

# Data Challenge

« Mobiliser la donnée environnementale en santé-environnement »



# Présentation générale du dispositif

# Une initiative portée par l'Ecolab (MTECT/CGDD) dans le cadre du Green Data for Health (GD4H)

Le Green Data for Health (GD4H) est un **espace commun de données** en santé-environnement incubé au sein de l'ECOLAB, laboratoire d'innovation par la donnée du Service de la Recherche et de l'Innovation du Ministère de la Transition écologique et de la Cohésion des territoires.

## Il s'agit d'une action phare du 4<sup>e</sup> Plan National Santé Environnement

- "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."

Le GD4H vise à contribuer à la production de connaissances par la donnée dans le domaine de la santé-environnement.

- Un Data Challenge qui aura lieu au premier semestre 2023 visera ainsi à développer des outils/projets visant à faciliter la mobilisation des données environnementales en santé-environnement.

### Animation

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

### 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.**

### Outils juridiques

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

# Pourquoi un Data Challenge ?

## MOTIVATION :



### CONSTAT

Les **données environnementales existantes restent peu mobilisées** pour faire émerger des associations entre expositions environnementales et pathologies humaines.



### ENJEU À ADRESSER

Améliorer les connaissances sur l'impact sanitaire des facteurs environnementaux par une **facilitation de la mobilisation des données environnementales pour la santé-environnement.**



### COMMENT ?

Par un **ancrage dans la communauté** de la donnée en santé-environnement animée par le Green Data for Health, afin d'**adresser des problématiques partagées.**

## OBJECTIF :

Fédérer les acteurs de la santé-environnement et faciliter la mobilisation des données environnementales pour la santé-environnement.

1/ 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.

2/ Produire des **outils transférables à d'autres cas d'usage et pérennisables.**

# Quel est le dispositif proposé ?

1

## Phase de **CADRAGE**

### **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.

- Grille de défis à forte plus-value pour l'écosystème santé-environnement

# Quel est le dispositif proposé ?

1

## Phase de CADRAGE

### 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.

➤ Grille de défis à forte plus-value pour l'écosystème santé-environnement

2

## Phase de PRODUCTION

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)

➤ Produits dérivés de données permettant de faciliter la mobilisation des données environnementales.

# Quel est le dispositif proposé ?

## 1 Phase de CADRAGE

### 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.

➤ Grille de défis à forte plus-value pour l'écosystème santé-environnement

## 2 Phase de PRODUCTION

### 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)

➤ Produits dérivés de données permettant de faciliter la mobilisation des données environnementales.

## 3 Phase de VALORISATION

**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**

➤ Jeux de données prêts à l'emploi et bien documentés (rédaction de « data papers »).  
➤ Outils remobilisables et pérennisables mis en accès libre.

# Quel est le dispositif proposé ?

<b>1</b>	<b>Phase de CADRAGE</b>  <b>1.a. Définition et sélection des défis</b> Recensement des besoins par le GD4H et sélection des défis par l'instance de pilotage du Data Challenge.  <b>1.b. Expertisation technique des défis</b> Avec l'appui notamment du prestataire, afin de garantir la solidité technique des défis présentés.	<b>2</b>	<b>Phase de PRODUCTION</b>  <b>Un déroulement sur 2 mois pour favoriser la qualité des productions.</b> <ul style="list-style-type: none"><li>• Constitution des équipes et appropriation des projets</li><li>• Production sur 6 semaines (3 sprints de deux semaines)</li></ul>	<b>3</b>	<b>Phase de VALORISATION</b>  <b>3.a. Finalisation des projets (avec l'appui d'un prestataire)</b> pour garantir la qualité des outils mis en ligne.  <b>3.b. Mise en ligne des productions</b> <ul style="list-style-type: none"><li>• Outils</li><li>• Données et data papers</li></ul> <b>3.c. Accompagnement post-Challenge éventuel pour quelques projets à fort impact</b>
➤ 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.	

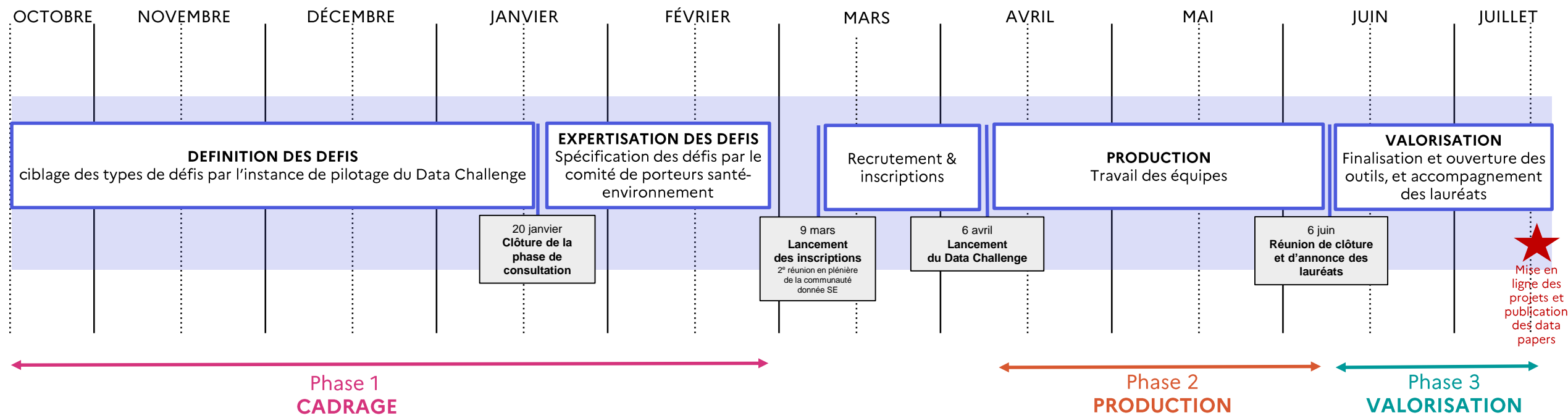
## Le « off » du Data Challenge : les MASTERCLASS

+ Des présentations interactives seront données par des personnalités de haut niveau, et ouvertes au public



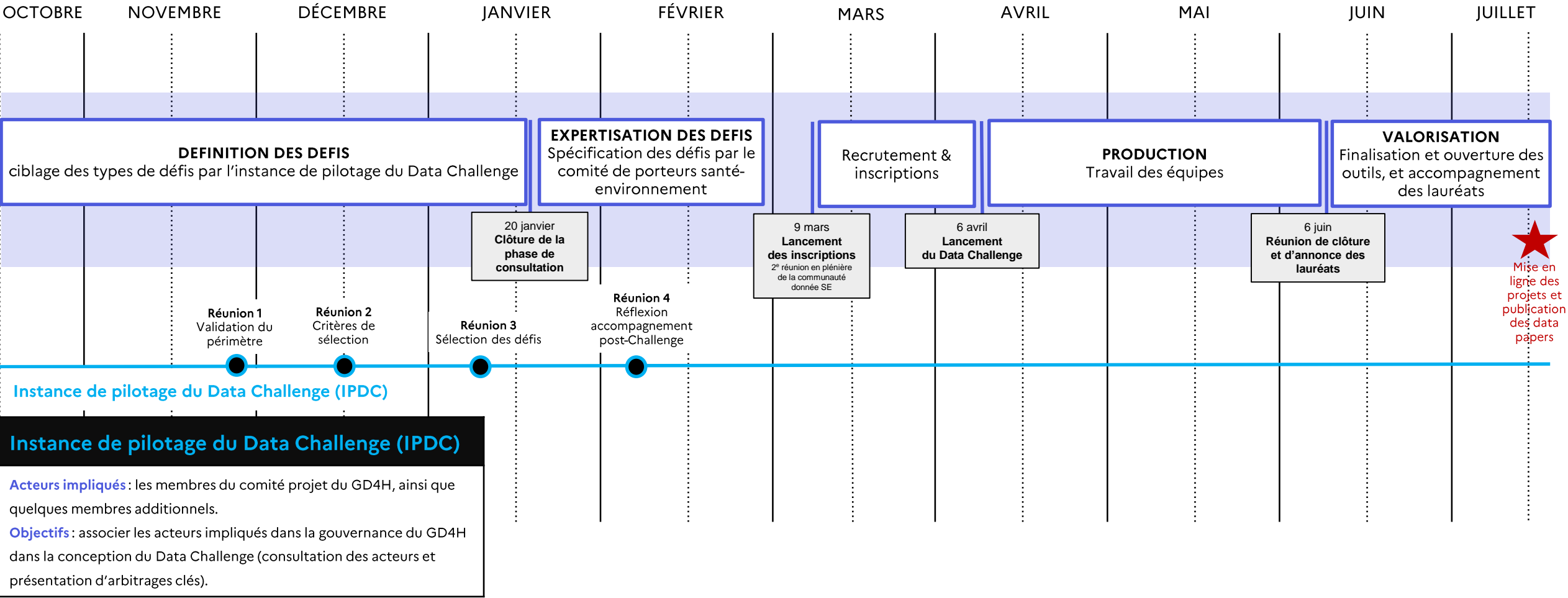
# Présentation générale

## CALENDRIER PREVISIONNEL :



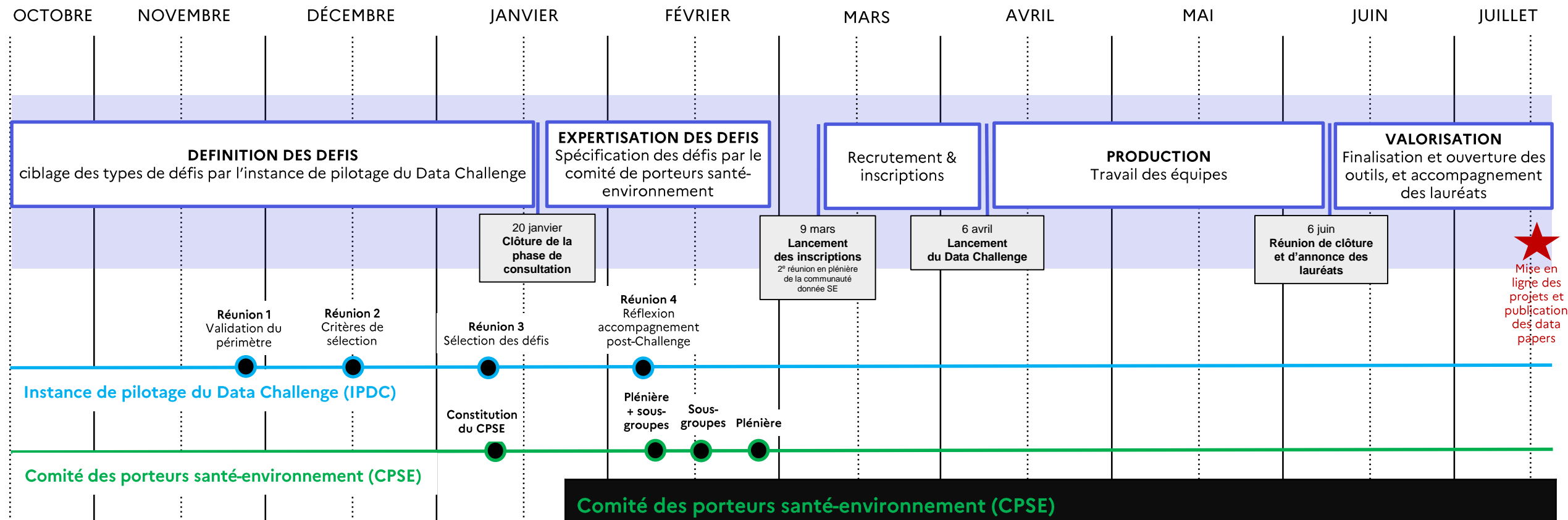
# Présentation générale

## CALENDRIER PREVISIONNEL :



# Présentation générale

## CALENDRIER PREVISIONNEL :



### Comité des porteurs santé-environnement (CPSE)

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

#### Objectifs :

- Proposer des défis utiles et pertinents (expertise) 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.



# Q&A : points soulevés fréquemment

# Q1 / Quelles sont les informations à renseigner pour proposer un défi ? Quel degré de maturité mon idée de défi doit-elle avoir ? Quelle date limite ?



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 !

Jusqu'au 20 janvier 2023

### Data Challenge "Mobiliser la donnée environnementale en santé-environnement" (MTECT/CGDD/ECOLAB)

Bonjour,

Vous souhaitez disposer de nouveaux outils pour mieux mobiliser la donnée environnementale en santé-environnement ?

Le Data Challenge organisé par le Green Data for Health<sup>1</sup> au premier semestre 2023 est fait pour vous ! Nous vous proposons ainsi de soumettre au sein de ce court formulaire vos propositions d'outils/de défis à développer dans le cadre de ce Challenge.

#### Pourquoi proposer des outils/défis ?

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.

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 pour le Data Challenge 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.

#### Quels types de défis/d'outils peuvent être proposés ?

Les défis/outils développés peuvent répondre à tout point de difficulté rencontré en termes de mobilisation de la donnée environnementale en santé-environnement. Les outils/défis pourraient notamment porter sur (liste non-exhaustive) :

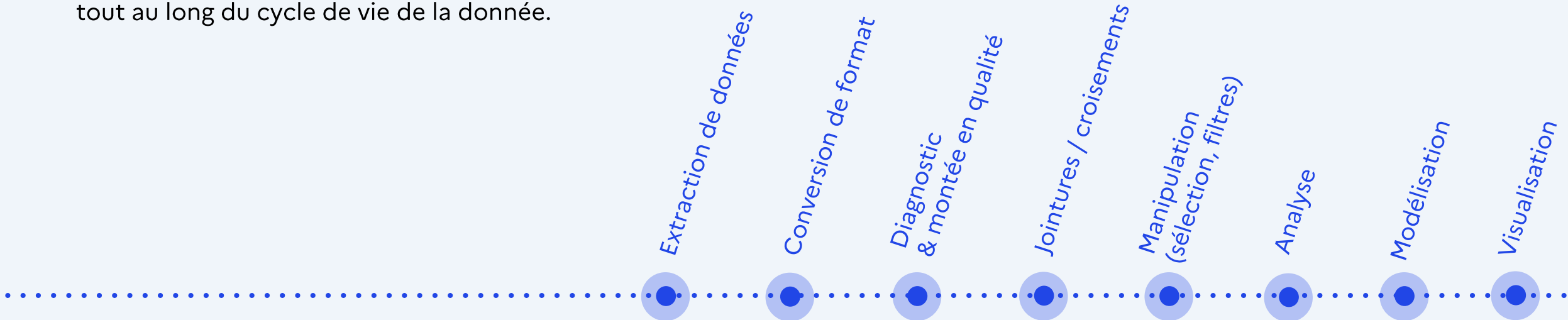
- La création ou la modélisation de nouvelles données.
- L'extraction de données.

<sup>1</sup> Le Green Data for Health (GD4H) est un espace commun de données incubé au sein de l'Ecolab, laboratoire d'innovation par la donnée du Commissariat général au développement durable (CGDD) du Ministère de la Transition écologique et de la Cohésion des territoires et du Ministère de la Transition énergétique. Le GD4H s'inscrit dans l'une des actions phares du 4e Plan National Santé Environnement (PNSE4), et vise à faciliter la mobilisation et la valorisation des données environnementales, par les acteurs de la recherche et de l'expertise, au service de la santé-environnement.

## Q2 / Mon idée de défi s'inscrit-elle dans le périmètre visé par ce Data Challenge ?

### PERIMETRE

Tout type d'outil permettant de **mieux mobiliser la donnée environnementale au service de la santé environnement** entre dans le périmètre du Data Challenge. Les outils peuvent ainsi viser différents freins rencontrés en termes de mobilisation de la donnée, tout au long du cycle de vie de la donnée.



### CRITERES

Des outils :

- \_ Utiles pour plusieurs acteurs
- \_ Partagés en ligne
- \_ Transférables entre cas d'usage
- \_ Pérennisables

## Q3 / Quel est le profil des équipes constituées pour les défis ?

### PROFIL DES EQUIPES

\_ **Socle minimal souhaité** : 1 chef de projet animateur, un porteur de projet métier (e.g. producteur de données, chercheur), 4-5 profils techniques (dont un profil expérimenté).

\_ Profils techniques recrutés :

> Un **recrutement des profils techniques modulable**, car effectué après stabilisation des cahiers des charges

> Une **participation bénévole** des profils recrutés.

\_ Autres profils techniques : mise à disposition de profils par les structures dans le cadre (i) d'un défi particulier ou (ii) de manière transverse/en appui aux différents défis.

.

## Q4 / Quel est le niveau d'implication attendu des parties prenantes ?

### NIVEAU D'IMPLICATION DES PARTIES PRENANTES

#### \_ Phase de cadrage

- > **sélection** : implication dans les travaux du CPSE (3 demi-journées sur janvier/février)
- > **préparation** : éventuellement (si pertinent), facilitation de l'accès à des données et outils détenues.

#### \_ Phase de production :

- > **En tant que porteur** : disponibilité pour répondre aux questions/arbitrages « métier » des profils techniques, suivi du projet.
- > **En tant que chef de projet animateur** : participation à l'organisation de l'équipe, suivi régulier et contrôle de l'adéquation au cahier des charges défini.

#### \_ Phase de valorisation : transfert facilité par l'investissement de ressources métier et techniques lors des phases précédentes.



# ● N'hésitez pas à communiquer sur le Data Challenge !

E-mail à diffuser + support de présentation



- + Sessions de webinaire
  - **Vendredi 16/12 à 12h**
  - **Vendredis 06/01 et 13/01 à 13h30**

- Nous souhaitons recueillir au maximum vos 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 !