Authentification biométrique et respect de la vie privée : points de vue et limitations

Estelle Cherrier-Pawlowski [1]

Le respect de la vie privée dans les systèmes d'authentification est intrinsèquement lié à la sécurité des données et à l'utilisabilité du système. Le respect de la vie privée englobe à la fois la protection des données personnelles, qui peut être reliée à leur sécurité, via la garantie de leur confidentialité et leur intégrité notamment. D'autre part, si les individus sont confiants dans le niveau (perçu) de protection de leurs données par un système (d'authentification biométrique), cette confiance contribue indiscutablement à l'utilisabilité de ce système.

Cet article propose d'appréhender la notion de vie privée selon différents points de vue : historique, éthique, informatique, réglementaire, avec un focus sur les données biométriques, données personnelles particulièrement sensibles. Une discussion sur les limitations des schémas actuels d'authentification biométrique protégeant la vie privée est proposée à la fin de l'article.

Différents points de vue sur la notion de vie privée

On peut distinguer deux types d'approches pour définir la protection de la vie privée : une première, qui tente de « protéger l'identité de l'utilisateur », et une seconde qui se concentre sur la « protection des données de l'utilisateur ». Dans la première approche, la protection de la vie privée est obtenue en empêchant un adversaire de faire le lien entre les données correspondant à un individu et son identité. Dans le

1. Maîtresse de conférences, HDR, laboratoire GREYC, Ensicaen.

DOI:10.48556/SIF.1024.20.145 **1024** – Bulletin de la SiF, numéro 20, Novembre 2022, pp. 145–160

second cas, l'identité du propriétaire des données est supposée être connue de l'adversaire. La protection des données est alors obtenue en modifiant les informations divulguées à l'adversaire, en les transformant, par exemple en les remplaçant par une valeur perturbée, en augmentant leur granularité, en y ajoutant des résultats fictifs, etc. On trouve des propositions pour les deux approches dans le cadre de la protection des données de localisation, particulièrement sensibles, dans la thèse de Nicolás E. Bordenabe [1].

La notion même de vie privée possède de nombreuses acceptions, qui toutes dépendent d'un contexte historique, social, culturel, légal. On distingue plusieurs types de vie privée. Chaque type est centré sur un aspect particulier : sur le corps (*bodily privacy*, correspondant à l'intégrité physique), sur les lieux de vie (*locational privacy*, correspondant à la protection du foyer, de la maison), sur les relations interpersonnelles (*relational privacy*, correspondant à la protection de la vie de famille, de la vie sociale) et enfin sur les données (*informational privacy*, correspondant à la protection des données personnelles ainsi qu'au secret des correspondances). Plus de détails sur ces différentes catégories peuvent être retrouvés dans le livre de B. Roessler [21].

La part de la vie privée impliquée dans l'authentification biométrique est par nature en lien avec sa dimension informationnelle. Une typologie plus approfondie de la vie privée au sens de la protection des données personnelles est illustrée par la figure 1, extraite de [15].

De façon générale, un défi se pose lorsqu'on réfléchit à la question : comment mesurer le respect de la vie privée dans un système ? L'article de Wagner et Eckoff [23] propose un tour d'horizon complet des différentes métriques pour évaluer la protection de la vie privée. Ces métriques sont classées selon quatre caractéristiques : les modèles de l'attaquant, les sources de données, les entrées et les sorties du système.

La vie privée est souvent définie par opposition à la vie publique. L'étude de cette notion déborde largement du cadre législatif. En effet elle intéresse de nombreux domaines, comme la sociologie, la médecine ou l'histoire, en passant par l'économie, l'éthique, et l'informatique bien sûr. Cette diversité de traitements rend quasi impossible l'établissement d'une définition universelle. On peut noter toutefois deux termes qui semblent constituer un socle commun, à savoir l'*accès* et le *contrôle*, comme on le voit sur la figure 1.

Dans le cadre de cet article, je vais m'attarder sur plusieurs éclairages reflétant la pluralité des définitions de la vie privée : l'évolution historique du concept de vie privée, des considérations éthiques, la vision informatique, puis réglementaire. Je m'intéresserai enfin aux données biométriques.

Point de vue historique

Cette section retrace en quelques dates et faits marquants les variations de la notion de vie privée au cours des siècles.

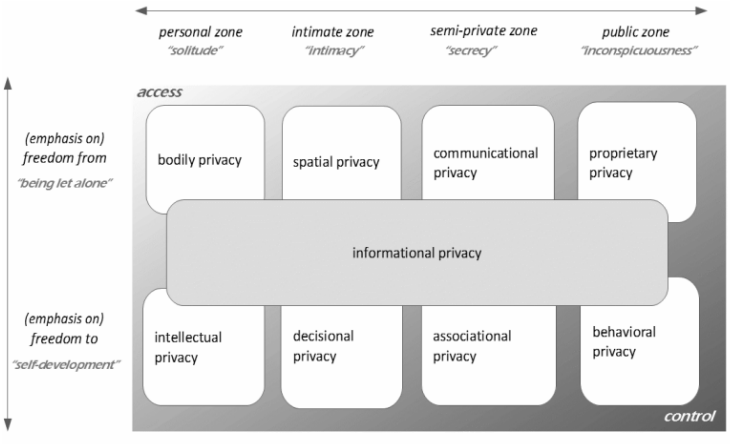


FIGURE 1. Typologie de la vie privée informationnelle [15]

Avant 1500, la notion même d'individualité n'existe pas. Seule la distinction entre la sphère familiale (donc privée, *oikos* en grec) et la sphère publique (principalement les activités politiques, *polis* en grec), développée par Aristote au IV^{e} siècle av. JC, existe.

A partir du XVI^{e} siècle, jusqu'à la Révolution française, le respect de la vie privée recouvre l'acception suivante : autonomie personnelle et individualité. L'émergence d'une classe moyenne ainsi que l'invention de l'imprimerie induisent le développement du travail intellectuel ainsi que la nécessité de disposer d'un endroit calme, privé. Même si la notion de vie privée revêt une connotation négative chez Shakespeare, en raison de son lien avec la solitude, ou encore avec les conspirations politiques, le philosophe anglais John Locke (1632-1704) — un des pères de l'*État de droit* — met à jour le but légitime des individus à protéger leur vie privée, leur liberté et leur propriété.

Au XIX^{e} siècle, les nouvelles technologies de l'époque sont en plein essor : courrier postal, télégraphe, téléphone, enregistrements audio, photographie, etc. On voit alors émerger deux visions de la notion de vie privée : (i) celle de l'état de surveillance, omniscient ; (ii) celle du citoyen, en tant qu'idéal et aspiration de chacun. La première menace directement la vie privée des citoyens, dans le sens où la discipline règne, avec comme point d'orgue le panoptique de Bentham. La seconde vision, celle des citoyens, fait suite aux différentes Révolutions (française, américaine) et Déclarations des droits (*Bill of Rights* en 1689 en Angleterre et Déclaration des droits de l'homme et du citoyen en France en 1789).

Cette tendance se confirme au XX^{e} siècle : en réaction aux deux guerres Mondiales et à la fin du colonialisme, les libertés et les droits fondamentaux sont réaffirmés. La Déclaration universelle des droits de l'homme est ainsi proclamée par l'Assemblée générale des Nations unies le 10 décembre 1948. Elle revêt uniquement une valeur déclarative. L'article 12 concerne la vie privée Un traité international est ensuite signé le 4 novembre 1950 par les états membres du Conseil de l'europe et entre en vigueur le 3 septembre 1953 : la Convention européenne des droits de l'homme (CEDH, aussi appelée Convention de sauvegarde des droits de l'homme et des libertés fondamentales). Afin d'assurer le respect des engagements de la CEDH, une juridiction internationale est instituée en 1959 à Strasbourg : la Cour européenne des droits de l'homme.

A partir des années 1970, la vie privée s'entend à l'ère des ordinateurs comme la protection, la confidentialité des informations (*information privacy*). Depuis les attaques du 11 septembre 2001 et la lutte contre le terrorisme qui s'est renforcée ensuite, la vie privée est un droit fondamental que les états s'autorisent à rogner sous prétexte de renforcer la sécurité des citoyens. Contre l'argument « *personne ne peut être contre les mesures de sécurité s'il n'a rien à cacher, rien à se reprocher* », Edward Snowden, lanceur d'alerte américain rétorque [2] : « *Dire que votre droit à la vie privée importe peu car vous n'avez rien à cacher revient à dire que votre liberté d'expression importe peu, car vous n'avez rien à dire. Car même si vous n'utilisez pas vos droits aujourd'hui, d'autres en ont besoin. Cela revient à dire : les autres ne m'intéressent pas* ». Il apparaît donc que la notion même de vie privée est un concept fluctuant au cours de l'histoire, en constante évolution donc, qui ne peut être étudié hors d'un contexte social, politique, économique et international.

Point de vue éthique

Comme mentionné dans le livre édité par Bart van der Sloot et Aviva de Groot [6, chapitre 5], un point de vue éthique sur la vie privée permet de répondre aux questions suivantes : quelle est la valeur de la vie privée ? Quelles normes concernant la vie privée doivent être respectées par les individus, la société, l'état ? Même si certains chercheurs renoncent à figer une définition précise — forcément réductrice — de la notion de vie privée, un grand nombre de définitions incluent les termes *accès* et *contrôle*.

Tout d'abord on peut citer la définition de Reiman [20] : « *privacy is the condition in which others are deprived of access to you* ». Le terme *accès* signifie soit un accès physique à une personne, soit un accès à des informations sur cette personne. Le point de vue éthique s'intéresse alors (i) à la manière dont cet accès est obtenu ; (ii) à quoi exactement on a accès. Cependant Fried [10] a montré que l'*accès*, seul, ne suffit pas. Il faut lui adjoindre le *contrôle*, ceci pour éviter le cas absurde d'un homme isolé sur une île déserte : effectivement, personne n'a accès à lui, mais la

2. *Nothing to hide*, documentaire produit par Mihaela Gladovic et Marc Meillassoux, 2017.

notion de vie privée n'a aucun sens dans ce contexte. La définition de Fried relie les deux notions : le respect de la vie privée consiste à contrôler l'accès à soi. En fait, l'accès à une personne ou à des informations personnelles n'est pas à rejeter en bloc : l'important est de pouvoir contrôler cet accès.

On peut remarquer que, d'une part, ces définitions se réfèrent à l'article *The Right to Privacy* de Warren et Brandeis [25] ; d'autre part, elles mettent en exergue la question éthique suivante : est-ce que les protections légales qui existent sont suffisantes pour protéger les droits individuels, en particulier le contrôle exercé par chaque individu sur sa vie privée ? Cette question est d'autant plus cruciale que les technologies actuelles tendent à défier les normes sociales. Les travaux de Beate Roessler [6, p. 137] proposent une réponse selon trois dimensions de la vie privée :

— la dimension locale : comment chacun contrôle l'accès à ses espaces physiques (par exemple, la clé sur la porte d'entrée de sa maison) ;
— la dimension informationnelle (le terme *informations* renvoie souvent aux *données*) : contrôle sur ce que les autres peuvent savoir de nous ;
— la dimension décisionnelle : contrôle sur l'accès (symbolique) à nos décisions personnelles, nos valeurs, nos objectifs, les raisons de nos décisions, etc.

Selon B. Roessler [21], ces trois dimensions de la vie privée doivent actuellement faire face à des défis.

Défis de la dimension locale. Les objets connectés sont par définition connectés à Internet, sphère publique et totalement ouverte. Simultanément, ils envahissent notre espace privé, ce qui n'est pas inquiétant en soi. Ce qui est potentiellement dangereux, c'est la collecte, le stockage, l'analyse, etc., à l'intérieur même de nos maisons, de nos données privées par des assistants personnels, des compteurs électriques ou autres radiateurs connectés, alors que notre espace physique privé était à peu près impénétrable jusque-là. Ces objets dits *intelligents* représentent également une menace potentielle pour les autres dimensions informationnelle et décisionnelle de notre vie privée.

Défis de la dimension informationnelle. Comme on vient de le voir, la collecte de données dans des domaines d'activité variés est de plus en plus importante. Couplée à des algorithmes d'apprentissage et de fouille de données toujours plus efficaces, cette collecte mène à la création de profils qui peuvent être utilisés dans différents contextes ; ainsi nos identités (numériques) ne peuvent plus être cloisonnées. La persistance des traces informatiques constitue une autre source d'inquiétude : dès qu'une donnée est créée, elle peut aussitôt être partagée très facilement et exploitée dans un contexte inconnu, dans un but inconnu. Dans son best-seller *L'Âge du capitalisme de surveillance* [26], Shoshana Zuboff retrace l'histoire de l'*économie de surveillance* : celle-ci repose sur un principe de subordination et de hiérarchie. Elle met en évidence le projet consistant à extraire une plus-value de nos agissements sur

Internet. L'auteur dénonce que Google a découvert que nous avions moins de valeur que les pronostics que d'autres font de nos agissements : « *L'industrie numérique prospère grâce à un principe presque enfantin : extraire les données personnelles et vendre aux annonceurs des prédictions sur le comportement des utilisateurs. Mais, pour que les profits croissent, le pronostic doit se changer en certitude. Pour cela, il ne suffit plus de prévoir : il s'agit désormais de modifier à grande échelle les conduites humaines.* »

Défis de la dimension décisionnelle. Pour continuer sur la même thématique, le constat suivant est un doux euphémisme : le commerce de nos données, à travers l'essor du *Big Data*, de l'*IoT*, des réseaux sociaux augmente le risque que ces technologies nous influencent, manipulent nos comportements. En 2014, Facebook a lancé une étude sur la manipulation des émotions, sans le consentement de ses utilisateurs. D'un point de vue éthique, le problème n'est pas le sujet de l'étude, mais les conditions dans lesquelles elle a été réalisée, sans respecter les normes scientifiques et encore moins éthiques, comme le montre l'analyse de Flick dans son article [9]. De la même façon, les média et réseaux sociaux aujourd'hui sont accusés de scléroser la pensée des internautes, en leur suggérant uniquement des contenus qui les intéressent déjà. On se retrouve piégé dans de véritables *bulles de filtres*. La curiosité intellectuelle n'est pas encouragée, impactant de fait la dimension décisionnelle de notre vie privée.

Point de vue informatique (TIC)

L'informatique fait partie des technologies de l'information et de la communication (TIC) en permettant notamment le traitement, le stockage et la communication de données, y compris de données personnelles. La multiplication des collectes, des flux, des traitements de ces données personnelles impacte nécessairement la vie privée des individus. En fait, c'est la relation faite entre des données capturées et l'identité des utilisateurs qui représente un risque pour la vie privée. La vie privée numérique est intrinsèquement liée à la sécurité numérique, dans le sens où elle relève de la confidentialité des données. Avec la prolifération des objets connectés, des applications mobiles, les risques sur la sécurité, et en particulier sur la vie privée, se multiplient. On est donc ramené au sujet plus général de la gestion des risques en sécurité informatique. Cependant, aucune méthode générique d'analyse des risques sur la vie privée n'existe, contrairement aux méthodes d'analyse des risques de sécurité. Il est donc très difficile d'évaluer le niveau de respect de la vie privée propre à un système, comme il est très difficile de comparer deux systèmes en matière de respect de la vie privée. L'article de Wagner et Eckoff [23] le prouve en recensant plus de quatre-vingt métriques disponibles sur le sujet.

Les données des utilisateurs peuvent être dans trois états selon [6, chapitre 5] : (i) les données stockées, ou statiques (i.e. *data at rest*, par exemple dans une archive, ou

une base de données peu consultée) ; (ii) les données utilisées, ou en cours d'utilisation (i.e. *data in use*, par exemple dans une base de données fréquemment modifiées par de multiples utilisateurs, ou des données en cours de traitement) ; (iii) les données en mouvement (i.e. *data in motion*, lorsque les données sont transférées via des réseaux informatiques à l'extérieur du système, ou sont utilisées par des services tiers). Les données ne sont pas soumises aux mêmes risques dans chacun des trois états.

Avec l'essor des ordinateurs personnels, puis le développement d'Internet et aujourd'hui l'essor des *smartphones* et des objets connectés dits *intelligents*, de véritables défis sont posés dans le domaine informatique, par la société ou la technologie, comme l'éthique de l'intelligence artificielle (dans les véhicules autonomes par exemple), l'éthique du *Big Data* (comment éviter les dérives dans les traitements de données massives), le développement de technologies respectueuses de la vie privée (ou *Privacy Enhancing Technologies*, PET).

On peut donc constater que le problème du respect de la vie privée est de plus en plus prégnant dans notre société : les nouvelles technologies représentent un défi toujours renouvelé pour les normes sociales. La question suivante se pose : est-ce que les protections légales existantes sont suffisantes pour protéger les droits des individus ? Le paragraphe suivant va donner quelques éléments de réponse.

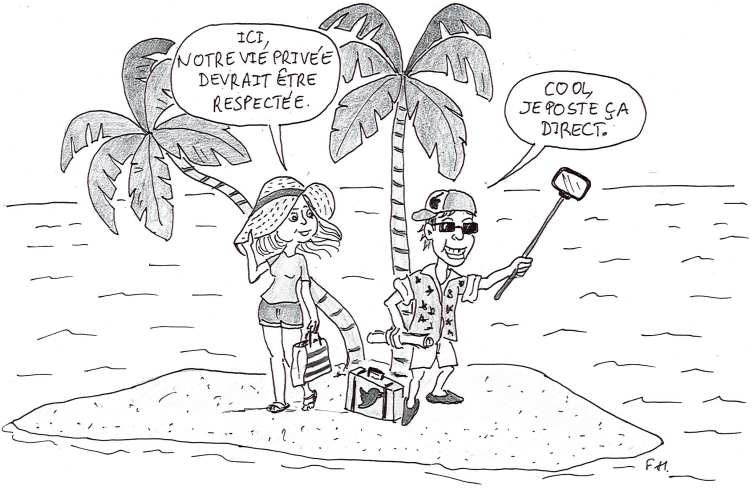
La réglementation sur le respect de la vie privée

Le respect de la vie privée, un droit fondamental

La définition de la vie privée constitue sans doute le plus ancien principe juridique, en tant que séparation entre la sphère publique et la sphère privée. Son étymologie vient du latin *privare* (priver de), et par extension signifie soustraire quelque chose du domaine public. Pendant longtemps, la sphère publique a été gérée par le roi, ou le législateur de façon générale. La sphère privée, quant à elle, était gérée par le *pater familias*, le père de famille, qui régnait sur le foyer, le ménage. Le droit au respect de la vie privée était inclus dans les lois constitutionnelles de certains pays dès le XIII^{e} siècle (en ce qui concerne l'intégrité physique, les lieux privés, les relations privées et le secret des correspondances). Au XX^{e} siècle, le droit à la vie privée figure dans l'article 12 de la Déclaration universelle des droits de l'homme des Nations unies de 1948. Ce droit figure également dans les articles 7 (*« respect for private and family life »*) et 8 (*« protection of personal data »*) de la Convention européenne des droits de l'homme de 1950 [3]. La Charte des droits fondamentaux de l'Union européenne, proclamée en 2000, comporte cinquante-quatre articles consacrant les droits fondamentaux des personnes, y compris les droits sociaux. Elle englobe le droit à la vie privée. Elle est contraignante pour les pays de l'UE dans le sens où elle se voit reconnaître une valeur constitutionnelle par le Traité de Lisbonne

3. https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/005.

en 2007. En France, c'est le Code Civil qui reconnaît le droit à la vie privée : selon l'article 9 alinéa 1, « *Toute personne a droit au respect de sa vie privée* ». Comme le droit à l'image ou le droit à l'honneur, le droit à la vie privée fait partie des *Droits de la personnalité* : ce sont des droits inhérents à la personne humaine, qui la protègent en interdisant toute atteinte à ses droits les plus fondamentaux (sa vie, sa dignité, son corps).



Dans la suite de cet article, seuls les aspects informatiques de la vie privée ainsi que la protection des données personnelles seront considérés, qui concernent directement les systèmes biométriques. C'est au niveau européen que la réglementation évolue : tout d'abord, le principe de *Privacy by Design* (PbD) (ou respect de la vie privée dès la conception), puis le règlement général sur la protection des données (RGPD).

Privacy by design et RGPD

Le *Privacy by Design* a été élaboré par Ann Cavoukian (commissaire à l'information et à la protection de la vie privée de l'état d'Ontario) vers la fin des années 1990 au Canada. Ce concept, détaillé dans le document [2], est une réponse à l'automatisation du traitement des données personnelles et à la quantité de plus en plus importante de données manipulées par les entreprises (ne nécessitant souvent pas une intervention humaine). Le *Privacy by Design* est une mesure préventive ayant donc pour but de limiter les risques d'abus et de violation des données : il implique une prise en compte de la protection de la vie privée des utilisateurs avant même la conception d'un système impliquant le traitement de données personnelles.

Le *Privacy by Design* énonce sept principes [4] :

— conception de mesures préventives et proactives ;

4. https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf.

— protection par défaut (*Privacy by default*) ;
— prise en compte des règles sur la protection de la vie privée dans la conception des produits et durant leur utilisation ;
— protection optimale et intégrale ;
— sécurité de bout en bout – protection du cycle de vie ;
— visibilité et transparence ;
— respect de la vie privée des usagers ou des cibles du service.

Le *Privacy by Default* est un principe qui vise tout produit ou service rendu public. Les standards en matière de protection des données personnelles s'appliquent alors par défaut, et ce sans recours à des procédures extérieures permettant cette protection. Le *Privacy by Default* est la garantie d'un niveau maximal de protection des données personnelles.

Les principes énoncés dans le *Privacy by Design* ont largement été repris dans le RGPD : en mai 2018, le règlement général sur la protection des données [8] vient remplacer la directive 95/46/CE du Parlement européen et du conseil du 24 octobre 1995, relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données. La définition d'une donnée personnelle est clairement posée par la CNIL [5] :

Définition 1 (Donnée à caractère personnel)**.** *Une donnée à caractère personnel est toute information se rapportant à une personne physique identifiée ou identifiable. Mais, parce qu'elle concerne des personnes, celles-ci doivent en conserver la maîtrise.*

Parmi les données personnelles, certaines données sont considérées comme particulièrement sensibles. Le RGPD interdit de recueillir ou d'utiliser ces données, sauf, notamment, si la personne concernée a donné son consentement exprès (démarche active, explicite et de préférence écrite, qui doit être libre, spécifique, et informée). Ces exigences concernent par exemple les données de santé, les opinions politiques, les convictions religieuses, ou encore les données biométriques. Par conséquent, en tant que donnée personnelle sensible, toute donnée biométrique devrait être protégée et ne devrait pas être stockée, ni transiter entre deux éléments d'un système biométrique sous sa forme brute, ni sous une forme qui permet de reconstruire (une partie de) la donnée originale.

La recherche en biométrie est particulièrement concernée par l'article 35 du RGPD, qui introduit la notion d'analyse d'impact relative à la protection des données (AIPD, ou DPIA pour *Data Protection Impact Assessment*). Cette analyse doit être réalisée dès lors qu'un traitement sur des données est susceptible d'engendrer un risque élevé pour les droits et libertés des personnes concernées. Toutefois, il est important de souligner que les données biométriques, dans un contexte de recherche

5. https://www.cnil.fr/fr/definition/donnee-personnelle.

publique, ont un statut spécial : le RGPD prévoit quelques options, offrant des latitudes aux pays membres de l'UE, notamment en ce qui concerne les données de recherche. La France a choisi d'assouplir les contraintes, pour respecter aux mieux les objectifs de recherche. Ainsi, la durée de conservation peut être allongée, la finalité peut évoluer au cours du temps... On trouve une liste de mesures appropriées à mettre en place pour pouvoir collecter des données biométriques dans un cadre de recherche sur le site de la CNIL, dans un document [6] intitulé Présentation du régime juridique applicable aux traitements poursuivant une finalité de recherche scientifique (hors santé) :

— conformément à l'article 5.1-c) du RGPD, le principe de pertinence et de minimisation des données traitées doit être respecté ;
— l'anonymisation des données pourrait être mise en œuvre (le G29 a publié des recommandations à ce sujet en 2014) ; elle est obligatoire lors de la diffusion des résultats, diffusion nécessaire à la présentation des travaux de recherche (cf. l'article 116 du décret n°2019-536 du 29 mai 2019) ;
— la pseudonymisation doit être mise en œuvre toutes les fois où cela s'avérerait pertinent (sachant que les informations permettant de faire le lien entre les données et l'identité des individus doivent être conservées séparément et soumises à des mesures techniques et organisationnelles) ;
— une logique d'accès sécurisé et contrôlé doit être développée, en fonction de la sensibilité des données et des finalités des réutilisations futures ;
— la réalisation d'une analyse d'impact sur la protection des données est quasi systématique.

L'article de Sanchez et al. [22] propose une adaptation de ces mesures dans la pratique. On constate donc que l'adoption du RGPD a permis de clarifier et de structurer le cadre à respecter pour la collecte de données biométriques : le fait même qu'à partir d'une bribe de donnée biométrique, on pourrait être en mesure d'extraire une information nuisible pour la vie privée de l'utilisateur rend la protection de ces données indispensable. C'est précisément l'objet des techniques de biométrie révocable.

Limitations

Même si le législateur a inévitablement un peu de retard sur les risques sur la vie privée dus aux nouveaux usages des *smartphones* et autres objets connectés, on constate que les protections des citoyens évoluent dans le bon sens, que ce soit au niveau du droit français, ou européen, comme le montre la progression du respect de la vie privée dans la hiérarchie des normes.

6. https://www.cnil.fr/sites/default/files/atoms/files/consultation_publique_-_presentation_du_regime_juridique_applicable_aux_traitements_a_des_fins_de_recherche.pdf.

Dans son article, Els Kindt pointe les insuffisances du RGPD en matière de protection des données biométriques [14], notamment à travers ce constat (Art. 4, définition 14 du RGPD[7]) : « *Les données biométriques sont les données à caractère personnel résultant d'un traitement technique spécifique, relatives aux caractéristiques physiques, physiologiques ou comportementales d'une personne physique, qui permettent ou confirment son identification unique* ».

Deux éléments sont donc nécessaires pour qu'une donnée biométrique entre dans ce cadre : le traitement spécifique et l'identification unique. Comme contre-exemple, E. Kindt propose une base de données de photographies d'empreintes ou de visages, simplement collectées et stockées : dans ce cas, ces images ne seraient pas des données biométriques (cf. le considérant 51 du RGPD). Elle poursuit son analyse en étudiant la distinction entre quatre catégories de « données personnelles en lien avec des caractéristiques physiques, physiologiques ou comportementales d'un individu, et qui permettent son identification ou authentification » :

— *les données personnelles ordinaires* (considérées comme non biométriques) : les données simplement stockées, les données de biométrie douce (qui ne permettent pas une identification unique). C'est le régime général du RGPD qui s'applique à ces données (Art. 6) ;

— *les données biométriques* : cette catégorie, qui concerne les données biométriques en général (qui subissent un traitement spécifique et qui permettent ou confirment une identification unique) sont sujettes au même régime que les données personnelles ordinaires ;

— *les données biométriques sensibles* : ce sont les données qui sont traitées dans le seul but d'identifier de façon unique des individus. Dans ce cadre, l'article 9 du RGPD s'applique également : en principe, l'identification biométrique est interdite, sauf exceptions. Il faut souligner qu'il s'agit ici uniquement d'identification. L'authentification pourrait, sous couvert d'une certaine interprétation du RGPD, tomber dans la catégorie précédente (le but n'est pas une identification unique). Cependant, E. Kindt recommande des conditions supplémentaires : l'authentification devrait être transparente pour l'utilisateur, avec des garanties légales et techniques (pas de stockage dans une base centralisée), et le recours obligatoire à une AIPD ;

— *les données biométriques dont le but est d'identifier des individus à grande échelle* : ces données sont sujettes à des contraintes réglementaires additionnelles (Art. 35 et 36 du RGPD).

L'article de Gellert [11] porte sur la notion de risque, telle qu'elle est définie dans le RGPD. Il propose également une réflexion très pertinente sur la distinction entre *Privacy Impact Assessment* (PIA) et *Data Protection Impact Assessment* (DPIA). L'auteur rappelle que la vie privée va bien au-delà des données personnelles. Par

7. https://gdpr-text.com/fr.

conséquent, mener un DPIA serait forcément réducteur par rapport à un PIA. Il s'appuie sur les travaux de R. Clarke [5] : *« A Data Privacy Impact Assessment is a study of the impacts of a project on only the privacy of personal data, whereas a PIA considers all dimensions of privacy »*. Les dimensions de la vie privée auxquelles il fait référence sont : les informations personnelles, l'intimité de la personne, le comportement individuel, les communications personnelles, les pensées et les émotions. L'auteur, R. Gellert, dépasse ce clivage traditionnel et analyse en profondeur la notion de risque présente dans le RGPD. Il admet que le texte du règlement contient la définition de l'événement (à l'origine du risque), les conséquences (du risque) et un certain nombre de critères liés aux facteurs de risque. Il note toutefois qu'il manque un certain nombre d'éléments constitutifs d'un risque : (i) aucune mention n'est faite concernant la vraisemblance (ou probabilité d'occurrence) du risque (seul son niveau de gravité est mentionné), (ii) il y a peu de critères concernant les conséquences des risques : cela signifie que le règlement laisse une totale liberté quant au calcul des « risques relatifs aux droits et libertés des personnes concernées » dans la pratique, et (iii) le RGPD ne préconise aucune méthodologie pour réaliser un DPIA. La dernière phrase de l'article est la suivante :« *Thus, to some extent, the way risk is defined and understood in the GDPR is pretty much irrelevant* ». Els Kindt et Raphaël Gellert sont des juristes, aussi leur point de vue externe au domaine de l'informatique permet une prise de recul nécessaire et salutaire, en élargissant les perspectives.

On trouve également dans la littérature des critiques plus techniques sur les algorithmes de biométrie révocable. Le concept de *biométrie révocable* [19] repose sur une transformation des données biométriques brutes, de telle sorte que les données transformées soient sûres et respectueuses de la vie privée, en accord avec les propriétés suivantes (cf. [16, 13] ainsi que la norme ISO/IEC 24745) : non-inversibilité, performance, diversité, révocabilité.

Dans l'article [18], Nandakumar et Jain soulignent que les propriétés de révocabilité et de non-inversibilité ne sont pas garanties intrinsèquement dans les schémas de biométrie révocable, et qu'il est nécessaire de développer des fonctions à sens unique pour rendre le calcul d'une pré-image plus difficile. D'autres attaques sont possibles visant les techniques à base de salage des données, elles sont détaillées dans l'article de Choudhury et al. [4] : ces attaques exploitent la présence de la donnée biométrique originale, de la clé, de la donnée transformée, ou l'interception de plusieurs données transformées. Dans l'article [18], les auteurs proposent trois raisons principales au fossé existant entre la théorie (notamment en ce qui concerne les propriétés attendues) et la pratique pour les systèmes de biométrie révocable :

— la méthode d'extraction des caractéristiques (*features*) ;
— le compromis entre la non-inversibilité et les performances de reconnaissance ;
— la garantie des propriétés de non-associativité et révocabilité.

Concernant les schémas reposant sur une transformation des données, souvent la non-inversibilité y est estimée de façon empirique à partir de la complexité de calcul de l'attaque par inversion. Les auteurs citent l'exemple de la courbe *Coverage-Effort* proposée par Nagar et al. [17], qui permet de mesurer le nombre de tentatives (ou efforts) requis pour récupérer une fraction (*coverage*) de la donnée originale. Ils soulignent également que cette approche empirique ne permet pas de traiter toutes les stratégies d'inversion des gabarits par des imposteurs. Dong et al. proposent d'exploiter la propriété de préservation des distances (inhérente aux techniques de biométrie révocable par construction) pour définir une classe d'attaques basées sur la similarité [7]. Les auteurs se placent dans l'hypothèse de Kerckhoff : l'attaquant a accès à un gabarit transformé et connaît la fonction de transformation ainsi que ses paramètres. Ils testent leur méthodologie d'attaque sur plusieurs schémas de biométrie révocable. Dans l'article [3], Chen et al. vont plus loin et démontrent, que lorsque la distance entre deux données biométriques non protégées augmente, la distance entre les données protégées correspondantes augmente également, inévitablement. La solution proposée repose sur le choix d'une fonction de transformation plus compliquée, non linéaire. Gomez-Barrero et al. proposent explicitement un cadre pour évaluer la non-associativité des données protégées par les schémas de biométrie révocable [12]. Les auteurs définissent deux fonctions — une globale et une locale — pour mesurer un score de non-associativité entre deux échantillons, selon qu'ils sont appariés ou non.

Aujourd'hui encore, il est quasi impossible de prouver qu'un schéma de biométrie révocable est totalement résistant aux attaques de sécurité et sur la vie privée. L'article de Wang et Baskerville [24], propose une vision divergente, qui critique les bénéfices de l'authentification à deux facteurs (par rapport à une authentification à un facteur), avec une analyse originale risques/bénéfices/coûts. Tous ces éléments laissent donc de nombreuses perspectives de recherche, dans des directions variées.

Conclusion

Pour conclure, voici un extrait d'une interview d'Alain Damasio (écrivain de science-fiction), diffusée sur France Culture en 2019[8] : « *Ce qu'il se passait en milieu fermé dans la prison où on était hypra surveillés, est devenu la norme d'existence aujourd'hui. Ça ne gêne personne au sens où tant qu'on n'a pas le retour de pouvoir sur nos vies, tant que quelqu'un ne s'empare pas de ça pour nous « emmerder », pour nous bloquer, pour nous refuser un job, on ne sent pas qu'on est pris dans cette nasse. Mais le fait est qu'on circule dans des prisons à ciel ouvert en fait.* »

Cette réflexion est également présente dans le livre de S. Zuboff [26], qui tente d'expliquer pourquoi on a « laissé » Google, Facebook, Instagram, etc., aspirer nos

8. https://www.franceculture.fr/societe/le-panoptique-a-lorigine-de-la-societe-de-surveillance.

données et nos comportements : cette révolution numérique est *sans précédent* (formule de l'auteure), par conséquent les utilisateurs n'ont pas de repères, pas de modèles auxquels se raccrocher. On a donc laissé faire au départ, de façon naïve et innocente, jusqu'à ce qu'on ne puisse plus se passer des services proposés, qui simplifient la vie quotidienne, et qu'on ferme les yeux sur le prix à payer.

L'Union européenne a réagi, sans doute un peu tard, en déployant le RGPD pour protéger les données de ses ressortissants. Ce règlement fait des émules, en Californie par exemple. Cependant, les entreprises s'adaptent très vite aux différentes contraintes : les lignes directrices du Conseil européen de protection des données (EDPB) sur le consentement valide publiées en mai 2020 stipulent que les *cookie walls* ne sont pas un moyen valide pour les sites web d'obtenir le consentement des utilisateurs au traitement de leurs données personnelles et à l'utilisation de *cookies*. Pourtant, on constate aujourd'hui que certains sites web ont trouvé la parade en quelques semaines en rendant leur consultation payante si les *cookies* sont refusés par les utilisateurs. Ce comportement présente au moins l'avantage de mettre fin à l'hypocrisie du « tout gratuit » sur internet. Le livre [26] met fin à la croyance classique « si c'est gratuit, c'est toi le produit » : les géants d'internet ne cherchent plus à créer des profils d'utilisateurs, ils recourent aujourd'hui à l'intelligences artificielle pour prédire leurs comportements futurs, et c'est cette prédiction qu'ils revendent à des annonceurs, des services marketing, des publicitaires, etc., qui à leur tour vont essayer d'influencer les comportements des utilisateurs.

Il est donc peut-être temps de renverser le modèle et de replacer l'utilisateur au cœur de la collecte de ses données, en appliquant la transparence, la minimisation, le consentement éclairé, imposés par l'application du RGPD. On peut également aller plus loin en redonnant à l'utilisateur une souveraineté sur ces aspects, une compréhension des enjeux et un réel pouvoir de décision, par exemple à travers le développement d'une application de collecte de données respectueuse de la vie privée : il s'agit de mon projet de recherche, exposé en détail dans mon manuscrit d'habilitation à diriger des recherches (HDR) [9].

Pour finir, je voudrais citer G. Bernanos, qui écrivait déjà en 1947 [10] :« *L'idée qu'un citoyen, qui n'a jamais eu affaire à la justice de son pays, devrait rester parfaitement libre de dissimuler son identité à qui il lui plaît, pour des motifs dont il est seul juge, ou simplement pour son plaisir, que toute indiscrétion d'un policier sur ce chapitre ne saurait être tolérée sans les raisons les plus graves, cette idée ne vient plus à l'esprit de personne.* »

Références

9. https://tel.archives-ouvertes.fr/tel-03326656.
10. La France contre les robots, Georges Bernanos, 1947.

[1] Nicolás E. Bordenabe. *Measuring Privacy with Distinguishability Metrics :Definitions, Mechanisms and Application to Location Privacy*. PhD thesis, Ecole Polytechnique, 2014.

[2] Ann Cavoukian. Privacy by design, 2011.

[3] Yanzhi Chen, Yan Wo, Renjie Xie, Chudan Wu, and Guoqiang Han. Deep Secure Quantization : On secure biometric hashing against similarity-based attacks. *Signal Processing*, 2019.

[4] Bismita Choudhury, Patrick Then, Biju Issac, Valliappan Raman, and Manas Kumar Haldar. A Survey on Biometrics and Cancelable Biometrics Systems. *International Journal of Image and Graphics*, 2018.

[5] Roger Clarke. An evaluation of privacy impact assessment guidance documents. *International Data Privacy Law*, 2011.

[6] Aviva de Groot and Bart van der Sloot, editors. *The Handbook of Privacy Studies : An Interdisciplinary Introduction*. Amsterdam University Press, 2018.

[7] Xing-Bo Dong, Zhe Jin, A. B. J. Teoh, M. Tistarelli, and K. Wong. On the reliability of cancelable biometrics : Revisit the irreversibility. 2019.

[8] European Parliament. Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. Available : `http://eur-lex.europa.eu/eli/reg/2016/679/oj`, 2016.

[9] Catherine Flick. Informed consent and the Facebook emotional manipulation study. *Research Ethics*, 2016.

[10] C. Fried. *Privacy : A Moral Analysis*, pages 203–222. Cambridge University Press, 1984.

[11] Raphaël Gellert. Understanding the notion of risk in the General Data Protection Regulation. *Computer Law & Security Review*, 2018.

[12] Marta Gomez-Barrero, Javier Galbally, Christian Rathgeb, and Christoph Busch. General Framework to Evaluate Unlinkability in Biometric Template Protection Systems. *IEEE Transactions on Information Forensics and Security*, 2018.

[13] Anil K. Jain, Karthik Nandakumar, and Abhishek Nagar. *Fingerprint Template Protection : From Theory to Practice*. Springer London, 2013.

[14] Els Kindt. Having yes, using no ? about the new legal regime for biometric data. *Computer Law & Security Review*, 34, 12 2017.

[15] Bert-Jaap Koops, Bryce Newell, Tjerk Timan, Ivan Skorvanek, Tom Chokrevski, and Maša Galič. A typology of privacy. *University of Pennsylvania Journal of International Law*, 2016.

[16] Davide Maltoni, Dario Maio, Anil K. Jain, and Salil Prabhakar. *Handbook of Fingerprint Recognition*. Springer London, 2009.

[17] Abhishek Nagar, Karthik Nandakumar, and Anil K. Jain. Biometric template transformation : a security analysis. In *Electronic Imaging*, 2010.

[18] Karthik Nandakumar and Anil K. Jain. Biometric Template Protection : Bridging the performance gap between theory and practice. *IEEE Signal Processing Magazine*, 2015.

[19] N. K. Ratha, J. H. Connell, and R. M. Bolle. Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal*, 2001.

[20] J. Reimann. Driving to the panopticon : A philosophical exploration of the risks to privacy posed by the highway technology of the future. *Santa Clara High Technology Law Journal*, 11(1) :27–44, 1995.

[21] Beate Roessler. *The Value of Privacy*. Cambridge : Polity Press, 2005.

[22] Raul Sanchez-Reillo, Ines Ortega-Fernandez, Wendy Ponce-Hernandez, and Helga C. Quiros-Sandova. How to implement eu data protection regulation for r&d in biometrics. *Computer Standards & Interfaces*, 2019.

[23] Isabel Wagner and David Eckhoff. Technical privacy metrics : A systematic survey. *ACM Computing Surveys*, 2018.

[24] Pengcheng Wang and Richard Baskerville. The case for two-factor authentication- evidence from a systematic literature review. In *Twenty-Third Pacific Asia Conference on Information Systems*, 2019.

[25] Samuel D. Warren and Louis D. Brandeis. The right to privacy. *Harvard Law Review*, 1890.

[26] Shoshana Zuboff. *Age of Surveillance Capitalism : The Fight for a Human Future at the New Frontier of Power*. New York : PublicAffairs, 2019.