulons accroître nos efforts pour garantir que nous fournissons les formations permettant aux utilisateurs finaux d'utiliser nos produits et nos services et de tirer parti du meilleur potentiel de notre technologie.

Merci pour ce dernier commentaire. J'ai omis dans ma présentation de dire que c'était exclusivement de l'anglais, mais en fait l'outil est disponible dans plusieurs langues. Et l'IA, elle est tellement efficace pour nous, car justement nous n'avons pas besoin d'une personne humaine pour lire et pour connaître toutes ces langues. C'est une composante très importante. Nous gérons non seulement une majorité de langues, mais toutes les langues. Nous voyons donc de l'hameçonnage qui est réalisé dans toutes les langues. L'IA nous permet de réaliser un travail que nous ne pourrions pas faire avec des personnes humaines.

GREG SHATAN Merci Jeff. Nous arrivons à la fin de notre réunion. C'est une discussion qui va se poursuivre. Je voudrais remercier tous nos

panélistes : Jeff Bedser, Tim Maurer, Ana Neves. Je remercie tout le monde d'avoir été à l'écoute. Et donc, merci à NARALO pour cette table ronde. Nous allons ensuite passer à la séance de débat et d'échanges avec NARALO. Merci.

[FIN DE LA TRANSCRIPTION]