Projet de recommandation technique

RELATIVE A L'UTILISATION DES **INTERFACES DE PROGRAMMATION APPLICATIVES (API)** POUR LE PARTAGE SECURISE DE DONNEES A CARACTERE PERSONNEL

Version soumise à consultation publique
jusqu'au 1er novembre 2022

CNIL.
COMMISSION NATIONALE
INFORMATIQUE & LIBERTÉS

Sommaire

1. Introduction

La Commission a pu observer 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. À 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.

2. 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 entendu 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 capacité 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 ») ou de leurs conséquences. Cependant, le respect de ces recommandations est de nature à contribuer grandement au respect par les acteurs de leurs obligations légales.

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 préjudice 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 proposent des modalités pour 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) dans le cas d'un partage de données impliquant une administration, ou encore les autres recommandations de la CNIL en vigueur.

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 (ou deux) 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.

Le schéma suivant présente l'organisation fonctionnelle entre ces trois types d'acteurs. L'annexe II présente plusieurs cas d'usages fréquemment observés, dans lesquels les rôles de chacun sont identifiés et expliqués.

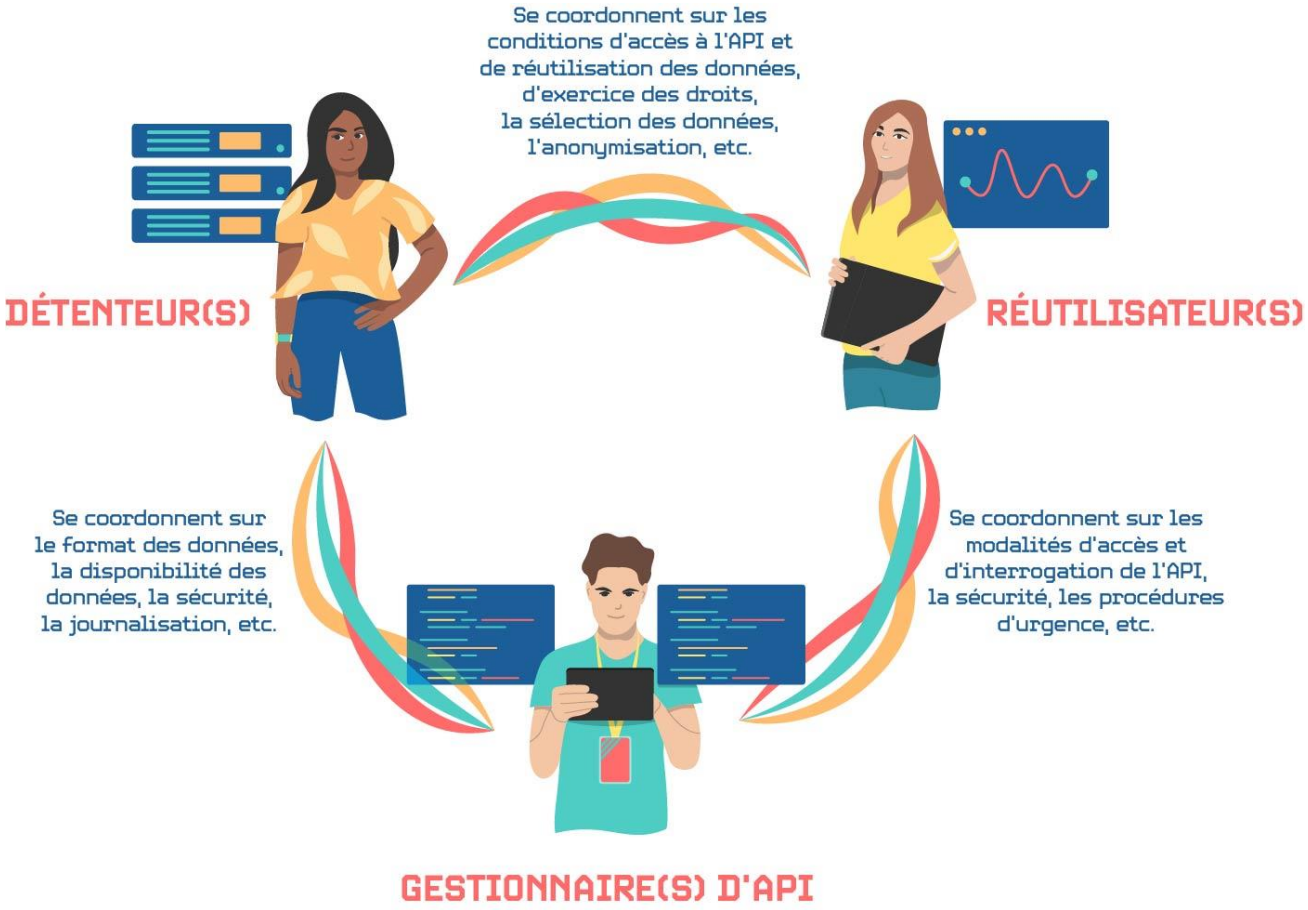

Figure 1 : Schéma relationnel entre les trois rôles de détenteur de données, gestionnaire d'API et réutilisateur

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 d'une exploitation de celles-ci pour son propre compte.

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 de le mettre en œuvre, à quelles fins et par quels moyens ; 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.

3. Recommandations générales

3.1. Sur les motifs justifiant le recours à une API

L'utilisation d'API est à favoriser lorsque :

- les données sont partagées à plusieurs réutilisateurs, et/ou
- elles 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 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).

En effet, la Commission a pu observer 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, la granularité des données accédées 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, etc.).

3.2. Sur les risques liés à l'utilisation d'une API

Le partage de données par API présente toutefois des risques, 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 : en facilitant le partage de données, les API augmentent les risques d'un détournement de finalité ou d'une perte de confidentialité des données ;
- l'exactitude des données source : l'automatisation du partage et des réutilisations des données permise par les API augmente la gravité d'un partage de données inexactes ;
- la traçabilité des accès : en permettant l'ouverture des données à davantage de réutilisateurs, les API peuvent causer une perte de visibilité pour les personnes concernées sur le parcours de leurs données à caractère personnel.
- la gouvernance : en permettant l'accès automatisé aux données et parfois leur modification, le partage de données par API entraîne un risque accru de perte de contrôle sur les données ou sur leur accès ;
- la sécurité : en augmentant la possibilité de partager des données, l'utilisation des API augmente la surface d'attaque sur les données et ainsi les mesures de sécurité à mettre en œuvre par chacun des acteurs.

Les risques précédents sont à considérer dans le contexte du partage de données afin d'évaluer leur vraisemblance et la gravité de leur impact. Les critères 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 afin de réaliser cette analyse :

- type d'accès à la base de données : en lecture seule ou en écriture ;
- conditions d'accès aux données : si l'accès est soumis à une demande, quelles vérifications sont apportées afin de valider les 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 ;
- type de données accessibles par l'API : des données sensibles ou bancaires sont plus susceptibles de faire l'objet d'attaques ;
- granularité des données : possibilité d'accéder uniquement à certaines informations ciblées.

Cette grille d'analyse devrait être considérée par chacun des acteurs concernés par le partage de données, quel que soit leur rôle. Les résultats de l'analyse devraient être documentés. 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 par l'analyse précédente 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.

3.3. 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 quel 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.

3.4. 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 granularité devrait être adaptée notamment à la gravité des conséquences pour les personnes que pourrait entraîner une utilisation détournée de l'API ou un accès illégitime aux données. Dans certains cas, les informations fournies à la personne devraient donc inclure la liste exhaustive des accès et modifications apportées aux données sur la période pertinente, ainsi que leur horodatage. 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 publiques 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 c'est pertinent. 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 significativement.

La Commission recommande que les informations précédemment citées soient présentées sur un site web connu 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, l'accès à ces informations ou à cette interface devra être sécurisé au moyen d'une méthode d'authentification conforme aux recommandations de la Commission, telle que France Connect dans la sphère publique.

3.5. 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, lorsque cela est possible, le détenteur de données à mener une réflexion avec les réutilisateurs des données sur les données strictement nécessaires aux réutilisations prévues, pour limiter le partage à ces seules données. À cet égard, elle souligne les éléments suivants.

Les catégories de données, leur format, leur historicité, 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.

Un cloisonnement physique ou logique devrait être réalisé entre les données dont le partage est prévu et la base contenant les données sources, lorsque celle-ci contient d'autres données collectées licitement mais dont le partage n'est pas prévu ou autorisé et dont la violation pourrait entraîner des conséquences pour les personnes concernées.

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, explicite 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 du paragraphe précédent 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, lorsque c'est nécessaire, des personnes physiques appartenant à l'organisme réutilisateur selon leur niveau d'habilitation. Ces mesures devraient être intégrées directement dans l'API par le gestionnaire lorsque cela est possible.

Enfin, une phase d'expérimentation en coopération avec le gestionnaire d'API et les réutilisateurs est recommandée, afin de confirmer la pertinence des choix faits et le respect des recommandations énoncées dans les deux 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.

3.6. Sur l'exercice des droits sur le partage des données

Les droits des personnes concernées sur 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.

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, en particulier de manière automatisée lorsque cela est possible et pertinent.

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.

3.7. Sur la gestion des accès

Lorsque l'API est soumise à une restriction d'accès (« API en accès restreint »), la Commission recommande qu'un formulaire de demande d'accès soit mis en place, dans lequel chaque réutilisateur apporterait les informations nécessaires à la vérification de la licéité de son accès à l'API. Dans ces cas, une mesure de contrôle d'accès complémentaire devrait être mise en place en limitant l'accès à l'API aux seules adresses IP déclarées par les réutilisateurs, qu'elles soient virtuelles ou non, lorsque cela est possible. 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, et des catégories de données nécessaires. La Commission recommande que le réutilisateur informe le détenteur du volume, de l'historicité, 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 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 évalué 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 à ce dernier. 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. 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 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 son service devrait lui être fourni.

3.8. Sur la gestion interne des API

La Commission recommande qu'une gouvernance dédiée soit mise en place 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 comporter également une description technique permettant l'intégration, le développement, la mise à jour et l'obsolescence 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, lorsque cela s'avère pertinent. 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. 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 de secrets devrait être privilégiée.

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.

4. Recommandations spécifiques à l'intention des détenteurs de données

4.1. 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 historicité, leur fiabilité et les procédés de pseudonymisation ou d'anonymisation utilisés. Le détenteur de données devrait s'assurer que les réutilisateurs aient 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, lorsque cela est pertinent, 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. À 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.

4.2. 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 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 utilisés.

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. En particulier lorsque la réutilisation des données partagées par voie d'API peut entraîner des conséquences pour les personnes, 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 de l'écriture.

4.3. 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 entrainer 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 cloisonnées des autres catégories de données. En particulier, lorsqu'un procédé de pseudonymisation ou d'anonymisation est prévu, les données brutes devraient être physiquement ou logiquement cloisonnées des données issues de ce procédé. La base source de l'API pourra alors être alimentée par un procédé d'export de données automatisé, dont la sécurité devrait être vérifiée.

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 nécessaire.

Journalisation

Le détenteur de données devrait s'assurer qu'une journalisation effective des accès et actions réalisées sur la base de données ait 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 et des informations contenues dans les traitements 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.

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. Dans ce cas, le détenteur des données pourrait ainsi alerter les réutilisateurs concernés. Les informations issues de la journalisation permettent également d'assurer une traçabilité sur les accès effectifs aux données qui pourrait être fournie 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.

5. Recommandations spécifiques à l'intention des gestionnaires d'API

Le gestionnaire d'API est le plus 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 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.

5.1. 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 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 chacun de signaler un problème concernant la sécurité de l'API. Un moyen de communication réactif devrait être choisi pour traiter les demandes urgentes.

Enfin, la Commission recommande que le code de l'API soit ouvert autant que possible, afin de recueillir les retours du public à son égard.

5.2. Sur la minimisation

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 aux 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 devraient porter sur le volume, l'historicité ou encore 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 de la base devraient être mises en œuvre par le gestionnaire d'API. La robustesse de ces méthodes devrait notamment être vérifiée, en particulier vis-à-vis de méthodes de réidentification telles que le 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.

5.3. 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 aux paragraphes 36 à 40.

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

5.4. 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 sur les recommandations de la CNIL en vigueur, telles que son guide de la sécurité des données à caractère personnel. La Commission recommande par ailleurs que le gestionnaire d'API se conforme aux normes techniques communément admises telles que le référentiel général de sécurité (RGS) de l'ANSSI 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 pour la mise en œuvre des API.

Communications

En particulier dans le cas d'une API en accès restreint, le chiffrement appliqué aux communications devrait garantir un haut niveau de confidentialité et d'intégrité. Pour rappel, la Commission considère que les règles et recommandations décrites dans les annexes B1 et B2 du référentiel général de sécurité (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 prémunissant 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 uniques 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 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. 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 aux paragraphes 61 à 64.

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 de plus permettre de détecter les surcharges du système et l'indisponibilité des données. Ces analyses devraient donner lieu à des signalement 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.

6. Recommandations spécifiques à l'intention des réutilisateurs

Le réutilisateur de données devrait prendre connaissance des instructions à sa disposition concernant l'utilisation et la sécurité de l'API. Il lui appartient d'appliquer rigoureusement ces instructions 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. Il devrait mettre en œuvre les mesures de sécurité adaptées afin de garantir la confidentialité et l'intégrité des données. Lorsque l'API est une brique technique visant à permettre des échanges de données nécessaires au fonctionnement d'un service, tel qu'un service d'authentification déporté reposant sur le protocole OAuth par exemple, le réutilisateur devrait s'assurer de l'efficacité du service et de la robustesse de son intégration dans son système.

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.

6.1. Sur l'information des personnes concernées

L'information des personnes concernée sur le partage de leurs données auprès de 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, lorsque cela est possible et pertinent, que ce dernier informe les personnes concernées sur la volumétrie et la fréquence des requêtes, sur les mesures de minimisation prises et le niveau de sécurité appliqué aux données.

6.2. 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.

De façon générale, et dans la mesure du possible, le réutilisateur de données devrait interroger l'API à 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 après leur traitement.

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.

6.3. 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 présentées aux paragraphes 61 à 64 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.

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 l'échange continu de données.
Détenteur de données : tout organisme public ou privé en possession de données, ici à caractère personnel, 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 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. 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 public ou privé ayant accès aux données partagées par l'API.
Partage de données par le biais d'une plateforme : utilisation d'un 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.
Partage de données par le biais d'un service de télécommunication : utilisation d'un service tiers, tel qu'un service de messagerie électronique, pour partager des données.
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 (voir les définitions correspondantes) ou encore PATCH, qui permet de mettre à jour partiellement des données, et DELETE, qui permet de supprimer des données.
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 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.
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 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 « méta-donnée ») 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.
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

Les cas d'usage présentés dans cette annexe visent à expliquer la répartition des rôles décrits dans cette recommandation. La liste proposée n'est pas exhaustive mais correspond aux cas d'usage les plus fréquemment rencontrés en pratique.

A. Trois organismes distincts pour les trois rôles

Dans ce premier cas d'usage, trois organismes différents tiennent les rôles de détenteur de données, gestionnaire d'API et réutilisateur de données. Les recommandations relatives à chacun de ces rôles s'appliquent respectivement à chacun des trois organismes. Toutefois, une coopération entre les organismes est attendue afin de permettre la meilleure mise en œuvre de ces recommandations.
Exemple de cas concret : accès aux données bancaires par un service d'information sur les comptes.
Sous réserve du respect du cadre juridique applicable, certains services d'informations sur les comptes C proposent aux clients de banques A d'accéder à leurs données bancaires pour en tirer certaines informations sur leurs dépenses. Le service d'information C peut dans ce cadre interroger la banque A sur les données bancaires de la personne par voie d'API. L'API en question peut être déployée par une entreprise tierce B, jouant le rôle d'agrégateur, qui réalise le lien entre les deux organismes. On retrouve le schéma suivant :
Détenteur de données : banque A.
Gestionnaire d'API : entreprise tierce B.
Réutilisateur de données : service d'information C.

B. Un organisme à la fois détenteur de données et gestionnaire d'API

Ce second cas d'usage illustre la situation où un organisme ouvre des données à sa disposition grâce à une API qu'il met en œuvre lui-même. Il dispose des moyens techniques lui permettant de mettre en œuvre l'API et se charge d'en donner l'accès aux réutilisateurs. Ce cas correspond à celui d'un réseau social ouvrant les données de ses usagers à des fins de recherche scientifique, ou encore d'une administration partageant ses données avec d'autres conformément aux textes relatifs aux échanges de données d'usagers entre administrations pour le traitement des demandes du public ou des déclarations transmises par celui-ci.
Exemple de cas concret : vérification de la situation d'un candidat pour son inscription à un concours soumis au contrôle de l'autorité publique vis-à-vis des obligations de service national.
L'API « Service national », mise en œuvre directement par le ministère des armées, permet de vérifier si un candidat est en règle vis-à-vis de ses obligations de service national. Les administrations en charge de la gestion des inscriptions à un concours nécessitant que les candidats aient effectué leur journée « défense et citoyenneté » peuvent ainsi être amenées à utiliser cette API lors de l'inscription des candidats. Le schéma est ainsi le suivant :
Détenteur de données : ministère des armées.
Gestionnaire d'API : ministère des armées.
Réutilisateur de données : toute administration chargée de la gestion des inscriptions au concours.

C. Un organisme à la fois gestionnaire d'API et réutilisateur

Dans ce troisième cas, un organisme principal cherche à collecter les données de plusieurs détenteurs et dispose des moyens techniques nécessaires pour la mise en œuvre d'une API. Il

est en charge de la conception de l'API et collabore avec les détenteurs de données pour que ceux-ci connectent leur base de données à l'API. Ce cas est rencontré lorsque les détenteurs de données sont structurellement de moindre taille que le réutilisateur, comme des collectivités, des succursales d'une entreprise, ou encore des particuliers.
Exemple de cas concret : collecte des données des utilisateurs d'ordiphones relatives à leur utilisation par les services d'exploitation.
Les fournisseurs de système d'exploitation des ordiphones utilisent des API afin de collecter les données de leurs clients relatives à leur utilisation de ces appareils. Les données des utilisateurs sont ensuite traitées et hébergées par le fournisseur du système d'exploitation en vue d'une réutilisation, par exemple à des fins d'amélioration des services proposés ou de ciblage publicitaire. Le schéma est ainsi le suivant :
Détenteur de données : les utilisateurs des ordiphones.
Gestionnaire d'API : le fournisseur du système d'exploitation.
Réutilisateur de données : le fournisseur du système d'exploitation.

D. Plusieurs gestionnaires d'API

Dans ce dernier cas d'usage, plusieurs gestionnaires d'API participent à l'élaboration des composants techniques sur lesquels repose le partage des données. Que chacun des organismes participe au développement du même outil technique ou qu'ils développent plusieurs outils qui, mis bout à bout, assurent le partage des données, chacun d'entre eux est considéré comme gestionnaire d'API dans la présente recommandation. Ce cas s'observe lorsqu'un organisme adapte une API tierce à son cas d'espèce, ou encore lorsque les réutilisations possibles des données nécessitent des traitements intermédiaires.
Exemple de cas concret : utilisation de statistiques issues des réseaux sociaux pour la recherche.
Dans le cas considéré, des chercheurs souhaitent utiliser les données des réseaux sociaux à des fins scientifiques et collectent des données statistiques auprès d'un prestataire missionné par les chercheurs. Le prestataire utilise l'API mise à disposition par le réseau social pour réaliser une nouvelle API spécifique, afin de permettre, de manière automatisée, une analyse des données collectées sans constituer une base de données sur son infrastructure technique.
Bien que ce schéma puisse parfois être divisé en deux partages (entre le réseau social et le prestataire, puis entre le prestataire et les chercheurs), le prestataire est ici uniquement considéré comme un gestionnaire d'API pour plusieurs raisons :

- il ne conserve pas de base de données ;
- les requêtes proviennent toujours des chercheurs ;
- il applique les règles d'accès aux données définies par le réseau social (en ajoutant éventuellement une souscription payante à ses services).

Ainsi, le schéma est le suivant :
Détenteur de données : le réseau social.
Gestionnaires d'API : le réseau social et le prestataire.
Réutilisateurs de données : les chercheurs.