Commission nationale de l'informatique et des libertés

Délibération n° 2023-050 du 25 mai 2023 portant adoption d'une recommandation technique relative à l'utilisation des interfaces de programmation applicatives (API) pour le partage sécurisé de données à caractère personnel

NOR : *CNIL2316369X*

La Commission nationale de l'informatique et des libertés,

Vu le règlement (UE) 2016/679 du Parlement européen 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) ;

Vu la loi n° 78-17 du 6 janvier 1978 modifiée relative à l'informatique, aux fichiers et aux libertés, notamment son article 8-I-2°-*b* ;

Après avoir entendu le rapport M. Claude CASTELLUCCIA, commissaire, et les observations de M. Benjamin TOUZANNE, commissaire du Gouvernement,

Formule les observations suivantes :

La Commission nationale de l'informatique et des libertés (la Commission) a observé au cours des dernières années une augmentation soutenue des partages de données à caractère personnel entre organismes, qu'ils soient publics ou privés. Cette tendance, expliquée par l'intérêt croissant dans la réutilisation des données pour diverses finalités, est confirmée par le souhait du législateur de renforcer la possibilité de ces échanges entre administrations mais également entre organismes publics et privés.

La Commission souligne que ces partages de données à caractère personnel doivent être accompagnés des mesures techniques adaptées pour garantir un niveau de sécurité dès la conception et maintenu dans le temps en adéquation avec les risques, et que les données partagées doivent être limitées au strict minimum. A cet égard, elle considère que le recours aux interfaces de programmation applicatives, communément appelées « API » en référence à leur nom anglais « *application programming interface* », peut fournir un cadre technique favorable à ces partages dans de nombreux cas, sous réserve du respect de certains principes.

Article 1er

Périmètre de la recommandation

La présente recommandation vise à identifier les cas dans lesquels l'utilisation d'une API est préconisée afin de partager de manière sécurisée des données à caractère personnel ou des informations issues de leur anonymisation, et à diffuser certaines bonnes pratiques concernant leur mise en œuvre et leur utilisation. Est ici entendue par « partage de données » la faculté offerte, à certains réutilisateurs identifiés ou bien au public, de récupérer des données détenues par un organisme, ou la faculté des détenteurs de données de transmettre celles-ci à des fins de réutilisation par des tiers, lorsque cela est autorisé ou demandé par la réglementation. La présente recommandation ne présente pas de caractère obligatoire, sauf lorsqu'elle rappelle les exigences découlant du règlement général sur la protection des données (RGPD) et de la loi du 6 janvier 1978 modifiée (ci-après loi « informatique et libertés »). Cependant, le respect de ces recommandations est de nature à contribuer grandement au respect par les acteurs de leurs obligations, en particulier l'obligation de protection des données dès la conception et de protection des données par défaut prévue à l'article 25 du RGPD.

Cette recommandation identifie les acteurs les plus à même de mettre en œuvre les différentes catégories de mesures nécessaires vis-à-vis de leur rôle fonctionnel, sans préjuger de leur qualification juridique. En pratique, cette qualification juridique devra être déterminée pour chaque cas particulier sur la base des critères définis par le RGPD, afin de déterminer les responsabilités et les obligations qui en résultent (voir les précisions fournies plus bas à ce sujet). Les bonnes pratiques retenues par la CNIL sont ainsi ventilées entre les détenteurs de données, les gestionnaires d'API et les réutilisateurs. Une définition de ces trois rôles est proposée en annexe I et les paragraphes suivants détaillent les modalités de leur coordination.

En outre, les mesures précisées ne visent pas l'exhaustivité, mais ciblent les points d'attention techniques les plus importants dans la mise en œuvre d'un partage de données par voie d'API. Certaines règles et bonnes pratiques sectorielles applicables aux partages de données par voie d'API devraient également être prises en compte le cas échéant, telles que le référentiel général de sécurité (RGS) (1) dans le cas d'un partage de données impliquant une administration, ou encore les autres recommandations de la CNIL (2).

La Commission souligne que les capacités techniques liées à chacun de ces trois rôles peuvent grandement varier en pratique. Par ailleurs, plusieurs rôles peuvent être tenus par le même organisme. Cela sera le cas, par exemple, lorsque le détenteur de données développe lui-même une API : il est alors également gestionnaire d'API, et toutes les recommandations visant le détenteur de données et le gestionnaire d'API s'appliqueront alors à cet organisme. *A contrario*, plusieurs organismes peuvent avoir le même rôle. Cette situation est courante pour le rôle de réutilisateur de données ; toutefois, elle peut également exister pour les autres rôles.

Le rôle de gestionnaire d'API peut être tenu par plusieurs organismes lorsque la gestion et le développement des outils techniques sur lesquels repose le partage sont répartis entre différents intervenants. Il arrive également que seul un des trois rôles existe à un certain stade du traitement. En particulier, lorsqu'un gestionnaire d'API développe un outil technique « sur étagère », c'est à dire dans la perspective de le mettre à disposition d'autres organismes, la présente recommandation devrait être prise en compte bien que ni le détenteur de données, ni les réutilisateurs ne soient encore identifiés.

Enfin, le partage peut schématiquement être séquencé afin que chacune de ses étapes principales corresponde à l'organisation proposée.

L'annexe II présente plusieurs cas d'usages fréquemment observés, dans lesquels les rôles de chacun sont identifiés et expliqués.

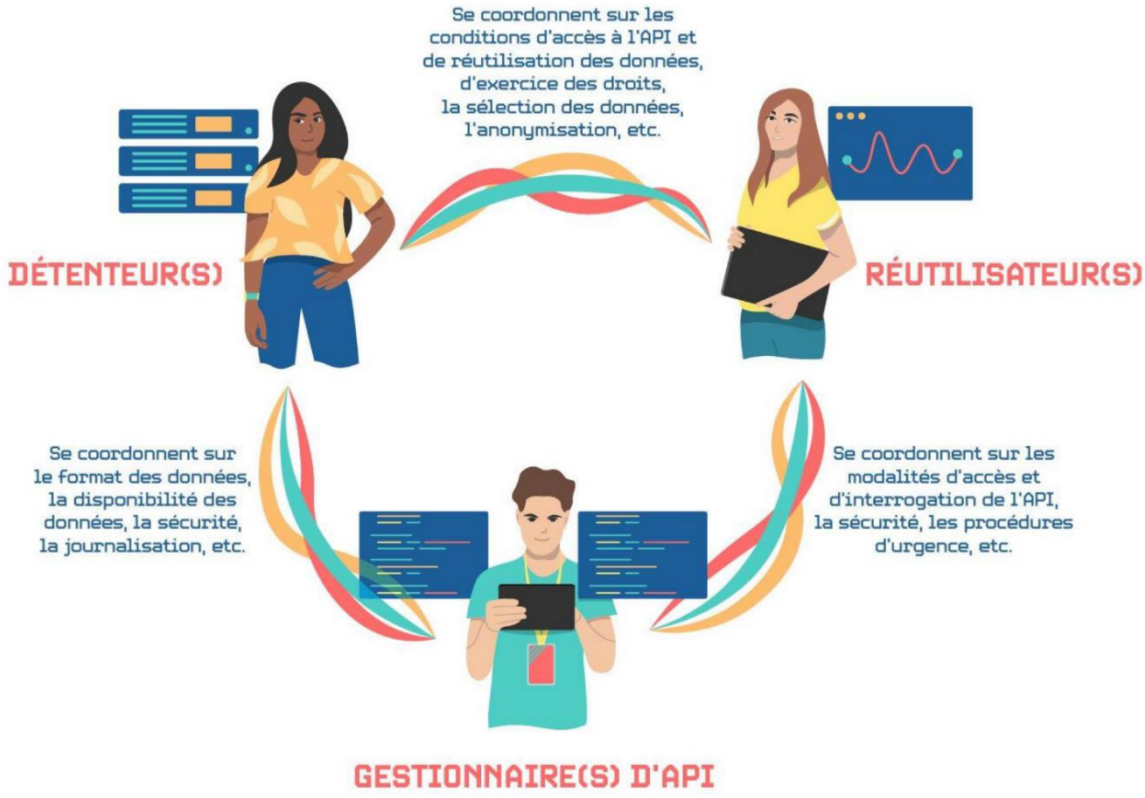

Figure 1. – Schéma relationnel entre les trois rôles de détenteur de données, de gestionnaire d'API et de réutilisateur

Un détenteur de données est caractérisé par le fait qu'il contrôle des données de manière technique ou organisationnelle. Un gestionnaire d'API est l'organisme en charge d'une partie ou de la totalité des composantes techniques sur lesquelles repose le partage de données. Enfin, le réutilisateur de données est tout organisme envisageant d'accéder ou recevant des données par voie d'API en vue de les exploiter pour son propre compte.

Ces trois rôles permettent de formuler des recommandations à l'intention des acteurs selon leur rôle technique dans le parcours des données, indépendamment de leur qualification juridique au regard du RGPD et des relations contractuelles pouvant exister entre eux, et ont de ce fait été privilégiés par rapport aux rôles habituels de fournisseur et de consommateur d'API (voir annexe I). Toutefois, l'articulation entre ces différents rôles est généralement la suivante :

- Le détenteur de données sera :
  - fournisseur d'API, lorsqu'il est également gestionnaire d'API. C'est le cas le plus courant ;
  - consommateur d'API, lorsqu'il transmet lui-même les données (ou réalise d'autres actions) via l'API ;
- Le gestionnaire d'API sera :
  - fournisseur d'API, dans la majorité des cas ;
  - ni fournisseur, ni consommateur lorsque son rôle est lié à la mise en œuvre technique du partage de données mais accessoire au fonctionnement de l'API (c'est le cas d'un site répertoriant l'API pour le compte du détenteur de données, par exemple) ;
- Le réutilisateur sera :
  - consommateur d'API, dans la majorité des cas ;

– fournisseur d'API, lorsqu'il est également gestionnaire d'API et reçoit les données des détenteurs par des requêtes en écriture.

Les API présentent une grande variété. La plupart des API ne sont ouvertes qu'à certaines personnes autorisées à accéder aux données (API en accès restreint), tandis que d'autres sont un moyen technique permettant de mettre des données à la disposition du public, et sont accessibles à tous (API ouvertes). Certaines API prévoient des requêtes permettant aux réutilisateurs d'accéder aux données (requêtes en lecture, définies en annexe I), alors que d'autres prévoient des requêtes permettant aux détenteurs de données de partager activement leurs données (requêtes en écriture, également définies en annexe I). Cette recommandation et les trois rôles précédents s'appliquent de façon identique à ces différentes situations.

S'agissant de la qualification juridique des acteurs, il peut être relevé que le détenteur de données sera généralement responsable du traitement de partage, dans la mesure où il aura librement décidé des finalités et des moyens du traitement ou lorsqu'il aura été contraint légalement de le mettre en œuvre ; par voie de conséquence, les recommandations exposées ci-dessous le concerneront pleinement. Il pourra aussi, en cas de détermination conjointe des finalités et moyens du traitement, en être « responsable conjoint » avec un ou plusieurs des autres acteurs, en particulier avec le réutilisateur si celui-ci a, en droit ou en pratique, exercé une influence déterminante sur les objectifs et conditions de mise en œuvre du traitement en cause. Toutefois, la qualification du réutilisateur correspondra souvent, et simplement, à celle de « destinataire » au sens du RGPD, sans préjudice de sa responsabilité à l'égard du traitement qu'il mettra en œuvre pour son propre compte dans la cadre de la réutilisation des données partagées. De son côté, le gestionnaire de l'API agira généralement en tant que sous-traitant, c'est-à-dire pour le compte et sous les instructions du détenteur de données et/ou du réutilisateur. Il n'est toutefois pas exclu que le gestionnaire de l'API mette en œuvre le traitement de partage pour son propre compte, en qualité de responsable de traitement, notamment lorsqu'il agit en tant qu'intermédiaire.

Il ressort de ce qui précède qu'une analyse au cas par cas est indispensable pour définir la qualification juridique des acteurs, et ainsi déterminer sur quel(s) acteur(s), à quel titre et dans quelle mesure, pèse la responsabilité de tenir compte des recommandations, ou obligations le cas échéant, exposées ci-dessous. La présente recommandation formule des « bonnes pratiques fonctionnelles » concernant chacun des trois rôles. Chaque recommandation est attribuée à l'acteur techniquement le plus à même de les mettre en œuvre, sans préjudice de la répartition juridique des responsabilités résultant du RGPD ou d'autres textes réglementaires.

Article 2

Recommandations générales

Sur les motifs justifiant le recours à une API :

L'utilisation d'API est à favoriser lorsque :

– les données sont fréquemment mises à jour, et/ou les réutilisateurs ont besoin d'y accéder régulièrement ; et/ou
– leur stockage par le réutilisateur n'est pas utile (utilisation unique ou traitement en continu sans nécessité d'historique) ; et/ou
– le ou les réutilisateurs n'ont pas systématiquement besoin d'accéder à l'intégralité des données mais seulement à un sous-ensemble des données non identifiable à l'avance ; et/ou
– les méthodes utilisées pour garantir leur sécurité sont susceptibles d'être mises à jour.

Dans tous les cas de la liste précédente, la Commission recommande l'utilisation d'une API pour le partage de données à caractère personnel. Elle le recommande d'autant plus si les données sont partagées à de nombreux réutilisateurs, voire mises à disposition du public. Elle déconseille en principe, dans ces cas, le recours à une plateforme de partage de données ou à un service de communication électronique (tels que définis en annexe I). Dans les autres cas, l'utilisation d'une API peut également être pertinente, mais son opportunité doit être évaluée en comparaison avec les autres techniques de partage de données possibles.

En effet, la Commission a observé que le niveau de sécurité apporté par les API est généralement plus élevé que celui relatif à ces autres méthodes, notamment en ce qui concerne la sécurité des communications par messagerie électronique ou encore la gouvernance des données une fois que celles-ci ont été partagées par messagerie électronique ou sur une plateforme de partage, ou que ces modes d'échange nécessitaient la transmission de grandes quantités de données, parfois inutilement.

Par conséquent, le partage par le biais d'une API permet une meilleure supervision du partage des données, d'une part en contrôlant les accès, le degré de précision des données transmises et, le cas échéant, les finalités d'utilisation des données et, d'autre part, grâce à la mise en place d'une interface d'échange standardisée entre détenteur, gestionnaire et réutilisateur, en permettant la transmission sécurisée d'informations associées à l'échange de données (durée de conservation, gestion de l'exercice des droits et notamment du droit à la portabilité, etc.).

Sur les risques liés à l'utilisation d'une API :

En facilitant et en automatisant le partage de données, l'utilisation d'API présente toutefois des risques accrus de détournement de finalité, de perte de confidentialité et de contrôle des données ou encore concernant la

transparence vis-à-vis des personnes concernées, qu'il est nécessaire de prendre en compte. Les objectifs suivants devraient être considérés comme prioritaires lors de la mise en œuvre de mesures visant à réduire les risques :

– la minimisation des données échangées ;
– l'exactitude des données source ;
– la traçabilité des accès ;
– la gouvernance et le respect des droits ;
– la sécurité.

Les objectifs précédents sont à considérer dans le contexte du partage de données, selon la vraisemblance et la gravité des risques associés. Les facteurs suivants, que la Commission considère comme particulièrement importants dans le cadre d'un partage de données par voie d'API, devraient être pris en compte :

– type d'accès à la base de données : en lecture seule ou en écriture ;
– délivrance des autorisations et conditions d'accès aux données : si l'accès est soumis à une autorisation, quelles vérifications sont apportées afin de valider ces demandes ? ;
– niveau de sécurité des techniques d'authentification utilisées ;
– nature des organismes impliqués dans le partage (maturité technique, gouvernance européenne ou non, capacités opérationnelles, etc.) ;
– autres mesures techniques (physiques ou logiques) et organisationnelles prévues pour améliorer le niveau de sécurité du système ;
– état des connaissances sur les techniques utilisées et les risques associés ;
– catégories de données accessibles par l'API : certaines données sensibles au sens de l'article 9 du RGPD ou encore des données à caractère hautement personnel (tels que des données bancaires ou des données de géolocalisation) sont plus susceptibles de faire l'objet d'attaques ou d'entraîner des conséquences plus graves ;
– degré de précision des données et des requêtes : possibilité d'accéder uniquement à certaines informations ciblées.

Cette liste de facteurs devrait être considérée par chacun des acteurs concernés par le partage de données, quel que soit leur rôle. Une documentation recensant les choix faits doit être constituée et conservée. Lorsqu'elles ne relèvent pas d'une obligation légale, les recommandations préconisées dans le reste de ce document seront à considérer au cas par cas, selon le niveau de risque déterminé et selon les moyens techniques susceptibles d'être mis en œuvre par l'organisme.

Lorsque le traitement de partage mis en œuvre est susceptible d'engendrer un risque élevé pour les droits et libertés des personnes concernées, cette analyse technique devra être intégrée à une analyse plus large, appelée analyse d'impact relative à la protection des données, conformément aux dispositions de l'article 35 du RGPD et aux lignes directrices en la matière (3).

Sur la coordination fonctionnelle des organismes :

Bien que la présente recommandation préconise des bonnes pratiques correspondant à chacun des rôles fonctionnels précédemment exposés (détenteur de données, gestionnaire d'API et réutilisateur des données), la mise en place d'une gouvernance efficace entre ces trois types d'acteurs représente un enjeu majeur pour le respect des grands principes « informatique et libertés ».

En particulier, et quelle que soit la répartition des responsabilités résultant des qualifications juridiques des acteurs précédemment évoquée, il est recommandé que les organismes concernés par le traitement se coordonnent sur les modalités de mise en œuvre de la présente recommandation. Cette coordination fonctionnelle devrait être formalisée sous la forme d'une documentation, définissant d'une manière générale les rôles et responsabilités de chaque acteur, et précisant les procédures mises en œuvre dans le cadre de cette recommandation afin de s'assurer de leur prise en compte par chacun des acteurs. Cette documentation, qui n'est pas nécessairement spécifique à chaque partage, devrait toutefois être suffisamment précise et complète pour prévoir et encadrer chaque cas d'usage.

Sur l'information des personnes :

Les organismes impliqués dans le partage de données devraient se coordonner pour fournir une information claire et complète aux personnes, concernant le traitement de mise à disposition de leurs données à caractère personnel. La Commission recommande que les mesures de traçabilité permises par les API, et en particulier la journalisation des accès et des actions, soient utilisées afin de collecter des informations statistiques concernant l'utilisation des données. Ces informations pourront être agrégées grâce à un procédé automatisé dont la précision devrait être adaptée notamment à la gravité des conséquences pour les personnes que pourraient entraîner une utilisation détournée de l'API ou un accès illégitime aux données. Lorsqu'une tentative d'accès illégitime aux données est particulièrement vraisemblable et qu'elle pourrait entraîner des conséquences particulièrement graves, comme cela est le cas pour certaines API accessibles par FranceConnect, les informations fournies à la personne devraient inclure la liste exhaustive des accès aux données sur la période pertinente, ainsi que leur horodatage afin de permettre à la personne concernée d'identifier un accès illégitime, en complément des autres mesures de sécurité mises en œuvre par le responsable de traitement. Dans le cas général, la Commission souligne que la liste des réutilisateurs devra être portée à la connaissance des personnes concernées, lorsqu'elle est connue. Elle considère également comme une bonne pratique le fait de rendre publics ou de fournir aux personnes, pour chacune des API

et éventuellement selon plusieurs niveaux d'information dont le premier serait compréhensible pour le grand public :

– une description détaillée des données partagées, de leur fréquence d'échantillonnage, des opérations réalisées en amont de leur partage, telles que des processus de pseudonymisation ou d'anonymisation ;
– des informations concernant les accès aux données réalisés via l'API, telles que leur fréquence, leur volume ou leur profondeur historique ;
– les objectifs de sécurité visés.

Ces informations devraient être mises à jour au gré de leur évolution, de manière automatisée lorsque le dispositif le permet. Les modifications devraient être portées à la connaissance des personnes concernées de manière individuelle lorsqu'elles changent substantiellement la nature du partage, notamment lorsque les données sont partagées pour une nouvelle finalité ou lorsque les restrictions d'accès les concernant évoluent de façon significative.

La Commission recommande que les informations précédemment citées soient *a minima* présentées sur un site web porté à l'attention des personnes ou qu'elles leur soient communiquées directement, sous un format accessible et compréhensible. Lorsque cela est pertinent au regard du niveau de risque évalué, en particulier au regard du besoin de confidentialité des données concernées, ainsi que de la criticité du traitement, l'accès à ces informations devra être sécurisé au moyen d'une méthode d'authentification conforme aux recommandations de la Commission, telle que FranceConnect dans la sphère publique.

Sur la sélection des données :

Pour assurer le respect des textes en vigueur encadrant d'ores et déjà le partage de données, et en application du principe de minimisation, la Commission encourage le détenteur de données à mener une réflexion avec les réutilisateurs des données sur les données strictement nécessaires aux réutilisations que chacun d'eux prévoit, pour limiter le partage à ces seules données dans le respect des éventuels textes encadrant le partage, en particulier lorsque celui-ci implique des organismes publics. A cet égard, elle souligne les éléments suivants.

Les catégories de données, leur format, leur profondeur historique, leur précision, leur fréquence d'échantillonnage, leur fréquence de mise à jour et les mesures de pseudonymisation ou d'anonymisation qui leur sont appliquées, devront notamment être choisies pour répondre strictement aux besoins relatifs aux réutilisations prévues.

Le détenteur de données devrait poursuivre cette réflexion avec les réutilisateurs après l'ouverture des accès, afin de recueillir leurs retours, notamment lorsque ces derniers n'ont pas pu être inclus dans la réflexion initiale.

Les données concernées par le partage devraient être fréquemment recensées, afin d'identifier celles dont le partage ne serait plus pertinent et d'y mettre fin. Le format des données choisi devrait être pérenne, univoque et documenté pour limiter les risques liés à une erreur d'interprétation humaine ou logicielle.

Afin de garantir que les données partagées par l'API soient au format attendu, l'utilisation d'un outil de validation des données est recommandée.

L'infrastructure technique, le format des données et les modalités d'interrogation de l'API, telles que le niveau de précision autorisé dans les requêtes, devraient être choisis pour respecter au mieux les recommandations précédentes et éviter le partage de données non pertinentes vis-à-vis de chacun des réutilisateurs, qu'il s'agisse des organismes ayant à connaître des données ou des personnes physiques appartenant à l'organisme réutilisateur lorsque plusieurs niveaux d'accès sont prévus au sein de cet organisme selon le niveau d'habilitation des personnes. Ces mesures devraient être intégrées directement dans l'API.

Enfin, lorsque les catégories de données pertinentes ne peuvent pas être précisément identifiées en amont du traitement, comme cela peut être le cas en matière de recherche, la Commission recommande qu'une phase d'expérimentation préalable soit mise en œuvre afin de vérifier la pertinence de certaines catégories sur des volumes restreints de données. Cette phase d'expérimentation, menée en coopération avec le gestionnaire d'API et les réutilisateurs, devrait permettre de confirmer la pertinence des choix faits et le respect des recommandations énoncées dans les paragraphes précédents. Cette phase d'expérimentation devrait utiliser l'infrastructure technique prévue en conditions réelles dans une version « bac à sable » et se limiter, autant que possible, à des données fictives ou altérées.

Sur l'exercice des droits sur le partage des données :

Les droits des personnes concernées par un traitement consistant à partager leurs données au moyen d'une API recouvrent :

– le droit d'accès, de rectification, d'effacement et de portabilité sur la base source utilisée par l'API ;
– le droit à l'information sur les partages opérés ;
– le droit à l'opposition ou au retrait du consentement au partage, ainsi que le droit à la limitation du traitement.

Recourir à une API permet d'automatiser un certain nombre d'opérations de traitement et de faciliter les demandes d'exercice des droits sur le partage, participant à leur prise en compte effective. Il est recommandé de limiter les opérations manuelles relatives au traitement de ces demandes.

Lorsqu'une personne concernée par le partage retire son consentement, exerce son droit d'opposition ou à la limitation, l'API devrait intégrer un dispositif technique permettant d'exclure automatiquement du champ du partage les données concernées.

Dans certains cas, comme lorsque la temporalité du traitement justifie une réponse rapide à une demande d'exercice des droits ou lorsqu'une erreur dans les données pourrait avoir des conséquences graves pour les personnes, l'API devrait également intégrer un dispositif spécifique permettant au détenteur de données d'informer chaque réutilisateur auquel les données ont été communiquées, de toute rectification, effacement de données ou limitation de traitement faisant suite à l'exercice des droits par les personnes concernées, cette information pouvant être obligatoire dans certains cas en vertu de l'article 19 du RGPD. De manière plus générale, ce dispositif devrait aussi permettre au détenteur d'informer activement les réutilisateurs sur les éventuelles restrictions aux réutilisations, qui pourraient notamment résulter de l'exercice des droits (p. ex. : opposition à certaines finalités de réutilisation). Il appartiendrait à ces derniers de les prendre en compte, de préférence de manière automatisée lorsque le dispositif prévu par le détenteur le permet.

Ce dispositif spécifique peut notamment reposer sur une API dédiée à la communication d'informations relatives à l'exercice des droits, un balisage des données (voir annexe I) ou sur l'association de métadonnées aux données lors de leur communication. En effet, l'utilisation d'un fichier indiquant les restrictions de réutilisation des données afin que celles-ci soient prises en compte par les réutilisateurs n'est pas à privilégier.

Il convient de noter que la base légale qui fonde le traitement de partage de données a des conséquences sur les droits des personnes concernées, certains droits pouvant être écartés en fonction de celle-ci ou en vertu de textes réglementaires.

Sur la gestion des accès :

Lorsque l'API est soumise à une restriction d'accès (« API en accès restreint »), la Commission recommande que le mécanisme de demande d'accès mis en place exige de chaque réutilisateur qu'il apporte les informations nécessaires à la vérification de la licéité de son accès à l'API. Lorsque ce n'est pas obligatoire, il est généralement recommandé au réutilisateur d'informer le détenteur des données de la finalité pour laquelle il accède aux données, ainsi que des catégories de données nécessaires. La Commission recommande que le réutilisateur informe le détenteur du volume, de la profondeur historique, de la fréquence et du type de requêtes envisagés afin de dimensionner les moyens techniques à mettre en œuvre.

Dans le cas d'une « API ouverte », la Commission recommande, à titre de bonne pratique, qu'un système de registre public facultatif et ne restreignant pas l'accès à l'API permette de connaître ces mêmes informations et l'identité des destinataires. Ce registre pouvant porter atteinte à la vie privée ou aux données à caractère personnel des destinataires, son opportunité doit être évaluée au cas par cas.

Le détenteur de données et le gestionnaire d'API devraient se coordonner pour mettre en œuvre une procédure de gestion des accès et, le cas échéant, d'attribution des habilitations, répondant aux objectifs de sécurité, traçabilité et minimisation. Lorsque l'accès à l'API est soumis à une validation préalable, les procédures correspondantes devraient être formalisées au sein d'une politique de gestion des habilitations précisant les procédures d'attribution des secrets d'authentification, les conditions de leur transmission, de leur sauvegarde, de leur révocation et de leur renouvellement. L'authentification donnant accès à l'API devrait être réalisée par un système robuste et éprouvé de vérification de clé reposant sur des protocoles cryptographiques conformes aux recommandations de la Commission.

Les habilitations devraient être accordées selon plusieurs niveaux, ne donnant accès qu'aux données nécessaires au traitement et n'accordant que les permissions nécessaires. Cette différenciation peut avoir lieu par réutilisateur, mais aussi être mise en œuvre en interne par le réutilisateur afin de limiter les accès de ses agents ou employés aux données strictement nécessaires. Les accès devraient être donnés pour une durée déterminée, cohérente avec les besoins du réutilisateur et avec la durée de validité des données (certaines données, comme celles concernant la situation géographique ou encore financière des personnes par exemple, pouvant nécessiter des mises à jour régulières). En particulier, des accès à usage unique ou à très courte durée devraient être fournis pour réaliser des expérimentations ponctuelles. La sécurité du mécanisme d'authentification utilisé devrait être vérifiée et ses instructions d'utilisation rigoureusement suivies. Des mesures de sécurité devraient notamment être mises en place afin de protéger le système d'information des intrusions et des attaques par déni de service, par exemple en bloquant temporairement un accès après un nombre trop important de tentatives de connexions infructueuses. La possibilité devrait également être laissée aux réutilisateurs de révoquer unilatéralement leurs accès, notamment en cas de détection de compromission de leurs secrets d'authentification.

Lorsque les détenteurs de données réalisent le partage de leurs données au moyen d'une requête en écriture, les mesures de sécurité décrites dans le paragraphe précédent devraient leur être appliquées. Les droits d'écriture qui leur sont accordés devraient être limités à ce qui est strictement nécessaire, afin d'éviter la compromission de données déjà sauvegardées. Lorsque plusieurs bases coexistent, l'accès ne devrait être fourni qu'aux bases auxquelles le détenteur a besoin d'accéder. Une distinction entre les privilèges de lecture, d'ajout, de modification des données et de l'architecture de la base et d'écriture devrait être faite, afin de limiter les privilèges accordés à ceux strictement nécessaires.

Pour garantir la disponibilité des services pour les réutilisateurs de données, ces derniers devraient être notifiés à l'avance quand leur secret d'authentification arrive à expiration et un moyen de le renouveler sans interrompre leur service devrait leur être fourni.

Sur la gestion interne des API :

La Commission recommande qu'une gouvernance dédiée soit mise en place chez chacun des acteurs pour le partage de données par voie d'API. Cette démarche devrait être documentée et faire l'objet d'un suivi régulier pour garantir son effectivité.

Une documentation facilement accessible à tous les intervenants ayant à en connaître devrait formaliser les procédures et, notamment, les protocoles d'urgence à mettre en œuvre en cas de survenance d'un événement concernant la sécurité des données. Cette documentation devrait également comporter une description technique permettant l'intégration, le développement, la mise à jour et l'interruption des systèmes liés aux API.

La Commission recommande qu'un outil de gestion des versions (voir annexe I) soit utilisé afin de suivre les modifications apportées au code source du système permettant le partage. Une procédure devrait permettre de revenir à une version antérieure du système lorsqu'un risque est identifié. Une attention particulière devrait être apportée au cloisonnement du code source sauvegardé sur cet outil et des données n'ayant pas à y figurer, telles que les clés de chiffrement, dont la présence peut être détectée grâce à un outil vérifié de recherche des secrets dans le code source (voir annexe I). D'une manière générale, l'écriture pérenne, ou « en dur », de secrets dans le code source, devrait être évitée et l'utilisation d'un gestionnaire ou d'un détecteur de secrets devrait être privilégiée pour en éviter la divulgation ou pour la détecter.

Plus généralement, la gestion des API devrait s'inscrire dans la politique de sécurité des systèmes d'information de chacun des acteurs. Leur intégration devrait être prise en compte dans les procédures de sécurité existantes et ces dernières devraient être adaptées pour tenir compte des risques spécifiques aux API.

Article 3

Recommandations spécifiques à l'intention des détenteurs de données

Sur l'information des réutilisateurs :

La Commission recommande que le détenteur de données tienne une documentation à jour à l'intention des réutilisateurs, concernant les données rendues accessibles. Cette documentation devrait fournir en termes clairs des informations générales telles que la provenance des données, la méthode de collecte utilisée, leur fréquence de mise à jour, leur format de transmission, leur profondeur historique, leur fiabilité et les procédés de pseudonymisation ou d'anonymisation utilisés. Le détenteur de données devrait s'assurer que les réutilisateurs ont pris connaissance de cette documentation avant tout accès effectif aux données.

La Commission recommande par ailleurs l'utilisation de métadonnées associées aux données ou à un groupe de données, indiquant par exemple la date de collecte des données, leur fiabilité ou encore leur durée de validité. Le détenteur de données devrait s'assurer de la fiabilité des métadonnées, en particulier lorsque celles-ci décrivent la qualité, le statut, la disponibilité de la donnée ou encore ses conditions de réutilisation.

Enfin, le détenteur de données devrait mettre en place un canal de communication avec les réutilisateurs de données, afin que ces derniers puissent signaler tout problème technique ou risque relatif à la confidentialité, l'intégrité et la disponibilité des données, notamment les risques de réidentification identifiés *a posteriori*. A cet égard, plusieurs points de contact devraient exister, en fonction du niveau d'urgence des signalements, permettant ainsi au détenteur de données de prendre connaissance des risques signalés dans les meilleurs délais.

Sur l'exactitude et l'intégrité des données :

Le détenteur des données, généralement responsable du traitement de mise à disposition, doit accorder une attention particulière à l'exactitude et à l'intégrité des données avant et pendant leur transmission et mettre en place les mesures techniques et organisationnelles permettant de garantir que les réutilisateurs accèdent à des données exactes et à jour. En particulier, une vérification régulière de l'exactitude des données devrait être menée. Lorsqu'un risque particulier relatif à l'intégrité des données existe, des mesures reposant notamment sur des procédés cryptographiques, tels que des empreintes cryptographiques ou condensats (dits « hachés » ou « *hash* »), devraient être utilisées pour la garantir.

Lorsque le détenteur de données réalise lui-même la transmission des données au réutilisateur au moyen d'une requête en écriture, il devrait s'assurer d'avoir suivi la procédure décrite pour garantir leur intégrité et éviter tout risque de compromission d'une partie ou de la totalité de la base existante. Le détenteur de données devrait porter une attention particulière aux messages retournés par le serveur distant, ceux-ci pouvant indiquer qu'une erreur a eu lieu lors d'une requête en écriture.

Sur la sécurité :

Objectifs généraux :

Le détenteur de données doit assurer la sécurité des données qu'il produit et qu'il confie au gestionnaire d'API pour mise à disposition aux réutilisateurs.

Le détenteur de données devrait étudier la sécurité du système prévu pour le partage de données en lien avec le gestionnaire d'API et informer ce dernier de tout risque de sécurité identifié. En particulier, la sécurité de l'interface entre l'API et la base de données devrait être rigoureusement vérifiée. En effet, un partage de données par voie d'API, en tant qu'interconnexion entre deux systèmes d'information, notamment lorsqu'un des deux appartient à une entité tierce, constitue une modification significative du système d'information. Ce changement devrait donc entraîner un renouvellement anticipé d'une éventuelle homologation du système d'information pour prendre en compte les menaces et risques résultant de cette interconnexion.

Cloisonnement et disponibilité :

La Commission recommande que les données dont l'accès est prévu par l'API (« base source » de l'API) soient séparées des autres catégories de données lorsque le partage de ces dernières n'est pas prévu et que leur divulgation pourrait entraîner des conséquences graves pour les personnes. En particulier, lorsqu'un procédé de

pseudonymisation ou d'anonymisation est prévu, les données brutes devraient être physiquement ou logiquement séparées des données issues de ce procédé. La base source de l'API pourra alors être alimentée par un procédé d'exportation de données automatisé et sécurisé, sous réserve que la latence introduite par ce procédé n'entraîne pas d'indisponibilité des données.

Lorsque les réutilisations prévues par les réutilisateurs sont des traitements critiques dont une indisponibilité (même temporaire) pourrait entraîner des conséquences graves, le détenteur de données devrait accorder une importance particulière à la disponibilité des données. Il devrait mettre en place les mesures techniques nécessaires pour éviter une compromission ou une indisponibilité de la base de données et pour en limiter l'impact le cas échéant, comme la redondance des infrastructures ou la mise en œuvre régulière de tests d'intégrité des données et de leurs sauvegardes.

Authentification :

Lorsque le détenteur de données s'authentifie pour accéder à l'API, ce qui peut être le cas notamment lorsqu'il réalise lui-même l'écriture sur la base du réutilisateur, il devrait mettre en place les mesures techniques et opérationnelles garantissant la sécurité des secrets d'authentification et, notamment, leur intégrité et leur confidentialité. Un système de sécurisation des secrets d'authentification adapté aux risques liés à une compromission de l'accès à l'API, tel qu'un coffre-fort numérique, devrait être utilisé. Lorsqu'un risque relatif à la sécurité des secrets est identifié, le réutilisateur devrait en informer le gestionnaire d'API dans les plus brefs délais, pour que celui-ci procède à leur révocation lorsque c'est nécessaire.

Journalisation :

Le détenteur de données devrait s'assurer qu'une journalisation effective des accès et actions réalisées sur la base de données a lieu : cette journalisation peut être techniquement réalisée par le détenteur des données et/ou par le gestionnaire d'API. Dans tous les cas, le détenteur des données devrait conserver une copie de ces traces, pour une durée conforme aux recommandations de la Commission.

La journalisation des accès externes, par les réutilisateurs autorisés à utiliser l'API, devrait être réalisée avec un niveau de détail dépendant de l'importance du risque que représente une intrusion dans la base de données ou une utilisation détournée du traitement. Dans le cas général, la Commission recommande que les opérations de création, consultation, modification et suppression des données à caractère personnel auxquels la journalisation est appliquée fassent l'objet d'un enregistrement comprenant l'auteur individuellement identifié, l'horodatage, la nature de l'opération réalisée ainsi que la référence des données concernées par l'opération avec un degré de précision adapté à la gravité des conséquences que pourraient entraîner pour les personnes une utilisation détournée de l'API ou un accès illégitime aux données.

Une analyse proactive et régulière des journaux internes et externes devrait être menée afin de vérifier la légitimité des actions réalisées. Celle-ci peut être automatisée (générant des alertes revues par des opérateurs) ou bien mise en œuvre par des mesures organisationnelles (par exemple, génération de rapports réguliers et contrôle humain des accès aux données). Plus particulièrement, dans les cas où un comportement inhabituel peut facilement être identifié, une journalisation et une analyse spécifiques devraient avoir lieu et faire l'objet d'un signalement permettant d'en vérifier la légitimité. Lorsqu'un détournement du traitement pourrait entraîner des conséquences importantes pour les personnes, tout accès supposé illégitime aux données devrait leur être signalé dans les plus brefs délais.

Les journaux pourraient également être utilisés afin de vérifier que les réutilisateurs ont bien pris en compte une éventuelle mise à jour des données. La précision et la durée de conservation des données doit être adaptée à chacune des finalités de journalisation. Dans ce cas, le détenteur des données pourrait ainsi alerter les réutilisateurs. Les informations issues de la journalisation permettent également d'assurer une traçabilité sur les accès effectifs aux données, dont les résultats pourraient être fournis aux personnes concernées, notamment à des fins de transparence ou dans le cadre de l'exercice de son droit d'accès. Le détenteur des données devrait s'assurer que ces informations leur sont restituées dans un format accessible et aisément compréhensible, notamment par le recours à des procédés statistiques.

Article 4

Recommandations spécifiques à l'intention des gestionnaires d'API

Le gestionnaire d'API est le mieux à même d'assurer en pratique la sécurité de la mise en œuvre de l'API, que ce soit en tant que responsable de traitement auquel cette obligation incombe, ou bien en tant que sous-traitant auquel le responsable de traitement confie contractuellement la mise en œuvre de cette obligation. Il réalise le lien entre le détenteur de données et les réutilisateurs et s'assure que le système est conforme à leurs besoins.

Sur la documentation :

La Commission recommande que le gestionnaire d'API crée une documentation à l'intention des détenteurs de données et des réutilisateurs, suffisamment détaillée et transparente pour réduire les risques liés à une utilisation imprévue de l'outil. Un générateur automatique de documentation (voir annexe I) reconnu pourra être utilisé à cet effet. Cette documentation pour être complète devrait reposer sur plusieurs supports (fichiers « readme », site web de type « wiki », commentaires apportés au code si celui-ci est ouvert, infolettre informant sur les mises à jour et nouveautés, etc.).

Cette documentation devrait décrire en premier lieu la procédure d'accès aux API. Lorsque l'accès à l'API est soumis à une validation préalable, les conditions à remplir pour y accéder devraient être clairement décrites.

Lorsque l'accès à l'API peut être donné selon différents niveaux d'habilitation, les accès prévus pour chacun de ces niveaux et les profils de réutilisateurs attendus, par leur finalité notamment, devraient être clairement décrits. Des indications concernant les modalités d'utilisation et de stockage des secrets d'authentification devraient être fournies.

En deuxième lieu, le format des données et des métadonnées et les décisions spécifiques prises concernant leur représentation devraient être décrits dans la documentation. Le gestionnaire d'API devrait indiquer la signification de chacune des variables décrivant les données, et en proposer des exemples d'utilisation clairs. Le gestionnaire d'API devrait décrire précisément les variables relatives à la sécurité des données ou décrivant leurs conditions techniques d'utilisation telles que leur durée de validité ou leur fiabilité. De plus, la documentation devrait préciser les catégories de requêtes et leur format. Elle devrait également décrire les paramètres devant ou pouvant être passés dans ces requêtes. Des exemples représentatifs des usages prévus de l'API devraient être fournis pour illustrer les informations précédentes.

En troisième lieu, la Commission recommande que la documentation décrive les aspects relatifs à la sécurité du système. Les limites de l'API devraient y être indiquées, en particulier le volume et la fréquence maximale des requêtes. Les capacités du système, notamment en cas de forte demande, devraient y être également décrites et, lorsque des périodes de forte demande peuvent être anticipées, les périodes à privilégier devraient être indiquées. Les standards, normes et certifications auxquels se conforme l'API devraient également être décrits.

En dernier lieu, la documentation devrait indiquer plusieurs points de contact permettant à quiconque de signaler un problème concernant la sécurité de l'API. Un moyen de communication réactif devrait être choisi pour traiter les demandes urgentes.

Sur la minimisation :

Le gestionnaire, en ce qu'il réalise l'implémentation technique de l'API, est le mieux à même de mettre en œuvre les mesures techniques de minimisation décidées lors de la concertation entre le détenteur de données et le réutilisateur détaillée à l'article 2.

Expérimentations dans un « bac à sable » :

La Commission recommande qu'une version jumelle de l'API, donnant accès à des données fictives, soit proposée afin de permettre aux réutilisateurs de données de réaliser des expérimentations et de sélectionner plus précisément les données nécessaires au traitement. L'accès à cette version devrait être facilité et une aide technique devrait être proposée aux réutilisateurs.

Limitations :

La Commission recommande que le gestionnaire d'API mette en œuvre des limitations au traitement des requêtes afin d'assurer la disponibilité du système et de prévenir toute utilisation détournée de l'API. Ces limitations devraient s'appliquer à chacune des requêtes, ainsi que par réutilisateur de données. Les limitations pourraient porter sur le volume, la profondeur historique et la précision des données. Ces limitations devraient prendre en compte les besoins réels des réutilisateurs et ne devraient pas empêcher l'accès à des données à jour, faute de quoi les réutilisateurs risqueraient d'utiliser des données inexactes.

Pour les usages d'API dont l'objectif est de ne pas révéler de données à caractère personnel (données anonymisées ou agrégées), des méthodes visant à préserver la confidentialité des données de chacune des personnes enregistrées dans la base devraient être mises en œuvre par le gestionnaire d'API. La robustesse de ces méthodes devrait être vérifiée, en particulier vis-à-vis de méthodes de réidentification telles que les attaques par corrélation. En effet, même lorsque les données ne sont pas directement identifiantes, un réutilisateur pourrait être en capacité de réidentifier une personne en réalisant des requêtes croisées sur la base de données, en effectuant un suivi longitudinal sur des données temporelles, ou en utilisant des informations tierces accessibles en dehors de cette base.

Sur l'exercice des droits concernant le partage de données :

La Commission recommande que le gestionnaire d'API mette en œuvre les mesures techniques nécessaires afin de permettre aux détenteurs de données et aux réutilisateurs, le cas échéant, de répondre aux demandes d'exercice des droits comme prévu à l'article 2.

Ces mesures devraient être précisées dans les documents décrivant les modalités de coordination entre détenteur de données, gestionnaire d'API et réutilisateur.

Sur la sécurité :

Objectifs généraux :

Le gestionnaire d'API s'assure du respect des obligations de sécurité résultant de l'article 32 du RGPD, en s'appuyant notamment sur les recommandations de la CNIL en vigueur, telles que son guide de la sécurité des données à caractère personnel. Pour la mise en œuvre des API, la Commission recommande que le gestionnaire d'API se conforme aux normes techniques communément admises, telles que le référentiel général de sécurité (RGS) (4) de l'Agence nationale de sécurité des systèmes d'information et ses recommandations applicables au type de système considéré, les références « RFC » (pour « *requests for comments* » en anglais) relatives à des protocoles standardisés, ainsi que des solutions éprouvées.

Communications :

En particulier, dans le cas d'une API en accès restreint, le chiffrement appliqué aux communications devrait garantir un haut niveau de confidentialité. Pour rappel, la Commission considère que les règles et recommandations décrites dans les annexes B1 et B2 du RGS sont la référence en ce qui concerne l'état de l'art de la cryptographie. Le gestionnaire d'API devrait mettre en œuvre les mesures nécessaires afin que ce niveau de chiffrement soit appliqué à toutes les communications, quel que soit le niveau de maturité technique du détenteur de données ou des réutilisateurs. En particulier, le gestionnaire d'API devrait définir et imposer des mesures de chiffrement minimales aux réutilisateurs.

Sécurité des systèmes d'information :

Le gestionnaire d'API devrait assurer la sécurité du système d'information dans sa globalité et dans le temps, en appliquant les normes et pratiques à l'état de l'art dans le domaine, telles que les normes ISO 9001 et ISO 27001. Le gestionnaire d'API devrait mettre en œuvre les mesures nécessaires pour se prémunir contre les attaques les plus connues et pouvant vraisemblablement être anticipées, telles que les injections de code ou les attaques exploitant des vulnérabilités entre sites de type « *cross-site request forgery* » (CSRF). La Commission recommande à cet égard l'utilisation de requêtes préparées et de mesures visant à prémunir contre les attaques de type CSRF. Les outils tiers devraient faire l'objet d'une analyse *a priori* afin de garantir leur sécurité et leur pérennité. Les composants logiciels tiers mettant en œuvre des API liées au traitement devraient être analysés et les instructions relatives à leur utilisation, et en particulier à leur sécurité, connues et appliquées. Les points d'accès non utilisés devraient être identifiés et révoqués. Lors de la mise à jour d'un outil logiciel, la version antérieure devrait être conservée pendant une période permettant d'assurer la compatibilité des accès pendant une période de recouvrement. Lors de la fermeture d'une API ou d'une de ses versions, une attention particulière devrait être apportée à la révocation effective des accès afin de garantir que les réutilisateurs n'ont accès qu'aux seules versions des API dont l'ouverture est effectivement prévue.

Le gestionnaire d'API devrait s'assurer de la fiabilité et de la robustesse des métadonnées et autres informations relatives à la sécurité du système. Des informations décrivant la qualité, la validité, et la disponibilité des données et des indications concernant le statut de l'API, telles que la charge instantanée ou des statistiques relatives à son utilisation, devraient être fournies aux réutilisateurs.

Les interruptions de la disponibilité de l'API devraient être prévues longtemps à l'avance, et les réutilisateurs de l'API devraient en être informés. Cette information devrait inclure les éléments nécessaires à l'information des personnes concernées, dans l'hypothèse où cette indisponibilité programmée aurait un impact sur celles-ci.

Lorsque les réutilisations prévues par les réutilisateurs sont des traitements critiques dont une indisponibilité, qu'elle résulte d'un acte malveillant ou accidentel, pourrait entraîner des conséquences graves, le gestionnaire d'API devrait accorder une importance particulière à la disponibilité de l'API et prévoir un dispositif alternatif à celle-ci garantissant un niveau de sécurité similaire. Il devrait prioriser les traitements en question et mettre en place les mesures techniques permettant de mesurer le niveau de disponibilité de l'API et de garantir que celui-ci reste suffisant.

Journalisation :

La Commission recommande vivement que le gestionnaire d'API mette en œuvre des outils permettant une journalisation des accès à l'API par les réutilisateurs conforme à la recommandation de la CNIL. Les fonctionnalités des API facilitant la traçabilité devraient être exploitées. Selon le niveau de risque évalué, la quantité d'information journalisée devrait être adaptée, selon le cadre général décrit à l'article 3.

Ainsi, une analyse régulière et proactive devrait pouvoir être réalisée par les outils mis en œuvre par le gestionnaire d'API, afin de vérifier la légitimité des actions réalisées. Les procédures prévues devraient permettre de détecter les surcharges du système et l'indisponibilité des données. Ces analyses devraient donner lieu à des signalements internes ou au détenteur de données. Il est recommandé que des informations statistiques relatives à la disponibilité du système et à son utilisation soient fournies au détenteur de données et aux réutilisateurs.

Article 5

Recommandations spécifiques à l'intention des réutilisateurs

Il appartient au réutilisateur de données d'appliquer rigoureusement les instructions à sa disposition concernant l'utilisation et la sécurité de l'API et de s'assurer de la licéité des usages qu'il fait des données. Lorsque le réutilisateur constate que certaines instructions sont obsolètes, inadaptées, incomplètes ou non conformes à l'état de l'art en matière de sécurité, il devrait en informer le détenteur de données ou le gestionnaire d'API.

Lorsqu'une charte ou licence de réutilisation est fournie, le réutilisateur devrait en prendre connaissance et la diffuser à chacun de ses agents ou employés ayant accès aux données.

Sur l'information des personnes concernées :

L'information des personnes concernées sur le partage de leurs données avec les réutilisateurs relève de la ou des entités qui endossent la responsabilité de traitement de ce partage, ce qui est en principe toujours le cas du détenteur, et peut être, dans certaines hypothèses, celui du réutilisateur. Outre les catégories de données et les finalités de leur accès par le réutilisateur, la Commission recommande que ce dernier informe les personnes concernées sur les mesures de minimisation prises et le niveau de sécurité appliqué aux données, et, à titre de bonne pratique, sur la volumétrie et la fréquence attendues des requêtes.

Sur la minimisation :

Lorsqu'un accès temporaire ou restreint est proposé dans le cadre d'une expérimentation de l'API, le réutilisateur devrait en profiter pour déterminer les données strictement nécessaires à ses besoins.

Pour des besoins ponctuels ou concernant des données peu volumineuses, le réutilisateur de données devrait interroger l'API chaque fois qu'il entend traiter les données partagées, c'est-à-dire sans les conserver dans ses propres systèmes informatiques. En requêtant systématiquement au moyen de l'API les données dont il a besoin, il s'assure ainsi d'obtenir les données les plus à jour, limite leur surface d'exposition et prend en compte au plus tôt toute modification faisant suite à une demande d'exercice des droits. Lorsque la duplication des données est inévitable, le réutilisateur doit notamment limiter celle-ci au strict nécessaire, définir une durée de conservation maximale et s'assurer que les conditions de sécurité des données sont adéquates. Lorsque la flexibilité de l'API ne permet pas une sélection suffisamment fine des données pertinentes, ou lorsque les données dont l'accès est donné par l'API sont utilisées pour réaliser un traitement permettant d'obtenir la donnée pertinente requise, les données sources non pertinentes ou devenues inutiles doivent en principe être supprimées sous les plus brefs délais.

Le réutilisateur doit utiliser les informations à sa disposition afin de s'assurer que les données traitées sont à jour et, le cas échéant, n'ont pas fait l'objet d'une opposition par les personnes pour la finalité correspondant à leur traitement.

Sur la sécurité :

Gestion des risques :

Lorsque le réutilisateur identifie un risque relatif à la confidentialité des données ou un risque de sécurité de l'API, il devrait en informer le détenteur de données et le gestionnaire d'API dans les plus brefs délais.

Sécurisation des clés :

Lorsque l'accès à l'API est sécurisé au moyen d'une technique d'authentification reposant sur un échange de clés, le réutilisateur de données devrait sécuriser les clés lui permettant d'accéder aux données en hébergeant ces dernières dans un répertoire sécurisé, voire un système de sécurisation des clés, tel qu'un coffre-fort numérique sécurisé, lorsque la gravité des risques liés à une compromission de l'accès à l'API le justifie. Lorsqu'il détecte ou suspecte une compromission de ses clés d'accès, il les révoque immédiatement et demande la génération de nouvelles clés au gestionnaire d'API.

Journalisation :

Dans le cas où le réutilisateur de données (que ce soit avec l'aide d'un gestionnaire d'API ou qu'il endosse également ce rôle) met à disposition l'API permettant aux détenteurs de données de partager activement leurs données par une requête en écriture, il doit mettre en œuvre les recommandations figurant à l'article 3 pour assurer une journalisation des modifications découlant de l'usage de l'API par le détenteur. Cette journalisation devrait permettre d'identifier en particulier les actions ou tentatives d'action visant à porter atteinte à l'intégrité des données ou de la base de données.

La présente délibération sera publiée au *Journal officiel* de la République française.

La présidente,
M.-L. Denis

---

(1) www.ssi.gouv.fr/rgs.
(2) La liste de ces recommandations peut être trouvée sur le site web de la CNIL.
(3) https://www.cnil.fr/fr/RGPD-analyse-impact-protection-des-donnees-aipd.
(4) https://www.ssi.gouv.fr/rgs.

ANNEXES

ANNEXE I

DÉFINITIONS

Interface de programmation applicative, ou *application programming interface* (API) : tout ensemble abstrait de fonctions, de procédures, de définitions et de protocoles qui permet la communication de machine à machine et la transmission de données dans un format structuré.

Détenteur de données : tout organisme ou personne public ou privé qui détient des données, ici à caractère personnel, c'est-à-dire qui en contrôle le cycle de vie et les modalités d'accès, ces données ayant vocation à être partagées à des tiers via l'utilisation d'une API, sous leur forme originale ou après l'application d'une transformation telle qu'un procédé d'anonymisation.

Gestionnaire d'API : tout organisme ou personne public ou privé en charge de l'opération et/ou du développement des composants techniques permettant le partage des données via l'API. Le rôle de gestionnaire d'API peut être tenu par plusieurs organismes lorsque la gestion et le développement des outils techniques implique plusieurs acteurs. En particulier, lorsque l'API est mise en œuvre au moyen d'un outil développé par un tiers, l'organisme qui détient la licence sur cet outil est gestionnaire d'API, tout comme celui qui est en charge du déploiement de cet outil dans le système d'information permettant le partage des données.

Réutilisateur des données : tout organisme ou personne public ou privé ayant accès aux données partagées par l'API.

Fournisseur d'API : tout organisme ou personne qui expose une API ouvrant la possibilité pour un consommateur d'utiliser un service que le fournisseur met à disposition, tel qu'un accès à des données.

Consommateur d'API : tout organisme utilisant l'API exposée par un fournisseur afin d'utiliser le service qu'il met à disposition, comme pour accéder à des données ou pour les transmettre.

Plateforme de partage de données : service reposant sur serveur partagé auquel le détenteur de données et le réutilisateur ont accès afin de partager des données. Les données sont généralement contenues dans des fichiers.

Service de télécommunication : service permettant l'échange d'informations par des procédés numériques, tel qu'un service de messagerie électronique.

Requête à l'API : toute demande reposant sur une interrogation de l'API et contenant des instructions décrivant l'action à réaliser. Les requêtes les plus fréquentes, dans le protocole HTTP, sont GET, POST (qui sont respectivement des exemples de requêtes en lecture et en écriture) ou encore PATCH, qui permet de mettre à jour partiellement des données, et DELETE, qui permet d'en supprimer.

Requête en lecture : interrogation de l'API permettant d'obtenir des données. Ce type de requête s'accompagne généralement de l'identifiant des données recherchées ou de règles permettant de les retrouver, et le cas échéant de paramètres destinés à sélectionner l'information retournée et nécessite un droit de lecture sur la base de données sous-jacente.

Requête en écriture : message envoyé via l'API contenant des instructions telles que l'inscription de données contenues dans le message. Ce type de requête nécessite un droit d'écriture sur la base de données.

Métadonnée : information permettant de décrire une donnée ou un ensemble de données (telle que sa date de création, son type, des informations relatives à sa provenance ou à sa validité, etc.).

Différenciation des accès et permissions par niveau : limitation des accès aux données à caractère personnel à celles strictement nécessaires pour chacun des agents ayant à en connaître et limitation des permissions (lecture, écriture, suppression, etc.) aux seules actions nécessaires pour chacun des agents. Il s'agit d'une mise en pratique du principe de moindre privilège.

Balisage : information annexe (ou « métadonnée ») liée à une donnée et indiquant un statut particulier, tel que sa validité, son origine, ou encore les conditions limitant sa réutilisation. Pour être pertinente, l'indication apportée doit être sans ambigüité et transparente.

Outil de gestion de versions : moyens techniques permettant de conserver les modifications successives d'un logiciel ou d'un document et leur historique, ainsi que d'en restituer toute version antérieure.

Outil de validation des données : outil permettant de vérifier la conformité des données échangées par API au format attendu (tel que la correspondance à la description de l'API, la correspondance au type de donnée attendu, l'appartenance à un ensemble de valeurs admises, etc.).

Générateur automatique de documentation : outil permettant de générer la documentation d'une API de manière automatique, à partir du code source de l'API. Ces outils permettent de mettre à jour automatiquement la documentation lors d'une modification apportée à l'API, et peuvent intégrer une gestion des versions. L'utilisation de ces outils permet d'éviter que les réutilisateurs ne se réfèrent à une documentation inexacte.

Outil de recherche des secrets dans le code source : outil permettant de détecter la présence de secrets (jetons d'accès, clés de chiffrement, mots de passe, etc.) par analyse du code source d'un logiciel. Les outils fonctionnant localement (ou *on premise*) et générant une alerte lors de la présence d'un secret dans le code, avant que celui-ci ne soit publié sur un serveur au moyen d'un outil de gestion des versions par exemple, devraient être privilégiés.

Requests for comments : série de documents de standardisation décrivant les spécifications techniques des protocoles mis en œuvre au sein de l'internet et des matériels informatiques sous-jacents, comme la RFC 2068 sur le protocole HTTP.

ANNEXE II

CAS D'USAGE PARTICULIERS

– **Exemple #1 : Partage restreint de données entre plusieurs organismes avec contractualisation**

– *Description du contexte :*

Lors de ce partage, un premier organisme privé (**le détenteur de données**) collecte des données personnelles auprès de ses clients. Il met ensuite les données de ses clients y ayant consenti à disposition de tiers agréés (**les réutilisateurs de données**). Ces derniers réutilisent les données pour proposer un nouveau service à ces mêmes clients. Interviennent alors des organismes apporteurs de solutions techniques (**les gestionnaires d'API**), dont le cœur de métier est la valorisation de ces catégories de données, par des méthodes d'analyse ou d'agrégation, et leur partage avec des organismes réutilisateurs. Le rôle de ces derniers sera de proposer des services personnalisés aux personnes reposant sur les données valorisées par l'apporteur de solution.

La quantité de données collectée concernant les clients est relativement importante. De plus, les catégories de données concernées par ce partage sont diverses et permettent de déduire de nombreuses informations à propos des personnes, par exemple sur leurs préférences dans un domaine ou encore leurs habitudes de consommation. Les apporteurs de solution proposent leur service à de nombreux détenteurs de données, et à de nombreux

réutilisateurs. Les détenteurs de données peuvent par ailleurs avoir recours à différents apporteurs de solutions techniques.

– *Recherche des facteurs de vulnérabilité :*

De par la nature de ce traitement, plusieurs facteurs de vulnérabilité sont identifiés :

– sur les catégories de données accessibles : les données des personnes sont en volume important et peuvent parfois révéler des informations particulièrement intrusives sur leurs habitudes de vie. Les objectifs à atteindre sont **la minimisation et la sécurité des données** ;
– sur la granularité des données et des requêtes : les réutilisateurs souhaitent accéder à un certain niveau de granularité des données pour mieux cibler les services qu'ils proposent. Les objectifs à atteindre sont **la minimisation et la sécurité des données** ;
– sur les conditions d'accès aux données : plusieurs gestionnaires d'API peuvent avoir accès à la base de données d'un détenteur, puis proposer leur API à plusieurs réutilisateurs. Les objectifs à atteindre sont **l'information des personnes, la traçabilité, la gouvernance, la sécurité des données et le respect des droits des personnes** ;
– sur la nature des organismes : les organismes réutilisateurs peuvent être nombreux, et parmi ces derniers figurent des entreprises de maturité variable, dont la compréhension de leurs obligations liées au RGPD peut être insuffisante et les moyens en ce qui concerne la sécurité des données limités. Les objectifs à atteindre sont **l'exactitude, l'information des personnes, la traçabilité, la gouvernance, la sécurité des données et le respect des droits des personnes**.

– *Recommandations applicables :*

D'après les facteurs de vulnérabilité identifiés, plusieurs objectifs sont à prioriser.

• **Objectif #1 : la minimisation des données.**

Compte tenu du caractère particulièrement intrusif des données partagées et de leur volume important, tous les organismes impliqués dans le partage devraient se coordonner afin de sélectionner les données strictement nécessaires à l'atteinte des finalités recherchées. La granularité des requêtes disponibles pour les réutilisateurs devrait également être adaptée pour limiter le contenu des réponses, et notamment les données identifiantes, à ce qui est strictement nécessaire, selon les modalités décrites dans la recommandation.

Le détenteur de données sera le plus à même de sélectionner les données pertinentes selon les besoins des réutilisateurs. De plus, le gestionnaire d'API peut être l'organisme le plus à même de mettre en œuvre techniquement ces recommandations.

• **Objectif #2 : l'information des personnes et la traçabilité des données.**

Etant donné le grand nombre d'acteurs impliqués et leur niveau de maturité variable, les mesures de journalisation prévues par la recommandation devraient être mises en œuvre afin que puisse être vérifiée la légitimité des accès et des actions réalisées sur la base de données.

Ces mesures devraient être complétées de celles prévues par la recommandation afin de fournir une information complète aux personnes concernées sur l'utilisation qui est faite de leurs données.

• **Objectif #3 : la gouvernance et le respect des droits des personnes.**

Au vu du nombre d'organismes avec lesquels les données sont partagées, la responsabilité du traitement et les modalités d'exercice des droits qui doivent être prévues en amont du traitement, devraient reposer sur des procédures robustes comme prévu par la recommandation.

• **Objectif #4 : la sécurité des données.**

La maturité des organismes impliqués pouvant être variable, les recommandations relatives à la procédure de fourniture des accès décrites dans la recommandation devraient être prises en compte.

Les autres mesures relatives à la sécurité devraient être considérées au cas par cas. En effet, bien qu'elles s'appliquent dans la majorité des cas, ces mesures doivent être mises en œuvre selon le niveau de risque évalué.

– **Exemple #2 : Partage de données entre réseaux sociaux et chercheurs**

– *Description du contexte :*

Dans le cadre d'une étude portant sur les mécanismes de propagation de certaines catégories d'information sur les réseaux sociaux, des chercheurs académiques souhaitent accéder à des informations liées aux activités des usagers de ces réseaux. Pour cela, un accès à une API leur est fourni par la plateforme qui utilise cet outil pour partager les données auprès de nombreux réutilisateurs. Par le biais de l'API, les chercheurs peuvent effectuer des requêtes et obtenir des informations comme les publications et les réactions des usagers. Parmi ces informations, figurent des données personnelles.

L'étude des chercheurs porte sur une période spécifique et sur des catégories d'information bien déterminées. Une fois l'étude publiée, la conservation des données en base active et d'un accès à l'API n'est plus utile. Les données sont stockées en archivage intermédiaire dans les serveurs sécurisés de l'université (5). Toutefois, durant leur recherche, certains chercheurs ont téléchargé les données sur leur machine locale pour avoir accès aux données lors de leurs déplacements.

– *Recherche des facteurs de vulnérabilité :*

De par la nature de ce traitement, plusieurs facteurs de vulnérabilité sont identifiés :

– sur les conditions d'accès aux données : les besoins des chercheurs sont temporaires et ciblés, une fois l'étude publiée, leur accès à l'API n'est plus nécessaire mais la procédure prévue par le réseau social ne prévoit une réévaluation des accès qu'après une durée plus importante, laissant un accès à l'API inutile et vacant. Les objectifs à atteindre sont **la traçabilité, la gouvernance, la sécurité des données et le respect des droits des personnes** ;
– sur la nature des organismes impliqués dans le partage : le réseau social partage des données avec de nombreux réutilisateurs, et n'est pas toujours en mesure de se coordonner avec chacun d'entre eux ou de journaliser les accès de chacun. Les objectifs à atteindre sont **l'exactitude, la traçabilité, la gouvernance, la sécurité des données et le respect des droits des personnes** ;
– sur les autres mesures techniques et organisationnelles prévues pour améliorer le niveau de sécurité du système : certains chercheurs stockant les données sur leurs machines, il est difficile de garder une vue d'ensemble sur les sauvegardes de données, et le niveau de sécurité appliqué aux machines n'est pas le même que celui appliqué aux serveurs de l'université. Les objectifs à atteindre sont **l'exactitude, la traçabilité et la sécurité des données** ;
– sur les catégories de données accessibles par l'API : les usagers du réseau social y publient des photographies, des contenus qui les intéressent ou encore des informations sur leurs activités. Ces données pouvant révéler des informations particulièrement intrusives sur leurs préférences et sur leurs habitudes de vie, un détournement de ces données pourrait entraîner des conséquences importantes pour les personnes. Les objectifs à atteindre sont **la minimisation et la sécurité des données** ;
– sur la granularité des données et des requêtes : l'API proposée par le réseau social est destinée à différentes catégories de réutilisateurs et leurs besoins propres peuvent difficilement être anticipés en amont. Les accès fournis par le réseau social ne sont pas toujours strictement limités aux besoins des réutilisateurs, qui peuvent accéder à des données dont ils savent en amont qu'elles ne sont pas pertinentes pour leur réutilisation. Les objectifs à atteindre sont **la minimisation et la sécurité des données**.

– *Recommandations applicables :*

D'après les facteurs de vulnérabilité identifiés, plusieurs objectifs sont à prioriser.

• **Objectif #1 : l'exactitude des données.**

Le téléchargement des données sur de nombreuses machines par les chercheurs représente un risque que les données ne soient pas mises à jour lorsqu'elles devraient l'être en raison de leur duplication, ou de modification compromettant leur exactitude.

De plus, étant donné la nature des données sur les réseaux sociaux, leur exactitude peut poser question d'une part en raison de la possibilité pour un utilisateur de les modifier après leur publication, et d'autre part en raison de l'absence de garanties de véracité sur ces données, bien que cela ne nuise pas nécessairement à l'objectif de la recherche.

• **Objectif #2 : la traçabilité des données.**

Le nombre de réutilisateurs prévu étant relativement important, et les conditions d'attribution et de révocation des accès étant généralement peu flexibles pour ce type d'API, des mesures devraient être mises en œuvre par le réseau social afin de vérifier la légitimité des accès et d'identifier les accès et utilisations imprévues de l'API. Ces traces de journalisation devraient être utilisées afin d'empêcher les utilisations détournées des données.

Les téléchargements de données par les chercheurs sur les machines locales devraient être évités autant que possible, et faire l'objet d'un suivi pour éviter toute perte de traçabilité sur les données. De plus, les jetons ou identifiants d'accès à l'API devraient être partagés aux seules personnes habilitées et faire l'objet d'un suivi lorsqu'il n'est pas possible d'obtenir des jetons propres à chaque personne.

• **Objectif #3 : la minimisation des données.**

Les besoins des chercheurs étant ciblés, les requêtes qui leur sont permises, et les données auxquelles ils ont accès devraient être limitées à ce qui est pertinent et strictement nécessaire. Si la traçabilité sur les données pertinentes n'était pas suffisante en amont du traitement, une version « bac à sable » expérimentale, reposant éventuellement sur des données fictives pourrait être utilisée afin de déterminer les catégories et la quantité de données nécessaires. Il revient par ailleurs au réutilisateur (l'organisme de recherche) de fixer des règles de minimisation sur les données obtenues par l'API.

• **Objectif #4 : la gouvernance et le respect des droits des personnes.**

Les données pouvant être partagées par le réseau social à un nombre important et à des catégories diverses de destinataires, les utilisateurs du réseau peuvent manquer de traçabilité sur les utilisations de leurs données. Dans un souci de transparence et pour remplir les obligations relatives à l'information des personnes, le réseau social devrait prendre les mesures nécessaires pour connaître les réutilisateurs et fournir ces informations aux personnes.

• **Objectif #5 : la sécurité des données.**

Les mesures relatives à la sécurité devraient être considérées au cas par cas. En effet, bien qu'elles s'appliquent dans la majorité des cas, ces mesures doivent être mise en œuvre selon le niveau de risque évalué. Ici, une attention particulière devrait être portée sur la sécurité des machines locales sur lesquels les chercheurs ont téléchargé les données. Ce téléchargement est à éviter autant que possible. Lorsqu'il est nécessaire, la sécurité des machines devient un objectif prioritaire.

– **Exemple #3 : L'ouverture de données de l'administration**

– *Description du contexte :*

Un texte prévoit l'ouverture par une administration (le détenteur de données) d'une certaine catégorie de données collectées auprès des personnes. Ce texte indique précisément les catégories de données concernées et les modalités d'exercice des droits des personnes. Il précise également que les données sont mises à disposition de tous, sans restriction d'accès et dans un format facilitant leur réutilisation. Les réutilisations possibles des données ne sont pas connues en amont par l'administration qui diffuse les données. Une administration tierce (le gestionnaire d'API) propose un outil afin de diffuser les données en question. Elle n'héberge cependant pas les données.

Après plusieurs études, il est prouvé que les données sont particulièrement utiles pour un certain secteur et un grand nombre d'organismes de ce secteur (les réutilisateurs) souhaitent alors collecter les données. Les données restent ouvertes, mais les besoins relatifs des réutilisateurs ont largement changé. Ce nouveau besoin conduit à une charge sur les serveurs du détenteur de données qui n'était pas prévue lors de la conception de l'API, causant des interruptions dans la disponibilité des données.

– *Recherche des facteurs de vulnérabilité :*

De par la nature de ce traitement, plusieurs facteurs de vulnérabilité sont identifiés :

– sur les conditions d'accès à la base de données : même lorsque les données sont pseudonymisées, et que le risque de réidentification est faible, l'évolution des attaques pourrait permettre la déduction de données personnelles supplémentaires dont l'ouverture n'est pas prévue par le texte et ayant des conséquences pour les personnes. L'ouverture des données à tous augmente la vraisemblance de ce type d'attaque. Les objectifs à atteindre sont **l'information des personnes, la traçabilité, la gouvernance, la sécurité des données et le respect des droits des personnes** ;
– sur la nature des organismes impliqués dans le partage : bien que les données ne soient initialement demandées que par certains réutilisateurs isolés, c'est finalement un grand nombre d'organismes qui en demande l'accès. Cette évolution peut avoir des conséquences importantes sur la disponibilité des données et leur réutilisation. Les objectifs à atteindre sont **l'exactitude, la traçabilité, la gouvernance, la sécurité des données et le respect des droits des personnes** ;
– sur la granularité des données et des requêtes : aucune restriction d'accès aux données n'est prévue, laissant ainsi la possibilité aux réutilisateurs de réaliser de multiples requêtes, dont les résultats peuvent être croisés entre eux. Cela rend la déduction d'informations à caractère personnel, dont l'ouverture n'était pas prévue, plus vraisemblable. Les objectifs à atteindre sont **la minimisation et la sécurité des données** ;
– sur l'état des connaissances des techniques utilisées et les risques associés : l'ouverture des données ayant pour objectif de permettre leurs réutilisations, l'administration en réalisant la diffusion manque de traçabilité sur le champ des possibles réutilisations et sur les dispositifs qui seront mis en œuvre par les réutilisateurs. Ce manque de traçabilité fait peser un risque sur la disponibilité des données. Les objectifs à atteindre sont **l'information des personnes, la traçabilité et la sécurité des données**.

– *Recommandations applicables :*

D'après les facteurs de vulnérabilité identifiés, plusieurs objectifs sont à prioriser.

• **Objectif #1 : la traçabilité des données.**

L'ouverture sans restriction d'accès n'est pas incompatible avec certaines mesures de traçabilité. En analysant le volume des données collectées par les réutilisateurs, le gestionnaire d'API peut lever une alerte en cas d'une demande qui risquerait d'excéder les limites de charge que peut supporter son système.

• **Objectif #2 : la gouvernance et le respect des droits des personnes.**

Dans le cas de données ouvertes, il peut être difficile de faire respecter les droits des personnes auprès des réutilisateurs. Des mesures devraient être prises en amont du partage pour anticiper les demandes d'exercice des droits des personnes.

• **Objectif #3 : la sécurité des données.**

Malgré l'ouverture sans restriction d'accès des données, l'exemple montre qu'en l'absence totale de planification du volume des demandes, un risque peut exister notamment sur la disponibilité des données. Le détenteur de données devrait s'assurer que les mesures adaptées sont prises afin de réduire ce risque en effectuant par exemple un suivi de la charge des serveurs ou du volume de requêtes reçues.

Les mesures relatives à la sécurité devraient être considérées au cas par cas. En effet, bien qu'elles s'appliquent dans la majorité des cas, ces mesures doivent être mises en œuvre selon le niveau de risque évalué.

• **Objectif #4 : la minimisation des données.**

Au vu du contexte du partage et des nombreuses inconnues portant sur la nature des organismes réutilisateurs, sur les réutilisations et sur les techniques d'attaque, des mesures de minimisation devraient être mises en œuvre par précaution.

– **Exemple #4 : Partage fermé de données entre services d'un organisme**

– *Description du contexte :*

Cette typologie de partage est retrouvée dans les secteurs public et privé, avec des cas d'usage particulièrement diversifiés. Dans le secteur privé, les services concernés peuvent être des filiales ou services d'un même organisme, coordonnées par une direction centralisée. Dans le secteur public, certaines données des administrés peuvent être partagées entre services ou établissements d'une administration, par exemple dans l'objectif de dématérialiser les démarches administratives et de faciliter l'accès à certains services et aides. La mise en œuvre de ces partages repose le plus souvent sur l'utilisation d'API.

Ainsi, des services (les détenteurs de données) en lien direct avec les usagers collectent certaines données personnelles et les inscrivent dans une base centralisée de l'organisme dont ils dépendent (le réutilisateur de données (6) et gestionnaire d'API). Cette inscription est réalisée par les services détenteurs de données par le biais des requêtes en écriture via une API modifiant directement la base centralisée de l'organisme central. Cet organisme héberge et traite les données de la base centralisée dans le cadre de ses missions. Elle met en œuvre et assure la gestion de l'API.

Les données de la base centralisée de l'organisme sont mises à disposition par le biais d'une seconde API permettant à d'autres acteurs privés ou publics de traiter les données. Ces acteurs sont par exemple des administrations qui instruisent ainsi des démarches sans avoir à collecter de pièces justificatives auprès des administrés.

– *Recherche des facteurs de vulnérabilité :*

De par la nature de ce traitement, plusieurs facteurs de vulnérabilité sont identifiés **pour le premier partage** :

– sur le type d'accès à la base de données : un accès en écriture dans la base centralisée est prévu pour que les services y inscrivent les données qu'ils collectent. Ce type d'accès permettant de modifier directement la base centralisée, il offre un cadre propice à l'introduction d'erreurs dans la base, ou à la suppression accidentelle de données et pose ainsi un risque sur l'intégrité des données. Les objectifs à atteindre sont **l'exactitude, la traçabilité, la gouvernance, la sécurité des données et le respect des droits des personnes** ;
– sur la nature des services impliqués dans le partage : ils sont nombreux et leur coordination peut poser problème, notamment en cas de perte d'intégrité des données : les alerter un à un présente une certaine difficulté. De plus, lorsque les données mises à disposition sont critiques pour les usagers, tout risque portant sur la disponibilité des données pourrait avoir des conséquences graves. Les objectifs à atteindre sont **l'exactitude, la traçabilité, la gouvernance, la sécurité des données et le respect des droits des personnes** ;
– sur les catégories de données accessibles par l'API : il s'agit ici des données personnelles des usagers, comme des informations sur leur situation personnelle leur permettant d'accéder à des aides et à des services de l'Etat. Tout risque portant sur leur confidentialité, leur intégrité ou leur disponibilité pourrait avoir des conséquences graves pour ces personnes. Les objectifs à atteindre sont la **préservation de la confidentialité et la sécurité**.

– *Recommandations applicables :*

D'après les facteurs de vulnérabilité identifiés, plusieurs objectifs sont à prioriser.

• **Objectif #1 : l'exactitude des données.**

Par nature les requêtes en écriture font peser un risque sur l'exactitude des données. Ce risque peut être accru lorsque les services réalisant l'écriture sont nombreux et peuvent manquer de maturité technique, ce qui est l'hypothèse que nous considérons dans cet exemple. Certains outils permettant de vérifier que les données correspondent au format attendu peuvent être utilisés pour automatiser des vérifications de sécurité

• **Objectif #2 : la gouvernance et le respect des droits des personnes.**

Ici également le nombre et la nature des services impliqués pourraient compliquer les procédures prévues pour l'exercice des droits, pour la notification d'une violation ou d'un nouveau risque, etc. La coordination des services, la documentation des procédures, la gestion des habilitations sont autant de mesures pouvant limiter ces risques. De plus, en prévoyant des techniques automatisées pour permettre aux personnes d'exercer leurs droits, le risque humain pourrait être réduit.

• **Objectif #3 : la traçabilité des données.**

La connaissance des détenteurs de données ici et des actions qu'ils réalisent peut permettre à l'organisme central d'agir rapidement en cas de réalisation d'un risque identifié. La traçabilité est ainsi à prendre en compte pour réagir rapidement à une menace comme une tentative d'intrusion ou une perte d'intégrité des données.

• **Objectif #4 : la minimisation des données.**

Dans le cas d'un partage dans le secteur public, les réutilisations étant généralement prévues par les textes, les catégories de données concernées par le partage sont généralement fixées et il peut sembler difficile d'appliquer le principe de minimisation dans ce contexte. Toutefois, lorsque les réutilisations sont bien connues, il est possible de jouer sur la granularité des données, leur profondeur historique, ou encore sur les mesures de pseudonymisation appliquées.

• **Objectif #5 : la sécurité des données.**

Les mesures relatives à la sécurité devraient être considérées au cas par cas. En effet, bien qu'elles s'appliquent dans la majorité des cas, ces mesures doivent être mises en œuvre selon le niveau de risque évalué.

Une autre analyse du même type devra être menée pour **le second partage** (se référer pour cela à l'exemple 1 qui lui est similaire).

– **Exemple #5 : Partage de données impliquant la personne concernée**

– *Description du contexte :*

Dans un objectif de dématérialisation, différents dispositifs permettent de faciliter la fourniture par un individu de ses informations à un organisme ayant besoin d'y accéder. Ces données peuvent être détenues par un premier organisme dans le cadre d'une relation préexistante, ou conservées dans un espace personnel sécurisé prévu à cet effet et géré par son titulaire. L'usager peut ainsi partager ses documents pertinents (tels que des documents administratifs, attestations de droits, justificatifs de résidence, etc.) avec des destinataires autorisés, par exemple à l'occasion de la fourniture d'un service. Une API est proposée par l'organisme détenteur de ces documents ou le fournisseur de l'espace personnel sécurisé (à la fois détenteurs de données et gestionnaires d'API) afin de permettre à leurs destinataires (les réutilisateurs de données) d'accéder aux documents suite à une demande expresse des personnes. Une authentification des personnes est nécessaire pour que l'accès aux documents soit accordé au fournisseur du service, dont l'identité est également vérifiée par l'organisme détenteur de données. Cet accès est permis par un procédé d'authentification robuste.

– *Recherche des facteurs de vulnérabilité :*

De par la nature de ce traitement, plusieurs facteurs de vulnérabilité sont identifiés :

– sur le niveau de sécurité des techniques d'authentification utilisées : l'authentification des usagers d'une part, mais aussi des organismes fournissant le service d'autre part sont nécessaires pour permettre l'accès aux données. Les mécanismes d'authentification utilisés doivent être suffisamment robustes pour empêcher une usurpation d'identité et un accès illégitime aux données personnelles. Les objectifs à atteindre sont **l'information et la traçabilité, la gouvernance, le respect des droits et la sécurité des données** ;
– sur la nature des organismes impliqués dans le partage : les organismes offrant un service dont l'accès peut être critique pour leurs usagers, tout risque sur la disponibilité de ce service pourrait avoir des conséquences graves. Les objectifs à atteindre sont **l'exactitude, la traçabilité, la gouvernance, la sécurité des données et le respect des droits des personnes** ;
– sur les catégories de données accessibles par l'API : si les documents concernés comportent des données sensibles au sens de l'article 9 du RGPD, toute perte de confidentialité, de disponibilité ou d'intégrité aurait potentiellement de graves conséquences pour les personnes. Les objectifs à atteindre sont **la minimisation et la sécurité des données**.

– *Recommandations applicables :*

D'après les facteurs de vulnérabilité identifiés, plusieurs objectifs sont à prioriser.

• **Objectif #1 : l'information des personnes et la traçabilité des données.**

Au vu de la gravité des conséquences que pourrait avoir une perte de disponibilité, d'intégrité ou de confidentialité des données dans cet exemple, des mesures importantes de traçabilité devraient être prises afin que le détenteur de données mais également la personne concernée puisse vérifier la légitimité des accès. Il est ici possible d'intégrer la personne concernée à cette vérification car les réutilisations sont théoriquement en faible nombre, et généralement initiées par la personne.

• **Objectif #2 : la gouvernance et le respect des droits des personnes.**

Les mesures de gouvernance devraient avoir pour objectif que seuls les organismes vérifiés puissent accéder aux données et cela, dans les conditions prévues. Les habilitations, les accès et la documentation devraient être à la mesure de la gravité des risques d'un accès illégitime aux données, d'un incident de sécurité ou d'une erreur lors du traitement.

Les mesures prises pour que les personnes puissent exercer leurs droits auprès du détenteur et des réutilisateurs devrait permettre à ceux-ci de vérifier que les utilisations de leurs données sont limitées à ce qui est prévu et de s'opposer à certaines utilisations lorsque le droit d'opposition n'est pas exclu.

- **Objectif #3 : l'exactitude des données.**

Au vu de l'impact pour les personnes d'une erreur dans leurs données lors de la fourniture du service essentiel, l'exactitude devrait être garantie.

- **Objectif #4 : la minimisation des données.**

Une perte de confidentialité des données partagées dans cet exemple pourrait avoir des conséquences graves pour les personnes. Ainsi en limitant les données partagées entre les organismes à ce qui est strictement nécessaire, ce risque peut être réduit.

- **Objectif #5 : la sécurité des données.**

Les mesures relatives à la sécurité devraient être considérées au cas par cas. En effet, bien qu'elles s'appliquent dans la majorité des cas, ces mesures doivent être mises en œuvre selon le niveau de risque évalué.

---

(5) Sur la distinction entre la « base active » et l'archivage intermédiaire, se référer à la page « Les durées de conservation des données » du site de la CNIL : https://www.cnil.fr/fr/les-durees-de-conservation-des-donnees.

(6) Le terme de « réutilisateur » fait ici référence au rôle technique introduit dans la recommandation API, et non pas la définition de l'article L. 321-2 du code des relations entre le public et l'administration.