

# **Catalogue de métadonnées**

## **Guide d'utilisation du catalogue de métadonnées de santé en marque blanche**

## 1. Informations du document

### 1.1. Auteurs et contributeurs

Auteurs	Rôle au Health Data Hub
Stéphanie Combes	Directrice
Louis Pery	Directeur technique
Samuel Hourì	Product Owner Data
Emilie Courtois	Directrice des Services aux bases de données de santé
Anis Hamroun	Développeur
Pierre-Casimir Ravel	Responsable des systèmes d'information internes

## 1.2. Sommaire

<b>1. Informations du document</b>	<b>2</b>
1.1. Auteurs et contributeurs	2
1.2. Sommaire	3
<b>2. Introduction</b>	<b>5</b>
2.1 Objectif du guide	5
2.2 Objectif du catalogue de métadonnées	6
<b>3. Modèle de métadonnées</b>	<b>6</b>
3.1. Introduction - Modèle DCAT-AP	6
3.2. Modèle Health DCAT-AP	7
3.3. Modèle du catalogue de métadonnées	8
3.3.1. Catalogue	8
3.3.2 Base de données (classe dataset)	10
3.3.3. Modalités d'accès (Distribution)	15
3.3.4. Tables (Sample distribution)	16
3.3.5. Variables (extension csvw)	17
<b>4. Gestion des métadonnées</b>	<b>18</b>
4.1. Gestion des métadonnées sans accès à l'admin	18
4.1.1 Créer un fichier de métadonnées	19
4.1.2. Base de données	20
4.1.3. Modalités d'accès	21
4.1.4. Tables et fichiers	21
4.1.5. Variables	22
4.1.6. Validation des métadonnées	24
4.1.7. Téléchargement du fichier	25
4.1.8. Importation d'un fichier	25
4.2. Gestion des métadonnées en tant qu'administrateur	26
4.2.1 Créer les métadonnées d'une base de données	26
4.2.2. Métadonnées du catalogue	28
<b>5. Déploiement de l'application</b>	<b>29</b>
5.1. Configurer les Helm Charts	29
5.2. Personnaliser l'application	30
5.2.1. Personnalisation des éléments de forme	30
5.2.2 Personnalisation du pied de page	30
5.3. Métadonnées du catalogue	31
5.4. Personnaliser Keycloak	31
<b>6. Gestion des utilisateurs</b>	<b>31</b>
6.1. Création d'un administrateur utilisateurs (admin keycloak)	32
6.2. Création d'un administrateur fonctionnel	33
6.3. Création d'un administrateur de bases de données	35
6.4. Modifier un rôle	36

6.5. Modifier un mot de passe	36
<b>7. Personnalisation du catalogue</b>	<b>37</b>
7.1. Personnalisation du thème (couleur, logo, police...)	37
7.2. Personnalisation des éléments du pied de page	38
<b>8. Interface utilisateur</b>	<b>39</b>
8.1. Bases de données	40
8.2. Tables, fichiers et variables	40
8.3. Ressources	41
<b>9. Annexes</b>	<b>42</b>
9.1. Aperçu du fichier settings.json	42
9.2 Diagramme des métadonnées	43
1. Informations du document	2
1.1. Auteurs et contributeurs	2
1.2. Sommaire	3
2. Introduction	5
2.1 Objectif du guide	5
2.2 Objectif du catalogue de métadonnées	6
3. Modèle de métadonnées	6
3.1. Introduction - Modèle DCAT-AP	6
3.2. Modèle Health DCAT-AP	7
3.3. Modèle du catalogue de métadonnées	8
3.3.1. Catalogue	8
3.3.2 Base de données (classe dataset)	10
3.3.3. Modalités d'accès (Distribution)	15
3.3.4. Tables (Sample distribution)	16
3.3.5. Variables (extension csvw)	17
4. Gestion des métadonnées	18
4.1. Gestion des métadonnées sans accès à l'admin	18
4.1.1 Créer un fichier de métadonnées	19
4.1.2. Base de données	20
4.1.3. Modalités d'accès	21
4.1.4. Tables et fichiers	21
4.1.5. Variables	22
4.1.6. Validation des métadonnées	24
4.1.7. Téléchargement du fichier	25
4.1.8. Importation d'un fichier	25
4.2. Gestion des métadonnées en tant qu'administrateur	26
4.2.1 Créer les métadonnées d'une base de données	26
4.2.2. Métadonnées du catalogue	28
5. Déploiement de l'application	29
5.1. Configurer les Helm Charts	29
5.2. Personnaliser l'application	30
5.2.1. Personnalisation des éléments de forme	30

5.2.2 Personnalisation du pied de page	30
5.3. Métadonnées du catalogue	31
5.4. Personnaliser Keycloak	31
6. Gestion des utilisateurs	31
6.1. Création d'un administrateur utilisateurs (admin keycloak)	32
6.2. Création d'un administrateur fonctionnel	33
6.3. Création d'un administrateur de bases de données	35
6.4. Modifier un rôle	36
6.5. Modifier un mot de passe	36
7. Personnalisation du catalogue	37
7.1. Personnalisation du thème (couleur, logo, police...)	37
7.2. Personnalisation des éléments du pied de page	38
8. Interface utilisateur	39
8.1. Bases de données	40
8.2. Tables, fichiers et variables	40
8.3. Ressources	41
9. Annexes	42
9.1. Aperçu du fichier settings.json	42
9.2 Diagramme des métadonnées	43

## 2. Introduction

### 2.1 Objectif du guide

Ce guide a pour objectif d'accompagner les organismes de santé et partenaires dans le déploiement de leur propre instance du catalogue de métadonnées, basé sur celui du Health Data Hub (HDH). Cette démarche leur permet de personnaliser et héberger localement leur catalogue de métadonnées, tout en conservant la possibilité de contribuer aux catalogues de métadonnées européens.

Le processus repose sur :

- la mise à disposition en open source du code du catalogue de métadonnées du HDH ;
- l'utilisation de la documentation technique fournie pour faciliter le déploiement, la configuration et la personnalisation du catalogue de métadonnées local.

Pour mener à bien ce déploiement, les partenaires doivent disposer :

- d'une capacité d'hébergement du catalogue de métadonnées dans leur propre environnement ;
- d'une équipe technique interne capable d'assurer la maintenance et les évolutions du catalogue de métadonnées local.

Le document est structuré de manière à répondre aux besoins des différents profils impliqués :

- **Parties 3 et 4** : elles s'adressent aux **administrateurs de bases de données**. Elles décrivent le modèle de métadonnées européen intégré dans l'outil et la gestion de ces métadonnées, notamment la création, la structuration et la publication des informations.
- **Partie 5** : elle accompagne les **administrateurs système** dans le déploiement technique de l'application à partir du code source.
- **Partie 6** : elle est destinée aux **administrateurs utilisateurs**, avec un focus sur la gestion des comptes et des droits d'accès.
- **Partie 7** : elle concerne les **administrateurs fonctionnels**, en détaillant les possibilités de personnalisation du catalogue de métadonnées (paramétrage visuel, contenus éditoriaux, ajustements organisationnels).
- **Partie 8** : elle fournit une vue d'ensemble destinée aux **utilisateurs finaux**, leur permettant d'explorer et de consulter les métadonnées d'une base via l'interface du catalogue.

Le guide a pour vocation d'offrir une approche progressive et adaptée aux différents rôles, afin de garantir une mise en œuvre réussie, une gestion cohérente et une utilisation efficace du catalogue de métadonnées dans chaque organisation.

## 2.2 Objectif du catalogue de métadonnées

Le catalogue de métadonnées est une application conçue pour documenter, gérer et publier les métadonnées des bases de données de santé. Il s'appuie sur le standard Health DCAT-AP, garantissant une conformité aux normes européennes et une interopérabilité avec les catalogues de métadonnées développés dans le cadre de l'Espace Européen des Données de Santé (EHDS).

Ses objectifs principaux sont les suivants :

- **structurer et normaliser la documentation des bases de données de santé**, en s'alignant sur les standards européens pour faciliter leur compréhension et leur valorisation ;
- **faciliter l'identification des bases** pertinentes par les chercheurs, sans exposer les données sensibles, grâce à des descriptions riches, accessibles et standardisées ;
- **améliorer l'expérience utilisateur** grâce à une interface intuitive permettant une navigation fluide dans les métadonnées ;
- **simplifier la contribution des partenaires**, via des formulaires guidés et des champs préconfigurés, réduisant les efforts techniques nécessaires ;
- **offrir une solution personnalisable en marque blanche**, permettant à chaque organisation d'intégrer le catalogue de métadonnées dans son propre environnement, avec son identité visuelle et sa structure organisationnelle ;
- **favoriser la diffusion des métadonnées à l'échelle européenne**, tout en maintenant une autonomie locale dans la gestion et la publication des informations.

Ce catalogue s'inscrit dans une démarche de gouvernance des données de santé, visant à :

- renforcer la transparence et la qualité des données ;
- répondre aux exigences réglementaires européennes ;
- accélérer les projets de recherche et d'innovation ;
- encourager la collaboration entre institutions, chercheurs et partenaires.

Déployé en marque blanche, il permet à chaque organisation de disposer d'une instance locale adaptée à ses besoins, tout en garantissant une interopérabilité complète avec les outils européens.

## 3. Modèle de métadonnées

### 3.1. Introduction - Modèle DCAT-AP

Dans le cadre de l'espace européen des données de santé (EHDS), le partage et la réutilisation des données de santé à des fins secondaires (recherche, innovation, pilotage des politiques publiques...) nécessitent un langage commun pour décrire les jeux de données, leurs conditions d'accès et leurs modalités d'utilisation. Ce langage prend la forme d'un **modèle de métadonnées**, structurant les informations de manière **Findable, Accessible, Interoperable and Reusable (FAIR)**.

Le modèle de référence utilisé est **Health DCAT-AP**, une extension sectorielle du standard **DCAT (Data Catalog Vocabulary)** développé par le W3C. DCAT repose sur les principes du **web sémantique** et permet de représenter les métadonnées sous forme de graphes RDF, facilitant leur exploitation automatisée et leur interopérabilité entre catalogues.

L'utilisation de DCAT présente plusieurs avantages. Elle facilite la compatibilité avec d'autres catalogues de données en Europe et permet l'alignement avec des vocabulaires complémentaires (Dublin Core, SKOS ou, PROV, etc.).

### **Dans DCAT, trois concepts fondamentaux organisent le modèle :**

- **Catalog** (*dcat:Catalog*) : collection de jeux de données publiée par une organisation.
- **Dataset** (*dcat:Dataset*) : entité conceptuelle représentant des données qui ont été préparées, publiées ou mises à disposition sous licence par un même organisme. Il s'agit d'une unité de description permettant de regrouper les informations essentielles sur une base ou un jeu de données, indépendamment de sa forme technique ou de son mode d'accès. Cette classe est traduite par "base de données" dans l'application catalogue.
- **Distribution** (*dcat:Distribution*) : une représentation de l'incarnation technique d'un dataset (jeu de données) dans un format donné. Il s'agit d'une version concrète du dataset, accessible ou disponible sur demande, selon qu'il s'agisse de données ouvertes ou protégées. La distribution décrit les modalités d'accès aux données (fichier téléchargeable, API, base de données, etc.) et permet aux utilisateurs de comprendre comment interagir avec le contenu réel du jeu de données. Cette classe est traduite par "modalités d'accès" dans l'application.

## **3.2. Modèle Health DCAT-AP**

Pour répondre aux besoins spécifiques liés aux données sensibles et aux exigences du règlement européen sur l'espace des données de santé (EHDS), le modèle est enrichi via le profil **Health DCAT-AP**.

Cette extension est conçue pour les métadonnées de santé. Là où DCAT propose un cadre générique, **Health DCAT-AP** introduit des règles et des champs supplémentaires adaptés aux enjeux du secteur : sensibilité des données, modalités d'accès, base légale (RGPD, réglementations nationales), niveau d'anonymisation, type de données (administratives, hospitalières, cohortes...), organisme responsable et conditions de réutilisation.

Le modèle existe en trois variantes selon le niveau de protection des données : **données ouvertes**, **données protégées**, et **données sensibles**. Il intègre également des métadonnées opérationnelles comme le nombre de patients inclus, la période de collecte, la fréquence de mise à jour ou la liste des établissements contributeurs. Ces éléments rendent le catalogue plus lisible et utile pour les chercheurs et les acteurs publics dans la santé.

Pour aller plus loin :

- La [spécification DCAT](#) est disponible sur le site du **W3C**.
- La documentation officielle de [Health DCAT-AP](#) est accessible via le **European Health Information Portal**.

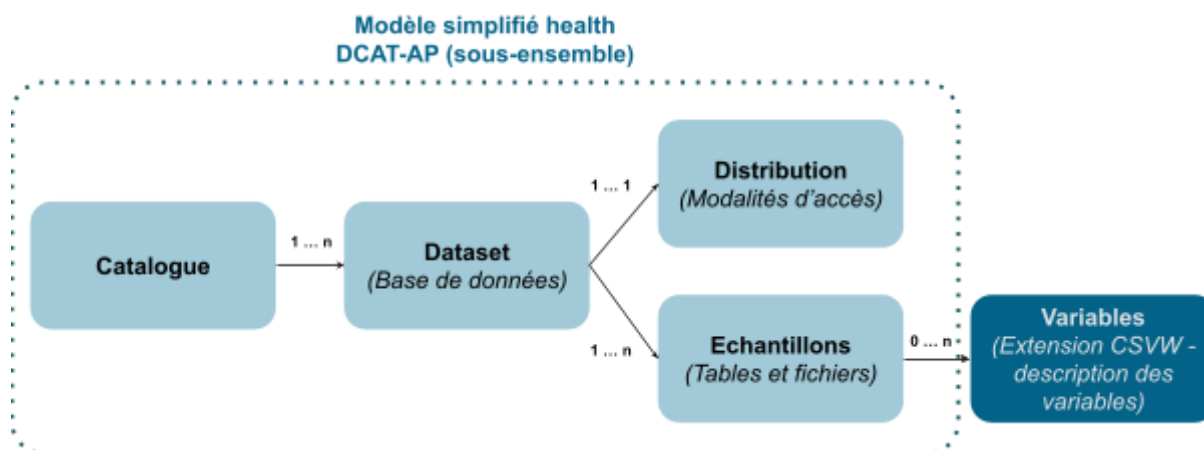


### 3.3. Modèle du catalogue de métadonnées

L'application du catalogue de métadonnées développée par le Health Data Hub respecte le standard européen **Health DCAT-AP**, qui définit comment structurer et documenter les métadonnées de santé. Chaque instance de l'application repose sur un **nœud central** appelé `dcat:Catalog`, qui regroupe l'ensemble des métadonnées. Ce catalogue peut contenir plusieurs jeux de données (`dcat:Dataset`), chacun décrit selon les champs définis par le standard, ainsi que leurs modalités de diffusion (`dcat:Distribution`).

Par défaut, le niveau de sensibilité des données est paramétré comme **sensible**, non pas parce que la description elle-même est confidentielle, mais parce que le catalogue vise des données partagées dans des environnements sécurisés, et non des données ouvertes en libre accès.

Cette version de l'application propose une sélection réduite de métadonnées par rapport au standard Health DCAT-AP, afin d'en faciliter l'usage. Le modèle reste évolutif et pourra être ajusté progressivement. Voir le diagramme détaillé en [annexe 9.2](#)



#### 3.3.1. Catalogue

Le **catalogue** correspond à une instance déployée par chaque utilisateur (*institution, région, ou même EDS privé*). Il regroupe l'ensemble des métadonnées. Plusieurs champs permettent de le décrire clairement : son nom, sa finalité, sa portée (locale, nationale...), sa langue, son éditeur, et ses dates de création ou de mise à jour.

D'autres métadonnées permettent d'assurer la traçabilité et l'interopérabilité du catalogue dans l'écosystème européen : l'identifiant unique (`dct:identifier`), les droits d'usage (`dct:rights`), ou encore les liens vers d'autres nœuds catalogues ou d'autres métadonnées des base de données (`dcat:dataset`, `dcat:catalog`).

Ces éléments doivent être renseignés dès le départ et conservés tout au long de la vie du catalogue. Il faut les saisir une première fois (cf. [partie 2](#) pour la méthode technique, [partie 2](#) via l'interface) et les conserver dans le formulaire toute la durée de vie du catalogue.

Nom de la métadonnée dans le catalogue	Description	Propriété Health DCAT-AP
<b>Titre</b>	Nom du catalogue de métadonnées.	dct:title
<b>Description</b>	Description du catalogue de métadonnées.	dct:description
<b>Site internet</b>	Site internet.	foaf:homepage
<b>Responsable du catalogue</b>	Organisation responsable du catalogue de métadonnées.	dct:publisher
<b>Nom</b>	Nom du responsable du catalogue de métadonnées.	dct:publisher / foaf:name
<b>Site internet</b>	Site internet du responsable du catalogue de métadonnées.	dct:publisher / foaf:homepage
<b>Adresse email</b>	Adresse email du responsable du catalogue de métadonnées.	dct:publisher / foaf:mbox
<b>Créateur du catalogue</b>	Créateur du catalogue de métadonnées.	dct:creator
<b>Nom</b>	Nom de l'organisation ayant créé le catalogue de métadonnées.	dct:creator / foaf:name
<b>Site internet</b>	Site internet de l'organisation ayant créé le catalogue de métadonnées.	dct:creator / foaf:homepage
<b>Adresse email</b>	Adresse email de l'organisation ayant créé le catalogue de métadonnées.	dct:creator / foaf:mbox
<b>Réglementation applicable</b>	Législation encadrant la création ou la gestion du catalogue de métadonnées.	dcatap:applicableLegislation (eli:LegalResource)
<b>Thème</b>	Système d'organisation des connaissances pour classifier les ensembles de données.	dcat:themeTaxonomy (skos:ConceptScheme)

<b>Langue</b>	Langue selon ISO 639-2, ISO 639-3 ou tag IETF BCP47.	dct:language
<b>Licence</b>	Licence du catalogue de métadonnées.	dct:license
<b>Couverture géographique</b>	Couverture géographique du catalogue de métadonnées.	dct:spatial
<b>Base de données</b>	Une entité conceptuelle représentant l'information publiée (correspond à la base).	dcat:dataset

### 3.3.2 Base de données (classe dataset)

La classe *Base de données (Dataset)* permet de décrire une base de données en tant qu'entité documentaire complète et cohérente. Elle centralise toutes les informations essentielles liées à une base : son contenu, son responsable, sa structure, ses modalités d'accès et ses usages.

Dans le catalogue, chaque base de données est ainsi décrite comme une entité complète via la classe **Dataset**. Pour aider les chercheurs à mieux identifier l'intérêt d'une base, nous fournissons également des métadonnées sur les **tables qui la composent** (voir ci-dessous). Cela permet de mieux comprendre la structure et le contenu des données disponibles. Chaque base doit disposer d'un titre clair (*dct:title*) et d'une description détaillée (*dct:description*) permettant à l'utilisateur de comprendre son périmètre, les thématiques couvertes, et les éventuelles restrictions d'usage. L'organisation responsable de la base est indiquée via la propriété *dct:creator*, tandis que la langue utilisée pour décrire les champs ou la documentation peut être précisée avec *dct:language*.

Il est également possible d'ajouter des dates clés telles que la date de création (*dct:issued*) ou de dernière mise à jour (*dct:modified*), ainsi que des informations sur le cadre réglementaire applicable (*dct:rights*) et les législations en vigueur (*dcatap:applicableLegislation*).

La classe base de données peut aussi être liée à des modalités d'accès (*dcat:distribution*) ou des services associés (*dcat:accessService*), facilitant son accès.

*NB - Il est aussi possible, dans certains cas, de modéliser une table comme un Dataset à part entière — par exemple si elle est diffusée indépendamment ou répond à des objectifs spécifiques. Ce choix dépend des pratiques de gestion et de publication de chaque organisation.*

Nom de la métadonnée dans le catalogue	Description	Propriété Health DCAT-AP
<b>Titre de la base de données</b>	Nom complet de la base de données.	dct:title
<b>Titre abrégé</b>	Nom abrégé de la base de données.	dct:alternative
<b>Description de la base de données</b>	Description de la base en texte libre : périmètre, type, thématique, critères d'inclusion, chaînage éventuel avec le SNDS.	dct:description
<b>Domaine médical</b>	Domaine médical associé à la base. Plusieurs thèmes possibles. Les URIs de Wikidata doivent être utilisés.	healthdcatap:healthTheme
<b>Mots-clés</b>	Liste de mots-clés associés, pour faciliter l'identification.	dcat:keyword
<b>Objectif initial de la base</b>	Contexte de la collecte des données (texte libre).	dct:purpose
<b>Population d'intérêt</b>	Description de la population dont les données sont contenues dans la base.	healthdcatap:populationCoverage
<b>Couverture géographique</b>	Région géographique couverte (pays, NUTS). Attention à éviter les contradictions si pays + sous-régions.	dct:spatial
<b>Profondeur historique – Date de début</b>	Date de début de la période temporelle (format MM/AAAA).	dct:temporal (startDate)
<b>Profondeur historique – Date de fin</b>	Date de fin de la période temporelle (format MM/AAAA).	dct:temporal (endDate)
<b>Sources et processus de collecte</b>	Source et processus de collecte (dossiers médicaux, dispositifs, questionnaires, qualité...).	dct:provenance

<b>Année de mise à disposition</b>	Date de mise à disposition de la base (après primo-accès).	dct:issued
<b>Fréquence de mise à jour</b>	Fréquence des mises à jour.	dct:accrualPeriodicity
<b>Langue de production</b>	Langue de la base (ISO 639 ou IETF BCP47).	dct:language
<b>Nombre d'entrées</b>	Nombre total de lignes dans la base.	healthdcatap:numberOfRecords
<b>Nombre d'individus uniques</b>	Nombre d'individus distincts dans la base.	healthdcatap:numberOfUniqueIndividuals
<b>Âge minimal en année</b>	Âge minimal moyen. Doit être arrondi pour l'anonymat.	healthdcatap:minTypicalAge
<b>Âge maximal en année</b>	Âge maximal moyen. Même règle d'anonymat.	healthdcatap:maxTypicalAge
<b>Catégories de données</b>	Catégorie de données selon le règlement EHDS.	healthdcatap:healthCategory
<b>Classification – Codes</b>	Codes CIM-10 et URL correspondante.	healthdcatap:codeValues
<b>Classification Identifier</b>	Identifiant de classification (CIM-10).	dct:identifier
<b>Classification Notation</b>	Code ou abréviation de la classification.	skos:notation
<b>Code Value Identifier</b>	URL de la classification ou sous-catégorie.	skos:Concept/@uri
<b>Code Value Notation</b>	Code ou abréviation de la valeur.	skos:notation
<b>Terminologies utilisées</b>	Terminologies médicales (CIM-10, CCAM, SNOMED CT) via URIs Wikidata.	healthdcatap:codingSystem
<b>Identifiant(s) secondaire(s)</b>	Identifiant secondaire (DOI, DataCite, etc.).	adms:identifier

<b>Ressources liées</b>	Ressource liée via URL et rôle associé.	dcat:qualifiedRelation
<b>Publications</b>	Publications liées à la base (DOI, identifiants pérennes).	dct:isReferencedBy
<b>Encadrement réglementaire</b>	Base légale justifiant le traitement des données personnelles.	dqv:hasLegalBasis
<b>Label de qualité</b>	Annotation de qualité, certificats, retours associés.	dqv:hasQualityAnnotation
<b>Responsable de données (créateur)</b>	Organisation responsable de la base.	dct:creator / foaf:name
<b>Site internet du responsable</b>	Site du responsable.	dct:creator / foaf:homepage
<b>Email du responsable</b>	Contact email du responsable.	dct:creator / foaf:mbox
<b>Responsable de la mise à disposition</b>	Organisation responsable de la mise à disposition.	dct:publisher
<b>Site internet du publisher</b>	Site web du publisher.	dct:publisher / foaf:homepage
<b>Email du publisher</b>	Email du publisher.	dct:publisher / foaf:mbox
<b>Description du publisher</b>	Description des activités du publisher.	dcat:publisherNote
<b>Statut juridique du publisher</b>	Type d'organisation selon vocabulaire contrôlé.	dcat:publisherType
<b>Organisme point de contact</b>	Organisme du point de contact.	dcat:contactPoint
<b>Email du point de contact</b>	Email du contact.	dcat:contactPoint
<b>Site internet du contact</b>	Site du contact.	dcat:contactPoint

<b>Modalités d'accès à la base</b>	Modalités d'accès, lien vers l'organisme d'accès aux données.	dcat:distribution
<b>Durée de conservation – début</b>	Date de début de conservation.	dcat:retentionPeriod (startDate)
<b>Durée de conservation – fin</b>	Date de fin de conservation.	dcat:retentionPeriod (endDate)
<b>Thème</b>	Thème imposé par défaut : Santé (et autres si nécessaires).	dcat:theme
<b>Législation applicable</b>	Texte de loi imposant la gestion des données (ex : EHDS).	dcatap:applicableLegislation
<b>Droits d'accès</b>	Niveau d'accès (par défaut sensible pour les données de santé).	dct:accessRights
<b>Identifiant principal</b>	Identifiant unique de la base (URI HTTP obligatoire).	dct:identifier
<b>Responsable de données légitime</b>	Détenteur de données fiable.	healthdcatap:trustedDataHolder
<b>Type de données</b>	Type de données contenues (par défaut : personnelles).	dct:type
<b>Plateforme nationale d'accès</b>	Plateforme nationale EHDS (par défaut : HDH).	healthdcatap:healthDataAccessBody
<b>Site internet de la plateforme</b>	Site web de la plateforme.	foaf:homepage
<b>Email de la plateforme</b>	Email de la plateforme.	foaf:mbox
<b>Pays de la plateforme</b>	Pays de rattachement de la plateforme.	dct:spatial
<b>Référentiels utilisés</b>	Référentiels, nomenclatures et standards utilisés.	dct:conformsTo
<b>Analyses disponibles</b>	URI vers analyses, indicateurs ou services liés.	healthdcatap:analytics

<b>Données personnelles</b>	Éléments clés représentant un individu (enum).	dpv:hasPersonalData
-----------------------------	--	---------------------

### 3.3.3. Modalités d'accès (Distribution)

Nom de la métadonnée dans le catalogue	Description	Propriété Health DCAT-AP
<b>Mise à jour de la base de données</b>	Statut de la base de données (collective active ou terminée).	adms:status
<b>Titre de la base de données mise à disposition</b>	Nom donné à la base mise à disposition (titre, éventuellement complété d'informations sur sa structure).	dct:title
<b>Description des modalités de mise à disposition</b>	Description des modalités d'accès (ex. accès sur autorisation ou via contact).	dct:description
<b>URL d'accès</b>	URL d'accès à la modalité d'accès (par défaut : site internet du responsable de la mise à disposition).	dcat:accessURL
<b>Taille de la base de données (en octets)</b>	Taille approximative de la base en octets.	dcat:byteSize
<b>Format des données</b>	Format des données contenues dans la distribution (par exemple csv, json, dicom).	dct:format
<b>URL de téléchargement</b>	Lien direct vers un fichier téléchargeable dans un format donné (non applicable pour des données sensibles).	dcat:downloadURL
<b>Cadre réglementaire</b>	Description du cadre réglementaire encadrant l'accès à la base.	dct:rights
<b>Service d'accès à la base de données</b>	Service sécurisé donnant accès aux données (portail sécurisé, API, documentation, support).	dcat:accessService
<b>Réglementation applicable</b>	Réglementation encadrant la mise à disposition de la base de données.	dcatap:applicableLegislation



### 3.3.4. Tables (Sample distribution)

Les **tables et fichiers** sont décrits à l'aide de la classe **Sample Distribution**. Cette classe est utilisée pour représenter de manière concrète le contenu physique d'une base de données, par exemple au travers de fichiers ou de tables.

L'objectif est d'offrir une vision réaliste de la structure des données. Il est également recommandé de renseigner un titre clair (*dct:title*) et une description libre (*dct:description*) afin de contextualiser chaque fichier ou table.

Les tables et fichiers décrits peuvent être rendus encore plus explicites grâce à la mise à disposition d'un **échantillon anonymisé**, qui peut être :

- un sous-ensemble anonymisé ou synthétique qui conserve les caractéristiques essentielles du jeu de données original sans révéler d'informations personnelles ;
- un fichier illustrant la structure du jeu de données.

Chaque échantillon doit alors inclure :

- une **URL de téléchargement directe** (*dcat:downloadURL*) pointant vers le fichier anonymisé ;
- le **format du fichier** (*dct:format*), tel que **CSV**, **JSON** ou **DICOM**, afin d'informer les utilisateurs sur sa nature technique.

A noter que si l'accès à des échantillons passe par une API ou un portail sécurisé, le service concerné peut être précisé via la propriété *dcat:accessService*, accompagné d'une *dcat:accessURL* donnant les modalités d'accès.

Nom de la métadonnée dans le catalogue	Description	Propriété Health DCAT-AP
<b>Titre de la table ou du fichier</b>	Nom de la table ou du fichier (si la base n'est pas tabulaire), saisi en texte libre.	dct:title
<b>Description de la table ou du fichier</b>	Description de la table issue de la base de données (ou du fichier, ex. JSON), en texte libre.	dct:description
<b>URL de téléchargement de l'échantillon anonymisé</b>	Lien direct vers un échantillon anonymisé de la base de données. Chaque base doit fournir au minimum un échantillon anonymisé.	dcat:downloadURL
<b>Format de l'échantillon anonymisé</b>	Format des données de l'échantillon (par exemple CSV, JSON, DICOM).	dct:format

<b>Service d'accès à la table (échantillon)</b>	Service sécurisé permettant d'accéder à la table ou au fichier (portail sécurisé, API, documentation).	dcats:accessService
<b>URL d'accès</b>	Lien décrivant les modalités d'accès à la base de données. Il ne s'agit pas forcément d'un lien direct vers les données, mais d'une URL donnant accès à des informations ou à la procédure d'accès.	dcats:accessURL
<b>Cadre réglementaire</b>	Explicitation du cadre réglementaire d'accès à la table ou au fichier échantillonné/anonymisé.	dct:rights
<b>Taille de l'échantillon anonymisé (en octets)</b>	Taille approximative de l'échantillon, exprimée en octets (ou gigaoctets).	dcats:byteSize
<b>Réglementation applicable</b>	Texte légal encadrant la mise à disposition de l'échantillon de la base.	dcats:applicableLegislation

### 3.3.5. Variables (extension csvw)

Pour faciliter la compréhension des bases de données, il est recommandé de fournir un dictionnaire de données permettant de décrire les variables contenues dans les tables. Ce dictionnaire peut être lié à une table ou un fichier via la classe décrite précédemment (Sample Distribution). Ce dictionnaire de données doit être fondé sur le vocabulaire [CSVW](#) et est facultatif dans le modèle de métadonnées HDCAT-AP.

L'onglet "Variables" permet de décrire chaque colonne d'une table :

- Un **titre métier** (`dct:title`) explicite le sens de la donnée, indépendamment du nom technique.
- Le **nom technique** (ex. : `num_patient`, `dt_naissance`) correspond à celui utilisé dans le système source.
- Une **description** précise le contenu, les valeurs possibles ou le contexte d'usage.
- Le **type de variable** (`dct:type`) indique son format (texte, entier, date...).
- Enfin, il est utile de préciser si la variable joue un rôle structurel (clé primaire, clé étrangère), pour mieux comprendre les relations entre tables.\*\*

Nom de la métadonnée dans le catalogue	Description	Propriété Health DCAT-AP
--	-------------	--------------------------

<b>Titre technique de la table ou fichier</b>	Nom technique de la table, tel qu'il est codé dans la base de données.	N/A
<b>Titre métier de la variable</b>	Nom métier qui explicite le sens du nom technique de la variable.	dct:title
<b>Type de la variable</b>	Type de la variable (nombre, binaire, date, etc.).	dct:type
<b>Nom technique de la variable</b>	Nom technique de la variable tel qu'il est codé dans la base.	N/A
<b>Description de la variable</b>	Description textuelle de la variable.	N/A
<b>Clé primaire</b>	Indique si la variable est une clé primaire.	csvw:primaryKey
<b>Clé étrangère</b>	Indique si la variable est une clé étrangère.	csvw:foreignKey

## 4. Gestion des métadonnées

### 4.1. Gestion des métadonnées sans accès à l'admin


Pour générer des métadonnées, il suffit d'accéder à l'éditeur en ligne via l'URL personnalisée suivante : **[https://editor.<nom\\_de\\_votre\\_domaine>](https://editor.<nom_de_votre_domaine>)**.

Cet éditeur propose un formulaire intuitif permettant de décrire une base de données et de produire automatiquement un fichier de métadonnées.

**Le formulaire peut être complété progressivement, à votre rythme.** À tout moment, vous pouvez **télécharger le fichier au format XLSX**, même si le formulaire n'est pas entièrement rempli. Le téléchargement au format TTL est disponible uniquement une fois le formulaire entièrement validé et conforme au standard européen HDCAT-AP.

**L'écran principal vous permet de saisir les informations directement dans les champs prévus. Une fois la saisie terminée** ou en cours, le bouton de téléchargement situé en haut à droite permet de récupérer le fichier dans le format souhaité. L'éditeur prend également en charge l'import d'un fichier XLSX déjà existant, que vous pouvez modifier, enrichir ou corriger directement dans l'interface.

La suite du guide détaille pas à pas comment utiliser cet éditeur pour créer un fichier de métadonnées complet et conforme.



#### 4.1.1 Créer un fichier de métadonnées

L'éditeur en ligne **ne permet pas de publier directement les métadonnées**, mais sert à générer un fichier structuré décrivant votre base de données. Pour cela, il vous suffit de compléter tous les champs du formulaire, en prêtant attention aux champs obligatoires indiqués par un astérisque (\*), et en veillant à ce que chaque groupe de champs soit correctement validé (ex. URL, menu déroulants...)

**Le respect du vocabulaire contrôlé et des contraintes imposées est essentiel :** certaines valeurs doivent être sélectionnées dans des menus déroulants, d'autres respecter un format précis (comme les URL). L'interface vous guide étape par étape grâce au bouton "Suivant", qui vous permet de naviguer facilement dans le formulaire en suivant un chemin logique de saisie.

**À noter :** l'accès à l'échantillon de la base ne peut pas être configuré depuis l'éditeur ; cette option est disponible uniquement via l'interface d'administration. Par ailleurs, les boutons "+" vous permettent d'ajouter plusieurs valeurs pour un même champ ou de saisir des **versions multilingues**, en cohérence avec les cardinalités du modèle (par exemple les champs de type 1...n).

## 4.1.2. Base de données

💡 Utilisation du formulaire en ligne pour générer un fichier de métadonnées

Ce formulaire vous permet de créer un fichier de métadonnées décrivant une base de données, aux formats XLSX ou TTL.  
Une fois complété (ou au cours de sa complétion), vous pouvez le télécharger grâce au bouton téléchargement situé en haut à droite de l'écran.

- ◆ Téléchargement au format TTL : disponible uniquement si le fichier est validé et conforme au standard européen de métadonnées de santé (HDCAT-AP)
- ◆ Chargement ou téléchargement d'un fichier XLSX : possible à tout moment, même si le fichier est partiellement rempli.

*Veillez noter que les traductions de l'anglais vers le français sont actuellement en cours de retravail. Certaines formulations ou termes peuvent encore évoluer afin d'améliorer la clarté et la cohérence des contenus.*

Base de données   Modalités d'accès   Tables et fichiers   Variables   Aperçu

✓ Enregistré le 10:43:36

**Section de la base de données**

- T** Vue d'ensemble de la base de données
- Caractéristiques de la base de données
- Ressources associées à la base de données
- Responsable de la mise à disposition de la base de données
- Responsable de données
- Points de contact
- Plateforme nationale des données de santé

### Vue d'ensemble de la base de données

**Droits d'accès \***

*Indication du niveau d'accès de la base de données ; les données personnelles de santé sont par défaut considérées sensibles, comme ayant un accès 'sensible'.*

Non public

**Titre de la base de données \***

*Nom complet de la base de données.*

Langue : Français   Titre de la base de données : titre

**Titre abrégé**

*Nom abrégé de la base de données.*

**Description de la base de données \***

*Description de la base de données (périmètre) en texte libre. Précisez le type de base, la thématique, le principal critère d'inclusion et si elle est chaînée ou non à la base principale du SNDS.*

Langue :   Description de la base de données :

Les champs de la section « **Base de données** » (classe *Dataset* du modèle Health DCAT-AP) sont organisés en 7 parties afin de structurer et simplifier la saisie :

- **Vue d'ensemble de la base de données** : informations générales du dataset - titre, description synthétique, droits d'accès, mots-clés, thématiques, couverture géographique et temporelle.
- **Caractéristiques de la base de données** : métadonnées descriptives spécifiques à la base : catégorie, granularité des données, description technique, nombre d'enregistrements ou d'individus, distribution par tranches d'âge, nature des données personnelles éventuellement présentes, etc.
- **Ressources associées à la base de données** : liste des ressources reliées au dataset : identifiants persistants, publications de référence, schémas, classifications ou autres jeux de données liés.
- **Responsable de la mise à disposition de la base de données** : organisation ou personne assurant la diffusion et la mise à disposition de l'ensemble de données.
- **Responsable de données** : producteur ou créateur de la base (organisation, service ou individu). Ce rôle peut être distinct ou identique à celui du point de contact ou du responsable de la mise à disposition. Dans le cadre du Health DCAT-AP, il est recommandé de réutiliser les mêmes URI (organisation ou personne) lorsque les rôles se recoupent.

- **Points de contact** : coordonnées des interlocuteurs pour échanger autour du dataset et de ses métadonnées (support, questions techniques, demandes d'accès).
- **Plateforme nationale des données de santé (Health Data Hub)** : informations spécifiques liées à l'hébergement et à la gestion de la base dans la plateforme nationale. Cette partie comprend plusieurs champs obligatoires et constitue l'une des sections les plus détaillées du formulaire.

Chaque partie regroupe plusieurs champs à renseigner. La section relative à la Plateforme nationale des données de santé est la plus détaillée et nécessite le plus de complétude.

#### 4.1.3. Modalités d'accès

Les informations d'accès à la base de données sont définies par les modalités d'accès (classe de distribution HDCAT-AP). Les champs obligatoires incluent le titre de la base et un lien vers la procédure d'accès.

Utilisation du formulaire en ligne pour générer un fichier de métadonnées

Ce formulaire vous permet de créer un fichier de métadonnées décrivant une base de données, aux formats XLSX ou TTL. Une fois complété (ou au cours de sa complétion), vous pouvez le télécharger grâce au bouton téléchargement situé en haut à droite de l'écran.

- ◆ Téléchargement au format TTL : disponible uniquement si le fichier est validé et conforme au standard européen de métadonnées de santé (HDCAT-AP)
- ◆ Chargement ou téléchargement d'un fichier XLSX : possible à tout moment, même si le fichier est partiellement rempli.

*Veuillez noter que les traductions de l'anglais vers le français sont actuellement en cours de retravail. Certaines formulations ou termes peuvent encore évoluer afin d'améliorer la clarté et la cohérence des contenus.*

Base de données Modalités d'accès Tables et fichiers Variables Aperçu

### Modalités d'accès

**Modalité d'accès**

- T Titre de la modalité d'accès \*
- Description des modalités de mise à disposition
- URL d'accès \*
- Réglementation applicable à la modalité d'accès \*
- Cadre réglementaire
- Détails du fichier
- Mise à jour de la base de données
- Service d'accès à la base de données

#### Titre de la modalité d'accès

Titre de la modalité d'accès \*

*Nom donné à la base de données mise à disposition (titre, complété d'informations sur sa structure si nécessaire).*

Langue  
Français





Titre de la modalité d'accès \*

PRÉCÉDENT SUIVANT

Étape 1 sur 8

#### 4.1.4. Tables et fichiers

La section « Tables et fichiers » (classe « echantillons hddatap ») répertorie toutes les tables ou tous les fichiers anonymisés correspondants. Vous devez créer autant de blocs qu'il y a de tables dans la base de données ou de fichiers à saisir. Pour commencer, saisissez votre première table comme indiqué ci-dessous.

 **Utilisation du formulaire en ligne pour générer un fichier de métadonnées**   

Ce formulaire vous permet de créer un fichier de métadonnées décrivant une base de données, aux formats XLS ou TTL.  
Une fois complété (ou au cours de sa complétion), vous pouvez le télécharger grâce au bouton téléchargement situé en haut à droite de l'écran.

- ◆ Téléchargement au format TTL : disponible uniquement si le fichier est validé et conforme au standard européen de métadonnées de santé (HDCAT-AP)
- ◆ Chargement ou téléchargement d'un fichier XLS : possible à tout moment, même si le fichier est partiellement rempli.

*Veillez noter que les traductions de l'anglais vers le français sont actuellement en cours de retravail. Certaines formulations ou termes peuvent encore évoluer afin d'améliorer la clarté et la cohérence des contenus.*

Base de données Modalités d'accès **Tables et fichiers** Variables Aperçu

---

**Modifier la table ou le fichier**

- T** Titre de la table ou du fichier \*
- Description de la table ou du fichier
- URL d'accès \*
- Réglementation applicable \*
- Cadre réglementaire
- Détails du fichier
- Service d'accès à la table (échantillon anonymisé)

### Titre de la table ou du fichier

Titre de la table ou du fichier \*


*Nom de la table (ou du 'fichier' de la base de données si non tabulaire) - en texte libre*

Langue  
Français


Titre de la table ou du fichier \*

ANNULER ENREGISTRER

Ajoutez ensuite autant de tables ou de fichiers que vous le souhaitez en cliquant sur le bouton « Ajouter une table ou un fichier » situé en haut à droite.

**Tables et fichiers** 

Ajouter une table ou un fichier associé à la base de données

Nom de la table  
Description de la table  
Aucune URL de téléchargement définie   
Tables: 0

MODIFIER SUPPRIMER

Pour ajouter un lien de téléchargement, cliquez sur le petit stylo et insérez l'URL d'un échantillon représentatif de la table ou du fichier. Veuillez noter qu'en tant qu'administrateur, vous avez la possibilité de charger le fichier directement même si vous ne disposez pas d'un lien.

#### 4.1.5. Variables

Chaque **table ou fichier** peut être associée à une ou plusieurs **variables** décrivant son modèle de données (nom du champ, description, type de donnée, clé primaire ou étrangère, etc.). Cette section ne fait pas partie du modèle **Health DCAT-AP**. Elle correspond à un **dictionnaire de données** permettant de décrire en détail toutes les variables présentes dans une base.

Pour renseigner une variable :

1. Accéder à la section « **Variables** ».

[Base de données](#)
[Modalités d'accès](#)
[Tables et fichiers](#)
[Variables](#)
[Aperçu](#)

- Indiquer en premier lieu le **nom technique de la table** ou du fichier déjà ajouté.



**Modifier les variables de**

- Titre technique de la table ou fichier
- Variables

**Titre technique de la table ou fichier**

Titre technique de la table ou fichier \*

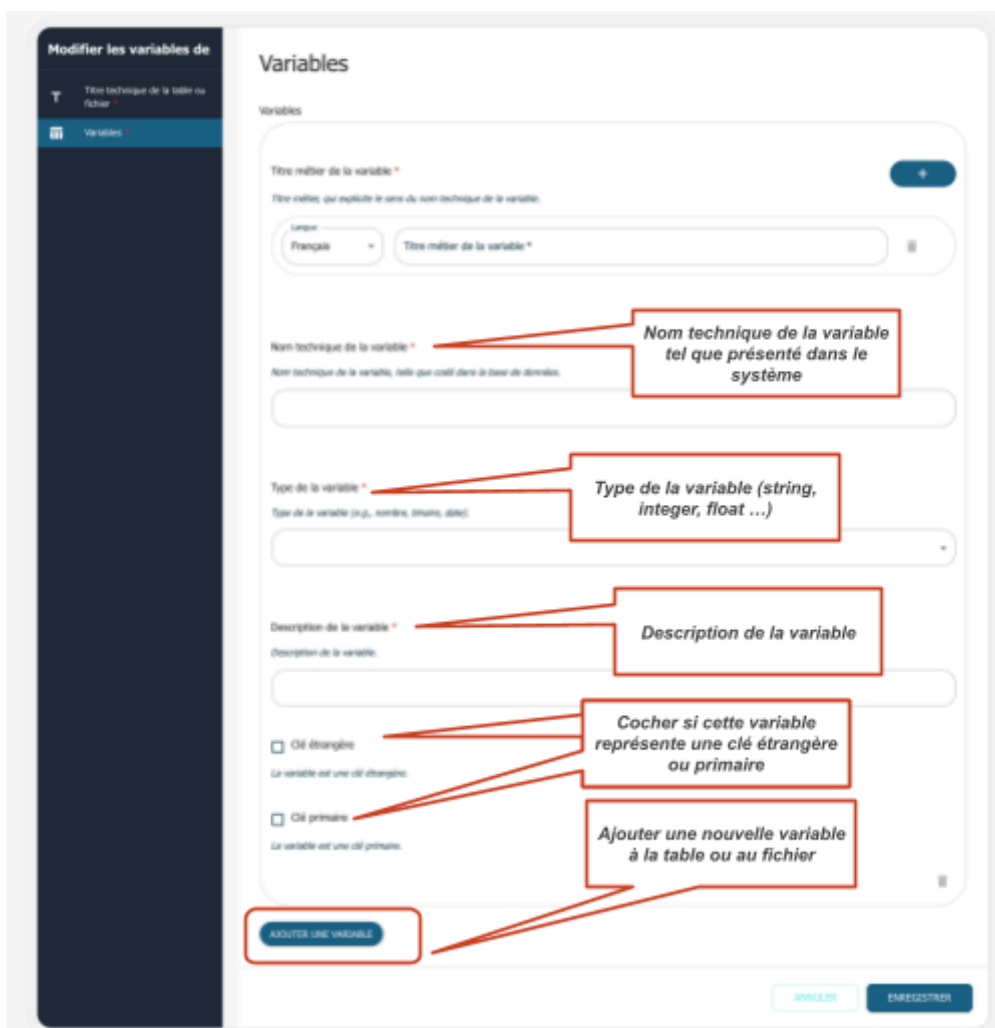
Nom technique de la table, telle que codée dans la base de données.

Langue: Anglais

Titre technique de la table ou fichier \*

ANNULER ENREGISTRER

- Ajouter ensuite les variables une à une en cliquant sur le bouton « **Ajouter une variable** » (en bas à gauche).



**Modifier les variables de**

- Titre technique de la table ou fichier
- Variables

**Variables**

Variables

Titre maître de la variable \*

Titre maître, qui explicite le sens du nom technique de la variable.

Langue: Français

Titre maître de la variable \*

Nom technique de la variable \*

Nom technique de la variable, telle que codée dans la base de données.

Type de la variable \*

Type de la variable (p.e., nombre, chaîne, date).

Description de la variable \*

Description de la variable.

☐ Clé étrangère

La variable est une clé étrangère.

☐ Clé primaire

La variable est une clé primaire.

Cochez si cette variable représente une clé étrangère ou primaire

Ajouter une nouvelle variable à la table ou au fichier

ANNULER ENREGISTRER






#### 4.1.6. Validation des métadonnées

Pour **valider un fichier dont la saisie est terminée**, il faut cliquer sur l'icône verte située en haut à droite de l'éditeur. Cette action déclenche automatiquement une validation complète du formulaire et génère un rapport d'erreurs listant les champs manquants ou incorrects.

Le rapport distingue deux niveaux de gravité : les **warnings** (avertissements), qui signalent des éléments recommandés **mais non bloquants pour publication**, peuvent être ignorés ; en revanche, les **violations** (erreurs critiques) doivent impérativement être corrigées. Ces dernières empêchent le téléchargement du fichier au format RDF (TTL) et bloquent toute possibilité de publication ou de partage.

Il est donc recommandé de procéder à une validation régulière lors de la saisie afin de corriger les erreurs au fil de l'eau.

Valider les métadonnées

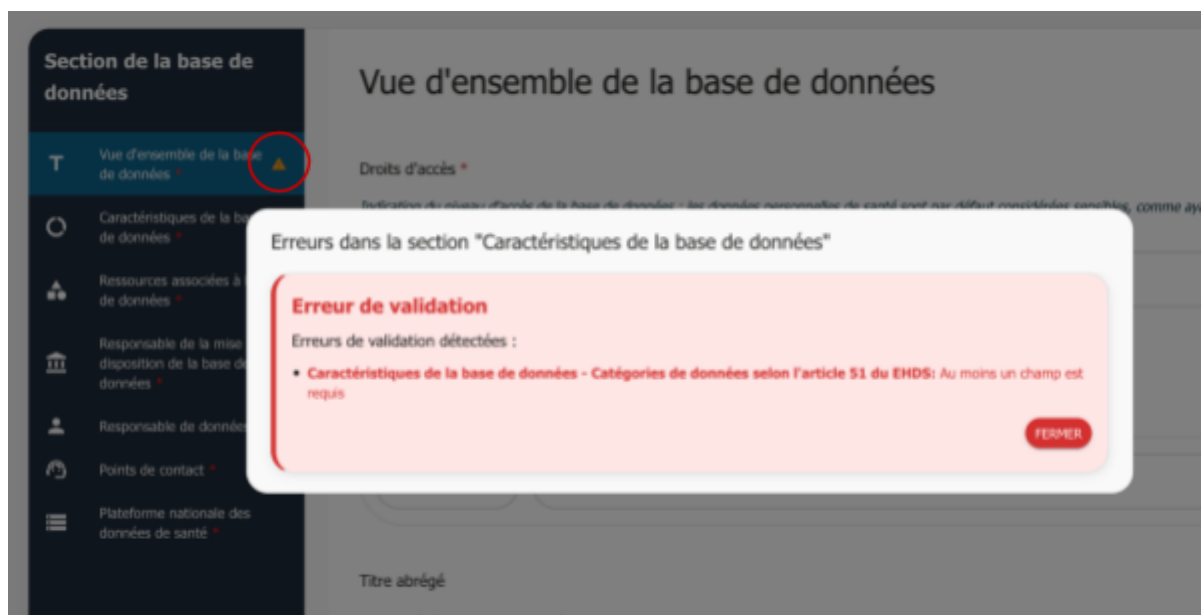
Résultat de la validation - hdcapUE

Statut de la validation ! Invalide

Gravité	Propriété	Message
Violation	Nom du responsable du Responsable de la mise à disposition de la base de données ou service d'accès d'une modalité d'accès ou d'une table et d'un fichier	Property needs to have at least 1 value
Violation	Site internet du responsable du Responsable de la mise à disposition de la base de données ou service d'accès d'une modalité d'accès ou d'une table et d'un fichier	Property needs to have at least 1 value
Violation	Adresse email du responsable du Responsable de la mise à disposition de la base de données ou service d'accès d'une modalité d'accès ou d'une table et d'un fichier	Property needs to have at least 1 value
Violation	Point de contact	Property needs to have at least 1 value
Violation	Couverture géographique	Property needs to have at least 1 value
Violation	Domaine médical	Property needs to have at least 1 value
Violation	Catégories de données selon l'article 51 du EHDS	Property needs to have at least 1 value
Violation	Âge minimal en année	Property needs to have at least 1 value
Warning	Responsable de la mise à disposition de la base de données	Value does not have node kind IRI

Afin de faciliter la saisie et de guider l'utilisateur dans la complétion du formulaire, une icône orange est disponible à côté de certains champs. En cliquant dessus, **vous pouvez afficher les erreurs de saisie spécifiques à ce groupe**, qu'il s'agisse d'un format incorrect, d'une valeur manquante ou d'un non-respect du vocabulaire contrôlé.

Cette option permet de **repérer rapidement les points à corriger**, sans avoir à attendre la validation finale du fichier. Il est particulièrement utile pour s'assurer, au fil de l'eau, que les données saisies sont conformes aux exigences du modèle.



#### 4.1.7. Téléchargement du fichier

Une fois la saisie terminée (ou en cours), deux formats de téléchargement sont proposés : le **format XLSX** et le **format Turtle (RDF)**. Le fichier XLSX peut être téléchargé à tout moment, même si le formulaire n'est pas entièrement complété. Il permet de reprendre la saisie plus tard ou de partager un brouillon.

Quelques règles de saisie de l'Excel:

	A	B	C	D	E	F	G
1	Droits d'accès						
2	Indication du niveau d'accès de la base de données : les données personnelles de santé sont par défaut considérées sensibles, comme ayant un accès restreint.	Nom complet de la base de données.	Nom abrégé de la base de données.	Description de la base de données (préliminaire) en texte libre. Précisez le type de base, la thématique, le principal critère d'inclusion et si elle est chahutée.	Domaine médical de la base de données. Un ensemble de données peut être associé à plusieurs thèmes. Les URIs de concepts Wikidata DOIVENT être indiquées.	Liste de mots-clés associés au catalogue de métadonnées et simplifiant son identification, à indiquer en texte libre.	Description en texte libre de la population de données contenues dans la base de données.
3	obligatoire	obligatoire	obligatoire	obligatoire	obligatoire	obligatoire	obligatoire
4	multilingue string	multilingue string	multilingue string	multilingue string	multilingue string	multilingue string	multilingue string
5	Vue d'ensemble de la base de données - Droits d'accès	Vue d'ensemble de la base de données - Titre de la base de données	Vue d'ensemble de la base de données - Titre abrégé	Vue d'ensemble de la base de données - Description de la base de données	Vue d'ensemble de la base de données - Domaine médical	Vue d'ensemble de la base de données - Mots-clés	Vue d'ensemble de la base de données - Description de la base de données
6	Sensible	Base de données sur la pathologie ABC (test)@fr	RDD ABC@fr	La base de données du programme ABC, gérée par un consortium hospitalier, est un entrepôt de données structurées issues des	Dermatologie	dermato@fr keyword@fr peau@fr	L'objectif du programme ABC est de constituer une base de données standardisée sur les patients atteints de
7							
8							
9							

Lorsqu'un champ peut contenir plusieurs valeurs (de type *1...n*), deux cas se présentent :

- **Si le champ correspond à un vocabulaire contrôlé** : ajoutez la nouvelle valeur sur la ligne suivante de l'Excel.
- **Si le champ correspond à un texte libre** : saisissez la nouvelle valeur dans la même cellule, en allant à la ligne (ex. mot-clés).

Pour **indiquer la langue** d'une valeur (afin de permettre au dictionnaire multilingue de la plateforme de la reconnaître), ajoutez le suffixe (ex. **@fr**, **@en**) — ou ne mettez rien si la valeur est en français par défaut.

En revanche, le téléchargement au format Turtle (fichier RDF) est soumis à une validation stricte : il n'est accessible que si l'ensemble du fichier respecte la *shape* européenne (HDCAT-AP). Ce mécanisme permet de garantir la qualité et la conformité des métadonnées, tout en évitant les erreurs de saisie ou de structure.

#### 4.1.8. Importation d'un fichier

Vous avez **la possibilité d'importer un fichier déjà généré afin de le modifier ou de le compléter**. Cela permet par exemple de commencer une saisie, de la sauvegarder en téléchargeant le fichier au format XLSX (ou TTL si finalisé), puis de la reprendre ultérieurement en le rechargeant dans l'éditeur.

Cette fonctionnalité est particulièrement utile pour travailler **en plusieurs temps ou collaborativement**, tout en conservant la structure et les données déjà saisies. Une fois importé, le fichier peut être modifié comme n'importe quel formulaire, avec accès à l'ensemble des validations et des options de téléchargement.



## 4.2. Gestion des métadonnées en tant qu'administrateur

Pour gérer directement les métadonnées du catalogue, il faut accéder à l'interface d'administration via l'URL suivante : **[https://admin.<nom\\_de\\_votre\\_domaine>](https://admin.<nom_de_votre_domaine>)**. Cet accès est réservé aux comptes disposant des droits administrateur. Si vous ne disposez pas encore de ces droits, rapprochez-vous de votre administrateur utilisateur.

L'interface admin offre plusieurs avantages : elle permet de **modifier les métadonnées directement depuis le formulaire**, sans avoir à générer ou manipuler de fichiers intermédiaires. Surtout, elle permet de **publier les métadonnées d'une base en un clic, une fois la saisie validée**, ce qui simplifie considérablement le processus par rapport à l'éditeur classique.

### 4.2.1 Créer les métadonnées d'une base de données

Pour créer une nouvelle base dans le catalogue, trois méthodes sont possibles selon vos besoins et votre niveau d'avancement :

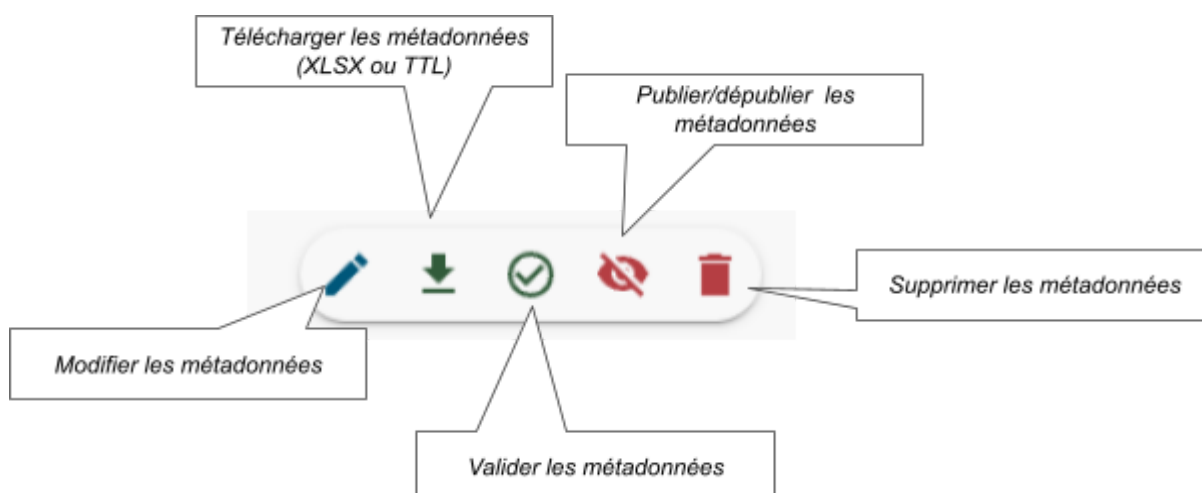
- **Créer manuellement** les métadonnées en remplissant le formulaire champ par champ.
- **Importer un fichier au format Turtle (RDF)**, si vous disposez déjà d'une version conforme au standard.
- **Importer un fichier XLSX**, généré depuis l'éditeur ou structuré selon le modèle attendu.



Une fois le fichier chargé, la base apparaîtra automatiquement dans la liste des jeux de données. Vous pourrez alors la modifier, la valider ou la publier depuis l'interface d'administration.



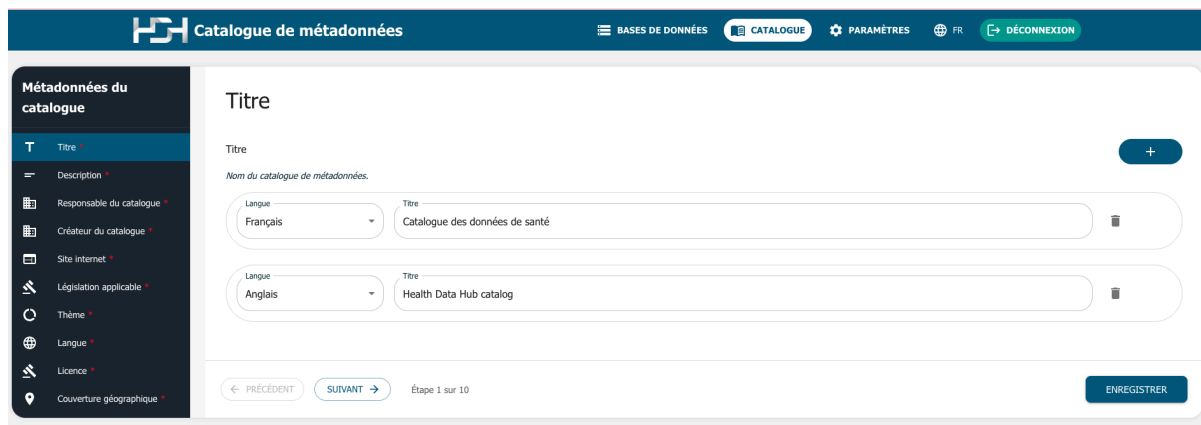
Une fois un jeu de métadonnées créé, **plusieurs actions sont possibles** depuis l'interface d'administration :



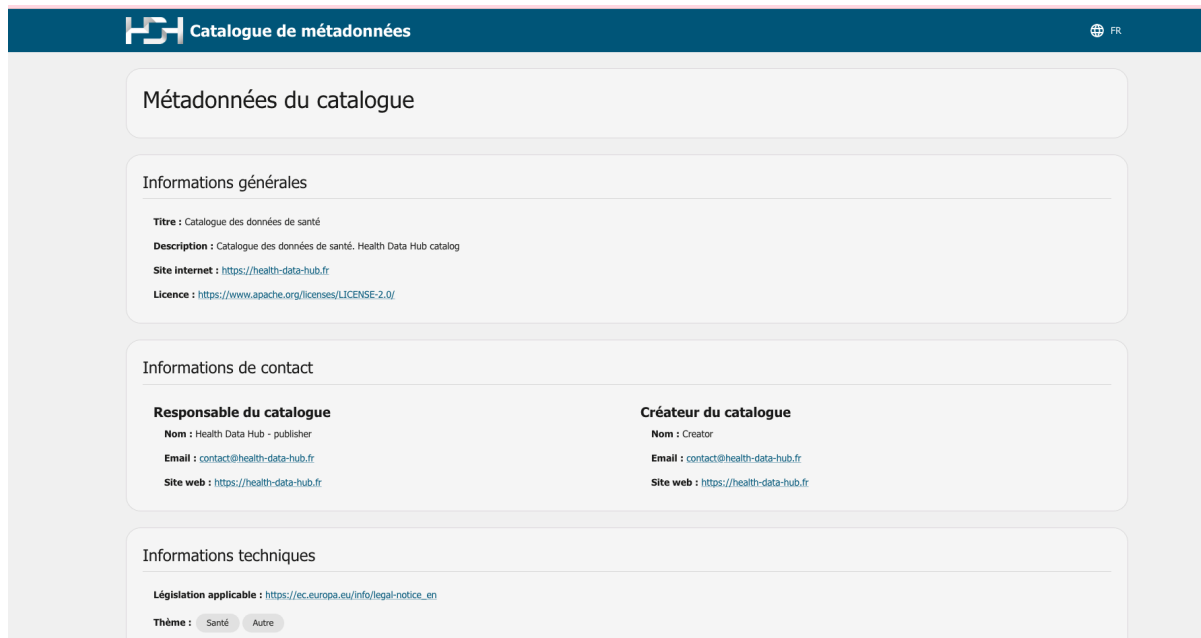
- **Modifier** : vous pouvez à tout moment éditer le contenu du fichier. Attention toutefois, **toute modification entraîne automatiquement la dépublication du jeu de données**, afin de garantir la cohérence des informations publiées.
- **Télécharger** : vous avez la possibilité de télécharger les métadonnées, aux formats XLSX ou TTL, comme depuis l'éditeur en ligne. Cela permet de conserver un historique de vos versions ou de s'en servir comme modèle pour une nouvelle base.
- **Valider** : cette étape permet de vérifier que le formulaire respecte bien les exigences du standard européen (HDCAT-AP). La validation est nécessaire pour pouvoir publier ou synchroniser le jeu avec le catalogue européen.
- **Publier** : en cliquant sur "Publier", le jeu de données devient immédiatement visible sur l'interface utilisateur. Il est alors accessible aux utilisateurs finaux, selon les modalités d'accès définies.
- **Supprimer** : cette action permet d'effacer définitivement le jeu de données de votre catalogue. Elle est irréversible.

## 4.2.2. Métadonnées du catalogue

Vous pouvez également **gérer les métadonnées du catalogue initial**, créées automatiquement lors du déploiement de la plateforme (cf. [partie 3](#)), sur **l'onglet "catalogue" du menu**. Ces métadonnées sont accessibles en bas de la page d'administration. Elles peuvent être modifiées, validées ou publiées comme n'importe quel autre jeu de données, afin d'assurer la cohérence et la qualité globale du catalogue.



### Affichage sur le catalogue (ex. Health Data Hub)



## 5. Déploiement de l'application

### 5.1. Configurer les Helm Charts

Le déploiement de l'application repose sur des **Helm charts** adaptés à l'API. Les étapes ci-dessous résument le processus de mise en place en production. Elles supposent une connaissance de Kubernetes et Helm.

- Consulter le **README** du code source pour les instructions détaillées.
- Compléter les modèles décrits dans *helm-charts/README.kubernetes.prod.md*.
- Vérifier attentivement les paramètres dans *values.yaml* (services, DB, ressources, secrets...).
- Documentation de référence : *helm-charts/README.kubernetes.prod.md*.

### 5.2. Personnaliser l'application

Avant le déploiement, la personnalisation de l'application s'effectue via le dossier **customization** (inclus dans le code source du repository *metadata-catalog-api*). Cette étape permet d'adapter la mise en forme du catalogue en amont. Par la suite, un rôle d'administrateur fonctionnel pourra toujours réaliser des personnalisations directement depuis l'interface (cf. [partie 5](#)).

#### 5.2.1. Personnalisation des éléments de forme

##### Couleurs du catalogue

- Se rendre sur le fichier *settings.json* du code source
- Modifier les variables `primary_color` et `secondary_color`





##### Police et type

- Se rendre sur le fichier *settings.json*
- Modifier la variable `font_family`

##### Logo

- Remplacer *logo.png* par votre logo
- NB : le fichier doit s'appeler également *logo.png*

#### 5.2.2 Personnalisation du pied de page

			
À propos de nous	Contactez-nous	Mentions légales	Politique de confidentialité

##### URL "À propos de nous"

- Se rendre sur le fichier *settings.json*

- Modifier la variable `about_us_url`

#### Email de contact

- Se rendre sur le fichier *settings.json*
- Modifier la variable `contact_email`

#### Mentions légales

- Se rendre dans le dossier *legal\_notice*
- Ouvrir le fichier markdown *fr.md* pour modifier les mentions légales en Français (*en.md* pour la version en anglais)

#### Politiques de confidentialité

- Se rendre dans le dossier *privacy\_policies*
- Ouvrir le fichier markdown "*fr.md*" pour modifier les mentions légales en Français (*en.md* pour la version en anglais)

## 5.3. Métadonnées du catalogue

Les métadonnées du « catalogue » peuvent être renseignées via une configuration JSON dans le fichier *settings.json* (cf. [partie 6](#)). Elles décrivent l'ensemble des bases de données ainsi que l'organisation qui en assure la gestion (voir annexe 1 pour un aperçu du fichier *settings.json*). **Veuillez noter que ces métadonnées peuvent être complétées post-déploiement par administrateur fonctionnel ou un administrateur de bases.**

## 5.4. Personnaliser Keycloak

Si un fournisseur d'authentification externe ou un fournisseur OIDC/SAML est utilisé, suivez les instructions de la documentation officielle de Keycloak disponible sur [keycloak.org](https://www.keycloak.org/docs/latest/server_admin/index.html#identity_broker).

[https://www.keycloak.org/docs/latest/server\\_admin/index.html#identity\\_broker](https://www.keycloak.org/docs/latest/server_admin/index.html#identity_broker)

Si une base de données utilisateur externe est utilisée, suivez les instructions de la section *User Storage Federation* de la même documentation.

[https://www.keycloak.org/docs/latest/server\\_admin/index.html#user-storage-federation](https://www.keycloak.org/docs/latest/server_admin/index.html#user-storage-federation)

## 6. Gestion des utilisateurs

**Trois types de rôle administrateurs sont définis dans l'application :**

- **L'administrateur utilisateurs** (ou administrateur Keycloak) a pour mission exclusive la gestion des comptes (ajout, suppression, modification, etc.). Il n'a pas accès aux fonctions de personnalisation ni à la gestion des métadonnées.
- **L'administrateur fonctionnel** est responsable de la personnalisation du site ainsi que de la gestion des métadonnées. En revanche, il n'intervient pas sur la gestion des utilisateurs.
- **L'administrateur de bases** se concentre sur la gestion des métadonnées. Il ne participe ni à la personnalisation du site ni à la gestion des utilisateurs.



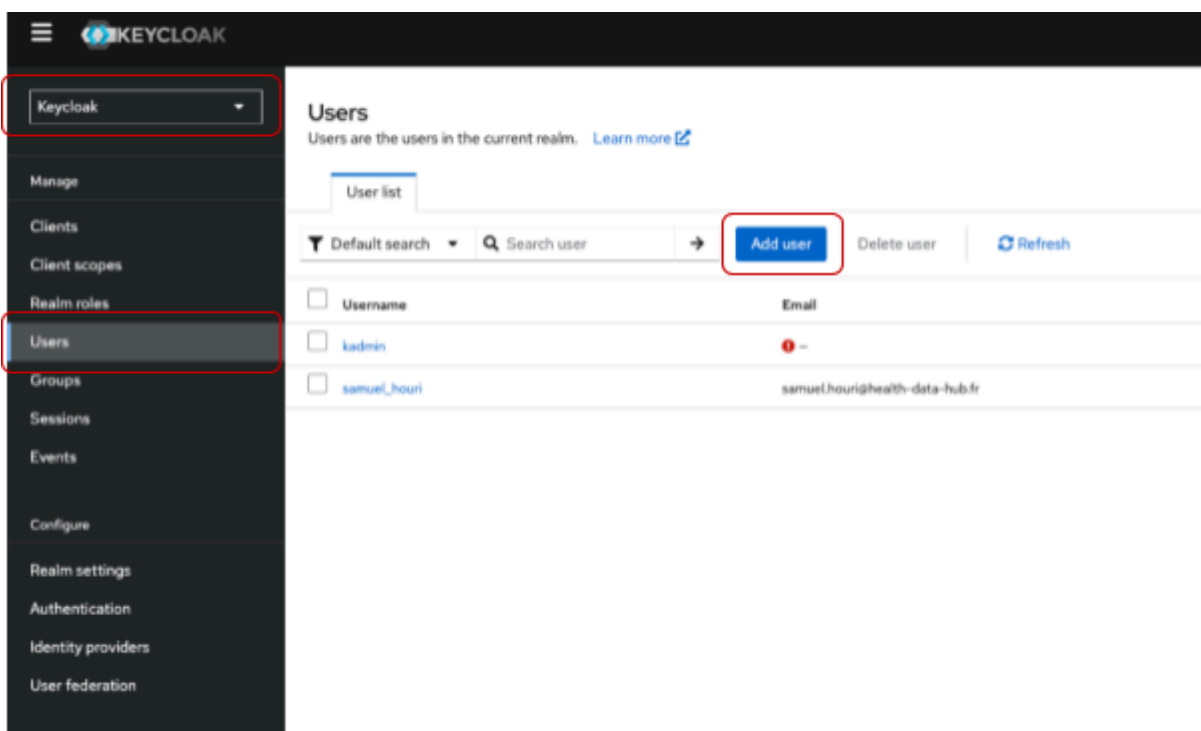
Rôle	Personnalisation du site	Gestion des métadonnées	Gestion des utilisateurs
Administrateur utilisateurs (keycloak)	!	!	✓
Administrateur fonctionnel	✓	✓	!
Administrateur de bases de données	!	✓	!

## 6.1. Création d'un administrateur utilisateurs (admin keycloak)

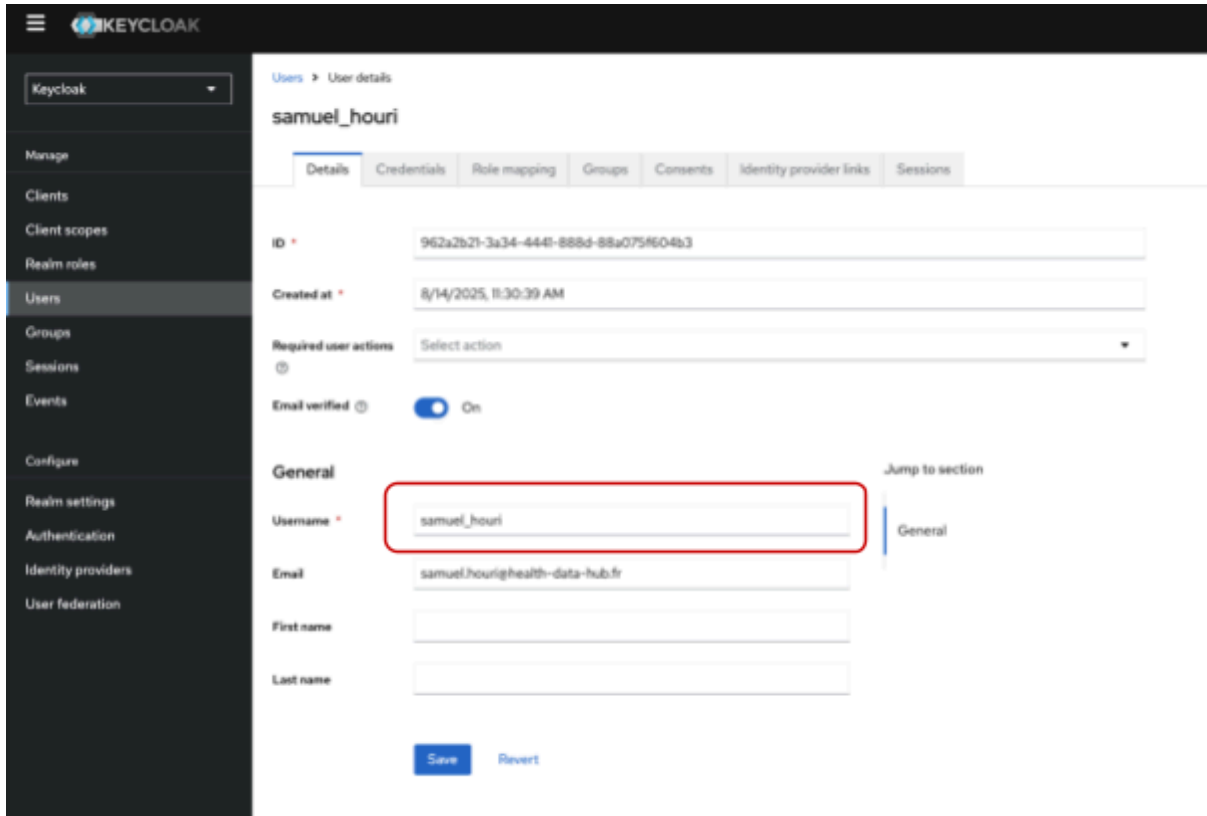
Le premier compte d'administrateur utilisateurs est créé lors du déploiement via les helm chart, en configurant les variables `KEYCLOAK_ADMIN` et `KEYCLOAK_ADMIN_PASSWORD`. Afin de gérer l'envoi et le partage des identifiants, veuillez vous référer à la documentation Keycloak concernant l'automatisation de l'envoi via un prompt ([guide d'administration du serveur](#)).

Néanmoins pour créer un nouvel administrateur fonctionnel il faut :

- **Se rendre sur `https://keycloak.<url_du_catalogue>` > Keycloak (realm) > users > Add users**

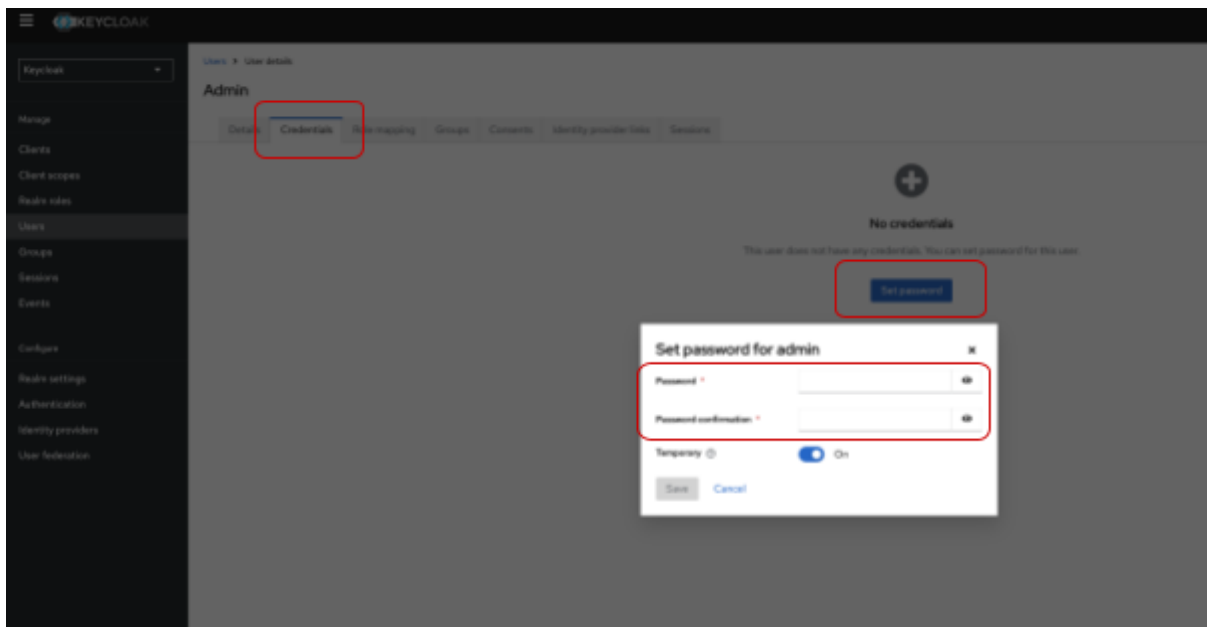


- **Créer l'utilisateur :** remplir le formulaire avec au minimum un nom d'utilisateur



The screenshot shows the Keycloak administration interface. On the left is a sidebar with navigation options: Manage, Clients, Client scopes, Realm roles, Users (selected), Groups, Sessions, Events, Configure, Realm settings, Authentication, Identity providers, and User federation. The main content area is titled 'Users > User details' and shows the profile for 'samuel\_houri'. The 'Details' tab is active, displaying fields for ID (962a2b21-3a34-444f-888d-88a075f604b3), Created at (8/14/2025, 11:30:39 AM), Required user actions (Select action), and Email verified (On). Below this is the 'General' section with fields for Username (samuel\_houri, highlighted with a red box), Email (samuel.houri@health-data-hub.fr), First name, and Last name. At the bottom are 'Save' and 'Revert' buttons. A 'Jump to section' sidebar on the right lists 'General'.

- **Mot de passe :** créer le mot de passe sur "Credentials"

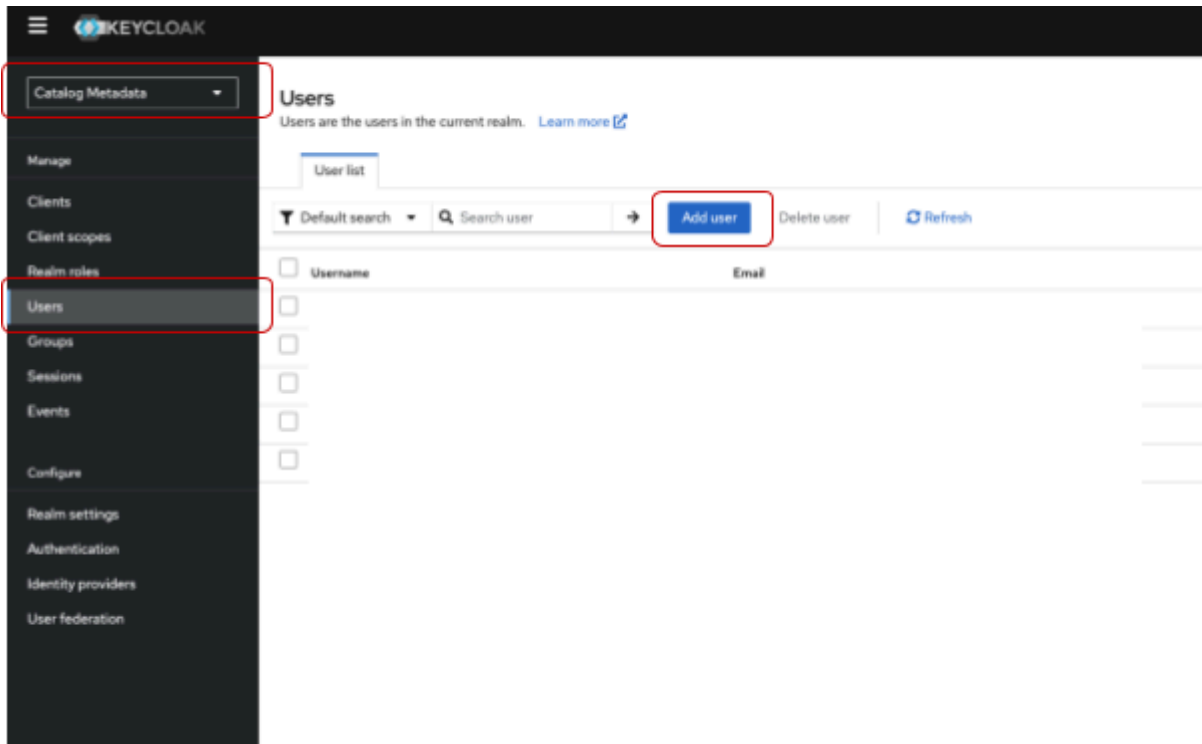


The screenshot shows the Keycloak administration interface for the 'Admin' user. The 'Credentials' tab is selected and highlighted with a red box. The main content area shows a message: 'No credentials. This user does not have any credentials. You can set password for this user.' Below this message is a 'Set password' button, also highlighted with a red box. A modal dialog titled 'Set password for admin' is open, showing fields for 'Password' and 'Password confirmation', both highlighted with red boxes. The 'Temporary' toggle is turned 'On'. At the bottom of the modal are 'Save' and 'Cancel' buttons.

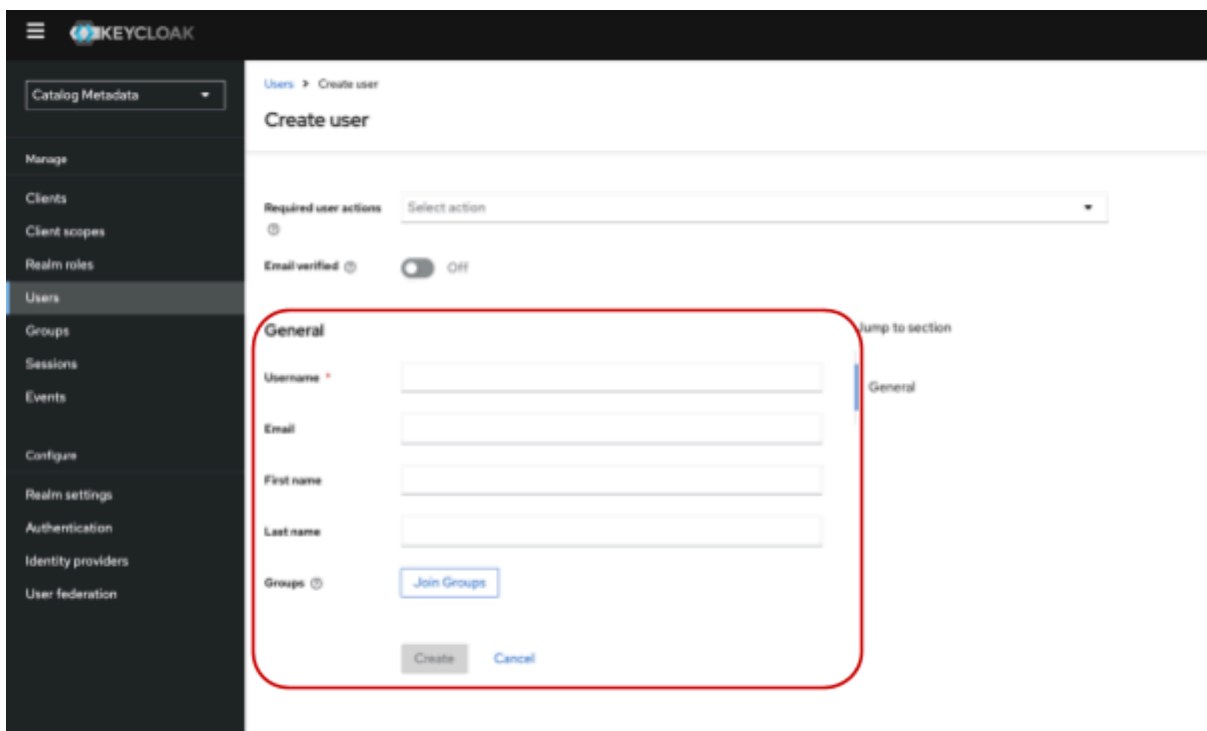
## 6.2. Création d'un administrateur fonctionnel

Pour créer un administrateur fonctionnel il faut :

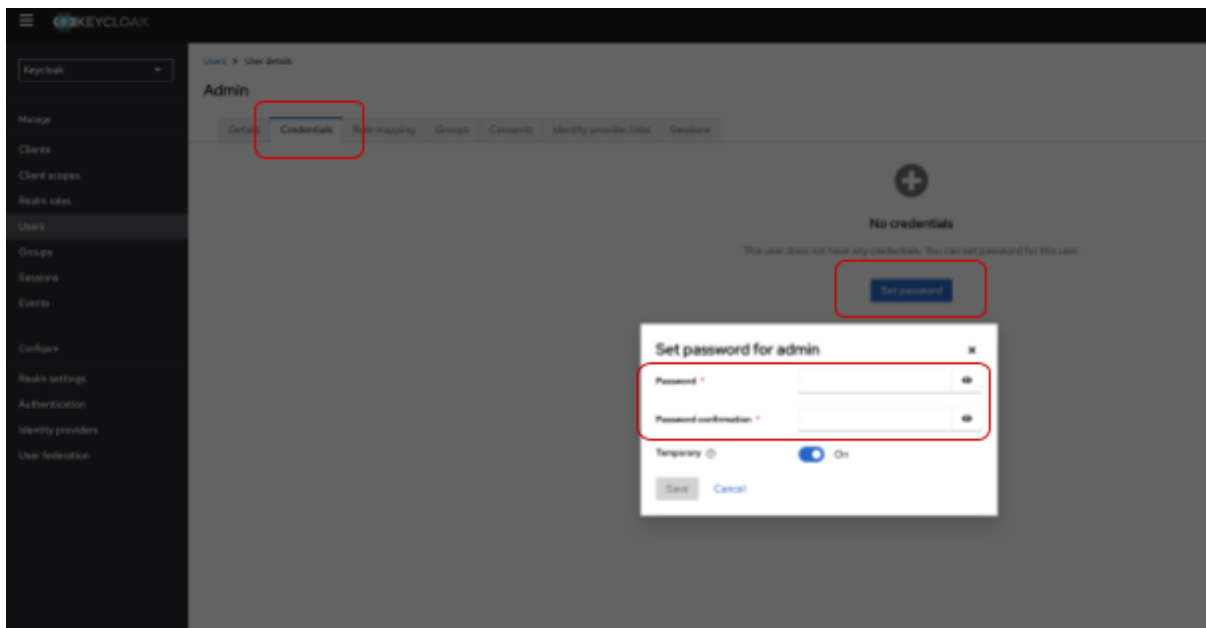
- **Se rendre sur** *[https://keycloak.<url\\_du\\_catalogue>](https://keycloak.<url_du_catalogue>)* > *Catalog Metadata (realm)* > *users* > *Add users*



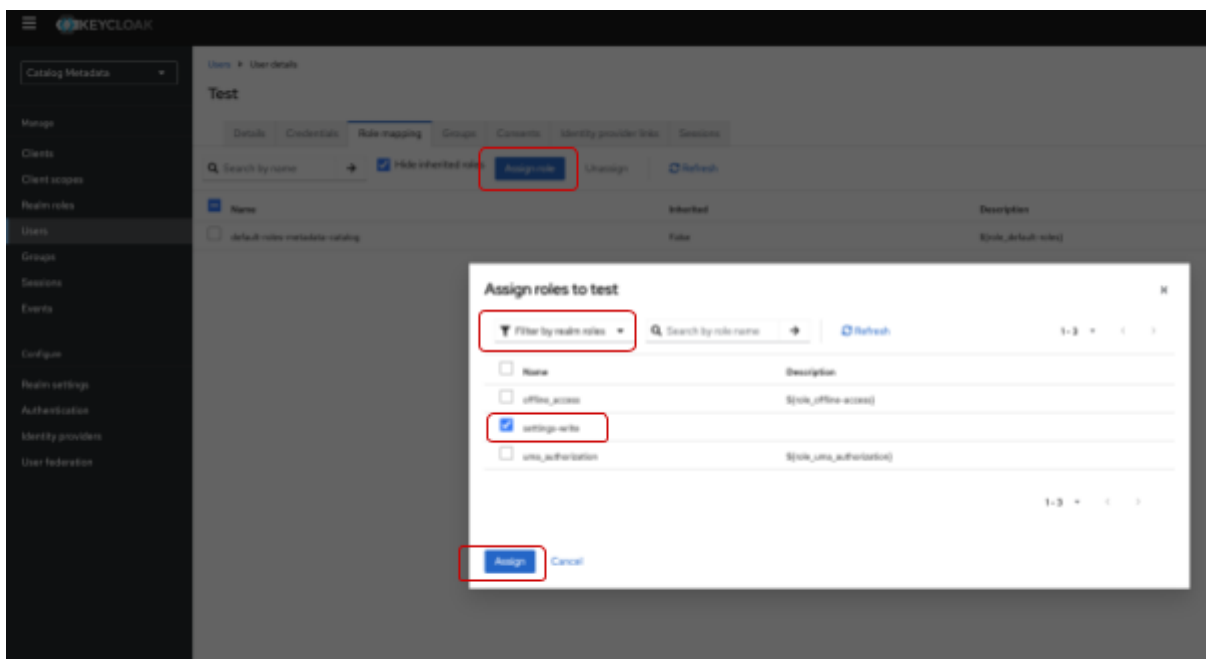
- **Remplir le formulaire** : compléter les informations de l'utilisateur avec au minimum le nom d'utilisateur



- **Mot de passe :** créer un mot de passe sur "Credentials"



- **Donner les droits du rôle :** aller sur "Role mapping" puis sur "Filter by realm roles" et sélectionner *settings-write*



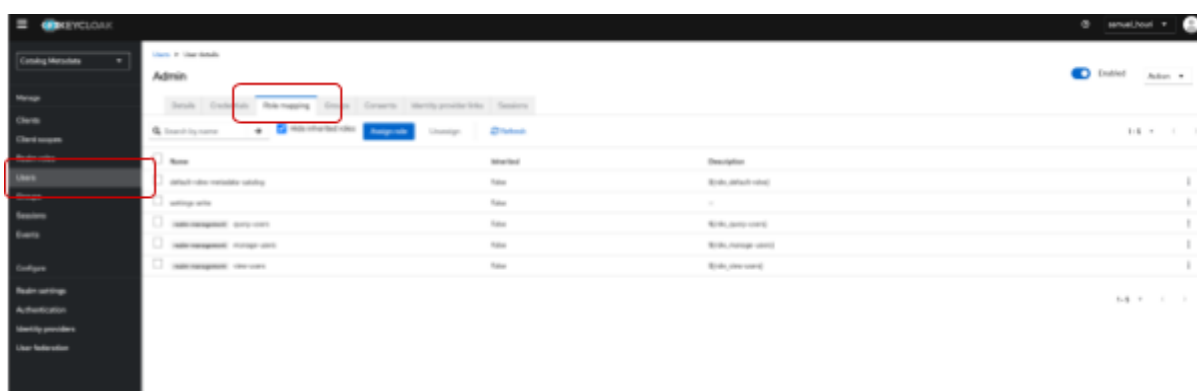
### 6.3. Création d'un administrateur de bases de données

Pour **créer un administrateur de bases de données**, la procédure est très simple. Il vous suffit de commencer par créer un nouvel utilisateur (comme expliqué dans la partie précédente). Lors de cette création, vous n'avez pas besoin d'assigner de rôle spécifique à cet utilisateur. En effet, par défaut, tout nouvel utilisateur dispose uniquement de droits sur les métadonnées des bases de données, sans autres privilèges additionnels. C'est ce

comportement par défaut qui lui permet d'agir comme administrateur des bases de données, puisqu'il peut gérer et superviser leur structure sans avoir accès aux données applicatives elles-mêmes.

## 6.4. Modifier un rôle

Pour modifier le rôle d'un utilisateur, vous devez d'abord accéder à la liste complète des utilisateurs depuis l'interface dans "Users". Une fois la liste affichée, identifiez et sélectionnez l'utilisateur concerné en cliquant directement sur son nom. Cela ouvrira sa fiche de gestion. Dans le menu situé en haut de la page, vous trouverez un onglet intitulé « **Role mapping** ». C'est à partir de cette section que vous pourrez consulter les rôles actuellement attribués et effectuer les changements nécessaires en ajoutant ou en retirant des rôles selon vos besoins. Pour connaître les rôles (realm) associées, lire les parties précédentes de la section.



## 6.5. Modifier un mot de passe

Pour modifier le mot de passe d'un utilisateur, commencez par accéder à la liste des utilisateurs en vous rendant dans la section « **Users** » de l'interface. Une fois la liste affichée, sélectionnez l'utilisateur concerné afin d'ouvrir sa fiche de gestion. Dans le menu associé à cet utilisateur, allez dans l'onglet « **Credentials** ». C'est à partir de cette section que vous trouverez l'option « **Reset Password** », qui permet de réinitialiser le mot de passe existant et d'en définir un nouveau.



## 7. Personnalisation du catalogue

La personnalisation de l'application s'effectue à partir de l'onglet **Paramètres** dans la barre de navigation. Pour y accéder, il faut se rendre sur le lien **[https://admin.<url\\_du\\_catalogue>](https://admin.<url_du_catalogue>)**.

Cette fonctionnalité permet à n'importe quelle organisation de personnaliser divers aspects de l'application, tels que le logo, les couleurs, la police de caractères, ainsi que les informations utiles situées en bas de la page de l'interface principale.

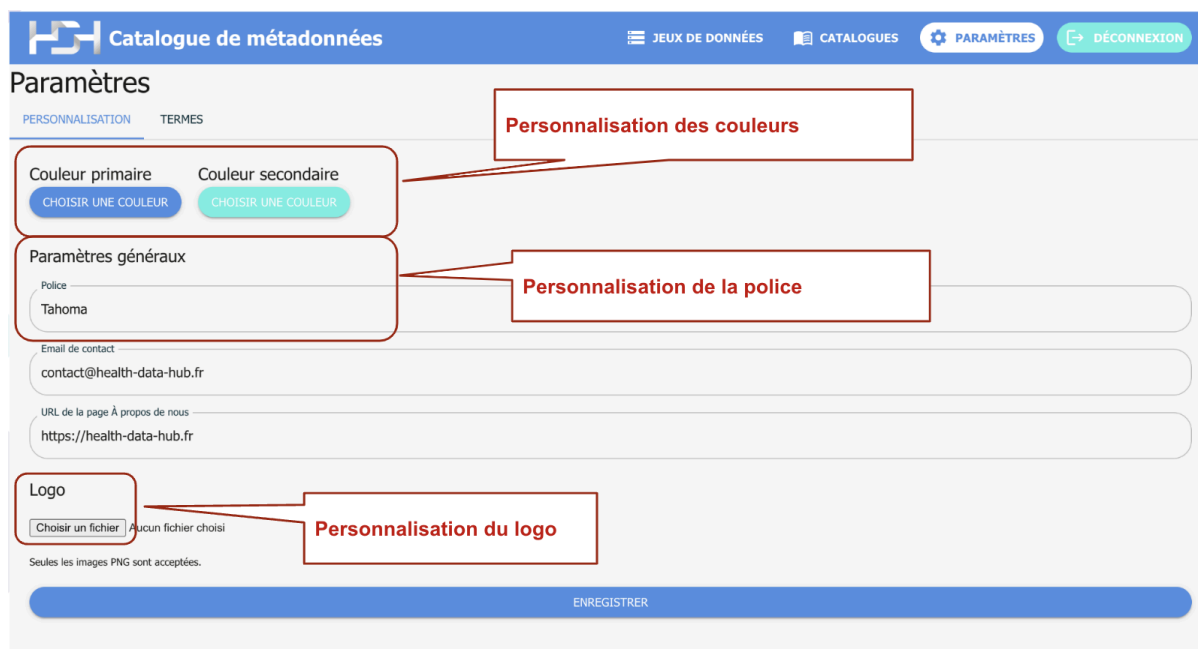
Dans le pied de page de la page principale des métadonnées, accessible par l'ensemble des utilisateurs, il est également essentiel de configurer des éléments supplémentaires comme les liens de contact, les mentions légales, et d'autres informations importantes pour l'organisation.

### 7.1. Personnalisation du thème (couleur, logo, police...)

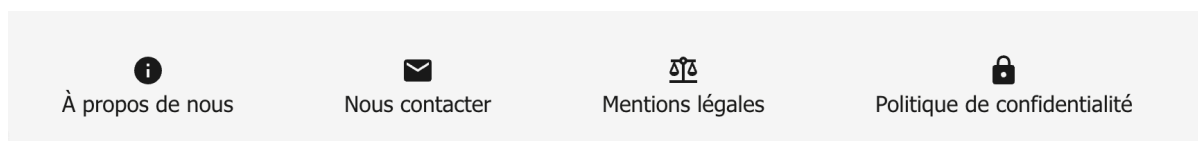
**Personnalisation des couleurs :** vous pouvez spécifier une couleur primaire et secondaire pour l'application qui correspondent aux couleurs de votre organisation, toutes les couleurs sont disponibles en Hexa ou en RGB.

**Personnalisation de la police :** vous pouvez choisir parmi les thèmes la police qui correspond à celle de votre organisme.

**Personnalisation du logo :** vous pouvez mettre à jour le logo avec celui de votre organisation. Les formats autorisés sont le *jpeg/jpg/png*.




## 7.2. Personnalisation des éléments du pied de page



**A propos de nous** : lien qui vous redirige vers le site internet de votre organisme.

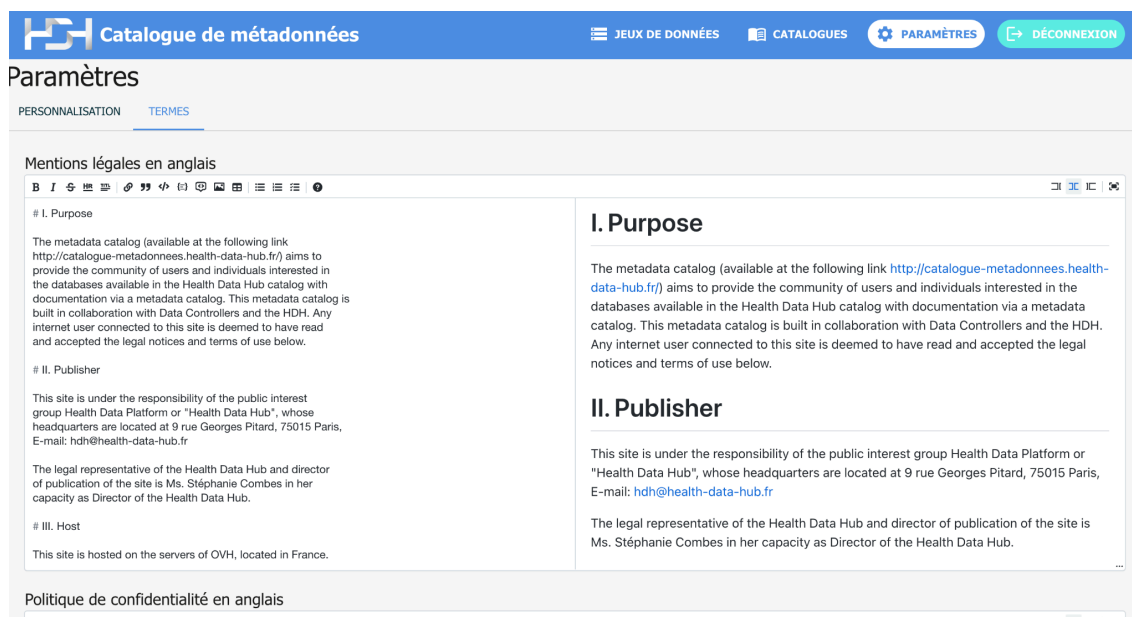
**Nous contacter** : adresse e-mail qui vous redirige vers une mise en contact.



**Mentions légales** : modification du contenu des mentions légales alignées à votre organisme. Vous pouvez vous rapprocher du service juridique de votre organisme pour les définir.

**Politique de confidentialité** : modification du contenu de la politique de confidentialité alignées à votre organisme. Vous pouvez vous rapprocher du service juridique de votre organisme pour les définir.

Les deux contenus précédents peuvent être écrits en français et en anglais.



## 8. Interface utilisateur

Une fois publiées, les **métadonnées** apparaissent dans l'**interface principale**, accessible depuis l'URL de base du catalogue.

L'interface principale comprend :

- **Une barre de recherche** : permet de rechercher un mot-clé ou le titre d'une base de données.
- **Une zone de filtres** : pour affiner la recherche selon différents critères (mots-clés, langue, droits d'accès, etc.).
- **La liste des datasets** : affichant les bases de données correspondant aux critères de recherche.

En haut à droite, il est également possible de **changer la langue**. Cette action modifie :

- les éléments du pied de page, si une version anglaise a été saisie lors de la personnalisation ;
- les champs des métadonnées, si vous avez complété le formulaire avec des traductions en anglais.





Ensuite, il suffit de **cliquer sur l'un des blocs** pour accéder aux métadonnées détaillées.

Les blocs présentés ont été **sélectionnés de manière pertinente**, en fonction des besoins les plus fréquents d'un utilisateur ou d'un chercheur.

Toutes les informations relatives au dataset sont **centralisées sur une page unique**, structurée en différents blocs descriptifs.

## 8.1. Bases de données




**Métadonnées de la section "base de données"**

**Contacts associés à la "base de données"**

**"Modalités d'accès" à la base**

## 8.2. Tables, fichiers et variables

Ce bloc regroupe les **métadonnées des échantillons** ainsi que les **variables associées**. L'utilisateur peut accéder au **détail et à leur description** en cliquant sur le bouton « Voir ». En cliquant sur « Accéder », l'utilisateur peut **télécharger l'échantillon anonymisé**, lorsque celui-ci a été mis à disposition par l'administrateur.

Tables, fichiers & variables				
Titre de la table ou du fichier	Description de la table ou du fichier	Voir les variables		Accès aux échantillons anonymisés
Table 2 - Consultations	Données sur les consultations médicales liées à la pathologie ABC:			
Titre métier de la variable	Description de la variable	Type de la variable	Clé primaire	Clé étrangère
Identifiant consultation	Identifiant unique de la consultation	string		
Identifiant patient	Référence vers le patient concerné	string		
Motif de consultation	Motif principal de la consultation	string		
Date de consultation	Date de la consultation	date		
Table 1 - Patients	Informations démographiques et cliniques des patients atteints de la pathologie ABC:			

En cliquant sur « **Accéder** », l'utilisateur peut **télécharger l'échantillon anonymisé**, lorsque celui-ci a été mis à disposition par l'administrateur.

### 8.3. Ressources

Enfin, la dernière partie concerne l'ensemble des **ressources associées à la base de données**. Elle regroupe tous les éléments liés au dataset : identifiants, publications, schémas, classifications ou tout autre document de référence.

Ressources associées à la base de données	
<b>Identifiant(s) relatif(s) à la base de données :</b> <ul style="list-style-type: none"> <li><a href="https://doi.org/10.1234/example.doi.2025%20%20%E2%80%8B">https://doi.org/10.1234/example.doi.2025%20%20%E2%80%8B</a></li> </ul>	
<b>Publications :</b> <ul style="list-style-type: none"> <li><a href="https://publication.hdh.fr/abc">https://publication.hdh.fr/abc</a></li> </ul>	
<b>Ressources en relation avec la base de données :</b>	
Ressources en relation avec la base de données	Rôle de la relation
<a href="https://www.health-data-hub.fr/">https://www.health-data-hub.fr/</a>	<a href="#">À propos</a>
<b>Encadrement réglementaire de la base de données :</b> <ul style="list-style-type: none"> <li>Autorisation CNIL n°2025-ABC pour traitement de données de santé</li> </ul>	
<b>Labels de qualité :</b>	
Cible	Certificat

## 9. Annexes

### 9.1. Aperçu du fichier settings.json

Couleurs

Police

Email de contact

A propos de nous

Métadonnées du  
catalogue

```
{
  "primary_color": "#036287",
  "secondary_color": "#5DEEE3",
  "font_family": "tahoma",
  "contact_email": "contact@health-data-hub.fr",
  "about_us_url": "https://health-data-hub.fr",
  "catalog": {
    "title": [
      {
        "content": "Catalogue des données de santé",
        "language": "fr"
      },
      {
        "content": "Health Data Hub catalog",
        "language": "en"
      }
    ],
    "description": [
      {
        "content": "Catalogue des données de santé",
        "language": "fr"
      },
      {
        "content": "Health Data Hub catalogue",
        "language": "en"
      }
    ]
  },
  "publisher": {
    "name": "Health Data Hub - publisher",
    "website": "https://health-data-hub.fr",
    "email": "contact@health-data-hub.fr"
  },
  "creator": {
    "name": "Health Data Hub - creator",
    "website": "https://health-data-hub.fr",
    "email": "contact@health-data-hub.fr"
  },
  "regulation": "https://ec.europa.eu/info/legal-notice_en",
  "themes": [
    "http://publications.europa.eu/resource/authority/data-theme/HEAL"
  ],
  "languages": [
    "http://publications.europa.eu/resource/authority/language/FRA",
    "http://publications.europa.eu/resource/authority/language/ENG"
  ],
  "licence": "https://www.apache.org/licenses/LICENSE-2.0/",
  "countries": [
    "http://publications.europa.eu/resource/authority/country/FRA"
  ],
  "website": "https://health-data-hub.fr"
}
```

## 9.2 Diagramme des métadonnées

