ÉTUDES
DE L'IFRI

PROGRAMME GÉOPOLITIQUE DES TECHNOLOGIES

JUILLET
2021

La gouvernance des données de santé

Leçons de la crise du Covid-19 en Europe, en Chine et aux États-Unis

Julie MARTINEZ
Clément TONON

L'Ifri est, en France, le principal centre indépendant de recherche, d'information et de débat sur les grandes questions internationales. Créé en 1979 par Thierry de Montbrial, l'Ifri est une association reconnue d'utilité publique (loi de 1901). Il n'est soumis à aucune tutelle administrative, définit librement ses activités et publie régulièrement ses travaux.

L'Ifri associe, au travers de ses études et de ses débats, dans une démarche interdisciplinaire, décideurs politiques et experts à l'échelle internationale.

ISBN : 979-10-373-0380-6

Comment citer cette publication :
Julie Martinez et Clément Tonon, « La gouvernance des données de santé : leçons de la crise du Covid-19 en Europe, en Chine et aux États-Unis », *Études de l'Ifri*, Ifri, juillet 2021.

Ifri

27 rue de la Procession 75740 Paris Cedex 15 – FRANCE

Tél. : +33 (0)1 40 61 60 00 – Fax : +33 (0)1 40 61 60 60

E-mail : accueil@ifri.org

Site internet : Ifri.org

Auteurs

Diplômée de trois masters européens de recherche en droit des affaires internationales et des nouvelles technologies, **Julie Martinez** est avocate au Barreau de Paris, spécialisée en droit des technologies de l'information et des données à caractère personnel.

Diplômé de l'École des hautes études commerciales (HEC) et de l'École nationale d'administration (ENA), **Clément Tonon** est haut fonctionnaire. Il est co-auteur de l'étude de l'Ifri « L'Europe : sujet ou objet de la géopolitique des données ? » parue en juillet 2018 et auteur de l'étude de l'Ifri « La GovTech, nouvelle frontière de la souveraineté numérique » parue en novembre 2020.

Résumé

La crise sanitaire a enclenché un mouvement tectonique dans la recomposition des modèles de gouvernance et de protection des données de santé dans le monde, tout en servant d'accélérateur à l'investissement des grandes entreprises du numérique dans le domaine de la e-santé. La présente étude propose une analyse comparative des modes de gestion des données de santé de l'Union européenne, des États-Unis et de la Chine dont les recompositions et les confrontations seront une composante essentielle des rapports de force futurs au sein de ce triangle stratégique. Les données étant la matière des flux numériques transnationaux, la compréhension de ces modèles permet de dépasser la simple dimension juridique pour concevoir leur articulation, leurs frictions et leurs influences mutuelles comme un enjeu de politique internationale à part entière entre ces trois pôles majeurs de l'économie numérique mondiale.

La pandémie de COVID-19 a révélé les carences des modèles de gouvernance préexistants dans chaque région du monde et la nécessité de tendre vers un modèle de gestion des crises de santé publique « par la donnée de santé ». Pour l'Europe et pour la France en particulier, la crise sanitaire a été en outre le révélateur d'une longue « innocence technologique » dont les conséquences sont en train d'être mesurées par les pouvoirs publics. Malgré ce réveil stratégique, des questionnements subsistent sur la capacité des Européens à coopérer efficacement et à avancer vers un espace commun du numérique en santé, sans disposer pour l'instant d'alternatives industrielles comparables aux géants du numérique américains et chinois.

Au terme de cette étude, un constat majeur semble s'imposer : l'une des ondes de choc de l'épidémie de COVID-19 sera numérique. En affectant durablement les rapports entre les États, les entreprises et les citoyens, la crise sanitaire a fait entrer les nations dans le XXIe siècle « géotechnologique », dans lequel les données de santé constitueront l'un des principaux vecteurs de puissance et de richesse.

Executive summary

The COVID-19 crisis has triggered a tectonic movement in health data governance and protection models around the world and has accelerated the investment of large digital companies in the field of e-health. This study offers a comparative analysis of the ways in which health data are managed in the European Union, the United States and China in this context. The ongoing reordering of and confrontations among the different governance models will be an essential component of the future balance of power within this strategic triangle. Since data is the subject of transnational digital flows, understanding these models allows to go beyond mere simple legal analysis to conceive of the articulation, frictions, and mutual influences. Health data has become full-fledged international policy issue between the three major poles of the global digital economy.

The COVID-19 pandemic has revealed the shortcomings of pre-existing governance models in every region of the world and the need to move towards a model of public health crisis management "through health data". For Europe and for France in particular, the health crisis has also revealed a long prevailing "technological innocence", whose consequences public authorities are now evaluating. Despite this strategic awakening, questions remain about the ability of Europeans to cooperate effectively and move towards a common digital space in health, without currently having industrial alternatives comparable to the American and Chinese digital giants.

At the end of this study, a major observation seems to be clear: one of the shock waves of the COVID-19 epidemic will be digital. As it has permanently affected the relationship between states, businesses and citizens, the health crisis has brought nations into the twenty-first "geo-technological" century, in which health data will be one of the main vectors of power and wealth.

Sommaire

Introduction

Le 20 mai 2021, l'émission Cash Investigation « Nos données personnelles valent de l'or ! », diffusée sur France 2, s'intéressait à la constitution d'un « entrepôt » de données de santé par la société américaine IQVIA à travers le réseau des pharmacies françaises. La mise en cause du fonctionnement de cette base de données a entraîné une réponse publique de la Commission nationale de l'informatique et des libertés (CNIL) pour rappeler le cadre légal européen et les garanties entourant cette exploitation[1]. Quelques jours plus tôt, le 17 mai 2021, le gouvernement français présentait sa « stratégie nationale pour le cloud » détaillant les outils (technologies sous licences, label cloud de confiance, investissements industriels…) qu'il compte mettre en place pour permettre aux entreprises et administrations françaises « d'avoir accès aux outils les plus performants du monde tout en garantissant un traitement des données respectueux des valeurs européennes[2] ». Cette stratégie a notamment pour ambition de protéger les données des Français « contre toute réglementation extracommunautaire », c'est-à-dire contre les lois, notamment américaines et chinoises, à portée extraterritoriale[3].

Comme l'illustrent ces événements récents, la crise du COVID-19 a placé les données – et surtout celles de santé, dont la production et les usages sont en plein essor –, au cœur du débat public français et européen sur l'articulation entre libertés fondamentales et souveraineté dans l'espace numérique. Elle a également servi de catalyseur à trois dynamiques structurelles qui sont au cœur de la présente étude.

D'une part, la pandémie a mis en valeur l'importance stratégique, pour les États, de comprendre et maîtriser les technologies de santé, ainsi que l'usage des données de santé des individus dans la gestion d'une crise majeure de santé publique. Ce faisant, elle interroge la pertinence des modèles de gouvernance et de régulation des données de santé élaborés ou en cours d'élaboration avant son déclenchement.

---

1. Commission nationale de l'informatique et des libertés, « *Entrepôt de données santé IQVIA : la CNIL rappelle les conditions et le cadre légal ayant permis son autorisation en 2018* », 17 mai 2021, disponible sur : www.cnil.fr.
2. Cédric O, Secrétaire d'Etat chargé de la Transition numérique et des Communications électroniques, lors de la conférence de presse de présentation de la « stratégie nationale pour le *cloud* » le lundi 17 mai 2021.
3. *Ibid.*

Plusieurs gouvernements ont d'ores et déjà engagé une réflexion sur une numérisation massive de leur système de santé et des modifications de leur réglementation pour tirer les conséquences de la crise : le gouvernement britannique a par exemple publié le 22 juin 2021 une ambitieuse stratégie intitulée « Data saves lives: reshaping health and social care with data[4] » contenant une liste de propositions et de modifications législatives à engager sur la période 2021-2030 « afin de libérer le potentiel des données dans le secteur de la santé et du soin, tout en conservant les plus hauts standards de protection de la vie privée, d'éthique et de confiance[5] ».

D'autre part, la crise sanitaire a démontré l'importance croissante, dans tous les pans de la vie sociale, de l'économie du numérique en santé, où se déploie un écosystème d'entreprises extrêmement dynamique et se positionnent la plupart des grands acteurs technologiques mondiaux. Entre 2013 et 2020, le volume de données de santé produites au niveau mondial a été multiplié par 15, passant de 153 exabytes à 2 314 exabytes[6]. Cette croissance a fait du marché du *big data* en santé l'un des secteurs les plus prometteurs au sein du marché de la e-santé. Selon certaines estimations, il devrait passer d'une valeur d'environ 11,5 milliards de dollars en 2016 à 70 milliards de dollars en 2025[7].

Enfin, la pandémie a exacerbé les rapports de force en construction au sein d'une véritable « géopolitique des données[8] » mondiale. Au-delà de la comparaison des modèles juridique et politique, il faut donc s'intéresser aux acteurs de cette compétition mondiale, c'est-à-dire aux complexes sanitaires, industriels, numériques et scientifiques qui permettent, au sein d'un espace donné, d'assurer la collecte et le traitement des données de santé en adéquation avec un corpus de règles et de valeurs. Dans la compétition numérique mondiale post-COVID, sans cet alignement entre un système de valeurs, un cadre juridique et un écosystème industriel, cette étude démontre qu'il ne peut réellement y avoir de puissance ni de souveraineté. Il existe ainsi, selon Cédric Villani et Gérard Longuet, un risque réel pour la France de « voir se développer l'usage d'algorithmes étrangers, qui ne seraient pas régis par les règles françaises, si notre pays ne se dote pas des moyens d'avancer au

4. Ce qui peut être traduit par « Les données sauvent des vies : repenser les soins de santé et la protection sociale grâce aux données ».
5. Department of Health and Social Care, « Data Saves Lives: Reshaping Health and Social Care with Data (draft) », 22 juin 2021, disponible sur : www.gov.uk.
6. Stanford Medicine Health Trends Report, « Harnessing the Power of Data in Health », 2017.
7. Statista, « Global Healthcare Big Data Market Size in 2016 and a Forecast for 2025 », disponible sur : www.statista.com.
8. A. Cataruzza, *Géopolitique des données numériques. Pouvoir et conflits à l'heure du* Big Data, Paris, Le Cavalier Bleu, 14 mars 2019.

rythme des autres[9] ». De même, pour le ministre allemand de la Santé, Jens Spahn, « la question [pour les Européens] est de savoir si l'offre doit venir uniquement de sociétés américaines ou si nous pouvons la développer chez nous, sur la base de critères européens de protection des données[10] ».

La présente étude propose une analyse comparative des modes de gestion des données de santé afin d'identifier des modèles juridiques et politiques dont l'enchevêtrement, l'opposition ou la mise en cohérence auront des répercussions diplomatiques significatives à moyen terme. D'abord, nous revenons sur les spécificités des données de santé au sein d'une typologie plus large des données numériques. Ensuite, nous présentons les modèles de gouvernance de l'Union européenne, des États-Unis et de la Chine, dont les recompositions et les confrontations seront une composante essentielle des rapports de force futurs au sein de ce triangle stratégique. Comme l'affirme le ministre de la santé allemand Jens Spahn, si le modèle européen s'est développé suivant une approche normative attachée à la protection des droits fondamentaux, il existe également des « modèles de surveillance policière ou capitaliste qui ne sont pas les nôtres[11] ». Enfin, nous analysons les changements dont la pandémie de COVID-19 a été un catalyseur : nous évaluons notamment les différences d'approches de la gestion de crise entre l'Union européenne, les États-Unis et la Chine, où les données de santé ont été au cœur des dispositifs de traçage des chaînes de contamination.

Cette étude démontre que la crise sanitaire a enclenché un mouvement tectonique dans la recomposition des modèles de gouvernance et de protection des données, largement centré autour de la donnée de santé, tout en servant d'accélérateur à l'investissement des grandes entreprises du numérique dans le domaine de la e-santé. Plus précisément, nous en tirons deux conclusions principales.

Premièrement, la pandémie de COVID-19 a révélé les carences des modèles de gouvernance préexistants dans chaque région du monde et la nécessité de tendre vers un modèle de gestion des crises de santé publique « par la donnée de santé ». Pour l'Europe et pour la France en particulier, la crise sanitaire a été en outre le révélateur d'une longue « innocence technologique[12] » dont les conséquences

9. C. Villani et G. Longuet pour l'Office parlementaire d'évaluation des choix scientifiques et technologiques, « L'intelligence artificielle et les données de santé », mars 2019.
10. Jens Spahn, dans une interview accordée aux Échos. Voir aussi N. Steiwer, « Données de santé : comment l'Allemagne tente d'échapper à l'emprise américaine », *Les Échos*, mars 2020.
11. *Ibid.*
12. T. Gomart, « Le COVID-19 et la fin de l'innocence technologique », *Politique étrangère*, vol. 85, n° 2, Ifri, juin 2020.

sont en train d'être mesurées par les pouvoirs publics. Malgré ce réveil stratégique, des questionnements subsistent sur la capacité des Européens à coopérer efficacement et à avancer vers un espace commun du numérique en santé, sans disposer pour l'instant d'alternatives industrielles comparables aux géants du numérique américains et chinois. Aux États-Unis, la pandémie a fait naître des tensions autour du modèle de gouvernance des données de santé face aux lacunes illustrées par la gestion technologique de la crise – notamment sur l'articulation entre le gouvernement fédéral et les autorités des États. Des changements sont annoncés par le Président Joe Biden, qui tendent vers un modèle de régulation global des données personnelles. En Chine, enfin, la pandémie a accéléré la transition vers un modèle de capitalisme de surveillance dont les réussites durant la crise sont à nuancer et qui pourrait occasionner à court terme une nouvelle lutte de pouvoir entre le régime et les géants chinois du numérique.

Deuxièmement, la crise du COVID-19 a accéléré l'irruption d'acteurs privés de dimension mondiale dans la gestion d'une crise mobilisant les compétences régaliennes des États[13]. Cette irruption, qui interroge à la fois le rôle de l'État et la notion même d'intérêt général, n'est pas neutre du point de vue des rapports entre la puissance publique et les acteurs privés. Les entreprises du numérique, qui se positionnent désormais sur toute la chaîne de valeur de la donnée de santé, de sa collecte brute via des objets connectés à leur traitement en masse à des fins assurantielles, ont désormais en main tous les leviers pour valoriser économiquement, et demain peut-être politiquement, cette collecte grâce au traitement des données massives et à l'aide de l'intelligence artificielle (IA).

Au terme de cette étude, un constat majeur semble s'imposer : l'une des ondes de choc durables de l'épidémie de COVID-19 sera numérique. En affectant durablement les rapports entre les États, les entreprises et les citoyens, la crise sanitaire a fait entrer les nations dans le XXIe siècle « géotechnologique », dans lequel les données de santé constitueront, sans aucun doute, l'une des « puissances du futur[14] ».

13. C. Tonon, « La GovTech, nouvelle frontière de la souveraineté numérique », *Études de l'Ifri,* Ifri, novembre 2020.
14. S. Grumbach et S. Frénot, « Les données, puissance du futur », *Le Monde*, 7 janvier 2013.

Spécificités des données de santé au sein de la « datasphère »

La centralité des données de santé dans les enjeux technologiques contemporains tient à leur double spécificité au sein de la typologie constitutive du monde numérique.

D'une part, leur définition est extrêmement large et évolutive : la notion recouvre à la fois les données de santé par nature (antécédents médicaux, prestations de soins réalisés, résultats d'examens, traitements, handicap, etc.), les données qui, du fait de leur croisement avec d'autres données, deviennent des données de santé (croisement d'une mesure de poids avec le nombre de pas quotidien, par exemple) ou encore les données qui deviennent des données de santé du fait de leur destination, c'est-à-dire de l'utilisation qui en est faite au plan médical[15]. À cette première distinction, il faut ajouter celle entre données personnelles de santé et données non personnelles de santé, c'est-à-dire notamment les données « anonymisées », dont le lien avec l'identité des personnes est totalement et durablement rompu.

À l'inverse, les données de santé ayant fait l'objet de « pseudonymisation » – procédé fréquemment utilisé dans la recherche médicale qui consiste à remplacer les données directement identifiantes (nom, prénom, etc.) par des données indirectement identifiantes (alias, numéro séquentiel, etc.), sont toujours considérées comme des informations concernant une personne physique identifiable et demeurent donc des données personnelles. Cette distinction est structurante pour le régime européen issu du Règlement général sur la protection des données de santé (RGPD) qui vient délimiter la frontière particulièrement mince, dans le cas des données de santé, entre leur caractère personnel et non personnel. Les données de santé faisant souvent partie intégrante de bases de données mixtes issues des dossiers médicaux électroniques, des essais

15. Comme le souligne le comité consultatif national d'éthique dans son avis précité du 29 mai 2019 : « Toute donnée primaire issue d'une activité humaine – même sans lien apparent avec la santé – peut contribuer – par son croisement avec d'autres données qui ne lui sont pas liées – à la création d'une information nouvelle relative à la santé d'une personne. Une donnée de santé ne peut plus se limiter aux seules données personnelles recueillies dans le cadre d'une prise en charge médicale (mesures biologiques, caractéristiques génomiques, données cliniques, etc.). »

cliniques ou des séries de données collectées par diverses applications de santé, le cadre normatif européen pose désormais que l'exécution des opérations de traitement initial et des opérations de traitement ultérieur des données est conforme à la réglementation sur les données personnelles[16].

D'autre part, le degré particulier de sensibilité des données personnelles de santé, consacré par la quasi-totalité des réglementations sur les données personnelles dans le monde, renforce les exigences de sécurité, de transparence et de responsabilité demandées par les sociétés démocratiques lors de leur collecte et de leur utilisation. En Europe, le RGPD fait ainsi une place à part aux données de santé, données génétiques et données biométriques au sein même de la catégorie des données sensibles figurant à son article 9, en disposant par ailleurs que « les États membres peuvent maintenir ou introduire des conditions supplémentaires, y compris des limitations, en ce qui concerne le traitement des données génétiques, des données biométriques ou des données concernant la santé ».

Cette sensibilité particulière, qui fait écho au dense et ancien corpus déontologique de la profession médicale en matière de protection du secret[17], est principalement causée par les risques particuliers que fait peser l'exploitation de ces données et les questions qu'elles soulèvent dans notre rapport à l'intime : leur croisement fragilise la frontière entre vie privée et vie publique et permet d'établir des profilages très précis pouvant aboutir à des discriminations ou des manipulations à des fins politiques ou commerciales. Leur collecte et leur analyse ouvrent ainsi la porte du champ de la santé à de nouveaux acteurs du numérique dont l'attachement aux principes de solidarité, d'équité et d'éthique au cœur des systèmes de santé occidentaux peut être plus distendu… Comme le souligne le comité consultatif national d'éthique pour les sciences de la vie et de la santé dans son avis « Données massives (*big data*) et santé : une nouvelle approche des enjeux éthiques » : « l'une des caractéristiques des données massives relatives à la santé est d'effacer les distinctions sur lesquelles repose la mise en œuvre des principes éthiques qui fondent la protection des droits individuels dans le champ de la santé […] soin et commerce deviennent plus

16. « Lignes directrices relatives au règlement concernant un cadre applicable au libre flux des données à caractère non personnel dans l'Union européenne », Commission européenne, 29 mai 2019.
17. Depuis le serment d'Hippocrate, la préservation du secret, la confiance entre le médecin et le patient et l'auto-contrainte dans l'exploitation d'un accès privilégié à l'information constituent des principes cardinaux de l'exercice de la profession de médecin comme de la recherche médicale.

difficiles à distinguer, conséquence de la transformation du soin et du marché de la santé[18] ».

De cette double particularité de la donnée de santé – son ubiquité et sa sensibilité[19] – peuvent à ce stade être tirées deux remarques dans le champ des rapports de puissance. Sur le plan de la sécurité, dans un contexte de concurrence internationale accrue entre acteurs publics et privés pour la maîtrise de la « datasphère », les données de santé sont devenues un terrain d'affrontement privilégié dans le cyberespace. Avec un doublement du nombre d'attaques en 2020[20], l'économie de la santé est désormais l'un des secteurs les plus ciblés par les cyberattaques dans le monde, avec des attaques de plus en plus fréquentes contre des structures hospitalières ou les systèmes d'information des systèmes nationaux de santé (dernière en date, le 14 mai 2021, la cyberattaque massive par « ransomware » ayant frappé le système de santé irlandais et causé sa paralysie pendant plusieurs heures[21]).

En France, cette montée en charge est très nette : le Secrétaire d'État au numérique, Cédric O, a affirmé en février 2021 devant le Sénat « que s'il y a eu 27 attaques majeures d'hôpitaux en 2020, il y en a une par semaine depuis 2021 », dont les attaques fortement médiatisées des hôpitaux de Dax et de Villefranche-sur-Saône[22]. On peut ainsi noter que l'Allemagne a connu en septembre 2020 la première cyberattaque mortelle, avec la paralysie du système informatique de l'hôpital universitaire de Düsseldorf, qui a provoqué le transfert des malades vers d'autres hôpitaux de la région et le décès en transit d'une patiente en état critique[23].

L'hypothèse d'un « Pearl Harbor numérique », occasionnant des décès massifs dans le monde physique via une attaque sur des infrastructures de santé, n'est plus à exclure. Ce phénomène s'est accentué dans le cadre de l'épidémie de COVID-19 et a poussé une quarantaine de diplomates et de personnalités de haut niveau comme Ban Ki-Moon, Madeleine Albright et Mikhail Gorbatchev à s'associer au comité international de la Croix-Rouge (CICR) pour appeler les

18. Comité consultatif national d'éthique pour les sciences de la vie et de la santé, Avis 130, « Données massives (*big data*) et santé : une nouvelle approche des enjeux éthiques », 29 mai 2019, disponible sur : MergedFile (ccne-ethique.fr).
19. « Concilier ubiquité et sécurité des données médicales », *Cahiers du CRID*, n° 32, INRIA Grenoble, 6 janvier 2011.
20. « Healthcare Cyberattacks Doubled in 2020 According to IBM X-Force », Compliance Home, mars 2021.
21. « Cyber Attack 'Most Significant on Irish State' », BBC, mai 2021.
22. Cédric O cité par *Le Figaro* le 17 février 2021 : Cybersécurité des hôpitaux : « 27 attaques majeures en 2020 et une par semaine en 2021 », disponible sur : lefigaro.fr.
23. F. Dèbes, « En Allemagne, une première cyberattaque mortelle », *Les Échos,* 19 octobre 2020.

États à s'accorder sur une régulation du cyberespace afin de prévenir les attaques contre les infrastructures de santé[24], chantier international majeur des mois à venir.

D'un point de vue politique et démocratique, la donnée de santé pose de façon abrupte la question de la confiance entre les citoyens et les gouvernements. Loin d'être un simple fait de politique interne aux États, cette problématique, illustrée notamment par la question des applications de traçage et du « pass sanitaire » pendant la crise du COVID-19, détermine en réalité leur capacité à collaborer efficacement au plan international et, pour les Européens, à avancer vers la création d'un véritable espace européen de santé. Comme l'affirmait Margareth Vestager, commissaire européenne à la société numérique, dans une interview du 7 mai 2021 : « Il est très important que nous puissions créer la confiance nécessaire à ce que les données de santé des citoyens européens puissent s'inscrire dans un espace européen des données de santé. En particulier, puisque nous savons que la majorité des Européens disent en fait non, nous ne sommes pas intéressés par le partage de nos données de santé à quelque fin que ce soit. »[25] Un sondage de janvier 2021 montre effectivement que 47 % des Européens interrogés s'opposent à ce que leur gouvernement partage leurs données de santé avec des entreprises privées telles que Google, alors que 45 % d'entre eux approuvent[26] (cf. schéma 1).

Les gouvernements sont ainsi confrontés à deux temporalités distinctes qu'il leur faut maîtriser : d'une part, la temporalité démocratique nécessaire pour conjurer la méfiance suscitée, notamment en Europe, par la collecte des données de santé par des États ou des entreprises privées ; d'autre part, l'urgence identifiée par les chercheurs et les industriels de ne pas passer à côté de la course mondiale autour des applications du numérique en santé. Comme le rappelle le rapport de l'Office parlementaire d'évaluation des choix scientifiques et technologiques intitulé « L'intelligence artificielle et les données de santé[27] », le seul vrai risque, pour un pays comme la France, « serait de ne pas s'ouvrir à [l'intelligence artificielle (IA)], au numérique et au pilotage par les données ». Le débat suscité outre-Manche par le projet du ministre britannique de la Santé, Matt Hancock, de centraliser les données de 55 millions de patients britanniques dans une seule base opérée par NHS Digital à des fins de

24. Comité International de la Croix Rouge, « Call to Governments: Work Together to Stop Cyber Attacks on Health Care », mai 2020.
25. M. Mc Mahon, « Margrethe Vestager Explains the EU's Position in the Global Battle for Data », Euronews, 17 mai 2021.
26. Center for the Governance of Change, « European Tech Insights 2021. Part I : How the Pandemic Altered our Relationship with Technology », 2021, disponible sur : www.ie.edu.
27. C. Villani et G. Longuet pour l'Office parlementaire d'évaluation des choix scientifiques et technologiques, « L'intelligence artificielle et les données de santé », mars 2019.

recherche médicale et d'amélioration de l'efficacité du système de santé, qui a provoqué un report de la date de l'« opt-out » (possibilité pour les patients de refuser la collecte de leurs données) du 1er juillet 2021 au 1er septembre 2021, révèle l'acuité de cette tension dans les démocraties occidentales dans le sillage de la crise sanitaire[28].

28. *The Guardian,* l'un des quotidiens les plus influents du Royaume-Uni, a par exemple pris position très officiellement pour un report de la réforme et l'organisation d'un débat démocratique sur la question, voir : « The Guardian View on Medical Records: NHS Data Grab Needs Explaining », *The Guardian*, 30 mai 2021 : disponible sur : www.theguardian.com.

Gouvernance et régulation des données de santé avant la crise du COVID-19

Avant la crise du COVID-19, l'Union européenne (UE), les États-Unis et la Chine ont élaboré des modèles de gouvernance et de régulation des données de santé distincts, dont la comparaison permet d'éclairer à la fois les différences structurelles, fondées sur des soubassements politiques et industriels hétérogènes, mais aussi les potentiels de rapprochement. Les données étant la matière des flux numériques transnationaux, la compréhension de ces modèles permet de dépasser la simple dimension juridique pour concevoir leur articulation, leurs frictions et leurs influences mutuelles comme un enjeu de politique internationale à part entière entre ces trois pôles majeurs de l'économie numérique mondiale, dont l'axe de gravité se déplace de plus en plus de l'Atlantique vers l'Asie[29].

L'Union européenne

Hétérogénéité de la gouvernance des données de santé en Europe

Trois types principaux de modèles d'organisation des données de santé en Europe peuvent être identifiés : le modèle décentralisé à l'allemande, le modèle ouvert nordique et le modèle centralisé à la française. L'analyse de cette hétérogénéité, dans les approches et dans la qualité des bases de données disponibles, permet d'identifier les points de convergence nécessaires pour tendre vers un véritable espace numérique de santé européen.

Allemagne

En novembre 2019, le marché de la e-santé en Europe était estimé à 234 milliards de dollars d'ici à 2023, soit une croissance de 160 % sur la période[30]. Si le marché de e-santé le plus dynamique d'Europe se trouve en Allemagne, avec une valeur estimée de 57 milliards d'euros

29. T. Gomart, J. Nocetti et C. Tonon, « Europe : sujet ou objet de la géopolitique des données », *Études de l'Ifri,* Ifri, juillet 2018.
30. BPI France, « E-santé : vers un marché de 234,5 milliards de dollars », novembre 2019.

en 2025 et une croissance de 19 milliards en un an[31], il pâtit de l'organisation fortement décentralisée du système de santé allemand. Celle-ci donne une place essentielle aux *Länder* dans la régulation de l'offre de soins et de la structure hospitalière, occasionnant fragmentation et hétérogénéité des bases de données. Au-delà d'un réel effet de rattrapage, le dynamisme allemand en matière de numérique en santé est favorisé depuis 2018 par le grand projet de numérisation porté par le ministre de la Santé, Jens Spahn. Ce dernier a clairement mis ce sujet au cœur de son agenda ministériel, avec pour ambition d'aligner les performances numériques du système de santé allemand sur les pays nordiques. À ce titre, le « Digitale Versorgung Gesetz », loi votée en novembre 2019 « visant à améliorer la numérisation et l'innovation », constitue une étape clef vers la numérisation du système de santé allemand.

Environ 73 millions d'assurés du régime d'assurance maladie obligatoire allemand ont désormais droit à la prescription à l'échelle nationale d'applications numériques de santé (DiGA) prises en charge par les caisses d'assurance maladie obligatoire. Depuis juin 2020, les entreprises de e-santé du monde entier peuvent demander en ligne une inscription accélérée de leur application de santé numérique (DiGA) à l'Institut fédéral des médicaments et des dispositifs médicaux (BfArM) en Allemagne. Dès son arrivée au ministère de la Santé, Jens Spahn a également lancé le projet de dossier électronique des patients ainsi que de délivrance électronique des ordonnances. Malgré ces investissements, les professionnels regrettaient avant le COVID-19 la précipitation avec laquelle la digitalisation des cabinets médicaux et des hôpitaux a été opérée dès 2019, au détriment des enjeux de sécurité. Harald Mathis, de l'institut Fraunhofer pour l'informatique (FIT), après avoir inspecté les installations d'une trentaine de cabinets, arrivait au constat que seul « un tiers des installations étaient sécurisées[32] ».

Pays nordiques

L'approche des pays du nord de l'Europe, comparable à celle du Royaume-Uni, est davantage centrée sur une gouvernance de la donnée de santé par l'*open data*, dont l'accès et l'utilisation sont laissés libres aux usagers et aux chercheurs. Très avancés sur le sujet de la gouvernance par la donnée de santé, la Norvège et la Suède comptaient respectivement 16 registres nationaux de données de santé (dont le premier a été créé en 1951) pour une population de 5 millions d'habitants et 73 registres nationaux pour 10 millions

31. R. Berger, « Future of Health - The Rise of Healthcare Platforms », septembre 2020.
32. N. Steiwer, « Données de santé : comment l'Allemagne tente d'échapper à l'emprise américaine », *Les Échos*, mars 2020.

d'habitants avant la crise sanitaire. Alliant outils analytiques, techniques d'Intelligence artificielle (IA) et de *machine learning*, la Suède, la Finlande, la Norvège et l'Islande partageaient avant la crise sanitaire un projet commun d'*open data* permettant de regrouper des données afin de créer une base commune de prescriptions de médicaments regroupant 25 millions de personnes[33].

France

Le système centralisé français, traditionnellement structuré autour du Système national d'information interrégime de l'assurance maladie (SNIIRAM), a permis de constituer de larges bases de données d'une qualité « sans équivalent en Europe, au regard du nombre de personnes concernées et de la diversité des données disponibles[34] ». Dans un rapport de 2016, la Cour des comptes regrettait toutefois que cette base « aux potentialités considérables en matière de santé publique, de recherche, d'efficience du système de soins et de maîtrise des dépenses[35] » ne soit pas exploitée au maximum de ses potentialités[36].

Si la loi du 26 janvier 2016 de modernisation de notre système de santé a permis de consolider la gouvernance des données de santé, celle-ci est encore en construction autour d'un partenariat entre l'État et les industriels du secteur. La loi du 26 janvier 2016 a également créé le Système national des données de santé (SNDS) qui constitue une base de données unique pour décrire la santé de la population à travers les recours au système de santé. Le SNDS vise ainsi à rassembler des bases de données médico-administratives en santé déjà existantes – telles que le SNIIRAM, la base de données « PMSI » (Programme de médicalisation des systèmes d'information, autrement dit, les données des hôpitaux) contenant les données d'analyse de l'activité des établissements de santé, et la base de données du CépiDc (Centre d'épidémiologie sur les causes médicales de décès).

La France a su structurer des bases de données spécialisées puis poser les fondements du SNDS. Toutefois, l'utilisation de ces bases demeure insuffisante[37]. En 2018, la Cour des comptes relevait que cette situation contrastait « avec des pays qui se sont le plus rapidement engagés dans cette voie, en particulier la Suède et à un

33. « Bridging Nordic Data », Deloitte, 2020, disponible sur: www.norden.diva-portal.org.
34. Cour des comptes, « Les données personnelles de santé gérées par l'assurance maladie », rapport, mars 2016.
35. *Ibid.*
36. « Les données de santé : une mine d'or pour l'économie française ? », Bignon Lebray, 15 janvier 2020.
37. E. Chapel, D. Gruson, D. Jaafar *et al.*, « La révolution du pilotage des données de santé : Enjeux juridiques, éthiques et managériaux », LEH Éditions, mai 2019.

moindre degré l'Italie et le Royaume-Uni » tout en soulignant l'absence de dispositif de prescription électronique de médicaments, de dispositifs médicaux ou prestations médicales occasionnant d'importantes pertes financières pour l'Assurance maladie[38]. Ce retard relatif de la France dans la valorisation des données de santé explique la tentative de relance de cette politique publique en 2019 avec le rattachement d'une Direction du numérique en santé au cabinet du ministre de la Santé, ayant pour mission de piloter la feuille de route ministérielle du numérique en santé jusqu'en 2022 et d'assurer la tutelle de l'Agence du numérique en santé en charge de l'établissement de référentiels, du développement de l'interopérabilité des bases, mais également des enjeux éthiques et de cybersécurité afférents.

Enfin, dans le sillage du rapport Villani sur l'Intelligence artificielle publié au mois de mars 2018[39] et du discours du Président de la République du 18 septembre 2018 sur le plan « Ma santé 2022 », la loi du 24 juillet 2019 relative à l'organisation et à la transformation du système de santé prévoit l'élargissement du SNDS à l'ensemble des données de santé associées à un acte bénéficiant d'un remboursement de l'Assurance maladie. Elle permet également la création du Health Data Hub (HDH), plateforme technologique ayant vocation à être ouverte à de nombreuses parties prenantes (monde de la recherche, associations de patients, entreprises…), dont la structuration a soulevé la question de la sécurisation de l'hébergement des données de santé en France et en Europe, étudiée dans la seconde partie de cette étude.

Les efforts pour créer un véritable espace numérique de santé européen

Avant le déclenchement de l'épidémie de COVID-19, l'Union européenne ne disposait pas d'un cadre réglementaire ou de gouvernance permettant le pilotage d'une crise sanitaire par la donnée de santé à une échelle communautaire. Cette carence était en réalité en germe dans la construction européenne, avec l'échec de la « Communauté européenne de la santé », présentée pour la première fois en 1952 par le gouvernement français et son ministre de la santé de l'époque, Paul Ribeyre[40]. Ce « pool blanc » – par complémentarité avec le « pool noir » constitué par la Communauté européenne du

38. Cour des comptes, « La Sécurité Sociale - Rapport sur l'application des lois de financement de la sécurité sociale », rapport, octobre 2018.
39. C. Villani, « Donner un sens à l'intelligence artificielle », mars 2018.
40. Voir A. Davesne, « L'Europe de la santé, une histoire française ? », *in* G. Coron (dir.), *L'Europe de la santé. Enjeux et pratiques des politiques publiques*, Rennes, Presses de l'EHESP, 2018, p. 19-40.

charbon et de l'acier – visait à créer une interdépendance entre les États en créant un régime commun d'assurance maladie, en harmonisant les politiques de santé publique et en mutualisant des efforts de recherche en médecine.

De façon significative, ce tournant manqué est concomitant de l'échec de la Communauté européenne de défense (CED), dont il avait été pensé comme le pendant fonctionnaliste, avec des ambitions stratégiques similaires. Comme l'affirmait Paul Ribeyre : « C'est à une nouvelle forme de Communauté européenne de défense que le gouvernement français vous demande de participer : il s'agit cette fois d'une Communauté européenne de défense contre la souffrance et la maladie[41] ». Butant sur les intérêts des industriels français et sur l'opposition des gaullistes, le projet est abandonné : les traités consacrent le fait que la santé n'est qu'une compétence communautaire partagée et que chaque État membre demeure responsable de sa propre politique et gouvernance de santé. Dans le silence des textes, l'espace européen de santé a été mis en sommeil pour des décennies (à l'exception notable du médicament) et émerge paradoxalement à nouveau sous l'angle de la donnée, avec les grandes régulations de la fin des années 1990.

Sans faire l'objet d'une régulation spécifique dans l'Union européenne, les données de santé sont particulièrement protégées au titre des régulations générales en matière de données personnelles depuis l'adoption, en 1995, d'une première directive 95/46/CE en ce sens[42]. Avec l'entrée en vigueur, le 25 mai 2018, du RGPD, les données personnelles de santé font l'objet d'une définition spécifique et extensive au niveau européen : constituent des données personnelles de santé toutes les données à caractère personnel « relatives à la santé physique ou mentale d'une personne physique y compris la prestation de services de soins de santé qui révèlent des informations sur l'état de santé de cette personne ». Cette définition large emporte une protection renforcée au titre des données « sensibles », fondée sur le principe d'une interdiction générale de la collecte des données de santé qui ne peut être levée que dans des circonstances très strictes et limitativement définies telles que par l'obtention du consentement éclairé et exprès du patient ou dans le cadre d'activités de la sécurité sociale, de médecine préventive et pour des intérêts de santé publique.

41. « Exposé de Paul Ribeyre à la conférence préparatoire à la Communauté européenne de la santé », Paris, 12 décembre 1952, *in* « Notes et documents concernant la Communauté européenne de la santé », *Notes et études documentaires*, n° 1718, série sociale XXV, 1953, p. 15.
42. 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.

Pour beaucoup d'acteurs du secteur de la e-santé, cette approche restrictive du RGPD est donc une entrave à la fois à l'exploitation commerciale des données de santé, en empêchant la constitution d'écosystèmes industriels européens de taille critique et en renforçant paradoxalement le monopole des acteurs installés[43], mais également au progrès de la recherche fondamentale[44]. Pour d'autres, le RGPD est la marque du modèle numérique européen et peut offrir un avantage comparatif à l'innovation de santé « à l'européenne[45] ».

Il faut également noter que le caractère protecteur du RGPD n'a pas empêché l'anticipation de la question de la résolution de crises sanitaires par le traitement de données de santé. En effet, le RGPD offre assez de souplesse pour mettre en œuvre des traitements de données liés à une pandémie : l'interdiction de traitement des données personnelles de santé peut être levée lorsque le droit de l'Union ou d'un État membre le prévoit, et lorsque l'intérêt public le commande, notamment à des fins de contrôle de maladies transmissibles et d'autres menaces graves pour la santé[46]. Malgré les difficultés pratiques soulevées par de nombreuses entités, l'approche européenne avait déjà intégré la nécessité d'une continuité des flux et l'importance d'une coopération internationale en matière de données de santé, ces dérogations permettant l'échange international de données entre services chargés des questions relatives à la santé publique, par exemple aux fins de la recherche des contacts des personnes atteintes de maladies contagieuses[47].

Sur le plan des données non personnelles de santé, l'Union européenne a lancé plusieurs initiatives visant à mettre en place une économie fondée sur le partage et la coopération : le règlement relatif à la libre circulation des données à caractère non personnel dans l'Union européenne[48], le règlement sur la cybersécurité[49], et la directive sur les données ouvertes[50] ont posé les

43. A. Lazarègue, « Là où le RGPD a échoué, le droit de la concurrence peut encore gagner », *Le Monde*, 14 juin 2019.
44. « International Sharing of Personal Health Data for Research », ALLEA, EASAC, FEAM, avril 2021.
45. L. Morlet et A. Templier, « Notre cadre juridique est un rempart contre l'appétit des GAFA pour nos informations médicales », *Le Monde*, 20 juin 2019.
46. RGPD, considérant 52.
47. RGPD, considérant 112.
48. Règlement (UE) 2018/1807 du Parlement européen et du Conseil du 14 novembre 2018 établissant un cadre applicable au libre flux des données à caractère non personnel dans l'Union européenne.
49. Règlement (UE) 2019/881 du Parlement européen et du Conseil du 17 avril 2019 relatif à l'ENISA (Agence de l'Union européenne pour la cybersécurité) et à la certification de cybersécurité des technologies de l'information et des communications, et abrogeant le règlement (UE) n° 526/2013 (règlement sur la cybersécurité).
50. Directive (UE) 2019/1024 du Parlement européen et du Conseil du 20 juin 2019 concernant les données ouvertes et la réutilisation des informations du secteur public.

premières pierres d'un véritable marché européen de la donnée de santé. En 2018, l'adoption du règlement 2018/1807 visait à favoriser la circulation des données non personnelles, dont les données de santé. L'Union européenne y encourage le développement de nouvelles applications technologiques en matière de santé et les dynamiques industrielles dans l'Union.

La stratégie commune pour l'Intelligence artificielle parue en 2018 prévoit des investissements massifs dans l'IA de santé qui visent à compenser le retard industriel européen perçu par rapport aux États-Unis et à la Chine, GAFAM (Google, Amazon, Facebook, Apple et Microsoft) et BATX (Baidu, Alibaba, Tencent, Xiaomi) en tête[51], mais également les avancées du Japon en matière d'IA et de robotique médicale, de la Corée du Sud, du Canada, d'Israël ou de la Russie, pays avancés ayant développé des stratégies ambitieuses d'investissement dans le numérique en santé. Afin de décliner ces stratégies de façon spécifique aux données de santé, la Commission européenne a lancé la construction d'une infrastructure de services numériques (eHealth) afin de fournir des services transfrontaliers fondamentaux pour la communication électronique des actes médicaux et des profils de patients. Le 23 mars 2017, l'Union européenne a publié un cadre pour l'interopérabilité des dossiers de santé électroniques existants, le « European Interoperabilty Framework », aboutissant à une recommandation du 6 février 2019 relative à un format européen d'échange des dossiers de santé informatisés, qui est cependant encore loin de connaître une adhésion massive par tous les États membres[52].

Enfin, le 19 février 2020, quelques jours après l'irruption de la pandémie sur le sol européen, la Commission européenne a présenté sa stratégie pour les données, avec un volet spécifique relatif aux données de santé. Consciente des lacunes en matière de coopération et de partage des données de santé, la Commission affirmait alors qu'« il est capital, pour l'innovation dans le secteur des soins de santé, de renforcer et d'étendre l'utilisation et la réutilisation des données de santé », notamment avec l'objectif d'améliorer le pilotage des systèmes de soin, de développer la compétitivité de l'industrie européenne et d'améliorer l'évaluation des produits médicaux[53]. L'UE définit dans ce document une ligne de crête entre son approche traditionnelle fondée sur la protection des libertés fondamentales des

51. E. Chapel, D. Gruson, D. Jaafar *et al.*, « La révolution du pilotage des données de santé : Enjeux juridiques, éthiques et managériaux », *op. cit.*
52. Commission européenne, « Services électroniques de santé transfrontières », non daté, disponible sur : www.ec.europa.eu.
53. « Une stratégie européenne pour les données », Communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au Comité des Régions, février 2020.

données de santé, en rappelant qu'elles doivent être traitées conformément aux dispositions du RGPD en vertu desquelles les données relatives à la santé méritent une protection spécifique, tout en affirmant pleinement leur importance économique : « la santé est un domaine dans lequel l'UE peut tirer parti de la révolution des données, en augmentant la qualité des soins de santé, tout en réduisant les coûts ». Avant le déclenchement de la crise sanitaire, la Commission observait déjà une fragmentation au sein des États membres et entre ceux-ci, ainsi qu'une diversité des modèles de gouvernance pour l'accès aux données que la stratégie dévoilée en février visait à combler.

Les États-Unis

Une réglementation des données de santé moins protectrice que dans l'Union européenne

Aux États-Unis, le traitement des données de santé est régi par plusieurs dizaines de réglementations, autant au niveau fédéral qu'étatique. Bien que le projet soit souvent évoqué, les États-Unis ne disposent pas à ce jour d'une réglementation globale de protection des données sur le modèle européen mais uniquement de législations sectorielles[54]. De cet ensemble réglementaire se dégage une approche de la protection des données différente de l'approche européenne, valable également pour les données de santé. Alors qu'en Europe, la protection de la donnée personnelle de santé est rattachée à un droit fondamental, les États-Unis envisagent d'abord la question du caractère commercial ou non de ces données. Même si certaines données – par exemple, collectées par les hôpitaux – bénéficient d'une protection élevée, les entreprises privées sont libres de les exploiter pour autant qu'elles ne commettent pas de « pratique déloyale[55] ».

La première grande loi sur la protection des données à caractère personnel, le *Privacy Act* de 1974, concernait les traitements de données effectués par le gouvernement fédéral et mettait en œuvre les « Fair Information Practices » développées par le ministère de la santé américain en 1973. Dans la lignée du *Privacy Act* de 1974, le

54. Le *Children's Online Privacy Protection Act* (COPPA) de 1998 pour les données personnelles des enfants ; le *Fair and Accurate Credit Transactions Act* (FACTA) de 2003 pour les données financières, etc.

55. W. J. Maxwell, « La protection des données à caractère personnel aux États-Unis : Convergences et divergences avec l'approche européenne », *in* B. Fauvarque-Cosson et C. Zolynski (dir.), *Le Cloud computing. L'informatique en nuage*, Actes de Colloque, Paris, Société de législation comparée, 2014.

législateur fédéral a développé une série de lois visant la protection des données à caractère personnel dans le secteur privé, notamment le *Health Insurance Portability and Accountability Act* (HIPAA) de 1996 qui constitue toujours le cadre principal de protection des données de santé. L'HIPAA exige des organisations du secteur de la santé qu'elles mettent en œuvre des mesures de sécurité techniques et organisationnelles détaillées, qu'elles divulguent les pratiques de traitement des données aux patients et qu'elles obtiennent leur consentement dans certaines situations limitativement énumérées.

L'HIPAA s'applique à différentes entités telles que les hôpitaux, les prestataires de soins de santé et leurs prestataires de services de facturation. Sa règle de confidentialité régit l'utilisation des informations de santé protégées (PHI) – définies, comme toute information détenue, par les entités couvertes concernant l'état de santé, la fourniture de soins de santé ou le paiement de soins de santé et pouvant être liée à un individu. La divulgation de ces informations est basée sur l'autorisation écrite expresse du patient, avec quelques exceptions pour des circonstances spécifiques, telles qu'une obligation légale de divulgation ou si la divulgation permet à l'entité réceptrice d'aider l'entité divulgatrice à remplir ses fonctions de soins. C'est d'ailleurs en règle générale sur cette base légale que les hôpitaux peuvent partager les informations des patients, sans obtenir leur consentement, à des entités privées telles que les GAFAM.

L'HIPAA a été complétée sous l'administration Obama par deux réglementations en matière de données de santé. En 2009, le Congrès adoptait le « Health Information Technology for Economic and Clinical Health » (HITECH), élargissant la portée des exigences de l'HIPAA en matière de protection des données, notamment en interdisant dans certains cas la vente d'informations protégées sans le consentement des patients[56]. En 2013, le *Genetic Information Nondiscrimination Act* (GINA) interdisait l'exploitation à des fins commerciales des données génétiques détenues par les acteurs de la santé et de l'assurance. L'élargissement des exigences de protection des données dans l'HIPAA était ainsi un moyen d'atténuer les risques liés aux nouvelles technologies en e-santé, qui ont mis en ligne des quantités importantes de nouvelles informations médicales. En plus de ces lois fédérales, chacun des États américains a adopté des lois visant la protection de certains aspects de la vie privée de leurs citoyens, dont le *California Consumer Privacy Act* (CCPA), entré en

56. L'HITECH a également augmenté la responsabilité légale en cas de non-respect et a renforcé les droits d'application de l'Office of Civil Rights (OCR, une division du Department of Health & Human Services), cf. P. Bailin, « Executive Summary: Evolution of Health Data Regulation », avril 2019.

vigueur le 1er janvier 2020, qui reprend, sans être aussi ambitieux, les principales caractéristiques du RGPD européen.

Enfin, un des objectifs principaux de l'HITECH, associé à un budget de 37 milliards de dollars inscrit dans le *American Recovery Act* de 2009 pour stimuler la numérisation du secteur médical, était de promouvoir une « utilisation significative » des dossiers de santé informatisés certifiés sur le territoire américain. Six ans après l'adoption de la loi, en 2015, force est de constater que les avancées étaient spectaculaires : 96 % des hôpitaux et 78 % des médecins utilisaient une technologie de dossier de santé électronique certifiée, contre respectivement 10 % des hôpitaux et 17 % des médecins au moment de la promulgation de la loi[57].

Une gouvernance de données fragmentée, ouvrant la voie aux ambitions des géants du numérique

Face au constat de la fragmentation du système de soin américain, de nombreux projets de transformation visant à améliorer la coordination des acteurs ont été engagés au niveau national par des acteurs publics et privés. D'une part, le développement de l'*open data* en santé, ancien aux États-Unis, s'est accéléré sous l'impulsion de l'administration Obama. Dès les années 1960, le gouvernement fédéral a pris le parti d'offrir un meilleur accès public aux données via la loi sur l'ouverture des données fédérales au public (le *Freedom of Information Act*). À partir de 2009, dans le sillage de l'« Open Government Initiative » lancée par Barack Obama au premier jour de son mandat puis de l'*Affordable Care Act* de 2010, le ministère de la Santé américain a permis l'ouverture de nouvelles données consultables sur un site dédié à l'open data en santé, « healthdata.gov[58] ».

D'autre part, plusieurs initiatives ont vu le jour afin de constituer des bases de données fédérales unifiées. En 2010, l'initiative « Blue Button » a fédéré l'administration des anciens combattants (VA), le ministère de la Défense (DoD) et l'assurance maladie fédérale (Medicare et Medicaid), les trois institutions publiques disposant des plus grandes bases de données de santé du pays, afin de permettre à leurs usagers de télécharger l'ensemble des données de leur dossier

57. « E-santé : augmentons la dose », Institut Montaigne, juin 2020.
58. « Panorama international de l'*open data* en santé : principaux enseignements », Département de l'information médicale, non daté, disponible sur : www.departement-information-medicale.com.

médical[59]. Le Vice-Président Joe Biden s'était fortement engagé à l'époque aux côtés de l'Institut national pour le cancer, dans le cadre de l'initiative « Cancer Moonshot », afin que les supercalculateurs du ministère de l'Énergie puissent analyser les données médicales détenues par l'administration des anciens combattants à des fins de recherche contre le cancer. Sous le second mandat de Barack Obama, 215 millions de dollars ont été consacrés à l'Initiative pour la médecine de précision visant à constituer une vaste banque de données concernant les gènes ainsi que d'autres éléments biologiques auprès d'un million de volontaires. Cet effort a abouti en 2017 avec le lancement du programme « All of Us » géré par les National Institutes of Health (« NIH[60] »), l'un des programmes de recherche les plus importants du monde dans le domaine.

Les initiatives du gouvernement fédéral se sont révélées insuffisantes pour unifier ou lier entre elles de façon efficace les bases de données de santé aux États-Unis, peu interopérables et alimentées par de nombreux agrégateurs publics et privés[61]. Face à cette fragmentation, de nombreux acteurs privés se sont engouffrés dans le segment des services de stockage, de coordination et d'agrégation. Durant les années 2010, les GAFAM ont ainsi investi le marché de la donnée de santé en nouant des partenariats avec des réseaux d'hôpitaux et de cliniques centrés autour, d'une part, de leurs solutions *cloud* pour héberger des volumes de données croissants et, d'autre part, de leur expertise dans l'IA pour accompagner les acteurs médicaux dans leurs projets de recherche[62]. Le Projet Nightingale de partenariat entre le gestionnaire de cliniques Ascension, Google et de nombreux hôpitaux, révélé en 2019 par le *Wall Street Journal*, visait par exemple à transférer les informations numérisées de millions de patients vers l'environnement Google Cloud[63].

De nombreux partenariats ont également été noués ces dernières années entre les GAFAM et des laboratoires pharmaceutiques tels que

59. D'après les propos d'un haut fonctionnaire français, spécialiste du numérique et des politiques de santé, recueillis par les auteurs en avril 2021, l'initiative « *Blue Button* » a ainsi changé la philosophie de gestion des données de santé en rendant la maîtrise de la donnée aux patients.
60. L. Determann, « Healthy Data Protection », *Michigan Technology Law Review*, vol. 26, n° 229, 2020.
61. Table ronde de l'Institute of Medicine (US) sur les soins de santé axés sur la valeur et la science, cf. *Clinical Data as the Basic Staple of Health Learning: Creating and Protecting a Public Good. Workshop Summary*, Institute of Medicine, National Academies Press, 2010, disponible sur : www.ncbi.nlm.nih.gov.
62. « Post COVID-19 : Comment la valorisation des données va-t-elle façonner la médecine de demain ? », BeraringPoint France, 2020.
63. R. Copeland, « Google's "Project Nightingale" Gathers Personal Health Data on Millions of Americans », *The Wall Street Journal,* 11 novembre 2019.

Novartis, Pfizer ou encore entre Sanofi[64], permettant à ces derniers de proposer, sur la base de la collecte et de l'analyse de données en vie réelle, des solutions allant désormais bien au-delà de la simple vente de médicaments (prévention de maladies, gestion des soins et propositions d'après-traitements). L'activisme des grandes entreprises du numérique dans le domaine des données de santé touche également des acteurs publics : en 2019, à l'initiative de Medicare, le système d'assurance maladie fédéral du gouvernement, un projet de centralisation des données de patients américains a été rejoint par Apple, Google et Microsoft. Les GAFAM ont offert le support logistique pour la création d'une interface de programmation dans laquelle chaque professionnel de santé peut déposer les documents de son patient[65].

Au-delà de ces partenariats institutionnels, l'ambition des géants du numérique d'intégrer de manière durable le marché des données de santé est facilitée par l'essor des objets connectés – montres connectées ou encore bracelets à usage médical[66] – permettant de collecter de manière massive et en temps réel des données de santé (ainsi que toute autre donnée personnelle utile en cas de croisement). Avec le rachat, le 15 janvier 2021, d'un des leaders mondiaux du secteur, Fitbit, pour 2,1 milliards de dollars, Google se positionne sur ce marché jusque-là dominé par Apple (46,4 % de part de marché) et Samsung (16 % de part de marché[67]). Une étude récente permet de montrer, sur la base de l'analyse de plus de 20 000 applications de santé mobile, que Google et Facebook sont également les principaux

64. Qui a développé la plateforme collaborative Darwin, regroupant une large variété de données de santé couvrant plus de 345 millions de patients, plus de 300 maladies et 48 études cliniques.

65. A. Burgat, « Les GAFA s'immiscent dans les systèmes de santé publics », *Les Échos Entrepreneurs*, 5 septembre 2019.

66. Pour illustrer quelques projets santé des GAFAM : Alphabet, la maison mère de Google, détient une filiale - Verily - qui collabore avec de nombreuses entités privées au développement de technologie médicale ainsi qu'avec de multiples groupes pharmaceutiques (on peut citer Novartis et Sanofi). Alphabet a mis en place différents projets, tels que le Project Baseline visant à « cartographier la santé humaine » et CityBlock afin de favoriser la prévention et l'éducation à la santé et Calico sur la prévention du cancer. Facebook a quant à lui acquis Moves, spécialisé dans le suivi de l'activité physique, puis Oculus, qui opère dans la réalité virtuelle. La plateforme propose également des encarts pour de la publicité ciblée pour les laboratoires pharmaceutiques. Amazon a lancé AmaronCare, etc.

67. Cette acquisition avait suscité les craintes du régulateur européen des données, affirmant dans une note du 19 février 2021 qu'« il est à craindre que la poursuite de la combinaison et de l'accumulation de données personnelles sensibles concernant les personnes en Europe par une grande entreprise de technologie puisse entraîner un haut niveau de risque pour les droits fondamentaux au respect de la vie privée et à la protection des données personnelles. » (EDPB, « Statement on privacy implications of mergers », 19 février 2020, disponible sur : www.edpb.europa.eu). En matière de santé mobile, si les Européens avaient dans un premier temps défendu le principe d'une régulation par un « code de conduite » élaboré par l'industrie, rejeté en 2018, ce sont désormais les dispositions du RGPD sur les données sensibles qui s'appliquent.

collecteurs de données par leur intermédiaire, via leurs différents services de publicité et de traçage directement intégrés dans ces applications[68].

La collecte sans précédant de volumes de données de santé permet enfin aux géants du numérique de se positionner sur des segments aval à forte valeur ajoutée. D'une part, la nouvelle frontière de l'ambition des GAFAM en matière de données de santé réside dans leur capacité à mener des recherches de haut niveau de façon indépendante et à investir le champ de la connaissance scientifique : l'application « Research » développée par Apple, propose par exemple de participer à des études médicales sur plusieurs années en collectant les données des utilisateurs d'un iPhone. Tim Cook, le président de l'entreprise, n'a pas hésité à affirmer que « dans le futur, quand on se demandera quelle a été la contribution d'Apple la plus importante pour l'humanité, la réponse sera la santé[69] ». Le 1er janvier 2020, une étude co-rédigée par des chercheurs de Google (Alphabet) et différents centres de recherche d'hôpitaux américains et britanniques analysait la lecture d'une mammographie par une IA pour dépister des cancers du sein[70]. Cette transformation accélérée du paysage de la santé par les GAFAM, qui investissent désormais massivement dans l'IA de santé, suscite désormais un débat intense dans les instances scientifiques internationales[71].

D'autre part, les géants du numérique ont fortement augmenté leurs investissements dans le secteur de l'assurance depuis 2015 : c'est notamment le cas de Google, avec par exemple un investissement de 375 millions de dollars dans la licorne de l'*insurtech* Oscar Health en août 2018 ou encore, en 2019, dans Clover Health, une solution d'assurance maladie qui s'appuie sur l'analyse prédictive de données pour offrir une couverture santé ciblant les seniors. Les autres géants du numérique ont suivi ce mouvement : Apple a noué un partenariat avec Aetna, un assureur américain qui développe une offre d'assurance maladie individualisée,

68. Ikram *et al.*, « Mobile Health and Privacy: Cross Sectional Study », *British Medical Journal*, 17 juin 2021 : « For most of the 20,000 medical and health and fitness apps analysed, we found that most can collect and potentially share data with third parties, including advertising and tracking services [...]. The analysis also revealed that mHealth apps were far from transparent when dealing with user data, with only about half being compliant with their declared privacy policies (if available at all). »
69. Entretien du 9 janvier 2019 avec Jim Cramer sur CNBC, disponible sur : www.cnbc.com.
70. S. M. McKinney, M., Sieniek, V. Godbole*et al.* International Evaluation of an AI System for Breast Cancer Screening. *Nature*, n° 577, 2020, p. 89-94, disponible sur : www.nature.com.
71. P. Corvol, médecin et chercheur en biologie, ancien administrateur du Collège de France : « L'usage de l'IA avec en corollaire la collecte des données est désormais un sujet suffisamment fort et évolutif pour nécessiter des mises au point régulières entre les différentes académies des sciences du G7 », cité dans l'article de L. Belot, « Les données de santé, un trésor mondialement convoité », *Le Monde*, 2 mars 2020.

reposant sur le recueil des données de l'Apple Watch et Amazon a investi dans l'*insurtech* indienne Acko. Cette pénétration des acteurs du numérique dans le marché assurantiel, qui inquiète les grandes entreprises du secteur[72], est également à l'œuvre en Chine où Ant Financial, filiale d'Alibaba et plus grosse *fintech* au monde, a racheté l'assureur Cathay Insurance China en 2015 et investi dans le géant de l'assurance en ligne Zhong An, valorisé à 10 milliards de dollars.

La Chine

Un contrôle renforcé sur les données pour numérisation accélérée du système médical

Les premières législations chinoises relatives au numérique en santé datent de 2014[73] mais c'est la loi sur la cybersécurité de 2017[74] qui est venue unifier pour la première fois le régime applicable aux données personnelles, avec comme principal objectif de sécuriser l'espace informationnel et numérique chinois. À ce titre, les données de santé sont considérées comme sensibles et ne peuvent pas faire l'objet d'un stockage hors du pays. Leur transfert à l'étranger est soumis au consentement de leur propriétaire, à moins qu'une raison impérieuse tenant à la sécurité ou aux intérêts économiques de la Chine ne s'y oppose. Pour veiller à ce que cette obligation soit remplie, le « Programme de Protection à Plusieurs Niveaux[75] » (le « Programme ») exige de certains opérateurs d'intérêt vital qu'ils mettent en place des normes de protection plus strictes si les violations potentielles de données présentent des risques graves pour la sécurité nationale. Selon la réglementation actuelle, les plateformes gérant des données de santé sont classées au niveau 3 (sur une échelle de 1 à 5 des risques), ce qui les incite à prévenir les incidents de sécurité sur toute la chaîne de valeur, de la production au stockage de la donnée. Bien que la participation au Programme soit volontaire, la pression implicite pour que les entreprises choisissent d'y adhérer semble s'accroître drastiquement depuis 2019[76]. Par ailleurs, les services de santé numériques basés sur des applications, qui s'appuient sur de vastes volumes de données pour fonctionner, sont

72. D. Cuny, « Le digital et les GAFA, premier risque pour les assureurs », *La Tribune*, octobre 2018.
73. Mesures administratives sur la gestion des informations sur la santé de la population, mai 2014.
74. Loi sur la Cybersécurité de la République de Chine, 1er juin 2017. Pour une traduction anglaise, disponible sur : www.newamerica.org.
75. Le Programme de Protection à Plusieurs Niveaux (MLPS), cité par H. Feldshuh et L. Yau, dans « The Growing Intersection of Digital Health and Data Processing in China », *China Business Review*, avril 2021.
76. *Ibid.*

régis par deux types de réglementations : les réglementations sur les dispositifs médicaux et les réglementations générales sur les applications numériques – dont aucune n'est spécifique aux entreprises de télésanté.

Parallèlement à cette évolution de la réglementation, le régime chinois a lancé depuis le milieu des années 2000 un développement à marche forcée de la e-santé, notamment dans le souci d'améliorer la couverture de vastes zones rurales, avec une mise à contribution importante des grands acteurs technologiques en appui des efforts des pouvoirs publics. Lancé par le Comité central du Parti communiste et le Conseil des affaires de l'État en octobre 2016, le plan « Healthy China 2030 », décrit par l'Organisation mondiale de la santé (OMS) comme « la plus vaste réforme sanitaire que le monde ait connue », vise notamment à développer fortement le marché de la e-santé, l'interopérabilité des bases face à un système fortement régionalisé et à centraliser les données dites de *big data* sanitaires et médicales, notamment *via* quatre centres nationaux dédiés au *big data* en santé, sitúés à Fuzhou, Xiamen, Changzhou et Nanjiang (où le « National Health and Medicine Big Data Center » devrait stocker les informations génétiques de plusieurs millions de citoyens chinois pour travailler sur des applications de pointe).

La même année, la Chine a mis en place son « plan national pour la médecine de précision » avec des investissements de l'ordre de 60 milliards de yuans (9,3 milliards de dollars) dans des projets de séquençage génétique et la constitution de bases de données dédiées. En avril 2018[77], le Conseil des affaires de l'État a publié ses avis sur la promotion du développement du numérique en santé, exigeant des gouvernements régionaux qu'ils se coordonnent pour faciliter la mise en place de la plateforme nationale unifiée d'informations sur la santé. Celle-ci sera progressivement connectée à la plateforme nationale de partage des données comptant déjà, sur la base du système national d'assurance maladie, les données de 600 millions de personnes (la moitié de la population chinoise). Dans cette même lignée, la Commission Nationale de la Santé a promulgué dès juillet et septembre 2018 une série de mesures visant, respectivement, à ce que tous les types d'institutions médicales à tous les niveaux se connectent aux plateformes régionales d'informations sanitaires nationales, transmettent et sauvegardent les données générées

77. A. Zhou et C. Huang, « Cybersecurity and Data Protection in Chinese Healthcare Industry », *Global Law Office and Practical Law China*, novembre 2020.

pendant les services de soins de santé, et fournissent un accès de surveillance aux départements administratifs de la santé[78].

Un capitalisme de surveillance qui se projette à l'international

Depuis presque 15 ans, le régime chinois a mené une stratégie de « cohabitation structurée » entre les secteurs public et privé en matière de santé, avec le double objectif de favoriser l'accès aux soins de base pour tous et de mettre en place un système de santé au niveau des meilleurs standards internationaux[79]. Les grandes entreprises du numérique, dont les BATX, coopèrent ainsi de manière étroite avec le Parti, les autorités locales, les médecins et les assureurs privés. Ainsi, depuis 2019, 450 000 médecins chinois sont par exemple devenus utilisateurs de Trusted Doctor, une start-up créée en 2018 et rachetée par Tencent, qui met en relation médecins et patients mais aussi cliniques et hôpitaux. Pour sa part, Alibaba Health Information Technology a développé le plus grand service de pharmacie en ligne du pays. LinkDoc, qui quatre ans seulement après sa création avait atteint le milliard de dollars de valorisation[80], utilise les bases de données collectées par les acteurs de santé afin d'entraîner des algorithmes sur le traitement du cancer. Wuxi NextCODE (issue de l'acquisition d'une entreprise américaine par un acteur chinois) et Huawei se sont quant à eux positionnés dès 2016 pour développer l'infrastructure de *cloud* requise pour stocker les quantités massives de données créées par le système chinois et développer la puissance de calcul adaptée.

L'investissement massif d'Alibaba, Baidu, Tencent et Huawei dans les technologies de santé, notamment d'analyse et de *big data*, renforce toutefois le caractère dual du système de santé chinois entre un secteur public peu performant et des acteurs privés détenteurs de masses critiques de données pour se maintenir à la frontière technologique. Il pose également la question de l'hybridation croissante entre finalités sanitaires et policières des bases de données ainsi constituées, au regard notamment des campagnes de collecte de

78. Mesures sur le *big data* sanitaire et médical de 2018, articles 4 et 38 cités par A. Zhou et C. Huang, dans « Cybersecurity and Data Protection in Chinese Healthcare Industry », *op. cit.*
79. J. D. Séval, « Santé et numérique : "L'ambition de la Chine est de mettre en place une offre globale et intégrée" », *Le Monde*, 26 novembre 2019.
80. F. Le Deu, « 8 Reasons Why China Is the Most Exciting Healthcare Story in the World Right Now », 2018, in *Étude de l'évolution des acteurs privés et publics pour faire face aux nouveaux défis de la Chine du XXIe siècle*, L'Industrie pharmaceutique en Chine, Mines ParisTech, 2018-2019.

matériel génétique visant des minorités ethniques ou permettant d'accroître la surveillance d'opposants politiques[81].

L'appétit en données du système de santé chinois a en outre conduit à une augmentation exponentielle, depuis 2014, des investissements des entreprises et acteurs parapublics chinois dans le secteur des technologies de santé à l'étranger. Selon un rapport de 2019 de la Commission d'examen de l'économie et de la sécurité entre les États-Unis et la Chine[82] – créée par le Congrès américain pour évaluer les implications en matière de sécurité nationale des relations commerciales et économiques entre les États-Unis et la Chine – les investissements directs à l'étranger chinois dans les biotechnologies sont passés de 100 millions de dollars en 2014 à 1,5 milliard de dollars en 2015 puis à un peu plus de 3,5 milliards en 2017. Les investissements en capital-risque ont atteint 3,8 milliards de dollars la même année. Le rapport souligne que le secteur des technologies de santé et des biotechnologies est devenu en 2018 le principal champ d'investissement des entreprises chinoises aux États-Unis. Avec 23 entreprises liées à la Chine ayant accès aux bases de données de santé les plus sensibles (y compris de séquençage génétique), ses rédacteurs estiment que « les efforts de la Chine pour détenir des données de santé de citoyens américains, combinés avec la faible protection offerte par le droit américain, posent une question de sécurité nationale[83] ». Le RGPD est explicitement cité comme permettant une meilleure protection des citoyens européens face à d'éventuelles velléités prédatrices d'acteurs chinois.

Pour accompagner ce mouvement d'internationalisation des acteurs chinois, la diplomatie numérique est devenue centrale aux yeux des autorités, notamment en matière de normalisation, si bien que « la coopération vers une harmonisation des règles de protection des données est devenue l'un des principaux sujets de discussion des diplomates chinois dans les instances internationales »[84]. L'harmonisation des standards techniques en matière de sécurisation des produits et des services médicaux constitue en outre l'une des dimensions des Nouvelles routes de la soie[85].

81. S.-L. Wee, « China Is Collecting DNA from Tens of Millions of Men and Boys, Using U.S. Equipment », *The New York Times*, 30 juillet 2020.
82. Commission d'examen de l'économie et de la sécurité entre les États-Unis et la Chine, « China's Biotechnology Development », février 2019.
83. *Ibid*.
84. R. Creemers, « Comment la Chine projette de devenir une cyber-puissance », *Hérodote,* vol. 177-178, n° 2, 2020.
85. J. Seaman, « China and the New Geopolitics of Technical Standardization », *Notes de l'Ifri,* Ifri, janvier 2020.

Leçons de la crise sanitaire : les données de santé comme enjeu de puissance à l'ère du COVID-19

L'Union européenne

L'échec d'une application mobile de traçage paneuropéenne

Avec la crise sanitaire, l'enjeu des données de santé s'est cristallisé au niveau européen sur la difficulté pour les États membres de l'Union de coopérer au déploiement commun d'une application mobile de traçage, qui a révélé la sensibilité politique de la question des données de santé et les différences d'approches entre les États membres.

Le 21 avril 2020, le Comité européen de la protection des données (CEPD) a publié un ensemble de lignes directrices relatives à l'utilisation des données de géolocalisation et d'outils de suivi des contacts suivant une vision « européenne harmonisée [...] : utilisation basée sur le volontariat des personnes, recours à la technologie Bluetooth, utilisation de données à caractère personnel pseudonymisées[86] ». L'entente générale semblait prévaloir au sein des États membres sur le déploiement des applications de traçage et le recours au Bluetooth, sans utilisation de techniques de géolocalisation. L'UE a néanmoins rapidement buté sur les modalités d'hébergement des données des « cas contact » collectées par les applications.

Très vite, deux alternatives techniques d'hébergement des données se sont imposées : une approche dite « centralisée » visant à héberger les données sur un serveur tiers contrôlé par l'État et une approche dite « décentralisée » visant à stocker les données directement sur le téléphone de l'utilisateur de l'application. Avec l'annonce d'un partenariat inédit entre Apple et Google dans le développement de telles technologies de traçage, cette alternative

86. « Lignes directrices 4/2020 relatives à l'utilisation de données de localisation et d'outils de recherche de contacts dans le cadre de la pandémie de COVID-19 », Comité européen de la protection des données, 21 avril 2020, disponible sur : www.cnil.fr.

technique est devenue un enjeu stratégique pour les États membres. Le 10 avril 2020, Apple et Google faisaient part de la création commune d'une interface de programmation d'application (API), baptisée « Exposure Notification », offrant un socle technologique pour le développement d'applications de suivi de contacts dans tous les téléphones mobiles dont Apple et Google assurent l'exploitation.

Cette solution logicielle décentralisée, fondée sur la technologie Bluetooth, a accéléré l'implosion de l'approche paneuropéenne. La France et son principal opérateur en matière de souveraineté numérique, l'Institut national de recherche en informatique et en automatique (INRIA), défendaient une approche centralisée avec l'Allemagne[87], qui avait une avance de deux mois sur ces développements, et avait accepté une approche commune de tests avec son voisin d'outre-Rhin[88]. Le gouvernement allemand avait en effet tout d'abord privilégié une application centralisée développée par un consortium rassemblant des acteurs européens et nationaux[89], parmi lesquels figuraient notamment les prestigieux instituts Fraunhofer (recherche fondamentale) et Robert Koch (équivalent allemand de l'Institut Pasteur). Cette solution centralisée a néanmoins suscité une opposition très vive dans les milieux politiques, chez certaines ONG et de la part d'une partie du monde académique[90]. Le point central des critiques portait sur le stockage des données à partir d'un serveur central, ce qui aux yeux des détracteurs du projet constituait un risque majeur pour les libertés publiques, en créant les conditions d'une surveillance étatique de masse de la population allemande. Face à l'ampleur des critiques, qui risquait de compromettre l'adhésion de la population au principe même d'une application numérique pour lutter contre l'épidémie de COVID-19, le gouvernement allemand a abandonné le projet et suivi la majorité des États membres de l'Union en portant son choix sur la solution décentralisée développée conjointement par Apple et Google.

Le débat suscité en Allemagne par le développement d'une application de traçage centralisée révèle les fractures conceptuelles entre pays européens sur la question des données de santé : focalisé sur les risques de surveillance étatique, il a mis au second plan le débat lié aux risques du développement de telles applications sur la

87. À l'origine du protocole ROBERT (ROBust and privacy-presERving proximity Tracing), finalement retenu par la France.
88. Entretien entre Claude Casteluccia, directeur de recherches en sciences et technologies du numérique à l'INRIA et les auteurs, en avril 2021.
89. Il s'agit du programme baptisé « Pan European Privacy Protecting Proximity Tracing » (PEPP-PT).
90. 300 chercheurs de 26 pays ont notamment signé une lettre ouverte demandant aux gouvernements d'opter pour la solution décentralisée d'Apple et Google afin de « créer la confiance », disponible sur : www.theguardian.com.

base d'un socle technique élaboré par des acteurs privés non européens qui, en France, a été prégnant. Ainsi, en s'opposant à une architecture reposant sur des serveurs centraux, sous le contrôle de leurs autorités nationales, les adversaires du projet ont paradoxalement renforcé le monopole des géants américains sur le traçage numérique dans plus d'une vingtaine de pays dans le monde. Cette situation souligne en creux l'absence d'acteur numérique européen de taille critique en mesure de faire jeu égal avec les GAFAM, que l'UE ne peut compenser en s'en remettant à la seule taille de son marché ou à la portée de sa puissance normative.

Débat français, enjeu européen : la question du stockage des données de santé en France

La crise sanitaire a également mis en évidence les carences du modèle industriel européen en matière de stockage des données de santé. En France, ce débat s'est concentré sur la question du « Health Data Hub » (HDH) et de son prestataire d'hébergement agréé « hébergeur de données de santé », l'américain Microsoft Azure[91]. La pandémie a en effet considérablement accéléré la mise en place du projet HDH avec un arrêté du 21 avril 2020 qui lui a permis de traiter de nombreuses catégories de données « aux seules fins de faciliter l'utilisation des données de santé pour les besoins de la gestion de l'urgence sanitaire et de l'amélioration des connaissances sur le virus COVID-19 ». À ce titre, le HDH était en mesure de centraliser et de croiser des données issues d'applications mobiles de santé, de télémédecine, des données du SNDS, des laboratoires et de pharmacies.

Le choix d'un prestataire d'hébergement américain a suscité un débat public intense en France, notamment du fait que Microsoft, en tant qu'entreprise américaine, puisse être soumis à des législations américaines telles que le *Clarifying Lawful Overseas Use of Data Act* (*CLOUD Act*) : entrée en vigueur le 23 mars 2018, cette loi fédérale prévoit la possibilité pour les autorités américaines de réquisitionner l'accès aux données détenues par des prestataires de stockage *cloud* américains, y compris lorsque ces données sont physiquement stockées en dehors des États-Unis. Les inquiétudes ont été renforcées par l'arrêt du 16 juillet 2020 de la Cour de Justice de l'Union européenne, dit Schrems II[92], invalidant l'accord de transferts de données entre l'Union

91. Les propos recueillis par les auteurs, respectivement auprès d'un haut fonctionnaire du ministère de la santé et d'un avocat spécialisé en sciences de la vie, soulignent que Microsoft était la seule entité offrant la totalité des fonctionnalités exigées par le déploiement du HDH.
92. Arrêt de la Cour de Justice de l'Union européenne, du 16 juillet 2020 rendu dans l'affaire C311/18 Data Protection Commissioner/Maximillian Schrems et Facebook Ireland du 16 juillet 2020.

européenne et les États-Unis (le « Privacy Shield ») en jugeant que la surveillance exercée par les services de renseignements américains sur les données personnelles des citoyens européens était excessive, insuffisamment encadrée et sans réelle possibilité de recours. Elle en avait déduit que les transferts de données personnelles depuis l'Union européenne vers les États-Unis étaient contraires au RGPD et à la Charte des droits fondamentaux de l'Union européenne, sauf si des mesures supplémentaires venaient à être mises en place ou si les transferts étaient justifiés au regard de l'article 49 du RGPD, qui prévoit des dérogations dans des situations particulières.

En raison de la sensibilité et du volume des données ayant vocation à être hébergées au sein du HDH, pour lesquelles le niveau de protection technique mais aussi juridique le plus élevé doit être assuré, y compris en matière d'accès direct par les autorités de pays tiers, la CNIL a fait part de son souhait que son hébergement et les services liés à sa gestion puissent être réservés à des entités relevant exclusivement des juridictions de l'Union européenne[93]. Le Conseil d'État, saisi de la question par plusieurs associations défenseures des libertés publiques sur internet, a rendu une ordonnance le 13 octobre 2020[94] reconnaissant l'existence d'un risque de transfert de données issues du HDH vers les États-Unis en raison de la soumission de Microsoft au droit américain et a demandé que des garanties supplémentaires soient mises en place. Le ministère de la Santé s'est engagé à recourir à une solution technique permettant de ne pas exposer les données hébergées par le HDH à d'éventuelles demandes d'accès illégales au regard du RGPD dans un délai compris entre 12 et 18 mois. Prenant en considération ces éléments, la CNIL a pu autoriser la réalisation de quatre nouveaux projets pilotes au sein du HDH.

Accéléré durant la crise sanitaire, le HDH a donc mis en évidence les vulnérabilités de l'approche française et européenne des technologies de santé, notamment la difficulté de faire émerger une alternative aux acteurs du *cloud* américain pour assurer un niveau de service conforme aux attentes des pouvoirs publics, des professionnels et des patients. D'autre part, il a exposé les failles du modèle normatif européen fondé sur le RGPD face à la carence industrielle du continent et sa dépendance à l'égard d'acteurs soumis au *CLOUD Act* américain : en France, le récent contentieux autour de l'hébergement par Amazon Web Services des données collectées par

93. La CNIL, « La Plateforme des données de santé (Health Data Hub) », février 2021.
94. Ordonnance du 13 octobre 2020, *Association Le conseil national du logiciel libre et autres*, n° 444937.

Doctolib dans le cadre de la crise sanitaire a une fois encore soulevé la question de l'articulation entre les deux systèmes normatifs[95].

Toutefois, le projet de « cloud de confiance » européen Gaia-X[96], lancé le 4 juin 2020 par les ministres de l'Économie français et allemand, fait preuve d'une véritable ambition pour une meilleure sécurisation des données de santé avec l'objectif de fédérer les 26 espaces numériques de santé nationaux parties prenantes à l'échelle européenne[97]. En France, la Stratégie nationale pour le cloud, présentée le 17 mai 2021[98], propose une nouvelle ambition pour associer compétitivité et souveraineté, en prônant l'exploitation de technologies américaines, sous licence, par des entreprises européennes, afin d'éviter l'application du droit américain ; en renforçant le référentiel de confiance « SecNumCloud » de l'ANSSI (que détiennent par exemple OVH Cloud et 3 D Outscale) ; et en investissant 107 millions d'euros dans le cadre du Plan France Relance pour soutenir des projets français innovants dans le *cloud*. Suite à ces annonces, Capgemini et Orange ont lancé « Bleu », une solution de cloud de confiance fondée sur un accord de licence avec Microsoft, pour permettre de concilier la double exigence de haut niveau de service et de sécurité des données hébergées par le HDH. Microsoft s'est pour sa part engagé à ce que les données des entreprises et services publics européens restent en Europe[99].

Vers une plus grande coopération européenne en matière de données de santé ?

Dans sa communication de février 2020 sur une stratégie européenne pour les données, la Commission européenne s'était engagée à soutenir la création d'un espace européen des données de santé, l'un des neuf « espaces de données » définis par la stratégie. La pandémie

95. D'après le juge des référés du Conseil d'Etat, les données personnelles hébergées par Doctolib dans le cadre de la vaccination contre le COVID-19 sont suffisamment protégées, bien qu'Amazon ait été choisie pour les héberger. Il justifie notamment sa décision par la mise en place d'une procédure de chiffrement reposant sur un tiers de confiance situé en France et le fait que les deux sociétés ont conclu un avenant permettant de lutter contre la mainmise des autorités américaines sur les informations des patients français (ordonnance du 12 mars 2021, *Société Interhop et autres*, n° 450163).

96. Ce projet, reposant sur la collaboration d'une vingtaine d'entreprises françaises et allemandes, notamment OVH Cloud et T-Systems, a pour ambition de construire une infrastructure de données fiable et sécurisée pour l'Europe.

97. Voir par exemple, Gaia-X : Le hub français invite les volontaires à rejoindre ses rangs – ZDNet, disponible sur : www.zdnet.fr.

98. « Stratégie nationale pour le cloud », dossier de presse, disponible sur : www.numerique.gouv.fr.

99. L. Barthélémy, « Où sont les données ? Microsoft tente de répondre aux inquiétudes européennes », AFP, 27 mai 2021.

de COVID-19 a révélé le besoin criant de coopération et d'échanges de données de qualité dans l'Union européenne pour concevoir la réponse politique appropriée face à ce type de menace globale et immédiate[100].

Face à l'hétérogénéité et au manque d'interopérabilité des bases de données de santé en Europe, ayant limité la possibilité d'une action véritablement coordonnée à l'échelle européenne, le projet de création de l'espace européen des données de santé a été accéléré par la crise. Le 11 novembre 2020, la Commission européenne déclarait vouloir davantage de coopération dans les soins de santé, la recherche et l'élaboration de politiques de santé en Europe[101]. Un premier pilier consisterait en un système solide de gouvernance des données et de règles de partage : pour la Commission, la crise du COVID-19 a légitimé l'approche européenne des données de santé, qui doivent être collectées et analysées de manière responsable pour lutter efficacement contre l'épidémie, au risque de porter atteinte à la vie privée et à la confiance du public. À ce dernier titre, la Commission semble avoir pris très au sérieux le débat public suscité par les applications de traçage et cherche à construire un climat de confiance quitte à ralentir la mise en place de l'espace européen de santé.

Ensuite, la Commission européenne estime que pour exploiter pleinement le potentiel des échanges de données relatives à la santé, il est essentiel de garantir leur qualité et de veiller à ce que les diverses sources de données de santé (par exemple les dossiers médicaux électroniques, les différents registres, les divers outils informatiques et numériques) puissent « dialoguer » les unes avec les autres. Ce dialogue nécessite une interopérabilité technique et sémantique à bâtir entre les différents systèmes informatiques et des infrastructures nouvelles permettant d'établir des liens entre les bases, le « eHealth Digital Service Infrastructure » (eHDSI) pour la portabilité des dossiers médicaux des patients, déployé à partir de 2025 dans les pays de l'Union, et le « Cross Border Health Information Services » (CBeHIS) pour les échanges de données entre institutions publiques[102]. Enfin, la Commission européenne décline les critères de son protocole relatif aux données concernant la santé, qui doivent être « faciles à trouver, accessibles, interopérables et réutilisables » (FAIR).

100. N. Iacob et F. Simonelli, « Towards a European Health Data Ecosystem », *European Journal of Risk Regulation*, vol. 11, n° 4, 2020.
101. Commission européenne, « Construire une Union européenne de la santé : une Europe mieux préparée aux crises et plus forte dans sa riposte », 11 novembre 2020, disponible sur : https://ec.europa.eu.
102. Communication de la Commission européenne, L'espace européen des données de santé, disponible sur : https://ec.europa.eu.

En sus de la constitution progressive d'un espace européen de la donnée de santé, le prochain défi des Européens résidera dans la mise en place effective d'un passeport sanitaire européen, validé le 29 avril 2021 et adopté le 9 juin 2021 par le Parlement européen sous la forme d'un « certificat COVID numérique de l'UE » permettant, à l'échelle de l'Union, de faciliter les trajets en Europe. Les documents, en versions papier ou numérique, attesteront qu'une personne a été vaccinée contre le coronavirus ou qu'elle a reçu récemment un test négatif ou encore qu'elle s'est remise de l'infection[103]. En matière d'infrastructure technique, un « service passerelle » permettant de vérifier les certificats « d'une manière sûre et en assurant le respect de la vie privée », a été mis en place par la Commission au niveau de l'UE. Au 21 juin 2021, 16 pays de l'Union avaient décidé de se connecter au service passerelle[104].

Les États-Unis

Les dysfonctionnements de la gestion technologique de la crise sanitaire

La crise sanitaire a révélé les dysfonctionnements du système américain de collecte et de traitement de données en matière de santé publique. Les Centres américains de contrôle et de prévention des maladies (CDC) utilisaient traditionnellement un système de surveillance sur des flux de données allant d'un niveau local vers le niveau fédéral pour suivre la propagation des épidémies. Efficace dans le cas d'épidémies de moindre ampleur comme celle de la salmonellose qui touche épisodiquement les États américains, cette méthode a rencontré de nombreuses difficultés avec des remontées lacunaires et une analyse trop lente face au rythme de propagation du virus[105].

Face à ces dysfonctionnements, l'administration Trump a décidé d'écarter les CDC du traitement des données sur les cas et les hospitalisations COVID, au profit d'un système *ad hoc* confié directement au ministère de la Santé, le Department of Health and Human Services. L'infrastructure de ce système était opérée par deux entreprises proches du Président, TeleTracking et Palantir, sur la base

103. « Le certificat européen COVID-19 doit faciliter la liberté de circulation sans discrimination », Communication du Parlement européen, avril 2021.
104. Les informations et le décompte des pays connectés sont disponibles sur : https://ec.europa.eu.
105. Au-delà des défaillances techniques, les causes de ces lenteurs seraient, en partie, à rechercher dans la rétention d'information de la part de responsables de l'Administration de l'ancien Président Trump à l'approche des élections de fin d'année. Disponible sur : https://thehill.com.

de plus de 200 jeux de données issues des trois quarts des 8 000 hôpitaux du pays. Le directeur de l'*American Public Health Association* à Washington D.C., Georges Benjamin, affirme qu'au lieu de rationnaliser la collecte des données de santé pendant la crise, ce changement a rendu cette dernière encore plus confuse[106] puisque de nombreux administrateurs d'hôpitaux ne savaient plus à quelles agences s'adresser en plus des difficultés liées aux partages obsolètes de données (par fax par exemple). Des épidémiologistes de la santé publique américaine (*Council of State and Territorial Epidemiologists*) avaient déjà fait état, dans un rapport publié le 25 septembre 2019[107], de la lenteur et du cloisonnement du système des données de santé aux États-Unis, avant tout manuel et basé sur le papier : en effet, si au début des années 2010, le gouvernement fédéral a dépensé plusieurs milliards de dollars pour encourager les médecins libéraux à remplacer les télécopieurs par des dossiers électroniques[108], le *HITECH Act* ne prévoyait pas de financement similaire pour les hôpitaux et cliniques publics. L'absence d'outil permettant la collecte et le partage rapide des données de santé avec les autorités de santé publique a ainsi rendu difficile le pilotage de la crise sanitaire et les débuts de la campagne de vaccination.

En matière d'application de traçage, le gouvernement fédéral est resté relativement discret sur le développement d'une application nationale et s'est largement reposé sur la solution développée par Apple et Google. Certains États ont profité de cette absence et n'ont pas attendu la publication de l'interface de programmation d'application des deux GAFAM mi-mai pour s'engager en ordre dispersé dans des solutions locales. Le Dakota du Nord, le Dakota du Sud et l'Utah ont ainsi développé des applications de traçage avec l'aide d'entreprises locales, centralisées, fondées sur le GPS et le Bluetooth et utilisées en appui d'un traçage manuel[109]. De manière générale, la réticence exprimée dans les sondages par les Américains envers les applications de traçage (environ 3 Américains sur 5 ont déclaré qu'ils ne pourraient ou ne souhaiteraient pas s'en servir[110]) ainsi que le manque d'impulsion au niveau fédéral ont conduit les États à se tourner vers des dispositifs humains parfois massifs (la

106. « The Trump Administration Must Stop Sidelining the CDC », *Nature,* 28 juillet 2020.
107. Fondation Beaumont, « White Paper: Driving Public Health in the Fast Lane », CSTE, 25 septembre 2019.
108. S. Kliff et M. Sanger-Katz, « Bottleneck for U.S. Coronavirus Response: The Fax Machine », *The New York Times,* 13 juillet 2020.
109. E. Setzer, « Contact-Tracing Apps in the United States », Lawfare Institute, Brookings Institution, 6 mai 2020.
110. « Most Americans Are Not Willing or Able to Use an App Tracking Coronavirus Infections. That's a Problem for Big Tech's Plan to Slow the Pandemic », *The Wahington Post*, 29 avril 2020.

Californie s'est appuyée sur un corps de 10 000 volontaires), plus efficace et politiquement moins coûteux.

Avec Google et Apple sur le terrain du traçage, et Palantir dans la constitution de bases de données de santé, les géants du numérique américains ont réussi à imposer l'image de réactivité de leurs services en ligne face à des systèmes publics jugés trop lents ou défaillants. Ils ont ainsi saisi l'occasion de la crise sanitaire pour « *changer la façon dont ils sont perçus par le grand public et les responsables politiques* » d'après William Kovacic, ancien de l'autorité américaine de la concurrence (Federal Trade Commission) et professeur de droit à l'université de Georgetown, à Washington[111]. Toutefois, ces nouveaux partenariats public-privé ont suscité de nombreuses inquiétudes devant la volonté des entreprises du numérique de se présenter comme les partenaires par défaut des gouvernements désireux de numériser l'accès aux services de santé pour leurs citoyens, et ce à leurs propres conditions[112].

Une convergence vers le modèle européen sous la présidence Biden ?

Du fait de son engagement fort et ancien dans la recherche contre le cancer *via* d'abord la « Cancer Moonshot Initiative » puis la « Biden Cancer Initiative » (cf. supra), le nouveau Président des États-Unis a, depuis longtemps, une vision très claire de la problématique des données de santé dans son pays. Dans une tribune importante en date du 19 mars 2018[113], Joe Biden développait sa vision des blocages et des réformes à mener dans ce domaine. Il y affirmait notamment que dans le sillage des mesures adoptées sous les mandats de Barack Obama, l'industrie médicale n'avait pas saisi l'opportunité de favoriser l'interopérabilité des bases et la maîtrise des patients sur leurs données. Il en déduisait la nécessité de l'intervention de la puissance publique dans le domaine des données de santé. Il suggérait ainsi l'idée d'imposer aux fournisseurs de soins de proposer aux patients leur dossier médical à jour dans un délai de 24 heures sous peine de sanction pénale, de constituer un portail unique sous la houlette du Center for Medicare and Medicaid Innovation permettant de stocker, de lire et d'échanger les données des assurés de façon unifiée et de renforcer le partenariat « Sync for Science (S4S) » entre le ministère de la Santé et les agrégateurs privés de bases de données

111. A. Piquard, « La crise du coronavirus va-t-elle améliorer l'image des GAFA ? », *Le Monde,* 10 avril 2020.
112. F. Guerrini, « The Dark Side of the Apple and Google Collaboration Against COVID-19 », Forbes, 13 avril 2020.
113. J. Biden, « Joe Biden: To Save and Improve Lives Using Data, Details Matter », Forbes, 19 mars 2018.

de santé (« electronic health record (EHR) vendors[114] ») afin d'améliorer l'accès des centres de recherche à ces ressources.

Dès le 21 janvier 2021, soit son premier jour en tant que président en exercice, Joe Biden a souhaité tirer les conséquences de la gestion de la pandémie par l'administration Trump en signant un décret présidentiel visant à garantir une réponse « fondée sur les données » au COVID-19 et aux futures menaces de santé publique à haut risque[115]. Entre autres mesures, le décret présidentiel met l'accent sur l'ouverture des données détenues par le gouvernement fédéral dans des formats lisibles et exploitables. Le décret désigne également le ministre de la Santé, en coordination avec les agences concernées, comme le responsable de l'examen de l'efficacité, de l'interopérabilité et de la connectivité des systèmes de données de santé publique fédéraux, mettant fin à l'enchevêtrement des compétences cultivées par l'administration Trump. Pour finir, le décret amorce un plan pour faire progresser l'innovation dans les données et les analyses de santé publique aux États-Unis, qui fera l'objet d'un financement massif dans le cadre du plan de relance[116]. Le président Biden souhaite en effet injecter 8,7 milliards de dollars dans les CDC afin de moderniser la collecte de données de santé publique à l'échelle nationale, en plus de soutenir l'amélioration des capacités de base en matière de santé publique dans les États américains.

Enfin, le mandat de Joe Biden sera vraisemblablement marqué par un débat renouvelé autour d'une grande loi fédérale de protection des données personnelles, sur le modèle européen. D'une part, la Vice-Présidente Kamala Harris s'est intéressée à la question lors de son mandat de sénatrice pour la Californie. Elle a notamment appuyé une proposition de loi déposée par Elizabeth Warren amendant l'HIPAA afin de mieux protéger les données de santé dans le cadre de la crise du COVID-19. Elle est en cela soutenue par la gauche du parti démocrate qui est également à l'origine de plusieurs propositions dans ce domaine. D'autre part, en 2020, deux propositions de loi ont été déposées par les Républicains au Sénat[117] et par les Démocrates à la Chambre[118] visant à améliorer la protection des données, avec un large consensus bipartisan sur le fond, notamment pour faire des données de santé une catégorie de données sensible à protéger de

114. Comme Allscripts, Cerner, eClinicalWorks, Epic…
115. *Executive Order on Ensuring a Data-Driven Response to COVID-19 and Future High-Consequence Public Health Threats*, Décret publié en ligne par la Maison-Blanche, 21 janvier 2021.
116. Plus d'informations disponibles sur : www.whitehouse.gov.
117. Les propositions de lois des Républicains en la matière sont disponibles sur : www.congress.gov.
118. Les propositions de lois des Démocrates en la matière sont disponibles sur : www.congress.gov.

façon spécifique, à l'exception d'une divergence notable sur la place à donner aux réglementations locales des États. Ce débat interne a des implications diplomatiques pour les relations transatlantiques : avec l'invalidation du « Privacy Shield » par la CJUE (cf. *supra*), un glissement même modéré du droit américain vers le droit européen du RGPD, et notamment la modification de certaines de ses dispositions les plus intrusives[119], pourrait permettre à la nouvelle administration américaine de donner des gages à ses alliés européens échaudés par quatre années de présidence Trump.

La Chine

Une gestion de la crise sanitaire renforçant le poids des BATX dans le capitalisme de surveillance chinois

L'épidémie de COVID-19 a permis de renforcer le contrôle des Centres chinois de contrôle et de prévention des maladies (CCDC), institutions gouvernementales responsables de la gestion technique du contrôle des maladies et de la santé publique, sur les acteurs de la santé. En connectant le système de surveillance et d'alerte précoce du CCDC au système de gestion électronique des dossiers médicaux des hôpitaux, les centres pouvaient surveiller en temps réel les ordinateurs des médecins en première ligne pour les inciter à vérifier l'exactitude des informations données par les patients[120]. Le système de surveillance des CCDC a ainsi permis de réduire le temps moyen nécessaire aux médecins pour déclarer un cas de COVID-19, qui est passé de 5 à 8 minutes à 40 secondes, et le temps nécessaire à la déclaration en ligne via le système de déclaration des maladies infectieuses des CDC, qui est passé de 2 à 3 minutes à quelques secondes[121].

Par ailleurs, le régime chinois a eu recours de manière extensive à des bases de données publiques et privées dans la lutte contre le COVID-19. La Chine a surtout fait le choix de suivre les déplacements de populations et de leur appliquer un régime strict en cas de contamination. Wu Zunyou, expert officiel en épidémiologie au centre national de la prévention et du contrôle d'épidémies[122], affirme avoir eu accès et recours à de nombreuses données pour suivre les

119. L'article 702 du *Foreign Intelligence Surveillance Act* et le décret présidentiel n° 12333.
120. J. Wu, J. Wang *et. al.*, « Application of Big Data Technology for COVID-19 Prevention and Control in China: Lessons and Recommendations », *Journal of Medical Internet Research,* vol. 22, n° 10, 2020.
121. J. Y. Shuangshuo, Technology, 28 mars 2020, disponible sur : www.sohu.com.
122. D. André, « La Chine utilise les données personnelles pour lutter contre le coronavirus », interview, France Inter, 31 janvier 2020.

déplacements – avec l'aide de China Railway – compagnie ferroviaire chinoise, et localiser les personnes malades : « Avec le *big data*, on sait où les cinq millions de personnes parties de Wuhan et de la province de Hubei sont allées. On peut capturer les données de localisation précises et tracer ces personnes. ».

Les géants chinois du numérique ont été mobilisés par le régime afin de traiter les quantités d'information disponibles. Un système numérique de prévention des épidémies, développé conjointement par Alipay, Dingding et Alibaba Cloud[123], a par exemple permis d'analyser les données médicales de cas diagnostiqués dans chaque hôpital du pays et d'identifier les personnes ayant acheté des médicaments contre la fièvre en pharmacie au cours du mois écoulé. Grâce aux données recueillies, ces entreprises ont été en mesure de modéliser la propagation de l'épidémie en fonction de zones géographiques et d'identifier les principaux foyers de transmission, comme les centres commerciaux de Baodi et de Tulong, ou encore l'hôpital de Tianjin. D'autre part, les BATX ont participé à la mise en place des différents systèmes de QR code sanitaire déployés par la plupart des provinces et des villes chinoises, opérés notamment par Alibaba (Alipay), Tencent (Wechat) et Baidu. Dès la fin février 2020, Alibaba hébergeait les applications de plus de 200 grandes collectivités contre plus de 300 pour Tencent[124]. La plupart de ces applications imposaient la fourniture des données biométriques, voire un enregistrement par reconnaissance faciale (par exemple à Pékin). Elles avaient un fonctionnement similaire, attribuant un code couleur (vert, jaune ou rouge), lisible via un QR code individuel, en fonction notamment de l'historique des mouvements de la personne et de son bilan de santé. Ce code déterminait le cas échéant la mise en quarantaine pendant 14 ou 7 jours, ainsi que l'accès à certains secteurs ou services.

Si le gouvernement central a dans un premier temps laissé se développer ces initiatives locales, l'incompatibilité de certaines applications entre elles a soulevé des difficultés lors de l'ouverture progressive du pays. Plutôt que de développer une application nationale, les autorités ont alors lancé au printemps 2020 la constitution d'une base de données à l'échelle du pays permettant le partage des données recueillies par les applications locales, favorisant ainsi leur interopérabilité et la reprise en main du pouvoir central. Si ce système de traçage est au cœur du « narratif » chinois de gestion

123. Wu, Wang *et. al.*, « Application of Big Data Technology for COVID-19 Prevention and Control in China », *op. cit.*
124. N. Gan et D. Culver, « China Is Fighting the Coronavirus with a Digital QR Code. Here's How It Works », *CNN*, 16 avril 2020, disponible sur : https://edition.cnn.com.

crise, son efficacité a été largement nuancée par les analystes[125] et la réalité semble plus proche des systèmes de traçage européens fondés sur une grande quantité d'agents humains permettant de contacter les cas suspects par téléphone et d'obtenir des données robustes, compensant l'absence de granularité des QR codes de santé.

Vers une mise au pas des BATX par le régime ?

Face à l'omniprésence des BATX durant la gestion de la crise sanitaire, la relation du pouvoir aux grandes entreprises technologiques semble avoir évolué dans le sens d'une restriction des marges de manœuvre de ces dernières. Ce regain de tensions intervenu dans le sillage de la crise du COVID-19 est notamment visible dans le durcissement, fin 2020, des réglementations antitrust, qui a obligé Alibaba à payer une amende de 2,3 milliards d'euros en novembre 2020, et de celle de la réglementation boursière qui a conduit au blocage de l'entrée en bourse à 34 milliards de dollars d'Ant Group, la filiale de paiements en ligne d'Alibaba[126]. Dans ce même mouvement, le gendarme numérique chinois, l'Administration du cyberespace de Chine (CAC) a récemment rappelé à l'ordre 33 applications mobiles, dont certaines détenues par Alibaba, Baidu et Tencent pour avoir collecté plus de données sur les utilisateurs qu'elle ne le juge nécessaire pour offrir leur service. Le 10 mai 2021, 84 applications supplémentaires étaient ciblées[127].

La volonté du gouvernement chinois de réduire la latitude des grandes entreprises du numérique s'est également traduite en octobre 2020 par le dépôt d'un nouveau projet de loi sur la protection des informations personnelles (PIPL), dont une nouvelle version est parue fin avril 2021[128], qui a pour ambition d'unifier la réglementation des données de façon plus stricte que la loi sur la cybersécurité de 2017[129]. En partie inspirée du RGPD, dont elle partage certains des effets extraterritoriaux (sur la protection des données de citoyens chinois détenues à l'étranger), le PIPL classifie les données personnelles de santé comme une donnée sensible (article 29 de la dernière version du projet) et devrait théoriquement assurer aux patients une protection supérieure à celle du droit

125. Entretien avec une chercheuse politologue et sinologue, et les auteurs, en avril 2021.
126. « What Is Behind China's Crackdown on Its Tech Giants: QuickTake », Bloomberg Business, 13 novembre 2020, disponible sur : www.bloomberg.com.
127. S. Sharwood, « Beijing Twirls Ban-Hammer at 84 More Apps It Says Need to Stop Slurping Excess Data », *The Register*, 12 mai 2021, disponible sur : www.theregister.com.
128. G. Webster, « Translation: Personal Information Protection Law of the People's Republic of China », DigiChina, University of Stanford, disponible sur : https://digichina.stanford.edu.
129. X. Fu-Bourgne, « Un nouveau projet de loi publié en Chine », *Expertise*, n° 467, avril 2021.

américain à l'égard des acteurs privés. Face à l'importance prise par les grandes entreprises du numérique dans la gestion de la crise sanitaire et à une certaine forme de demande sociale pour une meilleure prise en compte du consentement des citoyens dans la collecte de leurs données, le gouvernement chinois semble bien décidé à mieux encadrer l'économie numérique du pays et à reprendre la main sur ses principaux acteurs[130].

130. « La Chine met au pas ses géants du numérique », AFP, 22 mars 2021.

Conclusion : après la crise sanitaire, faire de l'Europe une puissance du numérique en santé

La géopolitique des données de santé, dont les lignes de force ont été largement recomposées avec la crise du COVID-19, complexifie le constat opéré en février 2018 par la Revue stratégique de cyberdéfense d'une « opposition fondamentale » entre la conception occidentale du cyberespace et les conceptions de la Russie et de la Chine. Cette étude a essayé de montrer qu'au sein même du monde occidental, les modèles de gouvernance et les corpus de valeurs qui entourent la gestion des données de santé, du fait de leur spécificité dans la typologie du monde numérique, donnent lieu à des frictions et des rapports de force.

Réciproquement, des influences mutuelles et des rapprochements ne sont pas à exclure dans les prochaines années entre les différents modèles analysés par cette étude. En ce sens, si la recomposition, d'une part, du modèle américain et, d'autre part, du modèle chinois de protection des données semble augurer une convergence vers un modèle à l'européenne, la différence majeure demeure le fait que les États-Unis et la Chine disposent d'ores et déjà d'un complexe « sanitaro-numérique » puissant, sorti renforcé par la crise sanitaire même si celle-ci en a également illustré les failles et les limites.

Loin de consacrer le triomphe de la puissance normative européenne, cette recomposition pourrait alors avoir pour effet, sans action résolue des Européens en matière industrielle et scientifique, de consacrer le duopole sino-américain sur les applications du numérique en santé, gisement de valeur et de pouvoir majeur du XXIe siècle. Comme le souligne Stéphane Grumbach, directeur de recherche à l'INRIA : « Nous allons inexorablement vers un modèle d'exploitation de plus en plus massive des données de santé car, dans un monde idéal, la promesse est extraordinaire pour le bien commun, le bien des personnes et des populations. Mais il se pose désormais des questions de souveraineté qui dépassent les seuls intérêts privé-public. Dépendre

de plateformes étrangères dans des domaines comme la santé ou l'éducation, c'est inhiber la capacité même de gouverner[131]. »

Avec les applications de traçage et le « pass sanitaire », un débat public européen important s'est engagé durant la crise sanitaire sur la question des données de santé, dont l'issue influencera l'avenir de la coopération européenne en matière de santé et la capacité des États membres à faire émerger un modèle industriel conforme à leurs intérêts, à leurs valeurs et à leur conception du soin. Si l'influence du modèle normatif européen actuel est réelle, il est concurrencé par le modèle chinois de surveillance et la vision capitalistique des géants du numérique, diffusant un narratif d'appropriation totale des données par les patients et de leur libre usage, ouvrant *de facto* la porte à leur marchandisation. Dans une interview publiée le 11 mai 2021, Éric Schmidt, l'ancien PDG de Google pressenti pour prendre la tête d'un organisme de recherche médicale en juin 2021, préconisait « que les règles de confidentialité [en matière de données de santé] soient modifiées pour que le consentement soit donné par défaut pour la recherche médicale[132] ».

Après les efforts de la présidence allemande en matière de souveraineté numérique fin 2020, la « Conférence sur l'avenir de l'Europe » lancée en mai 2021 et censée aboutir au printemps 2022 sous la présidence française du conseil de l'Union européenne, pourrait devenir le lieu d'une réflexion collective sur la question des données de santé, qui cristallise les enjeux démocratiques, économiques et politiques au cœur de la souveraineté numérique du continent à l'ère du COVID-19.

131. L. Belot, « Les données de santé, un trésor mondialement convoité », *Le Monde*, 2 mars 2020.
132. « Former Google CEO Schmidt Opens Up on Health Data Privacy and the Future of Tech in Medicine », Stat+, 12 mai 2021, disponible sur : www.statnews.com.

27 rue de la Procession 75740 Paris cedex 15 – France

Ifri.org