



HEALTH  
DATA HUB

**Bilan 2024  
et programme de  
travail 2025**

# Bilan

## du programme de travail 2024



# Des résultats concrets et des partenariats stratégiques



## Projets de recherche et d'innovation

Le **Health Data Hub (HDH)** accompagne **158 projets**, dont 54 % sont portés par des hôpitaux et 28 % par des industriels. 27 projets ont déjà produit des résultats intermédiaires ou finaux.



Parmi les projets utilisant la plateforme technologique du HDH, 90% des projets mobilisent la base principale du SNDS ; 75% nécessitent des appariements de données ou plusieurs bases.

## Partenariat avec l'Assurance Maladie

Le partenariat prévoit une autonomie accrue du HDH dans les ciblage et appariements, avec pour conséquence une **réduction des délais** de mise à disposition des données de la base principale.



Le partenariat stratégique avec la CNAM porte ses fruits avec plus de 300 extractions livrées et une trentaine d'appariements réalisées en autonomie par le HDH en 2024.

## Outils innovants

Le HDH a procédé au développement du **"concentrateur"** qui permettra d'améliorer significativement la réalisation des appariements directs avec le NIR, et rendra enfin possible ceux avec reconstitution.



Le projet SMPP de l'Université Paris-Cité s'appuiera sur les services du concentrateur pour réaliser un chaînage entre la cohorte Progène (portant sur les maladies prostatiques) et la base principale du SNDS, par un appariement direct incluant la reconstitution du NIR.

# Un catalogue du SNDS en expansion pour un meilleur partage des données de santé

Le **catalogue du SNDS** est une collection de bases de données non figée et construite itérativement. Les **10 premières bases ont déjà été inscrites au catalogue** via l'arrêté du 12 mai 2022, auxquelles se sont **ajoutées 12 bases identifiées comme pertinentes** pour l'enrichir.

## Mise en oeuvre opérationnelle du catalogue

Dans l'attente de la publication du référentiel de sécurité (mai 2024) et de la stabilisation des derniers éléments de doctrine de partage (février 2025), **des projets pilotes ont été lancés pour préfigurer la mise au catalogue des 10 premières bases.**

## Élargissement du catalogue

**12 nouvelles bases de données ont été ajoutées (mars 2025)**, dont l'entrepôt de médecine de ville **P4DP** qui sera disponible et croisé avec la base principale du SNDS, pour 8 premiers projets dès le premier semestre 2025.

## Ingestion de données non issues du catalogue

**Une trentaine de producteurs sont déjà connectés au HDH** pour plus d'une centaine de processus d'ingestion. La **mise en place de canaux de transfert peut être extrêmement rapide** selon les partenaires (deux jours).



Le HDH poursuit ses travaux pour **croiser les données de santé avec des données contextuelles**, comme en témoigne la relance de l'appel à projets début 2024 sur l'intégration des données de santé-environnement, porté conjointement avec le Green Data for Health.

# Le développement d'une expertise en standardisation des données et la mise à disposition de services adaptés



## Catalogue de métadonnées

Le HDH a développé un catalogue de métadonnées, bientôt accessible en *open source* et en marque blanche, pour offrir aux dépositaires de données la possibilité de fournir aux utilisateurs un **accès complet aux informations disponibles dans les bases de données**.

Le HDH déploie ce catalogue, en suivant le modèle européen "**HealthDCAT-AP**", et poursuit son enrichissement avec les données du **catalogue du SNDS** et celles du projet européen de Direct Grant "**French\_HealthData@EU**".



## Standardisation OMOP

Le modèle "OMOP-CDM" est une norme internationale de standardisation pour l'exploitation des données à des fins d'étude, permettant une meilleure interopérabilité des données en France, ainsi qu'au niveau européen ou international.

Le HDH a développé son expertise sur ce standard en travaillant notamment sur la transformation au format OMOP de :

- **3 millions d'enregistrements de patients pour l'EMA et le réseau DARWIN EU®**,
- **de dizaine de variables cliniques composant EMC2**, le premier entrepôt multicentrique clinique autorisé par la CNIL.



**Le projet QUANTUM** : le HDH est partenaire du projet QUANTUM lancé en février 2024 pour 3 ans, qui réunit **30 acteurs de 14 pays européens** et vise à développer un **label de qualité pour l'utilisation secondaire des données de santé** dans l'UE dans le cadre de l'Espace européen des données de santé (EHDS).

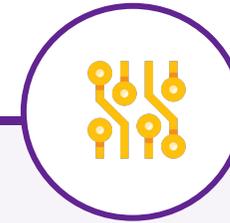
# Une offre technologique à l'état de l'art satisfaisant les plus hautes exigences de sécurité



## Une infrastructure de pointe

Le HDH s'appuie sur une **plateforme technologique cloud** permettant aux projets en données de santé d'avoir accès à **un stockage élastique et à de hautes capacités de calculs**.

L'offre technologique disponible a été renforcée en 2024 avec notamment la facilitation de **l'import de paquets et modèles pré-entraînés** sur la plateforme, l'augmentation de la **capacité de traitement de grands volumes de données**, la mise à disposition d'un **terminal** ou encore la connexion possible à **un système fils**, notamment pour le projet EMC2.



## Actualité de la plateforme

Le HDH continue sa démarche de **migration de sa plateforme technologique vers un environnement "souverain"**. Dans cette perspective, il intensifie ses échanges avec les différents acteurs du cloud français et a **contribué au lancement de l'appel à projets France 2030 opéré pour le compte de l'État par Bpifrance**, visant à renforcer l'offre de services cloud française et européenne.

À la suite à l'arrêté du 6 mai 2024 relatif au **référentiel de sécurité applicable au SNDS**, le HDH a actualisé **l'homologation de sa plateforme technologique** conformément aux nouvelles exigences et a collaboré avec le ministère de la Santé et de l'Accès aux soins pour élaborer un **guide destiné aux acteurs concernés** dont la publication est attendue pour la fin d'année.



**Les projets accompagnés par le HDH peuvent être très intensifs en technologie.** La plateforme stocke désormais près de 1 pétaoctet de données, avec certains projets nécessitant jusqu'à 300 téraoctets de données, plus de 500 gigaoctets de RAM, un processeur graphique (GPU), ou jusqu'à 512 unités de calcul (CPU) pour leur exécution.

# Un travail approfondi concernant la structuration et le contenu de l'accompagnement des utilisateurs



## Outils à destination des porteurs de projets

### Pour faciliter les démarches réglementaires

Le HDH met à disposition plusieurs kits pédagogiques pour accompagner les porteurs de projets dans l'appropriation des démarches réglementaires. En 2024, le HDH a publié les **kits relatifs aux MR-004, MR-007 et MR-008**, les deux dernières étant conçues pour faciliter l'accès aux données de la base principale du SNDS. Ces outils, élaborés en collaboration avec la CNIL et la CNAM, ont été présentés au cours de 3 webinaires ouverts. Des éléments du kit ont été traduits en anglais.

### Pour former les acteurs

**16 sessions de formations** ont été programmées tout au long de l'année, suivies par près de **1100 personnes**.

Plusieurs **MOOC sur l'architecture et les composantes de la base principale du SNDS** sont également accessibles en ligne sur le site internet du HDH. Ils font l'objet de 100 consultations par mois en moyenne.

### Pour partager la connaissance et les outils

Le HDH a mis en ligne **une Bibliothèque Ouverte d'Algorithmes en Santé (BOAS)** qui met à disposition librement des **codes et algorithmes dédiés aux données de santé**. 23 contributions sont accessibles en *open-source*. Un **guide utilisateur** a également été mis en place afin de faciliter la contribution et le référencement dans cette bibliothèque.

En 2024, le HDH a également publié **8 nouvelles fiches thématiques sur la base principale du SNDS** et mis à jour son **dictionnaire interactif** sur les métadonnées 2022 du SNDS.



**Les outils, actualités et activités du HDH sont régulièrement diffusés**, notamment lors de la **Journée de l'open science en santé** du 30 avril 2024, qui a réuni la communauté autour des travaux en *open source* et *open data*.

# Un hub national au sein de l'EHDS

## Une communauté de HDABs en construction

Le HDH interagit avec l'ensemble des organismes responsables de l'accès aux données de santé (HDABs) européens au sein de la "**Community of Practice**".

## La suite de la préfiguration de l'EHDS avec TEDHAS II

Le HDH contribue à l'**accompagnement de la mise en œuvre du règlement EHDS** aux côtés de 11 pays européens.

## La construction de services nationaux

Le HDH a progressé sur la mise en place de services d'intérêt aux niveaux national et européen.

- Environnement de traitement sécurisé,
- **Catalogue de métadonnées**,
- Registre public de projets,
- POC d'**infrastructure technique** HealthData@EU Pilot,
- **Concentrateur** pour l'exercice des droits et appariements,
- **Alignements OMOP**.

## L'intégration des acteurs français au réseau européen

Les partenaires français, en participant à des consortia pilotés par le HDH et sélectionnés lors de 4 appels à projets européens, se sont vus allouer près de **5 millions d'euros**.

Le HDH a par ailleurs initié l'**intégration de 19 bases FR au réseau de l'EHDS** via la mise en œuvre du Direct Grant "French\_HealthData@EU" et a facilité la **mobilisation de plus de 25 partenaires FR** pour des projets et cas d'usage européens.

- **5 cas d'usage** dans le cadre du projet-pilote à l'EHDS,
- **Mobilisation de données de vie réelle** pour l'Agence européenne du médicament (réseau DARWIN EU®),
- **Facilitation de la mise sur le marché de solutions IA** via le projet TEF-Health,
- **Mise en place d'un système de surveillance hospitalier en France** via l'initiative Orchidée menée par Santé publique France.

# Un positionnement en soutien des établissements de santé dans la construction d'une politique publique d'ampleur (1/2)



**Les établissements de santé doivent occuper une position centrale** en tant que producteurs de données, en pouvant gérer leurs propres données localement tout en alimentant des entrepôts nationaux. **Depuis 2022, plusieurs initiatives d'ampleur ont été mises en place par la puissance publique** et auxquelles le Health Data Hub est étroitement associé, pour soutenir le déploiement d'un véritable réseau. Des réflexions ont par ailleurs été entamées pour assurer la pérennisation financière cruciale du dispositif.



## Un appel à projets national pour financer les entrepôts de données de santé

- 75 millions d'euros,
- 16 projets d'entrepôts de données de santé lauréats,
- 47 établissements engagés, dont 29 CHU.

## dont les lauréats sont accompagnés de bout-en-bout par le HDH

- Kit pédagogique dédié aux entrepôts de données de santé, sessions collectives,
- Panorama avec l'ANAP des logiciels
- Élaboration avec le comité stratégique des données de santé d'éléments de doctrine (contrats type, standard, grilles de redevance, gouvernance, etc.)

## renforcé par l'appel à projets DAtAE visant à compléter le financement de l'infrastructure des entrepôts de données de santé

- Appel à projets de la DGOS opéré par le HDH,
- 15 projets pour 3,5 millions d'euros sélectionnés en 2024,
- Appel à projets renouvelé pour accompagner les prochaines étapes de développement des entrepôts de données de santé.



# Un soutien actif pour la recherche et l'innovation en santé...

## Travaux avec les organismes de recherche



### Soutien au partage des données collectées par les industriels

Lancement d'un appel à manifestation d'intérêt "**Partage de données industrielles**", en collaboration avec **Roche** et **MEDIPATH**.

Le HDH contribue par ailleurs activement aux travaux menés par le CSF sur les enjeux relatifs aux données de santé..

Le HDH poursuit ses travaux avec les **grands organismes de recherche (CNRS, Inria, Inserm)** en organisant des événements pour faire connaître ses services, en apportant son soutien au montage des **PEPR**, et en participant à de grands consortiums européens comme **Invents**.



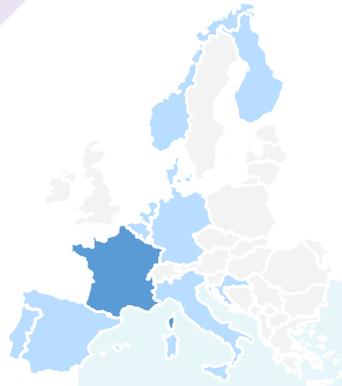
### Développement de partenariats internationaux

**Proposition d'événement commun avec la Commission européenne** pour le sommet pour l'action sur l'IA.

Organisation d'une **délégation au Royaume-Uni** pour rencontrer l'écosystème britannique des données de santé.



# ... et un investissement sur les enjeux d'intelligence artificielle



## Projet SHAIPED

Lauréat d'un appel à projets de la Commission européenne, le consortium piloté par le HDH regroupe 30 partenaires issus de 11 pays, et bénéficie d'un financement européen de 4 millions d'euros.

**Objectif :** développer des solutions d'IA en santé en s'appuyant sur les organismes responsables de l'accès aux données de santé (HDABs) et créer des synergies entre l'AI Act et l'EHDS.



## Projet PARTAGES

Lauréat d'un appel à projets France 2030, le consortium piloté par le HDH regroupe 20 hôpitaux et 10 équipes de recherche en IA, avec des partenaires industriels (ex. Mistral AI), et bénéficie d'un financement Bpifrance de 6,9 millions d'euros.

**Objectif :** utiliser l'IA générative pour analyser les comptes-rendus médicaux.



**Programme Data Challenge :** le HDH est chargé du programme Data Challenge dans le cadre de France 2030 qui a permis de mettre en oeuvre 5 challenges et réuni 1800 participants dans le monde depuis le début de l'initiative.

# Une feuille de route nationale pour favoriser le partage des données de santé



La mise en place d'une feuille de route nationale des données de santé centrée sur l'usage secondaire des données est cruciale. Elle doit **garantir que les efforts déployés jusqu'ici par le Comité stratégique des données de santé soient consolidés.**



En 3 ans, le **Comité stratégique des données de santé a réalisé plusieurs avancées :**



**Définition du catalogue du SNDS.** Les 10 bases en font partie et 12 autres ont été ajoutées.



**Standardisation d'un ensemble commun de données** issues des entrepôts de données de santé hospitaliers.



**Gouvernance de la mise à disposition** des données par les entrepôts de données de santé hospitaliers (contrats, redevances, CSE).

**4 axes d'actions prévus par la stratégie :**



**Axe 1 :** Favoriser la transparence et la confiance des citoyens.



**Axe 2 :** Constituer des bases données d'intérêt réutilisables.



**Axe 3 :** Réunir les conditions nécessaires à la mise à disposition des données de santé.



**Axe 4 :** Faciliter et simplifier l'utilisation des données.

# Les travaux autour de l'ouverture à la société



## Favoriser la transparence et l'exercice des droits

Le HDH a enrichi le répertoire public des projets par la **mise en ligne des résultats des projets terminés**. Il a également livré **un formulaire centralisé** pour l'exercice des droits des utilisateurs et mis en ligne **un portail patient**.



## Promouvoir l'Open Science

Le HDH a relancé **l'appel à projets "Data Challenges en santé"** en lien avec la DNS, le SGPI, Bpifrance.

Il a participé à **l'ouverture de 4 bases de données sur data.gouv** issues des *data challenges* et à l'ouverture de deux bases de données synthétiques.



## Transmettre une information accessible, et adaptée à chacun

Le HDH a développé ses initiatives pour **sensibiliser, informer, former les citoyens à la seconde vie des données** :

- **Webinaires** et supports de formation en libre accès sur une variété de thèmes,
- Organisation d'un **débat lycéen au CESE**, avec l'espace éthique d'Île-de-France, la DNS, le CNPEN, la DGOS et la CNIL.

Le HDH suit l'évolution des connaissances citoyennes via son **baromètre** en partenariat avec France Assos Santé, la CNIL, la DNS et PariSanté Campus et dont la seconde édition a été diffusée en 2024.



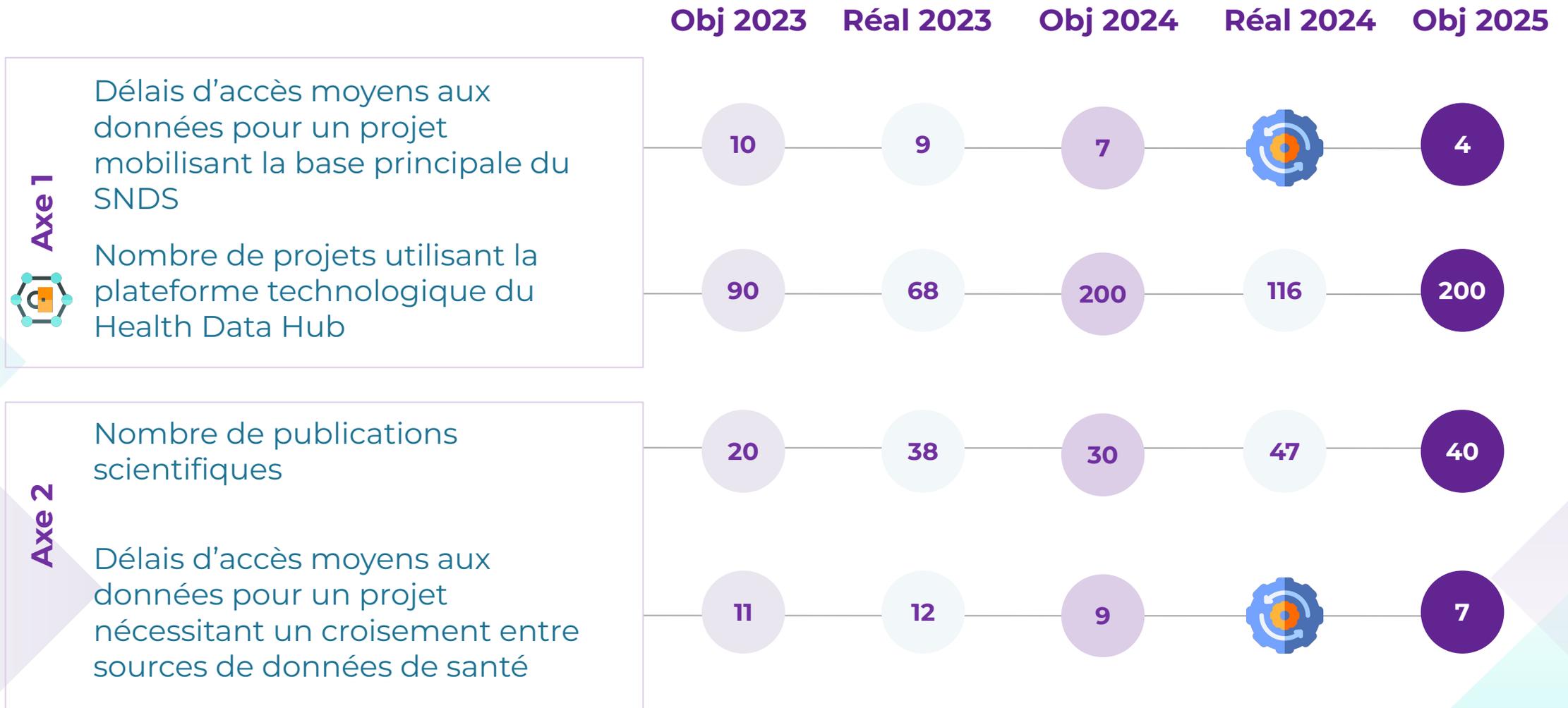
## Inclure les associations dans la recherche

**Le HDH et France Assos Santé ont constitué un Guichet de demande d'accès à des indicateurs du SNDS** sur sa base principale et ont déjà démarré l'accompagnement de trois associations.

**Le HDH s'est associé à la Ligue contre le cancer** pour le financement des plusieurs projets de recherche.

**Le HDH a lancé l'accompagnement de l'association Corasso** pour son projet Errance Cancer tête et cou, avec chaînage aux données de la base principale du SNDS.

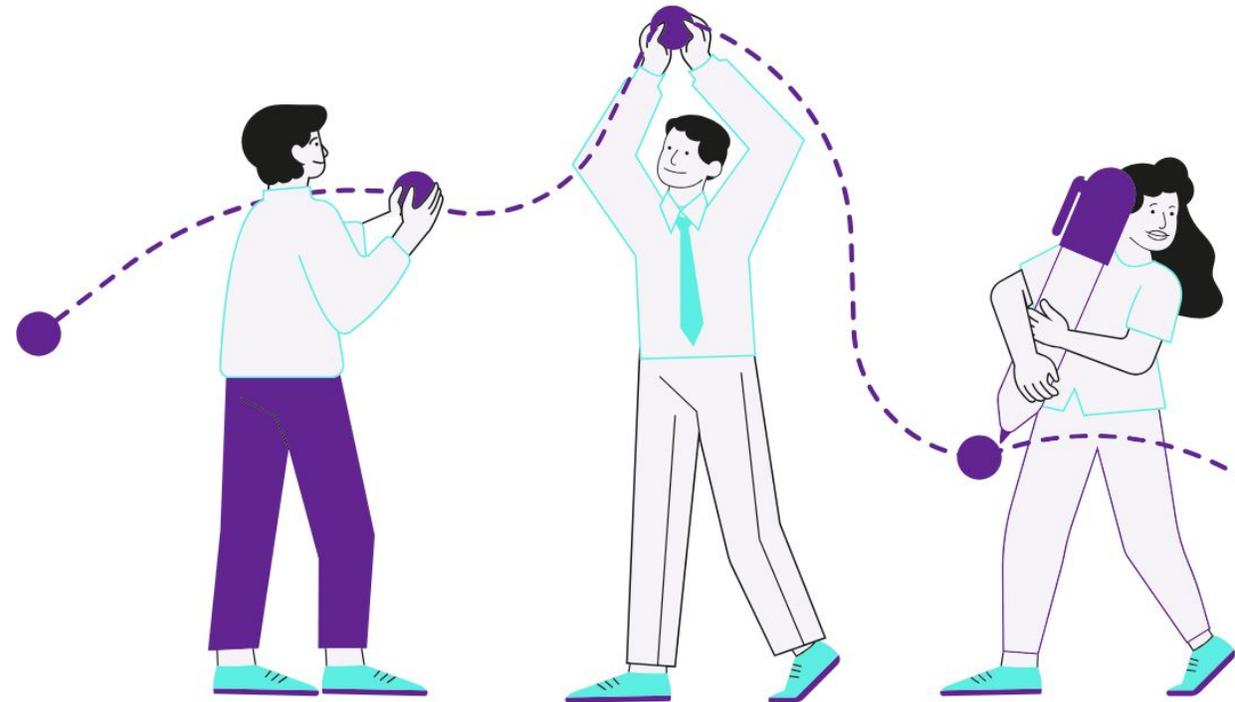
# Etat d'avancement des indicateurs de pilotage de la feuille de route pluriannuelle 2023 - 2025 (1/2)



# Etat d'avancement des indicateurs de pilotage de la feuille de route pluriannuelle 2023 - 2025 (1/2)

|   | Obj 2023 | Réal 2023 | Obj 2024 | Réal 2024 | Obj 2025 |
|---|----------|-----------|----------|-----------|----------|
| <b>Axe 3</b><br> Nombre de partenariats avec des infrastructures locales ou mutualisées au niveau national<br><br>Nombre de partenariats avec des infrastructures locales ou mutualisées au niveau européen    | 17       | 17        | 25       | 30        | 30       |
|   | 8        | 14        | 10       | 18        | 12       |
| <b>Axe 4</b><br> Nombre d'acteurs de la société civile participants / ayant été impliqués à des actions avec le Health Data Hub<br><br>Nombre de visites sur la page du répertoire public du Health Data Hub | 18       | 23        | 25       | 28        | 30       |
|   | 30 000   | 93 002    | 35 000   | 91 406    | 40 000   |

# Grandes orientations du programme de travail 2025



# Orientations du programme annuel de travail 2025

1



Réduire les délais  
d'accès aux  
données de santé  
et multiplier les  
projets impactants

2



Mettre à disposition  
les données de la  
base principale,  
l'enrichir et faciliter  
sa réutilisation

3



Renforcer les  
connexions du HDH  
aux acteurs de l'  
écosystème

4



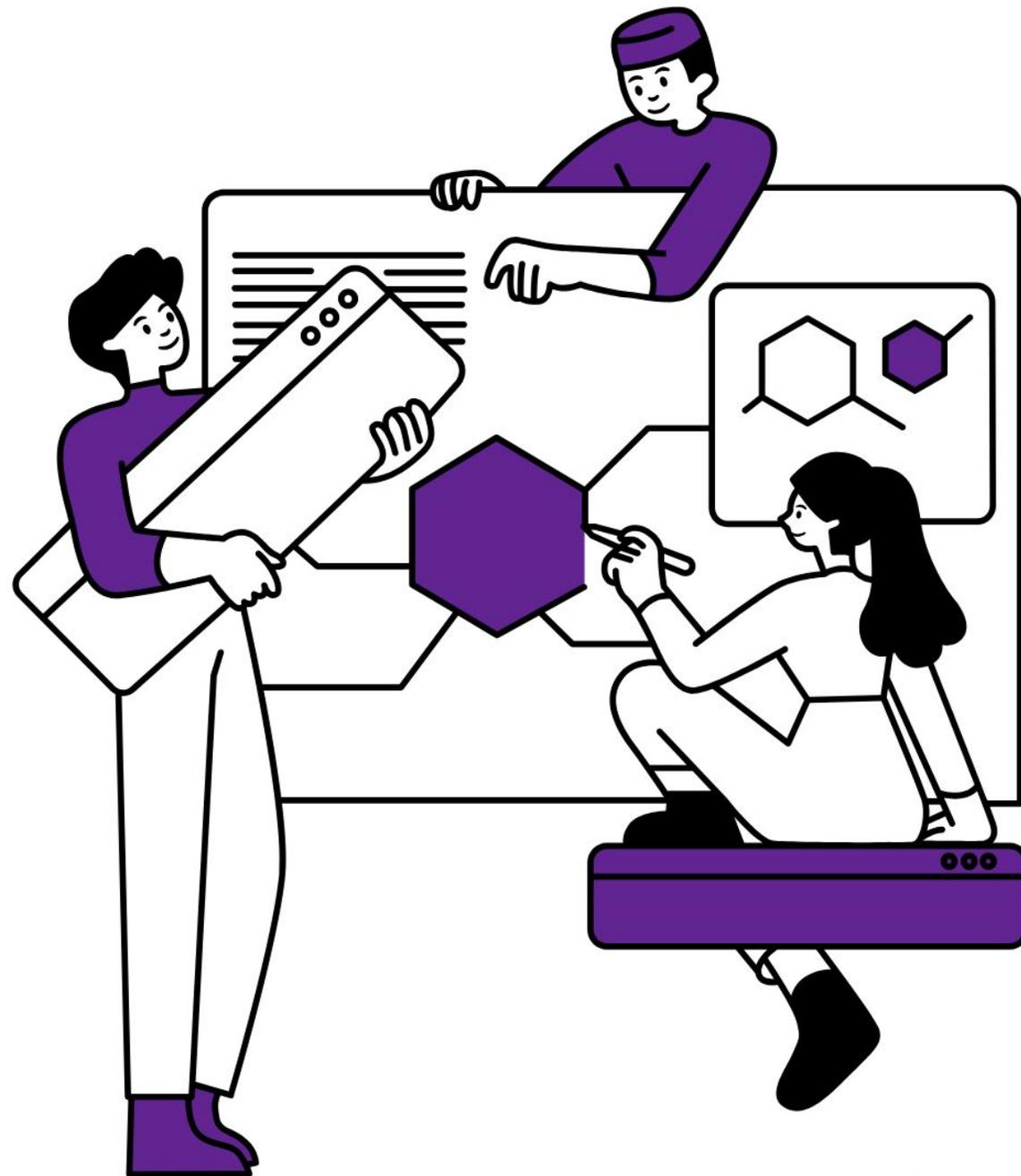
Etre à l'écoute de la  
société civile et  
co-construire une  
culture de la  
donnée de santé



Ces quatre axes stratégiques sont ceux de la feuille de route pluriannuelle 2023-2025, présentée en assemblée générale et en conseil d'administration le 9 juin 2022 et approuvée à l'unanimité par ce dernier.

Les axes stratégiques du HDH sont en cohérence avec les axes de la Stratégie interministérielle pour construire notre patrimoine national des données de santé 2025-2028.

# Programme de travail 2025

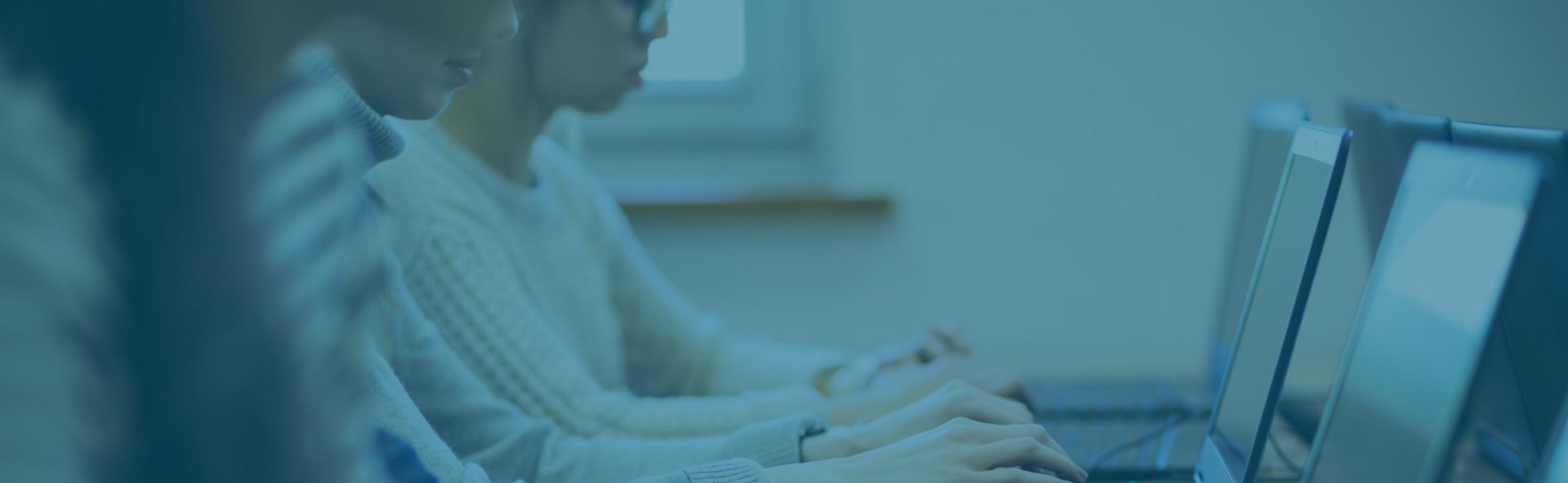




#1

## Grande orientation

Réduire les délais d'accès aux données de santé et multiplier les projets impactants



## Réduire les délais d'accès aux données de santé et multiplier les projets impactants

L'attractivité des données de santé pour la France et l'international dépend notamment des **délais d'accès**.

Le Health Data Hub poursuivra donc, avec ses partenaires tels que le ministère de la Santé et de l'Accès aux soins et la Caisse Nationale d'Assurance Maladie (CNAM), **l'industrialisation des processus clés d'accès aux données** et la **simplification des démarches administratives**.

# Objectif 1.1 - Simplifier et clarifier les procédures de création, d'accès et de réutilisation de données

## Réalisé 2024

- ✓ Concernant les démarches réglementaires, le HDH a participé à la consultation de la CNIL pour recueillir les besoins de l'écosystème en termes de référentiels simplifiés.
- ✓ Concernant les démarches contractuelles, en lien avec le Comité stratégique des données de santé :
  - ✓ Poursuite des travaux sur les redevances avec le lancement d'une **phase pilote pour tester les grilles de redevance** en conditions réelles et prise en compte des retours dans une deuxième version des grilles. Lancement d'un comité de suivi pour suivre l'intégration de ces grilles dans les processus des établissements de santé et leur généralisation,
  - ✓ Lancement d'un groupe de travail sur les **contrats-types de fourniture de données par des entrepôts de données de santé** dans le cadre du Comité stratégique des données de santé.

Ces travaux s'articulent par ailleurs avec les travaux nationaux de préfiguration de la mise en place de l'EHDS et les travaux menés par la Commission européenne avec les États membres, notamment au sein de l'action conjointe TEHDAS 2 (cf axe 3 et infra).

- ✓ Participation active du HDH au sein du **European Data Innovation Board** pour soutenir la Commission européenne dans la mise en œuvre du *Data Governance Act* et du *Data Act*.



## Plan de travail 2025

- Élaboration d'un **programme de travail pour le déploiement des référentiels simplifiés** en lien avec la CNIL et le Comité stratégique des données de santé, et **contribution à l'élaboration de ces référentiels**.
- Animation, sous l'égide du Comité stratégique des données de santé, du **comité de suivi de la mise en oeuvre opérationnelle de la tarification** et poursuite des travaux relatifs aux **contrats-types** dans le cadre de collaborations scientifiques.
- **Rédition du baromètre des projets**, visant à identifier et qualifier les difficultés rencontrées par les acteurs de l'écosystème et à analyser des **résultats**.
- Contribution à **l'analyse d'impact du règlement EHDS** sur les futures fonctions des organismes responsables de l'accès aux données de santé (HDABs) et ses implications sur le HDH et aux **travaux de TEHDAS 2**.

**CNIL**  
COMMISSION NATIONALE  
INFORMATIQUE & LIBERTÉS

  
**MINISTÈRE  
DE LA SANTÉ  
ET DE L'ACCÈS  
AUX SOINS**  
*Liberté  
Égalité  
Fraternité*

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

**TEHDAS**  
Towards  
European  
Health  
Data  
Space



Commission européenne

# Objectif 1.2 - Mettre à disposition des données d'intérêt et de qualité

## Réalisé 2024

- 🔍 Approbation par le Comité stratégique des données de santé de **12 bases de données visant à intégrer la seconde version du catalogue du SNDS** et saisine de la CNIL relative à l'instruction du projet d'arrêté listant ces bases.
- ✅ Réalisation d'un **kit documentaire à destination des responsables de données** qui souhaitent inscrire leur base de données au catalogue du SNDS, déclinaison des **outils de calcul des redevances** et définition des conditions **d'un primo-accès pour les bases de données du catalogue du SNDS** en lien avec le Comité stratégique des données de santé.
- ✅ Livraison du catalogue de métadonnées aligné avec le **modèle européen "HealthDCAT-AP", enrichissement du catalogue avec les partenaires du catalogue du SNDS** et du projet de Direct Grant "French\_HealthData@EU" et mise à disposition en **open-source** d'une version en marque blanche afin d'en permettre la réutilisation par d'autres acteurs et ouverture du code.
- ✅ Étude en cours visant à articuler le catalogue de métadonnées **FRESH**, piloté par l'Inserm, avec celui du HDH.
- ✅ Stabilisation du **label européen de qualité pour l'utilisation secondaire des données de santé dans le cadre du projet européen QUANTUM**, qui regroupe 30 partenaires de 14 pays. Une phase de test est prévue avant son lancement officiel.



## Plan de travail 2025

- Publication de la **mise à jour de l'arrêté portant sur le catalogue du SNDS (7 février 2025)**.
- **Intégration opérationnelle des premières bases au catalogue du SNDS**, chaînage avec la base principale du SNDS, déploiement des DOI et des catalogues de métadonnées associés.
- Mise en oeuvre des **grilles de redevances** pour les demandes d'accès aux bases de données du catalogue du SNDS.
- Poursuite de l'**enrichissement du catalogue de métadonnées** avec les partenaires du catalogue du SNDS et du projet de Direct Grant "French\_HealthData@EU".
- Connexion entre le **catalogue de métadonnées du HDH et le catalogue de métadonnées européen**.
- Coordination des partenaires français du HDH pour **tester le label de qualité et d'utilité des données**.



**Inserm**

La science pour la santé  
From science to health

# Zoom : 22 bases seront au catalogue du SNDS en début d'année 2025

Le **catalogue du SNDS** est une collection évolutive de bases de données, enrichie de manière itérative. L'arrêté du 12 mai 2022 a officialisé les **dix premières bases** inscrites, auxquelles se sont ajoutées **les douze bases sélectionnées** pour intégrer le catalogue.

**Base e-sis** : cancer du sein  
**Base ATU** : accès anticipé  
**BNDMR** : maladies rares  
**EPICOV** : Covid-19  
**ESME** : cancer du sein  
**E-must** : infarctus du myocarde  
**HEPATHER** : hépatite B ou C  
**MDO** : liste des maladies obligatoires  
**Memento** : maladie d'Alzheimer  
**OSCOUR** : surveillance des urgences



**SNDS**  
Système national des données de santé

## 12 bases ont intégré le catalogue en février 2025 suite à la publication de l'arrêté après avis de la CNIL



### Bases issues d'appel à projets/ appels à manifestation d'intérêt

**ARAC** : reste à charge

**SEDAAR** : ophtalmologie

**APSOREN, TARPON** : médecine d'urgence, traumatologie

**APRIORICS, NetSARC, PRECISION PREDICT** : oncologie



### Bases soutenues dans leur construction

**P4DP** : création d'un entrepôt de données de médecine de ville

**ISIS, REALIGIST** : oncologie



### Autres bases de données d'intérêt

**UroCCR** : registre sur le cancer du rein

**INTEGRA** : cohorte hospitalière sur les maladies métaboliques

# Zoom : le catalogue de métadonnées du HDH

Un catalogue de métadonnées granulaire, adapté aux données de santé et inscrit dans une démarche européenne :

- **Structuration des métadonnées en trois niveaux :** base, collection, variables.
- **Utilisation d'ontologies standards** (MeSH : "Medical Subject Headings").
- **Alignement avec les programmes européens :** recommandations TEHDAS, création d'une extension de DCAT-AP pour le domaine médical.
- **Conçu pour l'écosystème en santé :** utilisation pour documenter les bases catalogue du SNDS, et également mis à disposition en marque blanche aux acteurs désireux de déployer leur propre catalogue de métadonnées.

| Général                           |   |
|-----------------------------------|---|
| Acronyme                          | OSCOUR  |
| Identifiant DOI du jeu de données |   |
| Titre                             | Organisation de la surveillance coordonnée des urgences   |
| Référence catalogue HDH           |   |
| Description                       | La base de données OSCOUR contient l'ensemble des passages aux urgences en France métropolitaine et outre-mer. L'objectif principal de ce système est la détection précoce d'événements sanitaires inhabituels. Il est également utilisé pour le suivi des épidémies de pathologies saisonnières.         |
| Spécialité médicale               | Environnement et santé publique   |
| Mots-clés                         | Environnement et santé publique, Données administratives, CSV   |
| Documentation de valorisation     | <a href="https://www.santepubliquefrance.fr/surveillance-syndromique-sursaud-R/reseau-oscour-R-organisation-de-la-surveillance-coordonnee-des-urgences">https://www.santepubliquefrance.fr/surveillance-syndromique-sursaud-R/reseau-oscour-R-organisation-de-la-surveillance-coordonnee-des-urgences</a> |
| Financement                       | Public  |
| Schéma de circulation des données |   |
| Valorisation scientifique         |   |
| Langue de production des données  | Français  |



Le catalogue, notamment sa version en marque blanche, sont en cours de livraison pour une mise à disposition en open source.

 <https://catalogue-metadonnees.health-data-hub.fr/>

# Objectif 1.3 - Soutenir les projets (1/2)

## Réalisé 2024

- ✓ Publication de **guides** destinés à :
  - accompagner les porteurs de projets dans la rédaction de leur protocole,
  - détailler les démarches pour modifier une recherche,
  - expliciter l'obligation de transparence associée à l'usage des données de santé en français et en anglais.
- ✓ Mise à jour du **kit documentaire RGPD** permettant de construire le dossier de demande d'autorisation auprès de la CNIL, en prenant en compte les évolutions apportées par le nouveau référentiel de sécurité du SNDS.
- ✓ Publication des **kits pédagogiques sur les MR-007 et MR-008** visant à simplifier l'accès aux données de la base principale du SNDS **co-construits avec la CNIL et la CNAM** et organisation de 3 webinaires.
- 🔍 Publication de plusieurs **MOOC** décrivant la base des causes médicales de décès (CépiDc), des concepts clés de la base principale du SNDS et les bonnes pratiques de *data-management*.
- ✓ Organisation de **webinaires** à destination de l'AFCROs, du Leem et du SNITEM pour présenter l'obligation de transparence et déploiement d'une formation dédiée.



## Plan de travail 2025

- Expérimenter la **mise en place d'un cycle de formations / ateliers** à destination d'acteurs **primo-accédants** pour les accompagner dans l'élaboration de leur projet.
- **Publication d'un marché d'aide à la rédaction de protocole**, qui sera mobilisable pour des porteurs néophytes.
- Amélioration continue du **kit documentaire RGPD** au fil des instructions de la CNIL et de l'évolution de sa doctrine, ainsi que celle du Contrôleur européen de la protection des données (CEPD).
- Poursuite et mise à jour de la **traduction en anglais de l'offre de services du HDH** afin d'accompagner les porteurs de projets internationaux.

# Objectif 1.3 - Soutenir les projets (2/2)

## Réalisé 2024

- ✓ Poursuite de la **stratégie d'enrichissement de l'offre technologique** de la plateforme du HDH visant à :
  - faciliter l'import de paquets ou modèles pré-entraînés utiles pour les projets réalisés sur la plateforme technologique, notamment en IA,
  - faciliter le traitement de grands volumes de données notamment dans un contexte multi-utilisateur et faciliter l'industrialisation des traitements
  - permettre la connexion à un système fils, notamment dans le cadre du projet EMC2.
- ✓ **Renforcement de la documentation** à destination des utilisateurs, sur Zendesk et directement au sein de la plateforme via des Notebooks.
- ✓ **Déploiement d'une offre de services *open data*** pour accompagner les porteurs de projets dans l'ouverture de leurs bases de données anonymisées sur la plateforme gouvernementale **data.gouv.fr**.
- ✓ **Mise à disposition de collaborateurs du HDH dans le cadre de 8 projets accompagnés par le HDH** afin d'effectuer une expertise ponctuelle en data ingénierie, data science, standardisation et expertise des données de la base principale.



## Plan de travail 2025

- Amélioration continue de **l'offre technologique** afin de :
  - faciliter l'authentification aux différents services de la plateforme,
  - permettre aux utilisateurs de mieux synchroniser leurs outils avec la plateforme,
  - mettre en place une architecture distribuée pour Jupyter, afin d'avoir une meilleure allocation des ressources par espace et par utilisateur.
- Poursuite de la documentation interne à la plateforme et de la mise à disposition de guides et notebooks.
- Poursuite de **l'accompagnement pour l'ouverture en *open data* de bases de données anonymisées** issues des *data challenges* TissueNet ou VisioMel sur data.gouv.fr.
- Poursuite des mises à disposition de collaborateurs du HDH aux projets nécessitant un appui technique ponctuel.

# Objectif 1.4 - Mutualiser les connaissances et les expertises

## Réalisé 2024

- ✓ Poursuite du **programme data challenge de la SASN** avec 1 projet lancé et 3 projets sélectionnés en 2024 et publication en *open data* de la base de données anonymisées du *data challenge* **“Allergen-Chip challenge”** portant sur l'amélioration du diagnostic des allergies.
- 🔍 Mise en ligne de **19 ressources sur la Bibliothèque Ouverte des Algorithmes en Santé (BOAS)** et poursuite des appels à manifestation d'intérêt pour favoriser sa complétion. Les porteurs de projets accompagnés par le HDH sont également encouragés à y partager leurs ressources d'intérêt pour la communauté.
- ✓ Participation à la réalisation de la **feuille de route des Administrateurs Ministériels des Données, des Algorithmes et des Codes sources (AMDAC)** du ministère de la Santé et de l'Accès aux soins et notamment aux travaux relatifs à la conception et la publication du **guide d'accompagnement des acteurs à l'open source**.
- 🔍 **Lancement du projet PARTAGES et du projet PFDS**, respectivement lauréats d'un appel à projets France 2030, visant notamment à constituer une grande base de compte-rendus médicaux synthétiques et à l'ouverture de LLM médicaux pour faciliter leur réutilisation large et lauréat de l'appel à manifestation d'intérêt “Compétences et Métiers d'Avenir” visant notamment à produire des données synthétiques en vue d'enrichir les programmes de formation.



## Plan de travail 2025

- **Poursuite de l'enrichissement des bibliothèques de données open data en ligne et d'algorithmes (BOAS)**
- **Participation du HDH aux 4 groupes de travail de l'AMDAC** : "règles de gestion sur la pseudonymisation", "procédures partagées d'anonymisation", "définition, maintien de critères de qualité communs" et "documents de référence pour des entrepôts de données de santé".
- **Coordination du projet PARTAGES incluant** la production d'un corpus de comptes-rendus fictifs, accompagné d'un cadre méthodologique et d'un guide d'annotation et le développement de premiers modèles de langage (LLM), incluant des modèles de type BERT ainsi que des modèles génératifs spécifiques au domaine médical
- Développement de **plusieurs jeux de données synthétiques dans le cadre du projet PFDS** piloté par l'Université Grenoble-Alpes, **notamment en lien avec P4DP** (données de soins primaires ...), et réflexion plus large sur les enjeux des données de synthèse pour la recherche en santé.

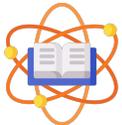
# Zoom : le développement d'une offre de services en science ouverte

Convaincu que l'intelligence collective permet une recherche plus transparente, inclusive et efficace, le HDH accompagne et soutient l'écosystème dans sa démarche de science ouverte.

À cette fin, le HDH a développé une **infrastructure dédiée à la mise en open data de données de santé anonymisées** articulée avec la plateforme data.gouv.fr dont les capacités en termes de volumétrie sont limitées.

## 4 bases de données publiées en open data sur data.gouv.fr

Deux de ces bases concernent des jeux de données synthétiques, tandis que deux autres sont issues de projets menés dans le cadre du programme "Data challenges en santé" co-porté par le HDH, la DNS et bpifrance. Les bases de 3 autres *data challenges* sont attendues en 2025.



The screenshot shows the 'Allergen Chip Challenge' dataset page on data.gouv.fr. The page includes a description of the challenge, the producer (Société Française d'Allergologie), the license (Licence Ouverte / Open Licence version 2.0), and a list of files for download. The files include a main CSV file, a dictionary PDF, and a dictionary XLS file.

Accueil > Jeux de données > ACC

Ajouter aux favoris Ouvrir sur explore.data.gouv.fr

### Allergen Chip Challenge acc

Description  
Doi 10.60597/j5fe-g420

L'augmentation constante des allergies nécessite d'améliorer le diagnostic des patients, qui repose sur l'histoire clinique, des tests cutanés (TC) et la détection d'anticorps IgE spécifiques (IgE) dans le sang. L'objectif est d'identifier ou d'exclure les allergènes potentiellement impliqués.

Nous disposons actuellement de tests IgEs multiplex basés sur des puces qui permettent une détection simultanée de centaines de spécificités allergéniques avec seulement 100 µL de sérum. Par exemple, Allergy Explorer ALEX<sup>2</sup> (Macro-Array Diagnostics, Vienne, Autriche) utilise 117 extraits allergéniques et 178 composants protéiques. Cet avantage à détecter les IgEs à de nombreux

Lire plus

Fichiers (2) Réutilisations et API (0) Discussions (0) Ressources communautaires (0) Informations

#### 1 FICHIER PRINCIPAL

allergenchipchallenge-data-corrected-final-hdh-sfa.csv  
Mis à jour le 26 juillet 2024 — csv (2,3Mo) — 50

#### 2 DOCUMENTATIONS

dictionnaire-acc-english.pdf  
Mis à jour le 26 juillet 2024 — pdf (55,7Ko) — 29

acc-dictionnaire-final.xls  
Mis à jour le 26 juillet 2024 — xls (43,5Ko) — 29

Les **compétitions en science des données ouvertes** "Data challenges en santé" visent à résoudre une problématique grâce à des solutions d'apprentissage automatique. Elles constituent des leviers innovants pour transmettre la culture de la science ouverte.



**2 compétitions** en science des données en 2024. 250 équipes ont participé au Data challenge DigiLut sur la détection du rejet de greffe pulmonaire organisé avec l'hôpital Foch.



**3 relances de l'appel à projets** "Data challenges en santé" permettant la sélection de 12 projets de compétitions depuis 2021.

# Objectif 1.5 - Renforcer l'attractivité du HDH

## Réalisé 2024

- ✓ **Poursuite des appels à projet pour soutenir des projets impactants sur des thématiques prioritaires avec :**
  - La sélection de **5 lauréats du second appel à projets "La donnée pour la recherche et l'innovation en santé-environnement"** du Green Data for Health (CGDD) et du HDH.
  - La sélection de **6 lauréats à l'appel à manifestation d'intérêt "Partage de données industrielles"** soutenu par Roche et MEDIPATH.
- ✓ **Poursuite de la présentation de l'offre de services du HDH auprès des acteurs de l'écosystème** avec plusieurs centaines de personnes rencontrées
- ✓ **Benchmark interne** de différentes structures nationales ou étrangères en vue de nourrir la réflexion sur la mise en place d'une fondation.



## Plan de travail 2025

- Lancement de la **troisième vague de l'appel à projets "La donnée pour la recherche et l'innovation en santé-environnement"** porté par le Green Data for Health (CGDD) et du HDH et de la deuxième vague de l'appel à manifestation d'intérêt **"Partage de données industrielles"**.
- Poursuite de la **présentation de l'offre de services auprès des acteurs de l'écosystème**.
- **Analyse du plan de création de la fondation**



#2

## Grande orientation

Mettre à disposition les données de la base principale, l'enrichir et faciliter sa réutilisation



## Mettre à disposition les données de la base principale, l'enrichir et faciliter sa réutilisation

La base principale du SNDS représente un potentiel énorme pour la recherche en France, sa mise à disposition dans les meilleures conditions continue de constituer un enjeu prioritaire du HDH.

Son exploitation par les acteurs de l'écosystème pourra être renforcée par **l'intégration de nouveaux périmètres de données d'intérêt majeur pour la recherche**, un meilleur **accompagnement des utilisateurs de la base** ainsi que par le maintien d'une **plateforme technologique sécurisée et fonctionnelle à l'état de l'art**.

# Objectif 2.1 - Développer l'expertise du HDH dans la mise à disposition des données

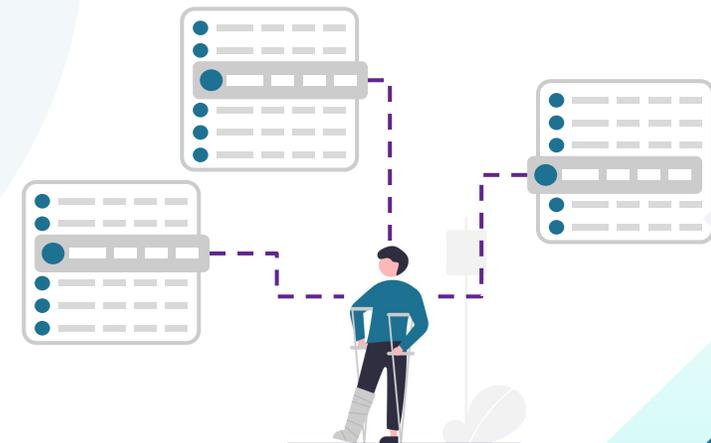
## Réalisé 2024

- ✓ Poursuite du **partenariat avec la CNAM** avec la réalisation de plus de **300 extractions**, ainsi qu'une trentaine de ciblage et d'appariements.
- 🔍 Mise en œuvre du **partenariat stratégique avec la CNAM pour l'autonomisation du HDH** dans la mise à disposition des données pour les projets qu'il accompagne.
- 🔍 Ouverture d'une partie des **services du concentrateur** pour industrialiser les appariements de bases diverses à la base principale, avec ou sans reconstitution du NIR, dans un contexte de hausse des demandes des utilisateurs pour ce type de croisement.



## Plan de travail 2025

- Poursuite du **partenariat avec la CNAM** avec la réalisation d'extractions, de ciblage et d'appariements et la mise à disposition des données en quasi autonomie pour les projets visés par la convention de cotraitance.
- **Travaux communs avec la CNAM sur les processus de mise à disposition des données** (stratégie de minimisation, conventionnement, évolution du *starter kit*, etc.).
- Retour d'expérience sur les **services du concentrateur** et évolutions si nécessaire.



# Zoom : la CNAM et le HDH travaillent de concert afin de faciliter l'accès à la base principale du SNDS

Dans le cadre d'un plan d'action conjoint conçu à la fin de l'année 2021 et destiné à multiplier par 2 la capacité de traitement des accès des projets de recherche aux données de la base principale du SNDS, **le HDH appuie la CNAM dans la réalisation des extractions à travers le portail de la CNAM.**



Depuis décembre 2023 et la **signature du convention de cotraitance entre la CNAM et le HDH**, les équipes du HDH bénéficient d'une autonomie quasi-complète sur la phase de préparation des données pour les projets accompagnés : ces opérations comprennent les **ciblages, appariements et extractions à opérer sur la base principale du SNDS** et, le cas échéant, **sur des bases externes apportées par les porteurs de projets**, ainsi que la phase de conventionnement avec les porteurs, qui ne fait plus intervenir la CNAM.



Cette collaboration, opérée depuis 2022 concerne :

- l'ensemble des projets utilisant des données de la base principale du SNDS,
- qu'ils soient ou non accompagnés par le HDH,
- qu'ils utilisent ou non la plateforme technologique du HDH in fine pour l'exploitation scientifique des données.



**Les objectifs fixés dans le plan d'action ont été atteints** et en pratique les livraisons de données de la base principale du SNDS ont concerné de l'ordre de **200 projets par an en 2022, 2023 et 2024** contre une centaine en 2020 et 2021 (soit le triple par rapport à 2018).



Des travaux pilotés par la CNAM d'optimisation des process sont en cours et des ateliers ont été organisés entre la CNAM et le HDH dans le cadre d'un objectif de meilleure maîtrise des délais d'accès sur projet aux données du SNDS.



Les projets traités en 2025 pourront bénéficier d'un pilotage plus fluide et plus visible des délais d'accès via la construction d'indicateurs pertinents suivant la typologie des projets.

# Zoom : le SI concentrateur, une solution pour simplifier l'appariement de jeux de données de différents responsables de données

Un accès simplifié des responsables de données à travers un portail web



Un suivi de l'avancement de l'appariement



Une gestion intégrée des projets multicentriques



La sécurité des données directement identifiantes au coeur des choix d'architecture



L'authentification à double facteur est obligatoire pour tous les utilisateurs

## Deux possibilités offertes aux responsables de données

### Appariement direct NIR

Le responsable de données sait associer chaque individu de sa base de données à **son NIR bénéficiaire**, ou à **son triplet** (NIR ouvrant droit, sexe et date de naissance du bénéficiaire).

Il peut alors fournir une table d'accrochage sous le format suivant :

| Id_acc | NIR ouvrant droit | Sexe bénéficiaire | Date naissance bénéficiaire | NIR bénéficiaire |
|--------|-------------------|-------------------|-----------------------------|------------------|
|        |                   |                   |                             |                  |
|        |                   |                   |                             |                  |

### Appariement avec reconstitution du NIR

Le responsable de données sait associer chaque individu de sa base de données à des informations de **son état civil** (nom, prénom, date de naissance, sexe).

Il peut alors fournir une table d'accrochage sous le format suivant :

| Id_acc | Nom | Prénom | Date naissance | Sexe |
|--------|-----|--------|----------------|------|
|        |     |        |                |      |
|        |     |        |                |      |

Toutes les données directement identifiantes sont **chiffrées en transit et au repos**



Toutes les données directement identifiantes **ne peuvent plus être consultées** mais seulement envoyées pour pseudonymisation

# Objectif 2.2 - Faciliter l'exploitation de la base principale

## Réalisé 2024

L'utilisation des données de la base principale du SNDS soulève un certain nombre de défis, sur lesquels, année après année, le HDH et ses partenaires tentent d'apporter solutions et outillages.

### Le partage de documentation et de programmes d'exploitation

- 🔍 Publication de **6 nouvelles fiches réalisées avec la société HEVA** et de **2 nouvelles fiches réalisées par le conseiller médical de la Drees sur la base principale du SNDS**. Réflexion sur la mise en place d'une gouvernance associée à la complétion et la maintenance de la documentation ouverte.
- 🔍 Mise en ligne de la **Bibliothèque Ouverte des Algorithmes en Santé avec 19 programmes référencés** et organisation des vagues 5, 6 et 7 de **l'appel à projets relatif à la Bibliothèque Ouverte des Algorithmes en Santé (BOAS)** pour l'enrichir.

### La conception et la diffusion de formations sur la base principale

- ✓ Déploiement des **MOOC sur la base principale du SNDS** en français et en anglais et mise en ligne d'une **formation sur les nomenclatures de la base principale**.
- ✓ Mise en place d'une **réflexion sur la stratégie de déploiement des formations en lien avec l'identification de concepts à maîtriser selon le niveau de l'utilisateur, élaboration de priorisations et d'un plan d'action**, dans le cadre d'un groupe de travail du Comité Stratégique des données de Santé lancé en juin 2024.

L'ensemble de ces travaux ont vocation à se poursuivre en 2025

### La mise au format OMOP des données pour des projets de plus en plus nombreux et à dimension européenne

- ✓ **Mise en pratique des scripts d'omopisation** en phase de préparation et rédaction des processus de test.

### L'amélioration des processus d'ingestion, de validation et également de gestion des variables sensibles avant la mise à disposition des données au sein de la plateforme

- ✓ **Industrialisation du code de validation des données** ingérées, à noter que cette industrialisation porte également sur les données autres que celles de la base principale du SNDS avec les défis techniques associés (par exemple les données d'imagerie médicale).
- ✓ Définition et développement d'une **pipeline standardisée de suppression des FINESS dans les extractions de la base principale du SNDS**, afin de limiter le risque de réidentification d'un établissement.

### La mise en place d'un support pour les utilisateurs

- ✓ Amélioration des configurations pour une utilisation fluide de Spark par de multiples utilisateurs et sur plusieurs To de données
- ✓ Mise à disposition d'un **premier guide exécutable sur la plateforme** pour répondre aux besoins des utilisateurs.
- ✓ Les travaux relatifs à la mise à disposition d'un logiciel de visualisation des parcours de soins sur la base principale en partenariat avec l'Inria se poursuivent pour aboutir en 2025

**Par ailleurs, le HDH a remis son rapport prévu par le décret relatif au SNDS** sur l'utilisation de son accès permanent.

# Zoom : la documentation collaborative du SNDS en ligne

## Une documentation collaborative autour des données riche en ressources et informations

- **> 70 fiches thématiques sur la base principale du SNDS**, intégrant pour certains des bouts de code réutilisables
- **Guide d'initiation** à destination de tout nouveau utilisateur souhaitant manipuler ces données
- **Nombreuses ressources partagées par nos partenaires** telles que des supports de formation, guides et autres documentations
- **Revois vers des ressources externes** comme des open data, des programmes, des données synthétiques, etc.
- Scripts de **standardisation du SNDS au format OMOP-CDM** avec documentation associée ainsi qu'une page dédiée au **standards de données de santé**



HEA  
DAT.

- **Une dizaine de nouvelles fiches thématiques publiées sur l'année 2024**
- **Une réflexion en cours sur la mise en place d'un comité éditorial regroupant les différents institutions partenaires et producteurs de données**

### Introduction

Qu'est-ce que le SNDS ?  
Les bases du SNDS, de l'alimentation à la restitution des données  
Comment accéder au SNDS ?  
La recherche avec le SNDS

### Fiches

### Glossaire

### Se former au SNDS

### Pour aller plus loin

### Open Data

## Documentation collaborative du SNDS

Bienvenue sur la documentation collaborative du Système National des Données de Santé.

Cette documentation est en construction, via ce [dépôt GitLab](#).

### Contributeurs

Cette documentation est maintenue par le Health Data Hub.

Elle résulte d'une mise en commun de documents et travaux par plusieurs organisations, dont :

- la Caisse nationale d'assurance maladie - Cnam
- le Health Data Hub - HDH
- Le Ministère des Solidarités et de la Santé: la Direction de la Recherche, des études, de l'évaluation et des statistiques - DREES
- les Agences régionales de santé - ARS

Requête SAS pour le repérage des détenus

```
proc sql;
drop table sasdata1.detenus ;
%connectors;
create table sasdata1.detenus as select * from connection to oracle (
select distinct /*t1.ben_nir_psa, t1.ben_rng_gen,*/ t2.ben_idt_ano,
t1.ben_sex_cod,
t1.ben_nai_ann
from er_prs_f t1 left join IR_BEN_R t2 on (t1.ben_nir_psa=t2.ben_nir_psa
and t1.ben_rng_gen=t2.ben_rng_gen)
where rng_cod in (651,652,653,654,655,656,657,658,659)
and RGM_GRG_COD in (1)
AND EXE_SOI_DTD BETWEEN to_date('01012018','DDMMYYYY') AND to_date('31122018','DDMMYYYY')
AND FLX_DIS_DTD BETWEEN to_date('01012018','DDMMYYYY') AND to_date('31032019','DDMMYYYY')
);
disconnect from oracle;
quit;
```

Documentation du SNDS & SNDS OMOP

### Introduction

### Documentation de l'ETL

Traitements préliminaires

Traitement préliminaire des référentiels

traitements

Traitement préliminaire des tables du

DCIR

Établissements de soins - PMSI

Établissements de soins et pharmacies

- DCIR

Table LOCATION

Patients

Table CARE\_SITE

Table PERSON

Table OBSERVATION\_PERIOD

Table PROVIDER

Table VISIT\_OCCURRENCE

STDL\_TABLE

Macros

À propos

Variables du format OMOP-CDM remplies

| Variable     | Type        | Description  | Clié primaire | Clié étranger | Table de la clé étrangère |
|--------------|-------------|--|---------------|---------------|---------------------------|
| location_id* | integer     | Identifiant unique pour chaque lieu géographique.  | Oui           | Non           |                           |
| address_1    | varchar(50) | Le champ d'adresse 1, généralement utilisé pour le nom de la rue, tel qu'il apparaît dans les données sources. | Non           | Non           |                           |
| city         | varchar(50) | Nom de la ville tel qu'il apparaît dans les données sources.   | Non           | Non           |                           |



<https://documentation-snds.health-data-hub.fr>

Visites/mois :

10 000

# Zoom : les premières contributions à notre Bibliothèque Ouverte d'Algorithmes en Santé (BOAS)



Depuis le lancement de la BOAS en avril 2024, **une vingtaine de ressources y ont été publiées**. Des algorithmes et programmes sont donc désormais **accessibles à tous** afin d'être **facilement réutilisables** par la communauté scientifique.



Des programmes pour **simplifier et standardiser les extractions de données** de la base principale du SNDS

- **Outil Magic Loop** pour simplifier l'extraction de données (DNUM/SCN)
- **Scripts de standardisation de la base principale du SNDS** selon le modèle OMOP (HDH)



**Des algorithmes de ciblage** dans la base principale du SNDS

- Un algorithme qui vise à produire une **estimation de la sévérité des AVC ischémiques** (Santé publique France)
- Un algorithme pour cibler les personnes **prises en charge pour un diabète** (CNAM)
- Un algorithme permettant le **repérage des individus avec des limitations motrices ou organiques** (IRDES)
- Un algorithme ciblant des **hépatites virales chroniques** (IPLesp et HDH)



**Des requêtes à la demande**, pour construire des indicateurs simples sur les données de la base principale du SNDS

- Prise en charge des **patients atteints de la maladie de Crohn** (HDH pour l'AFA Crohn RCH France)
- Estimation de la population de plus 40 ans ayant un **statut d'ayant droits** (HDH pour l'université Paris-Dauphine, LEDa UMR 8007)
- **Prise en charge des allergies aux pollens** dans la région Normandie (Groupe Vyv)

# Zoom : le catalogue de formation du Health Data Hub

## Des formations sur la base principale du SNDS disponibles en accès libre

- **Formations à destination du grand public**
  - SNDS CITY - 001 : Formation citoyenne au Système National des Données de Santé
- **Formations relatives aux procédures réglementaires**
  - REGL - 001 : Présentation des procédures d'accès aux données personnelles de santé
  - REGL - 002 : Formation sur les procédures simplifiées MR-007 et MR-008
  - REGL - 003 : Formation sur les modalités réglementaires d'accès à la base principale du SNDS
- **Formations à destination des utilisateurs de la base principale du SNDS**
  - SNDS TECH - 001 : Formation sur les données et ses modalités d'accès
  - MOOC SNDS - 001 : Tout ce que vous avez voulu savoir sur la base principale du SNDS

The screenshot shows the 'Toutes nos formations' page on the Health Data Hub website. The page features a navigation bar with user roles (Citoyen, Utilisateur de données, Responsable de données, Candidat, Start-up) and menu items like 'EXPLORER LE HDH', 'DÉPOSER VOTRE PROJET', 'NOTRE OFFRE DE SERVICE', 'DÉCOUVRIR LE CATALOGUE', and 'ACTUALITÉS'. A search bar and a filter sidebar are visible. The main content area lists training courses, including 'REGL - 003 : Formation sur les modalités réglementaires d'accès à la base principale du SNDS' and 'REGL - 002 : Formation sur les procédures simplifiées MR-007 et MR-008'. Each course entry includes a publication date, a 'Supports disponibles' icon, and a 'TOUT PUBLIC' label.

<https://health-data-hub.fr/formations>

### Mooc SNDS

Visites/mois :

100



*En 2024, le HDH a achevé le déploiement des MOOC sur la base principale du SNDS, disponibles en français et en anglais, et a lancé une formation en ligne dédiée aux nomenclatures de cette base.*



# Zoom : relever les défis des projets transfrontaliers, l'exemple du projet Coagulopathie



Dans le cadre du projet pilote pour l'**Espace Européen des Données de Santé**, le HDH participe à l'étude "Natural history of coagulopathy in COVID-19 patients and persons vaccinated against SARS-CoV-2 during the Omicron period" conduite par l'Agence Européenne des Médicaments (EMA). Cette étude réunit quatre pays européens ainsi qu'un centre de coordination (CC).

L'objectif est de **cartographier les étapes et les délais nécessaires à la mise à disposition de données formatées selon le modèle OMOP-CDM** en vue de la concrétisation d'un projet pan-européen.



Chaque structure doit exécuter localement des packages R fournis par le CC sur une base de données préalablement formatée selon le modèle OMOP-CDM. Les résultats agrégés sont ensuite transmis au CC pour analyse centralisée.



Le HDH participe à ce projet en mobilisant la **base principale du SNDS**. Les données de la population d'étude ont été échantillonnées et extraites depuis le portail CNAM.



Celles-ci ont été transférées sur la plateforme du HDH, **converties au format OMOP-CDM**, puis déposées dans un environnement technique sécurisé afin de réaliser les analyses.



La préparation des données brutes a nécessité **15 jours pour l'échantillonnage de la population** (10 millions de bénéficiaires) et **2 mois pour l'extraction** des données (468 fichiers, 6,4 To).



**L'ingestion et la validation** des données sur la plateforme HDH ont duré 17 jours. La conversion au format OMOP-CDM (16 fichiers) a duré 3 mois.



Une phase **d'alignement des vocabulaires** a permis de mapper les terminologies sources (CIM-10...) vers les vocabulaires standards définis par OMOP-CDM.



Un **environnement technique** a été mis en place comprenant des outils adaptés au traitement de grandes volumétries de données (Spark et Hive), avec une **allocation dynamique de ressources**.

# Objectif 2.3 - Contribuer à enrichir la base de données de la base principale

## Réalisé 2024

- ✓ Un certain nombre **d'initiatives menées de manière continue vont permettre de faciliter l'utilisation de la base principale enrichie avec d'autres sources de données** dans le cadre d'appariements facilités par les partenariats et la mise en place du SI concentrateur (cf supra) : l'enrichissement du catalogue du SNDS, le partenariat santé-environnement avec GD4H (cf infra), etc.
- ✓ Le Comité stratégique des données de santé avait également déjà mené des premiers travaux en 2023 sur un ensemble commun de données au sein des EDS hospitaliers. Il ambitionne de **poursuivre ces travaux avec une envergure plus large** en adressant des questions diverses telles que celles entourant le lien avec les données de santé scolaire, la santé au travail, ou l'autonomie.

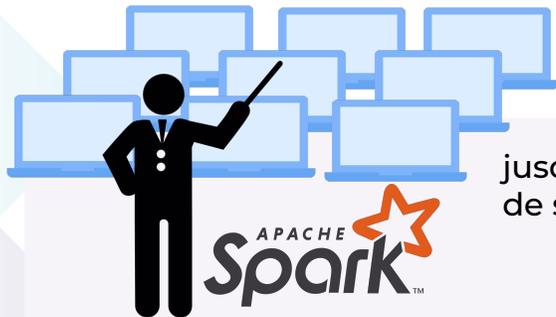


## Plan de travail 2025

- Suivi de **l'implémentation et la disponibilité du socle commun des entrepôts de données de santé hospitaliers** dans le cadre de l'accompagnement du HDH aux lauréats de l'appel à projet "accompagnement et soutien à la constitution d'entrepôts de données de santé hospitaliers" soutenu par France 20230
- Mise en oeuvre par le Comité stratégique des données de santé de la **stratégie d'enrichissement du SNDS centralisé**.

# Objectif 2.4 - Maintenir la plateforme à l'état de l'art fonctionnellement et en matière de sécurité

Les **machines virtuelles de la plateforme technologique** du HDH sont **ajustables** en fonction des besoins spécifiques des projets (volumétrie, outils, format des données, etc.) et peuvent être **modifiées en cours d'étude**. La plateforme propose actuellement une **dizaine** de types de machines différents.



jusqu'à x30  
de scaling

L'outil Spark permet d'ajuster la puissance disponible en distribuant le calcul sur plusieurs machines.

La puissance maximale utilisée sur la plateforme par un projet à un moment t :

**480 Coeurs (CPU)**

**1 pétaoctet de données**  
stockées sur la  
plateforme



jusqu'à **512 unités de calcul (CPU)** par projet.

**2 à 64 coeurs (CPU)**  
nombre de processeurs

**1 à 4 GPU**  
processeurs graphiques,  
(pour les projets de *machine learning* ou  
*deep learning*)

## Réalisé 2024

- ✓ **Implémentation des exigences eIDAS** requises par le nouveau référentiel de sécurité du SNDS concernant les moyens d'identification électroniques et **mise à jour de l'homologation de la plateforme technologique** au regard du nouveau référentiel de sécurité du SNDS.
- ✓ **Amélioration continue des opérations techniques et de sécurité et poursuite de la prise en compte des retours d'expérience.**



## Plan de travail 2025

- **Amélioration continue des opérations techniques et de sécurité et poursuite de la prise en compte des retours d'expérience.**

The background of the slide features a close-up of two hands shaking, symbolizing agreement or partnership. The hands are positioned in the center-right of the frame. The background is a blurred European Union flag, with its characteristic blue field and yellow stars. A large, light blue triangular graphic element is on the left side of the slide, pointing towards the center. The text is overlaid on this graphic and the background.

**#3**

## **Grande orientation**

Renforcer les connexions du HDH aux acteurs de l'écosystème



## Renforcer les connexions du HDH aux acteurs de l' écosystème

Depuis la mise en œuvre du HDH, l'articulation entre tous les niveaux - **local, national et européen** - et les conditions d'un **partage efficace entre ces acteurs**, en termes de **standards, de gouvernance mais aussi en termes de normes de sécurité** sont travaillées. Cet enjeu devient d'autant plus crucial avec la mise en œuvre du règlement européen des données de santé dit EHDS.

# Objectif 3.1 - Renforcer le travail partenarial avec les acteurs locaux (1/4)

## Réalisé 2024

- ✓ Poursuite de l'accompagnement du **projet EMC2** : préparation des données par les établissements partenaires en vue de leur transformation au format OMOP et leur croisement avec la base principale par le HDH.
- ✓ Poursuite de l'accompagnement des projets multicentriques et de l'analyse des enseignements associés :
  - Accompagnement des projets lauréats du **programme Unibase**.
  - Accompagnement du projet multicentrique **French Paradise** et intégration de deux nouveaux partenaires : CHU Dijon et CHU Besançon.
- ✓ Lancement du **plus ambitieux projet partenarial** mené par le HDH impliquant les deux tiers de la communauté des établissements de santé autour des enjeux de l'IA générative en santé (**PARTAGES**).



## Plan de travail 2025

- **Poursuite de l'accompagnement du projet EMC2** : transfert à l'Agence européenne du médicament (EMA) et mise à jour régulière des données.
- Poursuite de l'accompagnement des projets multicentriques avec les établissements de santé.
- **Coordination du projet PARTAGES** : production d'un corpus de comptes-rendus fictifs, élaboration d'un cadre méthodologique et d'un guide d'annotation, développement de premiers modèles de langage (LLM) tels que BERT et modèles génératifs spécifiques au domaine médical.



EUROPEAN MEDICINES AGENCY  
SCIENCE MEDICINES HEALTH

# Zoom : le projet PARTAGES a été sélectionné dans le cadre de l'appel à projets "communs numériques"



Le projet est né à la suite des travaux menés dans le cadre de la stratégie nationale pour l'IA (SNIA) sur les opportunités offertes par les traitements d'analyse de données textuelles pour la santé. Soumis le 24 octobre 2023 à l'appel à projets "communs numériques pour l'IA générative" de la SNIA soutenu par France 2030, **le projet a été sélectionné le 11 octobre 2024.** Le projet a également été soumis au sommet international pour l'action sur l'intelligence artificielle (IA).

Le projet PARTAGES réunit des acteurs représentatifs du territoire français mobilisés pour leurs compétences de pointe dans le secteur de la santé et de l'IA.



**10 équipes de recherche à la pointe du traitement automatisé du langage (TAL) en santé**



**20 établissements de santé**



**2 partenaires DeepTech (reciTAL et Mistral)**

**17 associations d'internes en médecine**

L'**UNESS**, l'Université Numérique En Santé et en Sport



# Zoom : le projet PARTAGES vise à libérer le potentiel du traitement automatisé du langage dans le domaine de la santé, en 4 étapes

Le développement de l'IA générative en santé permettra des **bénéfices majeurs en termes de qualité des soins et de compétitivité.**



Le coeur de ce potentiel réside dans la **compréhension, la génération et la structuration automatiques du texte libre en santé.**



La matière fondamentale ? Les **comptes-rendus médicaux** qui contiennent toutes les informations des parcours patients.

Pour cela, le projet PARTAGES met en oeuvre **4 étapes** :



## Étape 1

**Un LLM génératif médical général open source en langue française** pour servir de base à tous les cas d'usage futurs de la communauté.

*Grands modèles open source  
Bases médicales ouvertes  
Calculateur Jean Zay*



## Étape 2

**Une base inédite de comptes-rendus médicaux fictifs annotés** qui permettra à tous de traiter des cas d'usage spécifiques en se basant sur le LLM génératif médical général.

*Professionnels de santé  
Méthodologie rigoureuse  
Data augmentation*



## Étape 3

**7 modèles spécialisés adressant des cas d'usage à fort impact** pour la recherche et l'innovation en santé ainsi que le système de soins.

*Cas d'usage techniques  
Cas d'usage médicaux  
Modèles d'automatisation  
Cas d'usage de formation*



## Étape 4

**Une plateforme nationale souveraine d'évaluation fédérée** pour tester sur des données réelles tout algorithme développé, en respectant le cadre réglementaire.

*Déployée sur 18 établissements de santé  
Joignable par tout établissement.*

# Zoom : 7 cas d'usage médicaux à fort impact et réutilisables par l'écosystème seront développés dans le cadre du projet PARTAGES

1



## Pseudonymisation de comptes rendus

Sécuriser l'écosystème dans cette étape obligatoire préalable à toute utilisation des données grâce à un outil de pseudonymisation automatique fiable des comptes rendus médicaux.

2



## Codage automatique d'informations médicales

Automatiser et uniformiser le codage des informations médicales (DIM) aujourd'hui effectué par les médecins, et qui constitue la base du système de financement des établissements de santé.

3



## Résumé automatique de comptes rendus médicaux

Générer automatiquement des résumés pertinents et structurés de comptes rendus, afin de faire gagner un temps précieux aux professionnels de santé.

4



## Génération de cas cliniques pour la formation

Générer automatiquement et à la demande des cas cliniques représentatifs et diversifiés pour les examens en études médicales, en partenariat avec l'organisme régulateur.

5



## Identification de biomarqueurs

Identifier automatiquement les biomarqueurs tumoraux dans les comptes rendus des patients, pour appuyer la recherche en oncologie et gagner un temps médical précieux.

6



## Analyse de la réponse aux traitements en oncologie

Évaluer l'efficacité d'un traitement donné en structurant à partir de comptes rendus la réponse au traitement dans les bases de données de soins, afin de faciliter la recherche en oncologie.

7



## Détection automatique en infectiologie

Déterminer automatiquement la provenance des bactéries responsables d'infections du sang chez des patients, pour mieux comprendre les infections et lutter contre l'antibiorésistance.

# Objectif 3.1 - Renforcer le travail partenarial avec les acteurs locaux (2/4)

## Réalisé 2024

- ✓ Poursuite de l'accompagnement et des travaux impliquant les lauréats de l'appel à projets "accompagnement et soutien à la constitution d'entrepôts de données de santé hospitaliers" par le biais de différentes actions :
  - ✓ Atelier organisé à l'occasion de la **journée anniversaire de la SASN** avec les lauréats de l'appel à projets "entrepôts de données de santé" ainsi que lors de la **journée nationale pour l'innovation en santé** sur le sujet de l'open source et la mutualisation des outils.
  - ✓ 3 sessions collectives organisées : présentation du kit EDS, présentation du cadre réglementaire (CNIL) et formation aux bonnes pratiques de gestion de projet numériques (DINUM)
  - ✓ Réalisation d'un **panorama des logiciels destinés aux entrepôts de données de santé** en collaboration avec l'ANAP.
  - ✓ Poursuite de l'animation des travaux du Comité stratégique impliquant les lauréats pour construire une **doctrine nationale en matière d'entrepôts de données de santé** (cf infra) et mise en place d'une articulation entre le comité et la Commission des data des CHU.
  - ✓ Co-construction d'une **convention entre les lauréats de l'appel à projets et le HDH** pour encadrer le partage des données.
  - ✓ Participation aux **10 étapes du tour de France des entrepôts de données de santé hospitaliers**, organisé par la Délégation au Numérique en Santé (DNS) dans le cadre de la 'Stratégie interministérielle pour construire notre patrimoine de données de santé'.
- ✓ Sélection de **15 lauréats pour la première édition** de l'appel à projets Données de Santé et Application (DAAtAE) qui finance des projets de recherche réutilisant les données des entrepôts de données de santé hospitaliers, lancé par la DGOS et opéré par le HDH et lancement de la deuxième **vague de l'appel à projets DAAtAE**.
- ✓ Poursuite des travaux de sensibilisation sur cette thématique clé avec l'organisation de **3 webinaires sur les entrepôts de données de santé** dans le cadre des "mardis de la donnée de santé".



## Plan de travail 2025

- **Poursuite de l'animation et de l'accompagnement des lauréats** de l'appel à projet par le biais d'ateliers, sessions collective et le suivi du partage des données.
- Poursuite du développement du **catalogue de formations collectives sur les sujets prioritaires**, identifiés au niveau national et par les acteurs locaux
- Poursuite des travaux dans le cadre du comité stratégique des données de santé et lancement d'un groupe de travail dédié à **l'offre de services des entrepôts de données de santé hospitaliers**.
- Finalisation des **signatures de convention** avec les lauréats de l'AAP France 2030.
- Sélection des lauréats pour la **vague 2 de l'appel à projets DAAtAE avec la DGOS**.

# Objectif 3.1 - Renforcer le travail partenarial avec les acteurs locaux (3/4)

## Réalisé 2024

- ✔ Signature de la **convention entre le HDH et le CNGE** pour l'inscription au catalogue du SNDS de la base P4DP et réalisation des travaux préparatoires à l'ingestion des données.
- ✔ Sélection de **8 lauréats de l'appel à projets "P4DP" (Platform for Data in Primary care)** lancé avec le **Collège national des généralistes enseignants (CNGE)** et le **Collège de la médecine générale (CMG)** visant à utiliser les données de la première plateforme nationale des données de santé pour la médecine de ville croisées avec la base principale du SNDS.
- ✔ Autorisation par la **CNIL** de **l'entrepôt de données de santé "P4DP"**.



## Plan de travail 2025

- Ingestion de la **base P4DP** dans le catalogue du SNDS.
- Accompagnement dans la **mise en oeuvre opérationnelle des projets lauréats et cas d'usage pilotes de l'appel à projets "P4DP"**.
- Lancement d'une **deuxième vague de l'appel à projets "P4DP"**.

8

**projets accompagnés dans le cadre de l'appel à projets "P4DP"**

lancé avec le Collège national des généralistes enseignants (CNGE) et le Collège de la médecine générale (CMG)



# Objectif 3.1 - Renforcer le travail partenarial avec les acteurs locaux (4/4)

## Réalisé en 2024 à poursuivre en 2025

### ✓ Poursuite d'échanges privilégiés avec des familles d'acteurs avec lesquelles l'intensification des liens est essentielle :

- ✓ **Santé-environnement** : sélection de 5 projets lors de la deuxième vague de l'appel à projets conjoint avec le **Green Data for Health** et réalisation d'un webinaire citoyen sur la santé-environnement.
- ✓ **Génomique** : Initiation d'une réflexion pour une articulation efficace avec le **CAD** (Collecteur analyseur de données, portant sur les données génomiques) et réalisation de 2 webinaires "Les Mardis de la donnée de santé" sur les données génomiques en lien avec les acteurs clés du domaine (en collaboration avec le CHU de Montpellier, Auragen, SeqOIA et le Plan France Médecine Génomique 2025).
- ✓ **Organismes de recherche** : organisation d'un webinaire de présentation de l'offre de services du HDH à destination des chercheurs du **CNRS**, travaux menés avec l'Inserm dans le cadre de l'élaboration des 3 Programmes et Équipements Prioritaires de Recherche (**PEPR**) soutenus par le Plan Innovation Santé 2030. En 2025, de nouveaux projets sont à prévoir avec l'Inria, notamment autour des données de synthèse.
- ✓ **Associations de patients** : partenariat avec la Ligue contre le Cancer pour soutenir deux projets d'intérêt autour de l'utilisation des **données de santé en cancérologie**.

### ✓ Poursuite d'échanges privilégiés avec des familles d'acteurs avec lesquelles l'intensification des liens est essentielle :

- ✓ **Acteurs de l'IA** : Production d'une formation citoyenne en cours ; échanges avec les IA clusters dans l'optique de mutualiser les travaux de veille et d'identifier des synergies ; **Lancement d'un partenariat avec l'ANS** pour tester les approches innovantes en traitement du langage naturel sur les données de SAMU.
- ✓ **Professionnels de santé** : audition par l'Académie de médecine et réflexion sur les synergies possibles avec des syndicats d'internes.
- ✓ **Start-up et industriels** : nouvelle édition de l'événement "**Start-up Connect**" co-organisé avec Parisanté Campus et en partenariat avec Bpifrance, France Digitale, la French Care et France Biotech pour favoriser le dialogue avec les start-ups et participation active du HDH aux travaux du **Contrat Stratégique de Filière Industries et Technologies de Santé (CSF ITS)**.
- ✓ Réalisation d'**événements en commun avec Parisanté Campus** pour **contribuer à diffuser une culture de la donnée de santé** et participation à **Campus Live**, l'événement annuel organisé par Parisanté Campus.
- ✓ **Agences sanitaires** : soutien à Santé Publique France (SPF) dans le projet Orchidée visant à développer des indicateurs de surveillance d'infections respiratoires



# Zoom : mobiliser les données industrielles pour la recherche

Afin de multiplier les opportunités d'exploration et de réutilisation de données collectées par des acteurs industriels, Roche, Medipath et le HDH ont lancé **un appel à manifestations d'intérêt pour le "Partage de données industrielles"**.

Cette initiative permettra de réfléchir aux **conditions de partage des données d'origine privée**, ainsi que d'identifier et d'accompagner les projets d'intérêt public promettant des résultats rapides et innovants à partir de ces données.



**Données industrielles.** Les projets mobilisent une source de données proposée par un partenaire industriel, croisée ou non avec la base principale du SNDS.



**6 projets lauréats bénéficiant jusqu'à 100k€ d'accompagnement financier.** Un appui humain et technique sera également mis en oeuvre par le HDH durant 24 mois.



Les lauréats pourront bénéficier d'un espace projet sur la **plateforme technologique** du HDH, pour traiter les données en toute sécurité et avec des outils à l'état de l'art.

## Sources de données proposées par les partenaires :



### Base PRM cancer du sein

Données de vie réelle sur l'efficacité et l'usage de médicaments utilisés dans le traitement du cancer du sein (25000 patients de 2015 à 2020).



### Base PRM cancer du poumon

Données de vie réelle sur l'efficacité et l'usage de médicaments utilisés dans le traitement du cancer du poumon (5000 patients de 2015 à 2020).



### Base cancer du poumon

Population ciblée ayant développé un cancer pulmonaire (cancer non à petites cellules), avec les résultats de tests d'immunohistochimie et de biologie moléculaire (500 patients de 2016 à 2023).

# Objectif 3.2 - Renforcer la présence du HDH au niveau européen et international (1/3)

## Réalisé 2024

🔍 Livraison des résultats du projet **HealthData@EU Pilot** avec notamment :

- ✓ Un prototype d'infrastructure de **nœuds** nationaux et européens.
- ✓ Un **standard de métadonnées spécifique au domaine de la santé** (Health DCAT-AP) permettant la mise en place d'un catalogue de métadonnées européen.
- ✓ Un **formulaire de demande d'accès unique** européen auprès des HDABs.
- ✓ Un rapport sur la **réalisation de 5 cas d'usage transfrontaliers**.
- ✓ Un événement co-organisé avec la Commission européenne réunissant plus de 90 participants issus de 21 pays afin de **préparer le passage de relais entre le projet Health Data@EU et les futurs acteurs de l'EHDS** et de la 3e assemblée générale du projet.



## Plan de travail 2025

- Clôture du projet **HealthData@EU Pilot** avec la soumission des rapports techniques et financiers du projet.
- Connexion du catalogue de métadonnées français au catalogue de métadonnées européen.

# Zoom : les résultats du projet-pilote de l'Espace européen des données de santé (EHDS)

**HealthData@EU Pilot** est un consortium de 17 partenaires européens, mené par le HDH visant à préfigurer l'Espace européen des données de santé (EHDS).

Après 2 ans de travail, le **consortium a été un véritable accélérateur de la mise en place de l'EHDS**. Grâce à une étroite collaboration avec la Commission européenne, le projet a créé la base du déploiement de **l'infrastructure HealthData@EU** qui va faciliter l'usage des données de santé à l'échelle européenne.

De nombreux livrables et enseignements concrets grâce aux travaux du pilote



Développement d'un **standard européen de métadonnées** (Health DCAT-AP) et coopération avec la Commission européenne pour la création d'un catalogue de métadonnées européen, alimenté par des catalogues nationaux afin de faciliter la découverte de données au niveau de l'UE.



Création d'un **formulaire unique** de demandes d'accès pour simplifier les procédures, très variables entre les pays européens, et réduire les délais d'accès. Ce formulaire a permis la mise en place d'un **portail de demande d'accès au niveau européen**.



Conception et déploiement d'une **infrastructure connectant les différents pays** ainsi que la Commission européenne dans un réseau sécurisé d'échange d'informations. Ce réseau permet la création de services à travers l'interaction avec le **portail central européen**.



Les 5 cas d'usage ont démontré la **faisabilité et valeur ajoutée** de l'utilisation de données au niveau européen pour différents types d'usage. Leur retour d'expérience est un atout précieux pour poursuivre la création de l'EHDS.



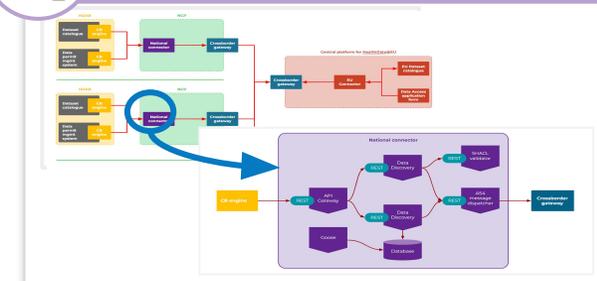
Standard de métadonnées Health DCAT-AP



3 ensembles de recommandations suite aux cas d'usage



Infrastructure IT HealthData@EU



Formulaire unique de demande d'accès



# Objectif 3.2 - Renforcer la présence du HDH au niveau européen et international (2/3)

## Réalisé 2024

- 🔍 Contribution au lancement des travaux de l'**action conjointe TEHDAS II**, pilotée par Sitra (Finlande).
- ✅ Lancement des travaux du **Direct Grant French\_HealthData@EU** pour préparer la mise en œuvre de l'EHDS à l'échelle nationale en France et renforcer les connexions avec l'écosystème des producteurs de données (19 partenaires français).
- ✅ Lancement du **projet européen QUANTUM**, réunissant 30 partenaires de 14 pays, pour développer un **label de qualité** sur l'utilisation secondaire des données de santé.
- ✅ Contribution à la **Community of Practice des HDABs**, mécanisme de coordination entre les HDABs initié par la Commission européenne : co-organisation avec la Commission européenne de la réunion de lancement à Bruxelles en janvier 2024.
- 🔍 Sélection du projet **SHAIPED** candidat à un appel à projets de la Commission européenne. Ce consortium regroupe 30 partenaires issus de 11 pays et est piloté par le HDH. Ce projet mobilise les organismes d'accès aux données de santé pour le développement, test et déploiement de l'IA en santé y compris à travers des cas d'usage.



Commission européenne



## Plan de travail 2025

- Poursuite de la participation à l'action conjointe TEHDAS II, notamment en tant que lead d'un Work Package
- Poursuite des travaux du **Direct Grant French\_HealthData@EU** impliquant notamment le déploiement du catalogue de métadonnées chez les partenaires et le soutien à des actions de standardisation des données.
- Poursuite des travaux dans le cadre du **projet européen QUANTUM** (CHU de Bordeaux, CHU de Toulouse, HCL de Lyon, CHU de Lille, AP-HM et CHU de Montpellier) , en testant et pilotant le label développé en 2024 avec des partenaires français.
- Lancement du **projet SHAIPED** avec un événement de kick-off en mars. Les premiers livrables du projet sont attendus pour 2025, incluant :
  - une **analyse du cadre juridique** applicable aux dispositifs médicaux intégrant de l'IA et une cartographie du paysage européen de certification des dispositifs médicaux,
  - une **description du cycle de vie des données** pour mener un projet intégrant de l'IA dans un dispositif médical.

# Zoom : mettre en place un label de qualité des données de santé au niveau européen avec le projet européen QUANTUM

Le projet **QUANTUM** est lauréat d'un appel à projets européen pour créer un label de qualité et d'utilité des données, dans le cadre de l'Espace européen des données de santé (EHDS). Ce label offrira aux porteurs de projet une meilleure visibilité sur les données exploitables pour leurs initiatives.

Le HDH participe aux *work packages* 1 et 3, portant respectivement sur la **conceptualisation puis le test d'un label de qualité et d'utilité des données**.



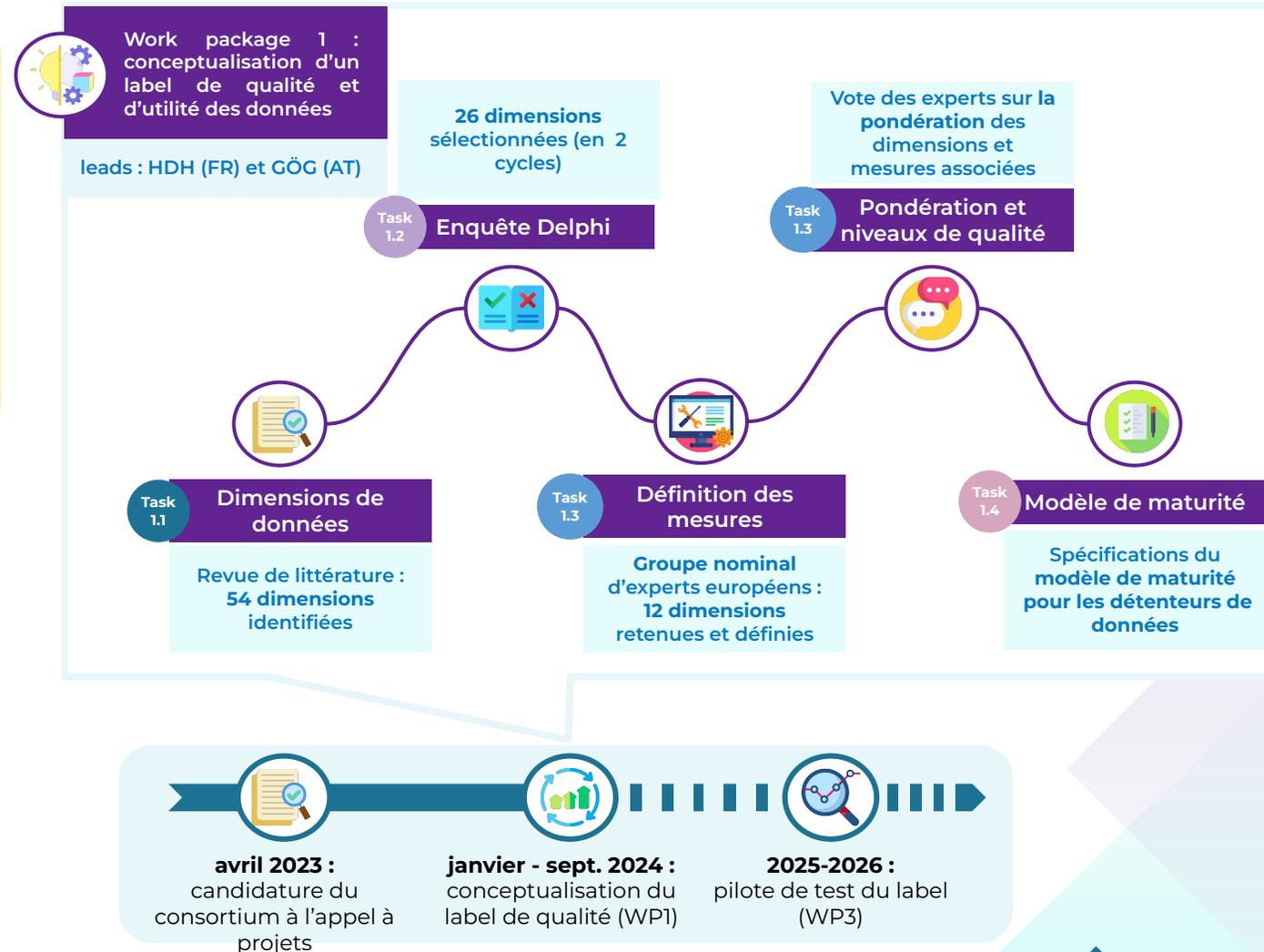
**4 millions d'euros de financement européen** au niveau du consortium.



**35 partenaires européens**, coordonnés par IACS (ES).



**7 partenaires français du HDH** dont l'Inserm, les HCL, la HAS, les CHUs de Montpellier, Bordeaux, Lille et Toulouse.



# Zoom : le projet SHAIPEd vise à développer l'IA en santé en s'appuyant sur les organismes responsables de l'accès aux données (HDAB)

Le HDH pilote le consortium SHAIPEd, qui mobilise l'écosystème français et **tous les organismes responsables de l'accès aux données de santé (HDABs) existants en Europe** pour promouvoir l'IA en santé en UE.

### 3 objectifs clés :

1. Renforcer l'offre de services des HDABs pour favoriser l'usage de l'IA.
2. Faire émerger des synergies entre l'EHDS et l'IA Act.
3. Réalisation de cas d'usage concrets.

4M€

de financements UE

3 ans

30

partenaires  
français et  
européens

11

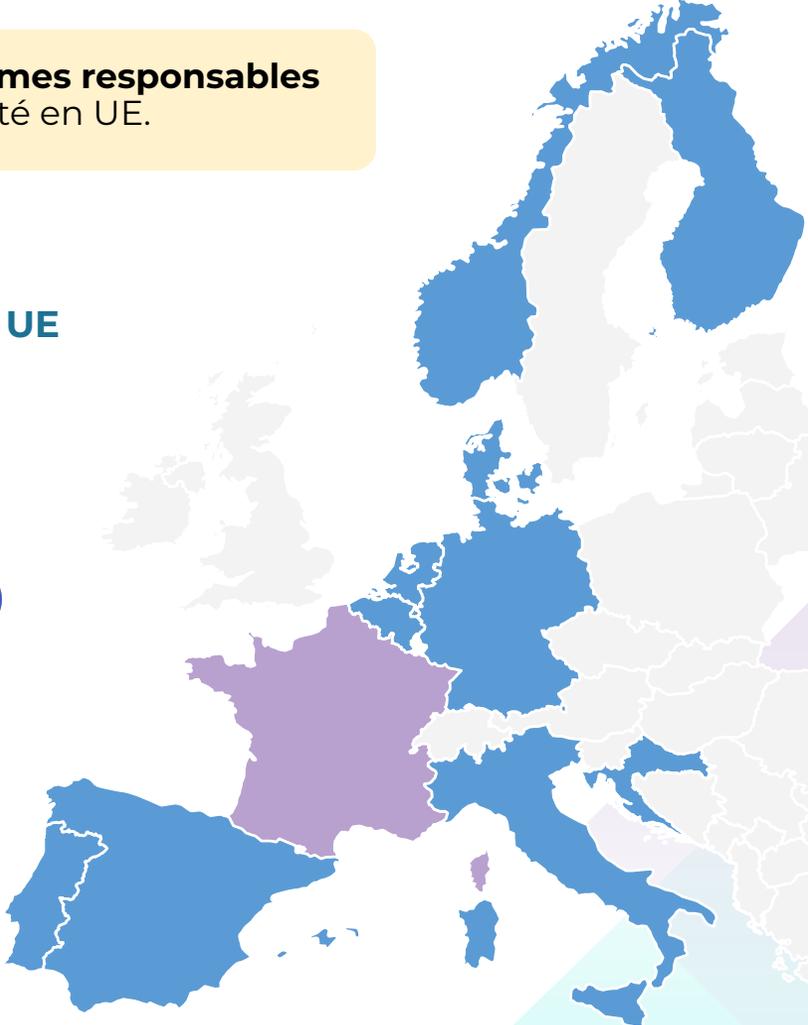
états membres  
impliqués



### 7 partenaires français



### Le HDH chef de file d'un réseau de 11 HDABs



# Zoom : le projet SHAIPEd met en oeuvre 3 cas d'usage renforçant la compétitivité européenne en matière d'IA en santé

## 3 cas d'usage pour tester les services des HDABs pour l'IA en santé :



### Cas d'usage sur la maladie rénale chronique (MRC) :

Exploration de la performance d'un algorithme existant sur des données provenant de divers pays.

Porté par l'hôpital universitaire d'Aarhus (Danemark) & participation du Finnish institute for health and welfare (Finlande), de CKD-rein et de France cohortes (France).



**Cas d'usage en cardiologie :** Évaluation en vie réelle d'un dispositif médical intégrant de l'IA impliqué dans la prise en charge de l'insuffisance cardiaque.

Porté par Implicit (France) & participation du Health Data Hub et du Health Data Lab (Allemagne).



**Cas d'usage en oncologie :** Mise en place d'une cohorte de validation pour évaluer la performance des outils d'IA dans la détection de métastases à partir d'images de scanner du poumon et de nodules à partir de mammographies.

Porté par le Centre Léon Bérard et le Health Data Hub (France)



### Partenaires du cas d'usage



### dont 2 start-ups pour évaluer des outils d'IA



Les cas d'usage répondent à l'enjeu de **renforcer la compétitivité européenne**, en ciblant des problématiques actuellement identifiées pour le **déploiement de l'IA en Santé en Europe**, principalement s'agissant des **dispositifs médicaux et des procédures réglementaires** (accès au marché, certification, remboursement, etc.).

# Objectif 3.2 - Renforcer la présence du HDH au niveau européen et international (3/3)

## Réalisé 2024

- 🔍 Intégration du HDH dans le **réseau DARWIN EU®**, **coordonné par l'Agence européenne du médicament** (EMA) afin de réaliser des études mobilisant des données de vie réelle pour soutenir l'EMA dans l'exercice de ses missions de surveillance des médicaments et soumission de plusieurs demandes d'autorisation unique.
- ✅ Poursuite des travaux au sein du projet européen **TEF-Health**, visant à développer des services de test et de validation de dispositifs médicaux d'IA. Le HDH participe notamment à l'identification de cas d'usage pour le projet.
- ✅ Soutien à **Santé publique France (SPF)** dans la préparation et le **lancement du projet Orchidée** visant le développement d'indicateurs de surveillance d'infections respiratoires. Le HDH a d'ores et déjà apporté une première contribution à travers une analyse des scénarios réglementaires et la mise en place du comité opérationnel du projet (CoTech)
- ✅ Poursuite de l'exploration de partenariats internationaux par le biais notamment de l'organisation **d'une délégation au Royaume-Uni** pour rencontrer l'écosystème britannique des données de santé et le développement d'un **partenariat France-Japon** ayant permis le dépôt d'un projet franco-japonais à un appel à projets de l'Inca sur la biologie du cancer.



## Plan de travail 2025

- Réalisation des premières études dans le cadre du **réseau DARWIN EU®**.
- Soutenir une ou plusieurs start-ups dans le cadre du **projet européen TEF-Health**.
- **Pilotage des travaux** du *work package* (WP) 3 (volet technique) **du projet Orchidée** pour accompagner le déploiement d'une première version d'indicateurs de surveillance d'infections respiratoires aiguës au sein des 25 établissements partenaires.
- Poursuite des **rencontres et délégations avec des acteurs internationaux d'intérêt** et en particulier, co-organisation d'un événement au Québec avec FRQS.
- Sélection d'un **projet de recherche IA en santé** en collaboration avec Technion/Rambam en Israël et **exploration d'un partenariat avec le MIT**.
- Réédition de **l'école d'été AI4Health co-organisé par les instituts 3IA** (Prairie, MIAI Grenoble, 3IA Côte d'Azur) permettant de faire intervenir des acteurs de renom

# Zoom : le HDH accompagne 4 études dans le cadre du réseau DARWIN EU®, visant à mobiliser des données de vie réelle au niveau européen



DARWIN EU® (*Data analysis and real world interrogation network*) est un **réseau coordonné** par l'Agence européenne du médicament (EMA) **visant à favoriser l'utilisation de données de vie réelle** afin d'étudier l'utilisation, la sécurité et l'efficacité des médicaments et des vaccins pour l'utilisation humaine.

En tant que **data partner** depuis juillet 2023, le HDH est responsable de **4 études** mobilisant des données de la **base principale du SNDS**.

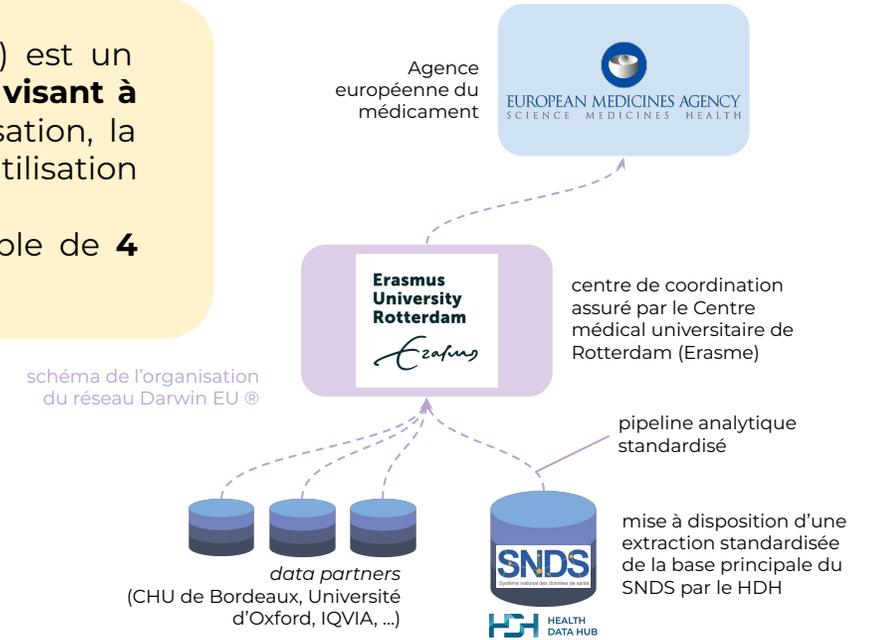


L'accès à la base principale du SNDS permettra la réalisation de 4 études

2 études, portant sur la prévalence et **l'incidence de pathologies**, ainsi que sur **l'incidence d'utilisation des médicaments** en France sont en cours d'autorisation CNIL.

2 autres études, portant sur la caractérisation des individus présentant une **condition médicale incidente**, ainsi que ceux **initiant des traitements médicamenteux** sont en cours d'instruction par le CESREES. D'autres possibilités d'études sont en cours de cadrage.

Ces études observationnelles se basent sur une **cohorte observationnelle mobilisant les données concernant 12 millions d'individus**, qui seront mises à jour annuellement.



**Assurer l'accès continu à une base de données à jour pour répondre aux missions de l'EMA**

Ces études font l'objet de **décisions uniques**, ce qui signifie qu'une fois autorisées, **l'EMA pourra utiliser ces données à chaque fois que cela est nécessaire**, pour la même finalité, tout au long du cycle de vie d'un médicament, sans autorisation supplémentaire. Un facteur d'efficacité primordial au regard des missions de l'EMA.

# Zoom : prévenir les maladies émergentes grâce au réseau de surveillance Orchidée mené par Santé publique France



Le **consortium Orchidée**, mené par Santé publique France et regroupant 28 partenaires nationaux, est lauréat d'un appel à projets de la Commission européenne visant à améliorer et renforcer les systèmes de surveillance à l'échelle du territoire.

L'objectif est de permettre aux autorités de disposer **d'indicateurs épidémiologiques de qualité**, grâce aux **données de santé des hôpitaux**, afin de connaître la situation épidémiologique en temps proche du réel.

## Le HDH est en charge d'un des axes de travail



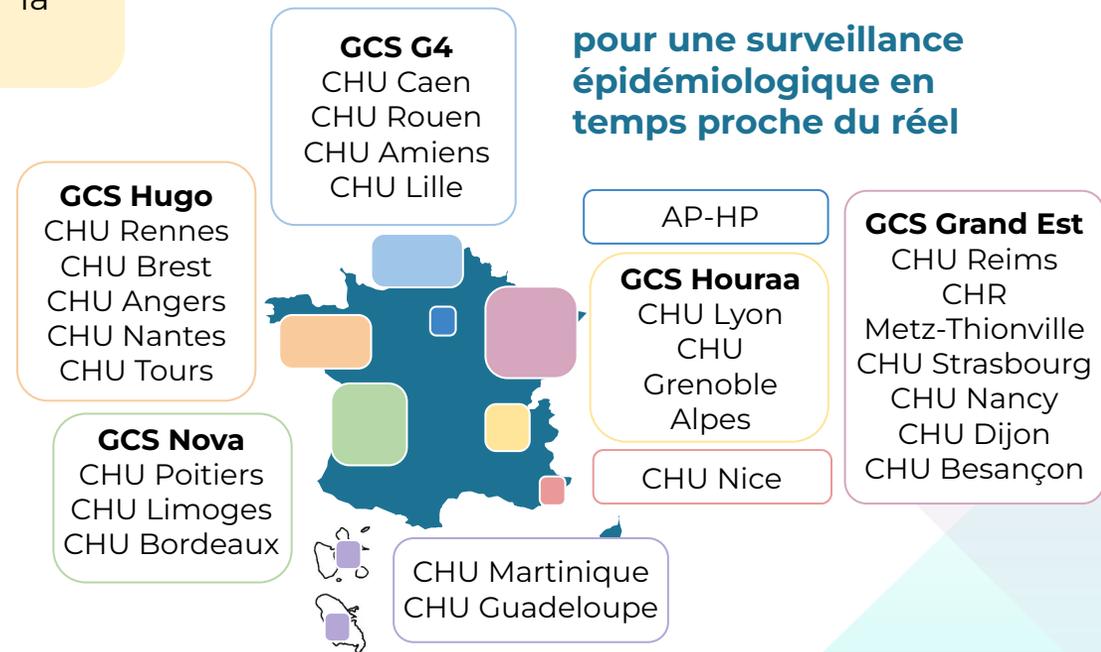
### WP3 : soutien des entrepôts de données de santé hospitaliers (EDSH) à la production des indicateurs

Dans ce cadre, le HDH organise :

- l'animation d'un comité technique à destination des EDSH dont l'objectif est de produire un kit technique pour faciliter le déploiement de scripts standardisés au sein des différents établissements partenaires
- la coordination et l'accompagnement à la standardisation et l'implémentation des indicateurs auprès des établissements

## la remontée d'indicateurs de 25 CHU

pour une surveillance épidémiologique en temps proche du réel



# Zoom : le Sommet pour l'action sur l'IA a abordé les enjeux de santé publique avec la participation du Health Data Hub

Un **Sommet mondial pour l'action sur l'intelligence artificielle s'est tenu les 10 et 11 février 2025 à Paris**, dans la continuité des sommets mondiaux sur l'IA de Séoul et de Bletchley Park.

5 thèmes principaux ont été abordés par le Sommet, parmi lesquels les enjeux de **l'IA au service de l'intérêt public**. Le Health Data Hub a contribué à porter la voix de la France lors de ce sommet à travers un data challenge en santé et un événement d'une journée co-organisé avec la Commission européenne.

## Le HDH a participé à deux initiatives dans le cadre du Sommet :



Depuis 2020, le **Health Data Hub** soutient l'organisation de Data Challenges sur des thématiques de santé à travers le dispositif "**Data Challenge en santé**", une initiative du plan **France 2030**. Près d'une dizaine de projets ont été soutenus jusqu'à présent et près de 2000 équipes ont participé. Ces projets favorisent une **collaboration interdisciplinaire** d'échelle **internationale**, le tout dans une démarche d'**ouverture de la science**.

**Le data challenge Cytologia** organisé par le Groupe Francophone d'Hématologie Cellulaire et Algoscope vise au développement d'une solution d'IA pour standardiser l'analyse des frottis sanguins, afin de diagnostiquer rapidement et avec précision des pathologies sanguines. Il est soutenu par le Sommet pour l'action sur l'IA.

**Animation d'une journée dédiée à l'IA en santé.** La Commission européenne et le HDH ont animé une session intitulée : "**IA en santé : des solutions pratiques pour des enjeux mondiaux urgents**", faisant intervenir citoyens, chercheurs et praticiens, décideurs ou industriels impliqués sur ces enjeux.

# Objectif 3.3 - Réunir les conditions d'un partage efficace en termes de standards, gouvernance d'accès et de valorisation dans un contexte national et européen (1/3)

## Réalisé 2024

- 🔍 Contribution à l'élaboration de la "**stratégie interministérielle** pour construire notre patrimoine national des données de santé".
- ✅ Contribution aux travaux européens d'harmonisation de la gouvernance des données de santé avec notamment :
  - La publication des livrables en lien et prévus dans le cadre du projet **HealthData@EU Pilot** dont le formulaire de demande d'accès et les recommandations issues des ateliers sur les portails de transparence.
  - L'initiation des travaux prévus dans le cadre de Tehdas 2 tels que la réalisation d'une étude comparative **des modalités tarifaires en Europe** et l'animation d'un groupe de travail pour établir des lignes directrices pour la tarification dans le cadre de l'EHDS.



## Plan de travail 2025

- Contribution à la **mise en oeuvre de la Stratégie interministérielle pour construire notre patrimoine des données de santé** pilotée par la DNS.
- **Poursuite de l'élaboration de modèles européens** pour le formulaire de demande d'accès et les permis d'accès dans le cadre de l'action conjointe TEHDAS II.

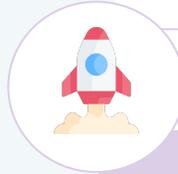


# Objectif 3.3 - Réunir les conditions d'un partage efficace en termes de standards, gouvernance d'accès et de valorisation dans un contexte national et européen (2/3)

## Réalisé 2024

🔍 Contribution et animation des différents travaux du **comité stratégique des données de santé** visant à favoriser l'harmonisation de la gouvernance des données de santé et notamment :

- Élaboration de **principes d'accès prioritaire au catalogue** du SNDS pour les détenteurs de données y contribuant ainsi que **des grilles tarifaires**.
- Élaboration de **principes de contractualisation et de tarification** pour les données hospitalières. Mise en place de comité de suivi de l'implémentation des grilles tarifaires et lancement de nouveaux travaux visant les contrats de fourniture de données de santé impliquant une **collaboration scientifique**
- Élaboration de spécifications technico-fonctionnelles pour la **standardisation de l'ensemble minimal de données** identifié en 2023 pour les entrepôts de données de santé hospitaliers et lancement d'un comité de suivi pour accompagner les acteurs dans la mise en oeuvre.
- Cadrage de nouveaux travaux à lancer en 2025 notamment s'agissant du **financement pérenne, de l'application du référentiel de sécurité du SNDS, ou de l'offre de service des entrepôts de données de santé** ou encore l'**enrichissement de la base principale**.



## Plan de travail 2025

- **Approbation en début d'année 2025 des travaux du comité stratégique des données de santé** en plénière
- Lancement ou poursuite des nouveaux groupes et comités de suivi pour suivre l'implémentation des livrables approuvés.

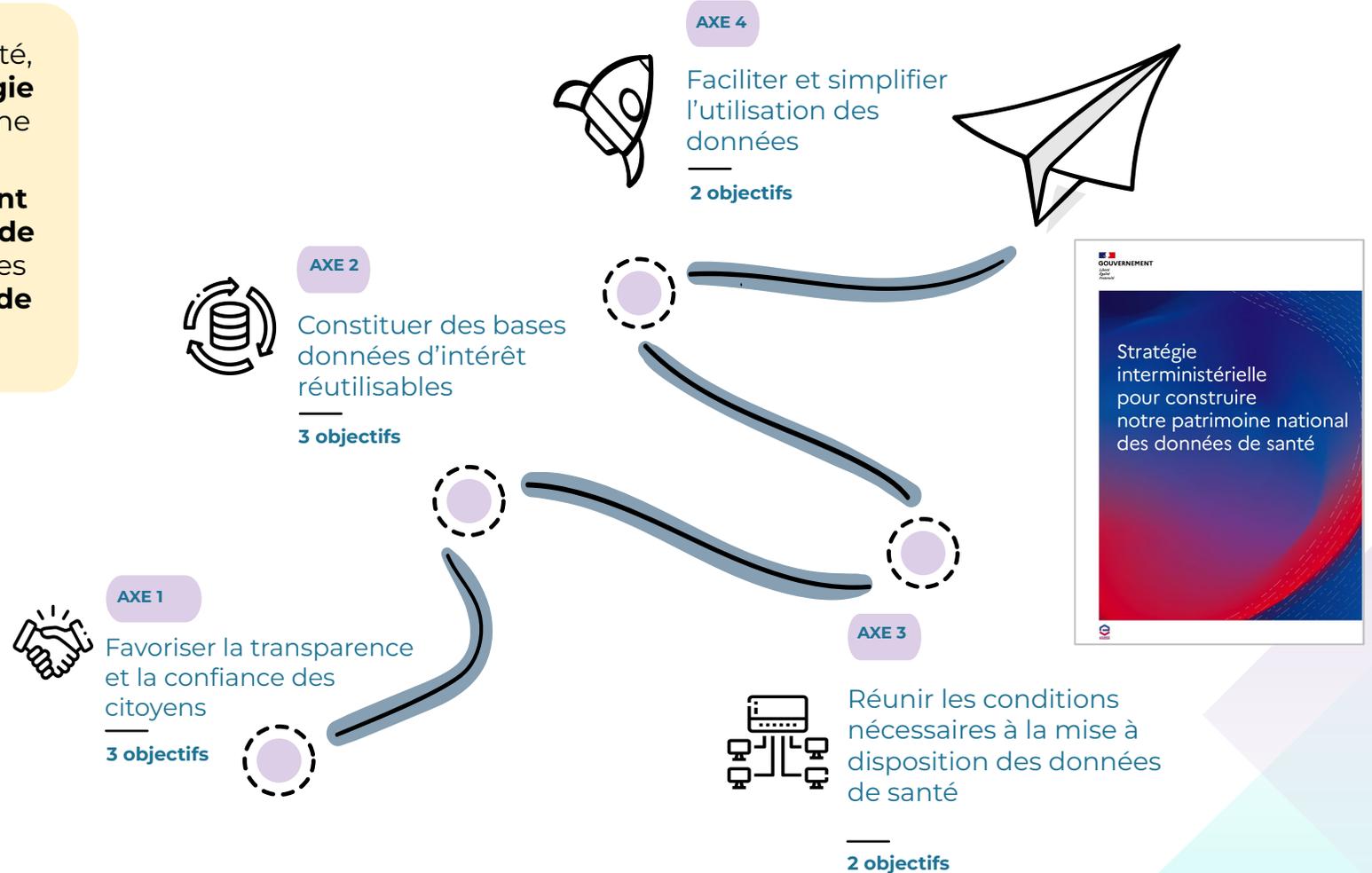
Un groupe de travail sur le financement pérenne des entrepôts de données de santé a été lancé dans la lignée du financement d'amorçage prévu par l'appel à projets France 2030. Il vise à définir les contours d'une mesure potentielle en termes de périmètre, modalités de calcul ou encore prérequis.

# Zoom : une stratégie nationale pour l'utilisation secondaire des données de santé

En tant qu'opérateur national des données de santé, le HDH est l'une des parties prenantes de la **stratégie interministérielle** pour construire notre patrimoine national des données de santé.

Celle-ci vise à mener les **chantiers permettant notamment de préparer la mise en oeuvre de l'EHDS**. Pour cela, la stratégie capitalise sur les **travaux du comité stratégique des données de santé**, auxquels le HDH contribue.

- L'harmonisation des pratiques de gouvernance et de gestion des données de santé et notamment hospitalières,
  - l'accélération de la mise à disposition des données,
  - l'identification et la mise en place des organismes responsables de l'accès aux données de santé (HDABs), etc.
- constituent quelques actions auxquelles le HDH contribue activement.



# Zoom : l'action conjointe TEHDAS 2 vise à mettre en oeuvre l'Espace européen des données de santé

Après la fin de TEHDAS 1, visant à préparer le règlement d'Espace européen des données de santé, l'action conjointe TEHDAS 2 se concentre sur la **mise en œuvre de l'EHDS avec deux objectifs** :

- Fournir des **lignes directrices** aux HDABs, aux détenteurs des données et aux utilisateurs des données.
- Élaborer des **spécifications techniques** pour la mise en œuvre de l'EHDS.



**29** pays  
**6** millions €



**WP4**

**Modèles de collaboration**

HDH : contributeur

**Implication des parties prenantes** (groupes d'experts nationaux, forums, ...).  
Lignes directrices concernant les **redevances** et **sanctions**, accès et transfert des données au niveau international.

**WP5**

**Découverte des données et métadonnées**

HDH : contributeur

**Catégories minimales de données de santé** pour l'usage secondaire,  
Descriptions et spécifications techniques des **catalogues de métadonnées et labels de qualité** (lien QUANTUM).

**WP6**

**Accès aux données**

HDH : contributeur

Rendre les données de santé disponibles, **garantir l'accès aux données**.  
**Procédures et formats** relatifs à la délivrance des autorisations d'accès aux données via le test du formulaire de demande d'accès commun et création de guide d'utilisation pour les Data Users



**WP7**

**Traitement sécurisé**

HDH : lead

**Environnements de traitement sécurisés** (interopérabilité, sécurité des données), CGU de la plateforme.  
Enjeux de **minimisation, d'anonymisation, de pseudonymisation, d'appariement ou de données synthétiques**.

**WP8**

**Au service des citoyens**

HDH : contributeur

Obligations envers les **personnes physiques (opt-out...)**.  
Obligations relatives aux résultats de recherche (**répertoire public, etc.**).



**31 janvier 2024**  
Soumission de la candidature

**1 mai 2024**  
Lancement des travaux

**déc. 2026**  
Fin des travaux

# Objectif 3.3 - Réunir les conditions d'un partage efficace en termes de standards, gouvernance d'accès et de valorisation dans un contexte national et européen (3/3)

## Réalisé 2024

- ✓ Un certain nombre de travaux menés à l'échelle européenne et nationale concernent les enjeux de standardisation, majeurs pour un travail collaboratif quelle que soit l'échelle retenue (locale, fédérée, centralisée etc.) :
  - ✓ Construction de spécifications technico-fonctionnelles de **standardisation** dans le cadre d'un groupe de travail du Comité stratégique des données de santé.
  - ✓ Poursuite de la réalisation des **alignements des terminologies**, en lien avec l'Agence du Numérique en Santé, notamment dans le cadre de la standardisation de la base principale du SNDS et de la mise en oeuvre de l'entrepôt EMC2.
  - ✓ Mise en ligne **d'une documentation sur les standards** de données de santé.



## Plan de travail 2025

- Animation d'un **comité de suivi** à la suite du groupe de travail relatif aux standards des données du socle commun mené par le Comité stratégique des données de santé pour accompagner les acteurs dans son implémentation.
- Accompagnement des partenaires du projet French\_HealthData@EU **dans la standardisation et la mise en qualité de leurs données.**
- Poursuite des travaux pour **l'omopisation de l'EDS EMC2 et de la base principale**, notamment dans le cadre de la mise à disposition pour les utilisateurs, ainsi que pour les études réalisées dans le cadre de la participation du HDH au réseau européen Darwin EU®.
- Dans le cadre du projet Orchidée, des travaux d'alignement pourront être pilotés dans l'optique de **produire des indicateurs standardisés.**
- Participation du HDH dans une **action conjointe européenne autour du dépistage de cancers**, dans le groupe de travail chargé de la standardisation des données dans ce domaine.

# Objectif 3.4 - Réunir les conditions d'un partage efficace en termes de norme de sécurité, souveraineté dans un contexte national et européen

## Réalisé 2024

- ✓ Élaboration, en collaboration avec le ministère de la Santé et de l'Accès aux soins, d'un **guide portant sur la refonte du référentiel de sécurité applicable au SNDS**, suite à l'arrêté du 6 mai 2024.
- ✓ Lancement d'un groupe de travail du Comité stratégique des données de santé pour identifier les **besoins d'accompagnement** des acteurs concernés par la refonte du référentiel de sécurité du SNDS.
- ✓ Poursuite **des rencontres avec les industriels** et veille sur les avancées du cloud souverain.
- ✓ Poursuite du travail avec la DINUM autour du **démonstrateur cloud de confiance** visant à soutenir le développement des services manquants. Un appel à manifestation d'intérêt a été lancé en mars 2024 pour accompagner les entreprises souhaitant proposer une offre cloud et d'IA répondant aux besoins de l'Etat.
- ✓ Travaux dans le cadre de la construction de la stratégie interministérielle pour construire notre patrimoine national des données de santé pour l'instruction d'une **solution transitoire**.



## Plan de travail 2025

- **Contribution aux travaux du comité d'audit** du SNDS en matière de partage des bonnes pratiques sur la mise en oeuvre des exigences du référentiel de sécurité.
- Animation du groupe de travail du Comité stratégique des données de santé relatif au référentiel de sécurité du SNDS et **conception d'une offre de service** le cas échéant.
- Poursuite de la **veille sur les avancées du cloud souverain**, et prise en compte de la publication du décret d'application de la loi du 21 mai 2024 visant à sécuriser et à réguler l'espace numérique (SREN).
- Poursuite des travaux préparatoires à la **migration sur une nouvelle infrastructure cloud** et notamment la mise en oeuvre de la **solution intercalaire**, laquelle fera l'objet d'une information régulière lors du franchissement des grands jalons.



#4

## Grande orientation

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



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

L'écoute, l'information, la formation et les partenariats avec les associations d'usagers du système de santé, les professionnels de la santé et de la recherche constituent une priorité pour permettre à l'ensemble des acteurs de s'adapter aux enjeux de la réutilisation des données de santé et de s'y impliquer.

# Objectif 4.1 - Informer largement sur les enjeux de la réutilisation des données de santé (1/2)

## Réalisé 2024

- ✔ **Publication de plusieurs articles** dans des ouvrages et revues, **et multiples interventions orales** dans des cours et séminaires.
- ✔ **Diffusion d'une infolettre citoyenne mensuelle HDH/France Assos Santé.**
- ✔ **Poursuite des webinaires mensuels "Les mardis de la donnée de santé".**
- ✔ **Lancement de travaux pour décrire les indicateurs / apports des bénéfices concrets de l'usage secondaire des données de santé**, en collaboration avec la Drees, et France Assos Santé.
- ✔ **Mise en téléchargement de deux affiches de sensibilisation** sur la seconde vie des données de santé pour les établissements de santé.
- ✔ **Élaboration d'une première version du set minimal de connaissances** sur la seconde vie des données pour les citoyens.



## Plan de travail 2025

- **Poursuite des initiatives d'information** : infolettres, publications dans des revues, interventions dans des cours, séminaires...
- **Poursuite de travaux autour des indicateurs mesurant le bénéfice des projets.**
- **Lancement d'une "mallette"** sur les données de santé à l'instar de la mallette "cybersécurité".
- **Publication d'un set minimal de connaissances** ayant fait l'objet d'une concertation.



# Zoom : des outils de sensibilisation pour construire une culture des données de santé

Initialement exposées au CHU de Rouen Normandie, deux affiches de sensibilisation à la seconde vie des données ont été conçues lors d'ateliers participatifs, réunissant citoyens et experts. **Les affiches sont maintenant téléchargeables et accessibles à tous. Elles peuvent constituer des outils pour les entrepôts de données de santé qui souhaitent améliorer l'information auprès du public.**

Mise en ligne, par France Assos Santé et le HDH, d'une **campagne de communication pédagogique sur les réseaux sociaux**, comprenant des vidéos et des supports pédagogiques.



Des affiches réalisées en partenariat avec...



- Où sont mes données de santé ?
- Mon identité est-elle partagée avec mes données ?
- Qui a le plus gagné quand mes données de santé sont partagées ?

# Zoom : les bénéfices de l'usage secondaire des données de santé

Éveiller les consciences des citoyens à propos de l'utilisation des données de santé pour la recherche et de leurs bénéfices constitue un enjeu fort, et un défi méthodologique et sociologique. C'est pourquoi France Assos Santé, la Drees, et le HDH ont commencé en 2024, et poursuivront en 2025, à explorer les nouvelles possibilités de **produire des narrations de ces bénéfices**. Dans un premier temps, ce travail a été réalisé grâce aux échanges avec une partie des porteurs de projets que le HDH accompagne.



## Les bénéfices

Innovation technologique

Éthique et régulation

Formation et éducation

Recherche scientifique et médicale

Engagement et information des patients

Optimisation des coûts et de l'efficacité des systèmes de santé

Évaluation de la qualité des soins et des systèmes de santé

Amélioration des politiques de santé publique

Amélioration des soins de santé et des pratiques cliniques



## Quels bénéfices sont les plus couverts ?

Après l'élaboration d'une liste de bénéfices (qui pourra évoluer au fil des échanges), il s'agira d'identifier, ceux que chaque projet couvre (au delà du "seul" bénéfice exposé lors du dépôt de la demande d'accès aux données). Après analyse de tous les projets, cela mettra en visibilité les bénéfices les plus couverts, et ceux qui le sont moins.

## Jusqu'où narrer les bénéfices d'un projet ?

La démarche ne vise pas l'évaluation des projets, mais à raconter leurs impacts (directs et indirects / sur plus ou moins long terme). Ce qui engage de confronter les perceptions des différents acteurs : porteur de projet, patient, professionnel de santé, et institutions (HDH, CNIL, CNAM...). *Un projet s'arrête-t-il à la seule découverte d'un algorithme et jusqu'où peut-on le narrer ?*

## Comment proposer des narrations des bénéfices ?

L'ambition est de restituer des récits de ce que chaque bénéfice signifie concrètement, appuyé par des exemples de projets (et non de narrer, uniquement, ce que tel ou tel projet va apporter comme bénéfice). Le défi n'est pas seulement la vulgarisation des bénéfices, mais de proposer des narrations porteuses d'espoir et convaincantes pour les citoyens, dans un langage non technique.



# Objectif 4.1 - Informer largement sur les enjeux de la réutilisation des données de santé (2/2)

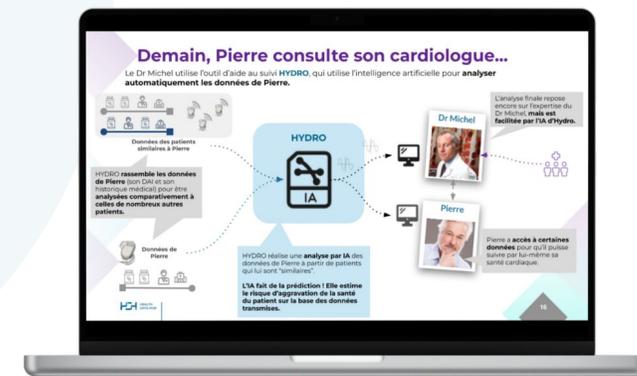
## Réalisé 2024

- ✓ **Publication des premiers résultats issus du répertoire public des projets.**
- ✓ **Évolution du répertoire public pour en faire un portail de transparence pour les projets**, prenant notamment en compte les retours de la CNIL.
- ✓ **Enrichissement des formations citoyennes avec de nouveaux modules :**
  - entrepôts de données de santé : Information/consentement et sensibilisation,
  - Données génomiques : pour le soin, et pour la recherche,
  - SNDS : Référentiel de sécurité
  - Cybersécurité dans le monde de la recherche ; Glossaire ; IA...
- ✓ **Préparation d'une première version d'un référentiel socle de compétences** sur la seconde vie des données.
- ✓ **Poursuite de l'implication auprès du consortium SN@SU** pour produire des contenus pédagogiques.



## Plan de travail 2025

- **Finalisation de l'implémentation des évolutions du répertoire public** pour les projets et intégration des entrepôts de données de santé.
- **Création de nouveaux modules** dont : "Santé mentale", "Données en vie réelle".
- **Consolidation, concertation puis publication du référentiel socle** de compétences sur la seconde vie des données.
- **Poursuite de l'implication auprès du SN@SU** pour des contenus pédagogiques.



# Zoom : le répertoire public des projets

**+9800**  
fiches projets publiées  
sur le répertoire...

...dont **1300**  
créées en 2024

## Tous les projets

Télécharger ↓

Filtrer VALIDER

Rechercher un intitulé 🔍

10 par page ▼ Les + récents ▼

|                                    |  |
|------------------------------------|--|
| Objectifs de l'étude ▼             | Estimation d'incidence et de prévalence de l'utilisation des médicaments dans la population générale en France dans le cadre du projet DARWIN EU® →  |
| Domaines Médicaux ▼                | OBJECTIFS <b>PRÉVENTION ET TRAITEMENT</b> (+1) DOMAINES <b>UROLOGIE, ANDROLOGIE ET NÉPHROLOGIE</b>   |
| Catégories des données utilisées ▼ | CATÉGORIES <b>INFORMATIONS RELATIVES AUX PATHOLOGIES DES PERSONNES CONCERNÉES</b>  |
| Type de responsable traitement ▼   | Etude portant sur le développement d'un outil numérique d'évaluation de la douleur aiguë par reconnaissance de l'expression faciale, intitulée "ETUDE-DEF-1" →   |
| Statut ▼                           | OBJECTIFS DOMAINES CATÉGORIES  |
|                                    | Etude Combo : Evaluation des nouvelles combinaisons de molécules en immuno-oncologie à partir des données de vie réelle. →   |
|                                    | OBJECTIFS <b>COMPRÉHENSION DES MALADIES</b> (+1) DOMAINES <b>CANCÉROLOGIE</b> CATÉGORIES <b>INFORMATIONS RECUEILLIES À L'OCCASION D'ACTIVITÉS DE PRÉVENTION, DE DIAGNOSTIC, DE SOINS OU DE SUIVI SOCIAL ET MÉDICO-SOCIAL</b> |
|                                    | Projet PsychoMG : Evaluation de l'impact d'un site internet d'aide à la prise en charge des principales pathologies mentales rencontrées en médecine générale →  |
|                                    | OBJECTIFS <b>PRÉVENTION ET TRAITEMENT</b> (+3) DOMAINES <b>PSYCHOLOGIE ET PSYCHIATRIE</b> CATÉGORIES <b>INFORMATIONS RELATIVES AUX PATHOLOGIES DES PERSONNES CONCERNÉES</b>  |

# Zoom : Les programmes de la formation citoyenne

## Base principale du SNDS

- Démarche de recherche
- Accès au SNDS
- Ressources pour les associations
- Chaînages de données
- Projet Hydro
- IA en santé
- Open data
- Projet Vivre-Covid19
- Cybersécurité dans le monde de la recherche



## Entrepôts de données de santé (EDS)

- Importance des EDS
- Construire un EDS
- Les humains derrière les EDS
- Information/consentement des patients dans les EDS
- Sensibilisation des patients dans les EDS



## Types de données

- Santé-environnement : c'est quoi ?



- Données génomiques pour le soin
- Données génomiques pour la recherche



Des formations citoyennes réalisées en partenariat avec....



**+2000** participants aux webinaires "Les mardis de la donnée de santé" en 2024

# Objectif 4.2 - Écouter et impliquer la société civile (1/2)

## Réalisé 2024

- ✓ **Clôture du débat sur l'IA en santé via l'organisation d'un événement au CESE** réunissant des lycéens et des membres de la CNIL, de la DNS, de la DGOS et du CNPEN.
- ✓ **Publication des résultats de la 2e édition du baromètre de connaissance** sur les données de santé, en partenariat avec France Assos Santé, Parisanté Campus, DNS et la CNIL.
- ✓ **Organisation d'un colloque en partenariat avec la Fédération Française des Diabétiques, et l'Espace éthique Ile-de-France**, portant sur l'utilisation des données de santé pour la recherche, avec un focus sur la prévention et le dépistage.
- ✓ **Animation et production en cours d'un livrable du Comité stratégique de filière (CSF) "industries et technologies de santé", visant à mieux sensibiliser la société civile** dans le cadre du groupe de travail Adhésion patients et citoyens avec deux groupes de dialogue.
- ✓ **Réalisation d'une première version de cartographie des "espaces/lieux citoyens"** pour sensibiliser et débattre avec les citoyens.
- ✓ **Réalisation d'une cartographie des acteurs de la médiation numérique.**



## Plan de travail 2025

- **3e édition du baromètre de connaissances** sur les données et leur usage pour la recherche.
- **Publication du rapport du groupe de travail Adhésion patients et citoyens** du CSF "industries et technologies de santé".
- **Finalisation de la cartographie des "espaces/lieux citoyens", et déploiement.**
- **Evolution de la direction en une direction de l'Ouverture.**
- **Organisation du premier colloque de la direction**



# Zoom : les résultats de la seconde édition du baromètre de connaissance

La **deuxième édition du baromètre de connaissances sur les données de santé** a été réalisée en 2024 par le HDH, en partenariat avec France Assos Santé, la CNIL, PariSanté Campus et la DNS. Diffusé auprès d'un échantillon représentatif de 2 000 Français, ce baromètre sera renouvelé chaque année pour suivre l'évolution des connaissances des citoyens sur ce sujet.

## Des Français mieux sensibilisés aux réglementations et à l'usage des données de santé...

85 %

ont entendu parler du **RGPD**.

35 %

savent de quoi il s'agit **précisément** (le RGPD).

83 %

savent ou se doutent que les **données de santé** sont **soumises à une réglementation et à une sécurisation plus stricte** que les autres types de données personnelles.

66 %

savent que leurs **données anonymisées** de santé peuvent être utilisées.

58 %

savent que leurs **données pseudonymisées** de santé peuvent être utilisées.

## ... et qui souhaitent approfondir leur connaissances sur l'utilisation secondaire des données de santé.

32 %

**estiment être suffisamment informés** sur les projets de recherche analysant les données de santé.

56 %

**souhaitent être mieux informés sur l'utilisation de leurs données de santé**. Ils préfèrent majoritairement obtenir ces informations via le site Ameli ou Mon Espace Santé.

53 %

**affirment connaître leurs droits concernant leurs données de santé**. Une grande partie de la population ignore encore les droits dont ils disposent, tels que le droit d'accès, d'opposition ou d'effacement des données.

# Objectif 4.2 - Écouter et impliquer la société civile (2/2)

## Réalisé 2024

- ✓ **Poursuite des groupes de dialogue**, avec les associations de patients et notamment autour de la stratégie nationale des données de santé, et les associations d'étudiants.
- ✓ **Diffusion par le groupe de dialogue étudiants de messages clés sur les données de santé** et la recherche.
- ✓ **Lancement des petits-déjeuners du numérique en santé avec Parisanté Campus et les associations étudiantes**, consacrés pour leur première édition aux métiers de demain avec le numérique, la e-santé / l'IA, et les entrepôts de données de santé.
- ✓ **Lancement de travaux avec l'ISNI** - InterSyndicale Nationale des Internes - **sur la sensibilisation et l'accompagnement à l'usage.**
- ✓ **Intégration d'un engagement dans le nouveau Plan d'action national 2024-2026** : construire et diffuser une culture numérique commune et partagée des données de santé.



## Plan de travail 2025

- **Poursuite des groupes de dialogue** institués via des rencontres régulières.
- **Poursuite des petits-déjeuners du numérique.**
- **Propositions de Café IA avec le CNum.**
- **Intégration d'experts SHS** dans cette nouvelle gouvernance de l'ouverture.
- **Construction d'une offre pour les internes avec l'ISNI et pour les enseignants professionnels de santé.**
- **Mise en oeuvre d'un pilote auprès d'une collectivité territoriale (associant les élus à la démarche) pour réaliser le premier jalon du PAN 2024-2026.**

### USAGE SECONDAIRE des données de santé

La seconde vie des données de santé, c'est le fait de **mettre en commun un ensemble de données** d'un grand nombre de patients, de manière pseudonymisées, pour une utilisation dans des projets d'intérêt public. Les données de tous font la santé de tous !

Le Health Data Hub, en tant que structure publique, accompagne les porteurs de projets pour faciliter l'accès aux données de santé pour les projets d'intérêts publics. 



Et à l'hôpital ? Les données de santé sont stockées dans des "entrepôts de données de santé" (EDS), pour bénéficier à des projets d'innovation et de recherche ou pour faciliter l'activité de l'établissement. L'EDS repose sur une infrastructure technique et sécurisée et une organisation humaine chargée d'alimenter l'entrepôt avec les données.

On parle donc d'utilisation secondaire, car vos données sont réutilisées pour la recherche !



# Objectif 4.3 - Développer les partenariats avec les associations d'usagers

## Réalisé 2024

- ✓ **Réalisation de campagnes de communication en partenariat avec France Assos Santé**, pour sensibiliser les citoyens.
- ✓ **Second accompagnement d'une association de patients dans l'accès aux données du SNDS** (base principale) : projet Errance Cancer Tête et Cou (**ECTC**) porté par Corasso, avec les expertises de l'Inserm et Seintinelles.
- ✓ **Mise en œuvre effective par le HDH et France Assos Santé du GUICHET ASSO - service d'accès sur demande à des indicateurs du SNDS** (base principale) et **diffusion d'une documentation** de type livret et fiches de recueil de besoins.
- ✓ **Finalisation de 2 requêtes de demandes d'accès aux indicateurs du SNDS dans le cadre du GUICHET ASSO :**
  - **AFM Téléthon** (patients bénéficiant d'une ventilation assistée pour trachéotomisés).
  - **Internes en médecine générale de l'Université de Lille** (dépistages des cancers du sein et du col de l'utérus chez les femmes résidant en Hauts-de-France).



## Plan de travail 2025

- **Poursuite de l'accompagnement Corasso pour le projet ECTC.**
- **Réalisation et publication de documents pédagogiques** pour permettre aux autres associations de réaliser des études.
- **Réalisation de retours d'expérience et adaptation du service.**
- **Dans le cadre du GUICHET ASSO, réalisation d'au moins 5 requêtes de demande d'accès aux indicateurs du SNDS**
- **Dans le cadre du GUICHET ASSO, automatisation de requêtes de demande d'accès aux indicateurs du SNDS** (développement de modèles de programmes réutilisables pour les cas d'usage fréquents comme le dénombrement de patients).



# Zoom : le service de guichet des requêtes à la demande pour les associations

France Assos Santé et le HDH ont constitué un guichet pour les requêtes à la demande. C'est un service d'accès sur demande à des indicateurs de la base principale du SNDS destiné aux **associations d'usagers du système de santé**. Ces dernières peuvent accéder à des **données agrégées utiles pour dénombrer et caractériser les patients selon différentes pathologies, des indicateurs chiffrés simples**. Grâce au Guichet, les associations **s'acculturent** aux données de la base principale du SNDS, des propositions de mutualisation des demandes sont réalisées, et les indicateurs sont systématiquement mis à la disposition de tous, en *open source*.

## L'équipe du Guichet des requêtes à la demande

Un **guichet** mis en place par le HDH avec France Assos Santé



## Les outils mis à disposition



Un **livret pédagogique** pour présenter le service



Pour effectuer une requête, une adresse mail unique de contact : **requetes.assos@health-data-hub.fr**

## Les 3 grandes étapes de la requête

**Présentation du service à l'association** : conditions et résultats possibles.

**Utilisation des données** : réception et valorisation des indicateurs.



**Construction de la requête** : accompagnement par l'équipe du guichet pour identifier les indicateurs souhaités.

# Objectif 4.4 - Accompagner les personnes concernées dans l'exercice des droits

## Réalisé 2024

- ✓ **Livraison d'un formulaire centralisé** pour l'exercice des droits des utilisateurs.
- ✓ **Contribution aux travaux nationaux** en lien avec :
  - la consultation portant sur les référentiels simplifiés lancée par la CNIL,
  - la préfiguration de la mise en oeuvre de l'EHDS.

### Explorer et comprendre vos données de santé

The screenshot shows a user interface with three main sections: 'Améliorer la santé' (with a heart icon), 'Données de santé' (with a person icon), and 'Projets de recherche utilisant les données de santé' (with a magnifying glass icon). Below these are 'Règles imposées sur l'utilisation des données de santé' and a 'Tout savoir sur' link.

### L'utilisation des données de santé

Découvrez l'importance de vos données de santé dans la recherche et apprenez comment exercer vos droits concernant ces données



#### Qu'est-ce qu'une donnée de santé ?

Quand vous allez chez le médecin, à l'hôpital ou à la pharmacie, des informations sont enregistrées : par exemple votre âge, votre poids, votre taille, les maladies que vous avez eues, les médicaments que vous prenez. Ce sont des données de santé.

Selon la Commission nationale de l'informatique et des libertés (CNIL) : "Les données à caractère personnel concernant la santé sont les données relatives à la santé physique ou mentale, passée, présente ou future, d'une personne physique (y compris la prestation de services de soins de santé) qui révèlent des informations sur l'état de santé de cette personne."



## Plan de travail 2025

- **Mise en ligne du portail patient : une ressource d'informations pédagogiques** pour les citoyens a été réalisée pour accompagner le formulaire d'exercice des droits.
- **Mise en oeuvre des adaptations nécessaires au formulaire national.**
- **Poursuite des travaux nationaux**



### Choisir une base de données

Veillez sélectionner les droits à exercer dans la ou les bases de données suivantes

#### Base principale du SNDS

Système National des Données de Santé (SNDS) historique

Opposition Accès

#### SI-VIC

Système d'information pour le suivi des victimes d'attentats et de situations sanitaires exceptionnelles (Covid-19)

Effacement



Votre demande a été transmise avec succès à nos équipes !

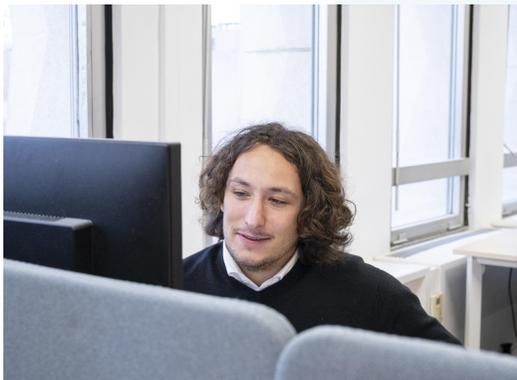
# Programme de travail du secrétariat général 2025



# Programme de travail 2025 du secrétariat général

## Réalisé 2024

- ✓ Renforcement des équipes en continu.
- ⚙️ En continu : revue des avancées du **plan d'actions de mitigation des risques issus de la cartographie des risques.**
- ✓ Participation aux **instances de PariSanté Campus** et aux réunions de préparation **Val de Grâce.**
- ⚙️ En continu : montée en compétences sur les **marchés complexes.**
- ✓ **Formations suivies par la majorité des salariés.**
- ✓ **Elections professionnelles.**
- ✓ **Sensibilisation des collaborateurs sur la problématique du bien-être au travail.**



## Plan de travail 2025

- Création de **6 nouveaux postes**
- Optimisation de la **comptabilité analytique**
- Poursuite du **développement des procédures de gestion et des contrôles internes.**
- Lancement d'une **politique handicap.**
- Lancement d'une **newsletter RH.**
- Application du **Plan de transformation écologique.**

# Health Data Hub - les membres du GIP

## L'État



Direction interministérielle  
du numérique



## Les caisses d'assurance maladie



## Les organismes d'assurance maladie complémentaires



# Health Data Hub - les membres du GIP

## La recherche et l'enseignement



## Les établissements de santé



## Les professionnels de santé



# Health Data Hub - les membres du GIP

## Les usagers



## Les agences



## Les industriels





# HD HEALTH DATA HUB

Suivez-nous sur les réseaux sociaux !