



# HEALTH DATA HUB

## Programme Grand Défi

---

**Appel à projet**  
**- Data Challenge en santé**

Juillet 2022



# Concept et organisation du Data Challenge

## Qu'est ce qu'un Data Challenge?



### UNE PROBLÉMATIQUE

Les Data Challenges accompagnés par le Health Data Hub, visent à répondre à des **problématiques médicales précises** à l'aide de **l'analyse des données** mises à disposition



### UNE BASE DE DONNÉES

Une base de **données de santé anonymisées** et **homogène** doit être constituée pour servir de source d'information à l'analyse. Une étape **d'annotation** des données peut être nécessaire pour constituer une vérité de terrain à fournir aux algorithmes



### DES PARTICIPANTS À LA COMPÉTITION

Des **data scientists** du monde entier accèdent à ces données via une **plateforme dédiée** et développent des **algorithmes d'apprentissage automatique** pour répondre à la question médicale posée

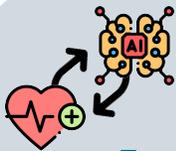


### UNE RÉCOMPENSE

Les meilleures équipes sont **récompensées à la fin de la compétition** par un prix

# Les objectifs de l'appel à projet

Le présent appel à projet vise à sélectionner des projets de Data Challenge qui bénéficieront d'un **accompagnement de dix-huit mois de bout en bout par le Health Data Hub** et d'un soutien financier de Bpifrance obtenu dans le cadre du programme Grand Défi. Cette démarche poursuit les intérêts et objectifs suivants :



Favoriser les interactions entre le monde de la **Santé** et le monde de la **Science des données**

Occasion pour les compétiteurs de se faire repérer par des acteurs majeurs du domaine et **dénicher des talents**



Favoriser le développement d'outils précis **d'assistance aux professionnels de santé**

Démontrer par des résultats rapides le **bénéfice concret de l'utilisation de la donnée**



Promotion de **l'innovation**

**Enrichir le catalogue de données** et d'algorithmes partagés à toute la communauté



# Les projets Data Challenges ciblés

Les projets sélectionnés lors de cet appel à manifestation d'intérêt devront répondre aux critères ci-dessous.



## Intérêt de la question clinique

La question médicale posée doit présenter un intérêt clinique et être originale.



## Pertinence de l'approche prédictive pour y répondre dans le délai imparti

La problématique doit pouvoir être traitée par de l'IA, le score de performance des algorithmes doit être mesurable et les données d'une qualité suffisante pour que les résultats soient fiables.



## Faisabilité de la collecte des données et de leur anonymisation

Les données doivent pouvoir être collectées en nombre suffisant dans un temps imparti à un ou des formats numériques homogènes. Les données doivent pouvoir être totalement anonymisées dans le respect du RGPD.



## Partage des données

Les données collectées et anonymisées dans le cadre du projet seront partagées en Open Data à l'issue du Data Challenge. Des modalités de citation des bases sont toutefois prévues. Les gagnants devront partager leur algorithme en Open Source pour obtenir le prix de récompense.



## Forte implication et faisabilité du projet

Une forte implication de votre part est requise sur toute la durée de l'organisation du data challenge. Plusieurs personnes de votre organisation devront s'impliquer dans le Conseil Scientifique et dans le Comité d'Organisation. Les grands jalons proposés pour le projet doivent le rendre réalisable en environ un an.



## Combinaison de plusieurs spécialités médicales (facultatif)

Le projet peut faire intervenir plusieurs spécialités médicales et combiner des données de natures différentes mais il faudra s'assurer que cela ne remette pas en cause leur caractère anonyme.

# L'accompagnement proposé par le HDH

Le Health Data Hub ou “plateforme des données de santé” est un groupement d'intérêt public mis en place fin 2019 après la publication de la loi “Organisation et Transformation du Système de Santé”. Sa mission principale est de permettre à des porteurs de projets de santé habilités **d'accéder à des données de santé non nominatives** via une plateforme technologique sécurisée tout en respectant les droits des citoyens et dans le but de servir un **intérêt public**.

Parallèlement, il admet un rôle **d'animation de l'écosystème** dans le but de promouvoir les initiatives innovantes et de favoriser le partage de connaissance entre les acteurs. C'est dans ce cadre, que le HDH est amené à organiser des Data Challenges.

**Le HDH accompagnera les projets sélectionnés à toutes les étapes de leur réalisation en apportant des moyens techniques, humains, et financiers adaptés.**



## Définition et cadrage du Data Challenge

- ✓ Aide au cadrage scientifique et organisationnel du projet
- ✓ Aide juridique sur la réglementation applicable au traitement des données de santé
- ✓ Coordination avec des prestataires spécialisés dans l'hébergement de Data Challenges



## Aide à la collecte et à l'anonymisation des données

- ✓ Accompagnement dans la collecte des données
- ✓ Accompagnement dans l'anonymisation des données
- ✓ Support technique dans la phase de labellisation des données
- ✓ Partage des bonnes pratiques



## Gestion de l'infrastructure d'hébergement des données

- ✓ Accompagnement par des data engineers & data managers pour constituer la base de données du Data Challenge
- ✓ Mise à disposition de serveurs sécurisés



## Mise en valeur des résultats

- ✓ Communication (pré et post challenge)
- ✓ Accompagnement dans la publication d'articles scientifiques
- ✓ Support dans les démarches Open Data et Open Source
- ✓ Mise en relation des acteurs de l'écosystème

# Possibilité de financement par le Grand Défi en IA et santé



Jusqu'à 200 000 euros de financement



En complément de l'accompagnement par le HDH, les projets sélectionnés pourront bénéficier d'un **financement du Grand Défi**:  
« Améliorer les diagnostics médicaux à l'aide de l'intelligence artificielle »

Ce régime d'aide peut financer les projets à hauteur de 200K€ maximum sur une période glissante de 3 ans.  
Les projets financés seront invités à déclarer les aides reçues au titre de-minimis au cours des 3 dernières années.

# Un travail collaboratif → Votre rôle

L'organisation du Data Challenge nécessitera une **forte mobilisation de votre équipe.**

Un Data Challenge réussi passe aussi par une **forte implication de l'ensemble des partenaires !**



## Définition et cadrage du Data Challenge

- ✓ Implication de plusieurs personnes dans le Comité d'Organisation et le Conseil Scientifique du Data Challenge sur toute la durée de l'organisation
- ✓ L'organisme organisateur du Data Challenge sera le responsable juridique du projet



## Aide à la collecte et à l'anonymisation des données

- ✓ Motivation et implication des organismes à partir desquelles les données sont collectées
- ✓ Mise en place de la stratégie de la collecte des données (numérisation, anonymisation...)



## Gestion de l'infrastructure d'hébergement des données

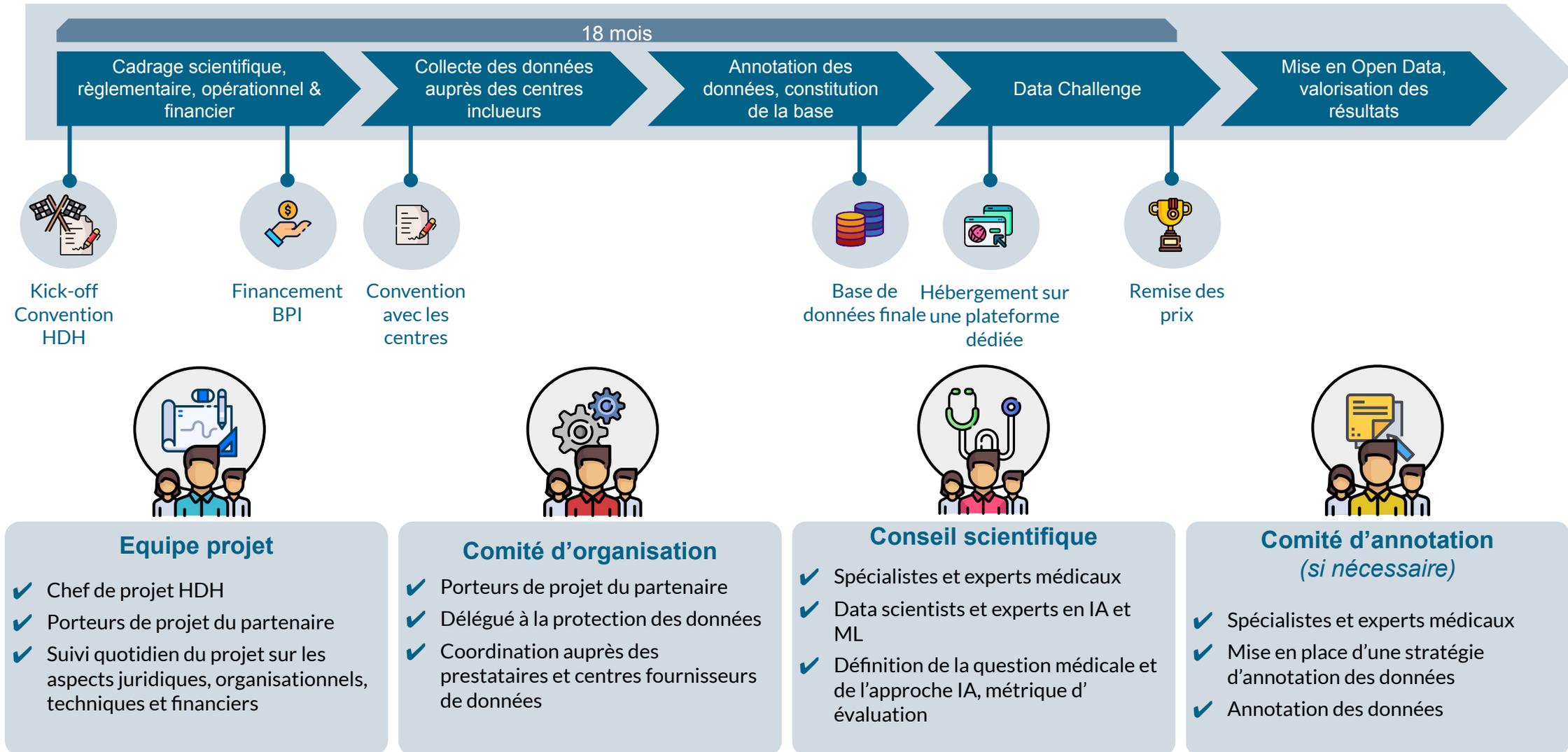
- ✓ Suivi de la gestion des données



## Mise en valeur des résultats

- ✓ Rédaction d'articles dans des revues scientifiques
- ✓ Collaboration avec le HDH dans les travaux de communication
- ✓ Collaboration dans la mise en œuvre de la politique de partage des données

# Organisation type



# Déroulement de la compétition

## Entre 1 et 4 mois de compétition en plusieurs phases...



### INSCRIPTION DES COMPÉTITEURS

Les compétiteurs s'inscrivent sur une plateforme dédiée d'hébergement du challenge.



### PHASE D'APPRENTISSAGE

Les compétiteurs entraînent leur algorithme à partir des données annotées.



### PHASE DE TEST INTERMÉDIAIRE

Les compétiteurs testent leur algorithme sur un jeu de données intermédiaire.



### PHASE DE VALIDATION FINALE

Soumission finale des algorithmes et obtention d'un score final de performances.

# Retour sur le challenge de la société française de pathologie



Data Challenge en pathologie du col de l'utérus - Société Française de Pathologie - Health Data Hub

<https://www.youtube.com/watch?v=Ue0Lt1RKaAE&t=502s>

Le challenge de la SFP en quelques chiffres...

**574** Participants du monde entier

**36** Équipes constituées

Des scores finaux très prometteurs allant jusqu'à **94%**

## Exemple du financement d'un challenge

	exemple de coûts	Porteur du coût
Plateforme Data Challenge (2 mois)	env. 100 000 €	BPI
Récompense gagnants Data Challenge (pris en charge par BPI uniquement pour les lauréats européens)	25 000 €	BPI
Prestation DPO	env. 20 000€	BPI
Rétribution inclusion données (attention ce coût ne peut être pris en charge ni par la BPI ni par le HDH)	env. 25 000 €	porteur de projet
Serveur d'hébergement HDH, disques durs, etc.	env. 1 000 €	HDH
Envoi courrier d'information aux patients	env. 1 000 €	HDH
Outil d'annotation	env. 6 000 €	porteur de projet / HDH
<b>TOTAL</b>	<b>Env. 180 000 €</b>	

## Constitution du dossier candidature

Les candidats à l'AAP devront constituer un **dossier** afin de décrire les différents aspects de leur projet. Le dossier de candidature devra notamment présenter les éléments suivants :



- ✓ Présentation du ou des **porteur(s) de projet** et description de **l'équipe projet**
- ✓ Domaine **médical** et **problématique envisagée**, intérêt et **caractère innovant**
- ✓ Description et justification des **données visées**, source, faisabilité de l'anonymisation
- ✓ Présentation du **design envisagé de la compétition** : approche IA, métrique d'évaluation...
- ✓ Modalités de **partage des données**
- ✓ Plan de **financement**

Afin d'aider les **candidats pré-sélectionnés** à cadrer les différents chantiers de leur projet **avant la phase de sélection finale**, le HDH souhaite leur offrir la possibilité de bénéficier de consultations techniques et juridiques\* dont les objectifs sont les suivants :



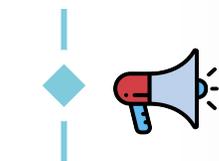
- ✓ Renforcer le **cadre réglementaire et scientifique** du projet
- ✓ Obtenir un **avis sur la faisabilité scientifique** du projet

Les candidats pré-sélectionnés pourront bénéficier de ces consultations pendant la période entre l'annonce de leur pré-sélection et la phase d'audition des projets.

\*Ces consultations seront réalisées par des prestataires externes au Health Data Hub. L'achat de ces prestations faisant l'objet d'un marché public, la mise à disposition des consultations se fera sous réserve d'un marché fructueux.

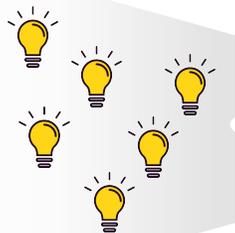
# Modalités de sélection et calendrier de l'AAP

Juillet 2022



Lancement

2 novembre 2022



Clôture de l'AAP

Février 2023



Annnonce des projets  
Sélectionnés

Mars 2023



Kick-off -  
signature des  
conventions de  
partenariat

2024



Data Challenge

## Dépôt des candidatures & consultations juridique et technique

Les porteurs de projets sont invités à déposer leur candidature sur la plateforme de Bpifrance d'ici le 02/11/2022.

## Auditions & pré-sélection

Les candidatures seront analysées par un jury pour établir une présélection. Les candidats présélectionnés désireux de consolider leur projet pourront bénéficier de consultations juridique et technique\*.

## Sélection des projets

Approfondissement des échanges avec les projets présélectionnés pour affiner leur cadrage et s'assurer de leur réelle faisabilité ; puis audition par un jury pour sélection des lauréats.

## Lancement de l'organisation Data Challenges

Lancement des projets sélectionnés. Selon la complexité de la collecte des données, la date de lancement et la durée du Data Challenge varieront.

## Accompagnement Post Challenge

Valorisation des résultats à travers des outils de communication  
  
Mise en Open Data des données & des algorithmes

\*Ces consultations seront réalisées par des prestataires externes au Health Data Hub. L'achat de ces prestations faisant l'objet d'un marché public, la mise à disposition des consultations se fera sous réserve d'un marché fructueux.

# Modalités de réponse à l'AMI



## Date limite de candidature

- 2 novembre 2022



## Plateforme de Candidature

- Déposez votre candidature sur le [site de la BPI](#)



## Contact

- [data.challenge@health-data-hub.fr](mailto:data.challenge@health-data-hub.fr)