HEALTH
DATA HUB

RAPPORT ANNUEL 2020

« **L’accès aisé et unifié, transparent et sécurisé, aux données de santé pour améliorer la qualité des soins et l’accompagnement des patients. »**

Sommaire

Préface de Bruno Maquart 3
Édito de Stéphanie Combes 5

1 LE HEALTH DATA HUB, UNE INFRASTRUCTURE CLÉ POUR LA RECHERCHE EN SANTÉ 7

- Le HDH, une mise en place progressive 8
- Le HDH : un guichet unique pour un partage plus large des données de santé 9
- L'évolution du HDH en un an 13
- L'organisation du HDH 16

2 RETOUR SUR LA FEUILLE DE ROUTE STRATÉGIQUE DU HEALTH DATA HUB 21

- Enjeu 1 – Décloisonner le patrimoine de données de santé et améliorer la qualité 23
- Enjeu 2 – Renforcer l'usage des données par une offre de service attractive 33
- Enjeu 3 – Positionner la France comme un leader dans l'usage des données de santé 41
- Enjeu 4 – Garantir la participation de la société civile et promouvoir les usages des données de santé 45

3 BILAN DE L'ACCÈS AUX DONNÉES 51

- Accès aux données et perspectives 52

Préface

Bruno Maquart, Président

⟨⟨ La finalité du Health Data Hub, faut-il le rappeler, est d'aider à mieux soigner nos concitoyens. ⟩⟩

Le Health Data Hub a fêté son premier anniversaire à la fin de l'année 2020.
Sa création répondait à la volonté du Président de la République de faire de la France l'un des leaders du numérique en santé. Grâce au travail soutenu des équipes historiques de l'Institut national des données de santé et désormais de la cinquantaine de personnes qui composent le Hub, les premières réalisations sont là.
L'année 2020 a été singulière à plus d'un titre. Dans ce contexte, le Health Data Hub a réussi à tenir les objectifs ambitieux de sa feuille de route, notamment dans la mise en place de son offre de service ; il est ainsi en mesure d'accompagner des projets qui apporteront des solutions concrètes demain pour améliorer la qualité des soins. Le Hub a su également s'adapter pour participer, dans le même temps, à l'effort collectif de lutte contre la pandémie.

L'attrait de l'écosystème pour l'offre de service du Health Data Hub est certain. Une quarantaine de projets sont déjà accompagnés, dont un quart a été autorisé par la CNIL. Ces projets sont le résultat de partenariats entre acteurs de l'écosystème, preuve supplémentaire du rôle déterminant de l'intelligence collective dans les progrès de la recherche en santé. Plusieurs promoteurs sont près de pouvoir traiter de manière effective les données sur la plateforme technologique, après l'aboutissement des démarches réglementaires associées : CoviSas, de l'Université de Grenoble-Alpes, aidée par la startup Sêmeia, et Hydro, de la startup Implicity, par exemple. Le premier se donne pour but d'étudier les parcours de soins des patients qui, atteints de plusieurs pathologies, sont particulièrement fragiles face à la Covid-19 ; le second vise à corréler les données produites par des pacemakers connectés avec des données d'hospitalisation pour prévenir les crises d'insuffisance cardiaque.

En réunissant des bases de données dans le catalogue, le souhait du Health Data Hub est d'aller beaucoup plus loin. Au-delà de l'accompagnement des cas d'usage, il s'agit de développer une offre différenciée s'adressant aux nombreux acteurs de l'écosystème : registres, cohortes, établissements de santé, agences, startups, etc. C'est dans cette même logique de rapprochement que le Hub participe à l'initiative PariSanté Campus.
Lancée il y a quelques mois par le Président de la République, elle ambitionne de réunir à horizon 2028, sur un même site, plusieurs acteurs de la donnée de santé : INRIA, INSERM, Université Paris Sciences et Lettres, Agence du numérique en santé et Health Data Hub. Ce dernier sera ainsi en mesure d'intensifier encore ses collaborations et de démontrer sa valeur.

La crise a illustré la place centrale qu'occupent aujourd'hui les données : elles sont indispensables pour mieux comprendre et être en capacité d'anticiper, que ce soit du point de vue scientifique ou de celui du pilotage des politiques publiques. Cette situation n'a fait que confirmer l'intérêt et l'urgence de la mission dont est investi le Health Data Hub. Dans un

environnement international extrêmement compétitif, il doit garder un rythme d'activité soutenu et tenir les échéances qui lui ont été fixées.

La compétition qui se joue est mondiale ; le Health Data Hub doit dès lors investir le champ international. Il le fera dès 2021, puisqu'il a été désigné autorité compétente pour la France et chargé de coordonner la réponse nationale à l'action conjointe de la Commission européenne autour de la création d'un espace européen des données de santé, action phare de la stratégie européenne des données. C'est un axe majeur d'investissement et de visibilité pour le pays.

J'en terminerai en remerciant les équipes du Health Data Hub, animées par sa directrice, Stéphanie Combes, qui n'ont pas ménagé leur peine au cours de l'année 2020. Ma reconnaissance va également aux membres du conseil d'administration et de l'assemblée générale, ainsi qu'à celles et ceux qui ont prêté leur concours aux travaux du Hub, dont la finalité, faut-il le rappeler, est d'aider à mieux soigner nos concitoyens.

Édito

Stéphanie Combes, Directrice

⟨⟨ Un an après sa création, plusieurs réalisations sont à mettre à l'actif du Health Data Hub, conformément aux engagements pris dans la feuille de route stratégique. ⟩⟩

L'année 2020 a été marquée par la **mise en production**, en avril dernier, de la plateforme technologique qui vise à mettre à disposition des utilisateurs autorisés les données nécessaires à leurs projets. Les premières bases de données sont ainsi déjà disponibles, et les premiers utilisateurs exploitent les données avec les outils mis à disposition par le biais de cette plateforme. Par ailleurs, le HDH **accompagne une quarantaine de projets** dans leur démarche d'accès aux données, dont une vingtaine a été sélectionnée par le biais d'appels à projets. Plusieurs **événements fédérateurs** ont été organisés comme un Data Challenge en imagerie médicale en partenariat avec la Société Française de Pathologie ou encore une école d'hiver faisant intervenir des experts de renommée internationale. Le HDH a également été choisi pour coordonner la réponse française dans le cadre de l'action conjointe de la Commission européenne pour la création d'un European Data Space dans le secteur de la santé. Enfin, un **travail a été mené avec les associations de patients et d'usagers** sur la politique d'information du HDH vis-à-vis des citoyens.

Parallèlement à cela, l'année 2020 a été fortement **marquée par la pandémie Covid-19,** ce qui a d'une part eu pour effet de ralentir certains chantiers, mais a d'autre part permis d'accélérer la mise à disposition de la plateforme dans le cadre de l'arrêté du 21 avril 2020 afin de lutter contre l'épidémie de Covid-19, permettant aux projets contribuant à la gestion de la crise sanitaire d'accéder à certaines données.

L'année 2021 sera une année de **poursuite et d'accélération de la mise en oeuvre du HDH,** en inscrivant dans la durée des actions prévues dans le cadre de la feuille de route stratégique publiée en janvier dernier autour de 4 enjeux :

› décloisonner le patrimoine de données de santé et améliorer sa qualité ;
› développer une offre de services performante pour les utilisateurs de données de santé ;
› placer la France comme un leader de l'utilisation des données de santé dans l'écosystème européen et international ;
› garantir la participation de la société civile à l'utilisation des données de santé.

Par ailleurs, plusieurs actions viennent **enrichir les jalons de la feuille de route initiale** notamment s'agissant des relations avec les associations et la société civile, le développement de partenariats avec des organismes de recherche ou des industriels, ou encore le développement de l'empreinte européenne et internationale.

Enfin, l'année 2021 va permettre la **structuration organisationnelle et financière** du HDH à travers notamment le renforcement de la gestion des ressources humaines et la poursuite de l'effort de recrutement. ⟩⟩

A
100
90
80
70
60
50
40
30
20
300
60

LE HEALTH DATA HUB, **UNE INFRASTRUCTURE CLÉ POUR LA RECHERCHE EN SANTÉ**

À la suite du rapport Villani sur l'intelligence artificielle rendu public le 28 mars 2018, le Président de la République a affirmé sa volonté de faire de la santé un des secteurs prioritaires pour le développement de l'intelligence artificielle en France.

Le Health Data Hub, une mise en place progressive

Prévu par la loi du 24 juillet 2019 relative à l'organisation et la transformation du système de santé, le HDH est une structure publique dont l'objectif est de permettre aux porteurs de projet d'**accéder facilement à des données non nominatives** hébergées sur une plateforme sécurisée, dans le respect de la réglementation et des droits des citoyens. Ils pourront y croiser les données et les analyser dans le but d'améliorer la qualité des soins et l'accompagnement des patients.

Issu de la transformation de l'**Institut National des Données de Santé** créé en 2016, le HDH qui a été pensé comme un « guichet unique » aura également pour rôle de réunir, d'organiser et de mettre à disposition les données de santé issues d'une multitude de sources, telles que les données hospitalières, les cohortes, les registres… et bien sûr, la base de données du Système national des données de santé (SNDS).

L'offre de service s'adresse aux différentes parties prenantes, elle vise notamment à :

- soutenir les responsables de données dans la collecte, la consolidation et la mise en valeur de leur patrimoine ;
- offrir à tous les porteurs de projets un accès simplifié et accéléré aux données ;
- garantir à la société civile le respect des droits des citoyens ;
- innover aux côtés des acteurs de la recherche et de l'industrie.

> **« Donner son sang, c'est aider son prochain ; il en est de même du partage de ses données de santé. Les partager à son équipe soignante, c'est être mieux soigné ; les partager pour tous, c'est permettre que l'ensemble de nos concitoyens le soient aussi. »**
>
> **Gérard Raymond**
> Président de France Assos Santé,
> Vice-président du HDH

Le Health Data Hub : un guichet unique pour un partage plus large des données de santé

LE HEALTH DATA HUB EST UN ACCÉLÉRATEUR DE LA RECHERCHE EN SANTÉ

Les nouveaux outils, comme l'intelligence artificielle, vont permettre d'importants progrès dans l'accompagnement des patients, l'évaluation et le choix des traitements et la gestion du système de santé. C'est par le traitement et le croisement d'un grand volume de données de qualité, que les recherches les plus impactantes pourront être menées : améliorer le dépistage et le diagnostic d'une maladie, analyser les effets secondaires des traitements, faire évoluer les essais cliniques.

Apporter des réponses aux patients atteints de pathologies rares.

Le sarcome est une catégorie de tumeurs rares pour laquelle l'efficacité des essais cliniques arrive à ses limites. Depuis près de 40 ans, ils ont été conduits sans pouvoir trancher sur la pertinence de la chimiothérapie avant ou après la chirurgie. Dans ce contexte, l'analyse des données collectées auprès de patients dans le cadre de cohortes, associées aux données de consommations de soin issues de l'Assurance Maladie offre une réelle opportunité d'évaluation des traitements et de leur performance.

Prendre en charge les patients le plus tôt possible

Les crises d'insuffisance cardiaque causent près de 5 % des hospitalisations en France. En corrélant les données issues des pacemakers avec les données relatives aux hospitalisations, il est possible de concevoir des outils d'alerte qui détectent en amont des signaux faibles et permettent de prendre en charge les patients le plus tôt possible.

Proposer aux patients les meilleurs traitements sur le long cours

Les informations concernant les liens entre la prise de médicaments et leurs effets à long terme sont aujourd'hui très lacunaires, dans le cas général comme en transplantation d'organes. Dans le cadre d'un traitement à vie, la connaissance de ces relations est indispensable pour optimiser les stratégies thérapeutiques, le choix des doses mais aussi des formules de ces médicaments.

Appuyer le professionnel de santé dans un contexte clinique qui se complexifie

Les logiciels de prescription utilisent déjà des outils d'alertes pour notifier le professionnel lorsque différents traitements pris par un même patient peuvent révéler des incompatibilités avec des risques pour sa santé. Largement perfectibles, ces systèmes génèrent aujourd'hui tant d'alertes qu'ils sont souvent désactivés ou ignorés par les professionnels. L'analyse des données permet de prioriser les alertes et de filtrer celles qui doivent impérativement être prises en compte.

Prédire les trajectoires individuelles des patients et améliorer les actions de prévention

Les données de cohortes de patients suivis dans les centres experts Parkinson rapprochées de celles de l'Assurance Maladie permettent d'étudier la progression de la pathologie et ainsi d'améliorer le suivi des patients. Des outils pourraient ainsi être développés et fournis aux neurologues.

Gagner du temps médical, améliorer le dépistage et réduire les délais de diagnostic

Avec 11 883 décès estimés en 2017, le cancer du sein tue plus de femmes que tout autre cancer en France. Le dépistage précoce permet une baisse de 21 % de la mortalité. Le recours à l'intelligence artificielle pourrait encore améliorer le dispositif de dépistage organisé du cancer du sein en réduisant le taux de faux positifs et de faux négatifs.

LA LOI DE 2016 A CRÉÉ LE SYSTÈME NATIONAL DES DONNÉES DE SANTÉ (SNDS)

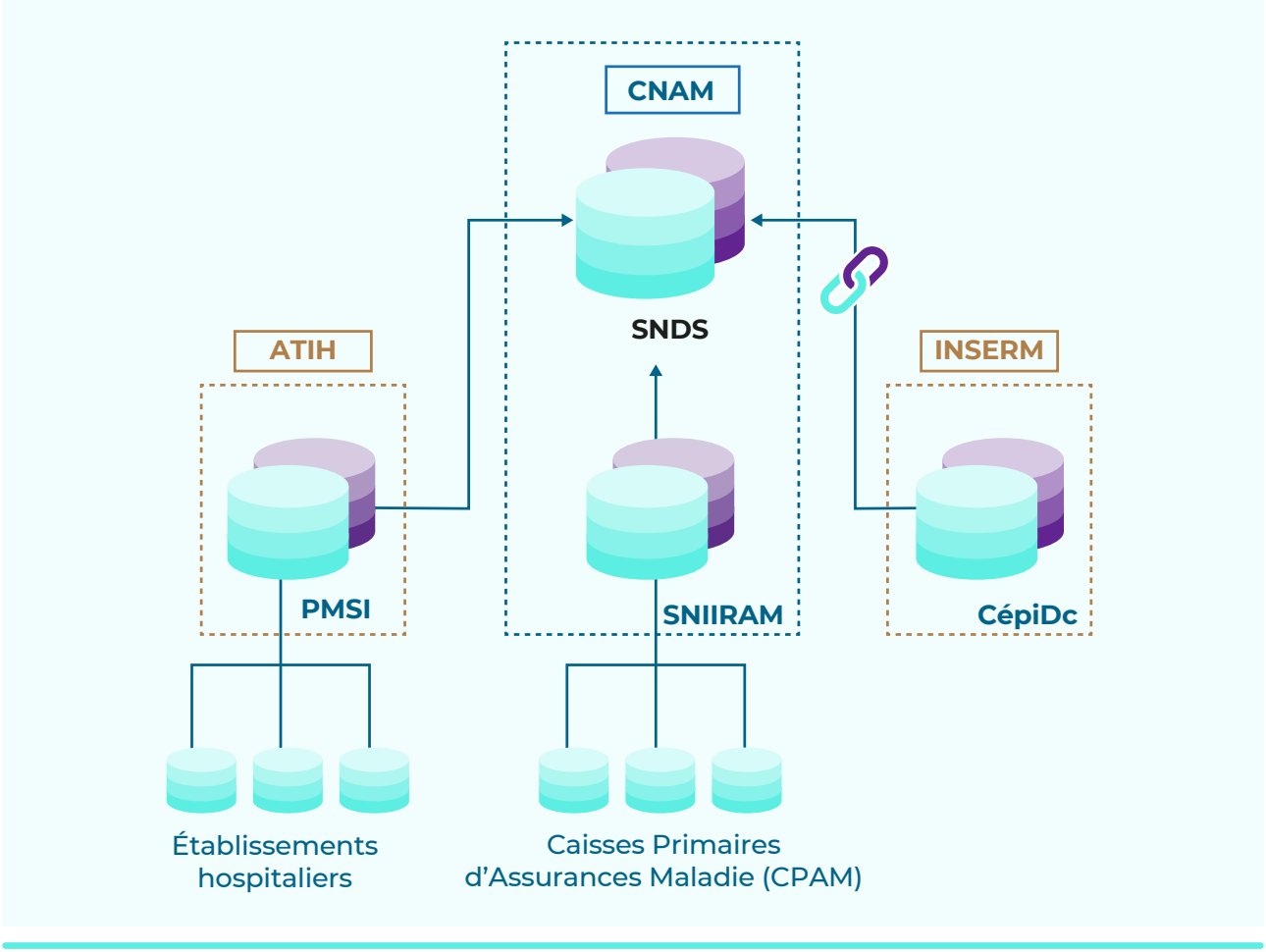
Créé en 2016 par la loi « modernisation de notre système de santé », le SNDS est initialement alimenté des trois bases de données suivantes rassemblées par la CNAM :

- **Le SNIIRAM :** l'ensemble des informations issues des remboursements de l'Assurance maladie.
- **Le PMSI :** les séjours hospitaliers, appelés résumé de sortie standardisé (RSS) classés dans des Groupes Homogènes de Malades (GHM) selon des critères médicaux et économiques réunis par l'ATIH.
- **Le CépiDC :** la base de Causes Médicales de Décès (BCMD) réunies par l'Inserm.

Ces données ont un potentiel important en recherche et sont utilisées depuis 2006 pour réaliser des études soit à caractère scientifique soit pour appuyer, déterminer et évaluer des politiques de santé et de dépenses.

En 2019, une nouvelle loi élargit le SNDS. On passe d'une base unique à un système de bases. Le SNDS élargi définit le cadre pour accueillir notamment des données cliniques : données de registres, de cohortes de recherche, d'entrepôts de données hospitalières, etc.

SYSTÈME NATIONAL DES DONNÉES DE SANTÉ



LE CATALOGUE VA RÉUNIR UNE COLLECTION DE BASES PSEUDONYMISÉES

La loi relative à l'organisation et à la transformation du système de santé, promulguée le 24 juillet 2019, a étendu le Système National des Données de Santé (SNDS). Les données médico-administratives y étaient initialement collectées, et seraient complétées par des données associées à un financement public dans le domaine de la santé afin d'élargir les usages au bénéfice de la santé de la population.
Parmis celles-ci, une collection de bases de données - le catalogue - sera mise à disposition par le HDH en plus du SNDS historique, renommé base principale, dont l'accès sera toujours possible par le portail de la CNAM par ailleurs.

Le catalogue rassemblera des copies de bases déjà existantes pour la plupart, jugées les plus pertinentes pour la recherche et l'innovation. C'est un comité stratégique piloté par l'État qui fixera les grandes orientations de ce catalogue qui se veut évolutif.
Cette collection sera construite de manière progressive et itérative, en partenariat avec les responsables de la collecte des données concernées.
La liste des bases sera portée par un arrêté pris après avis de la CNIL, lui-même découlant

d'un décret fixant le rôle de responsable de traitement du HDH et de la CNAM, dont la publication est attendue pour les prochaines semaines.

En créant ce système de bases de données, ce que la loi permet c'est une meilleure visibilité du patrimoine commun pour tout l'écosystème, une homogénéisation des règles d'accès et une facilitation de ces derniers avec la mise en place du HDH comme service sécurisé de mise à disposition du catalogue.
Les accès sont réglementés et se font dans le respect des droits des personnes.

LE HEALTH DATA HUB REPREND LES MISSIONS DE L'INDS TOUT EN LES ÉLARGISSANT

En prévoyant la création d'une plateforme des données de santé ou « HDH », la loi relative à l'organisation et à la transformation du système de santé entend élargir les missions de l'INDS en proposant davantage de services pour les demandeurs d'accès aux données de santé.

Comme l'INDS avant lui, le HDH assure le rôle de guichet unique pour faciliter l'accès aux données de santé dans le respect du droit des personnes et en assurant la transparence avec la société civile.

En complément de l'offre du portail de la CNAM, ce dernier met également à disposition des porteurs de projet, une plateforme sécurisée à l'état de l'art, offrant des capacités de stockage, de calcul, de rapprochement et d'analyse des données, permettant la conduite de projets innovants parfois considérés comme impossibles.

Sur cette plateforme, les données du catalogue seront mises à disposition ainsi que des données destinées à un projet en particulier.

Afin d'aider les porteurs à explorer le patrimoine de données disponibles, un catalogue de métadonnées, de la documentation ouverte et des bibliothèques de programmes sont également mis à disposition par la communauté.

L'évolution du Health Data Hub en un an

LA PREMIÈRE ANNÉE DU HDH A ÉTÉ RICHE DE RÉALISATIONS ET DE PROJETS POUR GARANTIR UN PARTAGE PLUS LARGE DES DONNÉES DE SANTÉ

La mise en œuvre de la plateforme et la préfiguration d'un catalogue de bases de données

L'année 2020 a été marquée par la mise en production de la plateforme technologique qui vise à mettre à disposition des utilisateurs autorisés les données nécessaires à leurs projets. Des premières bases de données, dans le cadre de l'arrêté du 21 avril 2020 autorisant le HDH à recevoir, stocker et mettre à disposition des données à caractère personnel concernant la santé nécessaires à la poursuite d'une finalité d'intérêt public en lien avec l'épidémie actuelle de Covid-19, sont ainsi déjà disponibles dans le catalogue. Les premiers utilisateurs exploitent les données avec les outils mis à disposition par le biais de cette plateforme.

L'accompagnement d'une quarantaine de projets, et l'ouverture à l'écosysteme

Le HDH accompagne une quarantaine de projets dans leur démarche d'accès aux données, dont une vingtaine a été sélectionnée par le biais d'appels à projets.

Au total, 8 de ces projets vont pouvoir démarrer, car ayant reçu une autorisation de la CNIL : 3 projets s'inscrivant dans la lutte contre la Covid-19 et 5 projets portant sur le cancer du sein, le sarcome, la greffe de rein et les interactions médicamenteuses.

Par ailleurs, le HDH et la CNAM collaborent pour partager l'expérience de la CNAM sur l'accompagnement et l'accès effectif aux données de la base principale du SNDS. Ces données sont mises à disposition par la CNAM. Il est nécessaire de maintenir cet investissement pour assurer l'accès pour tous les projets durant cette phase de montée en charge.

Une démarche de valorisation scientifique par l'organisation de quatre grands événements

Plusieurs événements ont été organisés dont un nouvel appel à projets en 2020, permettant aux 10 lauréats de bénéficier d'un accompagnement opérationnel et réglementaire du HDH et un accompagnement financier du Grand Défi. Un Data Challenge en imagerie médicale en partenariat avec la Société Française de Pathologie et le Grand Défi a également eu lieu, contribuant à l'avancée de la recherche en santé, et favorisant l'open data.
Le HDH a également été choisi pour coordonner la réponse française dans le cadre de l'action conjointe de la Commission européenne pour la création d'un European Data Space dans le secteur de la santé.

Il contribue ainsi avec des partenaires comme l'ANS, les Hospices Civils de Lyon, l'Inserm, l'Université Toulouse III et l'Université Aix-Marseille, à la structuration du paysage européen autour de l'usage des données de santé. La fin de l'année a aussi été marquée par la deuxième édition du colloque « Données de santé et intelligence collective », organisé avec le Grand Défi « Améliorer les diagnostics médicaux à l'aide de l'intelligence artificielle », et la Délégation ministérielle du Numérique en Santé mettant notamment en avant les initiatives internationales et européennes autour du partage des données.

La création d'une direction citoyenne pour être à l'écoute de la société civile et des associations d'usagers du système de santé

Parce que les données de santé concernent tous les citoyens, une direction dédiée a été créée. Elle a permis notamment la restitution d'un travail mené avec les associations de patients et d'usagers sur la politique d'information du HDH vis-à-vis des citoyens. C'est aussi l'accompagnement de projets associatifs, dont Vivre Covid-19 de France Assos Santé.

L'année 2021 poursuivra la même ambition en accélérant le déploiement, en enrichissant le catalogue de bases de données, en permettant le développement d'usages innovants à fort impact de la donnée de santé et en plaçant la France comme un des leaders de l'utilisation des données de santé dans l'écosystème européen et international.

LES TEMPS FORTS ET CHIFFRES CLÉS DU HDH

1 APPEL À PROJET
CO-ORGANISÉ AVEC LE GRAND DÉFI « AMÉLIORATION DES DIAGNOSTICS MÉDICAUX PAR L'IA » EN 2020

27 PROJETS PILOTES ET 9 PROJETS COVID
ACCOMPAGNÉS

8 PROJETS AUTORISÉS
PAR LA CNIL

1 VINGTAINE DE PARTENARIATS
EN COURS DE DISCUSSION AVEC DES RESPONSABLES DE DONNÉES

2 VERSIONS DE LA PLATEFORME
MISES EN PRODUCTION

2 BASES SUR LA PLATEFORME

3 PROJETS SUR LA PLATEFORME

1 CINQUANTAINE DE COLLABORATEURS

1 NOUVEL ESPACE DE TRAVAIL

547 PARTICIPANTS AU DATA CHALLENGE CO-ORGANISÉ AVEC LA SFP ET 4 VAINQUEURS

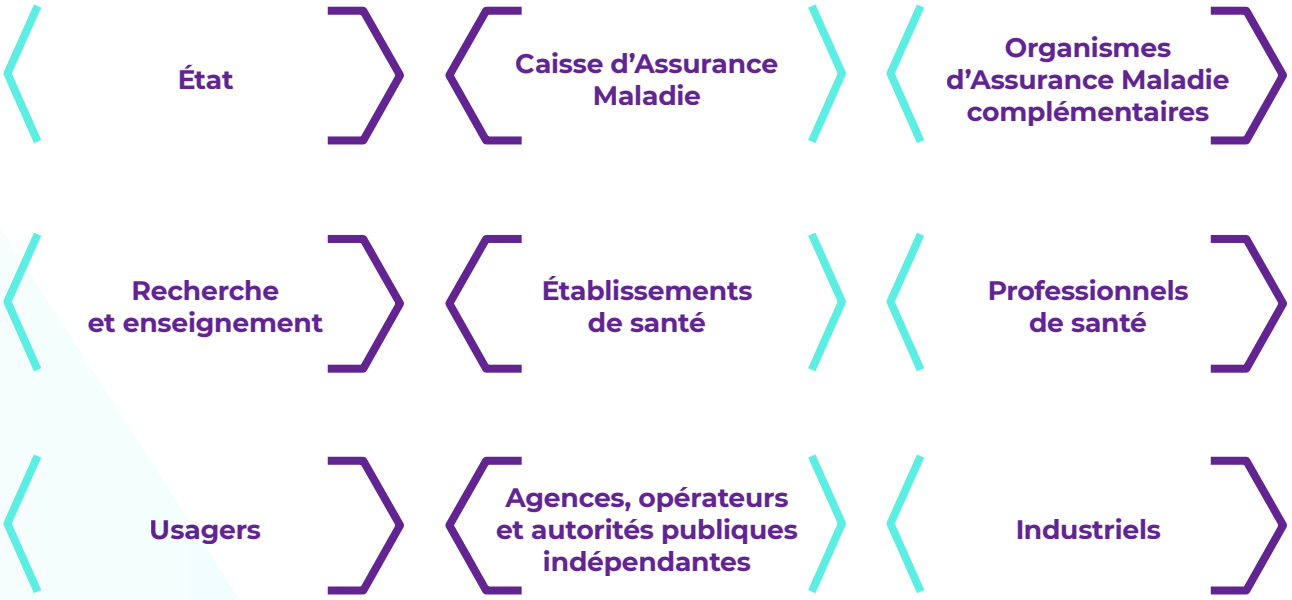
438 PARTICIPANTS
À L'ÉCOLE D'HIVER « AI4HEALTH »

12 ACTEURS FRANÇAIS MOBILISÉS
DANS LE CADRE DE L'ACTION CONJOINTE EUROPÉENNE « TOWARDS A EUROPEAN HEALTH DATA SPACE »

L'organisation du Health Data Hub

LA GOUVERNANCE

Le HDH réunit 56 parties prenantes siégeant à son assemblée générale, dont l'État, les organismes représentant les usagers du système de santé, les professionnels de santé, les producteurs de données de santé, les utilisateurs privés ou publics de données de santé, etc.
Les membres sont répartis en neuf collèges représentant la diversité de l'écosystème des données de santé. Les représentants de ces collèges siègent au conseil d'administation. Pour l'État, c'est le ministère des Solidarités et de la Santé et le ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation qui participent.
Le HDH est présidé par Bruno Maquart et son vice-président est Gérard Raymond, président de France Assos Santé.
Le HDH met en œuvre en partenariat avec la CNAM les grandes orientations stratégiques relatives au Système National des Données de Santé (SNDS) fixées par l'État.



LES MEMBRES DU HEALTH DATA HUB

L'ÉTAT

MINISTÈRE DES SOLIDARITÉS ET DE LA SANTÉ
Liberté Égalité Fraternité

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR, DE LA RECHERCHE ET DE L'INNOVATION
Liberté Égalité Fraternité

MINISTÈRE DE L'ÉCONOMIE DES FINANCES ET DE LA RELANCE
Liberté Égalité Fraternité

MINISTÈRE DES ARMÉES
Liberté Égalité Fraternité

MINISTÈRE DU TRAVAIL DE L'EMPLOI ET DE L'INSERTION
Liberté Égalité Fraternité

MINISTÈRE DE L'ÉDUCATION NATIONALE, DE LA JEUNESSE ET DES SPORTS
Liberté Égalité Fraternité

MINISTÈRE DE L'ACTION ET DES COMPTES PUBLICS
Liberté Égalité Fraternité

PREMIER MINISTRE
Liberté Égalité Fraternité

LES CAISSES D'ASSURANCE MALADIE

L'Assurance Maladie – Caisse Nationale

msa santé famille retraite services

UNRS

LES ORGANISMES D'ASSURANCE MALADIE COMPLÉMENTAIRES

CTIP

Fédération Française de l'Assurance

MUTUALITÉ FRANÇAISE

unocam

LA RECHERCHE ET L'ENSEIGNEMENT

cea

cnrs

CPU Conférence des Présidents d'Université

EHESP

Inria

Inserm

IRDES Institut de recherche et documentation en économie de la santé

LES ÉTABLISSEMENTS DE SANTÉ

FEHAP Santé Social - Privé Solidaire
FHF Fédération Hospitalière de France
FHP Fédération de l'Hospitalisation Privée
HAD Fédération Nationale des Établissements d'Hospitalisation à Domicile
Assistance Publique Hôpitaux de Paris
CME CHU Commission Médicale d'Établissement Conférence des Présidents
unicancer
CHU France Directeurs Généraux
CHU Centre Hospitalier Universitaire Limoges
Hôpitaux de Toulouse
CHUGO Hôpitaux Universitaires Grand Ouest

LES PROFESSIONNELS DE SANTÉ

Fsm Fédération des Spécialités Médicales
UNPS Union Nationale des Professionnels de Santé
Ordre National des Médecins Conseil National de l'Ordre
Ordre national des pharmaciens

LES USAGERS

France Assos Santé La voix des usagers

LES AGENCES

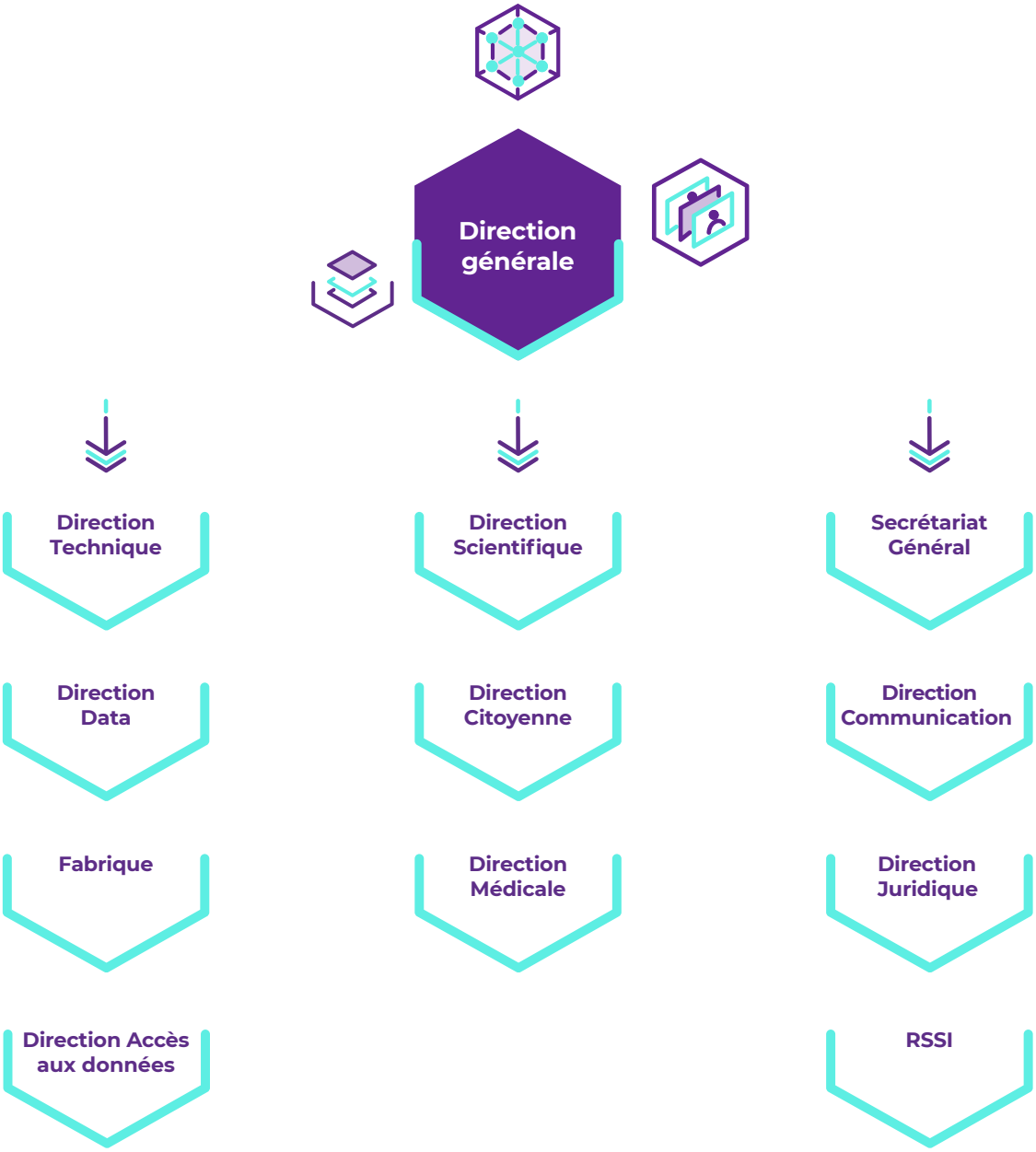
atih Agence Technique de l'Information sur l'Hospitalisation
HAS Haute Autorité de Santé
ansm Agence nationale de sécurité du médicament et des produits de santé
Santé publique France
Institut National du Cancer
agence de la biomédecine
Agence du Numérique en Santé
cnsa Caisse nationale de solidarité pour l'autonomie

LES INDUSTRIELS

leem les entreprises du médicament
AFCROs Les entreprises de la recherche clinique
Le Dispositif Médical SNITEM Pour faire avancer la santé
SN syntec numérique
France Digitale

L'ÉQUIPE DU HEALTH DATA HUB

L'équipe du HDH est composée d'une cinquantaine de personnes motivées qui partagent la volonté d'accompagner les acteurs qui œuvrent à trouver les solutions de demain pour améliorer la santé des personnes.



55 **55 COLLABORATEURS AU 31/12**
DONT 32 FEMMES

40 RECRUTEMENTS (CDI ET CDD) **40**

2

RETOUR SUR LA FEUILLE DE ROUTE STRATÉGIQUE DU HEALTH DATA HUB

QUATRE ENJEUX INSCRITS DANS LA FEUILLE DE ROUTE STRATÉGIQUE 2020-2022

En janvier 2020, le HDH a publié sa feuille de route stratégique.
Elle s'articule autour de quatre grands enjeux stratégiques :

1. Décloisonner le patrimoine de données de santé et améliorer sa qualité.
2. Développer une offre de services performante pour les utilisateurs de données de santé.
3. Placer la France comme un leader de l'utilisation des données de santé dans l'écosystème européen et international.
4. Garantir la participation de la société civile à l'utilisation des données de santé.

ENJEU 1

Décloisonner le patrimoine de données de santé et améliorer sa qualité

L'une des missions du HDH est de mettre à disposition des porteurs de projets, **un catalogue de bases de données**.
Ce catalogue réunit une collection de bases pseudonymisées issues du SNDS que le HDH est habilité à mettre à disposition. Des échanges sont en cours avec une vingtaine de partenaires pressentis tels que Santé Publique France, l'Agence de Biomédecine, l'Inserm, l'Université de Bordeaux, Unicancer, etc.

ACTION 1.1 INSCRIRE AU CATALOGUE DES JEUX DE DONNÉES PRIORITAIRES

Le HDH est chargé par la loi de « réunir, organiser et mettre à disposition les données du système national des données de santé (SNDS) mentionné à l'article L. 1461-1 du code de la santé publique et de promouvoir l'innovation dans l'utilisation des données de santé ». Élargi à un grand nombre de sources de données de santé, notamment cliniques, le système national des données de santé (SNDS) pourra comprendre au-delà des données médico-administratives, des données de registres, de cohortes de recherche, d'entrepôts de données hospitalières, etc.

Assembler les bases de données dans un même lieu donnera au HDH la capacité de servir un grand nombre de projets innovants d'intérêt général, en particulier lorsque ces derniers auront besoin de croiser différentes sources.
Cette collection de bases composera le « catalogue », **construit de manière progressive et itérative**, en partenariat avec les responsables de données concernés. Le contenu du catalogue sera **décrit sur le site internet du HDH**, ainsi que l'ensemble des projets qui seront réalisés grâce à lui dans le répertoire complet des projets prévu par la loi.

Le catalogue ne vise pas à rassembler toutes les données au sein d'un fichier – ou même d'une plateforme – unique. Ce catalogue se remplira progressivement. Concrètement, le HDH fera remonter selon les besoins des projets, des copies de bases déjà existantes, dans lesquelles les informations personnelles directement identifiantes auront été supprimées, avec l'appui de la CNAM pour assurer la pseudonymisation dans le cas où les bases du catalogue pourront être appariées avec la base principale du SNDS.

Dans le contexte de la crise sanitaire, le catalogue a été préfiguré par l'identification et le rassemblement de premières bases relatives à la Covid-19. En effet, l'arrêté du 21 avril 2020, prolongé par l'arrêté du 10 juillet 2020, prescrivant les mesures d'organisation et de fonctionnement du système de santé nécessaires face à la Covid-19 et publié après avis de la CNIL, a autorisé le HDH à recevoir, stocker et mettre à disposition des données à caractère personnel concernant la santé afin de permettre le lancement de plusieurs projets en lien avec l'épidémie.

9 PROJETS COVID
ACCOMPAGNÉS EN 2020

ZOOM SUR LES BASES DE DONNÉES DE SANTÉ RELATIVES À L'ÉPIDÉMIE

Bases de Données	Usages déjà identifiés
OSCOUR (Santé Publique France) : Résumés individuels des passages aux urgences depuis le 1er janvier 2020 à date	› Suivi de l'évolution de la crise › Analyse de la variation des motifs de passage aux urgences pendant le confinement › Analyse du non recours aux soins pendant la crise de la Covid-19
SNDS Fast Track (CNAM, ATIH, Établissements de santé) : Données issues du PMSI MCO concernant les séjours Covid-19 ("PMSI Fast Track") appariées avec les données du SNDS sur la consommation de soins (période de janvier à mai 2020)	› Analyse des parcours de soin des patients hospitalisés pour Covid-19 › Identification des facteurs de risque de formes plus ou moins sévères de la Covid-19 › Compréhension de l'impact des formes sévères de la Covid-19 sur les patients et anticipation des soins en sortie d'hospitalisation / de réanimation
SIVIC (DGS) : Données de prise en charge hospitalière des patients soignés pour la Covid-19 (transfert en cours)	› Suivi de la propagation de l'épidémie › Analyse des inégalités sociales de santé grâce au croisement avec d'autres sources
SIDEP (DGS) : Résultats des tests de dépistage de la Covid-19 remontant depuis le 16 mai 2020 (à venir)	› Suivi de la propagation de l'épidémie et développement de modèles prédictifs › Étude des antécédents médicaux des personnes atteintes de la Covid-19 grâce au croisement avec d'autres bases (par ex. SNDS) liés à l'infection à la Covid-19 › Analyse de l'impact à long terme de la Covid-19 sur la santé grâce au croisement avec d'autres bases (par ex. SNDS)

Les partenariats avec de premiers responsables de données

Sont considérées comme « responsables de données » les entités publiques ou privées assurant aujourd'hui la collecte, l'agrégation, la mise en forme et la mise en qualité des données de santé dans le cadre de cohortes, registres ou toute autre base relevant du SNDS et ayant vocation à être inscrite au catalogue.

Le HDH s'engage à appuyer financièrement et humainement les responsables des données dans la conduite des différentes actions nécessaires au transfert et à la mise à disposition des données : extraction, déidentification, standardisation, documentation, création d'échantillons, mise en place des flux d'ingestion. Les modalités précises d'accompagnement seront fixées au cas par cas dans les conventions qui encadrent la collaboration avec chacun des responsables de données.

Les investissements humains, technologiques et financiers consentis dans le cadre du HDH permettent la mise en œuvre d'une procédure d'accès harmonisée, l'augmentation de la capacité à réaliser des extractions et des appariements, la mutualisation de la gestion du support, de la sensibilisation et de l'information des utilisateurs ainsi que des coûts de stockage, de calcul et de mise à disposition des données via une plateforme technologique à l'état de l'art. Certaines de ces activités peuvent être tarifées.

Aux côtés des responsables des données, le HDH s'engage à contribuer à la diffusion de standards internationaux et à une meilleure interopérabilité, afin de permettre des agrégations et chaînages de qualité. Il promeut enfin les initiatives visant à coordonner les acteurs locaux d'un territoire dans le but de créer de nouvelles bases.

En partageant leurs données par le biais du HDH, les responsables de données peuvent gagner en visibilité, et peuvent enrichir leurs données avec celles de la base principale du SNDS. Levier de partage, le HDH participe à la démarche science ouverte portée par le ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation.

En 2020, le HDH a initié des échanges avec une vingtaine de responsables de données partenaires potentiels. Ces échanges permettront d'aboutir à la signature de premières conventions de partenariat en 2021.

>20 PLUS DE 20 PARTENARIATS
EN COURS DE DISCUSSION AVEC
DES RESPONSABLES DE DONNÉES

RESPONSABLES DE DONNÉES
NOTRE OFFRE DE SERVICE

	OFFRE	PRINCIPE
	Convention type	Une convention type, prête à l'emploi, précise **les grands principes de partenariat** entre le HDH et le responsable des données pour la mise au catalogue de la base.
	Modalité technique de partage	Le HDH propose un transfert sécurisé à 3 niveaux : l'établissement d'un réseau **virtuel privé (VPN)**, la mise en place d'un flux de **données chiffrées (SFTP)** et le chiffrement des données transmises par le responsable des données.
	Open science	Le HDH a pour objectif de favoriser le **partage**, sous **licence ouverte**, des **connaissances** et **ressources** sur les bases de données. Il contribue à l'émergence de **standards de documentation** et de **normes de qualité**.
	Publicité de la base via le méta catalogue	Le HDH met en avant l'équipe du responsable de données, sa gouvernance, ses travaux et vise à améliorer sa visibilité auprès de la communauté scientifique. Il encourage les **collaborations scientifiques**.
	Tarification et appui humain et financier	Le HDH accompagne financièrement ses partenaires pour la mise au catalogue.

RESPONSABLES DE DONNÉES
NOTRE OFFRE DE SERVICE

	OFFRE	PRINCIPE
	Conformité réglementaire de la base	Le HDH accompagne le partenaire dans la vérification de la conformité de sa base.
	Information patient	Le HDH fournit aux responsables de données des mentions d'information concernant la transmission des données. En parallèle, il délivrera un second niveau d'information sur son site internet.
	Procédure du droit d'opposition	Le HDH renvoie vers le responsable de données qui est seul à (potentiellement) avoir l'identité du citoyen. La donnée peut être retirée à la source et le responsable de données envoie l'identifiant concerné au HDH pour qu'il mette à jour la matrice d'exclusion pour les futures extractions.
	Mise en relation avec le CESREES	Les responsables de données alimentant le catalogue du HDH sont associés à l'instruction du CESREES afin notamment de lui faire bénéficier de leurs expertises pour juger de la faisabilité, de la méthodologie, de la pertinence ou encore du caractère éthique d'une étude mobilisant les données dont ils sont à l'origine de la collecte. Le secrétariat de ce comité assuré par le HDH organise cette mise en relation.

ACTION 1.2 INSTALLER UNE PROCÉDURE NATIONALE AUTOMATISÉE POUR FACILITER LA SÉLECTION ET LA REMONTÉE DES DONNÉES AU CATALOGUE ET LEUR APPARIEMENT

Un certain nombre de bases préexistantes ont été identifiées pour être partagées au sein du catalogue du Health Data Hub. Néanmoins, l'articulation entre les établissements de santé, leurs entrepôts de données de santé et le Health Data Hub revêt un potentiel très important pour la constitution de bases de données multicentriques thématiques ou pour la remontée de données à la demande.

Partenariat avec Unicancer pour prototyper l'alimentation du HDH en données des centres de lutte contre le cancer (CLCC)

Le projet « UNIBASE » permettra un partage inédit de données de cancérologie et pourrait devenir rapidement l'une des briques venant enrichir le catalogue du HDH, ce qui permettrait de poursuivre et amplifier une stratégie ambitieuse en matière de données de santé en oncologie, avec cinq objectifs communs entre UNICANCER et le HDH :

- la centralisation des données et le croisement de sources multiples lorsque cela est nécessaire ;
- le partage des données entre les CLCC et d'autres structures hospitalières ou la ville ;
- le contrôle de la qualité des données ;
- le lien entre les données cliniques, d'imagerie et « omiques » ;
- l'administration de bases de données de grand volume, la sécurité des bases de données et la protection des données personnelles.

L'ambition du projet UNIBASE est de mettre en œuvre en trois ans la base de référence en cancérologie à partir de dix CLCC répartis sur le territoire. Cette ambition pourra être concrétisée par le biais de projets pilotes.

Un des premiers projets pilotes du HDH, en partenariat avec le GCS Hugo, devrait également permettre de préfigurer l'articulation entre un regroupement d'entrepôts de données de santé hospitaliers et le HDH à l'échelon national.

ACTION 1.3 CRÉER UNE « COHORTE SOCLE » RASSEMBLANT DES DONNÉES CLINIQUES ET MÉDICO-ADMINISTRATIVES FINEMENT LOCALISÉES SUR UN ÉCHANTILLON REPRÉSENTATIF DE LA POPULATION

Les données de la base principale du SNDS ont démontré depuis des années un potentiel très important pour la recherche comme pour éclairer les décisions publiques en santé, ces données médico-administratives ont tout intérêt à être complétées de données plus médicalisées, notamment en ville pour ouvrir des champs nouveaux d'exploitation des données au bénéfice de la santé de tous.

Accompagnement du cadrage du projet d'Observatoire de médecine de ville

Afin d'améliorer le croisement et le traitement des données de santé, le HDH a participé au cadrage d'un projet d'Observatoire de médecine de ville.

La médecine de ville désigne l'ensemble des professionnels de santé libéraux et des soins réalisés en dehors de l'hôpital ou des établissements d'hébergement médico-social.

L'objectif étant de pouvoir consolider une base de données nationale de médecine de ville permettant, à terme, de disposer d'une vision complète et plus médicalisée des parcours de soin et de participer notamment à la gestion d'épidémies type Covid-19.

ACTION 1.4 PROMOUVOIR LES PRODUCTEURS PARTICIPANT À LA DYNAMIQUE DE PARTAGE DES DONNÉES

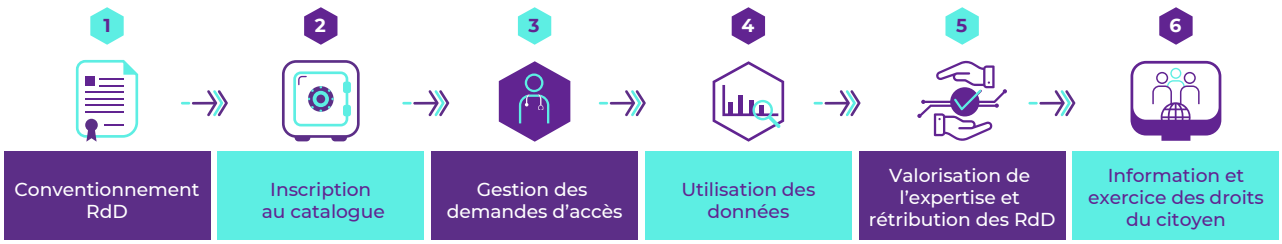
Le HDH vise à rassembler l'écosystème des données de santé : les organismes de recherche, les établissements et professionnels de santé, les associations de patients, les institutions, les industriels, les starts-up… Dans une démarche de co-construction, différents partenariats peuvent être développés, qui doivent prévoir la mise à disposition de jeux de données, de documentations, d'outils ou d'algorithmes pouvant servir à l'ensemble de la communauté.

La publication des engagements vis-à-vis des responsables des données

Fonctionner ensemble nécessite de définir des règles et des principes communs, qui vont régir les rôles et responsabilités de chacun. Établi en collaboration avec des représentants de l'écosystème comme l'AP-HP, l'INCA, l'INSERM, la CNAM, Santé Publique France, le CEA, etc, le HDH a publié ses engagements vis-à-vis de ses partenaires responsables des données. Ils prévoient, entre autres, de fournir :

- un appui à la collecte, la standardisation et la documentation des données ;
- un hébergement des données à l'état de l'art et hautement sécurisé ;
- un accompagnement dans la mise en conformité au RGPD ;
- la valorisation et la rétribution du travail et de l'expertise du responsable de données.

Ce document, disponible sur le site internet du HDH, explicite également le processus de mise au catalogue des données.



Instauration de règles de valorisation scientifique travaillées avec un panel de 16 experts

Le HDH s'engage notamment à augmenter la visibilité des bases de données grâce à un processus de citation fiable et pérenne, et à donner la possibilité aux responsables de données d'apparaître en tant que co-auteurs dans le cas d'une publication lorsque certaines conditions sont vérifiées. Ces règles ont été discutées avec un panel d'experts regroupant des chercheurs de différents domaines et des responsables de données.

En particulier, chaque porteur de projet devra citer la base de données dans les références grâce au DOI (identifiant numérique unique) mis en place par le HDH. Ce même porteur de projet devra citer la publication de référence de la base (quand celle-ci est communiquée par le responsable de données). Avec le DOI, la réutilisation des données peut être tracée et des indicateurs peuvent être construits pour qualifier la notoriété des bases et négocier des financements par exemple.

En cas de collaboration scientifique entre le responsable de données et le porteur de projet, les règles de signature de la publication seront exclusivement définies par les deux parties.

Réédition du colloque « données de santé et intelligence collective »

Fort du succès de l'édition de 2019, le HDH a souhaité rééditer son colloque annuel « Données de santé et intelligence collective » organisé avec la Délégation Ministérielle du numérique en Santé, porteuse de la feuille de route du numérique en santé et le directeur du programme « Amélioration des diagnostics médicaux par l'intelligence artificielle », Olivier Clatz. Ce rendez-vous a été l'occasion de mettre en valeur des cas d'usages en matière d'utilisation des données de santé en France et à l'échelle européenne.

Les intervenants ont montré les bénéfices concrets de l'utilisation des données de santé pour développer les usages visant à améliorer la santé.

Pour vivre ou revivre l'évènement :
https://health-data-hub.fr/page/colloque-2020-2

526

526 PARTICIPANTS AU COLLOQUE DONNÉES DE SANTÉ ET INTELLIGENCE COLLECTIVE

La politique de valorisation et de tarification

Le HDH s'engage à favoriser la mise en place de partenariats entre les responsables de données et les utilisateurs des données. Dans certains cas, l'accès direct aux données peut être tarifé à des utilisateurs privés.

Le HDH instruit les modalités de la mise en place de principes de tarification dans le cadre d'un pilote avec la cohorte I-share et d'une instruction juridique. Il s'agira en 2021 de conduire d'autres pilotes pour couvrir les différentes configurations possibles et de produire des principes directeurs. Évidemment, en cas de partenariat entre le responsable de données et l'utilisateur, ils fixent eux-mêmes les termes du partage de la valeur créée ensemble.

ENJEU 2

Renforcer l'usage des données par une offre de service attractive

Guichet unique des demandes d'accès aux données de santé, le HDH réceptionne l'ensemble des dossiers des projets mobilisant ces données nécessitant une autorisation de la CNIL. Il assure à ce titre le secrétariat du CESREES, comité d'expertise en charge de rendre un avis préalablement à l'autorisation CNIL. À noter que certains projets s'appuient sur des procédures simplifiées ou encore sur les méthodologies de référence.

ACTION 2.1 METTRE EN PLACE DES PROCÉDURES SIMPLIFIÉES ET UN PROCESSUS D'ACCÈS INDUSTRIALISÉ

Le HDH accompagne les porteurs de projet à chaque étape de leur demande d'accès avec l'appui indispensable des équipes des producteurs et notamment de la CNAM pour tous les projets utilisant la base principale du SNDS qui constituent la part la plus importante des projets. Interlocuteur privilégié, le secrétariat assure le lien entre les porteurs de projet, les responsables de données, le CESREES et la CNIL afin de permettre l'accès aux données dans les meilleurs délais. Tout au long de ces étapes, les chargés de mission, juristes et experts des données du HDH et des producteurs de données (comme les équipes chargées de la mise à disposition des données de la CNAM), répondent aux questions des porteurs de projet, les appuient dans leurs démarches, et ce, jusqu'à la mise à disposition des données.

À cet effet, un important travail de refonte de l'outil de gestion des procédures administratives pour accéder aux données a été mené avec les acteurs concernés (HDH, CESREES, CNAM, Santé Publique France, utilisateurs...) qui devrait aboutir au printemps 2021.

Le HDH est également force de proposition pour l'élaboration de procédures simplifiées d'accès aux données en étroite collaboration avec la CNIL. En 2020, cette action s'est traduite par la mise à jour de la procédure d'accès simplifié à l'échantillon généraliste des bénéficiaires (EGB).

Retour sur la mise en place du CESREES

En 2020, le CESREES (Comité Éthique et Scientifique pour les Recherches, les Études et les Évaluations dans le domaine de la Santé) s'est substitué au CEREES (Comité d'Expertise pour les Recherches, les Études et les Évaluations dans le domaine de la Santé).

Il se prononce sur :

- la finalité et la méthodologie de la recherche ;
- la nécessité du recours à des données de santé à caractère personnel ;
- la pertinence éthique ;
- la qualité scientifique du projet ;
- le cas échéant sur le caractère d'intérêt public que présente le projet.

Il a été créé à la suite de la parution du décret d'application de la loi informatique et libertés du 15 mai 2020, et de l'arrêté de nomination de ses membres du 9 juin 2020.

495 DOSSIERS EXPERTISÉS
PAR LE CEREES
PUIS CESREES EN 2020

Le HDH est secrétaire du CESREES. À ce titre, il assurera entre autres le lien avec les comités scientifiques des bases du catalogue afin de leur permettre de formuler toutes leurs remarques sur un dossier précis qui demanderait accès à ces données au sein du HDH.

L'utilisation et/ou la réutilisation des données

En France, la réglementation est précise concernant la réutilisation des données personnelles de santé à des fins d'étude et de recherche, elle reste néanmoins complexe à appréhender pour les porteurs de projets. C'est pourquoi, le HDH travaille avec la CNAM sur des kits pédagogiques pour faciliter la compréhension des démarches pour les porteurs de projet et assurer la cohérence entre l'accès aux données *via* le portail de la CNAM et *via* la plateforme du HDH..

ACTION 2.2 RENFORCER ET MAINTENIR À L'ÉTAT DE L'ART L'OFFRE TECHNOLOGIQUE DU HDH

Le HDH accompagne les chercheurs, entrepreneurs, associations de patients pour leur permettre de mener et d'accélérer le développement de leurs projets d'intérêt général, en mettant à leur disposition un « espace projet » sécurisé sur une plateforme technologique. Cet espace permet de recevoir, de stocker, de traiter et de visualiser les données de santé nécessaires aux projets sans les exporter.

3 PROJETS
SUR LA PLATEFORME

2 BASES
SUR LA PLATEFORME

Cet espace projet est isolé de l'extérieur : il est dédié uniquement au projet en question, sans transfert des données sur un poste local.
Fondé sur une infrastructure dite « élastique », il permet notamment d'offrir des ressources de calcul permettant de développer les usages les plus innovants.
Le contexte de sécurité de la plateforme technologique du Health Data Hub a été défini dans le cadre d'une démarche conduite en collaboration avec l'ANSSI (Agence nationale de la sécurité des systèmes d'information) et le Haut fonctionnaire de défense et de sécurité du ministère des Solidarités et de la Santé. Dans le cadre du suivi de l'homologation, la sécurité de la plateforme technologique est testée et revue fréquemment par le biais d'audits techniques menés par des prestataires indépendants qualifiés par l'ANSSI. L'ANSSI et la CNIL participent en outre à l'homologation de chaque nouvelle version de la plateforme technique. Deux homologations ont été réalisées en 2020.
Afin de répondre aux forts enjeux de sécurité, le HDH a également mis en place un certain nombre d'actions en lien avec ses partenaires : la mise en place d'une procédure de sécurisation des transferts de données par les responsables de données validée par l'ANSSI et un guide de bonnes pratiques sur la pseudonymisation partagé avec la CNIL.
Le HDH a également fait l'objet d'un audit de la DINUM (Direction interministérielle du numérique) au titre de l'article 3 du décret n° 2019-1088 relatif au système d'information et de communication de l'État. Cet audit a souligné le haut niveau de sécurité de la plateforme technologique.
À l'issue de phases de tests, l'année 2020 a été marquée par l'ouverture des services de la plateforme aux projets pilotes et aux projets relatifs à l'épidémie de la Covid-19. Fin 2021 ou début 2022, les services pourront être ouverts plus largement.

L'actualisation des écarts de fonctionnalités entre les services de la plateforme et les services proposés par les industriels français

Entièrement administrée par des agents du HDH, aux rôles et responsabilités bien définis, la plateforme technologique permettant l'analyse innovante des données mobilise des composants industriels choisis pour leur performance, leur sécurité et leur élasticité. L'offre logicielle à l'état de l'art permettra le développement des codes et algorithmes sans contrainte (R, python, spark.). Une « marketplace » est également prévue à terme pour permettre aux éditeurs de partager les outils pertinents avec la communauté et ainsi favoriser la mutualisation.
Les enjeux de la plateforme technologique sont respectivement :

1) **La sécurité :** assurer la protection des données de santé hautement sensibles stockées sur la plateforme technologique.
2) **Les fonctionnalités et la performance :** répondre aux besoins des porteurs de projets d'intérêt public, faisant appel aux approches issues de la science des données.
3) **Les coûts et délais :** mettre en œuvre rapidement une solution pour répondre au risque de perte de chance des citoyens français en lien avec la stratégie « intelligence artificielle nationale » et dans un contexte de compétition internationale.

Les choix technologiques ont été faits début 2019 pour répondre à ces enjeux. Certains partenaires internationaux ont été retenus en l'absence de solution souveraine mais la plateforme est développée selon une logique d'infrastructure programmable, à l'aide de langages indépendants de la solution d'hébergement retenue, permettant de la redéployer sur une autre solution du même niveau de maturité le moment venu (réversibilité).

Néanmoins, pour pouvoir migrer la plateforme, il faut qu'une solution française ou européenne équivalente existe. Un travail sur les écarts de fonctionnalités entre Microsoft et OVH a été mené fin 2019 et a été partagé avec la DINUM et la DGE. Il a été actualisé avec l'appui de la DINUM mi-juin 2020. Sur les 14 services essentiels sur lesquels s'est concentré l'exercice de juin, l'écart est toujours considéré comme très important par les experts du domaine.
Le HDH souhaite prendre toute sa part dans la réflexion sur le cloud souverain et continue, fin 2020 et début 2021, à travailler avec la DINUM sur l'actualisation de son étude de réversibilité.

ACTION 2.3 DÉVELOPPER UNE OFFRE DE SERVICE PERMETTANT UN USAGE PERFORMANT DES DONNÉES

Si l'ambition du HDH est bien de permettre un accès plus rapide aux données de santé, il n'est pas toujours garanti que les données obtenues permettent d'aboutir à des résultats probants, fiables, valorisables. Afin de tirer le meilleur parti des données, un travail considérable est nécessaire pour comprendre leur organisation, la façon dont elles sont collectées, les biais et les limites qu'elles présentent.

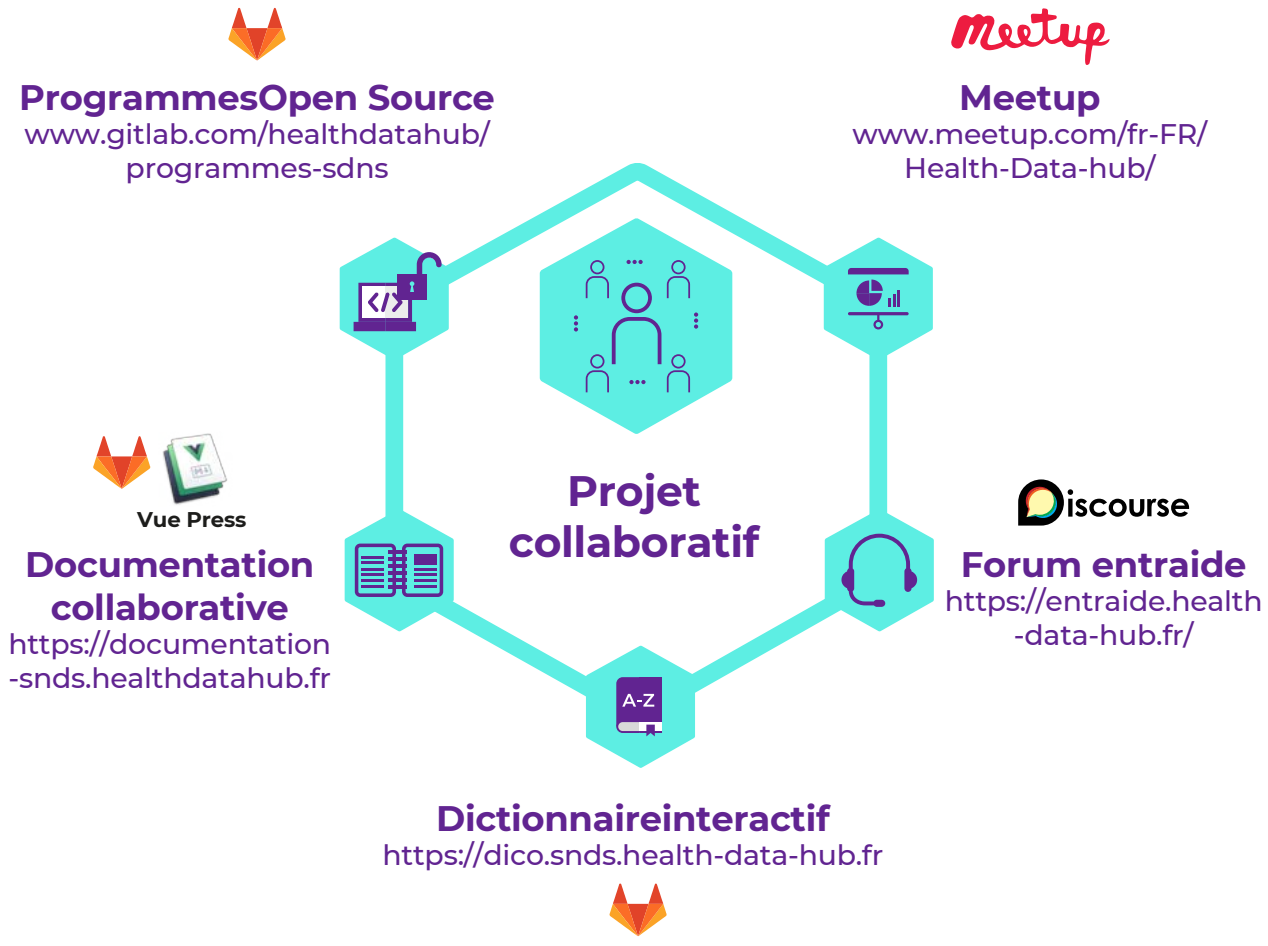
Ainsi, pour aider les porteurs de projet et utilisateurs à explorer le patrimoine de données disponible, le HDH a, par exemple, contribué à l'élaboration d'un catalogue de métadonnées. Ces travaux ont notamment été partagés dans le cadre d'un groupe de travail de l'Institut de Recherche en Santé Publique (IReSP) sur les études en santé et auprès d'un panel d'experts spécifique. C'est pourquoi, le HDH travaille avec la CNAM sur des kits pédagogiques pour faciliter la compréhension des démarches pour les porteurs de projet et assurer la cohérence entre l'accès aux données *via* le portail de la CNAM et via la plateforme du HDH.
Le HDH a également participé aux réflexions

d'un groupe de travail piloté par l'INRIA et CEA, aux côtés de l'INSERM, le Cancéropôle Lyon, 3DS, Roche, Sanofi, France Digitale Health, l'Institut Pasteur et le CNRS. Ce groupe de travail avait pour objectif de produire des recommandations en matière de logiciels à proposer de manière prioritaire dans la plateforme technologique du HDH. Ont été ainsi recensés des logiciels pour l'imagerie médicale, l'oncologie, le traitement automatique du langage, la génomique et bioinformatique, le machine learning, et les biostatistiques. Ces travaux seront poursuivis et prolongés courant 2021 notamment pour aborder les problématiques de certification des logiciels et outils de pseudonymisation (NLP).

La démarche open source

Le HDH propose à tous, en accès libre, des informations, de la documentation, et des formations. En lien avec ses partenaires, il met à disposition une série d'outils visant à faciliter la prise en main des données. De nombreuses actions ont été menées autour de la documentation et de l'utilisation de la base principale du SNDS avec le concours de la CNAM. En effet, comprendre les données – leur organisation, mais aussi leurs biais, potentiels, limites – représente un effort considérable pour chaque porteur de projet, c'est la raison pour laquelle le HDH propose de mutualiser les connaissances : c'est l'objet de la démarche open source.

LES OUTILS OUVERTS MIS EN PLACE



ACTION 2.4 DIFFUSER L'OFFRE DE SERVICE AUPRÈS DE L'ENSEMBLE DES ACTEURS : ORGANISMES DE RECHERCHE, INDUSTRIELS, PARTENARIATS PUBLICS PRIVÉS, ETC.

Dans le cadre de sa mission d'animation de l'écosystème, le HDH est moteur pour l'organisation d'événements fédérateurs, au rayonnement international et favorisant l'échange et les rencontres d'acteurs divers (entrepreneurs, académiques, personnel médical) pour promouvoir les partenariats et l'innovation. À titre d'exemple, un premier appel à projets a été lancé en 2019 pour identifier les premières initiatives qui bénéficient aujourd'hui de l'accompagnement du HDH pour leur réalisation. Trois objectifs principaux motivaient cet appel à projets : soutenir les usages innovants de la donnée de santé, participer à la construction et au développement du HDH, et démontrer par des résultats rapides les bénéfices de l'utilisation des données.

Zoom sur le second appel à projets du HDH « L'intelligence artificielle pour une expérience améliorée du système de santé »

Fort de ce premier succès, le HDH a lancé en 2020 un second appel à projets co-organisé avec Le Grand Défi « amélioration des diagnostics médicaux par l'IA » et Bpifrance. Cet appel à projets a eu pour vocation de sélectionner une deuxième vague de projets innovants consacrés cette fois-ci à l'amélioration de l'expérience du système de santé par l'IA. Les candidats se sont positionnés sur un des deux axes suivants : **« développement d'application à base d'IA à destination des professionnels ou des patients »** ou **« développement de modèles de populations pour la prévention ou la thérapie fondés sur des techniques innovantes d'analyse de données »**.

La procédure était ouverte à tous : organismes de recherche, établissements et professionnels de santé, associations de patients, industriels, starts-up… dès lors qu'ils (1) pouvaient faire état d'un usage innovant de la donnée de santé, (2) répondaient à une finalité d'intérêt public et (3) ne rentraient pas dans le champ des finalités interdites par le code de la santé publique.

Conformément à la volonté de co-construction du HDH, les porteurs de projet devaient prévoir la mise à disposition de jeux de données, de documentation, d'outils ou d'algorithmes créés pouvant servir à l'ensemble de la communauté.

9 PARTENAIRES À LA DÉMARCHE
OPEN SOURCE ET 2 JEUX DE DONNÉES SYNTHÉTIQUES POUR LE CATALOGUE

138 initiatives se sont portées candidates à cet appel à projets qui s'est clôturé le 1er juin. Ce nombre, ainsi que la diversité des porteurs de projets (structures publiques/laboratoires de recherche (48 %), PME (43 %), grandes entreprises (5 %) et associations (4 %)) reflètent les attentes fortes de l'écosystème vis-à-vis du HDH.

Chaque projet retenu dans le cadre de cet appel à projets fait l'objet : d'un soutien financier provenant du Grand Défi et opéré par Bpifrance d'un montant maximum de 300 000 € sous forme de subventions pour des dépenses de recherche et développement et pour une durée de projet comprise entre 12 mois et 24 mois ; et d'un accompagnement opérationnel de la part du HDH portant par exemple sur les démarches visant à accéder aux données, l'appui à la collecte et l'organisation des données, la mise à disposition de moyens de calcul et de stockage, la mise en relation avec d'autres acteurs de l'écosystème.

Ainsi, fin 2020, le HDH accompagne 27 projets pilotes et 9 projets Covid.
Certains projets ont déjà franchi des étapes significatives :
› 15 projets ont reçu un avis favorable du CESREES (6 projets pilotes et les 9 projets Covid) ;
› 8 projets sont autorisés par la CNIL (dont 3 projets Covid).
En 2020 également, 10 conventions de mise à disposition ont été signées dans l'objectif d'accompagner les porteurs de projets tant sur le plan humain que technique.

138

138 CANDIDATS À L'APPEL À PROJETS
600 PERSONNES ONT SUIVI L'ÉVÉNEMENT D'ANNONCE DES LAURÉATS DE L'APPEL À PROJETS EN DIRECT

Photo issue de l'annonce des lauréats de l'AAP du 9 juillet

Les projets lauréats de l'appel à projets l'IA pour une expérience améliorée du système de santé

Projet	Porteur	Objectif
AIDA-Ultrasound	**E-Scopics**	Développer un outil d'aide à la prise en main d'un dispositif d'échographie visant à démocratiser son utilisation.
APRIORICS	**CHU de Toulouse**	Développer un outil de description morphologique d'images microscopiques de cancer du sein pour mieux appréhender la variabilité de ces cancers.
APSoReN	**CHU de Toulouse**	Fournir un outil permettant de repérer les patients victimes de traumatisme crânien qui ne bénéficieraient pas de l'ensemble des soins qui leur seraient nécessaires.
DAICAP	**AP-HP**	Développer un algorithme qui permettra l'obtention d'un compte rendu d'IRM standardisé avec une prédiction de l'agressivité tumorale.
DeepMap	**DAMAE Medical**	Développer un dispositif permettant de fiabiliser le diagnostic du dermatologue, de visualiser l'étendue de la pathologie, et de fournir une analyse quantitative et compréhensible de la lésion.
INNNERVE	**Quantmetry**	Développer un logiciel de diagnostic précoce des neuropathies des petites fibres.
PRECSION-PREDICT	**Institut Curie**	Mieux comprendre l'hétérogénéité de réponse aux thérapies ciblées de patients atteints d'un cancer broncho-pulmonaire.
SEDAAR	**Hôpital Fondation Rotschild**	Développer des algorithmes afin de développer un service d'aide à la décision et à l'interprétation à destination des orthoptistes et ophtalmologistes.
TAMIS	**TELECOM SudParis**	Développer un algorithme de caractérisation automatique fine sur frottis sanguin des principales cellules sanguines, hématies, leucocytes et plaquettes.
TARPON	**Université de Bordeaux**	Construire un outil de surveillance et de recherche sur le traumatisme (accidents, violences, suicides) basé sur le traitement automatisé de visites annuelles aux urgences.

Le Health Data Hub dans sa mission d'animation de l'écosystème

De nombreux groupes de travail se sont tenus en 2020 malgré les conditions sanitaires.
À titre d'exemple, un groupe de travail organisé en février 2020 par la Fédération Hospitalière de France et la conférence des DG de CHU a permis de riches échanges entre le HDH et les établissements hospitaliers. Santé publique France a également organisé un temps entre le HDH et les porteurs de registre afin qu'ils expriment leurs attentes vis à vis du HDH. Des échanges ont également été menés avec les membres du collège des industriels du HDH pour recueillir leurs besoins.

Par ailleurs, le HDH a mis en place et consolidé un réseau de chercheurs affiliés. Ce réseau s'adresse aux chercheurs français (et internationaux) exprimant un intérêt pour les problématiques et questions de recherche liées à la santé. Il vise à renforcer le lien entre le HDH et la communauté de la recherche, en diffusant des informations et opportunités en lien avec la donnée de santé et en proposant des mises en relation autour des thématiques d'intérêt.

D'une manière générale, le HDH soutient les initiatives locales qui visent à vulgariser et promouvoir la réutilisation des données de santé. Il intervient dans de nombreux événements. À titre d'exemple, le HDH est partenaire du Data Lab Normandie qui a lancé en début d'année 2021 un appel à projets « Données et technologies appliquées à la santé ».

En 2020, le HDH s'est également positionné comme membre fondateur de PariSanté Campus, aux côtés de l'Inserm, l'Université PSL, l'Inria et l'Agence du Numérique en Santé. Ce projet, porté par le ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation, et le ministère des Solidarités et de la Santé, est destiné à conforter la place de notre pays comme acteur majeur de la recherche et du développement en santé numérique. Situé à terme dans un lieu emblématique, l'ancien hôpital d'instruction des armées du Val de Grâce, PariSanté Campus s'inscrit dans la stratégie globale du gouvernement visant à renforcer les moyens dédiés à la recherche et à l'innovation en santé numérique, en réunissant les acteurs du numérique et de la santé en un lieu unique.
Ainsi, PariSanté Campus vise à stimuler la production et l'exploitation des données de santé pour la recherche et l'innovation ; soutenir la formation d'excellence autour de la santé numérique et la création des viviers d'expertises ; développer des applications à fort impact économique, créatrices d'entreprises et d'emplois ; améliorer les stratégies de prise en charge médicale et l'efficience du système de soin dans l'intérêt permanent du patient.

ENJEU 3

Positionner la France comme un leader dans l'usage des données de santé

Le HDH est un service à destination de l'écosystème, des acteurs à l'origine de la collecte de données, des utilisateurs et de la société civile. Pour atteindre son plein potentiel, le HDH doit être au rendez-vous des attentes de l'écosystème et participer à sa structuration au niveau européen.

ACTION 3.1 ALIMENTER LA COMMUNAUTÉ SCIENTIFIQUE INTERNATIONALE

Le HDH organise avec ses partenaires des événements fédérateurs. Le but d'un Data Challenge est, pour ses participants, de répondre en un temps imparti, et à partir de données spécifiques, à une question posée. Le Data Challenge s'adresse à des data scientists (chercheurs, industriels, étudiants...) de diverses nationalités. Les compétiteurs développent leurs algorithmes à partir d'un jeu de données anonymisées. Dans une ultime étape, l'accessibilité des données (anonymisées) et des algorithmes issus du Data Challenge est encouragée afin de permettre à la recherche de poursuivre sa progression dans l'intérêt de tous.

2020, le premier Data Challenge organisé par le Health Data Hub

Pour favoriser le développement d'outils d'IA d'aide au diagnostic, le premier Data Challenge, en partenariat avec la Société Française de Pathologie (SFP) et le Grand Défi « Améliorer les diagnostics médicaux à l'aide de l'intelligence artificielle » a été organisé en 2020 dans le but d'améliorer la détection des lésions précancéreuses et cancéreuses pouvant survenir sur le col de l'utérus.

Plus de 500 compétiteurs ont pu développer leurs algorithmes à partir d'un jeu de données de plusieurs milliers de lames virtuelles annotées.

En effet, un jeu de données anonymes de près de 5 000 lames d'histopathologie (biopsies et conisations) a été constitué pour cette occasion. Le challenge s'est déroulé sur la plateforme de la société DrivenData, spécialisée dans les actions Data4Good. Les compétiteurs ont pu tester leurs algorithmes et évaluer en temps réel leur classement sur un leaderboard.

La SFP, grâce à son appel à manifestation d'intérêt, a permis de mobiliser une vingtaine de centres anatomopathologiques français, publics et privés, qui se sont portés volontaires pour inclure des lames. Ces derniers ont anonymisé chaque lame, avant de les numériser.

5 000

5 000 LAMES NUMÉRISÉES /
547 COMPÉTITEURS /
DES SCORES ALLANT JUSQU'À 0,94

Le HDH a apporté une aide organisationnelle et logistique dans la collecte de ces données et a également assuré le stockage des lames sur une infrastructure hautement sécurisée.
Les résultats des algorithmes soumis sont impressionnants. Douze équipes ont atteint un score de plus de 0,9 avec un vainqueur culminant à plus de 0,93. La compétition a été une réelle opportunité pour les candidats de se faire repérer par des acteurs majeurs du domaine. Elle a permis aux trois meilleures équipes, récompensées notamment lors du Carrefour Pathologie 2020, de recevoir des prix pouvant aller jusqu'à 12 000 €. Les médecins et statisticiens organisateurs de la compétition travaillent désormais à l'interprétation clinique des résultats de la compétition.
Pour poursuivre sur cette lancée, le HDH a déjà initié les échanges avec diverses sociétés savantes afin de renouveler l'expérience en 2021.

ACTION 3.2 STRUCTURER ET NORMALISER LES DONNÉES DE SANTÉ POUR FACILITER LEUR USAGE DANS LE CADRE DE COLLABORATIONS INTERNATIONALES

Si l'ambition du HDH est bien de permettre un accès plus rapide aux données de santé, il n'est pas toujours garanti que les données obtenues permettent d'aboutir à des résultats probants, fiables, valorisables. Afin de tirer le meilleur parti des données, un travail considérable est nécessaire pour comprendre leur organisation, la façon dont elles sont collectées, les biais et les limites qu'elles présentent. Cet investissement ne saurait être entrepris par chacun des porteurs de projet. Dans ce contexte, une des missions du HDH consiste à promouvoir le partage des connaissances et des outils entre les projets.

Le HDH est un des lauréats de l'appel à projet organisé par EHDEN, le réseau européen des données de santé et de preuves en matière de santé, financé par l'Innovative Medical Initiative. À travers ce nouveau partenariat, le HDH transforme des données de la base principale du SNDS dans un format qui a vocation à devenir une référence internationale : le format OMOP-CDM (Observational medical outcomes partnership - Common Data Model). La conversion des bases de données de santé dans le format OMOP-CDM renforcera leur interopérabilité et facilitera le partage d'outils, de méthodes, d'études et de programmes. Le HDH a initié sur la mise au format des données du SNDS sur la Covid-19, 62 tables sont mappées (PMSI et SNIIRAM) à ce jour.

ACTION 3.3 CONTRIBUER À LA STRUCTURATION DU PAYSAGE EUROPÉEN AUTOUR DE L'USAGE DES DONNÉES DE SANTÉ

Le Health Data Hub nommé autorité compétente pour la France

Depuis 2019, la Commission européenne a désigné l'espace européen des données de santé comme l'une des **priorités de la politique de santé européenne**. Il importe en effet de faciliter l'accès (technique, réglementaire, etc.) aux différents types de données disponibles au sein des systèmes de santé variés des États membres, pour développer la recherche et améliorer ainsi la santé des populations de l'Union européenne.
C'est dans ce contexte qu'elle a lancé une action conjointe pour un espace européen des données de santé (TEHDaS).

Ce mécanisme, financé en partie par la Commission européenne, permet aux 26 États Membres ayant candidaté conjointement, de travailler ensemble pendant deux ans à la production de recommandations.

En France, ce sont l'ANS, les Hospices Civils de Lyon, les unités de service US 14 (CépiDC) et US 10 (Orphanet) de l'Inserm, l'Université Toulouse III et l'Université Aix-Marseille qui contribuent aux work package aux côtés du HDH en tant qu'entités affiliées.

La France est particulièrement présente sur le work package portant sur l'engagement citoyen qu'elle copilote. Dans trois pays pilotes (Royaume-Uni, Belgique et France) seront menées une étude sociologique et une consultation citoyenne sur la question du partage des données de santé.

Ce projet d'envergure vise à renforcer l'utilisation secondaire des données de santé en Europe et à mettre en place un modèle commun de gouvernance des données de santé. Cet espace aura le potentiel de devenir un puissant moteur d'innovation et de fédération des acteurs européen de la santé.

Le HDH a également été nommé « national node » dans la poursuite de l'action conjointe InfAct et apporte son soutien à l'infrastructure européenne fédérée de structures de données de santé « Dipoh » (A Distributed Infrastructure on Population Health).

8 % DU BUDGET DE L'ACTION CONJOINTE EUROPÉENNE
« TOWARDS A EUROPEAN HEALTH DATA SPACE » POUR LA FRANCE

ACTION 3.4 PROMOUVOIR LE HEALTH DATA HUB À L'INTERNATIONAL

L'une des missions du HDH est de contribuer au rayonnement européen, voire international de la France dans l'utilisation des données de santé, notamment en valorisant l'excellence des acteurs français. Pour mener à bien cette ambition, le HDH a déjà réalisé un certain nombre d'actions en 2020 notamment par la définition d'un plan de communication internationale et la réalisation d'une cartographie des interlocuteurs dans le cadre d'un parangonnage. Ce parangonnage permet de positionner la France dans la dynamique internationale et d'identifier des partenaires potentiels.

Le HDH a également contribué à la mise en place d'un partenariat entre la Finlande et la France pour explorer la mise en œuvre de projets transfrontaliers mobilisant les données de santé.

Organisation de l'école d'hiver - AI4Health

L'année 2020 a été marquée par l'organisation de l'école d'hiver AI4Health, initialement prévue durant l'été 2020. Cette école virtuelle dédiée à l'intelligence artificielle appliquée à la santé a été co-organisée avec les Instituts Interdisciplinaires en intelligence artificielle (Prairie, 3IA Côte d'Azur et MIAI Grenoble). Elle a réuni **438** participants aux profils variés (60 % d'étudiants, 20 % d'académiques et 20 % de professionnels du secteur privé) répartis sur 30 pays. Pendant trois jours, des chercheurs de renommée internationale, ont présenté leurs travaux, puis des travaux pratiques se sont tenus pendant deux jours.
Une table ronde sur les applications concrètes de l'Intelligence Artificielle en santé pour le bénéfice des patients a complété le programme.

438

438 PARTICIPANTS À L'ÉCOLE D'HIVER/ AI4HEALTH

ENJEU 4

Garantir la participation de la société civile et promouvoir les usages des données de santé

Pour que le HDH atteigne son plein potentiel, il est essentiel que son objectif soit compris par l'ensemble des parties prenantes et en particulier la société civile. Dans ce contexte, le HDH a mis en place une direction citoyenne pour écouter la société civile, produire des contenus pertinents au regard de leurs attentes mais également concevoir des partenariats.

ACTION 4.1 ÉCOUTER LA SOCIÉTÉ CIVILE ET COMPRENDRE SES ATTENTES POUR ÉTABLIR UNE RELATION DE CONFIANCE

La direction citoyenne mène en priorité des actions visant à recueillir les besoins et attentes des citoyens. Pour ce faire, elle met en place des groupes de travail associant des représentants d'usagers et des citoyens. Ces groupes ont permis d'élaborer des premières études sur la perception des citoyens vis à vis du HDH et du partage des données. Ils peuvent également réagir à des propositions de contenus qui sont ainsi calibrés pour répondre à leurs attentes. C'est d'ailleurs ce qui a amené le HDH à s'engager dans une démarche Facile A Lire et À Comprendre, dite FALC. Cette dernière vise à traduire dans un langage accessible à tous les contenus tels que les engagements du HDH vis à vis des citoyens.
De manière similaire, le HDH participe aux travaux de vulgarisation de la Délégation du Numérique en Santé (DNS), notamment dans le cadre du groupe de travail autour de l'éthique du numérique en santé.

Toutes ces démarches seront poursuivies en 2021, avec l'organisation entre autres d'une étude sociologique complète et de grande envergure dans le cadre de l'action conjointe sur le thème du partage des données.

Focus sur la conception d'un questionnaire à destination des étudiants en pharmacie sur le rapport à la donnée de santé, et les besoins en termes d'information / de formation

Au-delà des usagers du système de santé, c'est également la perception des professionnels de santé qu'il s'agit de recueillir pour réunir les conditions les plus favorables au partage des données de santé. Un partenariat avec l'ANEPF (Association Nationale des Etudiants en Pharmacie de France) est initié pour co-construire un questionnaire diffusé auprès de ces étudiants sur leur perception de la donnée, leur niveau de connaissance et leurs besoins (en termes d'information / de formation). Les réponses au questionnaire permettront d'orienter les outils de communication à construire.

>1 950 **+DE 1 950 INSCRITS**
À L'ÉTUDE VIVRECOVID-19

Ce travail s'effectue au sein d'un groupe de travail réunissant des étudiants en pharmacie, et co-animé par la direction citoyenne du HDH. Enfin, ce questionnaire pourra permettre de guider la DNS dans son plan de formation au numérique destiné aux professionnels de santé.

ACTION 4.2 EXPLIQUER ET DIFFUSER LES ENGAGEMENTS DU HEALTH DATA HUB AUTOUR DU RESPECT DES PROCÉDURES, DU CONTRÔLE DES USAGES, DE LA MAÎTRISE DES DONNÉES, DE L'EXPLICITATION ET L'ACCOMPAGNEMENT DANS L'EXERCICE DES DROITS

Le HDH a publié ses engagements envers les citoyens, qui ont été établis en concertation avec des représentants de la société civile et des associations de patients dans le cadre d'un groupe de travail présidé par Gérard Raymond, président de France Assos Santé.

Ces engagements ont été conçus autour de quatre axes prioritaires : l'intérêt général des projets d'utilisation des données, la protection des données, le respect des droits individuels et la transparence. Deux vidéos pédagogiques accompagnent cette publication. Une version FALC de cette note d'engagement a également été réalisée.

S'agissant des droits des citoyens, une note pédagogique a également été co-construite avec le groupe de travail.

Sur le plan opérationnel, la direction juridique et le délégué à la protection des données ont travaillé sur les modalités pratiques d'exercice des droits pour les premières bases partagées au sein du HDH.
Pour garantir la transparence des projets réutilisant les données de santé, le répertoire public originellement maintenu par l'INDS a été fortement refondu. Cette liste de projets sera prochainement complétée par les éléments de résultats.

1 GROUPE DE TRAVAIL
POUR CO-CONSTRUIRE LES VIDÉOS PÉDAGOGIQUES

1 FOCUS GROUPE
POUR TESTER LA NOTE D'ENGAGEMENT

15 ENTRETIENS
SEMI-DIRIGÉS/INDIVIDUELS

2 FOCUS GROUPES
POUR TESTER LES MESSAGES DES VIDÉOS ET DE L'INFOGRAPHIE

ACTION 4.3 METTRE EN PLACE UNE POLITIQUE DE COMMUNICATION ET DE DIFFUSION DE LA CULTURE DE LA DONNÉE DE SANTÉ

Au-delà des engagements pris pour garantir la protection des données personnelles de santé, le HDH doit contribuer à vulgariser les enjeux relatifs à cette réutilisation.

Pour ce faire, le HDH conduit des actions de formation et a notamment initié une formation d'« introduction au SNDS » dont le premier volet a été réalisé avec des associations de patients en partenariat avec l'AFCROS. Des formations plus avancées sont également en cours d'élaboration et un partenariat avec l'Education Nationale a été discuté.

Le HDH met en œuvre une politique de communication qui va permettre, sur son nouveau site institutionnel, de valoriser les finalités des projets accompagnés et les résultats concrets. Le conseil scientifique dont le HDH s'est doté permettra également d'être plus ambitieux encore en matière de stratégie de diffusion de la connaissance.

Enfin, le HDH met en œuvre des partenariats avec les associations de patients pour leur donner l'opportunité d'être un véritable acteur du dispositif.

C'est ainsi qu'il a accompagné France Assos Santé dans la réalisation de l'étude Vivre Covid19, dont l'objectif est de recueillir le ressenti des usagers du système de santé durant l'épidémie. Plus de 4000 personnes, chaque mois, font remonter leur vécu via un questionnaire personnalisé. De cette manière, France Assos Santé va pouvoir suivre pendant deux ans leur ressenti, et disposer de connaissances robustes pour faire de l'information ciblée et orienter son plaidoyer.

Le HDH accompagne France Assos Santé de bout en bout dans ce projet et il participe ainsi à la mise en œuvre de l'étude VivreCovid19, à la coordination du projet avec les associations membres, à la valorisation du design et des résultats notamment lors de congrès, et à garantir la protection éthique des répondants. Il contribue, avec le soutien de la CRO de cette étude, à l'animation d'ateliers, l'un d'eux portant sur la conception et la mise en oeuvre d'une doctrine éthique et d'open source, prochainement en ligne, et actant différents principes dont la co-construction du questionnaire, le libre accès au protocole d'étude, et les modalités d'accès aux résultats et aux données de l'étude pour les associations membres et les autres acteurs intéressés.

1 611 ABONNÉS
À LA NEWSLETTER
MENSUELLE DU HDH

26 ACTUALITÉS PUBLIÉES
SUR LE SITE, AVEC UNE NOUVELLE
ACTUALITÉ EN MOYENNE PAR SEMAINE

ACTION 4.4 DÉVELOPPER LA MISE À DISPOSITION DE DONNÉES STRICTEMENT ANONYMES EN LIBRE ACCÈS

La démocratisation de la réutilisation des données de santé est facilitée par l'accès totalement ouvert à certaines données.
Pour cela, et compte tenu de la sensibilité des données, il faut produire des données parfaitement anonymisées ou bien des données fictives.
Dans cette optique, le HDH prévoit de rendre disponibles avec son catalogue de métadonnées des données de synthèse pour accompagner chacune des bases de données par ailleurs disponibles de manière pseudonymisée dans la plateforme hautement sécurisée. Le HDH a noué un partenariat avec l'équipe REPERES pour la mise à disposition de données de synthèse représentatives de la base principale du SNDS et créé un générateur permettant de produire des données synthétiques de manière automatique.
Plus facilement utilisables car en libre accès, ces données de synthèse facilitent l'appropriation des données réelles par les utilisateurs.
Le HDH fait également la promotion de l'ouverture des données réelles lorsqu'elles peuvent être parfaitement anonymisées. Ainsi, dans la foulée du Data Challenge organisé avec la SFP, une stratégie d'ouverture des données mobilisées a été élaborée.

+5 000 LAMES DE BIOPSIES DU COL DE L'UTÉRUS
NUMÉRISÉES DANS LE CADRE DU DATA CHALLENGE AVEC LA SFP

3

BILAN DE L'ACCÈS AUX DONNÉES

Les données de santé doivent être partagées au plus grand nombre pour améliorer la qualité des soins, mais cela doit se faire dans le respect des droits des personnes concernées. En effet, ces données, même pseudonymisées, restent des données personnelles à moins qu'elles ne soient parfaitement anonymisées. Dans ce dernier cas, leur utilisation est libre et aucune démarche n'est à entreprendre. À défaut, l'accès à ces données est réglementé.

Accès aux données et perspectives

UN ACCÈS RÈGLEMENTÉ

Dans le cas général, il est nécessaire d'obtenir une autorisation auprès de la CNIL pour traiter des données personnelles de santé. En amont de sa décision, deux institutions sont mobilisées dans le cadre de la procédure :

› Le HDH, point d'entrée unique des projets mobilisant des données de santé. Près de 495 projets y ont été déposés en 2020.

› Le CESREES, en charge de rendre un avis notamment sur la finalité, la pertinence et le sérieux des projets.

Afin de faciliter l'accès aux données et de permettre aux porteurs de projet de réaliser leur traitement dans les meilleurs délais, le temps imparti à chacune de ces institutions est encadré : le HDH dispose ainsi de 7 jours pour transmettre le projet complet au CESREES pour expertise. Ce dernier a un mois pour rendre un avis éclairé. Une fois saisie à la demande du porteur de projet, la CNIL dispose de 2 mois éventuellement reconductibles pour émettre une autorisation. En cas de non respect des délais, les avis ou décisions sont réputés tacitement favorables ou acquis.

3 PROCÉDURES SIMPLIFIÉES

1 MÉTHODOLOGIES DE RÉFÉRENCE
TELLES QUE MENTIONNÉES DANS L'ARTICLE 73 DE LA LIL

2 DÉCISIONS UNIQUES
TELLES QUE MENTIONNÉES DANS L'ARTICLE 66 DE LA LIL

3 ÉCHANTILLONS ET DONNÉES AGRÉGÉES
TELS QUE MENTIONNÉS AU II DE L'ARTICLE 66 DE LA LIL

Porteur de projets demandant l'accès aux données de santé

Ces méthodologies de références s'appliquent à des types d'études spécifiques et sont disponibles sur le site de la CNIL.

Pour obtenir une décision unique auprès de la CNIL, la procédure standard est requise. Une fois que la CNIL a délivré une décision unique, les traitements sont effectués sans examen préalable par le HDH ou le CESREES

Le HDH peut approuver sous certaines conditions la mise à disposition d'échantillons et de données agrégées. Cette procédure peut être utilisée uniquement pour les données, finalités et utilisation entrant dans ce cadre

GUICHET DU HDH
Réception et approbation

HDH
Enregistrement des études dans le registre du HDH

DONNÉES DE SANTÉ | **EGB**

SCHÉMATISATION DU CIRCUIT DES DEMANDES D'ACCÈS AUX DONNÉES DE SANTÉ

APPARIEMENT DE DONNÉES DE SANTÉ AVEC DES DONNÉES NOUVELLES AU SENS DE LA LOI JARDÉ

ACCÈS STANDARD À DES DONNÉES EXISTANTES

PROCÉDURES SIMPLIFIÉES

Porteur de projets demandant l'accès à des données de santé

1. Méthodologies de références
2. Décisions uniques de la CNIL
pour plusieurs traitements à finalités, catégories de destinataires et catégories de données identiques.
3. Approbation par le HDH
de demande d'accès à des échantillons ou données agrégées.

HDH

HDH

CPP
Émet un avis sur les conditions de validité de la recherche (délai de 45 jours max)

CESREES
Émet un avis sur la méthodologie de l'étude (délai de 1 mois max)

CNIL
Autorise le traitement des données

UTILISATION DES DONNÉES DE SANTÉ

Il existe des procédures d'accès dites simplifiées, dans le cadre desquelles l'autorisation de la CNIL n'est pas requise, c'est notamment le cas si le traitement envisagé est conforme à une méthodologie de référence (dite MR). Afin de promouvoir l'accès aux données et favoriser leur usage, **le HDH est force de proposition pour le déploiement de ce type de procédures en étroite collaboration avec la CNIL.**
Enfin, le HDH peut promouvoir l'accès facilité à des jeux de données agrégées ou des échantillons, dans des conditions préalablement autorisées par la CNIL. La procédure permettant au HDH sous 15 jours d'approuver des projets souhaitant accéder aux données de l'Échantillon Généraliste des Bénéficiaires du SNDS, adaptée pour tenir compte de la modification du cadre juridique a ainsi été validée par la Commission en juillet 2020.

L'accès aux données de santé en 2020 en quelques chiffres

Nombre total de projets déposés au guichet	**506**
Nombre de projets déposés au guichet en vue d'une autorisation CNIL	**495**
• dont transmis au CESREES	357
• dont réorientés (MR ou CPP)	97
• dont non complets	41
Nombre de demandes d'accès à l'EGB simplifié	**11**
Nombre de projets relevant d'une MR	**1 221**
• MR004	1 125
• MR005	35
• MR006	61

UNE OFFRE DE SERVICES À DESTINATION DES PORTEURS DE PROJET

Favoriser l'utilisation des données de santé nécessite de documenter ce qu'elles contiennent, les possibilités qu'elles offrent, de rendre lisible les procédures d'accès et de favoriser le partage des connaissances. À cette fin, une collection d'outils opérationnels et pédagogiques à destination des porteurs de projets pour les accompagner à toutes les étapes de la mise en œuvre de leur projet, ont été élaborés dès la préfiguration du HDH à la DREES, en utilisant également la documentation existante fournie par les producteurs. Il s'agit par exemple de :

› Une documentation collaborative pour permettre au porteur de projet de se familiariser avec les données de la base principale du SNDS.
› Un forum permettant d'obtenir des réponses rapides d'experts concernant les modalités d'accès aux données ou le contenu du SNDS.
› La mise à disposition de jeux de données de synthèse pour comprendre le format et la structure du SNDS et se former à sa manipulation avec des exercices dédiés notamment dans le cadre d'un partenariat avec l'équipe REPERES.
› L'organisation d'événements fédérant la communauté des utilisateurs afin de partager des expériences et des connaissances. Près d'une vingtaine de sessions a été organisée sur l'année 2020.
› Avec la mise en place du CESREES, un kit pédagogique, pour aider le porteur dans la préparation de son dossier de demande d'autorisation auprès de la CNIL a également été élaboré.

USAGE DES DONNÉES DE SANTÉ ET TRANSPARENCE

Le répertoire public des projets, prévu par la loi, a par ailleurs vu le jour. Près de 4500 projets mobilisant des données personnelles de santé y étaient répertoriés fin 2020. Ce répertoire rend compte des travaux menés à partir de ces données.
En complément, un tableau de bord a été mis en ligne : il propose des visualisations permettant de caractériser les projets soumis et de restituer leur état d'avancement.

UNE AMÉLIORATION EN CONTINUE DE L'OFFRE DE SERVICE EN 2021

Afin de favoriser la montée en qualité des protocoles, les kits seront enrichis dès le premier trimestre 2021 :

› Les modèles de protocole et résumé intègreront des éléments d'aide au remplissage parmi lesquels : des définitions, des points de vigilance, des conseils, des exemples et un protocole complet fictif pour permettre aux porteurs de projets d'identifier concrètement les éléments majeurs attendus.

› Une check list à destination des porteurs de projet sera ajoutée aux éléments constitutifs du dossier pour que ces derniers s'assurent que les items majeurs attendus par le CESREES et la CNIL sont bien présents.

Par ailleurs, pour optimiser les délais de mise à disposition des données de la base principale du SNDS une fois l'autorisation de la CNIL délivrée et pour bien identifier et apprécier le périmètre des données nécessaires, un guide pédagogique pour les expressions de besoin réalisé en collaboration avec la CNAM sera mis à disposition dès le dépôt du projet. Cette aide permettra aux porteurs de projet de préparer au mieux les informations utiles pour traduire leur besoin en extraction de données.

L'année 2021 verra aussi l'arrivée d'un nouvel outil de soumission des projets, l'actuel répondant insuffisamment aux besoins des acteurs intervenants dans la procédure d'accès. À titre d'exemple, cet outil permettra aux porteurs de projet de suivre l'état d'avancement de leur projet, de disposer de statistiques de suivi, de transmettre leur résultat d'étude ou encore de valider leur mise à disposition auprès du grand public. Les membres du CESREES, quant à eux, pourront accéder plus aisément aux pièces nécessaires à leur expertise, échanger entre eux si besoin ; ils pourront par ailleurs être informés du devenir des études qu'ils ont expertisées. Enfin, les responsables de données pourront *via* cet outil disposer d'éléments d'appréciation pour estimer le recours à leur données et les gestionnaires de plateformes pourront anticiper les travaux relatifs à l'expression de besoins au plus tôt afin de réduire les délais de mise à disposition.
Dans le cadre des obligations de transparence associées à l'utilisation des données de santé, cet outil sera interfacé avec le répertoire public des projets afin de restituer les résultats transmis par les porteurs de projet.

HEALTH
DATA HUB

L'accès aisé et unifié, transparent et sécurisé,
aux données de santé pour améliorer la qualité
des soins et l'accompagnement des patients.

9 rue Georges Pitard, 75015 Paris

E-mail : hdh@health-data-hub.fr

www.health-data-hub.fr

Édition : Health Data Hub
Conception et réalisation : Parimage
Avril 2021