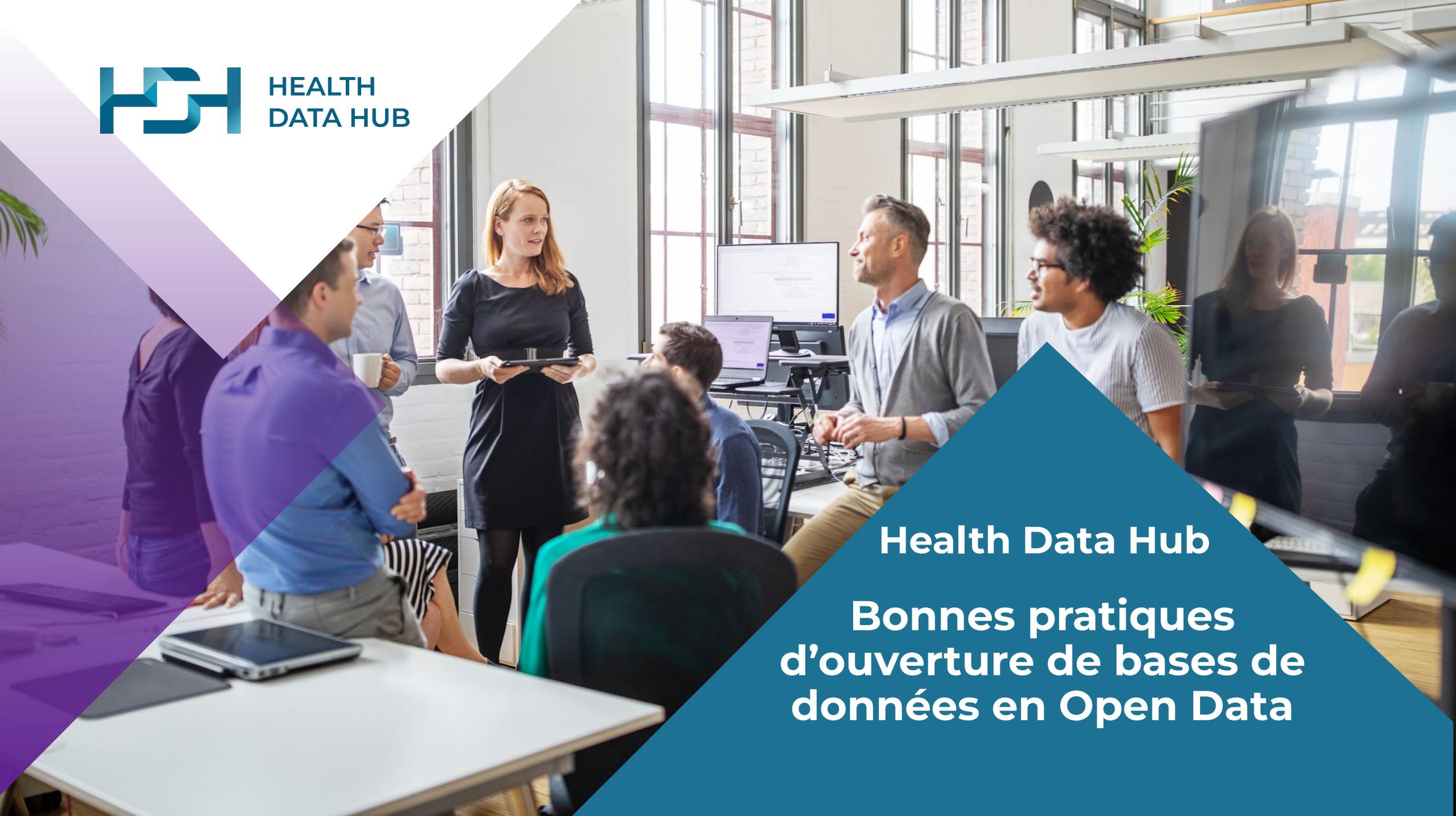




HEALTH  
DATA HUB



**Health Data Hub**  
**Bonnes pratiques  
d'ouverture de bases de  
données en Open Data**

## Plan

# Ouvrir une base de données de santé anonymisées

- I. Pré-requis à l'ouverture
- II. Étapes à suivre et modalités d'ouverture

Conclusion

# I - Pré-requis à l'ouverture



Le **caractère anonyme** de la base de données est un **prérequis obligatoire** pour l'ouverture de la base

Une solution d'anonymisation doit être construite **au cas par cas** par le partenaire et adaptée aux usages prévus. Pour aider à évaluer la bonne anonymisation d'un jeu de données, le **G29** a dégagé **trois critères** :



**L'individualisation** : est-il toujours possible d'isoler un individu ?



**La corrélation** : est-il possible de relier entre eux des ensembles de données distincts concernant un même individu ?



**L'inférence** : peut-on déduire de l'information sur un individu ?

- Dans le cas où il n'est pas possible de conclure parfaitement à l'absence de ces trois critères, le responsable de traitement qui souhaite anonymiser un jeu de données doit démontrer, via une **analyse des risques d'identification**, que le risque de ré-identification avec des moyens raisonnables est nul.
- **Le caractère anonyme des données est sous la responsabilité juridique du partenaire, qui doit être en mesure de le démontrer.**

**Documentation** :



- [Qu'est-ce qu'une donnée anonyme en santé ?](#)
- [Guide d'évaluation du caractère anonyme d'un jeu de données dans le cadre d'un projet de recherche](#)

## II - Étapes à suivre et modalités d'ouverture

*Étapes successives pour l'ouverture*



### Garantir l'interopérabilité des données

Utiliser des **formats de données standardisés** et largement acceptés pour **faciliter l'accès et l'utilisation** des données par le plus grand nombre :

- **Tabulaires** : csv
- **Imageries** : tiff

Les données peuvent nécessiter des manipulations avant ouverture garantissant l'**homogénéité** et la **lisibilité** des formats.

## II - Étapes à suivre et modalités d'ouverture

*Étapes successives pour l'ouverture*



Garantir  
l'interopérabilité  
des données

Fournir une documentation  
détaillée et les  
métadonnées

La **documentation** et les **métadonnées** doivent être **complètes** et expliciter :

- La **signification** des données,
- L'**origine**, le **contexte** et la **méthodologie** de collecte des données,
- La **structure** et le **format** des données,
- La **qualité** et **fiabilité** des données (mais aussi les erreurs potentielles et les **limites** du jeu de données),
- Les **mises à jour** et **évolutions**.

Un **template** de **documentation standardisé** est proposé pour vous guider dans la rédaction à cette étape.



## II - Étapes à suivre et modalités d'ouverture

*Étapes successives pour l'ouverture*



Garantir  
l'interopérabilité  
des données

Fournir une  
documentation  
détaillée et les  
métadonnées

Attribuer un DOI à la base  
de données pour garantir la  
parentalité

### DOI : Digital Object Identifier

- Fournit un **lien stable** vers un objet scientifique et sa description
- Il s'agit d'un **identifiant numérique unique et pérenne** pour un objet scientifique que l'on souhaite rendre **citabile**

Un DOI peut être attribué à  
l'algorithme à ouvrir via la  
plateforme fabrica de DataCite sur  
laquelle le HDH a un compte

## II - Étapes à suivre et modalités d'ouverture

Étapes successives pour l'ouverture



Soumettre un contenu à plusieurs licences permet de:

- ✓ **élargir les compatibilités**
- ✓ **profiter des communautés** utilisatrices des deux licences
- ✓ **protéger la distribution** du contenu dans l'éventualité où une juridiction nationale annule une des deux licences

La licence définit **les modalités de réutilisation** des jeux de données et de **distribution** des versions modifiées.

Nous recommandons l'**utilisation conjointe** de **deux licences permissives**, autorisées aux **administrations** et permettant une **réutilisation libre, gratuite et sans restriction** dans le **respect de la citation des auteurs** :

**Etalab 2.0 & CC-BY**

# II - Étapes à suivre et modalités d'ouverture

Plateforme de référencement des bases de données ouvertes - data.gouv.fr



Données Réutilisations Organisations Commencer sur data.gouv.fr Actualités Nous contacter

## Plateforme ouverte des données publiques françaises

Thématiques à la une

Données relatives aux **Énergies** →

Données relatives à **la Culture** →

Données relatives à **l'Éducation** →

Toutes les thématiques >

Actualités



### Quel suivi des usages des données ouvertes ?

Publiée le 28 novembre 2023

Venez plonger au coeur des réutilisations de données ouvertes !

Toutes les actualités >

**46 124**  
Jeux de données

**263 539**  
Ressources

**3 549**  
Réutilisations

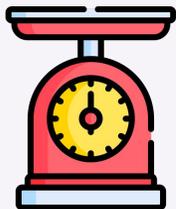
**108 715**  
Utilisateurs

**4 235**  
Organisations

**12 457**  
Discussions

## II - Étapes à suivre et modalités d'ouverture

### Modalités d'ouverture



Les **modalités** de dépôt dépendent de la **volumétrie** de la base de données :

- ❑  **$\leq 1$  Go** : le dépôt peut être **directement fait par le partenaire sur la plateforme data.gouv.fr**
- ❑  **$> 1$  Go** : la base doit être **hébergée sur une infrastructure propre** mise à disposition par le Health Data Hub, puis être **référéncée sur data.gouv.fr** via un **lien URL** fourni par un membre de l'équipe Open Science du Health Data Hub



Cas de figure 1 : la base a une volumétrie  **$\leq 1$  Go** : le dépôt peut être **directement fait par le partenaire sur la plateforme data.gouv.fr**

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 1 : la base a une volumétrie  $\leq 1$  Go : le dépôt peut être directement fait par le partenaire sur la plateforme [data.gouv.fr](https://data.gouv.fr)

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données



## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 1 : la base a une volumétrie  $\leq 1$  Go : le dépôt peut être directement fait par le partenaire sur la plateforme data.gouv.fr

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. Créer un **compte administrateur** sur data.gouv.fr

The image shows the data.gouv.fr website interface. At the top left is the French Republic logo and the text 'data.gouv.fr'. Below it is a navigation menu with items: Données, Réutilisations, Organisations, Commencer sur data.gouv.fr, Actualités, and Nous contacter. The main content area features a search bar with the text 'Recherche' and a magnifying glass icon. To the right of the search bar are two buttons: 'Se connecter' and 'S'enregistrer'. The 'S'enregistrer' button is highlighted with a yellow box, and a yellow arrow points from it to a registration form on the right side of the image. The registration form is titled 'S'enregistrer' and includes the following elements:

- A link: [Pourquoi se créer un compte?](#)
- A note: Les champs précédés d'une étoile (\*) sont obligatoires.
- An 'Email Address \*' field with the format 'Format attendu : nom@domain.fr'.
- A 'Password' field.
- A 'Retype Password' field.
- A 'Prénom \*' field.
- A 'Nom \*' field.
- A checkbox:  J'ai lu et j'accepte les conditions générales d'utilisation du service.
- A CAPTCHA image showing the word 'P5RKSE'.
- A label: Saisissez les caractères de l'image \*

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 1 : la base a une volumétrie  $\leq 1$  Go : le dépôt peut être directement fait par le partenaire sur la plateforme data.gouv.fr

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. Créer un **compte administrateur** sur data.gouv.fr
3. Créer **l'organisation** avec laquelle déposer le jeu de données :

### A titre d'exemple :



The screenshot shows the Health Data Hub website interface. At the top left is the HDH logo. Below it, the text 'Health Data Hub' is displayed. A navigation bar contains four tabs: 'Présentation', 'Jeux de données<sup>(1)</sup>', 'Réutilisations<sup>(2)</sup>', and 'Informations'. The 'Jeux de données<sup>(1)</sup>' tab is currently selected. Below the navigation bar, the 'Description' section is visible, containing three paragraphs of text.

**Description**

Les données de santé peuvent être utiles pour améliorer le système de santé, le rendre plus efficace, plus personnalisé et plus humaniste. Chaque citoyen le mérite ! C'est par le traitement et le croisement d'un grand volume de données de qualité que les recherches les plus impactantes pourront être menées : améliorer le dépistage et le diagnostic d'une maladie, analyser les effets secondaires des traitements, faire évoluer les essais cliniques ...

Nous proposons un ensemble de services pour accompagner les acteurs qui contribuent à trouver les solutions de demain pour améliorer la santé des citoyens : accompagner les porteurs de projet, construire et opérer une plateforme technologique pour leur offrir les meilleurs outils, réunir et mettre en forme les données au plus grand potentiel pour la recherche.

Créé par la Loi du 24 juillet 2019 relative à l'organisation et la transformation du système de santé, notre structure est un groupement d'intérêt public (GIP) qui associe 56 parties prenantes, en grande majorité issues de la puissance publique (CNAM, CNRS, Haute Autorité de santé, France Assos Santé, liste complète). Nous mettons en œuvre les grandes orientations stratégiques relatives au Système National des Données de Santé (SNDS) fixées par l'Etat et notamment le ministère des Solidarités et de la Santé. Notre financement est majoritairement public.

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 1 : la base a une volumétrie  $\leq 1$  Go : le dépôt peut être directement fait par le partenaire sur la plateforme data.gouv.fr

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Créer un compte administrateur** sur data.gouv.fr
3. **Créer l'organisation** avec laquelle déposer le jeu de données
4. **Déposer le jeu de données** sous cette organisation

Eléments du **template** de **documentation standardisé**

**Test\_20231212** Test\_1

**Description**

**Overview**  
Melanoma is a cancer of the skin which develops from cells responsible for skin pigmentation. In 2020, over 325,000 people were diagnosed with skin melanoma, with 57,000 deaths in the same year.<sup>1</sup> Melanomas represent 10% of all skin cancers and are the most dangerous due to high likelihood of metastasizing (spreading).<sup>2</sup>

Patients are initially diagnosed with melanoma after a pathologist examines a portion of the cancerous tissue under a microscope. At this stage, the pathologist assesses the risk of relapse—a return of cancerous cells after the melanoma has been treated—based on information such as the thickness of the tumor and the presence of an ulceration. Combined with factors such as age, sex, and medical history of the patient, these microscopic observations can help a dermatologist assess the severity of the disease and determine appropriate surgical and medical treatment. Preventative

[Lire plus](#)

**Producteur**  
Health Data Hub

**Dernière mise à jour**  
13 décembre 2023

**Licence**  
[Licence Ouverte / Open Licence version 2.0](#)

**Qualité des métadonnées**

**Fichiers (1)** **Réutilisations (1)** **Discussions (0)** **Ressources communautaires (0)** **Informations**

**1 FICHIER PRINCIPAL**

**bdd-test-melanoma.xlsx**  
Mis à jour le 13 décembre 2023 — xlsx (29.0Ko) — 0 téléchargements

[Voir les données](#)  

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 1 : la base a une volumétrie  $\leq 1$  Go : le dépôt peut être directement fait par le partenaire sur la plateforme data.gouv.fr

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Créer un compte administrateur** sur data.gouv.fr
3. **Créer l'organisation** avec laquelle déposer le jeu de données
4. **Déposer le jeu de données** sous cette organisation

### Test\_20231212 Test\_1

#### Description

##### Overview

Melanoma is a cancer of the skin which develops from cells responsible for skin pigmentation. In 2020, over 325,000 people were diagnosed with skin melanoma, with 57,000 deaths in the same year.<sup>1</sup> Melanomas represent 10% of all skin cancers and are the most dangerous due to high likelihood of metastasizing (spreading).<sup>2</sup>

Patients are initially diagnosed with melanoma after a pathologist examines a portion of the cancerous tissue under a microscope. At this stage, the pathologist assesses the risk of relapse—a return of cancerous cells after the melanoma has been treated—based on information such as the thickness of the tumor and the presence of an ulceration. Combined with factors such as age, sex, and medical history of the patient, these microscopic observations can help a dermatologist assess the severity of the disease and determine appropriate surgical and medical treatment. Preventative

[Lire plus](#)

Fichiers <sup>(1)</sup>

Réutilisations <sup>(1)</sup>

Discussions <sup>(0)</sup>

Ressources communautaires <sup>(0)</sup>

Informations

#### 1 FICHIER PRINCIPAL

 **bdd-test-melanoma.xlsx**

Mis à jour le 13 décembre 2023 — xlsx (29.0Ko) — 0 téléchargements

Voix les données



#### Producteur

 Health Data Hub

#### Dernière mise à jour

13 décembre 2023

#### Licence

[Licence Ouverte / Open Licence version 2.0](#)

 Qualité des métadonnées



## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 1 : la base a une volumétrie  $\leq 1$  Go : le dépôt peut être directement fait par le partenaire sur la plateforme data.gouv.fr

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Créer un compte administrateur** sur data.gouv.fr
3. **Créer l'organisation** avec laquelle déposer le jeu de données
4. **Déposer le jeu de données** sous cette organisation

**Test\_20231212** Test\_1

**Description**

**Overview**  
Melanoma is a cancer of the skin which develops from cells responsible for skin pigmentation. In 2020, over 325,000 people were diagnosed with skin melanoma, with 57,000 deaths in the same year.<sup>1</sup> Melanomas represent 10% of all skin cancers and are the most dangerous due to high likelihood of metastasizing (spreading).<sup>2</sup>

Patients are initially diagnosed with melanoma after a pathologist examines a portion of the cancerous tissue under a microscope. At this stage, the pathologist assesses the risk of relapse—a return of cancerous cells after the melanoma has been treated—based on information such as the thickness of the tumor and the presence of an ulceration. Combined with factors such as age, sex, and medical history of the patient, these microscopic observations can help a dermatologist assess the severity of the disease and determine appropriate surgical and medical treatment. Preventative

[Lire plus](#)

**Producteur**  
 Health Data Hub

**Dernière mise à jour**  
13 décembre 2023

**Licence**  
[Licence Ouverte / Open Licence version 2.0](#)

**Qualité des métadonnées**  


**Fichiers** <sup>(1)</sup> **Réutilisations** <sup>(1)</sup> **Discussions** <sup>(0)</sup> **Ressources communautaires** <sup>(0)</sup> **Informations**

**1 FICHIER PRINCIPAL**

 **bdd-test-melanoma.xlsx**  
Mis à jour le 13 décembre 2023 — xlsx (29.0Ko) — 0 téléchargements

[Voir les données](#)  

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 1 : la base a une volumétrie  $\leq 1$  Go : le dépôt peut être directement fait par le partenaire sur la plateforme data.gouv.fr

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Créer un compte administrateur** sur data.gouv.fr
3. **Créer l'organisation** avec laquelle déposer le jeu de données
4. **Déposer le jeu de données** sous cette organisation

The screenshot shows the 'Informations' tab of a dataset page on data.gouv.fr. The page is titled 'INFORMATIONS' and contains the following details:

- Tags:** data-challenge, hdh, health-data-hub, melanoma, visiomet
- Licence:** Licence Ouverte / Open Licence version 2.0
- ID:** 6578aaf9b29ddfeb445ea50a
- TEMPORALITÉ:**
  - Création:** 12 décembre 2023
  - Fréquence:** Ponctuelle
  - Couverture temporelle:** 2012 à 2016
  - Dernière mise à jour:** 13 décembre 2023

Éléments du **template** de **documentation** standardisé

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 1 : la base a une volumétrie  $\leq 1$  Go : le dépôt peut être directement fait par le partenaire sur la plateforme data.gouv.fr

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Créer un compte administrateur** sur data.gouv.fr
3. **Créer l'organisation** avec laquelle déposer le jeu de données
4. **Déposer le jeu de données** sous cette organisation



Le choix des tags est important pour que la base soit aussi référencée lors des recherches utilisant ces mots clés



Fichiers<sup>(1)</sup> Réutilisations<sup>(1)</sup> Discussions<sup>(0)</sup> Ressources communautaires<sup>(0)</sup> Informations

**INFORMATIONS**

Tags  
data-challenge hdh health-data-hub melanoma visiomel

Licence ID  
Licence Ouverte / Open Licence version 2.0 6578aaf9b29ddfefb445ea50a

**TEMPORALITÉ**

Création	Fréquence	Couverture temporelle
12 décembre 2023	Ponctuelle	2012 à 2016
Dernière mise à jour		
13 décembre 2023		

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 1 : la base a une volumétrie  $\leq 1$  Go : le dépôt peut être directement fait par le partenaire sur la plateforme [data.gouv.fr](https://data.gouv.fr)

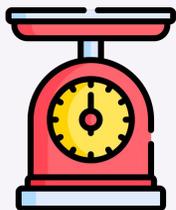
1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Créer un compte administrateur** sur [data.gouv.fr](https://data.gouv.fr)
3. **Créer l'organisation** avec laquelle déposer le jeu de données
4. **Déposer le jeu de données** sous cette organisation



**NB** : un membre de l'équipe Open Science peut vous accompagner dans toutes ces étapes en tant que co-administrateur

## II - Étapes à suivre et modalités d'ouverture

### Modalités d'ouverture



Les **modalités** de dépôt dépendent de la **volumétrie** de la base de données :

- ❑ **≤ 1 Go** : le dépôt peut être **directement fait par le partenaire sur la plateforme data.gouv.fr**
- ❑ **> 1 Go** : la base doit être **hébergée sur une infrastructure propre** mise à disposition par le Health Data Hub, puis être **référéncée sur data.gouv.fr** via un **lien URL** fourni par un membre de l'équipe Open Science du Health Data Hub



Cas de figure 2 : la base a une volumétrie **> 1 Go** : la base doit être **hébergée** sur une **infrastructure propre** mise à disposition par le Health Data Hub puis être **référéncée sur data.gouv.fr** via un **lien URL** fourni par un membre de l'équipe Open Science du Health Data Hub

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 2 : la base a une volumétrie **> 1 Go** : la base doit être hébergée sur une infrastructure propre mise à disposition par le Health Data Hub, puis être référencée sur [data.gouv.fr](https://data.gouv.fr) via un lien URL fourni par un membre de l'équipe Open Science du Health Data Hub

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données



## II - Étapes à suivre et modalités d'ouverture

*Modalités d'ouverture - Cas de figure 2 : la base a une volumétrie **> 1 Go** : la base doit être hébergée sur une infrastructure propre mise à disposition par le Health Data Hub, puis être référencée sur data.gouv.fr via un lien URL fourni par un membre de l'équipe Open Science du Health Data Hub*

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Partager la base de données à ouvrir** avec un membre du **Health Data Hub** pour **upload sur l'infrastructure** mise à disposition

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 2 : la base a une volumétrie **> 1 Go** : la base doit être hébergée sur une infrastructure propre mise à disposition par le Health Data Hub, puis être référencée sur data.gouv.fr via un lien URL fourni par un membre de l'équipe Open Science du Health Data Hub

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Partager la base de données à ouvrir** avec un membre du **Health Data Hub** pour **upload sur l'infrastructure** mise à disposition
3. **Créer un compte administrateur** sur data.gouv.fr

The image shows a screenshot of the data.gouv.fr website. At the top left is the logo of the République Française. The main navigation bar includes links for Données, Réutilisations, Organisations, Commencer sur data.gouv.fr, Actualités, and Nous contacter. A search bar with the text 'Recherche' and a magnifying glass icon is present. A yellow box highlights the 'S'enregistrer' button, with an arrow pointing to a registration form on the right. The form is titled 'S'enregistrer' and includes a link for 'Pourquoi se créer un compte?'. Below this is a note: 'Les champs précédés d'une étoile (\*) sont obligatoires.' The form fields are: 'Email Address \*' (with format 'nom@domain.fr'), 'Password', 'Retype Password', 'Prénom \*', and 'Nom \*'. There is a checkbox for 'J'ai lu et j'accepte les conditions générales d'utilisation du service.' and a CAPTCHA image with the text 'P5RKSE'. The bottom of the form says 'Saisissez les caractères de l'image \*'.

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 2 : la base a une volumétrie **> 1 Go** : la base doit être hébergée sur une infrastructure propre mise à disposition par le Health Data Hub, puis être référencée sur [data.gouv.fr](https://data.gouv.fr) via un lien URL fourni par un membre de l'équipe Open Science du Health Data Hub

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Partager la base de données à ouvrir** avec un membre du **Health Data Hub** pour **upload sur l'infrastructure** mise à disposition
3. **Créer un compte administrateur** sur [data.gouv.fr](https://data.gouv.fr)
4. **Créer l'organisation** avec laquelle déposer le jeu de données :

A titre d'exemple :



The screenshot shows the Health Data Hub website interface. At the top left is the HDH logo. Below it, the text 'Health Data Hub' is displayed. A navigation bar contains four tabs: 'Présentation', 'Jeux de données<sup>(1)</sup>', 'Réutilisations<sup>(2)</sup>', and 'Informations'. The 'Jeux de données<sup>(1)</sup>' tab is currently selected. Below the navigation bar, the 'Description' section is visible, containing two paragraphs of text. The first paragraph discusses the utility of health data for improving the healthcare system. The second paragraph describes the services provided to support actors in finding solutions for improving citizen health. At the bottom of the screenshot, a paragraph explains the legal basis for the organization, citing the July 24, 2019 law and its role as a public interest group (GIP) with 56 members, including CNAM, CNRS, Haute Autorité de santé, France Assos Santé, and the Ministry of Solidarity and Health.

**Health Data Hub**

Présentation **Jeux de données<sup>(1)</sup>** Réutilisations<sup>(2)</sup> Informations

**Description**

Les données de santé peuvent être utiles pour améliorer le système de santé, le rendre plus efficace, plus personnalisé et plus humaniste. Chaque citoyen le mérite ! C'est par le traitement et le croisement d'un grand volume de données de qualité que les recherches les plus impactantes pourront être menées : améliorer le dépistage et le diagnostic d'une maladie, analyser les effets secondaires des traitements, faire évoluer les essais cliniques ...

Nous proposons un ensemble de services pour accompagner les acteurs qui contribuent à trouver les solutions de demain pour améliorer la santé des citoyens : accompagner les porteurs de projet, construire et opérer une plateforme technologique pour leur offrir les meilleurs outils, réunir et mettre en forme les données au plus grand potentiel pour la recherche.

Créé par la Loi du 24 juillet 2019 relative à l'organisation et la transformation du système de santé, notre structure est un groupement d'intérêt public (GIP) qui associe 56 parties prenantes, en grande majorité issues de la puissance publique (CNAM, CNRS, Haute Autorité de santé, France Assos Santé, liste complète). Nous mettons en œuvre les grandes orientations stratégiques relatives au Système National des Données de Santé (SNDS) fixées par l'Etat et notamment le ministère des Solidarités et de la Santé. Notre financement est majoritairement public.

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 2 : la base a une volumétrie **> 1 Go** : la base doit être hébergée sur une infrastructure propre mise à disposition par le Health Data Hub, puis être référencée sur [data.gouv.fr](https://data.gouv.fr) via un lien URL fourni par un membre de l'équipe Open Science du Health Data Hub

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Partager la base de données à ouvrir** avec un membre du **Health Data Hub** pour **upload sur l'infrastructure** mise à disposition
3. **Créer un compte administrateur** sur [data.gouv.fr](https://data.gouv.fr)
4. **Créer l'organisation** avec laquelle déposer le jeu de données
5. **Créer la page de dépôt** de la base avec la **documentation**

Eléments du **template**  
de **documentation**  
standardisé



**Test20240110** BDD\_voluineuse

**Description**  
Melanoma is a cancer of the skin which develops from cells responsible for skin pigmentation. In 2020, over 325,000 people were diagnosed with skin melanoma, with 57,000 deaths in the same year.<sup>1</sup> Melanomas represent 10% of all skin cancers and are the most dangerous due to high likelihood of metastasizing (spreading).<sup>2</sup>

Patients are initially diagnosed with melanoma after a pathologist examines a portion of the cancerous tissue under a microscope. At this stage, the pathologist assesses the risk of relapse—a return of cancerous cells after the melanoma has been treated—based on information such as the thickness of the tumor and the presence of an ulceration. Combined with factors such as age, sex, and medical history of the patient, these microscopic observations can help a dermatologist assess the severity of the disease and determine appropriate surgical and medical treatment. Preventative treatments can be administered to patients with high likelihood for relapse. However, these are costly and expose patients to significant drug toxicity.

Assessing the risk of relapse therefore a vital but difficult task. It requires specialized training and

[Lire plus](#)

**Producteur**  
 Health Data Hub

**Dernière mise à jour**  
16 janvier 2024

**Licence**  
[Licence Ouverte / Open Licence version 2.0](#)

**Qualité des métadonnées**

Fichiers<sup>(0)</sup> Réutilisations<sup>(0)</sup> Discussions<sup>(0)</sup> Ressources communautaires<sup>(0)</sup> Informations



Il n'y a pas encore de fichier pour ce jeu de données.

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 2 : la base a une volumétrie **> 1 Go** : la base doit être hébergée sur une infrastructure propre mise à disposition par le Health Data Hub, puis être référencée sur [data.gouv.fr](https://data.gouv.fr) via un lien URL fourni par un membre de l'équipe Open Science du Health Data Hub

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Partager la base de données à ouvrir** avec un membre du **Health Data Hub** pour **upload sur l'infrastructure** mise à disposition
3. **Créer un compte administrateur** sur [data.gouv.fr](https://data.gouv.fr)
4. **Créer l'organisation** avec laquelle déposer le jeu de données
5. **Créer la page de dépôt** de la base avec la **documentation**

**Test20240110** BDD\_volvineuse

**Description**

Melanoma is a cancer of the skin which develops from cells responsible for skin pigmentation. In 2020, over 325,000 people were diagnosed with skin melanoma, with 57,000 deaths in the same year.<sup>1</sup> Melanomas represent 10% of all skin cancers and are the most dangerous due to high likelihood of metastasizing (spreading).<sup>2</sup>

Patients are initially diagnosed with melanoma after a pathologist examines a portion of the cancerous tissue under a microscope. At this stage, the pathologist assesses the risk of relapse—a return of cancerous cells after the melanoma has been treated—based on information such as the thickness of the tumor and the presence of an ulceration. Combined with factors such as age, sex, and medical history of the patient, these microscopic observations can help a dermatologist assess the severity of the disease and determine appropriate surgical and medical treatment. Preventative treatments can be administered to patients with high likelihood for relapse. However, these are costly and expose patients to significant drug toxicity.

Assessing the risk of relapse therefore a vital but difficult task. It requires specialized training and

[Lire plus](#)

Producteur  
 Health Data Hub

Dernière mise à jour  
16 janvier 2024

Licence  
[Licence Ouverte / Open Licence version 2.0](#)

Qualité des métadonnées

Le **référencement** de la base de données se fait dans un étape ultérieure via un **lien URL**

Fichiers (0) Réutilisations (0) Discussions (0) Ressources communautaires (0) Informations



Il n'y a pas encore de fichier pour ce jeu de données.

[Ajouter un fichier](#)

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 2 : la base a une volumétrie **> 1 Go** : la base doit être hébergée sur une infrastructure propre mise à disposition par le Health Data Hub, puis être référencée sur [data.gouv.fr](https://data.gouv.fr) via un lien URL fourni par un membre de l'équipe Open Science du Health Data Hub

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Partager la base de données à ouvrir** avec un membre du **Health Data Hub** pour **upload sur l'infrastructure** mise à disposition
3. **Créer un compte administrateur** sur [data.gouv.fr](https://data.gouv.fr)
4. **Créer l'organisation** avec laquelle déposer le jeu de données
5. **Créer la page de dépôt** de la base avec la **documentation**

Éléments du **template de documentation standardisé**

Fichiers<sup>(0)</sup> Réutilisations<sup>(0)</sup> Discussions<sup>(0)</sup> Ressources communautaires<sup>(0)</sup> Informations

### INFORMATIONS

**Tags**  
artificial-intelligence data-challenge hdh health health-data-hub melanoma pathologie prediction relapse sfp visiomet

**Licence**  
Licence Ouverte / Open Licence version 2.0

**ID**  
65a6878fea60c4ddbc95535d

---

**TEMPORALITÉ**

<b>Création</b> 16 janvier 2024	<b>Fréquence</b> Sans régularité	<b>Couverture temporelle</b> 2012 à 2016
<b>Dernière mise à jour</b> 16 janvier 2024		

---

**COUVERTURE SPATIALE**

<b>Couverture territoriale</b> France	<b>Granularité de la couverture territoriale</b> Pays
------------------------------------------	----------------------------------------------------------

---

**ACTIONS**

Intégrer sur votre site

```
<div data-udata-dataset="65a6878fea60c4ddbc95535d"></div><script data-udata="https://demo.data.gouv.fr/" src="https://demo-static.data.gouv.fr/static/oembed.js">
```

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 2 : la base a une volumétrie **> 1 Go** : la base doit être hébergée sur une infrastructure propre mise à disposition par le Health Data Hub, puis être référencée sur [data.gouv.fr](https://data.gouv.fr) via un lien URL fourni par un membre de l'équipe Open Science du Health Data Hub

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Partager la base de données à ouvrir** avec un membre du **Health Data Hub** pour **upload sur l'infrastructure** mise à disposition
3. **Créer un compte administrateur** sur [data.gouv.fr](https://data.gouv.fr)
4. **Créer l'organisation** avec laquelle déposer le jeu de données
5. **Créer la page de dépôt** de la base avec la **documentation**



**Le choix des tags est important pour que la base soit aussi référencée lors des recherches utilisant ces mots clés**

The screenshot shows the 'Informations' tab of a dataset page on data.gouv.fr. The 'Tags' section is highlighted with a yellow box and contains the following tags: artificial-intelligence, data-challenge, hdh, health, health-data-hub, melanoma, pathologie, prediction, relapse, sfp, and visiomet. Below the tags, the 'Licence' is 'Licence Ouverte / Open Licence version 2.0' and the 'ID' is '65a6878fea60c4ddbc95535d'. The 'TEMPORALITE' section shows 'Création' on 16 janvier 2024, 'Fréquence' as 'Sans régularité', and 'Couverture temporelle' from 2012 to 2016. The 'COUVERTURE SPATIALE' section shows 'Couverture territoriale' as 'France' and 'Granularité de la couverture territoriale' as 'Pays'. At the bottom, the 'ACTIONS' section includes a code snippet for embedding the dataset on a website.

INFORMATIONS		
<b>Tags</b>		
artificial-intelligence data-challenge hdh health health-data-hub melanoma pathologie prediction relapse sfp visiomet		
<b>Licence</b>	<b>ID</b>	
Licence Ouverte / Open Licence version 2.0	65a6878fea60c4ddbc95535d	
<b>TEMPORALITE</b>		
<b>Création</b>	<b>Fréquence</b>	<b>Couverture temporelle</b>
16 janvier 2024	Sans régularité	2012 à 2016
<b>Dernière mise à jour</b>		
16 janvier 2024		
<b>COUVERTURE SPATIALE</b>		
<b>Couverture territoriale</b>	<b>Granularité de la couverture territoriale</b>	
France	Pays	
<b>ACTIONS</b>		
<b>Intégrer sur votre site</b>		
<code>&lt;div data-udata-dataset="65a6878fea60c4ddbc95535d"&gt;&lt;/div&gt;&lt;script data-udata="https://demo.data.gouv.fr/" src="https://demo-static.data.gouv.fr/static/oembed.js"&gt;</code>		

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 2 : la base a une volumétrie **> 1 Go** : la base doit être hébergée sur une infrastructure propre mise à disposition par le Health Data Hub, puis être référencée sur [data.gouv.fr](https://data.gouv.fr) via un lien URL fourni par un membre de l'équipe Open Science du Health Data Hub

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Partager la base de données à ouvrir** avec un membre du **Health Data Hub** pour **upload sur l'infrastructure** mise à disposition
3. **Créer un compte administrateur** sur [data.gouv.fr](https://data.gouv.fr)
4. **Créer l'organisation** avec laquelle déposer le jeu de données
5. **Créer la page de dépôt** de la base avec la **documentation**
6. **Préparer les données** en vue de l'ouverture :



Afin que l'utilisateur n'ait pas besoin de télécharger la totalité de la base pour vérifier qu'elle correspond à ses besoins il est recommandé de :

- **Mettre à disposition un échantillon de données**
- **Scinder la base de données en fichiers selon un critère pertinent**

Exemple : Dans le cas d'une base de données sur le thème du mélanome incluant des patients avec 4 stades pronostiques, il est pertinent de scinder la base selon les critères pronostiques

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 2 : la base a une volumétrie **> 1 Go** : la base doit être hébergée sur une infrastructure propre mise à disposition par le Health Data Hub, puis être référencée sur data.gouv.fr via un lien URL fourni par un membre de l'équipe Open Science du Health Data Hub

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Partager la base de données à ouvrir** avec un membre du **Health Data Hub** pour **upload sur l'infrastructure** mise à disposition
3. **Créer un compte administrateur** sur data.gouv.fr
4. **Créer l'organisation** avec laquelle déposer le jeu de données
5. **Créer la page de dépôt** de la base avec la **documentation**
6. **Préparer les données** en vue de l'ouverture
7. **Référencer le jeu de données** sur la page de dépôt via un **lien URL**



La base à ouvrir étant hébergée sur l'infrastructure mise à disposition par le Health Data Hub, le référencement de la base sur la page de dépôt sera réalisé via un lien URL fourni par un membre de l'équipe Open Science du Health Data Hub

## II - Étapes à suivre et modalités d'ouverture

Modalités d'ouverture - Cas de figure 2 : la base a une volumétrie **> 1 Go** : la base doit être hébergée sur une infrastructure propre mise à disposition par le Health Data Hub, puis être référencée sur [data.gouv.fr](https://data.gouv.fr) via un lien URL fourni par un membre de l'équipe Open Science du Health Data Hub

1. **Contractualisation** avec le Health Data Hub pour encadrer le rôle des parties dans l'ouverture des données
2. **Partager la base de données à ouvrir** avec un membre du **Health Data Hub** pour **upload sur l'infrastructure** mise à disposition
3. **Créer un compte administrateur** sur [data.gouv.fr](https://data.gouv.fr)
4. **Créer l'organisation** avec laquelle déposer le jeu de données
5. **Créer la page de dépôt** de la base avec la **documentation**
6. **Préparer les données** en vue de l'ouverture
7. **Référencer le jeu de données** sur la page de dépôt via un **lien URL**



**NB** : un membre de l'équipe Open Science peut vous accompagner dans toutes ces étapes en tant que co-administrateur

# Conclusion

## Bonnes pratiques

- Publier les données sous une **organisation** préalablement créé et décrite pour gagner en visibilité (ne pas en votre nom propre)  
Exemple : “Consortium VisioMel” ou “Société Française de Pathologie”
- Si la base de données est **volumineuse**, nous recommandons de la **scinder** en fichiers et de mettre à disposition un **échantillon de données**  
Exemple : Dans le cas d’une base de données sur le thème du mélanome incluant des patients avec 4 stades pronostiques, il est pertinent de scinder la base selon les critères pronostiques
- Être exhaustif sur les **mots clés** à utiliser :
  - Sigles de **l’organisation** dépositaire : “sfp” ; “sfd” ; “gcc” ; “cnpath”
  - **Health Data Hub** avec les nomenclatures suivantes : “health-data-hub” ; “hdh” ; “health data hub”
  - Mots clés sur la **thématique** et les **applications** de la base de données  
Exemple : “data-challenge” ; “melanoma” ; “IA” ; artificial intelligence” ; “relapse” ; “prediction”



**Avez-vous des questions ?**



Suivez-nous sur les réseaux sociaux !

