Stratégie

Les leçons de la fraude au président en deepfake évitée par Ferrari

le 14 Février 2025

En 2024, un pirate se fait passer pour le DG de Ferrari, Benedetto Vigna, dans une conversation audio Whatsapp. La méfiance et la maîtrise des deepfakes du cadre contacté a évité le pire au constructeur. (Photo Ferrari)

Le constructeur de voitures de sport Ferrari a évité de justesse une fraude au président dopée à l'IA, mi-2024. Les pirates se sont servis d'un deepfake audio qui aurait pu coûter cher à l'entreprise si un dirigeant n'avait flairé l'arnaque. De quoi tirer quelques leçons de cybersécurité face à une technique de plus en plus utilisée.

En juillet 2024, le DG de Ferrari demande dans la messagerie Whatsapp à certains de ses cadres de se préparer pour une opération importante de rachat, opération confidentielle. Si le scénario ressemble comme deux gouttes d'eau à celui d'une arnaque au président et interpelle les dirigeants, difficile pour eux de ne pas accéder à la requête. Surtout que le dirigeant, Benedetto Vigna, lance ensuite une conversation vocale avec l'un de ses interlocuteurs. La photo de profil du DG devant le logo de la marque, et surtout sa voix et son accent du nord de l'Italie ne laissent que peu de doute quant à son identité.

Pourtant, des sons métalliques et un rythme d'élocution inhabituel intriguent un des cadres concernés. Ce dernier décide de poser une question à laquelle seul le DG peut répondre. Il lui demande de lui rappeler le titre du livre qu'il lui a récemment recommandé. Résultat : aucune réponse et interruption de l'échange. Le pirate est démasqué. Ferrrari a bien failli être victime d'un deepfake sonore, une version dopée à l'IA de la fraude au président. Et c'est la connaissance du sujet et la prudence d'un de ses cadres qui l'a probablement sauvée d'une arnaque.

Quelques mois plus tôt, en février 2024, un comptable de l'antenne hongkongaise du bureau d'étude industriel Arup a fait la une de la presse mondiale après avoir versé 25 M$ à un hacker. Il

venait de participer à une réunion en visioconférence dans laquelle tous les autres participants se sont révélés, après coup, être des deepfakes incarnant plusieurs cadres de l'entreprise.

Près d'une organisation sur deux victime de deepfake

Dans une étude menée en mai 2024 par le cabinet d'études Cap Gemini Research Institute auprès de 1000 organisations, 45% des répondants ont admis avoir été victimes d'une attaque par deepfake dans les deux années précédentes. Ces créations par l'IA (GenAI, machine learning, generative adversarial networks) de contenus imitant quasi parfaitement la réalité dans une vidéo, une image, un texte ou une voix, deviennent de plus en plus crédibles et « simples » à réaliser.

Résultat, des avatars hyperréalistes se font passer pour des dirigeants d'entreprise au cours de réunions en visioconférence des équipes de l'entreprise. Comme Katarzyna Kapusta, pilote Friendly Hackers du CortAIx Lab du groupe Thalès, l'a expliqué à CIO l'an dernier, il s'agit pour les pirates de préparer l'intervention du double numérique du dirigeant sur le plan technique, mais aussi en récupérant suffisamment de données en amont afin d'être crédible durant la conversation en temps réel.

Comme les fraudes au président téléphoniques classiques ou le phishing, les deepfakes tirent leur efficacité de l'exploitation des biais cognitifs et de confiance qui empêchent de se comporter lucidement, et d'identifier une erreur, en particulier en situation de stress ou d'urgence et lorsqu'une autorité hiérarchique est impliquée. S'y ajoutent les biais de confirmation - un employé ne se méfie pas d'une situation crédible, à laquelle il s'attend - et le fait que le cerveau soit « éduqué » à croire ce qu'il voit et ce qu'il entend.

Installer une culture du scepticisme

Dans un article sur le sujet, la revue de management Sloane du MIT s'inspire du cas Ferrari pour lister les précautions à prendre et à ajouter aux démarches de cybersécurité face aux deepfakes. À commencer par la sensibilisation et l'acculturation des équipes à ces sujets. Chez le constructeur automobile, c'est bel et bien la prudence d'un cadre, qui s'est méfié rapidement et a ainsi repéré des défauts dans la voix de son DG, et sa connaissance de ce type d'arnaque qui se sont révélées essentielles pour éviter l'escroquerie.

Ainsi, comme pour d'autres types d'attaques, celles basées sur l'IA - les deepfakes en particulier - et la manipulation psychologique qui va de pair devraient faire l'objet de programmes de sensibilisation spécifiques afin d'augmenter le niveau de vigilance des employés, tout comme c'est déjà le cas la plupart du temps avec le phishing. Le magazine Sloane conseille de favoriser une culture du scepticisme vis-à-vis de demandes inhabituelles, même si elles semblent crédibles. Reste à trouver le juste niveau de confiance. Sur un sujet certes moins grave, le Ciso de Somfy groupe expliquait récemment à l'occasion du Cybershow Paris comment certains employés avaient refusé d'ouvrir des mails venant de la direction de la communication, après une campagne de sensibilisation au phishing particulièrement efficace !

Vérifier les identités et prohiber les applications tierces

Sloane évoque également l'installation de protocoles rigoureux de « vérification d'identité en plusieurs étapes pour toutes les communications de haut niveau ou sensibles ». Cela comprend

l'affectation de lignes de communication directes pour les décisions importantes et l'interdiction d'applications tierces non sécurisées, comme Whatsapp en particulier, pour des échanges sensibles - ce qui aurait sans doute dû éveiller en premier lieu les soupçons chez Ferrari - , au profit de solutions d'échanges chiffrés. Comme le montrent les deux exemples d'arnaques visant le constructeur automobile et Arup, les connaissances en matière de biais cognitifs apparaissent par ailleurs tout aussi importantes que la culture minimale de l'IA et des deepfakes.

Autre bouclier de protection possible contre les deepfakes, la surveillance permanente des activités à haut risque par des systèmes de détection des fraudes, couplée à des protocoles d'arrêt rapide et de notification immédiate d'activités suspectes. Les entreprises peuvent ainsi ajouter la détection de deepfakes à leur politique et à leurs systèmes de détection de fraude.

La protection contre les attaques de plus en plus sophistiquées, et de plus en plus crédibles, par deepfake impose ainsi des actions à plusieurs niveaux : acculturation et sensibilisation, évolution de la gestion des risques et de la prévention des fraudes, technologies de protection des communications sensibles et de vérification d'identité, etc.

A lire sur le même sujet

- Des deepfakes sophistiqués à l'assaut des finances des entreprises

- Les deepfakes frappent aux portes de l'entreprise

Article rédigé par

Emmanuelle Delsol, Journaliste
Suivez l'auteur sur Linked In,