

Plateforme des données de santé

Fonctionnalités de la plateforme technologique

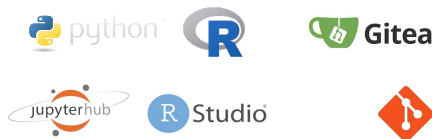


Quels outils sont disponibles sur la plateforme ?

Outils
disponibles

Utilisation

Développer



Langages :

- Python 3.6 / 3.8 / 3.10 / 3.14
- R 3.6 / 4.2 / 4.5

Environnements :

- Jupyter (Python et R)
- Rstudio (R)

Suivi du développement

- Versionnement des travaux

Package manager

- Installation en autonomie de librairies disponibles sur CRAN et PyPi

Stocker et traiter vos données



Base de données relationnelles :

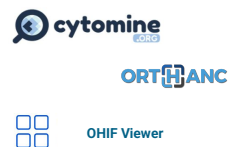
- PostgreSQL

Traitement volumétrique :

- Spark / Hive

Visualisation des données

Outils d'imagerie



Imagerie :

- Cytomine
- Orthanc
- OHIF Viewer

Outils spécifiques

NONMEM7™



- **Pharmaco :**
NONMEM / PsN

Les axes de développement de la plateforme

Agrandir le panel d'outils
pour traiter de nouveaux
cas d'usages

Faciliter la mise à
disposition des données



Améliorer l'expérience
utilisateur

Etre respectueux du RGDP

Simplifier l'opérabilité
de la plateforme



**PLATEFORME
DES DONNÉES
DE SANTÉ**
FRENCH HEALTH DATA HUB



Les fonctionnalités disponibles

Une connexion aux outils simplifiée



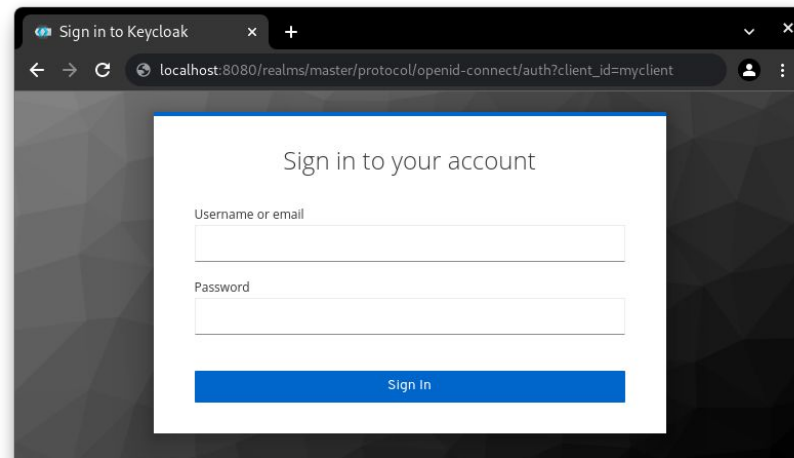
Contexte

Précédemment, les utilisateurs se **connectaient aux applications de leur espace** via :

- l'interface d'authentification de GitLab pour JupyterLab et GitLab. Les utilisateurs s'authentifiaient une fois via cette interface et pouvaient ensuite accéder aux deux outils.
- l'interface de connexion de chaque outil pour les autres outils.

Fonctionnalité disponible

- Un **portail d'authentification unique** est mis en place afin de permettre aux utilisateurs de se connecter une fois et d'accéder ensuite à **toutes les applications**.



Importer et utiliser des librairies (Python, R...)

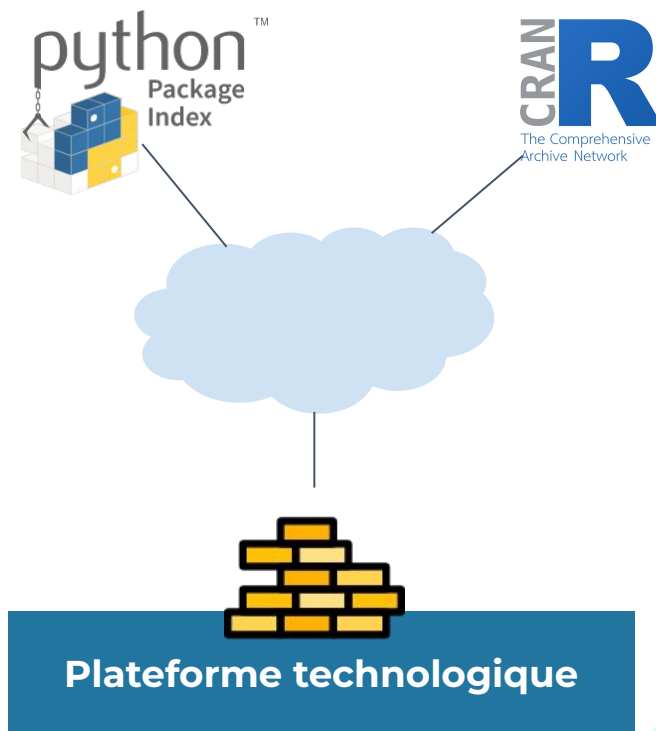


Contexte

- Par mesure de sécurité, les espaces d'analyse sur la plateforme technologique sont **coupés d'internet**
- Les projets devaient demander **l'installation sur leur espace de librairies** (outils de traitement de données) répondant à leurs besoins.
- La gestion des demandes, validation et installation de ces librairies nécessitait beaucoup d'actions manuelles.

Fonctionnalité disponible

- Un **gestionnaire de librairies** est disponible et permet aux utilisateurs d'installer **en autonomie** des librairies depuis les dépôts officiels **PyPI (Python) et CRAN (R)**, de gérer **les versions** à installer et de **vérifier automatiquement les vulnérabilités connues (CVE)**.



Importer des modèles Hugging Face

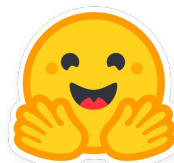


Contexte

- Les projets avaient remonté des besoins d'utilisation de **différents modèles présents sur Hugging Face**.
- Des solutions existaient mais présentaient des contraintes importantes qui limitaient notre capacité à répondre rapidement à ces besoins.

Fonctionnalité disponible

- En s'appuyant sur le package manager, les utilisateurs sont en capacité d'importer dans leur espace n'importe quel modèle présent sur Hugging Face **en autonomie** (à condition que le modèle passe **le filtre de sécurité**)



Hugging Face

Accéder à la documentation

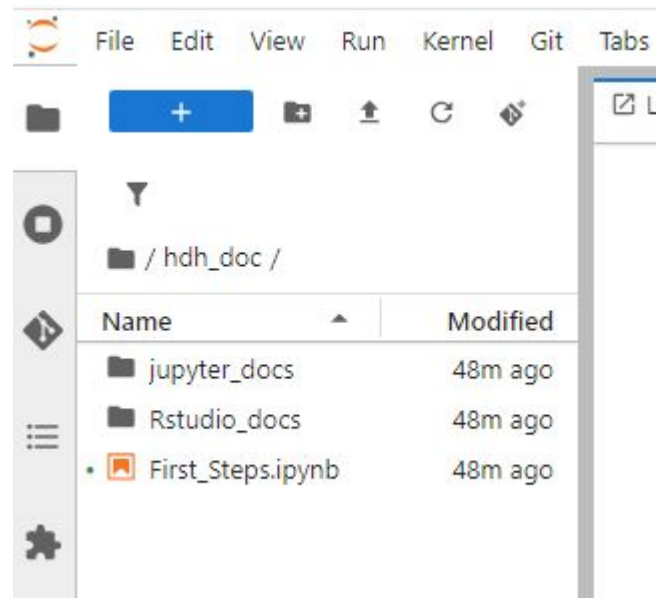


Contexte

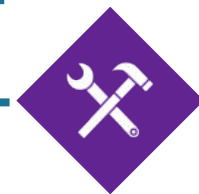
La documentation existante, disponible en dehors de la plateforme technologique, reste **difficile à mobiliser lors des travaux de recherche** des utilisateur.

Fonctionnalité disponible

- De la documentation, sous forme de **notebooks**, est directement accessible et exécutable sur l'espace de travail
- Cette documentation propose un accompagnement pas-à-pas pour **des cas d'usage spécifiques**, facilitant ainsi la prise en main de la plateforme et de ses outils.



Disposer d'un environnement de travail dédié par utilisateur

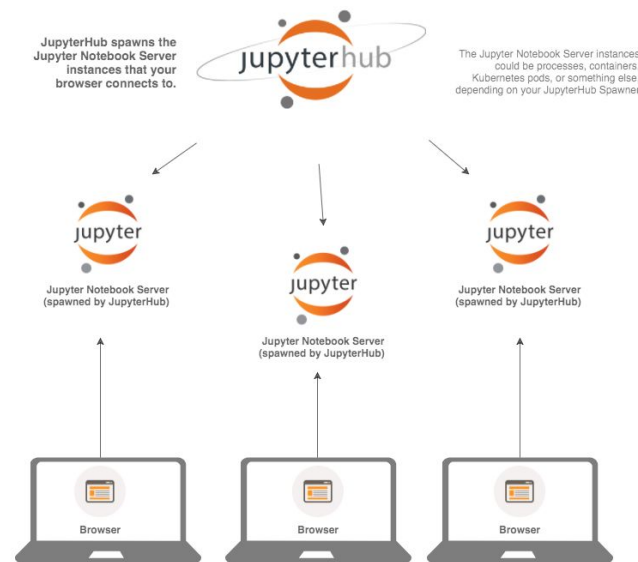


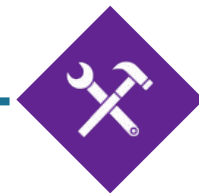
Contexte

- Actuellement, l'application Jupyter sur la plateforme impose aux utilisateurs de **partager un même espace de travail**.
- Ce mode de fonctionnement peut être assez contraignant en termes **d'expérience d'utilisation** (modifications de notebook par erreur) mais également en termes **de performance** (crash de Jupyter dans un contexte multi-utilisateurs).

Fonctionnalité disponible

- Mise en place de l'outil **Jupyter Hub** permettant que chaque utilisateur dispose de son environnement de travail dédié.





Contexte

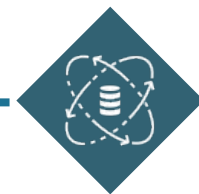
Certains projets qui ont un usage sur **de grandes volumétries de données** nécessite des outils performants pour les traiter.

Fonctionnalité disponible

- **Spark** est un puissant outil open source utilisé pour **traiter de grandes quantités de données**. Il permet de **diviser les tâches entre plusieurs machines** pour accélérer le traitement des données.



Transférer des données au format Parquet



Contexte

Dans le cadre de transfert de données volumineuses, des responsables de données pourraient être intéressés par l'utilisation du **format parquet**.

Fonctionnalité disponible

- Laisser la possibilité au Responsable de données, d'**importer directement les données au format Parquet** sur la plateforme technologique.
- Application **des automates de masquage** sur le format Parquet.



Parquet

Synchroniser du code entre la plateforme et l'externe

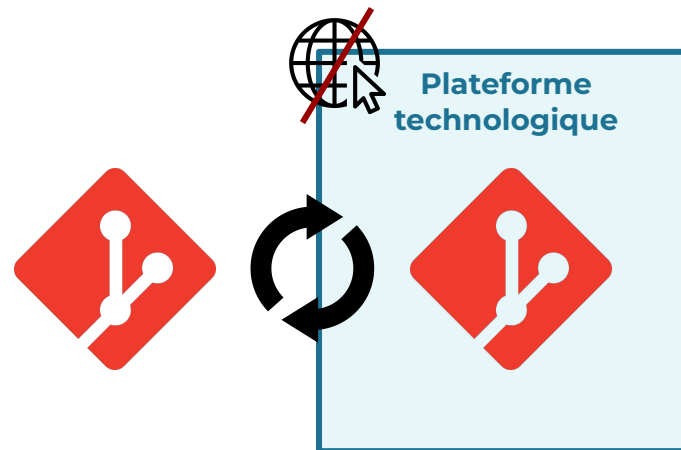


Contexte

- Par mesure de sécurité, les espaces d'analyse sur la plateforme technologique sont **couqués d'internet**.
- Les utilisateurs souhaitent **échanger du code** entre leur espace d'analyse et leur dépôt de code de travail usuel.

Fonctionnalité disponible

- Il sera possible de **synchroniser des dépôts Git externes et internes** à la plateforme, afin d'échanger du code **rapidement** et **facilement**.
- Des **contrôles automatiques** des imports et exports de code seront mis en place afin de limiter le recours à des interventions manuelles et ainsi réduire les délais.



Améliorer le masquage DICOM

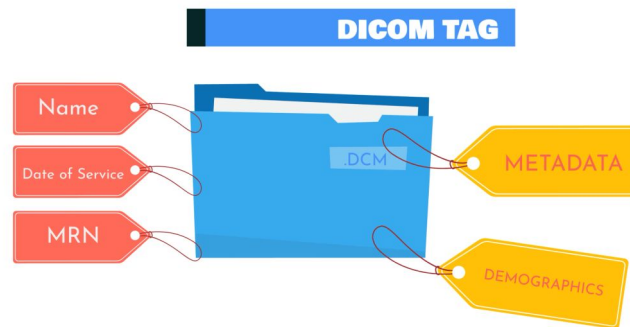


Contexte

Les automates de masquage ne permettaient pas de gérer de nouveaux cas d'usage sur les DICOM, comme **la structure des tags imbriqués**.

Fonctionnalité disponible

- Les automates de masquage sont maintenant en capacité de **gérer le masquage de tags imbriqués**.





Les fonctionnalités en cours d'analyse / développement

Migration de la plateforme vers un cloud souverain



Objectif

Mettre à disposition des acteurs de la recherche et de l'innovation en santé **un espace sécurisé sur un cloud souverain**, capable d'héberger les données de remboursement de l'Assurance maladie et d'autres bases de données de santé d'intérêt.

Les équipes techniques de la PDS travaillent pour atteindre un périmètre fonctionnel identique, voire enrichi, sur cette nouvelle plateforme tout en assurant **la continuité de service de la plateforme existante**.





Objectif

Proposer une **page d'accueil centralisée** sur la plateforme afin de mettre en avant :

- la documentation disponible
- les outils sur l'espace
- les dernières releases

et permettre d'accompagner les utilisateurs dans leurs premiers pas sur la plateforme.

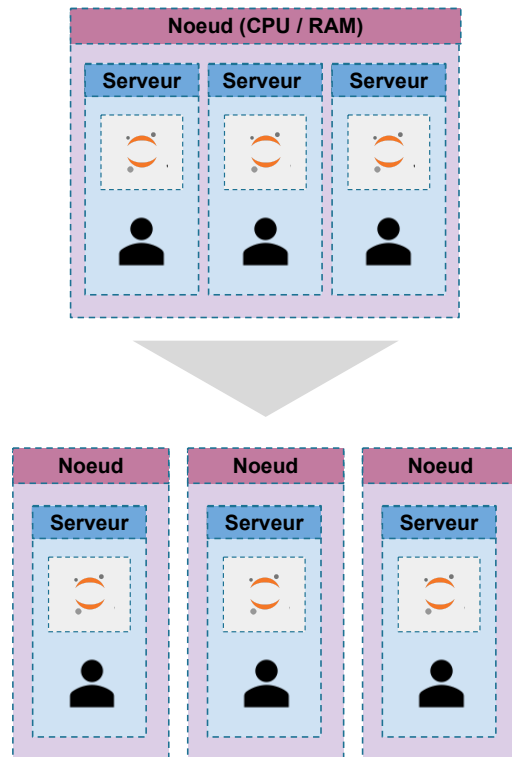


Gérer l'allocation des ressources selon les besoins utilisateurs

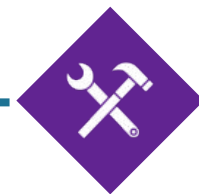


Objectif

S'appuyer sur la mise en place de **Jupyter Hub** afin d'avoir **une meilleure allocation des ressources** par espace et par utilisateur, et d'être en capacité de garantir un niveau de ressources pour chaque utilisateur.

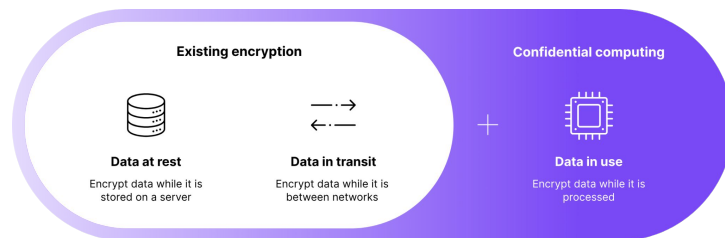


Confidential computing



Objectif

Expérimenter **l'architecture Confidential Computing** qui permet d'**encadrer l'accès et l'usage des données** tout en garantissant **leur confidentialité** et leur **traçabilité** au sein de la plateforme technologique.



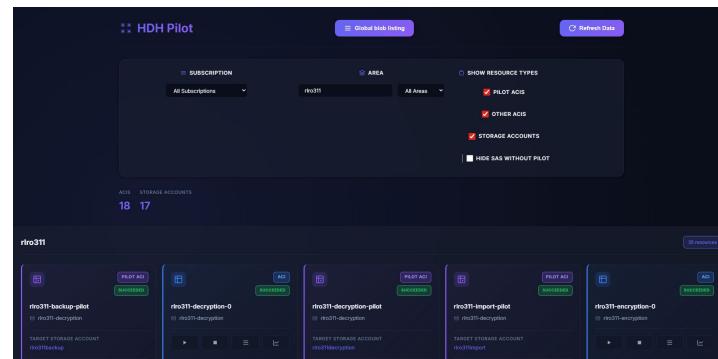
Amélioration du suivi des ingestions



Objectif

Mise en place d'un **tableau de suivi des ingestions** permettant de :

- Simplifier les opérations à réaliser dans le cadre **des ingestions et déplacements de données** sur la plateforme
- **réduire le temps de mise à disposition des données** au sein de vos espaces



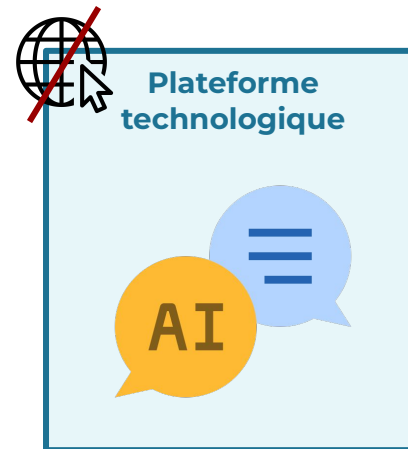
LLM intégré pour assister les utilisateurs



Objectif

Intégrer un **LLM directement** sur la plateforme pour permettre, dans un environnement déconnecté, un accompagnement continu à :

- la création et la fiabilisation du code
- la résolution d'erreurs
- l'accès à une documentation contextualisée





Suivez-nous sur les réseaux sociaux !

