KAIA-OR · LIVRE BLANC · INGÉNIERIE SOCIALE

Un second regard

Protéger les familles et les organisations contre les fraudes, la compromission de courriels d'affaires et l'usurpation d'identité, grâce à un second regard avant que quiconque réponde, paie ou ouvre quoi que ce soit, la décision restant entre les mains des personnes.

IDENTIFIANT DU DOCUMENT	WP-3
VERSION	1.0
DATE	3 novembre 2026
PUBLIC	Familles et particuliers; propriétaires de petites entreprises; responsables des finances, de la sensibilisation à la sécurité et des risques; équipes de sécurité des entreprises et des gouvernements

Table des matières

1. Sommaire 3

2. Le problème et le fonctionnement des leurres 4
   1. L’ampleur du problème 4
   2. Le fonctionnement des leurres modernes, de façon abstraite 4

3. Le modèle du second regard 6
   1. Pourquoi les verdicts ponctuels ne suffisent pas 6
   2. Le fonctionnement du second regard 6
   3. Le contexte compte 7

4. Familles et organisations 8
   1. Les familles et les proches âgés 8
   2. Organisations : fraude par fausse facture, usurpation de l’identité de dirigeants et personnel à l’étranger 8

5. Protéger l’assistant lui-même, et limites en toute transparence 10

6. Un guide pratique pour les personnes et les organisations 11

7. Ce que cela signifie pour vous 12
   1. Glossaire 12

© Kaia Technologies, Inc.

Sommaire

Bien des pertes liées à la criminalité en ligne ne commencent pas par une intrusion technique. Elles commencent par une personne que l'on persuade : de répondre, d'approuver, d'ouvrir, de transférer de l'argent. L'ingénierie sociale fonctionne parce qu'elle vise le jugement plutôt que le logiciel, et parce qu'elle emprunte les canaux mêmes que les gens utilisent pour le travail et la famille. Elle atteint le proche qui vit seul, la personne qui traite les factures des fournisseurs, le cadre en voyage et le fonctionnaire en poste à l'étranger.

Beaucoup d'outils répondent par un verdict ponctuel sur un seul élément, ou ne font rien avant que la perte survienne. KAIA-OR propose autre chose : **un second regard avant que quiconque réponde, paie ou ouvre quoi que ce soit**.

- **Collez un message, ou interrogez Hey KAIA à voix haute,** et voyez ou entendez les signaux d'alerte que KAIA-OR reconnaît, ainsi qu'une prochaine étape suggérée.
- **Vérifiez un lien sans l'ouvrir.**
- **Voyez ce qui semble anormal, et les cas où la vérification ne peut pas se prononcer.** La prochaine étape suggérée consiste à confirmer par un canal auquel vous faites déjà confiance. La décision vous revient.
- **Dans votre propre langue.** La vérification de message reconnaît les signaux d'alerte en anglais, en français, en turc et en hébreu.
- **Pour chaque personne.** Chaque adulte d'une famille, et chaque membre d'une organisation, dispose de son propre second regard sur son propre appareil. L'adhésion ne donne jamais accès aux messages d'un autre membre.

Ce livre blanc présente l'ampleur du problème, le fonctionnement des leurres modernes décrit de façon abstraite, les raisons pour lesquelles les verdicts ponctuels ne suffisent pas, le fonctionnement du modèle du second regard pour les familles et les organisations, la façon dont l'assistant lui-même est protégé, les limites de cette approche en toute transparence, ainsi qu'un court guide pratique pour les personnes et les organisations.

> **À PROPOS DE CE LIVRE BLANC**
>
> Ce livre blanc décrit la conception et la portée engagée du produit KAIA-OR. La disponibilité des fonctionnalités dépend du forfait et du déploiement. Pour éviter d'enseigner ou d'imiter de vrais leurres, il ne reproduit jamais de texte frauduleux, de vraies marques ni d'exemples d'adresses Web.

© Kaia Technologies, Inc.

1. Le problème et le fonctionnement des leurres

L'ampleur du problème

Les rapports publics de plus d'un pays décrivent la même tendance : les pertes sont importantes, elles augmentent, et la persuasion, plutôt que l'intrusion technique, est au cœur du phénomène. Les chiffres ci-dessous servent de contexte du marché. Ils décrivent des pertes et des incidents déclarés, et non l'effet d'un produit.

CHIFFRES PUBLICS SUR LA FRAUDE ET L'INGÉNIERIE SOCIALE

Sujet	Ce qui a été déclaré
Pertes déclarées liées à la criminalité sur Internet, États-Unis	Des pertes déclarées de 20,877 milliards de dollars US en 2025, en hausse de 26 % par rapport à 2024.[1]
Type de plainte le plus signalé, États-Unis	L'hameçonnage et l'usurpation d'identité, avec 191 561 plaintes en 2025.[1]
Pertes déclarées des personnes de 60 ans et plus, États-Unis	Des pertes de 7,748 milliards de dollars US en 2025, en hausse de 59 % par rapport à 2024, dans 201 266 plaintes, elles-mêmes en hausse de 37 %.[1]
Pertes déclarées liées à la compromission de courriels d'affaires, États-Unis	Environ 3,05 milliards de dollars US de pertes déclarées en 2025.[1]
Pertes déclarées liées à la fraude visant les consommateurs, États-Unis	Les consommateurs ont fait 3 millions de signalements de fraude en 2025 et déclaré 15,9 milliards de dollars US de pertes.[2]
Atteintes et attaques déclarées par les entreprises du Royaume-Uni	43 % des entreprises du Royaume-Uni ont déclaré une atteinte à la sécurité ou une attaque au cours des 12 mois précédents, et 85 % de ces entreprises ont subi de l'hameçonnage.[3]

Deux tendances se dégagent. Les personnes âgées subissent une part disproportionnée des pertes déclarées, souvent loin des proches qui les auraient aidées. Et les organisations sont souvent atteintes d'abord par un message qui persuade quelqu'un d'agir.

Le fonctionnement des leurres modernes, de façon abstraite

Les leurres changent constamment de formulation. Leur structure change beaucoup moins. Les schémas ci-dessous sont décrits de façon abstraite, sans reproduire aucun message réel.

- **La précipitation.** Le message ne laisse pas le temps de réfléchir, avec une échéance ou une conséquence qui semble imminente.
- **L'autorité.** Il semble venir d'une personne ou d'une entité qui a du pouvoir sur vous : un gestionnaire, une autorité publique, une institution financière, un fournisseur de services.
- **La familiarité.** Il semble venir de quelqu'un que vous connaissez : un proche en difficulté, un collègue, un fournisseur de longue date.

© Kaia Technologies, Inc.

- **De l’argent vers une nouvelle destination.** Une demande d’envoyer de l’argent vers une nouvelle destination, ou de changer la destination d’un montant habituel.
- **Des demandes d’identifiants de connexion.** Une demande de communiquer des identifiants de connexion, ou d’approuver une connexion que vous n’avez pas lancée.
- **Un lien ou un fichier.** Une page qui imite une page réelle, ou une pièce jointe qui demande à être ouverte.
- **Un changement de canal.** Le passage à un nouveau numéro, à une conversation privée ou à un appel téléphonique, loin des canaux où la demande pourrait être vérifiée.
- **Des voix et des visages synthétiques.** Du son ou de la vidéo générés qui imitent une personne que la cible connaît.

La plupart des leurres combinent deux ou trois de ces éléments. Leur but commun est d’éloigner la personne du moment où elle exerce son jugement. Une défense utile ne dépend donc pas de la reconnaissance de chaque message. Elle doit donner aux gens un moyen fiable de revenir à ce moment, avant de répondre, de payer ou d’ouvrir quoi que ce soit.

© Kaia Technologies, Inc.

2. Le modèle du second regard

Pourquoi les verdicts ponctuels ne suffisent pas

La réponse habituelle à un message suspect est un outil qui rend un seul verdict sur un seul élément. Cette approche présente trois faiblesses.

- **Un faux sentiment de sécurité.** Aucun outil de vérification ne reconnaît tous les leurres. Un verdict favorable sur un nouveau leurre est perçu comme une permission, et c'est exactement ce qu'un leurre bien conçu cherche à obtenir.
- **Aucun contexte.** Un verdict sur un seul élément collé ne sait rien des voyages de la personne, des événements qui l'entourent ni de la langue dans laquelle elle pense.
- **Aucune aide à la décision.** Un verdict dit « mauvais » ou « correct », mais pas quoi faire ensuite, alors que c'est à la prochaine étape que l'issue se décide.

Le fonctionnement du second regard

KAIA-OR remplace le verdict par un second regard. La personne garde la main; le produit l'aide à voir ce qui aurait pu lui échapper.

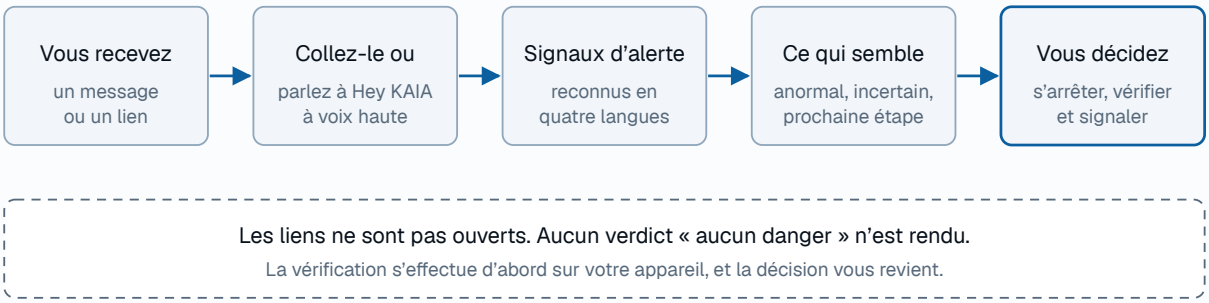


Figure 1. Le modèle du second regard. La figure est conceptuelle et ne décrit pas une mise en œuvre.

- **Collez-le, ou dites-le.** La personne colle un message dans KAIA-OR, ou interroge Hey KAIA à voix haute, ce qui aide celles qui préfèrent parler plutôt qu'écrire.
- **Voyez les signaux d'alerte.** KAIA-OR montre les signaux d'alerte qu'il reconnaît, comme une demande d'identifiants de connexion ou une demande d'argent, un changement de canal, ou un expéditeur qui ne correspond pas à l'identité annoncée.
- **Les liens ne sont pas ouverts.** Une personne peut vérifier un lien sans l'ouvrir.
- **Voyez ce qui semble anormal, et les cas où la vérification ne peut pas se prononcer.** Lorsque les données probantes ne suffisent pas, KAIA-OR le dit clairement au lieu de deviner.
- **Une prochaine étape suggérée.** Habituellement, confirmer par un canal auquel vous faites déjà confiance : un numéro que vous avez déjà, une adresse que vous connaissez, un collègue à qui vous pouvez poser la question en personne. La décision vous revient.

© Kaia Technologies, Inc.

Les vérifications s’effectuent d’abord sur l’appareil. Les messages et les liens qu’une personne vérifie restent sur son appareil, et rien n’est transmis à un service d’IA infonuagique sans son accord préalable.

Le contexte compte

Un leurre est le plus convaincant lorsqu’il colle au moment. Un message sur un trajet perturbé ne se lit pas de la même façon pendant que vous voyagez. Un message sur une urgence locale ne se lit pas de la même façon lorsqu’une urgence est signalée près de vous. Comme le second regard se trouve dans le même produit que l’image des voyages et des événements propre à chaque personne, celle-ci peut comparer ce qu’affirme un message avec ce que montre son image des événements, y compris le degré de corroboration de chaque événement. Le contexte ne décide pas à sa place. Il lui donne un élément de plus à peser avant d’agir.

© Kaia Technologies, Inc.

3. Familles et organisations

Les familles et les proches âgés

La personne la plus exposée est souvent celle qui risque le moins de demander de l'aide à temps : un parent âgé qui vit dans une autre ville, un proche gêné d'avouer qu'il a un doute, un adulte pour qui l'anglais est une langue seconde. KAIA-OR est conçu autour de leur situation plutôt que de celle de l'acheteur.

- **Chaque adulte a sa propre vérification.** Chaque adulte que vous invitez, même s'il ne vit pas sous votre toit, obtient son propre KAIA-OR et son propre second regard, sur son propre appareil.
- **Dans sa propre langue, de vive voix.** Lorsqu'un message se fait passer pour un proche, une institution financière ou un service de livraison, chaque membre peut le coller, ou interroger Hey KAIA à voix haute, et entendre les signaux d'alerte et une prochaine étape suggérée, en anglais, en français, en turc ou en hébreu.
- **Personne ne lit les messages des autres.** L'adhésion ne donne jamais accès aux messages, aux appareils, à la position ni aux lieux enregistrés d'un autre membre. L'aide vient du second regard propre à chaque personne, et non de quelqu'un qui regarde par-dessus son épaule.

Les familles peuvent renforcer cette protection par une simple entente convenue à l'avance : une demande d'argent inattendue de la part d'un proche se vérifie en rappelant ce proche à un numéro déjà enregistré, jamais à un numéro fourni dans la demande.

Organisations : fraude par fausse facture, usurpation de l'identité de dirigeants et personnel à l'étranger

Pour les organisations, les leurres les plus dommageables visent le moment où quelqu'un a le pouvoir de transférer de l'argent ou des données.

- **Fraude par fausse facture et fraude au fournisseur.** Un message qui semble venir d'un fournisseur connu annonce de nouvelles instructions de paiement, ou une facture arrive, presque identique à la facture habituelle, mais pas tout à fait.
- **Usurpation de l'identité de dirigeants.** Un message qui semble venir d'un membre de la haute direction demande un virement inhabituel, un document confidentiel ou une faveur qui contourne le processus habituel. C'est le cœur de la compromission de courriels d'affaires.
- **Personnel à l'étranger.** Le personnel en déplacement reçoit des messages qui exploitent son voyage, la situation locale ou l'éloignement des collègues qui pourraient confirmer une demande.

KAIA-OR donne à chaque membre d'une organisation son propre second regard avant tout paiement. Une instruction de fournisseur modifiée ou une demande de virement inhabituelle peut être collée pour voir les signaux d'alerte et une prochaine étape suggérée. Les liens ne sont pas ouverts. Chaque personne en déplacement obtient son propre second regard, en anglais, en français, en turc ou en hébreu, en plus de ce qui a changé dans les conseils officiels aux voyageurs pour son propre voyage.

L'organisation garde le contrôle de l'adhésion sans regarder dans les appareils des gens. Les rôles gèrent l'adhésion, pas les personnes. Les invitations sont à usage unique et liées à un compte, et les revues des accès sont datées. L'organisation voit ses sites, et non l'endroit où se trouvent ses membres, et elle ne lit pas les messages que ses membres vérifient.

© Kaia Technologies, Inc.

LE PROCESSUS L'EMPORTE SUR LA PERSUASION

La défense la plus solide d'une organisation contre la fraude par fausse facture et l'usurpation de l'identité de dirigeants est un processus qu'aucun message ne peut contourner : tout changement de destination des fonds est confirmé auprès d'une personne-ressource déjà inscrite au dossier, et les virements importants ou inhabituels exigent l'intervention d'une deuxième personne. Un second regard appuie ce processus; il ne le remplace pas.

© Kaia Technologies, Inc.

4. Protéger l’assistant lui-même, et limites en toute transparence

L’ingénierie sociale visant l’IA. À mesure que les assistants lisent une part croissante du contenu que les gens reçoivent, les attaquants cherchent à persuader l’assistant plutôt que la personne. Un message, une page ou un document peut contenir des instructions dissimulées destinées à amener un assistant à ignorer ses règles, à révéler de l’information ou à poser un geste. Hey KAIA est conçu pour y résister. Le contenu qu’on lui demande de vérifier est traité comme des données à examiner, jamais comme des instructions à suivre. Les résultats produits par le modèle n’accordent aucune autorité : rien de ce que produit un modèle d’IA ne peut approuver une action, modifier une autorisation ou élargir ce qu’un agent peut faire. Les réponses sont vérifiées au regard des données probantes avant que vous les voyiez, et elles les citent. Aucune conception ne peut exclure toute manipulation; le dernier rempart est donc le même qu’ailleurs dans KAIA-OR : la décision revient aux personnes.

Limites, en toute transparence. Un second regard est utile parce qu’il est franc sur ce qu’il est.

- **Des indicateurs, pas des verdicts.** KAIA-OR montre des signaux d’alerte et suggère une prochaine étape. Il ne déclare jamais qu’un message est inoffensif, et l’absence de signaux d’alerte ne veut pas dire qu’il n’y a aucun danger.
- **Il vérifie ce que vous lui donnez.** Il n’intercepte pas les messages ni les liens avant que vous les voyiez, et il ne prétend pas reconnaître toutes les fraudes.
- **Ce n’est pas un antivirus.** Les vérifications de fichiers et de logiciels espions fournissent uniquement des indicateurs de risque, jamais un verdict.
- **Lorsqu’il ne peut pas se prononcer, il le dit.** Certains messages ne peuvent pas être évalués à partir de leur seul contenu. KAIA-OR le signale alors, et la prochaine étape suggérée consiste à confirmer par un canal auquel vous faites déjà confiance.
- **Il ne peut pas réparer une perte.** Le second regard intervient avant qu’une personne réponde, paie ou ouvre quoi que ce soit. KAIA-OR ne peut pas récupérer l’argent déjà envoyé.

© Kaia Technologies, Inc.

5. Un guide pratique pour les personnes et les organisations

Trois étapes couvrent la plupart des situations : s'arrêter, vérifier par un canal auquel vous faites déjà confiance, et signaler. Il s'agit uniquement de prochaines étapes en matière de cybersécurité. Pour les risques naturels et les événements de sécurité publique, suivez les consignes de l'émetteur.

S'ARRÊTER, VÉRIFIER, SIGNALER

Étape	Pour une personne ou une famille	Pour une organisation
S'arrêter	Considérez une demande qui ne laisse pas le temps de réfléchir comme un signal d'alerte en soi. Ne répondez pas, ne payez pas et n'ouvrez rien tant que vous avez un doute.	Faites en sorte qu'il soit normal, et bienvenu, que le personnel s'arrête et pose des questions avant de donner suite à une demande inhabituelle, quelle qu'en soit la provenance apparente.
Vérifier	Collez le message dans KAIA-OR pour obtenir un second regard, ou interrogez Hey KAIA à voix haute. Vérifiez un lien sans l'ouvrir. Communiquez avec l'expéditeur à un numéro ou à une adresse que vous avez déjà, jamais à ceux fournis dans le message.	Confirmez tout changement de destination des fonds auprès d'une personne-ressource déjà inscrite au dossier. Exigez l'intervention d'une deuxième personne pour les virements importants ou inhabituels. Donnez à chaque membre son propre second regard, y compris au personnel à l'étranger.
Signaler	Prévenez l'organisation dont l'identité est usurpée par ses canaux officiels, et signalez la fraude aux autorités de votre pays.	Signalez la situation à votre personne-ressource en sécurité, communiquez les signaux d'alerte à vos collègues et passez régulièrement les accès en revue au moyen de revues des accès datées.

Un bon guide pratique est assez court pour qu'on s'en souvienne sous pression. Le but n'est pas de faire de chaque personne une experte en leurres. Il est de faire de l'arrêt, de la vérification par un canal de confiance et du signalement une habitude, avec un second regard à portée de main.

© Kaia Technologies, Inc.

Ce que cela signifie pour vous

Particuliers et familles. Avant de répondre, de payer ou d'ouvrir quoi que ce soit, collez le message ou interrogez Hey KAIA à voix haute, et voyez les signaux d'alerte et une prochaine étape suggérée dans votre langue. Chaque adulte de votre famille garde son propre KAIA-OR.

Équipes et petites entreprises. Un second regard avant tout paiement : une instruction de fournisseur modifiée ou une demande de virement inhabituelle peut d'abord être vérifiée, sans que les liens soient ouverts.

Équipes de sécurité des entreprises et des gouvernements. Chaque membre du personnel, y compris le personnel à l'étranger, dispose de son propre second regard en quatre langues. Les résultats produits par le modèle n'accordent aucune autorité, et l'organisation voit ses sites, et non l'endroit où se trouvent ses membres. Pour évaluer KAIA-OR, demandez un accès anticipé ou choisissez Nous joindre sur le site Web de KAIA-OR.

Glossaire

Terme	Sens dans ce livre blanc
Ingénierie sociale	Le fait de persuader une personne d'agir contre son propre intérêt, par exemple en répondant, en approuvant, en ouvrant quelque chose ou en transférant de l'argent.
Compromission de courriels d'affaires	Fraude dans laquelle un message usurpe l'identité d'un membre de la direction, d'un collègue ou d'un fournisseur pour détourner de l'argent ou obtenir des données.
Second regard	La vérification d'un message ou d'un lien par KAIA-OR : les signaux d'alerte qu'il reconnaît, ce sur quoi il ne peut pas se prononcer et une prochaine étape suggérée.
Signal d'alerte	Une caractéristique qu'ont souvent en commun les messages servant de leurres, comme une demande d'identifiants de connexion ou une demande d'argent.
Instructions dissimulées	Texte placé dans un contenu pour manipuler un assistant d'IA. KAIA-OR traite ce contenu comme des données, jamais comme des instructions.

Sources

1. FBI Internet Crime Complaint Center (IC3), 2025 Internet Crime Report, avril 2026. Contexte du marché uniquement.
2. Federal Trade Commission (FTC), FTC Testifies before the Joint Economic Committee on Agency's Efforts to Combat Fraud, mars 2026. Contexte du marché uniquement.
3. UK Department for Science, Innovation and Technology (DSIT) with the Home Office, Cyber Security Breaches Survey 2025, avril 2025. Contexte du marché uniquement.

© Kaia Technologies, Inc.

WP-3 · Version 1.0 · 3 novembre 2026

La disponibilité des fonctionnalités dépend du forfait et du déploiement. Les fonctionnalités décrites correspondent à la portée engagée du produit KAIA-OR.