La stéganographie : l'art de cacher l'existence même du secret

Stéphane FOSSE
fosse.fr

12 janvier 2026

Résumé

Il existe deux manières de protéger un message. La cryptographie le rend illisible. La stéganographie dissimule jusqu'à son existence. Cet article retrace l'histoire complète de cette discipline, depuis les tablettes de cire d'Hérodote au V^{e} siècle avant notre ère jusqu'aux malwares contemporains utilisant des images pour exfiltrer des données. Il examine les techniques employées à chaque époque, les outils disponibles aujourd'hui, les usages légitimes par les journalistes et lanceurs d'alerte, ainsi que les détournements par les groupes APT. Une perspective d'architecte sur le problème originel que cette technologie résout : communiquer sans que cette communication soit détectable.

1 Le problème originel

Il existe deux manières de protéger un message. La première consiste à le rendre illisible pour quiconque ne possède pas la clé de déchiffrement — c'est la cryptographie. La seconde, plus subtile, consiste à dissimuler jusqu'à l'existence du message lui-même. Cette seconde voie porte un nom que peu de gens connaissent, mais dont les applications traversent notre quotidien numérique sans que nous en ayons conscience : la stéganographie.

Le terme vient du grec *steganos*, « couvert » ou « caché », et *graphein*, « écrire ». L'écriture couverte. Là où le cryptographe dit « voici un message que vous ne pouvez pas lire », le stéganographe ne dit rien du tout. Son message n'existe pas aux yeux du monde. Et c'est précisément cette invisibilité qui constitue sa force [1].

Pour comprendre cette discipline, il faut remonter loin. Très loin. Avant les ordinateurs, avant l'électricité, avant même le papier tel que nous le connaissons. Car le problème que résout la stéganographie est aussi vieux que la communication humaine elle-même : comment transmettre une information sensible quand on sait que des oreilles indiscrètes écoutent, que des yeux hostiles surveillent, que le messager lui-même peut être intercepté et fouillé ?

Au V^{e} siècle avant notre ère, la Grèce antique vit sous la menace permanente de l'Empire perse. Les cités grecques, perpétuellement en conflit les unes avec les autres, doivent pourtant s'échanger des informations stratégiques quand l'ennemi commun se fait menaçant. Le problème est simple à énoncer : les routes sont surveillées, les messagers peuvent être capturés, et un message visible — même chiffré — attire l'attention. Un rouleau de parchemin portant des caractères incompréhensibles signale immédiatement son importance. Le porteur sera torturé jusqu'à ce qu'il révèle la méthode de déchiffrement. Ou exécuté par précaution.

Hérodote, dans ses *Histoires*, rapporte deux anecdotes qui comptent parmi les plus anciennes traces écrites de stéganographie [2]. La première concerne Histiée, tyran de Milet, retenu prisonnier à Suse par le roi perse Darius. Histiée voulait faire parvenir un message à son gendre Aristagoras pour l'inciter à la révolte contre les Perses. La solution qu'il trouva témoigne d'une ingéniosité brutale : il fit raser la tête de son esclave le plus fidèle, tatoua le message sur son crâne, attendit que les cheveux repoussent, puis envoya l'esclave vers Milet avec pour seule instruction de se faire raser la tête à son arrivée. Le messager traversa les lignes perses sans encombre. Personne ne cherche de message sur un homme qui n'en porte visiblement aucun.

La seconde anecdote met en scène Démarate, roi spartiate en exil à la cour perse. Apprenant que Xerxès préparait une invasion de la Grèce, il voulut prévenir ses compatriotes. Les tablettes de cire utilisées pour écrire à l'époque posaient un problème évident : n'importe quel garde pouvait les lire. Démarate gratta la cire d'une tablette, grava son message directement sur le bois, puis recouvrit le tout d'une nouvelle couche de cire vierge. La tablette semblait inutilisée. Elle passa les contrôles sans difficulté. À Sparte, la reine Gorgô eut l'intuition de faire gratter la cire, révélant l'avertissement qui permit aux Grecs de se préparer à l'invasion.

Ces deux récits illustrent le cœur du problème et sa solution. Le problème : la surveillance rend dangereuse toute communication visible. La solution : rendre la communication invisible. Ne pas la chiffrer, la cacher. La meilleure sécurité n'est pas un message illisible, c'est un message dont personne ne soupçonne l'existence.

2 L'évolution des techniques antiques et médiévales

Pendant les siècles qui suivirent, les techniques de stéganographie évoluèrent au rythme des supports disponibles et des menaces à contourner [3]. Les Romains utilisaient les œufs durs : en écrivant sur la coquille avec un mélange de vinaigre et d'alun, le message traversait la coquille et se fixait sur le blanc, invisible tant que l'œuf n'était pas écalé. Énée le Tacticien, stratège grec du IV^{e} siècle avant notre ère, décrivait une technique consistant à percer de minuscules trous au-dessus de certaines lettres dans un texte anodin, les lettres marquées formant le vrai message.

Les encres invisibles apparurent très tôt. Le jus de citron, le lait, l'urine même — tout liquide devenant visible sous l'effet de la chaleur fut exploité. Pline l'Ancien mentionne le latex de certaines plantes. Au Moyen Âge, les alchimistes perfectionnèrent ces recettes. L'acide gallotannique, extrait des galles de chêne, permettait d'écrire un message révélable uniquement par application de sulfate de cuivre. Durant la Seconde Guerre mondiale, des espions allemands utilisèrent une variante : un message écrit au sulfate de cuivre sur un mouchoir ne devenait lisible qu'exposé aux vapeurs d'ammoniaque.

Mais la technique qui fascina le plus les services de renseignement du XX^{e} siècle fut le micropoint. Inventée par les Allemands pendant la Première Guerre mondiale et perfectionnée pour la Seconde, elle consistait à photographier un document et à réduire l'image jusqu'à la taille d'un point typographique — le point final d'une phrase dans une lettre ordinaire. J. Edgar Hoover, directeur du FBI, qualifia le micropoint de « chef-d'œuvre de l'espionnage ennemi ». Une seule page de texte pouvait être réduite à un point d'un millimètre de diamètre, collé dans une correspondance banale. L'œil nu ne voyait qu'une lettre personnelle parlant de la météo et de la santé de tante Gertrude. Un microscope révélait des plans de fortifications.

Ces techniques partageaient une caractéristique commune : elles exploitaient les limites de la perception humaine. L'œil ne voit pas un message tatoué sous des cheveux. Il ne distingue pas un micropoint d'un vrai point. Il ne soupçonne pas une tablette de cire vierge. Le cerveau humain, confronté à un objet apparemment ordinaire, ne cherche pas de message caché. C'est cette confiance dans l'apparence que la stéganographie exploite.

3 La naissance du mot

Le terme « stéganographie » lui-même n'apparut qu'à la fin du XV^{e} siècle, sous la plume d'un moine bénédictin allemand dont l'œuvre allait marquer durablement l'histoire de la cryptologie : Johannes Trithemius [4]. Abbé de Sponheim puis de Würzburg, conseiller de l'empereur Maximilien, Trithemius était un personnage complexe, à la fois humaniste, historiographe, réformateur monastique et passionné de savoirs occultes.

Vers 1499, il commença la rédaction d'un ouvrage en trois volumes qu'il intitula *Steganographia*. Le livre présentait en apparence un système de magie angélique permettant de communiquer à distance par l'intermédiaire d'esprits. Les pages étaient remplies de conjurations, de noms d'anges, de tables astrologiques. Trithemius lui-même hésita à le publier, détruisant certaines parties qu'il jugeait trop dangereuses. L'ouvrage circula sous forme manuscrite pendant plus d'un siècle avant d'être finalement imprimé à Francfort en 1606, quarante-cinq ans après la mort de son auteur.

La réaction fut immédiate et violente. L'Église catholique inscrivit le livre à l'Index des ouvrages interdits en 1609. Protestants et catholiques s'accordèrent pour y voir un traité de magie noire, une invocation de démons. Cette réputation sulfureuse perdura pendant près de quatre siècles. Ce n'est qu'en 1996 que le cryptologue américain Jim Reeds démontra définitivement ce que certains soupçonnaient depuis longtemps : les formules magiques du troisième livre n'étaient pas de la magie, mais des *cover text* — des messages chiffrés dissimulant d'autres techniques de chiffrement [5]. L'apparence démoniaque du livre était elle-même une stéganographie. Trithemius avait caché un traité de cryptographie sous les atours d'un grimoire, sachant que les curieux chercheraient des sorts là où se trouvaient des chiffres.

L'ironie est savoureuse. Le premier ouvrage portant le mot « stéganographie » était lui-même stéganographié. Trithemius pratiquait ce qu'il prêchait. David Kahn, dans son monumental *The Codebreakers* publié en 1967, considère Trithemius comme l'un des fondateurs de la cryptologie moderne, aux côtés de Leon Battista Alberti [6]. L'abbé allemand avait inventé le terme qui donnerait son nom à une discipline entière, tout en démontrant par l'exemple sa puissance.

Dans la lignée de Trithemius, d'autres techniques de dissimulation textuelle se développèrent. Le « chiffre de Bacon », inventé par Francis Bacon en 1605, permettait de cacher un message dans un texte apparemment innocent en utilisant deux polices de caractères légèrement différentes. Les lecteurs non avertis voyaient une prose banale. Ceux qui savaient distinguaient les deux polices et reconstituaient le vrai message. Giovanni Battista della Porta, contemporain de Trithemius, proposa quant à lui d'utiliser les positions de mots ou de lettres dans un texte couverture pour encoder une information secrète.

4 Le tournant numérique

L'avènement de l'informatique transforma radicalement la stéganographie [7]. Non pas dans son principe — cacher un message dans un support d'apparence innocente — mais dans ses moyens et son échelle. Les fichiers numériques présentaient des caractéristiques que les supports physiques n'avaient jamais eues : une redondance massive et une facilité de modification imperceptible.

Pour comprendre ce tournant, il faut saisir ce qu'est un fichier numérique. Une image, par exemple, n'est qu'une longue suite de nombres représentant la couleur de chaque pixel. Une photographie de bonne qualité peut contenir des millions de pixels, chacun défini par trois valeurs (rouge, vert, bleu) codées généralement sur huit bits. Huit bits permettent 256 niveaux par couleur. L'œil humain, lui, ne distingue pas la différence entre un rouge codé 200 et un rouge codé 201. Cette limitation perceptive ouvre une brèche considérable.

La technique dite du « bit de poids faible » — *Least Significant Bit* ou LSB en anglais — exploite précisément cette brèche [8]. Elle consiste à remplacer le dernier bit de chaque octet de couleur par un bit du message secret. Visuellement, l'image reste identique. Statistiquement, elle a été modifiée. Mais la modification est si minime qu'elle se confond avec le bruit naturel de tout capteur photographique. Une image de huit mégapixels peut ainsi dissimuler un mégaoctet de données — de quoi stocker un roman entier ou les dix premiers chapitres de *Lolita* de Nabokov, comme le démontra un chercheur de Kaspersky dans une présentation technique [9].

Cette technique LSB, apparue dans les années 1990 avec la démocratisation des images numériques, reste aujourd'hui la plus répandue. Sa simplicité la rend accessible à n'importe quel programmeur débutant. Son efficacité la rend redoutable. Mais elle n'est pas la seule. Les images JPEG, compressées différemment des formats non compressés, nécessitent des approches spécifiques agissant dans le domaine fréquentiel via la transformée en cosinus discret (DCT). Les fichiers audio permettent de cacher des données dans les fréquences inaudibles ou dans les variations d'amplitude imperceptibles. Les vidéos combinent les deux. Les protocoles réseau eux-mêmes peuvent être détournés pour faire transiter de l'information dans des champs normalement inutilisés ou dans les variations temporelles entre paquets de données. Krzysztof Szczypiorski proposa en 2003 le terme de « stéganographie réseau » pour désigner cette dernière catégorie.

La puissance de ces techniques numériques tient à un facteur que les stéganographes des siècles passés n'auraient pas imaginé : le volume. Chaque jour, des milliards d'images sont publiées sur les réseaux sociaux, échangées par messagerie, téléchargées depuis des sites web. Chaque image est un vecteur potentiel. Retrouver une aiguille dans une botte de foin est difficile. Retrouver une aiguille dans un océan de bottes de foin est impossible si l'on ne sait même pas qu'il existe une aiguille.

5 Les usages légitimes

Avant d'aborder les dérives, il convient de rappeler que la stéganographie sert d'abord des causes légitimes [10]. Le tatouage numérique — *digital watermarking* en anglais — en est l'application la plus répandue [11]. Quand un photographe vend une image, il peut y insérer une signature invisible identifiant la transaction. Si l'image se retrouve sur un site pirate, l'analyse révélera quel client a rompu son contrat de licence. Les studios de cinéma utilisent des procédés similaires pour tracer les copies-écran de leurs films. Chaque projection en avant-première contient un marquage unique permettant, en cas de fuite, d'identifier la salle fautive. Netflix et d'autres plateformes de streaming appliquent des techniques comparables pour leurs contenus exclusifs.

La protection des lanceurs d'alerte et des journalistes représente un autre usage légitime. Dans les pays où la surveillance électronique est omniprésente, où le simple fait d'utiliser un logiciel de chiffrement peut attirer l'attention des autorités, la stéganographie offre une alternative. Un journaliste peut envoyer une photographie de vacances à un contact, photographie contenant en réalité un article compromettant pour le régime en place. Le Guardian Project, organisation à but non lucratif dédiée à la protection des militants et reporters, a développé Pixelknot, une application Android permettant de cacher des messages dans des images partagées sur les réseaux sociaux. L'image circule au vu de tous. Seul le destinataire sait qu'elle contient autre chose qu'un coucher de soleil.

Cette capacité à communiquer sous le radar des systèmes de surveillance intéresse naturellement les défenseurs des libertés civiles. Car la cryptographie, aussi robuste soit-elle, a une faiblesse structurelle : elle signale sa propre présence. Un fichier chiffré dit « je cache quelque chose ». Dans certaines juridictions, refuser de fournir une clé de déchiffrement constitue un délit. Au Royaume-Uni, le *Regulation of Investigatory Powers Act* prévoit des peines de prison pour quiconque refuse de déchiffrer des données sur demande des autorités. La stéganographie contourne ce problème. Comment exiger le déchiffrement d'un message dont on ignore l'existence ?

6 Les usages malveillants

Cette même invisibilité séduit évidemment des acteurs moins recommandables [12]. Les chercheurs en sécurité informatique observent depuis le début des années 2010 une utilisation croissante de la stéganographie par les groupes de cyberespionnage et les développeurs de logiciels malveillants [13]. Les raisons sont simples à comprendre : les antivirus et les pare-feu analysent le contenu des fichiers à la recherche de signatures connues de malwares. Un code malveillant dissimulé dans une image échappe à cette détection. L'image semble saine. Elle traverse les défenses. Une fois sur la machine cible, un programme extracteur récupère le code caché et l'exécute.

Le cas Duqu, découvert en 2011, fit date [9]. Ce malware sophistiqué, considéré comme un cousin de Stuxnet, utilisait des images JPEG pour exfiltrer les données volées. Les informations sensibles étaient chiffrées puis dissimulées dans des fichiers image apparemment anodins, que le malware téléchargeait vers des serveurs de contrôle. Les analystes de Kaspersky, qui étudièrent Duqu en détail, notèrent que cette technique rendait l'exfiltration quasiment indétectable par les outils de surveillance réseau traditionnels.

Depuis, la liste n'a cessé de s'allonger. Le groupe APT Turla, attribué aux services de renseignement russes, cacha des instructions pour ses implants dans les commentaires de photos publiées sur des comptes Instagram contrôlés. Le malware vérifiant régulièrement ces comptes récupérait ses ordres sans jamais établir de connexion suspecte vers un serveur de commande identifiable. APT32, aussi connu sous le nom OceanLotus et attribué au Vietnam, utilisait des images BMP contenant des charges utiles chiffrées. Le groupe Platinum ciblait des entités diplomatiques avec des documents Office contenant des images stéganographiées. LokiBot, un voleur d'identifiants largement diffusé, stockait ses fichiers de configuration dans des images JPEG.

En 2022, des chercheurs de Symantec (aujourd'hui Broadcom) découvrirent une campagne du groupe Witchetty — une émanation du collectif TA410 lié à APT10/Cicada — ciblant des gouvernements du Moyen-Orient [14]. L'attaque utilisait une image représentant l'ancien logo Windows, hébergée sur GitHub, pour dissimuler une porte dérobée. Le code malveillant était décodé par une opération XOR triviale mais suffisante pour tromper les outils de détection. L'image avait été téléchargée des milliers de fois sans que personne ne soupçonne sa vraie nature.

Ces exemples illustrent une tendance lourde. Les équipes de Kaspersky identifient trois raisons principales à cette adoption croissante de la stéganographie par les attaquants. Premièrement, elle masque non seulement le contenu mais le fait même qu'une communication a lieu. Deuxièmement, elle contourne les systèmes de *deep packets inspection* (DPI) déployés par les entreprises et les États. Troisièmement, les outils de sécurité ne peuvent pas analyser toutes les images transitant sur un réseau — il y en a trop, et le coût computationnel serait prohibitif.

7 La question terroriste

Au lendemain des attentats du 11 septembre 2001, une hypothèse fit les gros titres : les terroristes d'Al-Qaïda auraient utilisé la stéganographie pour coordonner leurs actions. Des articles de presse, notamment dans USA Today, affirmèrent que des messages codés avaient été dissimulés dans des images pornographiques postées sur des sites grand public. L'idée fascinait autant qu'elle effrayait. Les conspirateurs auraient communiqué au vu de tous, sous le nez des services de renseignement, en cachant leurs plans dans des photographies téléchargeables par n'importe qui.

Cette hypothèse mérite un examen critique. Maura Conway, chercheuse à Dublin City University, publia en 2003 une analyse rigoureuse intitulée « Code wars : Steganography, signals intelligence, and terrorism » [15]. Ses conclusions tempèrent considérablement le sensationnalisme médiatique. Conway souligne d'abord que les preuves tangibles de l'utilisation de stéganographie numérique par Al-Qaïda sont minces, voire inexistantes. Les affirmations de USA Today reposaient sur des sources anonymes et n'ont jamais été confirmées par des documents judiciaires ou des analyses techniques publiées.

Plus fondamentalement, Conway argumente que l'utilisation de stéganographie numérique par des organisations terroristes est à la fois techniquement complexe et opérationnellement risquée. Elle exige des compétences informatiques que tous les membres d'un réseau clandestin ne possèdent pas. Elle nécessite une discipline stricte dans le choix des images-support et des outils utilisés. Et surtout, elle impose aux destinataires de savoir où chercher les messages — problème non trivial quand les canaux de communication sont eux-mêmes surveillés.

Des chercheurs de l'Université du Michigan tentèrent en 2001 de vérifier empiriquement l'hypothèse médiatique. Ils analysèrent plus de deux millions d'images récupérées sur des sites populaires comme eBay, à la recherche de traces stéganographiques. Leur conclusion : aucune preuve de communication cachée à grande échelle. Cela ne signifie pas que la stéganographie n'a jamais été utilisée par des terroristes — prouver un négatif est impossible. Mais cela suggère que le phénomène, s'il existe, est marginal plutôt que systématique.

En revanche, des documents découverts lors de l'arrestation d'un militant autrichien lié à Al-Qaïda en 2011 à Berlin contenaient des manuels de formation dissimulés par stéganographie dans des fichiers vidéo. Ces PDF en allemand, anglais et arabe n'étaient pas chiffrés, simplement cachés. Les services allemands mirent plusieurs

semaines à les extraire. Cet exemple isolé prouve que la technique est connue dans ces milieux, sans pour autant démontrer un usage généralisé.

La prudence s'impose donc. La stéganographie est un outil. Comme tout outil, elle peut servir des desseins variés. Mais la tentation de l'invoquer pour expliquer des communications terroristes « indétectables » relève parfois davantage du fantasme technologique que de l'analyse rigoureuse.

8 La stéganalyse

Face à ces usages — légitimes ou non — une discipline miroir s'est développée : la stéganalyse [16]. Son objectif est double. D'abord, déterminer si un fichier contient ou non un message caché. Ensuite, si possible, extraire ce message. La première tâche est la plus difficile. Prouver l'absence est toujours plus ardu que prouver la présence.

Les premières méthodes de stéganalyse reposaient sur des analyses statistiques. La technique LSB, par exemple, modifie légèrement la distribution des valeurs de pixels dans une image. Un histogramme des couleurs peut révéler des anomalies caractéristiques. La méthode dite « attaque RS » (pour Regular-Singular), développée par Jessica Fridrich et ses collègues à l'Université de Binghamton, exploite les régularités introduites par l'insertion de données dans les bits de poids faible. L'attaque Chi-carré détecte les modifications en analysant les paires de valeurs adjacentes dans l'histogramme.

Ces méthodes fonctionnent raisonnablement bien contre les outils stéganographiques simples. Mais elles échouent face aux techniques adaptatives qui choisissent intelligemment où insérer les données — privilégiant les zones de l'image déjà bruitées ou texturées, où les modifications seront moins détectables. HUGO (Highly Undetectable Steganography) et UNIWARD (Universal Wavelet Relative Distortion) représentent l'état de l'art de ces approches adaptatives. Elles minimisent les altérations statistiques en concentrant l'insertion dans les régions où elle causera le moins de perturbation détectable.

La réponse des stéganalystes fut de se tourner vers l'apprentissage automatique (*machine learning*), puis vers l'apprentissage profond (*deep learning*) [17]. Les réseaux de neurones convolutifs, entraînés sur des millions d'images porteuses et non porteuses, apprennent à détecter des motifs trop subtils pour être formalisés en règles explicites. Le réseau SRNet, développé en 2018, représente une avancée significative dans cette direction. Des travaux plus récents explorent les réseaux génératifs antagonistes (GAN) pour améliorer à la fois la stéganographie et sa détection, dans une course aux armements permanente.

Car c'est bien d'une course aux armements qu'il s'agit. Chaque amélioration de la stéganalyse incite les stéganographes à perfectionner leurs techniques. Chaque nouvelle méthode de dissimulation pousse les analystes à développer de nouveaux outils de détection. Cette dynamique rappelle celle du chiffrement, où les concepteurs de codes et les briseurs de codes se poursuivent depuis des siècles. Mais la stéganographie ajoute une dimension supplémentaire : le détecteur ne sait même pas s'il cherche quelque chose. L'asymétrie est fondamentale.

9 Les outils disponibles

Quiconque souhaite expérimenter la stéganographie dispose aujourd'hui d'une palette d'outils accessibles [18], [19], [20]. Steghide, développé par Stefan Hetzl, reste une référence pour les images JPEG et BMP ainsi que pour les fichiers audio WAV. En ligne de commande, il permet de dissimuler n'importe quel fichier dans un support et de protéger l'accès par une phrase de passe. La commande est d'une simplicité déconcertante : `steghide embed -cf image.jpg -ef secret.txt`. L'extraction suit le même schéma. Steghide compresse et chiffre automatiquement les données avant insertion, offrant une double protection. Son principal inconvénient est son âge, la dernière version datant de 2003.

OpenStego, développé par Samir Vaidya, propose une interface graphique conviviale pour ceux que la ligne de commande rebute [21]. Disponible sous Windows, macOS et Linux grâce à son implémentation en Java, il gère les formats PNG, BMP, GIF et JPEG. Au-delà de la simple dissimulation de données, OpenStego offre une fonction de tatouage numérique permettant d'insérer une signature invisible dans une image pour en revendiquer ultérieurement la paternité.

Pour les fichiers audio, DeepSound se distingue en permettant de cacher des données dans des fichiers WAV et FLAC. L'interface graphique facilite l'utilisation, tandis que l'option de chiffrement AES renforce la sécurité. Outguess, plus ancien, cible spécifiquement les images JPEG en exploitant les coefficients DCT inutilisés après compression. F5, développé par Andreas Westfeld, utilise une technique de codage matriciel pour minimiser les modifications apportées à l'image JPEG.

Ces outils sont gratuits, souvent open source, et téléchargeables en quelques clics [22]. Leur existence même illustre la démocratisation de la stéganographie. Ce qui exigeait autrefois une expertise pointue est désormais à la portée de n'importe quel utilisateur motivé. Cette accessibilité a des implications ambivalentes. Elle permet

aux journalistes et aux militants de protéger leurs communications. Elle permet aussi aux acteurs malveillants d'échapper à la surveillance.

10 Le paradoxe de la visibilité

Un aspect souvent négligé de la stéganographie mérite réflexion : son efficacité repose sur le secret de son utilisation. Si tout le monde sait qu'Alice envoie des messages cachés dans ses photos de chat, ces photos seront analysées avec attention. Le simple soupçon annule une partie du bénéfice. C'est ce que le cryptologue Bruce Schneier appelle le paradoxe de la stéganographie : elle fonctionne tant qu'on ne sait pas qu'elle est utilisée.

Cette caractéristique la distingue fondamentalement de la cryptographie. Un algorithme de chiffrement peut être public, analysé par des milliers de chercheurs, et rester sûr si sa conception est solide. La sécurité repose sur la clé, pas sur le secret de la méthode. Kerckhoffs avait formulé ce principe dès 1883. La stéganographie, elle, requiert un certain degré de confidentialité sur les techniques employées. Un outil stéganographique connu et largement utilisé devient une cible prioritaire pour les stéganalystes. Ses signatures caractéristiques sont documentées. Les détecteurs sont calibrés pour le repérer.

Cela explique pourquoi les attaquants sophistiqués développent souvent leurs propres outils plutôt que d'utiliser des logiciels publics. Les implants découverts dans les campagnes APT utilisent rarement Steghide ou OpenStego. Ils implémentent des variantes sur mesure, plus difficiles à détecter car non référencées dans les bases de signatures. L'obscurité, ici, n'est pas une faiblesse mais une composante de la sécurité opérationnelle.

11 L'avenir de la discipline

L'intelligence artificielle transforme actuellement le paysage de la stéganographie comme celui de tant d'autres domaines [23]. Des chercheurs explorent l'utilisation de réseaux de neurones pour générer des images-support optimisées, conçues dès le départ pour dissimuler efficacement des données. D'autres travaillent sur des techniques dites *coverless* — sans support préexistant — où le message détermine directement la création de l'image qui le contient. L'image n'est pas modifiée pour cacher un message ; elle est générée à partir du message.

Les deepfakes, ces vidéos truquées par intelligence artificielle, ouvrent des perspectives nouvelles et inquiétantes. Une vidéo falsifiée peut servir de vecteur stéganographique à grande capacité. Qui soupçonnerait un message caché dans les artefacts de compression d'une vidéo déjà connue pour être artificielle ? La confusion devient stratégique.

Du côté de la stéganalyse, les réseaux de neurones continuent leur progression. Les modèles récents atteignent des taux de détection impressionnants sur les techniques classiques. Mais le problème du *mismatch* demeure : un détecteur entraîné sur des images d'une certaine source perd en efficacité face à des images d'une source différente. Les caractéristiques des capteurs photographiques, les chaînes de traitement, les formats de compression varient. Chaque variation perturbe les modèles. L'universalité reste un objectif distant.

La dimension juridique évolue également. Certaines législations commencent à s'intéresser spécifiquement à la stéganographie, souvent dans le sillage des lois sur le chiffrement. La question de la « perquisition » d'une image — peut-on exiger d'un suspect qu'il révèle si une image contient un message caché ? — n'a pas encore trouvé de réponse claire dans la plupart des juridictions. Le droit peine à suivre les évolutions techniques.

12 Une perspective

Pour revenir à la question initiale — quel problème cette technologie résout-elle ? — la réponse tient en une phrase : la stéganographie répond au besoin de communiquer sans que cette communication soit détectable. Ce besoin existait il y a vingt-cinq siècles quand Histiée tatouait le crâne de son esclave. Il existe toujours aujourd'hui, sous des formes différentes mais avec la même urgence pour ceux qui le ressentent.

Les moyens ont changé. Les tablettes de cire sont devenues des fichiers JPEG. Les micropoints photographiques sont devenus des modifications de bits de poids faible. Mais le principe reste identique : exploiter les limites de la perception — humaine hier, algorithmique aujourd'hui — pour faire passer un message là où personne ne pense à regarder.

La stéganographie incarne une vérité souvent oubliée dans notre monde de surveillance généralisée : ce qui n'est pas cherché n'est pas trouvé. Les systèmes de détection, aussi sophistiqués soient-ils, ne peuvent analyser que ce qu'ils sont conçus pour analyser. Une image parmi des milliards reste une image parmi des milliards. La masse devient protection. L'insignifiance devient camouflage.

Pour l'architecte de systèmes d'information, cette discipline rappelle une leçon essentielle : la sécurité n'est jamais absolue, et la menace n'est jamais là où on l'attend avec certitude. Celui qui veut vraiment cacher finira par trouver un moyen de le faire. La question n'est pas de rendre la dissimulation impossible — elle ne l'est pas.

La question est de comprendre les techniques existantes, leurs forces et leurs limites, pour adapter les défenses en conséquence. Ou, dans l'autre sens, pour protéger légitimement ce qui doit l'être contre des regards indiscrets.

Entre l'esclave au crâne tatoué et le fichier PNG téléchargé depuis GitHub, vingt-cinq siècles se sont écoulés. Le problème, lui, n'a pas changé.

Références

[1] G. C. Kessler, An Overview of Steganography for the Computer Forensics Examiner, Forensic Science Communications, vol. 6, no. 3, 2004.

[2] Wikipedia contributors, Steganography, 2026.

[3] CompTIA, The Ancient Practice of Steganography: What Is It, How Is It Used and Why Do Cybersecurity Pros Need to Understand It, 2021.

[4] Wikipedia contributors, Steganographia, 2026.

[5] J. Reeds, Solved: The Ciphers in Book III of Trithemius's Steganographia, *Cryptologia*, t. 22, n° 4, p. 291-317, 1998.

[6] D. Kahn, The Codebreakers: The Story of Secret Writing, 2e éd. Scribner, 1996, Première édition 1967, isbn : 978-0684831305.

[7] N. Subramanian, I. Cheheb, O. Elharrouss, S. Al-Maadeed et A. Bouridane, Image Steganography Approaches and Their Detection Strategies: A Survey, *ACM Computing Surveys*, t. 56, n° 10, p. 1-45, 2024.

[8] M. Hameed, A. Abdelmged, O. Zakaria et A. Youssif, Digital image steganography: A literature survey, *Neurocomputing*, t. 509, p. 73-97, 2022.

[9] Kaspersky Global Research and Analysis Team, Steganography in contemporary cyberattacks, 2017.

[10] E. Walia, P. Jain et N. Kanwal, Recent Advances in Steganography, in *Recent Advances in Steganography*, IntechOpen, 2022.

[11] Wikipedia contributors, Digital watermarking, 2026.

[12] J. Fruhlinger, Steganography explained and how to protect against it, 2021.

[13] T. Pevný, T. Filler et P. Bas, Framework for Malware Triggering Using Steganography, *Applied Sciences*, t. 12, n° 15, p. 7577, 2022.

[14] Symantec Threat Hunter Team, Witchetty: Group Uses Updated Toolset in Attacks on Governments in Middle East, 2022.

[15] M. Conway, Code wars: Steganography, signals intelligence, and terrorism, *Knowledge, Technology & Policy*, t. 16, n° 2, p. 45-62, 2003.

[16] R. Apau, E. Gyamfi et E. Akowuah, Image steganography techniques for resisting statistical steganalysis attacks: A systematic literature review, *PLOS ONE*, t. 19, n° 5, e0302765, 2024.

[17] H. Kheddar, L. Laouamer et M. Benyettou, Deep Learning for Steganalysis of Diverse Data Types: A Review of Methods, Taxonomy, Challenges and Future Directions, *Neurocomputing*, t. 573, p. 127 221, 2024. arXiv : 2308.04522.

[18] D. Lerch, Steganography Tools, 2024.

[19] 0xRick, Steganography - A list of useful tools and resources, 2019.

[20] D. Breuker, stego-toolkit, 2023.

[21] S. Vaidya, OpenStego, 2023.

[22] EC-Council, A Guide to Steganography: Meaning, Types, Tools, & Techniques, 2023.

[23] S. Chaudhary, M. Dave et A. Sanghi, A Systematic Review of Computational Image Steganography Approaches, *Archives of Computational Methods in Engineering*, t. 29, p. 4775-4797, 2022.

[24] HawkEye, Using Steganography to Hide Malware - Witchetty APT Case Study, 2022.