PROJET DE RECOMMANDATION

SUR LE DEPLOIEMENT D'UNE SOLUTION DE FILTRAGE WEB

Version soumise à consultation publique jusqu'au 30 septembre 2025

CNIL.
COMMISSION NATIONALE
INFORMATIQUE & LIBERTÉS

1. Introduction

Objet et contexte

Cette recommandation a pour objectif d'accompagner les responsables de traitement dans la mise en œuvre d'une solution de filtrage web (filtrage d'URL et détection et blocage des charges malveillantes[1]) conforme au règlement général sur la protection des données (RGPD). Elle vise également à accompagner les fournisseurs de solutions dans la mise en œuvre de bonnes pratiques concernant ces outils de sécurité.

Ces solutions de filtrage visent à répondre à des enjeux cruciaux de cybersécurité, en permettant notamment :

- de prévenir l'accès à des sites illicites ou inappropriés, qu'il s'agisse de contenus illégaux (pédopornographie, terrorisme), frauduleux et malveillants (hameçonnage) ou de sites web dont l'accès serait interdit par l'employeur ;
- d'empêcher les cyberattaques, notamment par hameçonnage ou rançongiciel, susceptibles de compromettre la sécurité des systèmes d'information des responsables de traitement.

Périmètre d'application de la recommandation

Cette recommandation s'applique aux responsables de traitement, employeurs publics ou privés, qui déploient une solution de filtrage web pour l'accès à internet professionnel de leurs employés, agents ou prestataires. Elle s'applique également dans le cadre d'un accès Wifi professionnel mis à disposition de visiteurs dans les mêmes locaux, même lorsque ces derniers ne peuvent être identifiés (consultants externes, partenaires, visiteurs du site du responsable de traitement).

Elle **ne s'applique pas** aux accès Wifi publics ouverts à tous[2], mis en place par des commerçants, des médiathèques ou d'autres organismes (publics comme privés) à destination du public.

Portée de la recommandation

Ce document détaille les recommandations de la Commission nationale de l'informatique et des libertés (CNIL) pour la mise en conformité au RGPD des responsables de traitement dans leur usage d'une solution de filtrage web. Elle n'a pas de caractère contraignant et ne préjuge pas des obligations légales applicables aux responsables de traitement qui existent par ailleurs, en particulier celles issues des textes suivants :

- Code du travail, notamment les articles L1121-1 et L2312-38 ;
- Code des postes et des communications électroniques (CPCE notamment l'article L32-3) ;

[1] Charge malveillante ou code malveillant : Tout programme développé dans le but de nuire à/ou au moyen d'un système informatique ou d'un réseau. Maliciel/logiciel malveillant.
[2] Fournir un accès internet public : quelles sont vos obligations ? https://www.cnil.fr/fr/fournir-un-acces-internet-public-quelles-obligations

- Loi n° 2000-719 du 1er août 2000 modifiant la loi no 86-1067 du 30 septembre 1986 relative à la liberté de communication (notamment les articles 43-7 et 43-9) ;
- Loi sur la Sécurité Quotidienne (LSQ 2001) ;
- Loi pour la Confiance dans l'Économie Numérique (LCEN 2004) notamment l'article 6 ;
- Loi anti-terrorisme de 2006 ;
- Directive 2006/24/CE sur la conservation de données générées ou traitées dans le cadre de la fourniture de services de communications électroniques accessibles au public ou de réseaux publics de communications ;
- Décret n° 2006-358 du 24 mars 2006 relatif à la conservation des données des communications électroniques (décret d'application de la directive 2006/24/CE).

2. Qu'est-ce qu'une passerelle web filtrante ?

Une passerelle web filtrante, souvent appelée proxy web filtrant, est un dispositif ou un service utilisé pour contrôler et surveiller les accès Internet en filtrant le contenu web selon des politiques prédéfinies. Son rôle principal est de bloquer l'accès à certains sites web ou catégories de contenus pour des raisons de sécurité et de conformité.

S'agissant des données collectées, celles-ci seront décrites dans la partie intitulée « Minimisation des données » de la présente recommandation.

De manière générale, les passerelles filtrantes répondent à deux catégories d'objectifs, à savoir :

Le filtrage en lui-même, d'une part :

- Le filtrage web pour sécuriser le système d'information (SI) du responsable de traitement (par exemple : limitation des accès aux sites web illicites, malveillants ou non conformes à la politique d'entreprise). En effet, il n'est ici nulle question d'analyse de flux sortants ni de solutions de prévention des pertes de données (solutions *DLP pour Data Loss Prevention* en anglais).
- La détection et le blocage des contenus des flux entrants en vue de sécuriser le SI contre les charges logicielles malveillantes.

La conservation des traces, d'autre part :

- La traçabilité des accès des utilisateurs aux sites web bloqués ou autorisés.

Ces deux catégories peuvent être déclinées en trois fonctionnalités précisées ci-après :

Filtrage web :

- Filtrage : bloquer ou autoriser l'accès à des sites spécifiques en se basant sur des listes de blocage (en anglais *blacklists*) ou des listes d'autorisation (en anglais *whitelists*) d'URL.
- Catégorisation : deux niveaux de catégorisation existent.
  - (1) La catégorisation préalable des sites web par l'éditeur (par exemple, commerce en ligne, voyage, réseaux sociaux, plateformes de jeux en ligne, contenu pour adultes, etc.).

- (2) l'application d'un profil à l'employé avec les catégories qui lui sont appliquées (autorisées ou interdites). Par exemple, les équipes achats ont un profil autorisant l'accès aux plateformes de vente en ligne.

Détection et blocage des flux entrants en réponse aux requêtes sortantes émises par les utilisateurs :

- Déchiffrement HTTPS : déchiffrer et inspecter le trafic HTTPS pour détecter les charges malveillantes dans le trafic web chiffré.
- Analyse antivirus/antimalware : scanner le trafic web entrant pour détecter et bloquer les logiciels malveillants, les virus et les tentatives d'hameçonnage (à savoir : redirection vers un site malveillant).

Traçabilité des accès dits de navigation web :

- Journalisation : enregistrer les requêtes web des utilisateurs pour des raisons de conformité, de sécurité (prévention d'incident ou légale).
- Génération de rapports : générer des rapports concernant les tentatives d'accès bloquées, et les menaces détectées.

3. Finalités du Traitement

La finalité du traitement est l'objectif principal de l'utilisation de données personnelles. Les données sont collectées pour un but bien déterminé et légitime. Les opérations relevant des fonctionnalités réalisées dans le cadre de la passerelle web filtrante, décrites ci-dessus, peuvent avoir pour finalités :

- le respect des obligations légales et réglementaires en matière de cybersécurité (notamment l'article 32 du RGPD) ;
- la protection des systèmes d'information contre les cybermenaces ;
- la prévention des accès à des sites web non conformes à la politique de l'organisme ou à des sites malveillants ;
- la conservation des journaux d'accès.

Les données traitées à ces fins ne doivent pas être réutilisées pour de nouvelles finalités qui seraient incompatibles avec ces objectifs initiaux.

4. Base légale du traitement

Pour pouvoir être mis en œuvre, tout traitement de données à caractère personnel doit se fonder sur l'une des bases légales prévues par le RGPD et en particulier les bases légales suivantes :

Le traitement qui vise à garantir les finalités décrites ci-dessus peut reposer sur l'**intérêt légitime** du responsable de traitement (article 6(1)(f) du RGPD).

Dans certains cas, particulièrement s'agissant de la conservation des journaux d'accès (tels que les décrets 2021-1361[3], 2021-1362[4] et 2021-1363[5]), l'**obligation légale** du responsable de traitement pourra être retenue (article 6(1)(c) du RGPD).

Il est important de souligner que la base légale de l'obligation légale n'est mobilisable que pour des textes prévoyant cette obligation de manière explicite. En particulier, les articles 5.1.f et 32 du RGPD qui prévoient une obligation de sécurité n'engendrent pas à eux seuls une obligation légale de mettre en place une passerelle web filtrante dans le cadre de traitements de données à caractère personnel.

5. Réalisation d'une AIPD

La nécessité d'une analyse d'impact relative à la protection des données (AIPD) pour une solution de passerelle web filtrante dépend de plusieurs facteurs liés aux fonctionnalités retenues et à leurs usages. D'une manière générale, il convient de se référer aux lignes directrices du CEPD pour déterminer les cas ou la réalisation d'une AIPD est obligatoire[6].

Le responsable de traitement devrait réaliser une AIPD si les traitements de la passerelle web filtrante remplissent au moins deux au regard des lignes directrices du Comité européen de la protection des données (CEPD), parmi notamment les critères[7] suivants :

- collecte de données sensibles ou données à caractère hautement personnel ;
- collecte de données personnelles à large échelle ;
- personnes vulnérables[8] ;
- surveillance systématique.

En tout état de cause, la réalisation d'une AIPD permettra de formaliser la justification des choix réalisés et, ainsi, d'identifier et traiter les risques juridiques et techniques.

[3] Décret n° 2021-1361 du 20 octobre 2021 relatif aux catégories de données conservées par les opérateurs de communications électroniques, pris en application de l'article L. 34-1 du code des postes et des communications électroniques, https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000044228877

[4] Décret n° 2021-1362 du 20 octobre 2021 relatif à la conservation des données permettant d'identifier toute personne ayant contribué à la création d'un contenu mis en ligne, pris en application du II de l'article 6 de la loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique, https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000044228912

[5] Décret n° 2021-1363 du 20 octobre 2021 portant injonction, au regard de la menace grave et actuelle contre la sécurité nationale, de conservation pour une durée d'un an de certaines catégories de données de connexion, https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000044228976

[6] Lignes directrices concernant l'analyse d'impact relative à la protection des données (AIPD) et la manière de déterminer si le traitement est « susceptible d'engendrer un risque élevé » aux fins du règlement (UE) 2016/679

[7] Critères pertinents issus des Lignes directrices concernant l'analyse d'impact relative à la protection des données (AIPD) et la manière de déterminer si le traitement est « susceptible d'engendrer un risque élevé » aux fins du règlement (UE) 2016/679, https://www.cnil.fr/sites/cnil/files/atoms/files/wp248_rev.01_fr.pdf

[8] Les lignes directrices du CEPD précisent que les employés peuvent être considérés comme des personnes concernées vulnérables en raison du déséquilibre des pouvoirs accru qui existe entre elles et le responsable du traitement (l'employeur).

6. Proportionnalité du traitement

Il incombe au responsable de traitement de s'assurer de la proportionnalité des opérations envisagées au regard des finalités poursuivies.

Cela implique une évaluation préalable de la nécessité et de l'adéquation de la solution de filtrage web, ainsi que de la solution de journalisation des accès et de la prise en compte de la protection des données dès la conception et par défaut (art.25 du RGPD).

Par ailleurs, afin de garantir la transparence et une prise en compte équilibrée des intérêts en jeu, la CNIL recommande aux responsables de traitement d'associer systématiquement les instances représentatives du personnel à ce questionnement. Une consultation préalable de ces dernières dans le cadre d'une analyse concertée de la solution de filtrage web envisagée (et de conservation des journaux d'accès) favorise l'évaluation de la proportionnalité du dispositif préalablement à sa mise en œuvre. Elle contribue également à identifier les éventuels risques et à déterminer les mesures d'atténuation appropriées (mise en place de listes blanches, modalités d'information des personnes concernées).

7. Minimisation des données

Conformément à l'article 5 du RGPD notamment, le responsable du traitement met en œuvre les mesures techniques et organisationnelles appropriées pour garantir que, par défaut, seules les données à caractère personnel qui sont nécessaires au regard de chaque finalité spécifique du traitement sont traitées.

Les données collectées doivent être strictement limitées aux éléments nécessaires aux fonctionnalités de filtrage et de journalisation. À titre d'exemple, les données suivantes peuvent être collectées et traitées :

Catégories de données collectées :

- l'identité de l'utilisateur (gestion des accès web des utilisateurs ou authentification sur le proxy web filtrant) ;
- l'adresse IP de l'utilisateur ;
- l'URL consultée, limitée au nom de domaine ;
- l'horodatage de l'accès ;
- la catégorie de site visité ;
- l'action (autorisé, bloqué, ou forcé (si cette possibilité est offerte aux utilisateurs)).

Listes blanches et déchiffrement HTTPS

Pour limiter les données collectées, les responsables de traitement devraient également configurer des listes blanches excluant le déchiffrement HTTPS sur les sites ne présentant pas de risque spécifique en vue de limiter les données personnelles qui pourraient être conservées lors de la journalisation des accès par la solution de filtrage (banques, santé, administrations publiques, etc.). À la création d'une liste blanche statique pourra s'ajouter un processus de validation et d'ajout de nouvelles URL faisant suite à une demande à l'initiative de l'employé (besoin d'ajout d'une URL spécifique nécessaire à l'exercice de ses fonctions).

Déchiffrement HTTPS

Les solutions de filtrage d'URL peuvent nécessiter le déchiffrement HTTPS pour détecter des fichiers malveillants.

Ces opérations, sensibles pour la vie privée, doivent être encadrées. En effet, la mise en place de mécanismes de déchiffrement HTTPS présente des risques dans la mesure où cette opération entraîne la rupture d'un canal sécurisé et expose des données en clair au niveau de l'équipement en charge de l'opération. Lorsqu'un tel déchiffrement est nécessaire, sa mise en œuvre doit s'accompagner de nombreuses précautions de sécurité et doit respecter certains principes de protection des données tels qu'un déchiffrement réalisé uniquement sur les domaines non listés dans les listes blanches, la non conservation des données échangées dans les requêtes HTTP. La présente recommandation ne couvre pas le cas d'analyses au-delà de la détection d'indicateurs de compromission. S'agissant des aspects techniques de sécurité, le responsable de traitement veillera à suivre les recommandations[9] de sécurité concernant l'analyse de flux HTTPS proposées par l'ANSSI.

8. Durée de conservation et journalisation

Le responsable de traitement devra définir une durée de conservation des journaux proportionnée au besoin. Une durée supérieure aux préconisations de la CNIL[10] (6 mois à 1 an) a vocation à être justifiée et documentée.

9. Information des personnes concernées

Le responsable de traitement doit informer les personnes concernées (employés ou visiteurs) des traitements mis en œuvre conformément à l'article 13 du RGPD.

S'agissant des passerelles web filtrantes, cette information pourrait notamment venir alimenter le règlement intérieur de l'organisme employeur, par exemple sous forme d'une charte informatique annexée, après un passage devant les instances représentatives du personnel.

L'information devant être individuelle, il est recommandé aux organismes de l'apporter *a minima* à chaque arrivée d'un utilisateur du SI. Cette information peut être réalisée par voie de messagerie électronique professionnelle, lorsque celle-ci est disponible, ce qui facilite par ailleurs la preuve de sa réalisation, cette preuve pouvant également se matérialiser par la signature de la charte informatique.

[9] ANSSI - Note technique - Recommandations de sécurité concernant l'analyse des flux HTTPS, https://cyber.gouv.fr/publications/recommandations-de-securite-concernant-lanalyse-des-flux-https
[10] Délibération n° 2021-122 du 14 octobre 2021 portant adoption d'une recommandation relative à la journalisation, https://www.cnil.fr/sites/cnil/files/atoms/files/recommandation_-_journalisation.pdf

Les informations à communiquer doivent comprendre l'ensemble des éléments énumérés dans l'article 13 du RGPD[11] dans un langage accessible. Cette information pourra être renouvelée régulièrement, notamment en cas de modification substantielle.

Par ailleurs et étant données les spécificités des passerelles web filtrantes, le responsable de traitement devrait s'assurer, au titre de l'article 12-1 du RGPD, de porter à la connaissance des personnes concernées non seulement l'existence du traitement de filtrage web, mais aussi des précisions sur le fonctionnement pratique du dispositif :

- les catégories de sites exclus du déchiffrement HTTPS via des listes blanches (par exemple, sites administratifs, bancaires, ou de santé) ;
- l'absence de conservation des contenus échangés (requêtes HTTP, données à caractère personnel contenues dans les URL tels que mot de passe, identifiant, etc.) ;
- les droits des personnes concernées et les modalités pour les exercer.

10. Modalités de déploiement et recommandations

Il est possible d'envisager le déploiement de la solution de passerelle web filtrante selon différentes modalités :

1. Infrastructure sous le contrôle du responsable de traitement (déploiement sur le site du responsable de traitement ou sur un cloud privé),
2. *Software as a Service (SaaS)* (déploiement sur le cloud en mode *SaaS* chez un fournisseur de services),
3. Hybride (Requêtes sur le cloud de la solution déployée en interne pour bénéficier en temps réel des dernières URL catégorisées par l'éditeur seulement dans le cas où une URL requêtée par l'employée ne serait pas présente dans la base locale). Cette solution est pertinente notamment pour obtenir les dernières URL des sites d'hameçonnage, très fréquemment dupliquées en masse.

Le déploiement d'une solution de passerelle web filtrante dans ces différentes configurations peut entraîner des différences notables en termes de protection des données et de conformité au RGPD.

En particulier, si **une solution en *SaaS*** peut présenter un certain nombre d'avantages, elle est susceptible d'entraîner des risques qu'il revient au responsable de traitement de maîtriser via des mesures adéquates.

À cet égard, la CNIL recommande en particulier de :

- vérifier et encadrer les transferts éventuels de données hors de l'UE ;
- prévoir les clauses prévues à l'article 28 pour encadrer le traitement des données par le fournisseur, agissant comme sous-traitant au sens du RGPD ;
- s'assurer de la possibilité de récupérer les logs en cas d'audit ;

[11] Plusieurs contenus ont été mis en ligne sur le site de la CNIL afin d'accompagner les responsables de traitement dans la rédaction de tels contenus tel que https://www.cnil.fr/fr/conformite-rgpd-information-des-personnes-et-transparence

- s'assurer de la sécurité du fournisseur pour prévenir le risque de violations de données (ex. fuite de logs contenant des informations personnelles des utilisateurs) ;
- vérifier que les logs collectés en local et les informations envoyées à l'éditeur respectent la minimisation des données. Notamment, en cas de déploiement d'un client sur les postes des utilisateurs (en mode SaaS, une solution de filtrage d'URL est susceptible de nécessiter l'installation d'un agent local) : mettre en place des mesures de pseudonymisation des données envoyées au sous-traitant, comme l'obfuscation des identifiants, ou des mécanismes de tokenisation des logs avant transmission au cloud ;
- clarifier dans la documentation interne les flux de données et les responsabilités entre l'entreprise et le fournisseur *SaaS*.

S'agissant d'un **déploiement hybride**, la CNIL recommande en outre de :

- sécuriser les échanges entre la solution déployée par le responsable de traitement sur son système d'information et l'éditeur de la solution *(*notamment dans le cas d'échanges nécessités par la récupération des mises à jours d'URL*)* : mise en œuvre d'un chiffrement de flux.

11. Sécurité de la solution de filtrage et de journalisation des accès

En vue de sécuriser la solution de filtrage web et la solution de journalisation des accès, le responsable de traitement devra notamment tenir compte des recommandations ci-dessous ainsi que de la recommandation journalisation[12] émise par la CNIL.

Filtrage

S'agissant du filtrage, la CNIL recommande :

- pour une solution déployée sur site, que seuls les noms de domaines des sites web bloqués et non catégorisés seront remontés à l'éditeur en vue de leur analyse, sans rattachement à la personne concernée par la tentative d'accès au site ;
- de limiter la possibilité pour le fournisseur d'identifier les utilisateurs de la solution de filtrage, cette capacité devant être autant que possible réservée au responsable de traitement lui-même. Le hachage des données des utilisateurs peut permettre de répondre à cet enjeu, notamment lorsque la solution est déployée en mode SaaS.

Journalisation

S'agissant de la journalisation, la CNIL recommande de mettre en place :

- la signature d'une clause de confidentialité stricte des administrateurs de la solution ;
- un stockage des journaux dans un environnement garantissant leur sécurité et leur confidentialité ;
- un accès aux journaux restreint à des administrateurs dûment habilités,
- une authentification multifacteur pour les accès administrateurs à la solution.

[12] Délibération n° 2021-122 du 14 octobre 2021 portant adoption d'une recommandation relative à la journalisation, https://www.cnil.fr/fr/la-cnil-publie-une-recommandation-relative-aux-mesures-de-journalisation