



RÉPUBLIQUE FRANÇAISE



**OFFICE PARLEMENTAIRE D'ÉVALUATION
DES CHOIX SCIENTIFIQUES ET TECHNOLOGIQUES**

Note à l'attention des membres de l'Office

Épidémie de COVID-19 – Point sur les technologies de l'information utilisées pour limiter la propagation de l'épidémie de COVID-19

Cette note a été présentée en réunion de l'Office parlementaire d'évaluation des choix scientifiques et technologiques le 9 avril 2020, par son premier vice-président, le député Cédric Villani¹. Elle a été mise à jour le 11 avril 2020 et validée pour publication.

Alors que des traitements sont en cours d'évaluation et que des vaccins sont en développement pour, à terme, immuniser la population, la limitation de la propagation de l'épidémie par des stratégies non pharmaceutiques reste un levier essentiel de la réponse sanitaire. Le confinement en est un des aspects, mais il n'est pas pratiqué partout, du moins pas dans les modalités extensives retenues à Wuhan ou celles qui ont effet en Italie et en France. À l'image de la stratégie déployée par la France au début de l'épidémie, certains pays poussent l'isolement des personnes infectées et des cas suspects et la recherche des cas contacts au maximum de leurs capacités. Les technologies de l'information peuvent contribuer à cette stratégie. Plus ou moins intrusives, ces technologies ont pour finalités principales l'évitement des personnes potentiellement contagieuses et l'identification de toutes les personnes ayant été en contact avec les porteurs du virus au cours de leur période de contagiosité. La présente note revient sur ces outils en trois temps : l'aspect technologique, l'aspect organisationnel et l'aspect éthique du traitement de ces données.

Aspects technologiques

Deux catégories de données personnelles sont mobilisées par les technologies de l'information pour lutter contre l'épidémie :

- les données à caractère médical, qui ont vocation à permettre de surveiller des symptômes pour décider ou orienter une prise en charge ou à préciser l'étiologie de l'infection ;
- les données permettant de localiser les individus ou de renseigner sur des positions relatives, qui ont vocation à conforter les mesures de distanciation sociale ou à étudier les mouvements de population².

1 Déclaration de liens d'intérêt : le député Cédric Villani a été membre du Conseil scientifique d'Orange entre 2015 et 2017.

2 « Distanciation sociale » est le terme le plus utilisé, mais « distanciation physique » est plus adapté, car le lien social, que permettent le téléphone, les outils de messagerie ou encore les réseaux sociaux, est, de l'avis général, nécessaire pour surmonter l'épreuve psychologique du confinement.

Secrétariats de l'OPECST :

Assemblée nationale : 01 40 63 26 81 – Sénat : 01 42 34 31 07

Les données à caractère médical, permettant de surveiller des symptômes ou de préciser l'étiologie de l'infection

L'aide à la surveillance des symptômes se développe par le biais d'applications pour smartphone. Sur la base de la déclaration de ses symptômes, l'utilisateur reçoit un message l'informant sur la démarche à suivre : rester chez soi, consulter son médecin traitant ou bien appeler le SAMU. Les objectifs peuvent être multiples :

- Décharger les services d'urgence et le SAMU de patients qui, étant donné l'afflux de cas graves à l'hôpital, n'y seraient pas admis ;
- Mieux anticiper les besoins hospitaliers ;
- Surveiller la survenue de symptômes chez des personnes placées en quarantaine préventive, les personnes entrant dans un pays, par exemple ;
- En couplant les données médicales aux données de localisation, connaître les zones géographiques regroupant des cas de maladie pour adapter les consignes à la population et, notamment, mieux orienter les campagnes de dépistage massif.

En France, l'Assistance publique – Hôpitaux de Paris a mis en place Covidom³, une application que les patients sont invités à utiliser après inscription par un médecin. En Corée du Sud, une application d'autodiagnostic est utilisée pour surveiller l'état de santé des personnes arrivant sur le territoire. Au Vietnam, l'application NCOVI⁴ a été mise en place à la fois pour mieux anticiper les besoins médicaux, mais aussi pour prévenir la population avoisinante de cas déclarés. De nombreuses applications de cette sorte ont été développées ; étant donné leur caractère médical, il semble important que les autorités sanitaires nationales se prononcent sur leur pertinence.

En plus des objectifs précités, ces applications peuvent avoir pour objectif de collecter des données médicales anonymisées afin de mieux caractériser la maladie. En effet, les patients ne présentant que des formes bénignes de la maladie – mal comprises à ce jour – ne font pas l'objet d'une hospitalisation et du suivi approfondi qui l'accompagne. L'Office souligne l'intérêt de la collecte de ces données, dans une perspective de développement de la science ouverte (*open science*), et invite les promoteurs de tels services à s'inscrire dans cette démarche.

D'autres initiatives visent, à partir de l'agrégation de données médicales hospitalières, à créer des outils prédictifs de l'évolution de l'état de santé des patients, par exemple à partir d'une image de scanner thoracique ; de tels outils peuvent aider à adapter les parcours de soins ou à mieux anticiper les besoins hospitaliers en structures lourdes (lit de réanimation et respirateur artificiel, par exemple). Ces projets, qui nécessitent calibrage et apprentissage automatique à partir de grandes quantités de données, se heurtent souvent à des difficultés pour accéder, dans l'urgence, à des bases de données fiables et bien structurées.

3 « Covidom : une solution de télésuivi à domicile pour les patients porteurs ou suspectés Covid-19 co-construite par l'AP-HP et Nouveal e-santé ». <https://www.aphp.fr/contenu/covidom-une-solution-de-tele-suivi-domicile-pour-les-patients-porteurs-ou-suspectes-covid-19> (consulté le 2 avril 2020).

4 « Vietnam launches health app to manage COVID-19 », OpenGov Asia, mars 10, 2020. <https://www.opengovasia.com/vietnam-launches-health-app-to-manage-covid-19/> (consulté le 30 mars 2020).

Les données permettant de localiser les individus pour conforter les mesures de distanciation sociale ou étudier les mouvements de population

La localisation des personnes contribue à limiter la propagation de la pandémie en poursuivant des objectifs variés mais concourant à l'efficacité des mesures de distanciation sociale :

- Veiller au respect du confinement d'une personne à son domicile ;
- Identifier les contacts établis par un malade durant sa période de contagion ;
- Connaître les mouvements de population au sein d'un pays.

Différentes technologies permettent cette localisation : l'utilisation de données GPS⁵, de données téléphoniques ou des données bancaires. Indirectement, la technologie *bluetooth*, le port de bracelets électroniques ou encore la reconnaissance faciale, le permettent également. En France et dans d'autres pays européens, plusieurs projets d'application sont actuellement à l'étude.

- Des technologies relativement simples : GPS, *bluetooth*, données téléphoniques...

Le respect du confinement de certaines personnes (ou quatorzaine⁶) est essentiel pour les pays n'ayant pas mis en place de confinement généralisé. Il s'agit des personnes malades, des cas suspects, de personnes ayant été en contact avec des porteurs du virus, mais n'ayant pas développé de symptômes, ou de personnes entrées sur le sol d'un pays n'ayant pas fermé ses frontières. Plusieurs pays sont dans cette situation, notamment Hong Kong, Singapour, la Corée du Sud, Taïwan, la Chine, à l'exception de la province du Hubei, ou encore Israël.

Il est, par exemple, demandé à tous les entrants sur le territoire de Hong Kong de porter un bracelet électronique pour vérifier la bonne application de la quatorzaine. Le bracelet communique avec un smartphone, dont le GPS renseigne la localisation de la personne aux autorités sanitaires ; toute déconnexion entre le bracelet et le smartphone est détectée et indique le non-respect de la quatorzaine⁷. La Corée du Sud, Taïwan et Singapour surveillent également le respect du confinement des cas suspects ou confirmés par ce que l'on pourrait appeler des « clôtures électroniques »⁸.

L'isolement des personnes malades est une mesure nécessaire, mais non suffisante, dans la mesure où la durée exacte de la période de contagiosité n'est pas encore clairement établie ; de nombreux scientifiques estiment qu'une personne infectée est contagieuse avant même que les symptômes ne surviennent. Ainsi, certains pays ont adopté une stratégie pour empêcher une propagation incontrôlée de la maladie qui consiste à retracer tous les cas contact, c'est-à-dire identifier toutes les personnes avec qui le malade a été en contact pendant le temps où il était contagieux, et imposer la quatorzaine ou des mesures strictes pour ces personnes.

5 Notons que la localisation par GPS utilise effectivement la technologie GPS lorsque le téléphone se situe en extérieur, mais utilise d'autres moyens, comme un « bornage » par les emplacements wifi les plus proches lorsque le téléphone est à l'intérieur. Quand on parle de « données GPS », il s'agit donc d'un abus de langage pour désigner la combinaison de plusieurs technologies de repérage géographique plus ou moins fin.

6 On utilisera dans cette note le terme générique de quarantaine, et parfois celui de quatorzaine lorsque la durée de celle-ci est de deux semaines.

7 « Smart devices help HK fight against new coronavirus », FutureIoT, févr. 14, 2020. <https://futureiot.tech/smart-devices-help-hk-fight-against-new-coronavirus/> (consulté le 17 mars 2020).

8 B. Betz, « Coronavirus: New tech being used to fight pandemic », Fox News, mars 25, 2020. <https://www.foxnews.com/tech/coronavirus-new-tech-fight-pandemic> (consulté le 30 mars 2020).

Différents moyens peuvent être utilisés pour retracer les contacts établis : via la collecte de la position GPS des téléphones portables, ou la détermination de leur position par l'étude des données téléphoniques, comme cela est permis en Israël⁹, dans le contexte de la crise sanitaire, ou par le biais du *bluetooth*. De ces différentes techniques, la dernière est celle qui permet de déterminer la proximité le plus finement, même s'il ne s'agit que d'une proximité relative, car le GPS a une résolution bornée à quelques mètres et la résolution spatiale atteinte avec le réseau téléphonique dépend de la densité d'antennes – elle est de plusieurs dizaines de mètres en ville. Une application smartphone singapourienne, Trace Together¹⁰, téléchargée sur la base du volontariat, permet d'identifier tous les téléphones ayant téléchargé l'application et ayant été à moins d'un mètre de distance pendant au moins 30 minutes¹¹. *A posteriori*, l'application prévient les personnes ayant été en contact avec une personne déclarée positive au virus, sans en révéler l'identité. L'application n'enregistre pas la localisation absolue des personnes puisqu'elle fonctionne uniquement *via* la détection d'une proximité relative ; elle est donc peu intrusive, même si le principe de l'alerte donnée *ex post* repose sur l'identification des personnes infectées et leur renseignement dans une base de données centralisée. En France, plusieurs projets d'application seraient actuellement à l'étude¹².

D'autres moyens sont utilisés à Singapour, qui a choisi de mettre en œuvre une recherche des cas contacts très poussée : ainsi, l'utilisation des données bancaires et de la vidéo-surveillance peut servir à retrouver les personnes qui auraient été dans la même salle de cinéma qu'une personne malade¹³.

L'étude des mouvements de population a été médiatisée en France à l'occasion de l'étude menée par l'équipe de recherche de l'Inserm dirigée par Vittoria Colizza, en partenariat avec l'opérateur téléphonique Orange. Elle a permis de montrer que près d'un million de Franciliens ont quitté la région parisienne juste avant le début du confinement et au début de celui-ci¹⁴. Alors que l'Île-de-France est aujourd'hui la région qui comptabilise le plus de cas confirmés de COVID-19 et de décès¹⁵, ces mouvements de population sont susceptibles de favoriser la propagation de l'épidémie. À l'échelle européenne, les sociétés de télécommunications partagent également des données de localisation avec les autorités sanitaires en Italie, en Allemagne et en Autriche, pour évaluer la diminution des déplacements¹⁶. Les données

9 « ARTE Journal », ARTE. <https://www.arte.tv/fr/videos/091306-028-A/arte-journal/> (consulté le 30 mars 2020).

10 « TraceTogether ». <https://www.tracetogogether.gov.sg/> (consulté le 2 avril 2020).

11 Ces critères sont déterminés pour repérer au mieux les situations de contact susceptibles de causer une contamination ; un temps de proximité trop court et une distance minimale trop importante risqueraient de générer une alerte pour un trop grand nombre d'évènements de contact, non pertinents en termes sanitaires, qu'on pourrait appeler « faux positifs ».

12 Grallet, Guillaume. « Géolocalisation : comment tracer sans cliquer ? » Le Point, 1^{er} avril 2020. https://www.lepoint.fr/sciences-nature/le-modele-technologique-de-singapour-contre-celui-de-taiwan-01-04-2020-2369746_1924.php.

13 « Interrogatoires serrés, vidéosurveillance, porte à porte : comment Singapour traque les contaminés », Les Echos, mars 16, 2020. <https://www.lesechos.fr/monde/asie-pacifique/interrogatoires-serres-videosurveillance-porte-a-porte-comment-singapour-traque-les-contamines-1185804> (consulté le 2 avril 2020).

14 Cette étude utilisant les données de téléphonie mobile garantit les critères d'anonymisation par l'utilisation du maillage territorial Iris : le nombre de personnes par maille est suffisant pour qu'il soit impossible de remonter aux personnes à l'origine de ces données. Salle de presse | Inserm. « Les statistiques issues du réseau de téléphones mobiles au service de la lutte contre la pandémie de Covid-19 », 27 mars 2020. <https://presse.inserm.fr/les-statistiques-issues-du-reseau-de-telephonies-mobiles-au-service-de-la-lutte-contre-la-pandemie-de-covid-19/38831/>.

15 <https://geodes.santepubliquefrance.fr/>

16 « 19 countries track mobile location to fight COVID-19 », GPS World, mars 26, 2020. <https://www.gpsworld.com/19-countries-track-mobile-locations-to-fight-covid-19/> (consulté le 30 mars 2020).

agréées sont anonymes, cartographiant les concentrations plutôt que les individus, afin de respecter les normes européennes sur la protection des données personnelles.

Le réseau téléphonique peut également être utilisé, à l'échelle d'un quartier, d'une commune, pour prévenir les habitants que des cas sont présents dans cette zone et qu'ils doivent modifier leurs habitudes en conséquence. Il s'agit de diffuser des alertes sous la forme de sms via la diffusion cellulaire localisée¹⁷.

- ... à la reconnaissance faciale ...

En plus de servir à vérifier si des personnes, au sein d'une foule, ne respectent pas des mesures telles que le port du masque, la reconnaissance faciale est utilisée à Moscou pour vérifier que les personnes revenant de pays touchés par la pandémie, respectent bien leur quarantaine¹⁸. En Chine, la reconnaissance faciale permettrait également d'attribuer aux citoyens une probabilité d'être porteur du virus, sur la base du côtoiement de personnes ne portant pas de masque, dans l'espace public¹⁹.

- ... et à l'intelligence artificielle pour analyser les déplacements

En Chine, l'intelligence artificielle est utilisée pour détecter des personnes ayant été exposées à des risques de contamination, pour mieux cibler les personnes à placer en quarantaine, ainsi que pour déterminer combien de cas sont présents dans une même zone. C'est notamment sur la base des données de navigation sur internet et sur les réseaux sociaux que ce système fonctionne²⁰. Ce genre de technologies permet à l'outil Healthmap de définir combien de cas sont présents par unité géographique²¹.

- Des outils dont l'efficacité technique est encore incertaine au regard des objectifs poursuivis

Toutes ces techniques voient cependant leur efficacité limitée par le fait que de nombreuses personnes ne déclarent pas, ou qu'elles n'ont que peu de symptômes, lorsqu'elles sont infectées par le virus SARS-CoV-2. La fraction de ces cas a- et paucisymptomatiques et leur contribution à la transmission de la maladie ne sont pas déterminées, mais elles semblent non négligeables.

On notera également que la proximité avec des personnes malades n'est qu'un déterminant très approximatif de la probabilité d'avoir été infecté, étant donné que le port du masque et l'adoption des gestes barrière peuvent prévenir une infection, malgré une proximité avec un cas avéré. A l'inverse, une « distance d'effet » paramétrée à un mètre ignore le risque des contaminations résultant d'une toux importante par une personne ne portant pas de masque. De

17 Cette technique, le *cell broadcast* en anglais, est largement utilisée dans certains pays pour prévenir les populations de dangers tels que des tsunamis. Elle agit via les antennes relais, touchant ainsi tous les téléphones connectés à ces antennes, permettant de cibler une zone géographique précise.

18 « Caméras, reconnaissance faciale : les yeux de Moscou au service du confinement », Le Courrier de Russie, mars 20, 2020. <https://www.lecourrierderussie.com/societe/2020/03/cameras-reconnaissance-faciale-les-yeux-de-moscou-au-service-du-confinement/> (consulté le 30 mars 2020).

19 Cette application est issue d'une collaboration entre la direction générale du Conseil des Affaires d'Etat, la Commission nationale de la santé et China Electronics Technology Group Corporations (CETC).

20 « How technology is helping to fight against CoronaVirus? | LeewayHertz », Software Development Company, févr. 24, 2020. <https://www.leewayhertz.com/how-technology-is-helping-to-fight-against-coronavirus/> (consulté le 26 mars 2020).

21 « HealthMap | Flu Map | Contagious Disease Surveillance | Virus Awareness ». <http://healthmap.org> (consulté le 2 avril 2020).

plus, le contexte est ignoré : le risque de contagion dans un environnement fermé est plus élevé que dans un environnement ouvert, or, le traçage de la proximité relative par *bluetooth* ne permet pas la contextualisation, contrairement aux traçages GPS.

L'efficacité d'un tel outil, permettant de tracer les contacts par la technologie *bluetooth*, s'il est utilisé sur la base du volontariat, dépendra également de son degré d'utilisation²². En France, un sondage effectué fin mars suggère que l'application pourrait être largement utilisée²³.

Aspects organisationnels

Les données de santé sont une ressource précieuse ; celles relatives au COVID-19 seront particulièrement utiles pour mieux connaître la maladie (symptômes, part des cas asymptomatiques, etc.) mais aussi pour développer des outils techniques qui permettront de mieux prendre en charge les patients ou de mieux contenir l'épidémie. Par exemple, entraîner des outils d'apprentissage algorithmique (intelligence artificielle) sur de telles données permettrait de concevoir des outils prédictifs de l'évolution d'un cas d'infection.

De telles données de santé peuvent provenir de l'utilisation des technologies décrites précédemment. Cependant, en France, les données de santé disponibles sont principalement :

- Les données médico-administratives du Système national d'information inter régimes de l'assurance maladie (SNIIRAM), les données relatives aux remboursements effectués par la Caisse nationale d'assurance maladie – c'est cet organisme qui gère ces données ;
- Les données hospitalières du Programme de médicalisation des systèmes d'information (PMSI), gérées par l'Agence technique de l'information sur l'hospitalisation (ATIH) ; ces données sont intégrées au SNIIRAM ;
- Les données relatives aux causes médicales des décès, élaborées par le CépiDc qui est un laboratoire de l'Inserm.

Ces données sont agrégées de façon pseudonymisée dans le Système national des données de santé (SNDS), au sein de la Plateforme des données de santé (*Health data hub*). Cette plateforme a pour vocation d'apparier les données issues des bases médico-administratives, reflétant le parcours de soins, et d'autres sources, telles que des bases de données d'imagerie médicale, et de les rendre disponibles pour la recherche, académique et privée.

Notons que le couplage des bases PMSI et SNIIRAM permettrait, dans le cadre du COVID-19, de rapprocher l'état du patient avec les éléments antérieurs de son parcours de soins, sur un

22 Différentes estimations font état d'une efficacité à partir du moment où l'outil serait utilisé par 25 à 60 % de la population. « Traçage des données mobiles dans la lutte contre le COVID-19 - Analyse des potentiels et des limites », Note parlementaire du député Mounir Mahjoubi, publiée le 6 avril 2020. L'application Trace Together a été téléchargée par un million d'habitants sur les 5,6 millions que compte le pays, entre le 20 mars et le 1^{er} avril ; https://www.francetvinfo.fr/sante/maladie/coronavirus/tracage-numerique-via-bluetooth-pour-eviter-les-contaminations-au-covid-19-lexemple-de-singapour-qui-pourrait-inspirer-la-france_3907331.html .

23 Le sondage, réalisé pour une équipe de recherche britannique sur un échantillon de population de 1 000 personnes, évalue à près de 48 % les personnes qui installeraient « sans aucun doute » l'application, et à 31 % celles qui l'installeraient « probablement » ; <https://osf.io/24uan/>.

grand nombre d'individus, pour révéler ainsi des facteurs de risque de la maladie, ou bien encore, des interactions médicamenteuses.

L'importation des données au sein du SNDS se fait à plusieurs échéances :

- Les données SNIIRAM, consolidées et définitives, sont intégrées dans le SNDS de façon annuelle, 6 mois après l'année écoulée ;
- Elles peuvent toutefois être débloquées de façon mensuelle, mais elles ne sont alors pas consolidées.

Dans le contexte de la crise actuelle, la mise au point de solutions contribuant à limiter la propagation de l'épidémie et à mieux prendre en charge les patients est urgente. Pourtant, le chaînage des données PMSI et SNIIRAM par la CNAM, qui prend environ un mois, et les différentes étapes de la procédure d'accès aux données étendent les délais de mise à disposition de ces données.

Lorsqu'un dossier de demande d'accès est soumis au *Health data hub*, un premier examen est effectué par le CEREES, le Comité d'expertise pour les recherches, les études et les évaluations dans le domaine de la santé, qui a pour mission d'examiner le caractère scientifique des demandes et leur méthodologie²⁴. Un deuxième examen est effectué par la Commission nationale de l'informatique et libertés (CNIL). La procédure d'accès aux données PMSI a été accélérée pour les demandes en lien avec le COVID-19. Ainsi, l'examen du dossier par le CEREES est effectué dans un délai de deux semaines, alors qu'il est habituellement d'un mois, et la CNIL instruit les dossiers COVID-19 en priorité. Notons toutefois que le CEREES, qui devait être remplacé par le CESRESS en application de la loi relative à l'organisation et à la transformation du système de santé²⁵, a dû continuer d'assurer cette fonction, le décret permettant la mise en place du CESRESS n'ayant pas été publié à ce jour²⁶ ; un tel retard est difficilement compréhensible, surtout dans un contexte de crise où la puissance publique a besoin d'être éclairée par des travaux de recherche qui doivent pouvoir s'alimenter aux données gérées par le *Health data hub*. Le système administratif doit montrer une plus grande réactivité et faire en sorte de lever les blocages à l'entrée en fonction du CESRESS.

D'autres freins à l'utilisation de ces données à court terme sont à prévoir : les principales données concernant les patients COVID-19 confirmés sont celles renseignées dans le PMSI par les « médecins DIM » (Département de l'information médicale) ; étant donné les contraintes actuellement subies par le système hospitalier, il est à prévoir un retard de codage pour la période mars-avril 2020, comme l'a justement permis l'ATIH²⁷.

Les autres données de santé, non relatives au COVID-19, seront également affectées par la pandémie car les soins non liés à celle-ci et nécessitant un passage en réanimation ont été en

24 <https://www.indsante.fr/fr/impact-de-la-loi-relative-lorganisation-et-la-transformation-du-systeme-de-sante-sur-les-donnees-de>.

25 Le CESREES, le Comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé, est prévu l'article 41 de la loi n° 2019-774 du 24 juillet 2019 relative à l'organisation et à la transformation du système de santé.

26 Il s'agit du décret d'application du 2° de l'article 76 de la loi Informatique et libertés dans sa rédaction issue de la loi Organisation et transformation du système de santé (décret en Conseil d'État, pris après avis de la Commission nationale de l'informatique et des libertés, qui doit fixer la composition du comité éthique et scientifique et définir ses règles de fonctionnement).

27 <https://www.atih.sante.fr/transmission-des-donnees-m2-des-donnees-pmsi>.

grande partie déprogrammés. Des répercussions inhabituelles sur les soins pourraient s'observer jusqu'à l'automne et s'observer dans toutes les données de santé.

D'autres bases de données sont amenées à venir alimenter le SNDS en données concernant des patients COVID-19 ; notamment celles recueillies par les applications à caractère médical. Ces contributions reposent cependant sur le volontariat. La crise actuelle devrait inciter tous les acteurs à adhérer au concept d'*open science* en partageant leurs données via le *Health data hub*. Les données de santé accumulées dans les entrepôts hospitaliers, telles que les résultats d'analyses biologiques ou d'imagerie, sont particulièrement intéressantes relativement au COVID-19, ainsi que d'autres initiatives émergeant des secteurs de recherche les plus en lien avec la maladie.

Aspects éthiques et juridiques

Un corpus juridique exigeant en matière de protection des données personnelles.

La collecte et le traitement de données personnelles font l'objet d'un encadrement juridique contraignant en France et en Europe reposant sur la loi Informatique et libertés de 1978, ainsi que sur le règlement général sur la protection des données (RGPD) de 2016 et la directive ePrivacy de 2002 sur la protection de la vie privée dans les communications électroniques.

Ces textes posent les principes suivants :

- Obligation de consentement explicite et positif à la collecte de données personnelles. En principe, il ne peut y avoir de collecte de données à l'insu des intéressés et pas de collecte forcée. Cependant, il peut exister d'autres bases légales que le consentement pour le recueil de données, notamment lorsque celui-ci est entrepris dans l'intérêt de la santé publique ou de la recherche. Les personnes publiques mentionnées à l'article 67 de la loi informatique et libertés peuvent par exemple mettre en place des traitements de suivi individualisé sous réserve de la seule réalisation d'une analyse d'impact ;
- Interdiction du profilage au moyen de systèmes automatisés dans le but d'ouvrir ou de restreindre des droits individuels ;
- Confidentialité du contenu des communications électroniques ;
- Droit à l'effacement des données personnelles collectées au bout d'un certain délai ;
- Absence de publicité des données individuelles collectées, impliquant lorsque des données sont mises à disposition de tiers une exigence d'anonymisation des informations.

A ces règles s'ajoutent des exigences générales de proportionnalité de la collecte de données aux finalités des traitements, de sécurité des données ou encore de non-détournement d'usages.

Ces règles s'appliquent aux traitements de données qui pourraient être utiles dans le cadre de la gestion de la pandémie du coronavirus : traitement des données de localisation des individus ou encore traitement des données de santé.

Les questions posées par le traçage (*tracing*) des déplacements à travers les smartphones.

Le traçage des déplacements pourrait être entrepris en phase de déconfinement pour identifier les contacts de chaque individu et obtenir une cartographie des contaminations possibles à travers les interactions volontaires ou involontaires entre individus.

Il ne semble pas y avoir d'interdiction générale de ce type de traitement de données, mais celui-ci ne pourrait être mis en œuvre qu'en respectant certaines conditions très strictes :

- Les données doivent être **anonymisées**, s'agissant des données de déplacement de population, ou le **consentement recueilli de manière explicite**, s'agissant des données individuelles, permettant d'établir la liste des contacts, et faire l'objet d'un chiffrement robuste pour éviter toute atteinte à la vie privée. Or, l'anonymisation est une exigence lourde car les scientifiques ont démontré que connaître 4 points de présence d'un individu dans un jeu de données, peut typiquement suffire à retrouver son jeu de données personnalisées de géolocalisation²⁸ ;
- Il conviendrait de **ne pas réduire les droits des personnes** qui ne disposent pas de smartphone, en matière de déplacement.

Concrètement, si la sortie du confinement devait être conditionnée à la collecte de données de géolocalisation visant à un traçage des contacts, il serait nécessaire de prendre une loi spéciale, car un consentement forcé à la collecte de données individuelles ne serait pas conforme aux textes en vigueur. L'article 15 de la directive ePrivacy permet justement aux États membres de légiférer afin de restreindre les droits des individus, notamment en matière de consentement, afin de sauvegarder la sécurité publique. En outre, le périmètre des données collectées devrait être limité aux seules données strictement nécessaires pour contrôler les interactions des personnes dans le cadre de la lutte contre la propagation du coronavirus.

Les questions posées par l'utilisation de données de santé.

Les données de santé sont des « données sensibles » qui font l'objet d'une protection juridique renforcée, en particulier en matière de confidentialité des données (secret médical) et de conditions de conservation, avec l'exigence d'un hébergement certifié.

La collecte de données de santé par les professionnels de santé est toujours possible dans le cadre de la fourniture d'un service de santé, même sans consentement explicite des intéressés.

L'utilisation d'applications stockant des données de santé individuelles (groupe sanguin, vaccinations, paramètres biologiques personnels, résultats des derniers tests Covid-19) est possible mais ouvre une série de questions sensibles non résolues :

- Pourrait-on exiger la **présentation d'informations médicales pour conditionner l'accès à l'espace public** en phase de déconfinement ? Le secret médical s'oppose à une utilisation directe de données médicales ;
- Devra-t-on **communiquer certaines de ces informations à l'employeur**, par exemple pour mettre en œuvre des mesures de protection particulières ? Là encore, le secret médical ne permet pas un accès direct de l'employeur aux données de santé de ses salariés.

À plus long terme, un risque d'« effet cliquet » et de recours à des technologies plus intrusives

L'usage des nouvelles technologies et l'exploitation des données dans le cadre de la lutte contre les pandémies pose toutefois des questions éthiques et juridiques qui s'étendent, au-delà de la

28 Montjoye, *et al.* « Unique in the Crowd: The Privacy Bounds of Human Mobility ». *Scientific Reports* 3, n° 1 (25 mars 2013): 1-5. <https://doi.org/10.1038/srep01376>.

crise actuelle du Covid-19, aux conditions du « retour à la normale » d'une part, et à la nécessité de prévenir les prochaines crises d'autre part.

En effet, les technologies évoluent, et les menaces aussi, de sorte qu'il est *a priori* impossible d'exclure qu'un usage aujourd'hui inacceptable soit jugé demain acceptable, par exemple pour faire face à un péril bien plus grand. À cet égard, toute réflexion prospective sur le sujet doit donc :

- Prendre en compte les technologies en développement et pas seulement les technologies d'ores et déjà matures ;
- Envisager leurs implications permanentes et à long terme, c'est-à-dire non seulement en période de crise mais aussi « en temps normal » ;
- Faire l'hypothèse du pire, c'est-à-dire d'une pandémie majeure dont le taux de létalité et la contagiosité excèderaient largement ceux du Covid-19, et pour lesquels aucun traitement ou vaccin ne serait envisageable à court ou moyen terme.

Dans ces conditions, deux « basculements » sont porteurs de risques particuliers : le passage d'un état de crise à un état d'alerte permanent, et la mise en place d'outils plus efficaces mais aussi plus attentatoires aux libertés publiques et droits fondamentaux.

- « L'effet cliquet » ou le risque d'un état d'urgence sanitaire permanent

Une fois la phase pandémique terminée (phase 6 selon la classification de l'OMS) se posera la question du maintien – partiel ou total – des mesures d'exception. Trois questions se posent plus précisément :

- À quel moment lever les mesures d'exception mises en place pendant une crise ? Doivent-elles se prolonger tant que des *clusters* seront identifiés dans le monde, et/ou si la maladie prend la forme d'une épidémie saisonnière ?
 - À quel moment réactiver les mesures d'exception pour prévenir une nouvelle épidémie ? La crise actuelle a en effet montré qu'une réaction forte et précoce était efficace : faut-il, dès lors, mettre en place des mesures générales (signalements, collecte de données, contrôle des déplacements par géolocalisation etc.) dès la détection des premiers cas ?
 - Dans l'intervalle, les données collectées doivent-elles être conservées, et si oui, par qui et dans quelles conditions ?
- Des innovations technologiques à venir qui poseront la question d'un arbitrage difficile entre « efficacité » et « libertés »

La prévention des futures épidémies implique de s'interroger en responsabilité non seulement sur les technologies actuelles, mais aussi sur leurs probables développements futurs, qui pourraient être d'une grande efficacité mais posent des questions éthiques majeures et apparaissent peu compatibles avec le cadre juridique actuel. Plusieurs risques peuvent à cet égard être identifiés :

- **Le passage de mesures volontaires à des mesures contraignantes.** Aujourd'hui, les mesures envisagées en France ont un caractère strictement informatif (envoi de notifications) ou volontaire (application de diagnostic Covidom de l'APHP, application de localisation relative Trace Together utilisée à Singapour ; une application de ce type

étant étudiée en France). En cas de crise grave, la question du passage à des outils contraignants (port d'un bracelet électronique, signalement obligatoire etc.), le cas échéant assortis de sanctions automatiques, se posera nécessairement. De tels outils sont d'ores et déjà employés en Chine, en Corée du Sud, à Taïwan ou encore en Israël.

- **Une primauté accordée à la préservation du groupe au détriment du respect de la vie privée.** Certaines solutions technologiques pourraient reposer sur une collecte bien plus intrusive de données personnelles (données médicales²⁹ comme non médicales³⁰), exploitées à des fins d'analyse prédictive³¹, voire rendues publiques pour assurer un contrôle social par les pairs (signalement des personnes infectées ou suspectes au sein du voisinage). S'il existe une plus grande sensibilité des pays occidentaux au respect des droits individuels et de la vie privée que dans les sociétés asiatiques, et notamment dans les sociétés à tradition confucéenne, cette différence culturelle ne semble pas constituer une garantie absolue³² – *a fortiori* en cas de crise majeure. Il existe par ailleurs un risque de réutilisation des données collectées dans le cadre de la lutte contre une pandémie à d'autres fins³³.
- **Une dépendance croissante des acteurs publics envers les acteurs privés.** La crise sanitaire actuelle, comme du reste plusieurs catastrophes naturelles ou attaques terroristes récentes, ont mis en lumière le rôle incontournable des grands acteurs numériques, qui maîtrisent les technologies, collectent et détiennent les données, et sont aussi force de proposition ou d'initiative en la matière. Il s'agit des opérateurs télécoms³⁴, des GAFAs³⁵, sans même parler des BATX³⁶ chinois. Or, en France du moins, il n'existe pas à ce jour de « plan de contingence » permettant à l'État de s'assurer, en cas de crise, de la coopération de ces acteurs, selon des modalités qu'il maîtriserait.

Ce dernier point met également en lumière l'écueil qui consisterait à ne pas prendre en compte les enjeux de souveraineté, associés aux données de santé, alors que de grands groupes internationaux mettent des solutions technologiques à disposition des États, comme cela a été

29 Par exemple la collecte en temps réel des données de température corporelle ou de fréquence cardiaque pour tous les utilisateurs de montres connectées, dont le nombre va considérablement augmenter ces prochaines années.

30 Par exemple les recherches sur Internet, les interactions sur les réseaux sociaux, la liste des contacts, achats, déplacement etc.

31 Par exemple pour prendre des mesures individualisées s'il était établi que le moindre respect des mesures de confinement était corrélé à la catégorie socio-professionnelle, à l'âge, au lieu de résidence, à un profil médical ou génétique etc.

32 Aux États-Unis, des applications telles que *Criminal Watchdog* ou *Offender Locator* permettent d'ores et déjà de géolocaliser les délinquants sexuels présents dans le voisinage.

33 On pourrait par exemple imaginer que le bon respect – ou non – des mesures de confinement fasse à terme partie des critères du système de « crédit social » mis en place dans certaines métropoles chinoises.

34 Voir notamment l'estimation par Orange de la part de la population ayant quitté l'Île-de-France, et plus généralement le rôle des FAI dans l'identification des déplacements ou l'envoi de messages d'alerte.

35 Google, Apple, Facebook, Amazon et autres. Voir notamment le rôle de la fonction « *Safety Check* » de Facebook en cas de catastrophe naturelle ou d'attaque terroriste, utile non seulement aux individus mais aussi aux autorités (sécurité civile etc.).

36 Baidu, Alibaba, Tencent, Xiaomi. En particulier, l'application *WeChat*, éditée par Tencent et qui compte plus de 1,5 milliard d'utilisateurs, sert tout à la fois de messagerie instantanée, de moyen de paiement et de titre de transport, ce qui en fait un outil « naturel » pour la mise en œuvre de mesures plus ou moins contraignantes dans le cadre de la lutte contre une pandémie.

rappelé par le Comité pilote d'éthique numérique du Comité consultatif national d'éthique³⁷. En effet, si l'épidémie perdure, l'utilisation de ces technologies sur le temps long générera de précieuses données. La France dispose d'entreprises suffisamment avancées technologiquement pour répondre aux attentes des autorités de santé, tout en garantissant leur bon usage, dans le cadre européen précité.

Conclusions

Les outils disponibles – ou qui le seraient dans un temps relativement court – sont multiples et permettent d'envisager de nombreux usages des technologies de l'information.

La crise actuelle a mis en évidence les difficultés administratives auxquelles est confrontée la plateforme nationale de données de santé (*Health Data Hub*) ; grâce à un expédient, celle-ci s'est cependant mise en mesure de répondre aux demandes qui lui sont adressées, mais il est impératif que les difficultés rencontrées soient levées au plus vite. Les organismes publics qui centralisent des données individuelles doivent participer à la réponse collective en s'adaptant à l'urgence, en mettant à disposition de la recherche académique et des travaux de R&D, notamment dans le domaine médical, les outils et données nécessaires, tout en assurant le respect de la réglementation relative à la confidentialité des données personnelles.

L'efficacité des outils de localisation sur la propagation de la maladie n'est pas démontrée dans le contexte français ; nous n'avons d'ailleurs pas connaissance d'une évaluation rigoureuse de leur utilisation dans d'autres pays. Mais ces outils apporteraient certainement une contribution utile pour limiter la propagation de l'épidémie, accélérer son contrôle et son déclin. Leur mise en œuvre devra néanmoins obligatoirement s'inscrire dans le cadre d'une stratégie globale, en complément des mesures d'ordre sanitaire ou logistique à impact direct (faisant intervenir des équipements, masques, tests, traitements, etc.). Quand bien même ces outils seraient imparfaits, leur déploiement maîtrisé permettrait de les améliorer pour faire face à une recrudescence de l'épidémie et mieux se préparer à une future pandémie.

La réglementation européenne instaure un cadre très strict pour la protection des données personnelles, mais celui-ci ménage une souplesse suffisante pour faire face à la crise actuelle. À cet égard, la CNIL³⁸, le Comité pilote d'éthique numérique³⁹ et le Comité européen de protection des données⁴⁰ ont souligné que les projets d'applications de traçage des contacts tels qu'étudiés en France répondaient aux exigences de cette réglementation et ne nécessitaient pas

37 Dans ses « recommandations aux services d'État », le Comité national pilote d'éthique du numérique, dépendant du Comité consultatif nationale d'éthique (CCNE), invite à « privilégier des solutions numériques souveraines pour la gestion logistique tout particulièrement en période de crise ». <https://www.ccne-ethique.fr/sites/default/files/publications/bulletin-1-ethique-du-numerique-covid19-2020-04-07.pdf>

38 Propos liminaires de Mme Marie-Laure Denis, présidente de la CNIL, lors de son audition par la commission des lois de l'Assemblée nationale le 8 avril 2020 (https://www.cnil.fr/sites/default/files/atoms/files/propos_liminaire-audition_commission_des_lois-assemblee_nationale-8-04-2020.pdf).

39 « Recommandations aux service d'État » du Comité pilote d'éthique numérique, dépendant du CCNE. <https://www.ccne-ethique.fr/sites/default/files/publications/bulletin-1-ethique-du-numerique-covid19-2020-04-07.pdf>

40 *Statement on the processing of personal data in the context of the COVID-19 outbreak. European data protection board*, publié le 19 mars 2020 ; https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_statement_2020_processingpersonaldataandcovid-19_en.pdf

de légiférer, si de tels outils ne sont pas rendus obligatoires. La mise en œuvre d'outils plus intrusifs dans un contexte sanitaire de crise est donc possible, à condition que les atteintes à la vie privée restent strictement proportionnées au but recherché et qu'ils ne soient utilisés que pour une durée limitée.

Personnalités consultées

- Emmanuel Bacry, directeur scientifique du *Health Data Hub* et directeur du projet Data/Santé à l'École Polytechnique ;
- Stéphanie Combes, directrice du *Health Data Hub* ;
- Nicolas Demassieux, directeur de la recherche d'Orange ;
- Pr. Laurence Devillers, membre du Comité pilote d'éthique numérique du CCNE ;
- Pr. Arnaud Fontanet, directeur de l'unité Épidémiologie des maladies émergentes à l'Institut Pasteur, membre du Conseil scientifique Covid-19 ;
- Pr. Bernard Nordlinger, membre de l'Académie de médecine et chair du groupe de travail « Santé et intelligence artificielle » de l'Académie nationale de médecine et de l'Académie des sciences ;
- Bertrand Pailhès, directeur des technologies et de l'innovation et Hélène Guimiot-Breud, cheffe du service santé, pour la CNIL.