



HEALTH
DATA HUB

Présentation de la
plateforme
technologique

Présentation de la plateforme technologique



Présentation de la plateforme technologique du Health Data Hub

La plateforme technologique du Health Data Hub permet de recevoir, de stocker, de traiter et de visualiser des données de santé, pour la réalisation de projets innovants et d'intérêt public dans le domaine de la santé. L'enjeu d'une telle plateforme est de permettre le traitement sécurisé de données hautement sensibles tout en répondant aux besoins fonctionnels de ses utilisateurs projet.

Présentation de la plateforme et de l'espace projet

Qu'est-ce que c'est ?

- Une plateforme sécurisée de mise à disposition des données pour les projets ;
- Un espace projet dédié à chaque projet, mis à disposition par le Health Data Hub, isolé de l'extérieur, constitué d'un ensemble d'outils de traitements de données.

Comment ça fonctionne pour l'utilisateur ?

- Je suis un utilisateur habilité travaillant sur un projet autorisé disposant d'un espace projet ;
- Je me connecte à distance, depuis mon poste de travail, à l'espace projet ;
- Je peux analyser les données sur l'espace projet, sans qu'elles passent sur mon poste local.

Quels avantages ?

- Un espace projet sécurisé accessible à distance, sans transfert des données sur mon poste local ;
- Une flexibilité des outils d'analyse des données ;
- Une flexibilité des capacités de calcul (CPU / GPU).

Quelles contraintes ?

- Une nécessaire homologation de l'espace projet par le Responsable de Traitement ;
- Un catalogue de services disponibles progressivement.

Exemples d'outils qui pourraient être mis à disposition dans l'espace projet

Orchestration de containers

- Ubuntu
- Windows 10

Langages et outils liés

- Python 3 (avec Anaconda)
- R (avec R Open)

Outils de développement

- Jupyter Notebook (noyaux R, Python, Julia, PySpark)
- R Studio Desktop
- Visual Studio Code, Vim, Emacs, Atom
- git

Visualisation et bureautique

- LaTeX
- navigateur (Firefox / Chrome)

Machine Learning & AI

- Packages standard machine learning (R, Python)
- Xgboost + LightGBM
- Weka
- Rattle
- Scikit learn
- Some plots (Python : seaborn, matplotlib, R : ggplot)

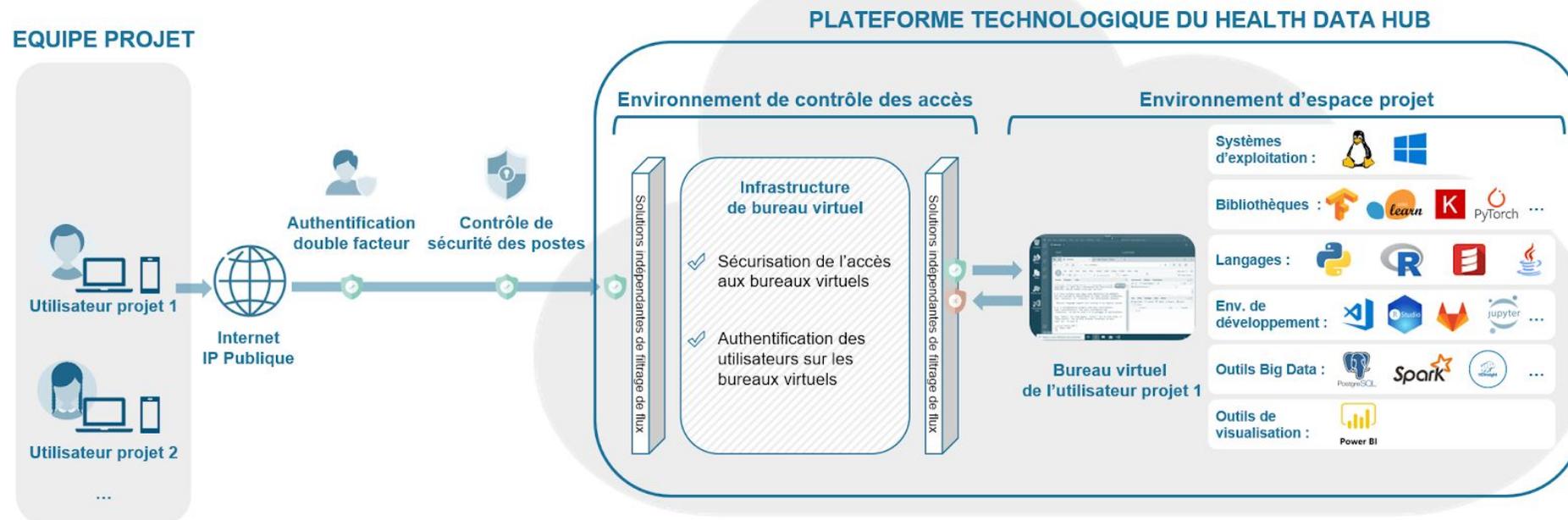
Deep Learning

- Tensorflow
- CUDA
- Keras
- Pytorch

Outils divers

- Serveur git interne
- PostgreSQL
- Cluster Spark

Présentation de la plateforme et de l'espace projet



Quels sont les services disponibles sur l'espace projet ?

- Plateforme la plus modulaire et évolutive possible, en s'adaptant aux besoins des équipes projet.
- Personnalisation possible des espaces projet et des outils mis à disposition.
- Ajustement de la puissance de calcul allouée à un projet à l'utilisation qui est en faite.

Présentation de la plateforme et de l'espace projet

Si vous hésitez à utiliser la plateforme technologique du HDH pour mener votre étude

Avec ou sans l'utilisation de la plateforme du Health Data Hub

Le responsable de traitement du projet doit s'assurer de la conformité de son étude au [RGPD](#).

Pour ce faire, **une Analyse d'Impact relative à la Protection des Données (AIPD)** doit être menée pour [tout traitement susceptible d'engendrer des risques élevés pour les droits et libertés des personnes](#) et ainsi, garantir le respect à la vie privée des personnes concernées.

L'AIPD doit être menée :

- En amont du démarrage du projet et peut être mise à jour tout au long de son cycle de vie ;
- Avec l'implication du Délégué à la protection des données (DPD/DPO) pour conseiller le responsable de traitement.

La liste des types d'opération de traitement pour lesquelles une analyse d'impact relative à la protection des données est disponible [ici](#) et les lignes directrices concernant l'analyse d'impact sont disponibles [ici](#).

Utilisation de la plateforme technologique du Health Data Hub

- Homologation de l'espace projet (slide 6) : analyse de risque



- Organisation du transfert des données si elles ne sont pas déjà inscrites au catalogue du Health Data Hub (slide 7)

Présentation de la plateforme et de l'espace projet

L'homologation de votre espace projet est nécessaire pour déposer à la CNIL

1

Pour se mettre en conformité

L'homologation de l'espace projet est un **prérequis à l'obtention des autorisations CNIL**.

En effet, les risques liés aux traitements de données de santé doivent être maîtrisés **pour protéger la vie privée des citoyens**. Cette démarche est rendue obligatoire par le RGPD pour tout traitement de données sensibles et est soumise au contrôle de la CNIL.

2

Pour mitiger les risques identifiés

L'homologation de votre espace projet **s'inscrit dans une démarche de sécurité** initiée par le Health Data Hub (HDH) et la complète.

Les mesures de sécurité mises en place par le HDH permettent de couvrir la majorité des risques identifiés par l'analyse réalisée sur la plateforme technologique. Afin de couvrir les risques subsistants, liés à l'action des utilisateurs projet, il est important que le Responsable de Traitement mette en place des mesures de sécurité complémentaires, et accepte les risques résiduels.

PÉRIMÈTRE SUR LEQUEL METTRE EN PLACE DES MESURES COMPLÉMENTAIRES

Plateforme technologique du Health Data Hub

Poste de travail des
opérateurs



Espaces dédiés aux
opérateurs



Votre espace projet



Poste de travail des
utilisateurs projet



Légende



Opérateurs du HDH



Utilisateurs projet

Mesures de sécurité mises en place par le HDH

Mesures de sécurité complémentaires à mettre en place

La démarche d'homologation est portée par le responsable de traitement et doit être menée conjointement avec le(s) responsable(s) de données (organisme possédant les bases utilisées pour l'étude et les transférant au Health Data Hub).

Présentation de la plateforme et de l'espace projet

Transfert des données vers la plateforme du HDH lorsque les données ne sont pas déjà au catalogue

Sécurité transfert de données

Processus à mettre en place pour assurer les transferts des données au sein de la plateforme technologique du HDH :

- Chiffrement des données
- Capacité à configurer un VPN IP SEC
- Mise en place d'un accès SFTP (serveur accessible par le HDH uniquement)

Nota bene étapes à confirmer avec le [Health Data Hub](#) selon les cas

Les interlocuteurs de votre organisation ou de celle du responsable de données à mobiliser

Afin de mener à bien les différentes étapes du parcours, nous vous recommandons d'**identifier** et de **vous rapprocher dès à présent** des interlocuteurs suivants dans votre organisation ou dans celle de l'équipe identifiée comme responsable de données, dont le rôle vous est précisé.

Interlocuteur	 Cadrage des données et du transfert	 Mise en œuvre du transfert
Chef de projet (équipe projet)	Suivi des échanges entre les équipes techniques du Responsable de Données et du HDH	Coordination de l'avancement des travaux en lien avec l'opérateur projet (interlocuteur HDH)
Experts infrastructure (responsable de données)	Définition avec le HDH du protocole de transfert des données vers la plateforme du HDH	Mise en œuvre des éléments d'infrastructure technique nécessaires au transfert et réponse aux sollicitations du HDH sur le sujet
Experts de la base de données (responsable de données)	Réponse aux sollicitations du HDH sur le contenu de la base, partage des informations permettant de sensibiliser aux données les futurs utilisateurs	
Responsable de la Sécurité des Systèmes d'Information (RSSI) du Responsable de Données	Validation du niveau de sécurité du protocole de transfert des données vers le HDH	

Besoin d'aller plus loin ?



A toutes les étapes de ma démarche, je peux :

- Consulter le [répertoire des projets](#) pour avoir une idée de ce qui est réalisé à partir des données ;
- Faire appel à la [communauté](#) ;
- Interroger le [forum d'entraide](#) pour échanger sur la faisabilité de mon étude ;
- Participer aux [matinales organisées](#) en collaboration avec la CNAM pour avoir un premier niveau d'information sur le SNDS historique et aux [meet-up](#) pour échanger avec tout type de porteurs de projet ;
- Me référer à la [documentation collaborative](#) ou utiliser un des programmes partagés par la communauté ;
- Consulter le [starter kit](#) présent sur le site du Health Data Hub ;
- [Contacter](#) le HDH



Suivez-nous sur les réseaux sociaux !

