LES DOSSIERS TECHNIQUES

ARCHIVAGE ET SUPPRESSION DE DONNEES

Novembre 2023

CLUSIF

Table des matières

Remerciements

Le Clusif tient à mettre ici à l'honneur les personnes qui ont rendu possible la réalisation de ce document, tout particulièrement :

Le responsable du groupe de travail :

Thomas	**VAN DEN HEUVEL**	Agence de la biomédecine
Afaf	**FAFI**	Banque de France

Les contributeurs :

David	**BLONDEAU**	Croix-Rouge française
Christophe	**GIRAULT**	IRP AUTO
René	**KOUM**	TEREOS
Thierry	**MATUSIAK**	Microsoft
Laurent	**PARIS**	GROUPE DIFFUSION PLUS
Fabrice	**POLLART**	Maisons & Cités
Jules	**SARRAT**	APF France handicap

Le Clusif remercie également les adhérents ayant participé à la relecture.

1 Introduction

1.1 Description du sujet

Quel que soit leur type, les données ne peuvent être conservées indéfiniment. Par exemple, le règlement européen du 27 avril 2016[1] et la loi Informatique et Libertés introduisent un principe de conservation pour une durée limitée et de minimisation. Cette durée doit être déterminée par le responsable de traitement, et plus généralement le propriétaire des données, en fonction de la finalité ayant conduit à la collecte de ces données. Se pose alors la question du sort de ces données à différents stades de leur cycle de vie. C'est dans ce cadre que nous aborderons dans ce document le cas de l'archivage et de la suppression.

1.2 Objectifs du document

Lorsque des données sont collectées, au-delà de la finalité de leur traitement, la question de leur « fin de vie » doit ainsi se poser. Que doit-on faire de ces données une fois que l'objectif premier de leur collecte est atteint ou qu'elles n'ont plus besoin d'être conservées ?

Le présent dossier technique cherche à répondre simplement à différents cas d'usage qui peuvent être rencontrés au sein des organisations, ainsi qu'aux enjeux légaux, réglementaires et normatifs en lien avec cette « fin de vie ». Seront ainsi abordées dans la suite de ce document les notions d'archivage et de suppression des données.

La notion de donnée mentionnée dans le présent document est à prendre au sens générique du terme. Il traite **autant les données à caractère personnel que les autres types de données** (techniques, financières, anonymisées, paramètres, etc.).

1.3 À qui s'adresse le document

Le document s'adresse à tout organisme, privé ou public, amené à recueillir, exploiter et archiver des données dans le cadre de leur activité.

Les points abordés seront vus selon le prisme de plusieurs acteurs centraux et transverses : **les métiers, les DSI, le RSSI ou encore le DPD/DPO.**

1.4 Limites du document

Ce document n'a pas pour vocation de définir des durées standards ou des critères de conservation des données. Les exemples pris dans ce document ne sont que des illustrations. Chaque organisme doit définir ses propres durées en fonction de la législation ou de la réglementation applicable, et de ses contraintes, notamment liées à son domaine d'activité.

Les données traitées qui entrent dans le cadre de l'archivage public[2] (faisant partie de l'archivage définitif défini dans le présent dossier) sont régies par des règles techniques et

---

[1] Règlement (UE) 2016/679 du Parlement et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données), JOUE L119/1 du 4 mai 2016. Voir aussi https://www.cnil.fr/fr/reglement-europeen-protection-donnees

[2] Les archives publiques sont « l'ensemble des documents, y compris les données, quels que soient leur date, leur lieu de conservation, leur forme et leur support, produits ou reçus par toute personne physique ou morale et par tout service ou organisme public ou privé dans l'exercice de leur activité » (art. L. 211-1 du code du patrimoine). Il s'agit des documents produits ou reçus par un organisme public ou un organisme privé chargé de mission de service public.

juridiques précises. Les informations fournies dans ce document **ne seront pas nécessairement alignées avec les contraintes de l'archivage public**[3] (législation contraignant tous les types d'organismes chargés d'un service public, agrément pour leur gestion par des tiers, agrément SIAF[4], etc.).

Ce document n'a **pas vocation à faire une analyse juridique exhaustive des lois et réglementation**. Notamment, sous l'angle du RGPD, pour les structures publiques et privées chargées d'une mission de service public, le respect du principe de limitation de la conservation des données à caractère personnel devra s'opérer dans le respect des obligations prévues par le code du patrimoine s'agissant des archives publiques[5]. Le cas échéant, le lecteur est invité à se rapprocher de son archiviste ou du service responsable au sein de son organisation.

Il importe par ailleurs de ne pas confondre les notions de sauvegarde et d'archivage des données. Ce document ne traite pas directement de l'aspect technique de la sauvegarde. Il convient de comprendre tout au long du dossier que la **sauvegarde n'est que la mesure de sécurité ayant pour objectif d'assurer la continuité d'un traitement par restauration de l'information**.

Le document **se focalise uniquement sur les données numériques**, c'est à dire facilement administrables selon des méthodes automatisées. Même si elles ne sont pas traitées dans ce dossier, il est important de conserver à l'esprit que les données dans un autre format doivent aussi être gérées selon des règles précises, définies et applicables.

[3] https://www.cnil.fr/sites/default/files/atoms/files/guide_durees_de_conservation.pdf

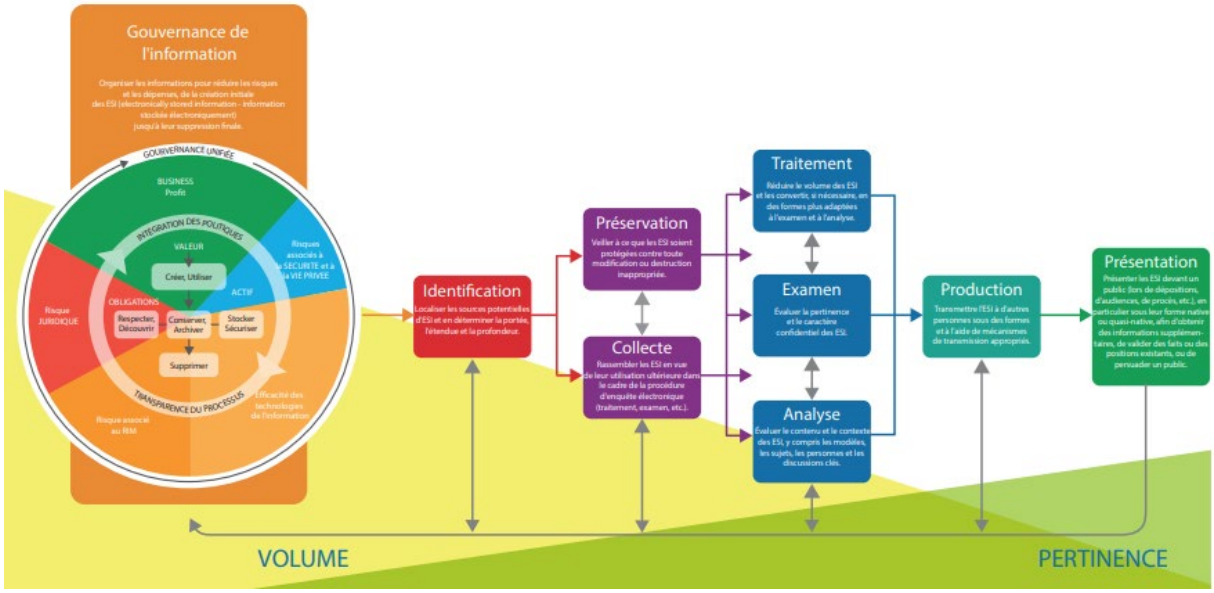
[4] https://francearchives.gouv.fr/fr/article/26287438

[5] https://www.cnil.fr/sites/default/files/atoms/files/guide_durees_de_conservation.pdf

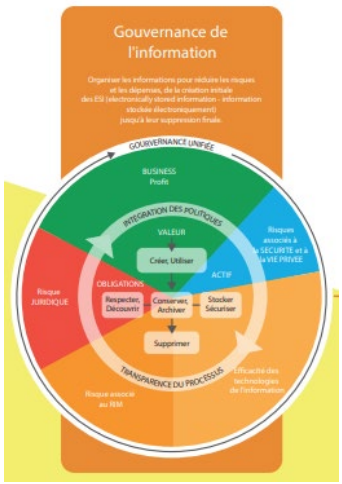
2 Définitions

2.1 Cycle de vie de la donnée

Le cycle de vie de la donnée correspond aux procédés et mécanismes **de gestion de la donnée depuis la création/collecte jusqu'à sa suppression**. Pour un même traitement, les données, à caractère personnel ou non, poursuivent des phases successives, qu'on peut regrouper en trois phases : active, intermédiaire et définitive. Il existe plusieurs modèles de référence publics sur la gouvernance des données et la gestion du cycle de vie de données. On peut citer par exemple le modèle EDRM[6] dont ce diagramme[7] – *Modèle de référence pour la découverte électronique*[8] – donne une vue synthétique.



Normes, lignes directrices et ressources pratiques pour les professionnels du droit et les spécialistes de la recherche de données électroniques (e-Discovery) - Copyright @2020, EDRM Global, Inc., Creative Commons Attribution 4.0 International, edrm.net



On s'intéresse ici en particulier à la section « Gouvernance de l'information ». Ce modèle donne une bonne vue d'ensemble de la gestion du cycle de vie de l'information, depuis sa création, son utilisation, sa protection, sa préservation jusqu'à son transfert ou sa suppression.

Il présente également **l'écosystème étendu des parties prenantes impliquées** dans ce processus : métier, IT, sécurité, protection des données, risques, légal et conformité. On voit que l'archivage, la préservation et la suppression des données y occupent une place importante, tout en étant connectés avec un écosystème plus large en charge de la gestion du cycle de vie complet des données.

[6] https://edrm.net/resources/frameworks-and-standards/edrm-model/

[7] https://edrm.net/wp-content/uploads/2020/04/EDRM-clean-poster-24x36-1.pdf

[8] Ce diagramme a été traduit en français dans le cadre de l'écriture de ce document.

2.2 Données actives

D'après la CNIL[9], il s'agit des **données nécessaires à *« l'utilisation courante »* par les services chargés de la mise en œuvre du traitement**. Il s'agit de données qui sont accessibles directement par les personnes habilitées dans les environnements de *« travail immédiat »* (voir chapitre 2.7).

2.3 Métadonnées

Pour permettre de traiter les informations et documents archivés et de garantir le respect du cycle de vie de la donnée (politiques d'archivage et de suppression), il faut prévoir d'associer **des informations complémentaires aux objets** (description, accès, recherche…). Il s'agit de métadonnées ou d'index.

Les métadonnées peuvent être classées en deux grandes catégories :

- **Gestion et technique** – par exemple : nom de fichier, format, durée de conservation… ;
- **Descriptives et fonctionnelles** (orientées métier) : indexation et type comme les factures, fiches de paie, etc.

La norme NF Z42-013:2020 (archivage électronique) propose les définitions suivantes :

- **Métadonnées** (3.41) : informations structurées ou semi-structurées permettant les traitements des documents numériques ;
- **Métadonnées de description** (3.42) : informations sur le document numérique et/ou son contenu permettant son classement et son indexation pour répondre à un besoin d'accès et de recherche et/ou contribuant à son intelligibilité ;
- **Métadonnées de gestion** (3.43) : informations concernant les attributs nécessaires à la gestion du cycle de vie des archives ainsi que les références aux documents spécifiant les caractéristiques appliquées à l'archive (politique d'archivage, règle de conservation, échéance d'élimination…) ;
- **Métadonnées techniques** (3.44) : informations concernant l'identification de l'objet numérique, son origine, son format, sa taille, sa datation, les éléments relatifs à son intégrité ainsi que les éléments relatifs au processus en amont du versement (numérisation…).

2.4 Conservation / rétention pendant la durée de vie du traitement

La notion de conservation de données peut être définie comme **le temps durant lequel la donnée est directement accessible aux personnes chargées de sa gestion courante**. Il existe donc une durée minimale et une durée maximale qui peuvent être définies sur la base de contraintes : légales ou réglementaires, métier, etc.

Il est **conseillé, voire obligatoire pour les données à caractère personnel, de définir une durée maximale de conservation**. L'article 5 du RGPD[10] définit le principe général d'*« imposer à chaque responsable de traitement de déterminer une durée de conservation des données personnelles cohérente et justifiée au regard de l'objectif de leur traitement… »*[11]. Les dates de collecte ou de dernières modifications d'une donnée peuvent être prises comme point de départ pour la durée de conservation des données actives.

---

[9] https://www.cnil.fr/fr/les-durees-de-conservation-des-donnees

[10] FAQ RGPD – Quels sont les principes du règlement ? – Clusif

[11] https://www.cnil.fr/sites/default/files/atoms/files/guide_durees_de_conservation.pdf

2.5 Préservation

La préservation renvoie à la **capacité technique d'assurer la récupération et la pérennité des informations pendant toute la durée de conservation définie**. C'est particulièrement le cas quand les durées de conservation (et/ou d'archivage) sont longues. Dans le cadre du numérique, il s'agit de se prémunir de la détérioration ou de l'obsolescence, en particulier des supports et versions de logiciels utilisés.

2.6 Archives

Il existe de nombreuses définitions d'archives. Un focus peut être fait sur deux d'entre elles qui permettent de donner une bonne compréhension de ce qu'est une archive :

- D'après le code du patrimoine (Livre II)[12] : *« Les archives sont l'ensemble des documents, y compris les données, quels que soient leur date, leur lieu de conservation, leur forme et leur support, produits ou reçus par toute personne physique ou morale et par tout service ou organisme public ou privé dans l'exercice de leur activité. »*
- D'après l'ISO 15489[13] et l'ISO 30 300[14] (en lien avec l'ISO 27001) : « *Information créée, reçue, et maintenue comme preuve, par une organisation, en cas de poursuites judiciaires, par obligation légale, à des fins de conduite des affaires. Dans un second temps, l'information archivée permet de documenter la recherche historique.* »

2.7 Archivage

L'archivage peut être vu comme **un état intermédiaire entre les bases de données actives et la suppression des données**. D'après la CNIL[15], il est possible de **distinguer deux niveaux d'archivage** dans les trois principales phases du cycle de vie de la donnée.

2.7.1 Conservation en base active

Il s'agit de la durée nécessaire à la réalisation de l'objectif (finalité du traitement) ayant justifié la collecte/l'enregistrement des données. Par exemple, dans une entreprise, les données d'un candidat non retenu seront conservées pendant 2 ans maximum (sauf s'il en demande l'effacement) par le service des ressources humaines. En pratique, les données seront alors facilement accessibles dans l'environnement de travail immédiat pour les services opérationnels chargés de ce traitement (ex. : le service des ressources humaines pour les opérations de recrutement).

2.7.2 Archivage intermédiaire

Les données personnelles ne sont plus utilisées pour atteindre l'objectif fixé (« dossiers clos »), mais présentent encore un intérêt administratif pour l'organisme (ex. : gestion d'un éventuel contentieux, etc.) ou doivent être conservées pour répondre à une obligation légale (par exemple, les données de facturation doivent être conservées dix ans en application du code de commerce, même si la personne concernée n'est plus cliente). Les données peuvent alors être consultées de manière ponctuelle et motivée par des personnes spécifiquement habilitées.

---

[12] https://www.legifrance.gouv.fr/codes/section_lc/LEGITEXT000006074236/LEGISCTA000006129161/#LEGISCTA000006129161

[13] https://www.iso.org/obp/ui/#iso:std:iso:15489:-1:ed-2:v1:fr

[14] https://www.iso.org/obp/ui/#iso:std:iso:30300:ed-2:v1:fr:sec:3.4.2

[15] https://www.cnil.fr/fr/les-durees-de-conservation-des-donnees

2.7.3 Archivage définitif

En raison de leur « valeur » et intérêt, certaines informations sont archivées de manière définitive et pérenne.

À la différence de la conservation en base active, les deux dernières étapes ne sont pas systématiquement mises en place. Leur nécessité doit être évaluée pour chaque traitement, et, pour chacune de ces phases, un tri sera opéré entre les données.

Dans le cadre de notre document, nous considérons principalement l'archivage intermédiaire.

2.8 Suppression des données

À l'autre « bout du spectre » du cycle de vie de la donnée, il **est souvent nécessaire de supprimer des données qui ne sont plus nécessaires et qui peuvent représenter un risque**. L'analyse de risque initiale, par exemple un Privacy Impact Assessment (PIA), dans le cas de données à caractère personnel, a également pu faire ressortir une durée de conservation nécessaire pour certaines données collectées, et la nécessité de les supprimer en fin de cycle.

La durée de conservation impose en conséquence, sauf indications contraires, la suppression de la donnée. Le RGPD rappelle, en miroir de la durée de conservation, qu'*« un organisme ne peut donc pas conserver des données personnelles de manière illimitée, sauf dans certains cas spécifiques et limités à ce qui est strictement nécessaire »*[16]. Il est donc **indispensable de supprimer la donnée qui n'est plus nécessaire** (échéance de sa durée de conservation ou d'archivage atteinte, par exemple).

Il faut également tenir compte de toutes les autres réglementations qui imposent aux organismes, privés comme publics, de conserver des données même si la finalité est réalisée (ce qui va imposer de les conserver dans une base intermédiaire) : déclarations fiscales, contentieux, RH avec cotisations sociales et calcul de retraites…

[16] https://www.cnil.fr/sites/default/files/atoms/files/guide_durees_de_conservation.pdf

3 Gouvernance de l'archivage et de la suppression des données

Comme toute étape en lien avec la gestion de l'information, l'archivage (et indirectement la suppression) des données doit être soumis à des **modalités de gouvernance préétablies**. Autrement dit, une définition claire des **activités et technologies que les organisations mettent en place dans le cadre de la gestion de leurs archives** doit être formalisée. L'objectif est de maximiser la valeur de leur archive, tout en minimisant les coûts et les risques associés, cela en respectant la réglementation.

Comme tout type de données, les archives doivent être sécurisées tout au long de leur cycle de vie. C'est particulièrement le cas si les données archivées sont des données sensibles ou des données qui pourraient, en cas de violation, avoir des impacts graves sur les personnes concernées, ainsi que directement sur l'organisation.

3.1 Dans quel cas conserver en base active ?

Il s'agit de la conservation courante des données par les services opérationnels en charge de leur traitement. Elle est **nécessaire à la réalisation de l'activité** – par exemple, conserver un contrat de prestation de service pendant toute la durée de la relation commerciale.

Les métiers peuvent donc effectuer différentes opérations au fil de l'eau :

- Consultation instantanée ;
- Modification ;
- Extraction, transmission, suppression, etc.

Exemple pratique : analyse et conservation d'un CV envoyé par e-mail par un candidat.

Le CV reçu est enregistré dans l'espace réseau du service recrutement de l'entreprise. L'objectif poursuivi est de permettre aux collaborateurs du service de consulter ces données instantanément au gré des besoins de recrutement, à court et moyen terme, de l'entreprise. La CNIL considère alors que ces données pourront être conservées en base active pendant 2 ans[17]. Elles pourraient aussi, pendant cette période, être transférées en archive intermédiaire dans le cas où le dossier serait clos, en particulier pour pouvoir gérer d'éventuels contentieux. Lorsque la durée de conservation est atteinte (dans notre exemple, au bout de 2 ans), les données doivent être supprimées ou anonymisées.

3.2 Dans quels cas archiver ?

Les **entreprises, organismes ou établissements publics ont l'obligation** (en application de la réglementation) **ou le besoin d'archiver** nombre d'informations très détaillées sur leurs activités passées ou en cours, en particulier au sujet des opérations effectuées avec leurs clients, fournisseurs ou salariés. Ces informations sont variées (documents internes, pièces comptables, déclarations sociales et fiscales, transactions bancaires, contrats, etc.) et peuvent comporter des données à caractère personnel. Elles sont, dès lors, protégées par les dispositions du règlement européen et de la loi du 6 janvier 1978 modifiée par la loi du 6 août 2004 relative aux fichiers, à l'informatique et aux libertés[18].

[17] https://www.cnil.fr/fr/les-durees-de-conservation-des-donnees

[18] https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000886460

On peut distinguer trois étapes liées à la gestion des données :

- Collecte/création des données ;
- Période de rétention de ces informations, pendant laquelle elles doivent être conservées ;
- Suppression effective à l'issue de cette période, si possible automatisée pour assurer la cohérence et respecter la durée de conservation définie initialement (la conservation d'une preuve de suppression peut être nécessaire pour garantir que les données ont effectivement été supprimées).

Une étape supplémentaire peut être envisagée à travers l'archivage dit « intermédiaire ». Il peut en effet être prévu d'archiver les données qui ne sont plus utilisées au quotidien, mais qu'il reste nécessaire de conserver pour d'autres raisons (obligations légales, preuve, contentieux). Par exemple, certains documents sont conservés afin d'être utilisés en cas de contentieux[19].

Toutes les données n'ont pas nécessairement à connaître cette phase d'archivage intermédiaire. Il importe de réaliser une analyse au cas par cas pour définir les données concernées, et leur durée. Une liste de plusieurs cas d'usage classiques que les organisations peuvent rencontrer dans leurs activités est proposée plus loin dans ce document. À noter que lorsque les données passent de la base active à la base d'archivage intermédiaire, elles ne peuvent être consultées par toutes les personnes initialement prévues, mais seulement par des personnes spécifiquement habilitées et ayant à les connaître en raison de leur fonction (service juridique pour la gestion des contentieux, service RH pour les cotisations sociales…).

3.3 Dans quel cas supprimer ?

Les données peuvent être amenées à être **supprimées (ou anonymisées** si celles-ci sont bien non réversibles) dans les systèmes opérationnels pour plusieurs raisons : **données temporaires, données devenues inutiles, données obsolètes**…

Dans le contexte de ce document, le cas retenu est celui dans lequel le propriétaire ou responsable de traitement a défini une durée de conservation au bout de laquelle elles doivent être effacées.

À noter que dans certains cas, le propriétaire de la donnée peut choisir de supprimer les données avant leur expiration (par exemple, des données marketing devenues inutiles et donc supprimées avant leur date d'expiration).

Dans d'autres cas, la suppression anticipée de ces données n'est pas possible (par exemple, un contrat en cours de validité).

Il faut donc envisager une **approche hybride** : l'opération de **suppression doit être programmée**, si possible automatisée, pour tenir compte des durées de conservation fixées en amont de la mise en œuvre du traitement. Les utilisateurs de ces données **peuvent également anticiper cette suppression tant qu'ils restent dans les limites décidées** par le propriétaire de la donnée et la réglementation.

Deux stratégies sont possibles pour identifier un fichier à supprimer (elle peut également s'appliquer à l'archivage s'il existe une volonté de l'automatiser) :

- Se baser sur son **emplacement** (ex : tous les fichiers partagés dans cette conversation auront une durée de conservation de 6 mois, sauf exception précisée par les participants) ;
- Analyser son **contenu** (ex. : tous les fichiers qui contiennent des numéros de sécurité sociale seront supprimés au bout de 5 ans).

[19] https://www.cnil.fr/sites/default/files/atoms/files/cnil_guide_securite_des_donnees_personnelles-2023.pdf

Le format des fichiers influence beaucoup la capacité à analyser leur contenu. Pour des fichiers texte, plusieurs options sont envisageables : patterns, liste de mots, classification sémantique, format de document, etc.

Pour des images ou des scans, la reconnaissance de caractères, la reconnaissance de forme ou la catégorisation (*clustering*) sont parfois possibles, mais restent difficiles à généraliser.

Il existe **plusieurs méthodes pour supprimer** une donnée numérique, selon la sensibilité de la donnée concernée : de la **simple suppression** d'un fichier, jusqu'à la **destruction physique et irrécupérable du support** par des procédures sécurisées. La suppression des données à caractère personnel peut parfois s'effectuer par un processus « logique » en les **rendant anonymes** pour détruire le lien entre la donnée et la personne concernée.

3.4 Précautions à prendre

Un certain nombre de précautions sont à prendre lors de la mise en place d'un système d'archivage :

- Définir **un processus de gestion des archives** : quelles données archiver, comment et où les stocker, comment gérer les métadonnées descriptives ? ;
- **Intégrer l'archivage dans le cycle de vie des données** en production ;
- Assurer une bonne **intégration dans la gestion des projets métier avec une attention particulière dans le cas de manipulation de données à caractère personnel** (Privacy by Design[20] et dans le PIA le cas échéant) ;
- Mettre en œuvre des **modalités d'accès spécifiques** aux données archivées, car l'utilisation d'une archive doit pouvoir intervenir de manière ponctuelle, exceptionnelle et justifiée ;
- Éviter les **effets de bord et préserver l'intégrité des données conservées** – par exemple, suppression de données dans les sauvegardes ou directement dans les bases actives ;
- Faire attention au format utilisé pour un système d'archivage. Il faut autant que possible **éviter des formats dits « propriétaires » (liés à un éditeur)**. En cas de changement ou disparition du logiciel, les documents seraient inexploitables… Il vaut donc mieux choisir des formats « ouverts » et non propriétaires. Il existe des formats normalisés tels que le PDF/A[21] spécialisés pour l'archivage ;
- Choisir un mode opératoire **garantissant l'intégralité de la destruction d'une archive**.

3.5 Qui est concerné ?

Les différents acteurs suivant la norme NF Z42-013 et leur relation dans le cadre de l'archivage sont :

- **Le propriétaire des archives**, comme son nom l'indique, est l'entité ayant la responsabilité juridique des archives. Selon les cas, le propriétaire peut être le service producteur, le service versant ou le service d'archives. Le propriétaire des archives se définit comme l'entreprise, l'organisation, le département, le service (RH, juridique, Direction financière…) qui va produire, recevoir, verser et/ou demander la prise en charge de ses archives par un service d'archives.

[20] https://clusif.fr/publications/privacy-by-design/

[21] https://fr.wikipedia.org/wiki/PDF/A

- Le **service d'archives**, responsable de la mise en place et de l'exploitation du système d'archivage. Le service d'archives peut être interne, externe ou mutualisé.
- Le **tiers archiveur** se définit comme le prestataire (entité) mettant à disposition un système d'archivage électronique (SAE) dont il a la responsabilité (maintien en conditions opérationnelles, hébergement, sécurité…) et pouvant offrir la gestion de ce système d'archivage électronique en tant que service d'archives si le propriétaire des archives ne souhaite pas gérer cette responsabilité.

4 L'archivage du point de vue du RSSI

Les besoins en matière de sécurité s'appliquent aussi bien aux activités d'archivage des données qu'aux données à supprimer. Une fois les données supprimées, les risques associés sont fortement limités, voire éliminés. Dans le même temps, il est indispensable de ne pas supprimer trop tôt, volontairement ou involontairement, les données dans les SI afin d'assurer le service rendu.

C'est pourquoi il est indispensable de se poser la question de l'archivage des données et de leur suppression au plus tôt dans la conception des systèmes. La mise en place de telles mesures dans le cycle de vie de la donnée participe largement de la sécurisation de l'ensemble du SI.

Pour tous les critères de sécurité (disponibilité, intégrité, confidentialité, traçabilité), le besoin dépendra du traitement, du contexte, de l'application métier ou encore de la criticité du service rendu ou des données traitées (notamment leur sensibilité). Cette réflexion devra être menée dans le cadre du processus de Privacy by Design[22]. Le cas échéant, la réalisation d'une analyse de risques Sécurité Systèmes d'Information (type EBIOS par exemple) et/ou d'un PIA permettra de définir précisément les mesures de protection adéquates.

4.1 Disponibilité

Pour la **base active**, les besoins en disponibilité **dépendent fortement du contexte dans lequel est traitée l'information**.

La particularité du besoin en disponibilité, dans le cadre de ce dossier technique sur l'archivage et la suppression, doit **être prise sous l'angle de la durée de conservation**. En effet, celle-ci doit être suffisamment définie en amont pour éviter que les données **ne soient supprimées trop tôt** (et donc que le besoin de disponibilité ne soit pas respecté). De fait, une suppression trop précoce de certaines données peut avoir un impact non négligeable pour l'organisation qui les traite (réglementaire, besoin métier, service rendu…).

Par opposition, les données ne doivent **pas être conservées trop longtemps** et doivent respecter le principe de **minimisation** pour éviter une accumulation de données qui ne sont plus utiles ou nécessaires au traitement. Si celles-ci venaient à fuiter, le préjudice pourrait être important pour les personnes ou l'organisation concernées.

Les accès à une **archive intermédiaire** sont par définition **plus rares et plus ponctuels** que pour les informations en base active.

L'enjeu autour de la disponibilité des archives intermédiaires se situe donc **autant au niveau de la disponibilité du service** (souvent lié à des obligations légales, réglementaires ponctuelles ou à un préjudice) qu'au niveau de la **pérennité du contenu**.

Dans le cas de la disponibilité, celle-ci ne doit pas nécessairement être permanente. Néanmoins, le système doit pouvoir répondre à des demandes ponctuelles avec des délais pouvant être contraints (quelques jours, voire quelques heures). Il peut être envisagé d'avoir des systèmes déconnectés qui ne sont sollicités qu'au moment où le besoin se présente (avantages de coût, de sécurité et de consommation de ressources).

Sur le long terme, les notions de pérennité du contenu et de son support sont essentielles, et prévalent notamment sur la notion de disponibilité du service.

En effet, les données nécessitent parfois d'être conservées de nombreuses années, ce qui implique un certain nombre de problématiques et d'aléas (environnementaux, cyber) qui doivent être gérés en conséquence. En parallèle, les supports informatiques évoluent

[22] https://clusif.fr/publications/privacy-by-design/

(problème de compatibilité entre plusieurs technologies de générations différentes) ou se dégradent (par exemple, les bandes de sauvegarde/archivage, les disques durs, etc. peuvent devenir illisibles avec le temps), nécessitant une attention particulière pour assurer la lisibilité des informations dans la durée.

Ces systèmes d'archivage doivent donc être pensés en tenant compte de ces contraintes.

4.2 Intégrité

L'intégrité fait partie des enjeux essentiels, si ce n'est premier, dans la protection des données archivées. Elle couvre la conservation, le traitement et l'accès aux données archivées (qui, quand, comment…). L'intérêt de l'archivage est en effet de **retrouver l'exactitude des données des années plus tard**. Il est donc nécessaire de se prémunir à la fois contre des modifications volontaires, involontaires ou malveillantes (principalement, accès illégitimes) et une altération de la donnée *(voir paragraphe précédent)*.

L'intégrité peut, au-delà de l'information portée directement par la donnée, concerner de manière plus indirecte le contexte d'accès à celle-ci (métadonnées, journaux de logs, registre), qu'il soit logique ou physique *(voir paragraphe sur la traçabilité)*. Par exemple, dans le cadre de l'accès à des archives papier ou sur bande, l'intégrité du registre d'accès (papier, biométrique, par badge…) doit être assurée.

Au-delà des cas d'usage classiques d'archive, les caractéristiques de sécurité des systèmes d'archivage (intégrité, horodatage…) peuvent être utiles pour la mise sous séquestre.

4.3 Confidentialité

Tout comme l'intégrité, il est crucial de préserver la confidentialité des données archivées. Le niveau de confidentialité **dépendra principalement du type de données archivées** et notamment de leur niveau de sensibilité (au sens du RGPD comme au sens de la classification interne de l'organisation).

4.4 Traçabilité

La traçabilité peut prendre plusieurs formes, et le système d'archivage n'échappe pas à la règle.

Les traces peuvent être :

- **techniques** (journalisation au niveau de l'infrastructure, des équipements de sécurité…) ;
- **applicatives** (authentifications, gestion des droits, erreurs…) ;
- **fonctionnelles** (consultation des actions effectuées, historisation…).

Par ailleurs, les traces jouent un rôle central dans la **gestion de la preuve**. C'est cette journalisation et son intégrité qui va permettre de **garantir la force probante des informations archivées**, ainsi que de contrôler les actions réalisées sur les systèmes hébergeant les archives.

4.5 Traitement des risques : notre « Top 5 »

Dans ce paragraphe, nous vous proposons les 5 bonnes pratiques qui peuvent contribuer à sécuriser son système d'archivage. Elles sont issues notamment de **la norme NF Z 42-013 qui peut constituer un référentiel utile pour aller plus loin dans la sécurisation de ses archives**. La CNIL, de son côté, propose des recommandations pour sécuriser les archives[23] (processus de gestion de l'archivage, gestion des accès et mesures d'intégrité).

- **Processus de gestion des archives / suppression**

  Dans le cadre du cycle de vie de la donnée, un **processus doit permettre de définir le « sort » des données** : doivent-elles être supprimées (notamment, purge à la fin de la durée de conservation définie), archivées *(voir contraintes décrites dans ce dossier)* ou anonymisées ?

  Ce processus doit également permettre de **définir les rôles et les responsabilités ainsi que les risques** dans le cadre de ces traitements, incluant les étapes de validation et de contrôle (notamment au moment de l'archivage, mais surtout au moment des demandes d'accès).

- **Accès et habilitations**

  La consultation (et donc l'accès) aux données doit être **limitée dans le temps à des personnes et à un périmètre bien définis,** car il s'agit souvent de cas de figure sensibles (par exemple : procédure judiciaire, contentieux…).

  Le système d'archivage pouvant être dissocié et indépendant du SI principal d'une organisation, il est possible que les comptes et accès soient suivis avec moins de rigueur. Une attention particulière devra donc être portée sur les **revues d'habilitation**.

- **Cloisonnement des environnements (notamment vis-à-vis de la production)**

  Le système d'archivage peut être **cloisonné logiquement** (réseau virtuel ou fonctionnellement via des clés de chiffrement au niveau du logiciel par exemple) voire **complètement isolé physiquement** du reste du SI notamment pour garantir son intégrité et sa confidentialité.

  Cette mesure permet de **limiter le risque en cas de compromission du reste du SI**, rendant plus compliquée pour un attaquant la possibilité de réaliser des mouvements « latéraux » et de compromettre le système d'archivage à son tour.

- **Solution garantissant l'intégrité**

  Considérant que l'intégrité des données est un des enjeux majeurs de l'archivage, il est nécessaire de mettre en place des **mesures garantissant leur intégrité** ainsi que celle des **métadonnées**. Ceci passe notamment par un renforcement des contrôles d'accès, mais aussi par **des technologies spécifiques** (par exemple : de grands acteurs du Cloud public proposent des offres pour « geler » la donnée), mais aussi par des solutions ou mécanismes dédiés (par exemple : checksum, audit de l'activité).

- **Externalisation chez un prestataire de confiance**

  Comme dans d'autres cas de figure, une solution peut être de **transférer le risque en externalisant vers un tiers de confiance** notamment certifié. L'organisation s'assure ainsi d'un niveau de gestion et de sécurité à l'état de l'art et respectant les bonnes pratiques en matière d'archivage.

  À noter toutefois qu'une attention particulière doit être portée à cette mesure qui implique de facto un risque inhérent qui peut être traité au travers d'une analyse de risque sécurité

[23] https://www.cnil.fr/fr/securite-archiver-de-maniere-securisee

ou d'un PIA. En effet, il est dans cette hypothèse indispensable de s'assurer que les données archivées peuvent **être récupérées à tout moment et plus particulièrement à la fin de la relation contractuelle**. Pour limiter ce risque fournisseur, il faut donc envisager, dans le cas d'une prestation avec un éditeur de récupérer le code source, et pour un tiers archiveur une clause de réversibilité effective avec des mesures de sécurité adaptées, d'annexer un Plan d'Assurance Sécurité et des clauses sur la protection des données à caractère personnel.

4.6 Normes et certifications

Plusieurs référentiels et normes décrivent les mesures de protection applicables aux systèmes d'archivage électronique (SAE) :

- **Norme française NF Z 42-013** ;
- **Certification NF 461** attribuée par l'AFNOR à un système d'archivage électronique[24] ;
- **Norme internationale ISO 14641-1** (en version 2018 au moment de la rédaction de ce dossier).

Le système d'archivage fait partie du système d'information et doit donc également respecter les bonnes pratiques de sécurité. Les activités liées à l'archivage doivent aussi s'inscrire dans une démarche d'amélioration continue.

[24] https://cdn.afnor.org/download/produits/FR/NF461.pdf

5 L'archivage du point de vue du DPO

Les activités d'archivage **portant sur des données à caractère personnel doivent, comme partie d'un traitement, être conformes au RGPD**. La suppression est quant à elle **au cœur de plusieurs principes essentiels** du règlement : durée de conservation, minimisation, droit à l'oubli… En effet, comme c'est le cas pour la cybersécurité, une fois les données à caractère personnel supprimées, les risques pour les personnes concernées sont fortement limités, voire éliminés.

C'est pourquoi il est indispensable de se poser la question de l'archivage des données et de leur suppression **au plus tôt dans la mise en place de traitements** de données à caractère personnel (notamment durée de conservation). Cette réflexion devra être menée dans le cadre du processus de Privacy by Design[25]. Le cas échéant, la réalisation d'un PIA permettra de **définir précisément les mesures de protection adéquates**. Il faudra par ailleurs se poser la question de **la compatibilité du système d'archivage avec les enjeux des droits** des personnes (accès, oubli, limitation, portabilité…).

5.1 Intégrer l'archivage dans le Privacy by Design et analyse d'impact

L'archivage est une étape dans le cycle de vie d'un traitement. Elle doit par conséquent être considérée au **moment de la conception du traitement (Privacy by Design)** et doit apparaître au **registre comme une étape du traitement**. En effet, de la même manière que le reste du traitement, les archives doivent être limitées, sélectives et sécurisées. Autrement dit, seules les informations nécessaires (**minimisation**) seront archivées du fait d'une obligation légale ou réglementaire (par exemple, l'article L. 3243-4 du Code du travail impose à l'employeur de conserver un double du bulletin de paie du salarié pendant 5 ans) ou d'un besoin métier, voire, dans certains cas, tout ou partie de l'archive devra être anonymisé. Elles devront être associées à une durée de conservation en lien avec les contraintes associées au besoin ou à la finalité.

Une attention particulière devra être portée aux **métadonnées associées aux archives** qui peuvent elles-mêmes contenir des données à caractère personnel. Les principes du RGPD s'y appliquent donc également : minimisation, anonymisation en fin de cycle de vie, etc.

Le responsable de traitement doit garantir la sécurité des données archivées en s'assurant également que les **sous-traitants offrent des garanties de sécurité équivalentes**. Les risques concernent la destruction, l'altération, la perte, la divulgation ou un accès non autorisé aux données. Si des dispositifs de suppressions automatisées sont utilisés, des contrôles devraient vérifier la destruction effective des données (en intégrant les contraintes ou évolutions technologiques ou techniques de récupération de données).

Enfin, si un traitement est soumis à un PIA, les archives doivent être intégrées à la réflexion et embarquer les mesures de protection de données identifiées lors de cette analyse.

[25] https://clusif.fr/publications/privacy-by-design/

5.2 Anonymisation

À l'issue de la durée de conservation, il faut choisir entre la suppression de la donnée ou l'anonymisation[26] de la donnée. Toute organisation fait face à des enjeux potentiellement contradictoires. Dans certains cas, elle a besoin d'archiver l'information pour la conserver, la garder à disposition et éventuellement pouvoir y accéder. En parallèle, elle peut décider d'anonymiser une partie des données archivées. Il faut alors **identifier quelles données peuvent être anonymisées, comment les anonymiser** (en préservant l'intégrité générale des données) et garder trace de cette anonymisation.

Trois stratégies sont possibles :

- **Anonymiser les données actives** (l'archive en découlant sera donc également anonyme). C'est souvent la meilleure solution pour minimiser les risques et assurer la cohérence des données sur le long terme, aussi bien pour les données actives que pour les archives.
- **Anonymiser les données au moment de l'archivage**. Cela nécessite que le processus d'archivage soit capable de gérer cette anonymisation.
- **Anonymiser les données après leur archivage**. Cela nécessite que les formats d'archivage supportent les mises à jour et représente un risque à quantifier sur la cohérence des données et la compatibilité des systèmes.

5.3 Droits des personnes concernées en lien avec l'archivage

Les droits des personnes concernées (accès, rectification, oubli, etc.) sur les données personnelles collectées et traitées par une organisation ont déjà été étudiés par le groupe de travail Protection des données du Clusif[27].

Le présent document se focalise ici sur les droits à l'accès, à la limitation, à l'oubli et à la portabilité des données personnelles, du point de vue d'archivage de celles-ci.

- **Droit à l'accès**

  Le droit d'accès par la personne dont les données ont été recueillies constitue un droit fondamental. Si la définition prévue à l'article 15.1 du RGPD ne précise pas explicitement le périmètre à inclure (on parle de l'accès « auxdites données » sans autres éléments), **la CNIL a depuis longtemps précisé que ce droit s'appliquait également aux données contenues dans les archives intermédiaires** (données présentant un intérêt administratif, en cas de contentieux par exemple) et définitives[28].

  Il convient donc de prévoir des modalités de mise en œuvre des systèmes d'archivage les plus larges possibles afin de répondre à l'exercice de ce droit. Les **données en archivage doivent donc être prises en compte dans la réponse**. À noter que la réponse au droit d'accès doit être accompagnée d'un certain nombre d'informations complémentaires, comme les finalités des traitements et les durées de conservation (décrites à l'alinéa 3 de l'article 15 du RGPD). Il sera donc pertinent d'inclure dans la réponse les données actives et archivées, ainsi que leurs durées de conservation.

---

[26] https://www.cnil.fr/fr/lanonymisation-de-donnees-personnelles

https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_fr.pdf

[27] https://clusif.fr/publications/faq-rgpd-exercice-des-droits-des-personnes/

[28] https://www.legifrance.gouv.fr/cnil/id/CNILTEXT000017651957/

- **Droit à la limitation**

  Le droit à la limitation **peut s'appliquer aussi bien sur les données actives que sur les données archivées**. Le système d'archivage doit donc le supporter.

  Le droit à la limitation peut conduire à geler les données temporairement, mais le système d'archivage n'est pas la bonne solution pour geler temporairement des données métier, sauf s'il propose cette fonction nativement. La création d'un traitement spécifique apparaît alors souvent plus adaptée.

- **Droit à l'oubli**

  Chaque entreprise peut déployer un ensemble de mesures et de processus pour respecter ce droit, comme les autres : cartographie des données collectées, capacité à retrouver ces données et à les effacer en cas de besoin, point de contact pour recueillir les demandes des utilisateurs… Dans certains cas, un **effacement direct pourrait représenter un risque pour les applications, voire être techniquement impossible** (par exemple pour conserver la cohérence d'une base de données). Le **remplacement** des données par des **valeurs anonymisées** *(voir paragraphe dédié)* peut être une solution envisageable pour l'organisation, comme une évolution des applicatifs pour implémenter un service d'oubli qui effacerait les données à un niveau fonctionnel (dont archives) et préserverait la cohérence des traitements.

  La problématique des droits des personnes et en particulier du droit à l'oubli dépasse le cadre de l'archivage et de la suppression des données. Trois dimensions liées aux archives peuvent cependant être identifiées ici :

  - La suppression régulière des données devenues inutiles participe des principes de minimisation[29] et de durée limitée de conservation des données. Elle participe également de l'implémentation du droit à l'oubli dans le traitement des données. L'industrialisation des processus de suppression régulière des données contribue dès lors au droit à l'oubli, ou en tout cas simplifie son implémentation. Elle passe principalement par l'analyse de risques qui doit être conduite, la sensibilisation des utilisateurs et l'automatisation dans les outils de la suppression des données devenues inutiles ;
  - À tous les niveaux, les archives incorporent des données personnelles. Il apparaît une contradiction potentielle entre la volonté de conserver les données et l'existence d'un droit à l'oubli pour les personnes concernées. De nombreuses technologies d'archivage ne permettent pas de modifier « finement » une archive. La résolution de ce conflit potentiel commence par une évaluation, d'un côté des données à passer de la base active à l'archive intermédiaire, et de l'autre des technologies d'archivage retenues permettant de modifier une archive existante ou de gérer une granularité suffisamment fine des données archivées ;
  - Si l'on conclut que les données archivées doivent être « oubliées », leur anonymisation reste une autre possibilité technique, dont la complexité est à mettre en parallèle de leur suppression pure et simple *(voir le paragraphe sur l'anonymisation dans ce document)*.

- **Droit à la portabilité**

  Si le traitement est basé sur un contrat ou un consentement, la personne peut dès lors demander l'exercice de son droit à la portabilité de ses données. De la même manière que pour le droit d'accès, même si la portabilité a un périmètre restreint, il faut envisager **la prise en compte des données archivées dans le droit à la portabilité**.

[29] RGPD Article 5 - Principe

6 Du point de vue métier ou DSI

Les Métiers (au sens « cœur de l'activité de l'organisme ») ont **un rôle central dans la gestion de l'archivage**. En effet, ils doivent **définir les directives** en matière d'archivage en lien avec leurs activités ou avec le **concours des directions concernées** (direction juridique ou service d'archivage), que ce soit pour des raisons légales ou réglementaires.

Pour la DSI, au-delà des bénéfices qui peuvent être **tirés des outils ou technologies de « sauvegarde » ou d'archive, il importe de garantir le fonctionnement attendu du processus d'archivage** pour permettre les opérations d'accès, de consultation ou de suppression. Les projets doivent prendre en compte, dès la phase de conception, des mesures pour identifier les catégories de données utilisées et construire un référentiel de conservation de données avec les différents acteurs de l'organisation (DPO, RSSI, Métiers) pour intégrer les opérations d'archivage ou de suppression dans les fonctionnalités des outils ou des applications.

Une cartographie des données et un référentiel de traitement sont à même de favoriser la gestion des informations : données dupliquées, échangées, version faisant « autorité », pour éviter les phénomènes de rémanence des données (restauration d'une donnée qui avait fait l'objet d'une purge) en associant des mesures de suppression sécurisée des données.

6.1 Archivage et/ou suppression pour diminution des coûts

Les données archivées sont par définition **des données qui ne nécessitent pas un accès régulier**. Les supports techniques d'archivage peuvent donc présenter des **performances moindres** (vitesse de stockage notamment). Par conséquent, leurs **coûts peuvent être tirés vers le bas** avec des tarifs inférieurs à ceux des moyens de stockage en production.

L'archivage permet ainsi une bonne planification de la capacité en production. De fait, lorsqu'une donnée n'est plus nécessaire en base active (production), elle est **soit supprimée, soit versée en archivage intermédiaire, libérant ainsi de l'espace sur les supports de stockage les plus performants** (et par conséquent plus chers).

6.2 Archivage et/ou suppression pour limiter le risque en cas de fuite de données

Il existe plusieurs manières de traiter un risque : le réduire, le transférer, l'accepter ou le supprimer. Il apparaît donc assez logique que la **mise en place d'une politique de conservation/suppression et d'archivage des données contribue à limiter les risques cyber**.

Si les données sont limitées au strict minimum, le risque en est d'autant réduit. Autrement dit, l'absence de données implique par nature une absence de risque de fuite. Par ailleurs, pour les données encore nécessaires, la politique d'archivage peut largement contribuer à diminuer l'impact du risque. Deux cas de figure existent :

- En cas d'archivage dans le SI interne de l'organisme, le **cloisonnement entre le système d'archivage et le reste du SI** peut permettre de conserver l'intégrité des données archivées en cas de compromission de la partie principale du système d'information (« latéralisation » d'une cyberattaque) ;
- En cas d'archivage chez un tiers de confiance (par exemple : tiers archiveur), **le risque est transféré** à une entité experte de l'archivage – pouvant être certifiée *(voir paragraphe « Normes et certifications »)* – avec pour obligation de garantir la sécurité des données qui lui sont confiées.

6.3 Archivage légal et réglementaire

Une organisation doit conserver tout document émis ou reçu dans l'exercice de son activité pendant une durée (minimale et maximale). Ce délai varie selon la nature des données, les finalités définies pour leur traitement et, dans certains cas, les obligations légales ou réglementaires.

Des **durées « standardisées » de conservation des documents peuvent être publiées de manière explicite et officielle**. C'est par exemple le cas de la Direction de l'information légale et administrative[30] (document civil et commercial, pièce comptable, document fiscal, document social, gestion du personnel). De son côté, la CNIL propose un guide des durées de conservation ainsi que des référentiels de durées de conservation.[31]

De manière plus spécifique, certaines durées de conservation des documents peuvent être soumises à des **contraintes légales et réglementaires sectorielles**.

Enfin, si aucune règle n'est précisée quant à la durée de conservation, celle-ci devra être définie de manière proportionnée au regard des finalités définies par le responsable du traitement.

Dans tous les cas, la politique d'archivage doit tenir compte de ces durées et ne pas conserver toutes les données en base active.

6.4 Préservation des logs

Les organismes sont régulièrement soumis à des **investigations numériques** (Forensic) en cas d'attaque informatique, fraudes ou incidents divers. La quantité de journaux accumulés et la nécessité de conserver les **logs comme valeur probante** (cas de certaines réglementations par exemple[32]) sont deux raisons pour lesquelles l'archivage des logs pour leur préservation se pose. En effet, comme vu précédemment, les coûts peuvent être réduits au niveau des systèmes d'archivage tout en garantissant un haut niveau de sécurité et notamment d'intégrité (valeur probante). Par ailleurs, la mise en place d'un politique de logs (dont leur archivage) peut permettre de réduire les responsabilités de l'organisme concerné[33].

En cas de besoin (par exemple en cas de cyberattaque), ces archives peuvent être mises à disposition des personnes chargées de l'investigation. La collecte des traces peut être réalisée par les ressources internes. Lors de la phase d'analyse des traces, celle-ci doit être réalisée de façon indépendante (ne pouvant être juge et partie – ISO 27 042[34]) avec des ressources qualifiées.

6.5 Suppressions de données structurées et non structurées

Il est souvent difficile de supprimer des données directement dans une base structurée (base de données relationnelle par exemple) sans déstabiliser les applications qui les utilisent. Il est alors nécessaire de prévoir un scénario métier de **suppression qui assurera la cohérence des opérations**. **L'anonymisation et la pseudonymisation peuvent constituer des moyens utilisés pour conserver l'intégrité et la cohérence des bases de données** en lieu

[30] https://entreprendre.service-public.fr/vosdroits/F10029

[31] https://www.cnil.fr/fr/les-durees-de-conservation-des-donnees

[32] Art L34-1 du Code des Postes et des Communications Electroniques ; art 6 Loi pour la Confiance dans l'Economie Numérique

[33] https://clusif.fr/publications/referentiel-des-logs-journalisation/

[34] https://www.iso.org/fr/standard/44406.html

et place de la suppression d'une partie des données.

Pour les données non structurées (fichiers), on peut distinguer deux scénarios :

- Considérant la suppression **d'une partie d'un fichier**, il s'agit d'un cas « structuré », car il est nécessaire de pouvoir identifier les données spécifiques à supprimer à l'intérieur dudit fichier. Cette approche granulaire doit être limitée à des besoins très particuliers, car elle peut devenir très coûteuse, techniquement complexe à mettre en œuvre et très difficile à automatiser, donc à faire passer à l'échelle. Pour des fichiers au format libre, elle peut même être pratiquement impossible à implémenter (par exemple, comment supprimer toutes les références à un client particulier dans un fichier texte libre de suivi commercial).
- Si le fichier est **considéré comme un tout « atomique »**, la cohérence est plus facile. Il est conservé ou supprimé dans son intégralité. L'enjeu est alors principalement un enjeu de cartographie / de référencement et d'automatisation de la suppression de l'information (surtout pour les traitements à grande échelle).

6.6 Suppression interdite

L'interdiction de la suppression peut être **associée à une durée de conservation « minimale »**, associée à un engagement de préservation des informations collectées : besoin métier, légal ou réglementaire… Elle aura donc une traduction en besoins de sécurité : Disponibilité, Confidentialité, Intégrité.

6.7 Suppression obligatoire

Il s'agit de la durée de **conservation « maximale » des informations**. C'est souvent une décision résultant d'une approche de **conformité au RGPD** « Privacy by Design[35] » ou du PIA s'il y a lieu. Dans le cas de données n'ayant aucun caractère personnel (informations financières, patrimoniales…), la suppression peut être perçue comme **un moyen d'optimisation** *(voir « Archivage et/ou suppression pour diminution des coûts »)* ou de réduction du risque.

Le RGPD ne définit pas les durées précises pendant lesquelles les données personnelles doivent être conservées. En revanche, plusieurs autres textes permettent de définir une durée à appliquer aux données collectées *(voir paragraphe « Archivage légal et réglementaire »)* :

- Les **dispositions légales ou réglementaires** (par exemple : Code du travail, code des postes et des communications électroniques, etc.). Certains textes imposent une durée minimale de conservation (par exemple, les données relatives aux bulletins de paie des salariés doivent être conservées au moins 5 ans en application de l'article L. 3243-4 du Code du travail) ou une durée maximale (par exemple, l'article L. 252-3 du code de la sécurité intérieure limite la durée de conservation des images de vidéo protection à un mois) ;
- Les **délibérations de la CNIL** : référentiels sectoriels de durées, « cadres de référence » de la CNIL (par exemple : référentiel relatif aux dispositifs d'alertes professionnelles) ;
- Les **références sectorielles** (par exemple : code de conduite, etc.).

Si aucune de ces sources ne permet de fixer une durée, il appartient au **responsable du traitement** de données à caractère personnel de définir cette durée. Pour cela, il devra se fonder sur la finalité (et sa base légale) pour laquelle le traitement des données personnelles est mis en œuvre, c'est-à-dire le but qu'il poursuit.

[35] https://clusif.fr/publications/privacy-by-design/

La conservation et la suppression des données personnelles doivent aussi s'inscrire dans un cadre plus large : conservation et suppression des données métier, client, industrielles, propriété intellectuelle, etc., pour assurer la cohérence des approches retenues. De plus, la qualification des données peut aussi dépendre de leur contexte d'utilisation. Par exemple, une montre connectée ou un compteur électrique rentrent dans la sphère personnelle une fois qu'ils sont associés à un propriétaire, un utilisateur ou une adresse, mais pas quand les données produites sont stockées séparément.

7 Glossaire

Acronyme	Définition
CNIL	Commission nationale de l'informatique et des libertés
DPD/DPO	Délégué à la protection des données
DSI	Directeur (ou Direction) des systèmes d'information
EDRM	Electronic Discovery Reference Model, organisme qui propose des standards, outils et guides sur des sujets tels que la protection de la vie privée, la sécurité, la gouvernance de l'information…
IT	Information Technology ou Technologie de l'information
PAS	Plan d'assurance sécurité
PIA	Privacy Impact Assessment ou EIVP en français (Étude d'impact sur la vie privée)
RGPD	Règlement général sur la protection des données
RIM	Records and Information Management, gestion des documents d'archives
RSSI	Responsable de la sécurité des systèmes d'information
SAE	Système d'archivage électronique

Tour Eria
5 rue Bellini
92821 Puteaux cedex

France

✆ +33 1 53 25 08 80

clusif@clusif.fr

clusif.fr