

ACTION 18 DU PNSE4

GD4H - Green Data for Health

**Pour la mobilisation et valorisation, par les acteurs de la recherche et expertise,
des données environnementales au service de la santé environnement.**

18.11.2022

● Ordre du jour

_ Présentation de l'offre de service

- Périmètre et objectifs
- Catalogue de données
- Outils juridiques
- Actions d'animation

_ Améliorer la mobilisation des données environnementales en santé-environnement pour les territoires

- SONDAGE : recueil des points de difficulté et des besoins
- FORMULAIRE : proposition de défis/d'outils
- Echanges

Partie 1

Présentation de l'offre de service

- Périmètre et objectifs
- Catalogue de données
- Outils juridiques
- Actions d'animation



Présentation de l'offre de service

➤ **Périmètre et objectifs**

Quoi ? - L'OBJECTIF DU GD4H (action 18 PNSE4)

Permettre une meilleure mobilisation et valorisation, par les acteurs de la recherche et de l'expertise, des données environnementales au service de la santé environnement.



Faciliter la **repérabilité** et l'**accès** aux données environnementales y compris au niveau des territoires



Décrire les données environnementales pertinentes et disponibles, et outiller les acteurs, pour un **appariement rigoureux** avec des données de santé



Améliore la **communication** et la **synergie** entre les acteurs de la santé environnement (notamment sur les enjeux d'interopérabilité)

Pour qui ? – LES UTILISATEURS DU GD4H

Producteurs et gestionnaires
de données environnementales

Consommateurs et réutilisateurs
de données : chercheurs et
experts en santé environnement

● Offre de service du GD4H en cours d'incubation

Animation

Fédérer l'écosystème des organismes responsables de données, experts et chercheurs en SE **autour de problématiques communes telles que l'interopérabilité et de cas d'usage.**

- _ Communauté de la donnée en santé environnement
- _ 4 cas d'usage sur des problématiques ciblées
- _ Appel à projets conjoint avec le Health Data Hub
- _ Data Challenge

Catalogue de données

Améliorer la repérabilité et l'accessibilité des bases de données environnementales grâce à la **mise à disposition d'un catalogue en ligne bénéficiant d'un moteur de recherche.**

- _ Catalogue de 129 ressources (JDD, BDD, portail) environnementales utiles pour la santé environnement
- _ Catalogue d'outils pour pré-traitement et analyse de données - cadrage en cours

Outils juridiques

Améliorer l'accessibilité des bases de données environnementales par les chercheurs - **acculturer et accompagner les producteurs dans l'ouverture de leurs données.**

- _ Mise en ligne de fiches pédagogiques sur les fondamentaux du droit - ressource juridique
- _ Outil autonome d'analyse juridique (FAQ dynamique) - cadrage et rédaction du contenu en cours

● Equipe projet du GD4H* incubée au sein de l'Ecolab (MTECT)



Marie Ramon-Daré
Cheffe de projet
Design de service
UX/UI



Paul Grignon
Chef de projet adjoint
Data Science
Sciences comportementales



Charly Dottin
Designer de service
Entrepreneur d'Intérêt Général

—
*Equipe projet épaulée par :

- _ une prestation de développement fullstack pour la réalisation du catalogue de données.
- _ une prestation de juriste pour la rédaction et la conception de l'outil d'analyse juridique autonome.



Présentation de l'offre de service

➤ Catalogue de données

Animation

Catalogue de données

Outils
juridiques

● Le catalogue du GD4H aujourd'hui

CONTENU DU CATALOGUE

129 ressources référencées* (JDDs, BDDs, portails).

54 producteurs et référents publics de données environnementales.

27 champs de métadonnées dont **7 spécifiques au domaine** santé-environnement

—

*Le catalogue référence et pointe vers les ressources, mais il ne les stocke pas.



Nous avons besoin de vous pour :

- Poursuivre l'identification des ressources (JDDs, BDDs, portails) manquantes sur votre domaine d'expertise.
- Identifier le(s) type(s) de filtre(s) le(s) plus pertinent(s) à mettre en place pour optimiser les recherches au sein du catalogue.

● **FOCUS – les métadonnées du catalogue**

La modélisation des données consiste à définir les champs qui décrivent une ressource - JDD, BDD, portail - avec des métadonnées descriptives qui :
_ doivent être utiles et pertinentes (**qualification**) afin d’être facilement identifiées (recherche et filtre) et utilisées (**accessibilité**)
_ doivent suivre des référentiels pour s’assurer de l’uniformité et la conformité des valeurs et de leur croisement entre elles et avec d’autres portails (**interopérabilité**)

129 jeux/bases/portails de données environnementales, 54 producteurs/référents publics de données environnementales

MÉTADONNÉES GÉNÉRALES

Titre de la ressource	Sigle de la ressource	Organisme(s) producteur(s)	Type de ressource	Description de la ressource
-----------------------	-----------------------	----------------------------	-------------------	-----------------------------

MÉTADONNÉES JURIDIQUES ET TECHNIQUES

Url d'accès à la ressource	Url de téléchargement de la ressource	Point de contact	La ressource est-elle accessible ?	Tarification	Nom de la licence	Format de la ressource
----------------------------	---------------------------------------	------------------	------------------------------------	--------------	-------------------	------------------------

MÉTADONNÉES TEMPORELLES ET GÉOGRAPHIQUES

Couverture temporelle	Millesime	Granularité temporelle	Fréquence de publication	La collecte est-elle terminée ?	La ressource est elle géoréférencée ?	Couverture spatiale	Granularité spatiale
-----------------------	-----------	------------------------	--------------------------	---------------------------------	---------------------------------------	---------------------	----------------------

MÉTADONNÉES SANTÉ ENVIRONNEMENT

Milieu	Voie d'exposition	Recherche, utilisation ou impact en SE	Type d'agents présents dans l'environnement	Nom de l'agent présent dans l'environnement	Jeux de données associés	Réutilisation associée
--------	-------------------	--	---	---	--------------------------	------------------------

● FOCUS – les métadonnées du catalogue / métadonnées santé-environnement

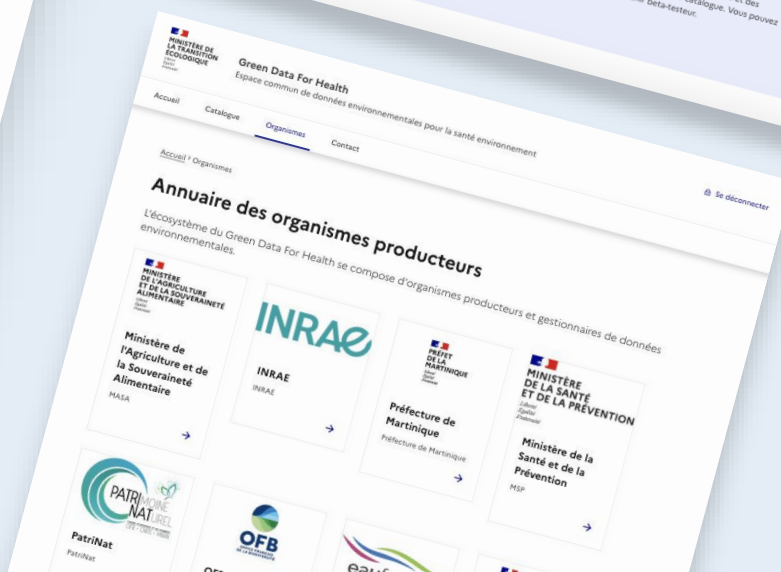
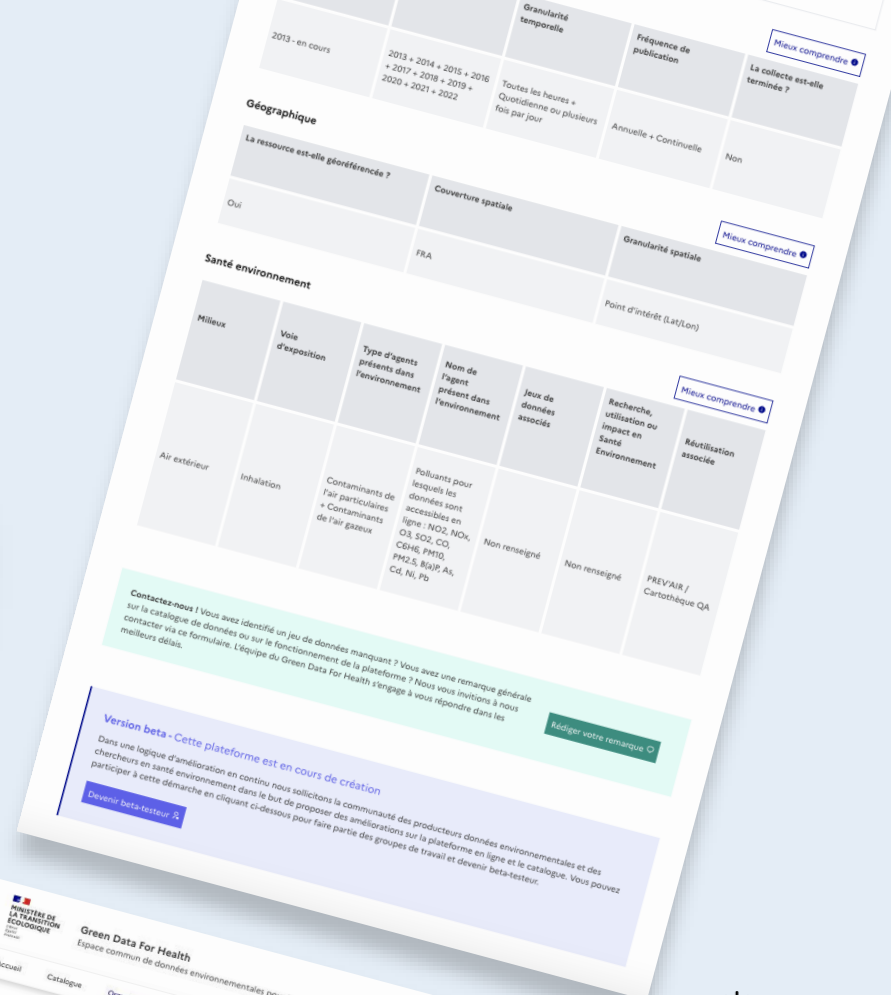
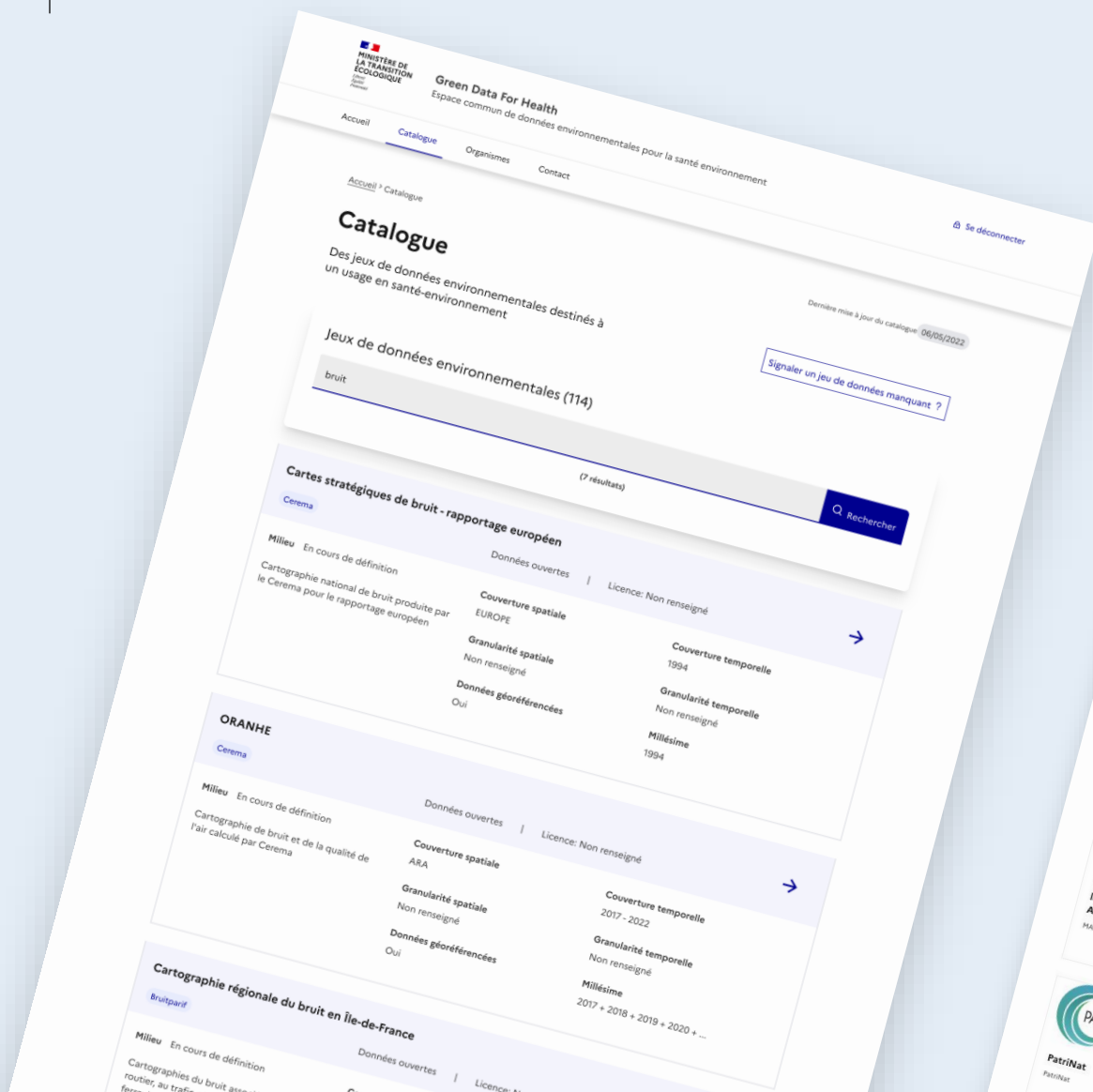
Milieu	Voie d'exposition	Recherche, utilisation ou impact en Santé Environnement	Type d'agents présents dans l'environnement	Nom de l'agent présent dans l'environnement
Ce champ permet de définir le milieu concerné par la ressource à partir d'un vocabulaire contrôlé interne.	Ce champ permet de définir la manière dont les personnes entrent en contact avec des substances dangereuses.	Ce champ permet de définir le type d'agent biologique, chimique ou physique présent dans l'environnement et susceptible d'affecter la santé des générations actuelles et futures. Il est caractérisé dans la ressource soit en fonction de sa famille, de ses usages ou de ses effets.	Ce champ permet de nommer l'agent biologique, chimique ou physique évoqué e mesuré dans la ressource à partir d'une entrée libre. Si pas applicable mettre NA.	Ce champ permet de définir quel impact, quelle utilisation, quel usage ou quel réemploi en Santé Environnement à partir de mots clés libres.
Choix multiple parmi : Air intérieur Air extérieur Poussières déposées Eaux Eaux brutes souterraines Eaux brutes superficielles Eau de mer Eaux usées Eau du robinet, eau minérale, eau thermale Aliment Sols Sols urbains Sols agricoles Sols forestiers Sols naturels Sols sites industriels Sédiments Occupation du sol, usage du sol Faune, flore, réservoir biologique Autre NA	Choisir parmi : Inhalation Ingestion Exposition cutanée et via les muqueuses Exposition externe aux rayonnements ionisants Autre NA	Renseigner mot(s) clé(s)	Choix multiple parmi : Ambiances thermiques Bruit Vibrations Rayonnements ionisants Rayonnements optiques Champs électromagnétiques Nanoparticules de carbone Amiante Bactéries Champignons microscopiques Virus Parasites Prions Pollens Toxines Végétaux Insectes Métaux lourds Pesticides (phytosanitaires et biocides) Contaminants de l'air particuliers Contaminants de l'air gazeux Polluants organiques persistants Hydrocarbures Perturbateurs endocriniens Autre NA	Champ texte libre



Nous avons besoin de vous pour :

➤ **Améliorer en continu les métadonnées santé-environnement développées.**

● La version beta du catalogue est en ligne !



Pour consulter le catalogue,
rendez-vous sur :

<https://gd4h.ecologie.gouv.fr/>

ID : betatesteur@gd4h.fr

MDP : betatesteurgd4h@2022



Présentation de l'offre de service

➤ Outils juridiques

Animation

Catalogue de données

Outils
juridiques

Conception de la brique offre juridique

OBJECTIFS

- _ Améliorer l'accessibilité des BDDs environnementales par les chercheurs et leur partage par les producteurs
- _ Permettre une montée en connaissance de la communauté sur les aspects juridiques liés à la donnée env.
- _ Construire un commun réutilisable pour d'autres projets du MTE et d'autres organismes publics et ministère

QUALIFICATION DES RESSOURCES

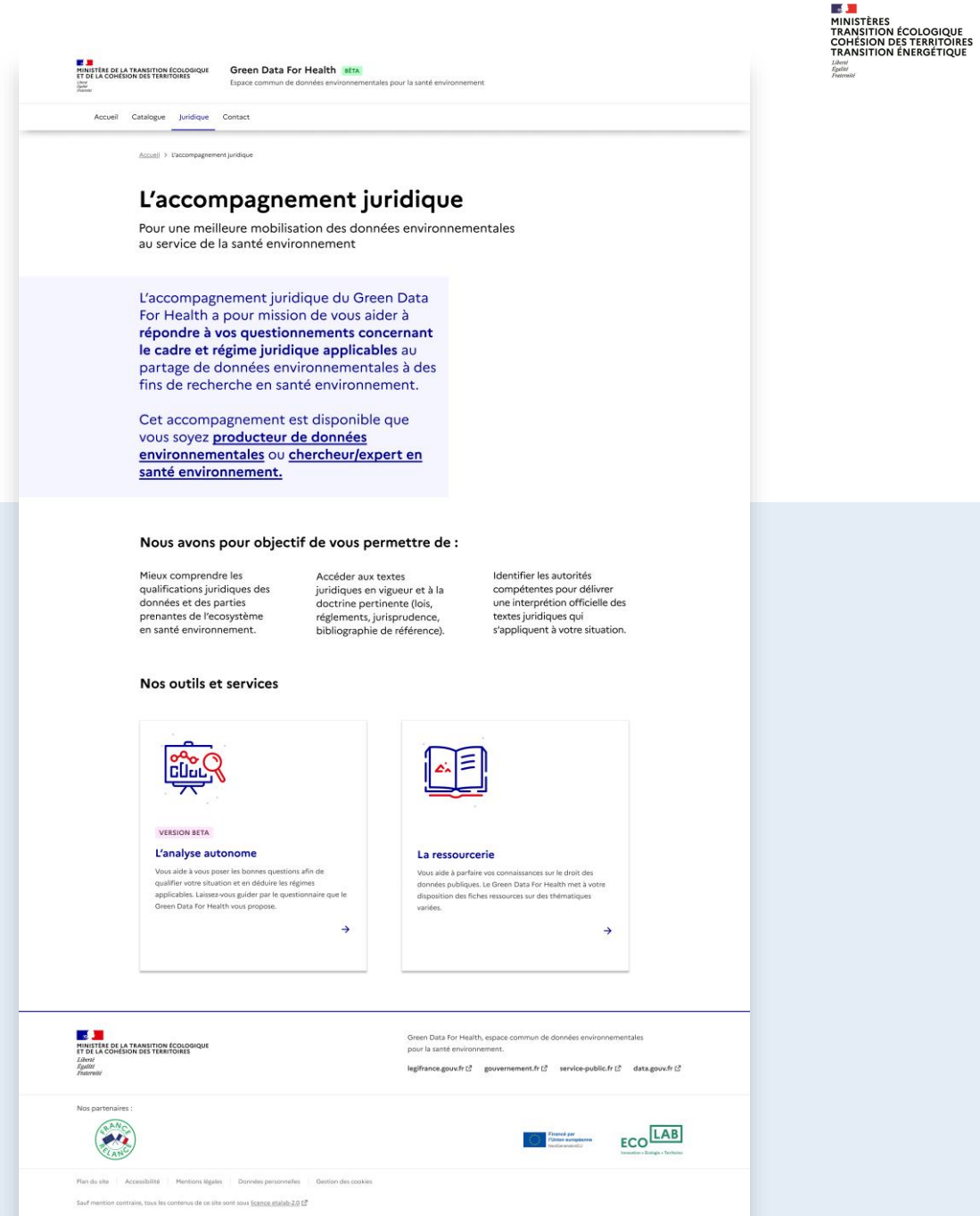
- _ Etat des lieux finalisés (4 niveaux d'accessibilité définis)
- _ Echange avec les producteurs pour affiner la qualification

RESSOURCERIE

- _ 20 fiches juridiques génériques ont rédigées et seront bientôt mises en ligne sur la plateforme

OUTIL ANALYSE AUTONOME

- _ V2 de la structure de l'arbre pour améliorer la clarté du parcours
- _ Rédaction en cours des fiches de régime applicable
- _ Investigation pour réalisation technique



Focus sur les fiches de la ressource

OBJECTIFS

Les fiches pédagogiques doivent permettre à un utilisateur sans connaissance juridique de :

- _ Mieux comprendre la nature juridique des données qu'il utilise
- _ Cerner le régime applicable à ces données
- _ Accéder rapidement à des informations juridiques (normes pertinentes, jurisprudences ou décisions d'autorités compétentes utiles...)

L'OFFRE

La rédaction comprend vingt fiches pédagogiques divisées en trois catégories :

Fiches « Qualifications »

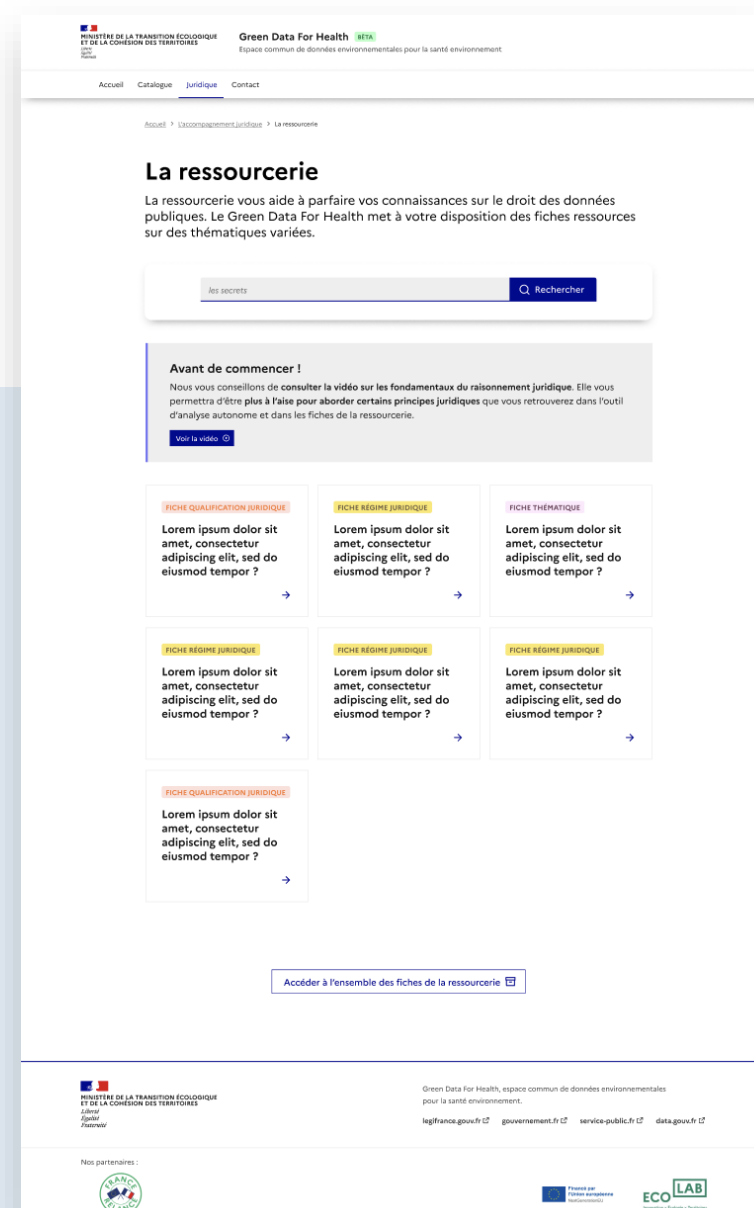
Elles permettent aux utilisateurs de comprendre la qualification juridique de différents catégories de données (exemple: La notion d'information relative à l'environnement, la notion de document administratif ...)

Fiches « Régime »

Elles permettent aux utilisateurs de comprendre le régime applicable à leurs différentes catégories de données (exemple: L'accès aux informations relatives à l'environnement ; La communication des documents des documents administration ...)

Fiches « Focus »

Elles permettent de faire le point sur des notions juridiques transverses dont la compréhension est nécessaire aux utilisateurs (exemple: Le secret industriel ...)





Présentation de l'offre de service

▸ Actions d'animation

Animation

Catalogue de données

Outils
juridiques

● Un socle : la communauté de la donnée en santé-environnement



QUELS ACTEURS ?

Chercheurs, décideurs et experts en santé environnement **référénts** de bases de données de santé ou environnementales pertinentes pour la santé-environnement

Responsables de données, gestionnaires, financeurs,



QUELS OBJECTIFS ?

- **Susciter des collaborations** en intégrant les besoins et priorités de l'écosystème.
 - Mettre en place des **projets concrets autour d'objectifs partagés**
- Faire monter en compétences la communauté** sur des sujets d'intérêt conjoints

● Un socle : la communauté de la donnée en santé-environnement

QUELLES ACTIONS ?



Réunion en plénière 2 fois par an
130 participants le 7 Juillet 2022

À VENIR

- _ Newsletter
- _ Deuxième réunion en plénière de la communauté début 2023
- _ Réflexion en cours sur la mise en place de nouveaux cas concrets



Autres modalités (GT, ateliers, webinaires...) à co-construire avec vous en fonction des besoins partagés

CAS CONCRETS

Objectif : renforcer les liens entre acteurs de la communauté autour de projets facilitant la mobilisation des données env. pour le champ de la santé-environnement, et éclairer le développement du GD4H.

Modalités : 4 cas concrets (eau, chlordécone, bruit, pesticides).

AAP CONJOINT HDH x GD4H

Objectif : renforcer les liens entre acteurs de la communauté autour de projets visant à faciliter un appariement rigoureux entre bases de données environnementales et de santé.

Modalités : accompagnement de 4 projets lauréats.

Calendrier : lancement fin 2022.

DATA CHALLENGE

Objectif : renforcer les liens entre acteurs de la communauté autour de projets visant à améliorer la mobilisation des données environnementales pour la recherche en santé-environnement.

Modalités : accompagnement des différentes équipes.

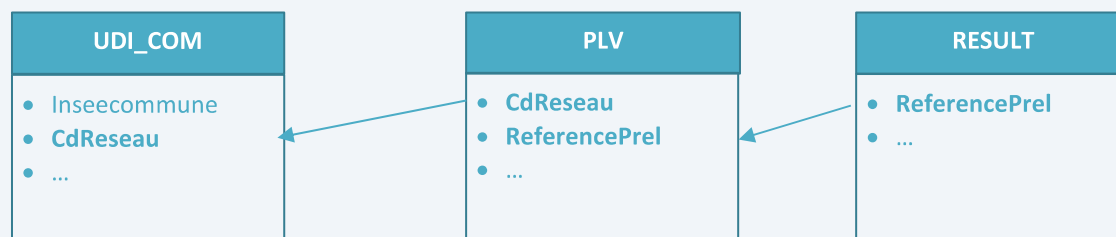
Calendrier : cadrage dès le T4 2022, lancement au S1 2023

Exemple de cas concrets : la qualité de l'eau potable

Le problème

Données de qualité de l'eau potable

- Collectées dans la base SISE-Eaux des Agences Régionales de Santé (ARS)
- Mises à disposition sur le site data.gouv.fr et via l'API [Hub'eau](#)



Données géospatialisées des unités de distribution

Les ARS produisent également des cartes des unités de distribution et le mettent à disposition sur le site [AtlaSante](#).

- Difficulté à obtenir les données de qualité d'eau potable de manière directe et à une maille fine (adresse, communes)
- Difficulté à visualiser facilement le périmètre couvert par les UDI sous-jacents.

L'objectif

Mettre à disposition des chercheurs une jointure entre les données de contaminants de l'eau par UDI (data.gouv.fr) et les données géographiques de ces dernières (AtlaSanté).

Parties prenantes

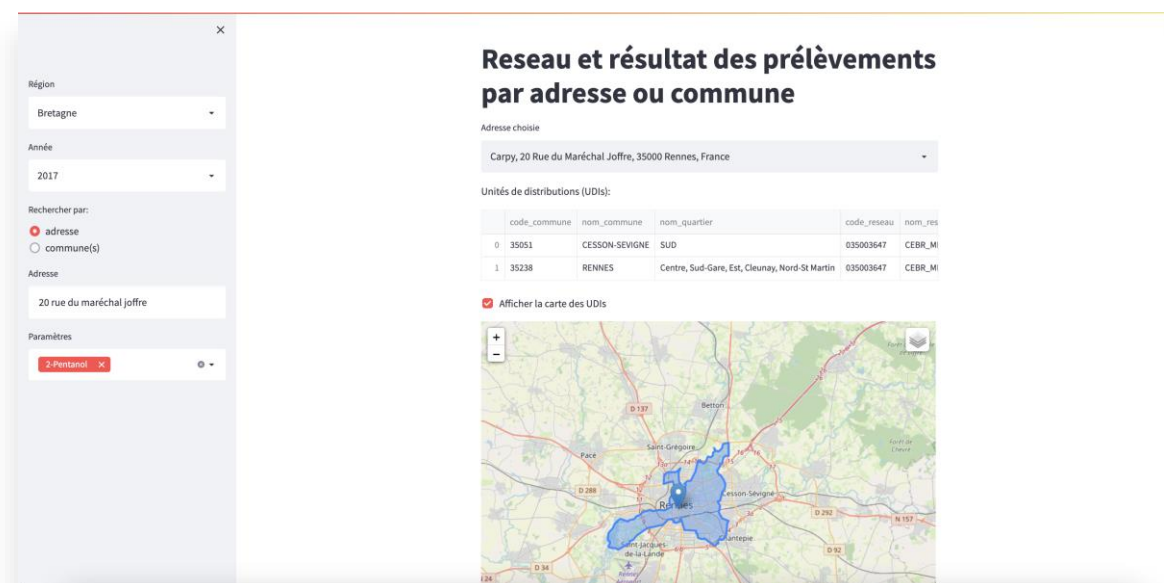
Chercheurs
x
DGS
x
AtlaSanté

Exemple de cas concrets : la qualité de l'eau potable

La solution

Une interface permettant de rechercher, pour une localisation donnée, divers paramètres sur la composition de l'eau potable (e.g. nitrates, herbicides, insecticides...).. L'outil permet de sélectionner une ou plusieurs communes, d'extraire un tableau avec les détails du réseau et les résultats des prélèvements d'une année, pour un ou plusieurs paramètres physico-chimiques choisis.

- Contribution à mieux mobiliser la donnée Environnementale : rapprochement de la couche UDI avec les données commune.



The screenshot shows a web application interface for water quality data. On the left is a sidebar with search filters: Région (Bretagne), Année (2017), Recherche par (adresse selected), Adresse (20 rue du maréchal Joffre), and Paramètres (2 Pentanol). The main panel is titled 'Reseau et résultat des prélèvements par adresse ou commune' and shows the selected address 'Carpy, 20 Rue du Maréchal Joffre, 35000 Rennes, France'. Below this is a table of distribution units (UDIs) and a map showing the location.

Unités de distributions (UDIs):					
	code_commune	nom_commune	nom_quartier	code_reseau	nom_res
0	35051	CESSON-SEVIGNE	SUD	035003647	CEBR_MI
1	35238	RENNES	Centre, Sud-Gare, Est, Cleunay, Nord-St Martin	035003647	CEBR_MI

Below the table, there is a checkbox 'Afficher la carte des UDIs' which is checked, and a map showing the location of the distribution units in Rennes.

Lien vers l'outil : <https://gd4h-eau.lab.sspcloud.fr/>

Appel à projets HDH x GD4H

OBJECTIF :

Stimuler et faciliter les recherches sous-tendant un appariement entre données environnementales et de santé en **accompagnant les chercheurs sur la résolution d'enjeux « données »** (aider à lever des freins tels que le manque de repérabilité, la difficulté d'accès, un déficit de qualité, des difficultés de traitement et d'analyse).

MODALITÉS :

QUI ?

Les projets de recherche en santé-environnement sur les différentes thématiques prioritaires identifiées par le groupe de travail dédié. 4 projets seront sélectionnés.

COMMENT ?

Le GD4H et le HDH proposent un accompagnement humain, technique et financier des lauréats (environ 135 000€ par projet) pour une durée de 18 mois.

QUAND ?

Lancement de l'AAP en octobre 2022.

Sélection des candidats puis annonce des lauréats début 2023.



Présentation du Data Challenge

OBJECTIF :

- Fédérer les acteurs de la santé-environnement et faciliter la mobilisation des données environnementales pour la santé-environnement.
- Mettre en œuvre des projets communs basés sur des cas d'usage utiles pour les parties prenantes du Data Challenge, et pour l'écosystème d'acteurs de la santé-environnement.
 - Produire des outils transférables à d'autres cas d'usage et pérennisables.

DISPOSITIF PROPOSE :

1	Phase de CADRAGE	2	Phase de PRODUCTION	3	Phase de VALORISATION
	1.a. Définition et sélection des défis Recensement des besoins par le GD4H et sélection des défis par l'instance de pilotage du Data Challenge. 1.b. Expertisation technique des défis Avec l'appui notamment du prestataire, afin de garantir la solidité technique des défis présentés.		Un déroulement sur 2 mois pour favoriser la qualité des productions. • Constitution des équipes et appropriation des projets • Production sur 6 semaines (3 sprints de deux semaines)		3.a. Finalisation des projets (avec l'appui d'un prestataire) pour garantir la qualité des outils mis en ligne. 3.b. Mise en ligne des productions • Outils • Données et data papers 3.c. Accompagnement post-Challenge éventuel pour quelques projets à fort impact
	➤ Grille de défis à forte plus-value pour l'écosystème santé-environnement		➤ Produits dérivés de données permettant de faciliter la mobilisation des données environnementales.		➤ Jeux de données prêts à l'emploi et bien documentés (rédaction de « data papers »). ➤ Outils remobilisables et pérennisables mis en accès libre.
	Le « off » du Data Challenge : les MASTERCLASS + Pour dynamiser le Data Challenge et l'ouvrir vers l'extérieur, des présentations interactives seront données par des personnalités de haut niveau, et ouvertes au public				

Phase de CADRAGE : définition des défis

➤ Partagez-nous vos souhaits et idées !

COMMENT SOUMETTRE VOS IDEES ?

Vous pouvez exprimer vos besoins d'outils et vos idées de défis ici :

➤ <http://enqueteur.cgdd.developpement-durable.gouv.fr/index.php/476527?lang=fr>.

NB : vos propositions n'ont pas besoin d'être totalement matures/abouties à ce stade : elles pourront être complétées ensuite avec notre appui pour aboutir à une liste de défis potentiels, utiles pour l'écosystème des acteurs mobilisant la donnée environnementale en santé-environnement.

POURQUOI PARTICIPER ?

Une fois la liste des défis potentiels finalisée, **15 défis** seront sélectionnés pour être implémentés dans le cadre du Data Challenge. Celui-ci permettra :

- ✓ **La constitution d'équipes dédiées pour chaque défi : des profils aux compétences techniques adéquates** seront recrutés et travailleront en équipe sur un défi spécifique pour une durée de 6 à 8 semaines ;
- ✓ **Un suivi de la qualité des produits développés pour chaque défi** par des experts de la donnée, afin d'aboutir à des produits de qualité, accessibles en open source et pérennisables ;
- ✓ Eventuellement, pour les projets à fort impact, un accompagnement post-challenge (humain et/ou technique) adapté selon les besoins spécifiques identifiés.

Phase de CADRAGE : sélection et expertisation des défis

Comité des porteurs santé-environnement (CPSE)

Acteurs impliqués : les porteurs de défis (acteurs dans le domaine de la santé-environnement*). Le comité pourrait être structuré en sous-comités (« familles de défis ») selon les types de facteurs environnementaux retenus.

Objectifs :

- Proposer des défis utiles et à forte valeur ajoutée pour les acteurs en santé-environnement.
- Inclure et valoriser les porteurs de défis afin d'assurer un portage de qualité durant et après le Data Challenge.

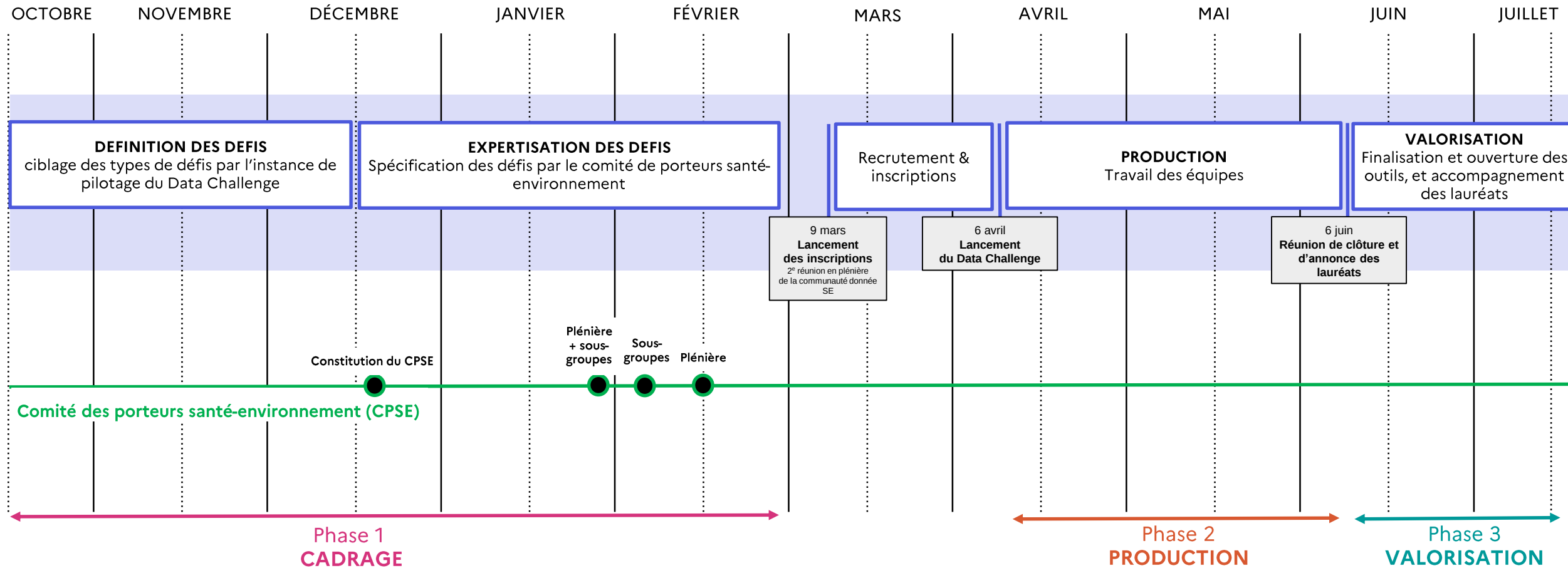
Fonctionnement : plusieurs sessions pour définir les défis.

- Session 1 : réunion en plénière pour présenter les objectifs du comité, puis travail en sous-groupes par thématique afin d'identifier des défis pertinents.
- Session 2 : travail en sous-groupe par thématique pour cadrer techniquement les défis.
- Session 3 : réunion en plénière avec le prestataire pour expertiser les défis d'un point de vue technique.

*Les porteurs de défis pourraient inclure des organismes producteurs/gestionnaires/référents de bases de données environnementales, des réutilisateurs de ces données (notamment organismes publics d'expertise, chercheurs...)

Calendrier prévisionnel du Data Challenge

GD4H



Partie 2

Améliorer la mobilisation des données environnementales en santé-environnement pour les territoires

- **SONDAGE** : recueil des points de difficulté et des besoins
- **FORMULAIRE** : proposition de défis/d'outils
- **Echanges**

SONDAGE : recueil des points de difficulté et des besoins



Nous avons besoin de vous pour :

- Consolider notre compréhension de vos points de difficultés et besoins en termes de mobilisation de la donnée environnementale pour la santé-environnement.

COMMENT ?

En participant au sondage accessible au lien suivant :

- <http://enqueteur.cgdd.developpement-durable.gouv.fr/index.php/213352?lang=fr>

N'hésitez pas à communiquer ce lien à tout acteur concerné par la mobilisation de données environnementales pour la santé-environnement !

FORMULAIRE : proposition de défis/d'outils



Nous avons besoin de vous pour :

- Définir les types d'outils et de défis à privilégier dans le cadre du Data Challenge.

COMMENT SOUMETTRE VOS IDEES ?

Vous pouvez exprimer vos besoins d'outils et vos idées de défis ici :

➤ <http://enqueteur.cgdd.developpement-durable.gouv.fr/index.php/476527?lang=fr>.

N'hésitez pas à communiquer ce lien à tout acteur concerné par la mobilisation de données environnementales pour la santé-environnement !

DISCUSSION : articulation du GD4H avec les PRSE4

?

**Quels enjeux de mobilisation des données
environnementales dans le cadre des PRSE4 ?**