

VERS UN NOUVEAU PACTE SOCIAL

Engager citoyens, élus,
décideurs, chercheurs et
industries dans une **culture
commune** des données de
santé



agence
de programmes
de recherche
en **santé**

SOMMAIRE

Préambule	p.05
Recommandations	p.08
Introduction	p.10

I. État des lieux

Entre promesse et contrainte : la révolution des données en mouvement	p.12
1. Un potentiel largement reconnu	p.12
2. Un écosystème français et européen en mouvement	p.14
3. Des freins structurels	p.18

II. Levier(s)

Un pacte social renouvelé grâce aux données de santé	p.24
1. Créer un élan concertatif et démocratique autour des données	p.24
2. Informer, acculturer, mobiliser, convaincre	p.27
3. Promouvoir le « citoyen-chercheur »	p.29

III. Impact

Bâtir une souveraineté scientifique, économique et démocratique	p.30
1. Restaurer la confiance démocratique dans la science	p.30
2. Assurer la soutenabilité économique	p.31
3. Piloter notre système de santé	p.35
Conclusion	p.38
Le rôle de l'APRS	p.39
Remerciements	p.41

PRÉAMBULE

Franck Mouthon & Vanina Paoli Gagin

« Données de santé : un actif sous-valorisé

Les données de santé constituent bien plus qu'un simple outil, dès lors qu'en elles réside la promesse d'une profonde transformation de notre système de soin, de son pilotage stratégique, de sa mue du curatif au préventif.

Plusieurs grands défis nous exhortent ainsi à aller plus loin pour libérer le plein potentiel des données. Ces défis sont sanitaires (augmentation de la prévalence des maladies chroniques, vieillissement), économiques (soutenabilité du système), politiques (inégalités territoriales d'accès aux soins, désinformation, défiance envers la science, retrait des grandes puissances des organismes multilatéraux de la santé), réglementaires (espace européen des données de santé).

Nous vivons aujourd'hui un moment unique pour faire converger et donner leur pleine puissance à toutes les forces actives de cette révolution en France — DGOS, Délégation au numérique en santé (DNS), Agence de l'Innovation en Santé, Agence de la biomédecine, etc. L'interconnexion données de soin / données de recherche est déterminante, et la France dispose d'atouts majeurs pour structurer et intégrer leur exploitation, que ce soit à travers la base médico-administrative de l'Assurance maladie, les grandes cohortes de recherche, les entrepôts hospitaliers, le système national des données de santé (SNDS), ou encore la plateforme des données de santé (PDS).

Au-delà du traitement et de la prise en charge, les données peuvent nous permettre de mieux analyser et comprendre les déterminants de santé, anticiper les risques, améliorer la prévention, garantir l'équité d'accès aux soins, dégager des priorités de recherche, renforcer l'attractivité de la France, accélérer l'innovation au service du bien commun. Pour cela, l'engagement de toute la société est nécessaire — État, soignants, chercheurs, industries, citoyens.

Promouvoir un nouveau pacte social

L'un des verrous trop souvent minorés est celui du lien, pourtant fondamental, entre données de santé et société. Or, ces données n'existeraient tout simplement pas sans les citoyens qui les produisent. Dans la course à leur structuration et leur exploitation technologiques, nous avons manqué l'occasion de définir un cap politique clair permettant d'embarquer la collectivité dans la compréhension des enjeux autour des données qu'elle produit. Faute de ce cadre partagé, l'appréhension du sujet a rapidement été réduite à la seule question des droits des usagers — un terrain restrictif qui nourrit paradoxalement, mais inévitablement, l'interprétation, l'incompréhension, voire la défiance.

Seule une gouvernance partagée et inclusive garantira l'appropriation collective et la confiance nécessaires à une valorisation efficace et vertueuse des données de santé. Une concertation nationale permettra de définir collectivement les enjeux, les valeurs et les priorités liés aux usages des données de santé. Cette gouvernance pourrait se fonder, de manière plus structurelle, sur un véritable programme d'éducation scolaire et de formation professionnelle, afin de faire émerger une authentique culture de la donnée. Des plateformes déjà bien connues des Français, telles que Mon Espace Santé ou le Dossier Médical Partagé, pourraient intégrer des espaces ergonomiques dédiés à l'information sur les usages des données, offrant des contenus clairs, accessibles, illustrés par des retours d'impact concrets et incarnés.

À l'heure des IA décentralisées et de l'intégration des données cliniques dans la vie réelle, la donnée de santé est le vecteur idoine pour lier recherche, santé et société, et enfin placer le citoyen au cœur de la production de la connaissance.

Renforcer la souveraineté démocratique, économique, sanitaire

Réengagées au cœur d'une appropriation collective, les données de santé peuvent enfin devenir un puissant moteur de souveraineté. Souveraineté démocratique d'abord, car à l'heure de la « société de la désinformation », il est vital d'acquérir collectivement une *immunité cognitive* et des facultés de discernement pour résister aux narratifs hostiles, renouer avec la valeur de la science et reconnaître la puissance de la connaissance. Souveraineté économique ensuite, car les données de santé peuvent accompagner le virage préventif de notre modèle sanitaire, optimiser la prise en charge, nourrir une filière de recherche attractive et compétitive : autant de cas d'usage pour atténuer l'inédite tension qui traverse notre système social. Souveraineté sanitaire enfin, puisque les données de santé ouvrent la voie à un pilotage du soin plus fin, qui réduira les inégalités d'accès, anticipera les besoins, et individualisera les parcours.

En bâtissant un nouveau pacte social fondé sur la confiance et sur la connaissance, nous pouvons faire de la donnée de santé un projet collectif sensé, porteur de certitude en temps de doutes. ➤➤

RECOMMANDATIONS

1 Intégrer pleinement les citoyens aux enjeux de gouvernance des données

Institutionnaliser la démocratie sanitaire des données, en associant patients, citoyens, industries, territoires, décideurs et élus à la définition des usages et à l'évolution des cadres. Leur participation devra également porter sur la priorisation des programmes de soin et de recherche liés à la donnée, sur l'évaluation continue de la stratégie nationale des données et son adossement aux débats budgétaires via la représentation nationale.

2 Lancer un cycle de concertation nationale autour des données de santé

Réunir citoyens, élus, chercheurs, employeurs et administrations pour définir collectivement les enjeux, les valeurs et les priorités liés aux usages des données de santé afin de construire une culture partagée de la confiance et de la souveraineté numérique. Une première consultation nationale pourrait porter sur le don de données par défaut, à l'instar du don d'organes.

3 Promouvoir une culture de la donnée responsable et solidaire

Élaborer une charte nationale de l'implication citoyenne fondée sur cinq principes : compréhension, équité, solidarité, sobriété, responsabilité, afin de faire de la donnée un commun au service de la santé publique.

4 Créer une culture commune de la donnée

Lancer un plan national de formation initiale et continue à la donnée et à l'intelligence artificielle (IA) des professionnels de santé et décideurs publics, et déployer des campagnes d'information et de sensibilisation auprès du grand public.

5 Faire de la donnée un bien scientifique collectif et souverain

En lien avec l'Espace européen des données de santé, adopter une architecture nationale des données au service de la recherche scientifique, la décision publique et la souveraineté économique et technologique nationale, lisibles par tous.

6 Définir de grands objectifs nationaux de santé fondés sur la donnée

Établir des objectifs clairs, mesurables et collectivement partagés, adossés à une stratégie interministérielle de planification, de prospective et d'anticipation, afin d'aligner les priorités de recherche, les moyens et les évaluations publiques à partir des données.

7 Faire de la donnée un levier stratégique national

Structurer une stratégie nationale des données de santé fondée sur l'interopérabilité, la qualité, la sécurité, la traçabilité et la souveraineté, en garantissant un financement pérenne des infrastructures, une valorisation de la science des données et une gouvernance partagée entre acteurs publics.

8 Redéfinir notre modèle sanitaire

Faire évoluer notre système de santé d'un modèle centré sur la consommation de soins vers une logique d'anticipation, de prévention, d'évaluation continue, et de soutenabilité, en mobilisant les données pour mieux comprendre, prévoir et orienter les besoins de santé.

INTRODUCTION

L'usage secondaire des données de santé¹ — leur réutilisation à des fins de recherche, d'innovation et de pilotage des politiques sanitaires — s'est désormais imposé comme un choix stratégique fondé, documenté, admis². Les récents rapports remis aux pouvoirs publics français et européens convergent vers un diagnostic clair : face aux défis majeurs de notre système de santé et à l'impérative nécessité de renforcer la souveraineté scientifique et technologique de la France, les données de santé constituent un potentiel immense pour anticiper les grandes transitions et les vulnérabilités à venir, orienter les priorités scientifiques pour y répondre, structurer les programmes prioritaires de recherche et éclairer les décisions relatives à la santé publique.

Si les progrès accomplis en ce sens ces dernières années sont significatifs et nombreux, l'exploitation des données reste freinée par des obstacles techniques, réglementaires, culturels et politiques.

[1] Sur la base des définitions proposées par le RGPD ou la CNIL, les données de santé sont des informations relatives à l'état physique ou mental d'une personne, passée, présente ou future, ainsi qu'aux soins qui lui sont prodigués, permettant de décrire, évaluer ou anticiper sa santé. Elles comprennent à la fois des données médicales directes (diagnostics, traitements, examens), des données génétiques ou biologiques, et des données (éducation, travail, environnement...) qui, par leur croisement ou leur utilisation à des fins médicales, révèlent des informations sur la santé d'une personne.

[2] Notamment avec le rapport de Jérôme Marchand-Arvier, Fédérer les acteurs de l'écosystème pour libérer l'utilisation secondaire des données de santé, (2023)

Ce constat appelle un changement de regard : et si les données de santé, longtemps enfermées — étant donné leur caractère sensible — dans une grammaire de la défiance, devenaient au contraire l'instrument d'une reconnexion du citoyen à la science, en lui offrant la possibilité de contribuer directement à la production de connaissances et à une recherche clairement tournée vers l'amélioration de sa santé, individuelle et collective ? Et si, face à une époque marquée par la post-vérité, minée par la guerre cognitive et la désinformation, les données devenaient le socle d'un nouveau pacte social et solidaire, l'armature d'une démocratie scientifique ? À l'aune de ces questions, c'est un changement de modèle qui s'opère pour garantir la soutenabilité de notre système de santé, et dont nous devons décider collectivement.

C'est dans cette perspective que l'Agence de programmes de recherche en santé (APRS), confiée à l'Inserm, publie ce plaidoyer. Ayant pour rôle de fédérer l'écosystème de la recherche-santé autour de l'identification, l'accompagnement et le rayonnement des grands programmes de recherche prioritaires, l'APRS s'appuie sur l'usage stratégique des données comme moteur de choix éclairés, de projection et d'anticipation. Ce présent travail s'appuie sur l'analyse approfondie d'une littérature scientifique et institutionnelle particulièrement riche, mais surtout sur l'audition de nombreux experts, chercheurs, décideurs publics, entrepreneurs et représentants des administrations, qui ont accepté de partager leurs réflexions, leur expérience et leur vision du sujet. Leurs contributions ont permis d'appréhender l'ensemble des enjeux liés à l'usage secondaire des données de santé et d'alimenter la réflexion stratégique conduite par l'Agence sur leur articulation avec le pacte social, dans une démarche d'écoute, de concertation et de co-construction.

L'ambition de l'Agence est ici de contribuer à dessiner les contours de ce nouveau pacte social de confiance, où données, recherche et participation citoyenne se conjuguent pour orienter la recherche et l'innovation au service d'un système de santé en pleine mutation.

I. ÉTAT DES LIEUX

ENTRE PROMESSE ET CONTRAINTE — LA RÉVOLUTION DES DONNÉES EN MOUVEMENT

1. Un potentiel largement reconnu

Les données de santé sont identifiées comme levier incontournable de transformation de la recherche, du soin et des politiques publiques. Leur potentiel se décline à plusieurs niveaux : capacité à prédire l'évolution des pathologies grâce à des modèles populationnels et à l'intelligence artificielle³; possibilité de renforcer la prévention, en détectant plus précocement des facteurs de risque individuels ou collectifs ; aptitude à personnaliser les traitements et à ajuster les prises en charge aux besoins spécifiques de chaque patient ; capacité à réduire les temps, coûts et risques de l'ensemble de la chaîne de valeur de la R&D au service de l'innovation diagnostique et thérapeutique ; faculté à évaluer les stratégies thérapeutiques et organisationnelles en vie réelle ; opportunité d'améliorer l'efficacité économique et la soutenabilité budgétaire des systèmes de santé⁴ en orientant plus finement les ressources grâce à une analyse en granularité des parcours de soin.

[3] Cédric Villani, Donner un sens à l'intelligence artificielle. Pour une stratégie nationale et européenne (2018)

[4] Ministère des Solidarités et de la Santé Les données de santé au service de la recherche et de la santé publique (2021)

Ainsi, qu'il s'agisse du cadre national ou communautaire, les données de santé ne sont plus considérées comme un outil strictement technique, mais une partie intégrante de la réponse pour relever les nombreux défis — augmentation de la prévalence des maladies chroniques⁵, vieillissement, inégalités territoriales⁶, sociales et environnementales, addictions⁷, hausse des troubles mentaux, hausse de la mortalité infantile⁸, contrainte budgétaire, désinformation et défiance dans la science⁹, retrait de grandes puissances des investissements pour une santé mondiale.¹⁰ À l'échelle européenne,¹¹ nationale,¹² territoriale,¹³ les données constituent une ressource stratégique au service de changements de paradigmes — passer du curatif au préventif, de l'individuel au populationnel, de la consommation de soins à l'anticipation des besoins, de l'opacité à la confiance.

[5] En 2021, en France, 12 millions de patients souffraient d'une maladie chronique reconnue dans le cadre du dispositif d'affection de longue durée (ALD). La prévalence de ces maladies est passée de 14,6 % en 2008 à 17,8 % en 2021, notamment à cause du vieillissement de la population – Drees, Mieux connaître et évaluer la prise en charge des maladies chroniques (2023)

[6] Les 10 % de la population les mieux dotés en médecins généralistes ont accès en moyenne à 5,6 consultations par an tandis que les 10 % de la population les moins bien dotés ont accès en moyenne à 1,4 consultation par an : l'accessibilité des premiers est ainsi 4,1 fois supérieure à celle des seconds. Ce rapport augmente de 5 % entre 2022 et 2023, ce qui témoigne de la croissance des inégalités territoriales d'accessibilité aux médecins généralistes – Drees, Accessibilité aux soins de premier recours en 2023 : dégradation de l'accessibilité aux médecins généralistes et aux infirmières, amélioration de l'accessibilité aux kinésithérapeutes, aux sages-femmes et aux chirurgiens-dentistes (2024)

[7] Rapport 2025 de l'AFDT, la cocaïne tue plus de 40 000 personnes par an, l'alcool est responsable d'environ 250 000 hospitalisations annuelles, pour un coût supérieur à 100 milliards.

[8] Le taux de mortalité infantile a augmenté, passant de 3,5‰ en 2011 à 4,1‰ en 2024 - Insee, Un enfant sur 250 meurt avant l'âge d'un an en France (2025)

[9] Nature, [Trust in scientists and their role in society across 68 countries](#) (2025)

[10] Le retrait américain de la solidarité internationale sous l'administration Trump notamment via le démantèlement de l'USAID aurait un impact considérable sur la santé notamment des enfants et des femmes dans de nombreuses régions du monde, [Amnesty International, Life at risk, 2025](#)

[11] Commission européenne, European Health Data Space – Proposal for a Regulation (2022).

[12] Institut Montaigne, Données de santé : libérer leur potentiel (2024)

[13] Institut Montaigne, Santé : faire le pari des territoires (2022)



2. Un écosystème français et européen en mouvement

C'est dans cette perspective que depuis une décennie, la France et l'Union européenne engagent des transformations structurelles majeures pour doter les politiques publiques et la recherche en santé d'infrastructures de données à grande échelle. En France, le Système national des données de santé (SNDS), qui rassemble des données de consommation de soins et données médico-sociales d'une ampleur inédite (assurance maladie, hôpital, causes de décès, etc.), a progressivement été élargi à d'autres sources, notamment les données issues des cohortes, des registres, des entrepôts hospitaliers, devenant l'une des plus importantes bases de données au monde. Cet élargissement a permis de renforcer la diversité et la richesse des informations mobilisables, tout en posant les bases d'une exploitation scientifique robuste et pertinente.¹⁴

[14] IQVIA, Why France is the Go-To Country to Conduct Real-World Studies, (2023)

Cette dynamique met également en lumière une dualité structurante : d'une part, les données de santé issues du soin produites dans le cadre habituel de la prise en charge médicale ; d'autre part, les données de recherche, souvent dérivées de ces mêmes sources mais enrichies d'annotations cliniques, biologiques ou comportementales complémentaires, permettant une compréhension plus fine des phénomènes étudiés. À cette articulation classique s'ajoute désormais un troisième ensemble en pleine expansion : les données de santé dites « périphériques », issues de l'utilisation d'applications numériques, de logiciels de gestion du parcours de soins, de dispositifs connectés (*wearables*), ou encore de plateformes d'assurance et d'acteurs privés. Ces données, produites en dehors du cadre strict du soin ou de la recherche clinique, contribuent à redéfinir les contours de ce que l'on entend par « information de santé », en ouvrant des perspectives nouvelles mais aussi des défis majeurs en matière de gouvernance, d'interopérabilité et d'éthique, comme détaillé infra.

En 2019, la création de la Plateforme des données de santé (PDS, autrement appelée Health Data Hub) a constitué une étape structurante : elle a permis de centraliser, documenter et faciliter l'accès aux données, en proposant un environnement sécurisé (avec de nombreuses réserves qu'il convient d'adresser notamment en termes d'hébergement) pour leur traitement et en jouant un rôle de guichet d'entrée et d'accompagnement des porteurs de projets, notamment de recherche. La PDS¹⁵ est devenue un acteur pivot, fédérant les initiatives et soutenant la diffusion progressive d'une culture de la donnée au sein du système de santé français, notamment à travers sa direction santé et de l'ouverture à la société. En parallèle, le développement de plateformes de données de médecine de ville à l'instar du Data in Primary Care (P4DP) et de multiples entrepôts de données de santé, principalement au sein des établissements de soins, et le développement des cohortes, représentent un progrès significatif.

[15] [Health Data Hub, Rapport annuel, \(2023\)](#)

Plus généralement, l'adoption récente de la Stratégie nationale de santé 2023-2033 et la Stratégie intelligence artificielle et données de santé offrent des nouveaux cadrages pour la France. L'implication d'acteurs clés — DGOS, Délégation au numérique en santé (DNS), Agence du numérique en Santé, Agence de la biomédecine sur les enjeux éthiques, etc. — dessinent un pilotage de plus en plus cohérent de l'usage secondaire des données de santé en France.

À l'échelle européenne, l'émergence du European Health Data Space¹⁶ (EHDS) marque une étape réglementaire et une impulsion politique décisives. Cet espace signe en effet un changement de paradigme : annoncé en 2022 par la Commission européenne, ce projet de règlement vise à créer un cadre commun et exigeant pour organiser l'usage primaire et secondaire des données de santé dans l'ensemble des États membres via des organismes responsables de l'accès aux données (ORAD) nationaux. L'ambition est double : d'une part, faciliter l'accès transnational aux données de santé pour la recherche (notamment en cadencant les délais d'accès), l'innovation et l'évaluation des politiques publiques ; d'autre part, harmoniser les pratiques et garantir un haut niveau de protection des droits fondamentaux des citoyens. Les projets financés par Horizon 2020, tels que HealthyCloud et FAIR4Health, ont de leur côté testé à l'échelle européenne la mise en œuvre des principes FAIR (Findable, Accessible, Interoperable, Reusable), confirmant que la « découvrabilité » et la réutilisation des données de santé demeurent des défis majeurs¹⁷.

En parallèle de ce mouvement, le cadre juridique encadrant l'accès aux données de santé s'est progressivement consolidé. Sur la base du règlement général sur la protection des données (RGPD), la réutilisation des données est guidée par l'intérêt public, la recherche scientifique ou historique ou à des fins statistiques.

[16] OECD, Health Data Governance for the Digital Age (2022).

[17] HealthyCloud, Survey on European Data Hubs and FAIRness (2023) ; FAIR4Health, Improving Health Research in the EU through FAIR Data (2021).

Dans ce cadre, la Commission nationale de l'informatique et des libertés (CNIL) conserve un rôle central d'autorité d'accompagnement à la mise en conformité, de contrôle et de sanction. Créé en 2020, le Comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé (Cesrees), instance dédiée à l'examen préalable à la CNIL de l'intérêt public des projets d'études nécessitant le recours à des données de santé a renforcé le contrôle sur la finalité de l'exploitation des données des santé, contribuant à sécuriser autant qu'à légitimer l'usage secondaire des données, notamment en recherche.

Ainsi, à la croisée des échelles nationale et européenne, se dessine un écosystème qui tend à renforcer la capacité des systèmes de santé à s'appuyer sur des données robustes, interopérables, accessibles, tout en plaçant la protection des individus et la confiance des citoyens au cœur de son architecture.



3. Des freins structurels

Malgré ces progrès manifestes, plusieurs verrous demeurent qui entraînent une sous-exploitation des données de santé. En premier lieu, la gouvernance reste insuffisamment lisible, avec des responsabilités éclatées et une faible coordination ou coopération des acteurs producteurs, qui développent régulièrement des réflexes propriétaires à l'échelle locale, nationale et européenne.¹⁸ On observe le glissement progressif d'un modèle jacobin et centralisé vers une logique d'appropriation locale : plusieurs acteurs envisagent de créer leurs propres entrepôts de données à partir de pans de données chaînées du SNDS, orientés selon leurs besoins opérationnels. Cette logique entraîne une fragmentation du dispositif national, une duplication de jeux de données et des vulnérabilités nouvelles en matière de sécurité et de protection de la vie privée. Malgré leurs qualités respectives, les bases de données restent morcelées et hétérogènes,¹⁹ mal appariées, entravant sérieusement l'interopérabilité,²⁰ pourtant fondamentale.

La nature encore largement médico-administrative des données disponibles, conjuguée à l'absence ou au retard d'intégration de sources cruciales (soins de ville, soins en vie réelle, imagerie, génomique, objets connectés, hôpitaux de proximité, bases de données françaises plus larges sur l'éducation, l'environnement, la fiscalité...), obère fortement leur potentiel scientifique et clinique. Il est cependant important de noter les récentes initiatives mises en place par l'État en ce sens — le Green Data for Health (GD4H), catalogues d'accès aux données environnementales de toute nature à même d'apporter des informations sur les déterminants environnementaux des inégalités de santé, appels à projets pour des pilotes de structuration de données HDH-GDti, etc.

[18] Cour des comptes européenne, Challenges in the use of big data for health research (2022)

[19] France Stratégie, Santé et données massives : un enjeu stratégique pour la France et l'Europe (2022).

[20] Sénat, Rapport d'information sur les données de santé, (2022).



À cela s'ajoutent des procédures d'accès encore longues, complexes, dissuasives pour les chercheurs, liées principalement aux incertitudes persistantes sur les conditions d'hébergement souverain. Une grande partie des infrastructures d'hébergement et de traitement des données repose en effet encore sur des technologies extra-européennes. Clouds américains, logiciels propriétaires, plateformes tierces : au-delà du problème majeur de « coincer » les utilisateurs de ces modèles, la plupart des acteurs s'accordent à dire que ces dépendances peuvent fragiliser notre autonomie scientifique, exposent nos données aux législations étrangères et limitent notre capacité à définir et à imposer nos propres standards de recherche, outre le fait qu'elles peuvent être exploitées à titre gratuit. Dans la perspective du déploiement d'une solution intercalaire conforme à la loi SREN et de la définition d'une infrastructure souveraine et certifiée SecNumCloud, la migration progressive de la PDS vers un hébergement national représente une opportunité majeure pour renforcer la souveraineté numérique et fluidifier l'accès des chercheurs aux données. Cette évolution permettra également de garantir l'indépendance technologique des patients et des chercheurs français, tout en assurant à la France la maîtrise de ses priorités scientifiques et stratégiques.

Car dans cette vulnérabilité technologique, il existe un risque réel d'appropriation : lorsque les données ou les traitements sont concentrés entre les mains de quelques acteurs dominants, ceux-ci peuvent imposer leurs conditions (monétaires, contractuelles ou techniques), modifier brutalement leurs tarifs, verrouiller les systèmes.

Un modèle économique flou, une culture du partage insuffisamment enracinée, un manque de connaissance, de reconnaissance et de valorisation d'étapes essentielles post-collecte (data management, analyse...), un manque de coordination et d'impulsion interministérielle viennent obérer plus généralement le potentiel des données pour la santé — récemment documentés par la mission Marchand-Arvier²¹—, qui appelle à dépasser ces blocages à travers une démarche de fédération, d'acculturation et d'ouverture des différentes parties prenantes. Une mise en cohérence du système, une standardisation imposée, des référentiels de qualité et de sécurité, le développement de catalogues de métadonnées, des critères d'auditabilité (journalisation systématique, audits de biais et de sécurité) et une gouvernance partagée paraissent fondamentaux.

Dans cette perspective, le financement pérenne, structurel et harmonisé des entrepôts de données de santé, l'atterrissage de la solution intercalaire sur un hébergeur conforme, et la garantie à terme d'une infrastructure souveraine, sécurisée et alignée sur l'EHDS, répondent à une attente politique forte. Cela permettrait enfin de structurer des guichets uniques pour les projets de recherche, aujourd'hui dispersés, afin de mieux faire remonter les besoins et de les exprimer dans un langage commun.

[21] Jérôme Marchand-Arvier, *Fédérer les acteurs de l'écosystème pour libérer l'utilisation secondaire des données de santé* (2023).

L'exemple GISAID

Plateforme mondiale principale des séquences génomiques du SARS-CoV-2, GISAID illustre les risques d'une dépendance à des infrastructures privées non transparentes. Créée en 2008 pour faciliter le partage mondial des données sur les virus grippaux, elle concentre plus de 15 millions de séquences mais fait l'objet de vives critiques (gouvernance opaque, règles d'accès arbitraires et pouvoir excessif de son fondateur). Cette crise révèle la fragilité d'un modèle mondial reposant sur des acteurs hors contrôle public, et souligne la nécessité de dispositifs souverains et ouverts pour sécuriser le partage des données scientifiques stratégiques.

Aussi, le manque d'interopérabilité et d'autonomie technologique contraint parfois les chercheurs français, dans le cadre de consortiums de recherche qui leur en ouvrent l'accès, à l'utilisation de données américaines, données qui reposent pourtant sur des profils épidémiologiques sensiblement différents : par exemple, le taux d'obésité²² y est bien plus élevé qu'en Europe, les types de comorbidité sont plus sévères, les complications plus fréquentes. Les maladies cardiovasculaires²³, certains cancers liés au mode de vie, ainsi que certaines infections sexuellement transmissibles, présentent des profils et des taux différents, tout comme les maladies respiratoires chroniques²⁴ ou les addictions.

[22] Prevalence and Correlates of Overweight and Obesity in 12 European Countries in 2017–2018

[23] Harvard, The state of cardiovascular disease in G20+ countries

[24] Geographic distribution of COPD prevalence in the world displayed by Geographic Information System maps (2019)

En introduisant ces données dans nos modèles, le risque d'entraîner des algorithmes biaisés, mal adaptés aux réalités européennes, et par conséquent de fausser la prise en charge des patients, est grand.²⁵ Cet état de fait ajoute de surcroît au régulateur en charge de la validation de tels dispositifs médicaux une vigilance particulière accrue. Une réflexion doit également être portée sur les pathologies des femmes et des minorités. Selon Élodie Germani, chercheuse Inserm, dont la thèse s'est intéressée aux applications de l'IA en imagerie médicale, la recherche est fortement obérée par les jeux de données disponibles.²⁶ À l'instar de l'endométriose, la recherche sur les pathologies féminines se voit par exemple confrontée à une surreprésentation des pelvis masculins dans les datasets d'entraînement existants. Obtenir des données variées et de qualité est donc essentiel pour progresser en recherche et en santé, mais un levier d'adhésion populaire auprès de toutes, tous et chacun(e).

Le développement de l'IA appliquée à la santé (et à la recherche en santé) met en évidence des besoins complémentaires :

- un encadrement éthique renforcé, pour améliorer la transparence, lutter contre les biais sociaux²⁷ et assurer la protection des droits des patients²⁸;
- une formation interdisciplinaire adaptée pour les décideurs publics, les professionnels de santé et les chercheurs pour analyser, maîtriser, interpréter, utiliser ces technologies de manière responsable et harmonisée.

[25] Plus généralement, les jeux de données actuels présentent souvent un déséquilibre géographique et phénotypique, avec une sur-représentation du "Global North" (USA, Europe, Asie), susceptible d'introduire des biais à l'échelle mondiale. L'enjeu se pose également en termes d'évaluation : les données doivent refléter la population cible et inclure la variabilité inter-individuelle pertinente. Il est crucial de mesurer les performances sur différents sous-groupes pour garantir que le modèle fonctionne de manière équitable pour tous et toutes. Audition d'Elodie Germani par l'APRS

[26] Audition Elodie Germani par l'APRS, 3 octobre 2025 ; [Exploring and mitigating analytical variability in fMRI results using representation learning](#)

[27] [Conseil de l'Europe, L'impact de l'intelligence artificielle sur la relation médecin-patient \(2022\)](#).

[28] [Sur ce point, nous pouvons noter l'importance de l'intégration d'une garantie humaine obligatoire dans l'IA act européen](#)

Surtout, l'adhésion citoyenne, indispensable aux usages secondaires des données de santé, reste limitée²⁹. Les patients se montrent généralement favorables à la mise à disposition de leurs données lorsqu'elles sont gérées par des acteurs publics ou académiques, mais demeurent réservés vis-à-vis des usages privés³⁰ et plus encore étrangers, comme le confirment les enquêtes de la CNIL³¹. Si certaines initiatives, à l'instar des comités citoyens déployés par la DNS, contribuent à sensibiliser les citoyens, la mobilisation demeure hétérogène et limitée sur l'ensemble du territoire, sans véritable dynamique globale pour une véritable culture de la donnée de santé. À l'échelle nationale, les efforts doivent être intensifiés pour massifier l'appropriation des enjeux, faire évoluer les mentalités et instaurer durablement la conscience collective que les données de santé, tout en étant sensibles, constituent une ressource précieuse pour la recherche et le soin lorsqu'elles sont utilisées dans un cadre sécurisé, consenti, et maîtrisé.

Les freins structurels identifiés mettent tous en évidence que la confiance reste un enjeu déterminant : elle est indispensable pour que les données de santé puissent pleinement déployer leur potentiel scientifique, clinique et social. Seule une gouvernance partagée intégrant citoyens, décideurs publics, professionnels de santé, employeurs et chercheurs garantirait l'appropriation collective et la confiance nécessaires à la libération du potentiel des données de santé.

[29] En 2025, 67% des citoyens se sentent insuffisamment informés, notamment sur les données de santé pour la recherche et 58% souhaitent en savoir plus - Résultats du baromètre de connaissance du HDH.

[30] CCNE, *Avis Plateformes de données de santé / enjeux d'éthique* (2023)

[31] CNIL, *Entre partage et protection, quelle éthique pour l'ouverture des données ?* (2021)

II. LEVIER(S)

UN PACTE SOCIAL RENOUVELÉ GRÂCE AUX DONNÉES DE SANTÉ

1. Créer un élan concertatif et démocratique autour des données

Les données de santé engagent le citoyen d'un bout à l'autre : à la fois donneur de la data et receveur du soin qu'elle a permis d'ajuster, à la fois contributeur économique et usager social. La participation citoyenne, souvent évoquée comme une promesse majeure de l'usage des données de santé, reste à ce jour encore embryonnaire dans sa mise en œuvre concrète. Il paraît pourtant indispensable d'associer les citoyens à des exercices concertatifs avec élus, chercheurs, professionnels, territoires, entrepreneurs, employeurs et syndicats, autour de la définition, du sens, des usages, de la gouvernance et de la destination des données de santé. Une première consultation nationale pourrait porter sur le don de données par défaut, et la possibilité de s'inscrire sur un registre de refus d'usage de ses données, comme sur le principe du don d'organes³². Il pourrait également s'agir de développer des concertations sur la priorisation des programmes de recherche dont l'identification est facilitée par les données populationnelles.

[32] Présumé sauf si la personne a exprimé de son vivant le refus de donner, en informant ses proches ou en s'inscrivant sur un registre national des refus.

Des outils tels que l'IA WeLearn, développée par le Learning Planet Institute ou des techniques comme le *broad listening*³³ – écoute large, multidimensionnelle et proactive permettant de détecter des opportunités ou menaces avant qu'elles ne deviennent évidentes, et de nourrir les objectifs stratégiques de manière éclairée – pourrait être utilement mobilisés dans ce cadre. L'adhésion citoyenne doit aussi être évaluée en continu via des enquêtes, consultations et études.³⁴

Dans un contexte marqué par la désinformation, la surinformation et la manipulation des contenus, amplifiées par les technologies d'intelligence artificielle, il devient essentiel de refonder nos décisions et nos débats publics sur la donnée objective et la preuve. Alors que certains pays, comme la Suisse³⁵ ou le Danemark³⁶, consacrent de nouveaux droits fondamentaux tels que l'intégrité cognitive ou le droit d'auteur sur son propre corps, la construction d'un pacte social de la donnée apparaît comme une condition de notre résilience collective : un engagement commun par et pour des données fiables.

[33] UN Global Pulse (UNGP) est une initiative des Nations Unies qui utilise les données massives et l'intelligence artificielle pour mieux comprendre les besoins des populations et anticiper les crises humanitaires ou sociales.

[34] Notamment le baromètre du numérique (Arcep, Arcom, CGE, ANCT / Crédoc)

[35] Protection de l'intégrité cognitive et de la liberté de penser

[36] Deepfakes et arnaques IA : le Danemark en passe d'accorder à ses citoyens un droit d'auteur sur leur propre corps

Initiatives existantes

L'**Arbre des connaissances** est une association qui promeut le dialogue entre chercheurs, enseignants et jeunes autour des sciences, en favorisant la curiosité, l'esprit critique et la compréhension de la recherche scientifique. Plus généralement, le rôle des associations est clé pour le maillage territorial et l'accès au public le plus défavorisé grâce à « l'aller-vers ».

À travers son Avis citoyen, la **Fondation Roche** et ses partenaires recueillent plusieurs dizaines de propositions concrètes issues du travail de 36 citoyens sur l'inclusion numérique en santé.

Le dispositif des **Assises citoyennes du Numérique en Santé** s'appuie sur une méthodologie éprouvée de démocratie participative inspirée des conventions citoyennes, visant à associer directement les citoyens à la réflexion et à la formulation de recommandations sur des enjeux stratégiques de santé et de recherche. Chaque édition réunit un comité citoyen composé d'environ trente participants représentatifs de la population française. En moyenne, près de 80 % des propositions formulées sont reprises ou intégrées dans les suites opérationnelles.

L'**Académie des données de santé**, lancée par la PDS en 2025, est un espace permettant d'accéder à des formations pour comprendre les enjeux et le fonctionnement de l'usage des données pour la recherche, et un Guichet associatif des données.

Depuis 2018, l'Etablissement français du sang a créé l'**EFS Social Lab**. Sur le modèle des « Labs » d'entreprise, cette structure a pour finalité d'analyser les évolutions sociétales et comportementales des parties prenantes de l'EFS, au premier rang desquels les donneurs, actuels ou potentiels, et d'émettre des préconisations opérationnelles ou d'axes de recherche. Ses travaux s'appuient sur les connaissances académiques en matière de sciences humaines et sociales, l'écoute des donneurs et des réseaux sociaux, et l'écoute sociétale au travers de publications ou d'enquêtes d'opinion.



2. Informer, acculturer, mobiliser, convaincre

Le « siècle de la donnée » exige un changement culturel. En son cœur, les missions de vulgarisation et de pédagogie doivent reposer sur des « ambassadeurs » de confiance de la donnée : médecins de ville, professionnels de santé, associations de patients, citoyens convaincus. Ces relais doivent avoir la capacité de donner à voir la destination des données et l'utilité concrète et immédiate de la recherche pour tous : celle de mieux soigner, dans un temps immédiat (recherche clinique, essais thérapeutiques, accès compassionnel) comme au long cours. Le développement de formations initiales et continues obligatoires sur les données et l'IA à destination des étudiants, des professionnels de santé et des décideurs publics paraît indispensable pour renforcer leurs compétences, adapter leurs pratiques et les aider à accompagner les patients et citoyens dans la compréhension et l'usage de leurs propres données, dans la continuité des efforts portés par la Stratégie Santé numérique (Appel à manifestation d'intérêt Compétences et Métiers d'Avenir notamment). Humaniser l'usage des données, en les présentant comme des leviers de solidarité par des incarnations de confiance³⁷ est impératif pour retisser les liens entre science et société.

[37] La cote de confiance des Français reste particulièrement élevée pour les professionnels de santé (91 %), les hôpitaux et cliniques privés (83 %) et les hôpitaux publics (82 %), Institut Viavoice

Cette ambition passe aussi par des supports d'information clairs, efficaces et accessibles, au sein de plateformes existantes (Dossier médical partagé, Mon espace santé, Doctolib) qui pourraient proposer des espaces dédiés à l'information sur les usages de la donnée, avec retours d'impact concrets et incarnés. Ces messages contribueraient à restaurer la confiance des citoyens, en montrant que leurs données sont, non seulement, protégées et encadrées par des garde-fous solides, mais aussi sont utiles et constituent des sources d'engagement. De grandes campagnes d'information nationales, simples et percutantes — à l'instar du célèbre « Les antibiotiques, c'est pas automatique », des campagnes vaccinales ou don du sang — aideraient, également, chacun à mieux comprendre le rôle des données dans la santé, notamment à destination des plus jeunes générations. Enfin, d'autres leviers étatiques existent pour acculturer les citoyens : l'éducation, à travers une sensibilisation scolaire et une éducation civique dès le plus jeune âge ; l'économie, à travers mobilisation de l'épargne citoyenne au service de la recherche, notamment sur les données.

Pour construire une **culture commune** des données de santé et dans l'anticipation de l'espace européen des données de santé, il pourrait être intéressant de développer un **guide de bonnes pratiques** et des indicateurs de performance sur l'implication citoyenne dans la collecte et l'utilisation des données, comme préalable aux grands projets de recherche. Ce guide pourrait utilement être intégré aux principes d'action du **futur ORAD national**.

- **Droit à la compréhension et à la maîtrise** : information claire, consentement éclairé et modulable, accès au suivi des usages
- **Droit à l'équité et à la représentativité** : inclusion des territoires et publics fragiles dans les bases
- **Devoir de solidarité** : partage des bénéfices issus des données pour améliorer la santé collective
- **Sobriété et souveraineté** : infrastructures responsables et européennes, non exposées à différentes lois extraterritoriales
- **Responsabilisation des acteurs** : citoyens co-acteurs, institutions garantes, entreprises redevables

3. Promouvoir le « citoyen-chercheur »

La donnée peut faire lien pour rapprocher le citoyen de la recherche, à condition que le citoyen en devienne un acteur à part entière, précisément par la donnée qu'il accepte de partager. Ce dernier pourrait donc n'être plus seulement « objet de données », mais *coconstructeur* de la connaissance et de la découverte, via des outils d'auto-suivi et d'encadrement médical favorisés notamment par l'essor de l'IA à l'instar des *patient-reported outcomes* et *patient remote monitoring*. L'enjeu est de commuer le patient passif en acteur central de sa santé, capable de comprendre et d'articuler ses propres données médicales, à l'image du *quantified self* chez les sportifs appliqué au suivi de santé. Ce rôle devient encore plus crucial à l'heure des IA décentralisées et de l'intégration des données cliniques et de vie réelle : le patient devient le « hub » qui relie parcours de soins, analyses et traitements, le mieux placé pour organiser et donner du sens à ses données. Avec l'accès généralisé aux IA génératives, la transmission du savoir se démocratise et s'horizontalise, offrant au patient de nouvelles capacités d'empouvoirement et de participation à sa santé.³⁸

En rétablissant la confiance et en redéfinissant un véritable pacte social autour des données de santé, nous pourrions lever les obstacles qui en limitent encore l'usage, transformer un potentiel encore sous-utilisé en orientations stratégiques et faire des données un levier de souveraineté scientifique, économique et démocratique.

[38] Audition d'Emmanuel Mussault, par l'APRS, 17 septembre 2025

III. IMPACT

BÂTIR UNE SOUVERAINETÉ SCIENTIFIQUE, ÉCONOMIQUE ET DÉMOCRATIQUE

1. Restaurer la confiance démocratique dans la science

Réengagées dans un récit positif, les données de santé viendraient aussi éclairer les décisions politiques et renforcer la crédibilité des institutions scientifiques, à une époque où la défiance envers la science et les savoirs fondés sur la preuve nous fragilise collectivement. À ce titre, la désinformation n'est plus seulement une dérive technologique : elle est devenue un instrument de puissance géostratégique, utilisé par certains États ou groupes organisés pour déstabiliser les systèmes démocratiques. La pandémie de Covid-19 a révélé combien le champ de la santé est un vecteur privilégié de manipulation, via la diffusion de traitements inefficaces, de campagnes antivaccinales ou de récits hostiles aux institutions scientifiques. Ces stratégies érodent la confiance des citoyens, compromettant simultanément la sécurité individuelle et collective. Les évolutions récentes aux États-Unis, en Europe et dans d'autres régions montrent que la contestation du rapport à la science et à la vérité est devenue un outil systématique de subversion politique.

Dans ce contexte, renforcer la résilience des sociétés suppose d'aller bien au-delà de la régulation technique des flux numériques. Il s'agit de développer une « immunité cognitive » collective : la capacité des citoyens à identifier et résister aux narratifs manipulateurs, à comprendre et valoriser les preuves scientifiques, et à maintenir un rapport éclairé à la donnée et à la connaissance. Dans cette perspective, la protection et l'usage responsable des données de santé, combinés à une communication claire et transparente, constituent un levier stratégique.

2. Assurer la soutenabilité économique

En réaffirmant la valeur collective des données comme bien commun au service de la connaissance et de la décision, en libérant leur potentiel, le nouveau pacte social offrirait les conditions de soutenabilité économique de notre système de santé, aujourd'hui en proie à une inédite tension.³⁹

D'abord en fondant et en accompagnant le virage préventif : le système de santé français demeure largement structuré autour du curatif, avec une logique de facturation et de consommation des actes médicaux. Ce modèle, passif, coûteux et fragmenté, traite les conséquences plus que les causes, et peine à infléchir les tendances épidémiologiques structurelles mettant à mal notre santé et la viabilité du système de soins — ainsi, 40 % des cancers et 80 % des maladies cardiovasculaires sont évitables. À l'inverse, des technologies fondées sur une connaissance fine et intégrée de l'individu ouvrent la voie à une approche radicalement différente.

[39] Le dernier rapport « Charges et produits » de l'Assurance maladie prévoit notamment un changement de paradigme majeur en faveur de la prévention, un effort inédit dans la dynamique de dépenses, des actions structurantes pour assurer la pérennité du système de santé - Améliorer la qualité du système de santé et maîtriser les dépenses : les propositions de l'Assurance Maladie pour 2026

L'IA offre déjà la capacité d'exploiter les données de manière pertinente⁴⁰ : une récente publication de *Nature* démontre une possible prédiction de plus de 1000 maladies, en fonction de l'historique médical de chaque individu, avec une précision comparable à celle des modèles existants spécialisés sur une seule maladie. L'outil Aire (AI-ECG risk estimation) serait capable d'identifier des risques de décès dans les 10 ans à venir à partir d'un ECG dans 78 % des cas.⁴¹ Plus largement, les modélisations populationnelles, les données longitudinales, l'approche personnalisée et territorialisée, l'évaluation socio-économique sont autant de données mobilisées pour cette impérieuse bifurcation, qui offrirait des perspectives économiques enthousiasmantes pour la France.⁴²

En termes de rationalisation médico-économique, les données pourront être utilement mobilisées pour optimiser le parcours de soin⁴³, améliorer la prise en charge et le suivi des patients, en cartographiant et en appariant plus finement l'offre et le besoin, notamment dans les territoires ruraux.

[40] Shmatko, A., Jung, A.W., Gaurav, K. et al. Learning the natural history of human disease with generative transformers. *Nature* (2025)

[41] Artificial intelligence-enabled electrocardiogram for mortality and cardiovascular risk estimation: a model development and validation study, Sau, Arunashis et al., *The Lancet Digital Health*

[42] La Suède a opéré ce virage il y a plus de vingt ans en investissant massivement dans la santé mentale, la nutrition, la lutte contre tabac et alcool, l'activité physique, l'accès précoce aux soins primaires, la digitalisation des parcours. Accompagnée d'une mobilisation coordonnée des collectivités locales et d'une forte implication des professionnels de santé, cette politique a été gagnante : on observe en effet un écart substantiel croissant avec la France depuis le milieu des années 2000, alors que les deux pays étaient comparables auparavant. Suède (2004) : 62 ans (hommes) / 60,8 ans (femmes) ; Suède (2020) : 72,7 années de vie en bonne santé ; France (2004) : 61,5 ans (hommes) / 64,3 ans (femmes) ; France (2020) : 64,6 années ; Source Eurostat – tableau interactif ; Source INSEE. A noter : A partir de 2020, la Suède baisse drastiquement, ce qui peut s'expliquer par le choix de gestion de la crise sanitaire (pas de confinement) avec un bilan épidémiologique contrasté. Audition de Pauline Martinot par l'APRS, 24 septembre 2025.

[43] CRAPS, L'exploitation des données de santé en France : un enjeu stratégique de transformation du système de soin et l'innovation en santé

À condition d’assurer une formation initiale et continue à la hauteur, apporter un soutien aux professionnels de santé dans leur travail quotidien,⁴⁴ éviter la redondance des actes et des prescriptions, gagner du temps de soin sur le temps administratif, objectiver les choix, éviter les biais humains dans les comptes rendus médicaux, mesurer l’impact, et guider la prise de décision médico-économique et budgétaire. Aussi, la montée en puissance de nouvelles méthodologies rendues possibles par les données massives offre de nouveaux potentiels économiques. Aujourd’hui basée sur des essais randomisés — modèle coûteux, sélectif, long et difficile à réaliser⁴⁵ —, l’évaluation thérapeutique pourrait gagner en efficience et en pertinence via les nouvelles méthodes pour les essais cliniques,⁴⁶ notamment les études observationnelles.

Ces méthodes peuvent être mises en place rapidement, sont faciles à répliquer, et sont, surtout, bien moins coûteuses pour le système. Plus généralement, l’objectif serait d’avancer vers une collecte plus systémique et mieux structurée des données de santé, afin de soutenir l’émergence d’un « système apprenant » permanent. Un tel système permettrait, à terme, d’améliorer la capacité à suivre l’efficience des dépenses de santé et à documenter plus finement l’efficacité et la sécurité des diagnostics et traitements, en s’appuyant davantage sur les données issues de la pratique clinique.⁴⁷ Cette approche continue permettrait de mieux orienter les ressources, d’identifier rapidement les sous-groupes de patients les plus répondeurs, de suggérer des combinaisons thérapeutiques pertinentes et de détecter précocement la perte d’efficacité de certains traitements, évitant ainsi des dépenses inutiles.

[44] Anact, Travail et numérique en santé, Les effets des transformations numériques sur les conditions de travail dans les établissements sanitaires et médico-sociaux (2024)

[45] SchorkNJ. Personalized medicine: Time for one-person trials. Nature (2015)

[46] État d’avancement des évolutions méthodologiques en recherche clinique De la recherche expérimentale au suivi d’évaluation en phase d’usage courant GT AIS/F-CRIN « L’évolution des méthodologies d’essais cliniques : nouveaux outils, nouveaux usages et conditions de recours »

[47] Audition de Philippe Ravaud par l’APRS, mercredi 1^{er} octobre 2025

Structurer une filière attractive et compétitive

C'est également en soutien de l'attractivité de la recherche que les données viendraient jouer un levier économique : historiquement leader, l'Europe désormais est au 3ème rang mondial en matière de recherche clinique,⁴⁸ avec une compétition croissante des États-Unis et de la Chine dans les domaines d'innovation les plus prometteurs. Les données pourraient servir à promouvoir une filière industrielle au service de la souveraineté économique, que ce soit pour développer de nouvelles approches fondamentales comme les jumeaux numériques⁴⁹ ou la médecine de précision, développer des métiers et des carrières (ingénieurs et data managers dans le public à travers des rémunérations compétitives et un crédit scientifique renforcé, notamment dans les publications), etc. Elles peuvent devenir à ce titre un nouveau levier compétitif, un nouvel instrument de *soft power* pour consolider la place de la France dans la compétition mondiale de la recherche-innovation, en plus d'augmenter les potentiels de publication.

Garantir la résilience socioéconomique

En qualifiant les besoins par territoire, les données peuvent aussi concourir à réduire les fractures sociales en santé, garantie d'une cohésion socio-économique à long terme : bien exploitées, la data permet de mieux cibler les populations les plus vulnérables, d'identifier les zones où l'accès aux soins est le plus fragile, et de concevoir des stratégies de prévention et de soin adaptées aux réalités locales.⁵⁰ Or un territoire où les écarts d'accès et de prise en charge sont réduits est un territoire plus robuste économiquement : moins de pertes de productivité liées à la maladie, moins de dépenses curatives évitables, et une population en meilleure santé pour soutenir l'activité collective.

[48] Devenir leader européen en recherche clinique : les propositions du Leem (2024)

[49] A l'instar de l'initiative juniore qui a été annoncée dans la Stratégie intelligence artificielle et données de santé 2025-2028

[50] Institut Montaigne, Santé : faire le pari des territoires (2022)

3. Piloter notre système de santé

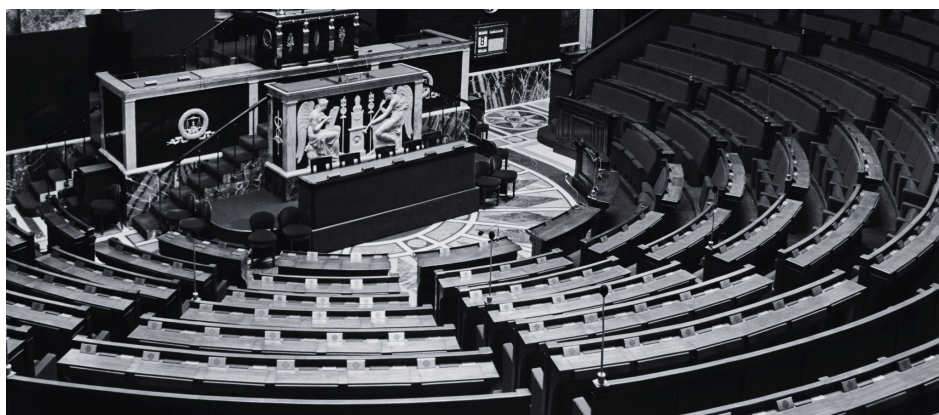
La libération des données permettrait enfin d'avoir une capacité de pilotage de notre modèle sanitaire, notamment via la définition de grands objectifs de santé, clairs, mesurables, partagés, et fondés sur la data. Ces objectifs doivent devenir le socle structurant à partir duquel s'orientent nos priorités (notamment de recherche), nos moyens et nos instruments d'évaluation. Une telle ambition suppose une stratégie interministérielle forte, intégrant planification, anticipation des besoins, et prospective.⁵¹ Dans ce cadre, la donnée de santé devient, non seulement, un outil d'observation, mais, surtout, un instrument de gouvernance, de pilotage et d'arbitrage stratégiques, permettant d'ajuster les politiques publiques au plus près des réalités territoriales et des évolutions sanitaires.

Pourrait découler de ces objectifs la structuration d'indicateurs sous-jacents et de mesures d'impact, d'évaluation des priorités et des budgets — comme il existe un « budget vert » ou un « budget sensible au genre », pourrait être envisagé un « budget sanitaire » capable d'appréhender l'ensemble des dépenses publiques (et pas seulement du PLFSS dédié) à l'aune de ces objectifs prospectifs, comme un nouveau contrat social de santé. Sur le modèle des objectifs de développement durable, désormais ancrés dans la culture commune, ou des *National Agreements* néerlandais,⁵² cette feuille de route collective offrirait un cadre de confiance, de lisibilité et de projection pour les citoyens, mieux à même alors de saisir l'intérêt immédiat et structurel de leur don de données au service de ces grands objectifs partagés.

[51] Exercices prospectifs de Haut-commissariat au plan, les travaux d'Horizon scanning de l'AIS, du HCAAM sur des ressources humaines en santé pour assurer l'équité d'accès aux soins sur tous les territoires...

[52] Government of the Netherlands, The National Prevention Agreement

La Stratégie nationale de santé, identifiée comme un outil particulièrement pertinent par l'écosystème,⁵³ pourrait ainsi constituer le cadre fédérateur de cette vision, la colonne vertébrale de ce pilotage de long terme. Sa mise en cohérence avec les PLF et PLFSS est un enjeu crucial pour sortir de la logique annuelle et court-termiste, et inscrire la santé dans un horizon de planification pérenne, prolongeant la généalogie de la protection sociale à la française : de la Sécurité sociale de 1945 à la CMU en 1999, jusqu'aux ambitions de pilotage et de planification fondées sur la donnée au XXI^e siècle.



[53] Institut Montaigne, Margaux Tellier-Poulain, Politique de santé : pour une stratégie pluriannuelle (2024)

CONCLUSION

Les données de santé constituent un levier majeur pour améliorer la connaissance, éclairer la décision publique et orienter le système de santé. Pour que ce potentiel se concrétise à l'heure du plus grand défi que traverse notre modèle sanitaire — celui de sa propre survie — il est nécessaire d'accompagner l'usage des données d'une transformation culturelle.

Nous sommes convaincus que celle-ci se trouve dans la définition et la promotion d'un nouveau pacte social de santé, qui associe pleinement citoyens, chercheurs, professionnels de santé, employeurs, décideurs et élus, autour des grands enjeux et des grands usages des données.

Seul ce pacte portera la connaissance, la confiance et la conviction nécessaires pour faire de la donnée un instrument de souveraineté et de cohésion sociale.

Le rôle de l'APRS

L'Agence de programmes de recherche en santé s'inscrit pleinement dans cette ambition collective. En fédérant les acteurs de la recherche en santé en France autour d'une gouvernance collégiale et en anticipant les grandes transitions sociétales, épidémiologiques, environnementales, technologiques, sanitaires et industrielles, l'Agence vise à identifier les priorités stratégiques de recherche en santé et d'élaborer sur cette base un nombre limité de grands programmes nationaux de recherche. L'Agence, confiée à l'Inserm, souhaite agir comme catalyseur et stratège au service de l'État : elle rassemble les communautés, maximise l'impact de la recherche pour la santé, le système de santé et la compétitivité économique et renforce la visibilité de la recherche française au niveau national et international. Les données de santé sont au cœur des 6 grands objectifs qui structurent son action.

1. Identification des priorités stratégiques

Les données de santé constituent un outil de prospective et d'anticipation au service de l'État, de la Nation, de la société : elles permettent d'identifier les priorités scientifiques à venir, mais également d'analyser les futurs usages, vulnérabilités et inégalités liés à la donnée.

2. Valorisation industrielle

En accompagnant la mise en œuvre des programmes, l'Agence évalue leur potentiel d'innovation et de transformation industrielles. Elle favorise la production de données robustes et exploitables, qui peuvent devenir un levier majeur pour la compétitivité économique du pays.

3. Lien science et société

Chaque programme stratégique est conçu comme un espace de co-construction avec les citoyens, les professionnels de santé, le secteur industriel et les décideurs publics. Cette démarche souhaite transformer les données générées par les programmes en bien commun, favorisant l'appropriation collective et renforçant la confiance sociale dans leur usage et leur exploitation.

4. Suivi des programmes

L'APRS a vocation à mettre en place des indicateurs rigoureux pour suivre l'impact scientifique, sociétal et économique des programmes qu'elle accompagne. Cette approche permet de capitaliser les données produites, d'identifier les succès et les axes d'amélioration, et d'assurer une transparence vis-à-vis des citoyens et des institutions.

5. Ancrage européen et international

L'Agence ambitionne de développer des partenariats internationaux avec les organisations programmatiques sœurs pour harmoniser les standards de production et d'exploitation des données des programmes prioritaires. Elle contribue ainsi à une dynamique multilatérale de la recherche, permettant d'anticiper les priorités émergentes et de renforcer la compétitivité et l'attractivité scientifique.

6. Impact

La donnée doit devenir un levier de transformation des politiques publiques, du système de soin et de la société dans son ensemble. En traduisant la connaissance issue des programmes en indicateurs d'impact, modèles et scénarios utiles à la décision, elle fait le lien entre la recherche et l'action publique. Cette capacité de rétroaction continue – nourrissant le décideur en amont pour éclairer les choix, et en aval pour évaluer les politiques mises en œuvre – est la clé d'un système apprenant, capable d'améliorer simultanément la qualité des soins, l'efficacité du système de santé et la souveraineté scientifique nationale.

REMERCIEMENTS

Personnes extérieures auditionnées

- **Régis Aubry**, Ancien membre du Comité Consultatif National d'Éthique, Président de l'Institut pour la prévention des vulnérabilités liées à la santé, co-coordonateur scientifique du Programme de recherche interdisciplinaire sur la fin de vie
- **Emmanuel Bacry**, Chercheur au CNRS et directeur scientifique au sein de la Plateforme des données de santé
- **Gabriel Bédubourg**, médecin militaire, Chef de la division Observatoire, Veille, Expertise de l'École du Val-de-Grâce, Délégué de l'Observatoire de la santé des vétérans
- **Jean François Brochard**, directeur général de Roche France
- **Lydie Catalano**, CEO IA Medical & Marketing
- **Hakim Chabane**, Directeur médical national sanitaire Croix-Rouge Française
- **Isabella de Magny**, CEO Inspiring Futures, Experte Santé de l'Institut Sapiens
- **Antoine Duclos**, Professeur des Universités, Praticien Hospitalier en Santé Publique, équipe METHODS au CRESS
- **Line Farah**, docteure en pharmacie, Conseillère prévention, santé publique et numérique, cabinet de la Ministre de la Santé, des Familles, de l'Autonomie et des Personnes handicapées
- **Jean Louis Fraysse**, co-fondateur de BOTdesign, Co-Président de SILICA, Vice-Président SoFia Santé
- **Élodie Germani**, chercheuse postdoctorale en intelligence artificielle pour l'imagerie médicale Inserm
- **Caroline Guillot**, Directrice santé et de l'ouverture à la société de la Plateforme des données de santé
- **Catherine Grenier**, Médecin-Conseil Nationale de la CNAM
- **Dominique Joseph**, Présidente du Groupe Santé et Citoyenneté du CESE
- **Julie Lagrave, Azilis Le Masne, Émile Moawad & Yann-Maël Ledouarin**, direction générale de l'offre de soins (DGOS)

REMERCIEMENTS

- **Franck Le Ouay**, co-fondateur et président de Lifen
- **Laurie Marrault**, Maîtresse de conférences à l'Ecole des hautes études en santé publique (EHESP)
- **Pauline Martinot**, Médecin & Docteur en Neurosciences et IA
- **Emmanuel Mussault**, Patient Partenaire en Recherche et Innovation
- **Etienne Minvielle**, Directeur du Centre de Recherche en Gestion de l'École polytechnique (i3-CRG), directeur de recherche CNRS
- **David Morquin**, Directeur Médical, Gouvernance des Données & Stratégie IA, Professeur Associé, Université de Montpellier
- **Céline Poucin**, directrice du mécénat scientifique de la Fondation Bettencourt Schueller
- **Fabienne Pioch-David**, Directrice filière sanitaire de la Croix-Rouge française
- **Philippe Ravaud**, directeur du CRESS, Professeur d'épidémiologie, Directeur du Centre EQUATOR, Chair du Cochrane Scientific committee
- **Gérard Raymond**, Président de France Assos Santé, membre du groupe Santé et citoyenneté du CESE
- **David Sainati, Christophe Mattler, Julie Richard**, Délégation au numérique en santé
- **François Taddei**, fondateur et président du Learning Planet Institute
- **Margaux Tellier Poulain**, Responsable de projets santé et protection sociale & **Lou Vincent**, chargée de projet santé, Institut Montaigne
- **Rodolphe Thiébaud**, directeur de l'École Universitaire de recherche Digital Public Health, coordinateur du Master « Public Health Data Science » à l'Isped

REMERCIEMENTS

- **Matthieu Tixier**, maître de conférence en Informatique à l'université de technologie de Troyes
- **Damien Vergé**, directeur de la stratégie, des études et des statistiques de la Cnam
- **Vincent Vuiblet**, premier vice-président délégué à l'IA et la communication de l'URCA ; PUPH CHU de Reims ; Coordonnateur réseau PAISaGE ; Président de la FNSDS
- **Marie Zins**, professeure des universités en santé publique / Unité de service Inserm « Cohortes épidémiologiques en population »



Inserm

La science pour la santé —————
————— From science to health