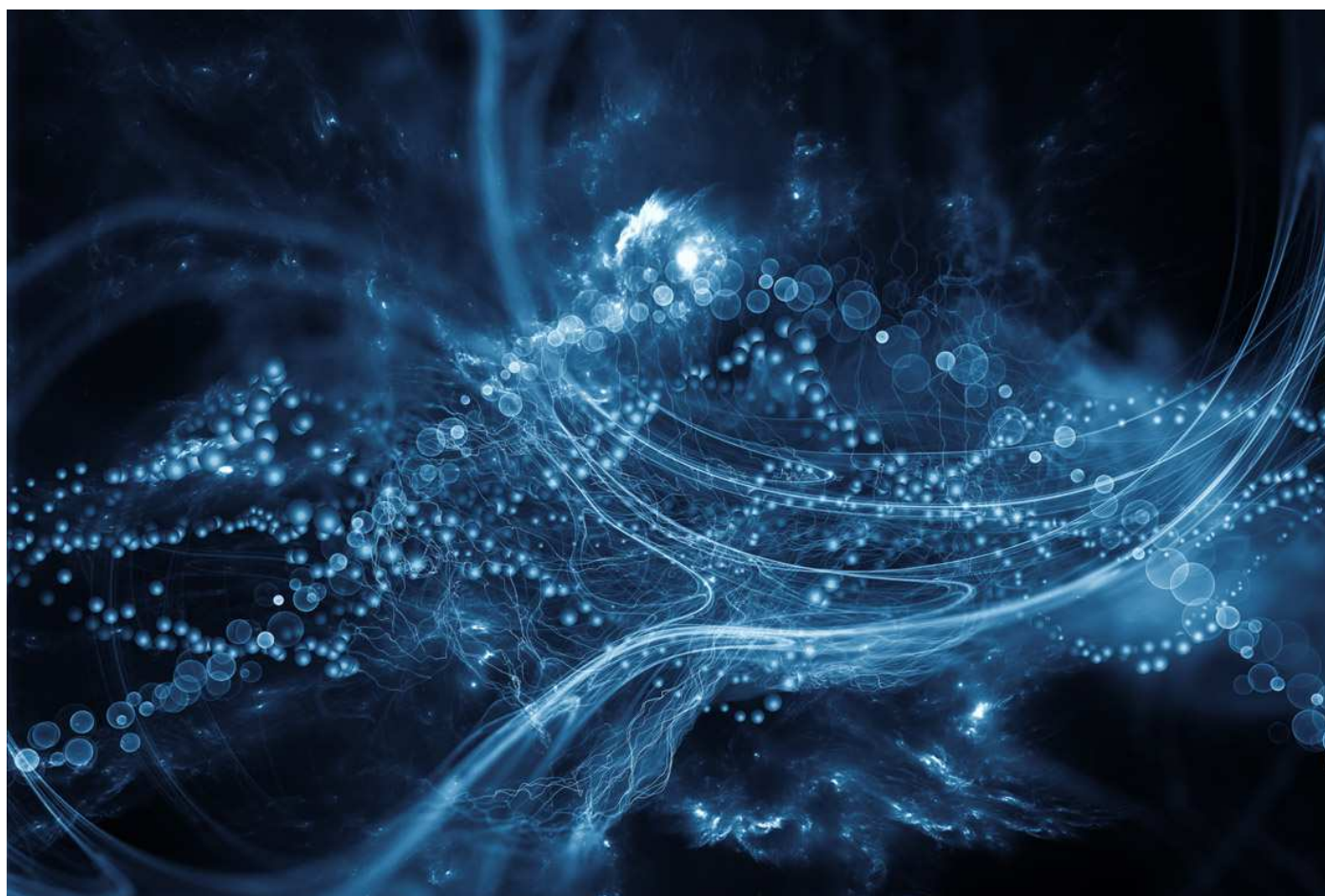




MINISTÈRES  
TRANSITION ÉCOLOGIQUE  
COHÉSION DES TERRITOIRES  
MER

*Liberté  
Égalité  
Fraternité*

# Feuille de route de la donnée, des algorithmes et des codes sources



Septembre 2021

# SOMMAIRE

Préambule .....	4
Pourquoi définir et déployer une feuille de route de la donnée, des algorithmes et des codes sources pour la transition écologique ? .....	5
1. Être au rendez-vous de la nouvelle politique publique de la donnée.....	5
2. Répondre à trois enjeux essentiels .....	7
2.1. Exploiter les données pour mener des politiques publiques plus pertinentes et efficaces .....	7
2.2. Partager les données, les algorithmes et les codes sources pour renforcer l'impact écologique des politiques publiques .....	8
2.3. S'inscrire dans la démocratisation de la gestion de la donnée, des algorithmes et des codes sources pour une plus grande transparence de l'action publique et davantage d'innovation .....	9
Quels sont les objectifs de la feuille de route?.....	12
1. Objectif : Rendre la donnée mobilisable .....	13
1.1. Augmenter la qualité des données en vue de leur réutilisation.....	14
1.2. Promouvoir et amplifier l'APIsation de la donnée de la transition écologique .....	16
1.3. Mettre en place un guichet d'accès à la donnée de la transition écologique et de la cohésion des territoires .....	16
1.4. Généraliser le partage des codes sources et algorithmes présentant un intérêt général en termes de connaissance et de transparence de l'action publique.....	17
2. Objectif : Développer la culture et l'usage de la donnée.....	19
2.1. Faire prendre aux décideurs le virage de la donnée .....	19
2.2. Faire monter en compétence les praticiens métiers ( <i>makers</i> ).....	20
2.3. Se doter de ressources humaines et d'ingénierie financière.....	20
2.4. Innover pour entretenir la compétence « donnée » dans le temps et fidéliser les ressources.....	21
2.5. Mobiliser des méthodes pragmatiques pour montrer la valeur de la donnée .....	22
2.6. Poursuivre les actions des start-up d'Etat .....	23
3. Objectif : Partager les données entre administrations.....	24
3.1. Animer le partage de données entre les administrations .....	24
3.2. Réussir les initiatives de coopération entre administrations .....	25
3.3. Poursuivre l'ouverture des données publiques, des algorithmes et des codes sources, en priorisant les usages .....	26
4. Objectif : Partager des données par la coopération public-privé.....	27
4.1. Promouvoir le partage de données entre le secteur public et le secteur privé .....	28
4.2. Réussir des premières initiatives concrètes, notamment les Espaces Communs de Données (ECD).....	29
4.3. Faciliter l'accès aux données publiques de la transition écologique via un guichet d'accès et des correspondants clairement identifiés .....	30

- 4.4. Faciliter l'accès aux données privées d'intérêt général pour la Transition écologique  
30

<b>5. Objectif : Positionner la France dans la stratégie de donnée de l'Union européenne .....</b>	<b>31</b>
5.1. Mettre en place un dispositif de coopération européenne .....	31
5.2. Faire de la France une force de proposition dans les négociations.....	32
5.3. Relayer auprès des services et des partenaires les initiatives européennes et favoriser les recherches de partenariat et de financement européens.....	32
<b>6. Objectif : Utiliser les données de manière conforme et sécurisée.....</b>	<b>32</b>
6.1. Cerner le périmètre juridique concerné.....	33
6.2. Vers un accompagnement juridique centralisé .....	34
<b>7. Faire vivre la feuille de route via une gouvernance dédiée.....</b>	<b>34</b>
7.1. Elaborer et mettre en œuvre la feuille de route de la donnée, des algorithmes et des codes sources.....	34
7.2. Piloter dans le temps la mise en œuvre de la feuille de route de la donnée, des algorithmes et des codes sources.....	35
7.3. Faire vivre la feuille de route autour de la Communauté de la donnée, des algorithmes et des codes sources.....	36
7.4. Animer la donnée au niveau territorial dans un lieu d'échange unique .....	37
7.5. Actionner la feuille de route de manière coordonnée avec les gouvernances extérieures .....	37
<b>8. Équiper le pôle ministériel avec des outils adaptés en matière de données.....</b>	<b>38</b>
8.1. S'inscrire au maximum dans l'écosystème technique et applicatif existant ou planifié 39	
8.2. Réussir la Géoplateforme : l'outil collaboratif de production et d'entretien de la donnée.....	40
8.3. Mettre à disposition une plateforme de services de la donnée et les capacités d'hébergement associées.....	40
8.4. Déployer des outils d'aide à la décision (visualisation de la donnée ou solutions décisionnelles) .....	41
8.5. Mettre en place un outil collaboratif communautaire.....	41
Annexe 1 - Méthode mise en œuvre et participants.....	42
Annexe 2 - Panorama des acteurs de la donnée du pôle ministériel .....	43
Annexe 3 – Répartition des rôles CGDD-SNum sur le volet data du Plan de Transformation Numérique .....	48
Annexe 4 – Extrait de l'offre de services du Service du Numérique (SNum) .....	49

# Préambule

La feuille de route de la donnée, des algorithmes et des codes sources du pôle ministériel<sup>1</sup> fixe le cap à atteindre en matière de politique générale et initie les actions correspondantes.

Un meilleur usage de la donnée publique et privée accélèrera la transition écologique. Tous les leviers à disposition du pôle ministériel sont mobilisés dans ce but, en lien avec le réseau interministériel et en coopération avec ses partenaires : collectivités territoriales, entreprises, associations, collectifs porteurs de "communs" numériques, écosystèmes d'innovation, chercheurs, instances européennes. L'environnement technologique, juridique, économique, institutionnel, stratégique de la donnée connaît une mutation profonde que le pôle ministériel accompagne et pour laquelle il se veut proactif, exigeant et innovant. Cela concerne tout autant les données intéressant les directions métier que les directions transverses. La donnée est l'affaire de toutes et tous, quelles que soient nos fonctions, que nous soyons déjà familiers ou non de la donnée.

Le commissariat général au développement durable (CGDD) est l'administrateur ministériel des données (AMD) des ministères de la transition écologique et de la mer (depuis mai 2021, le ministère de la cohésion des territoires et des relations avec les collectivités territoriales dispose de son propre administrateur ministériel des données, au sein de l'agence nationale de la cohésion des territoires [ANCT]). Il a piloté, en lien avec le secrétariat général, la préparation et la mise en œuvre de la feuille de route via son laboratoire d'innovation Ecolab.

Une concertation avec les directions générales, les services déconcentrés et les opérateurs a permis de recueillir leur vision, leurs usages de la donnée et leurs besoins<sup>2</sup>.

Un groupe de travail, composé du CGDD, via son Ecolab, et du secrétariat général, via son service du numérique (SNum), a croisé l'ensemble des résultats collectés et proposé de dessiner une ambition collective, à la fois pragmatique et innovante. Cette ambition s'articule autour de 6 objectifs et 13 actions phare. Elle constitue une réponse à la nouvelle politique publique de la donnée, des algorithmes et des codes sources impulsée par le premier ministre le 5 février 2021.

Pour autant, la feuille de route n'a pas vocation à retracer de façon exhaustive les initiatives du pôle ministériel. Un choix de recommandations et de thématiques a été effectué en fonction des besoins identifiés par les acteurs rencontrés. Du fait de ce parti pris, il s'agit d'un document vivant. Même si les actions identifiées sont exprimées à horizon 3-4 ans, la feuille de route est amenée à évoluer et à être enrichie au fil de l'eau en fonction de l'avancement des projets et des changements de contexte et de priorités en matière de politiques publiques et d'usages.

Une synthèse de la feuille de route est disponible. Il est conseillé de s'y référer dans un premier temps puis de se plonger dans ce document plus complet. La synthèse met en avant les 13 actions phare de la feuille de route, tandis que le présent document détaille l'intégralité des actions prévues.

La présente feuille de route contient une action phare dédiée à l'intelligence artificielle, renvoyant vers une feuille de route dédiée. Celle-ci a été élaborée avec la même méthodologie pour organiser la mobilisation de l'intelligence artificielle au service de la transition écologique : « *Feuille de route intelligence artificielle et transition écologique du pôle ministériel 2021 - 2024* ».

---

<sup>1</sup> 1 Ministère de la Transition écologique - Ministère de la Cohésion des territoires et des Relations avec les collectivités territoriales - Ministère de la Mer.

<sup>2</sup> Le détail des acteurs rencontrés est présenté en annexe n°1.

# PARTIE 1

## Pourquoi définir et déployer une feuille de route de la donnée, des algorithmes et des codes sources pour la transition écologique ?

### 1. Être au rendez-vous de la nouvelle politique publique de la donnée

Comprendre le réel et l'anticiper suppose d'accéder à des données. Celles-ci, de sources multiples et disparates, peuvent être valorisées par croisements, traitements massifs, visualisations, expertises. Cela nécessite au préalable que les données circulent, soient aisément accessibles et puissent être resituées dans leur contexte.

C'est là l'ambition portée par la **feuille de route** : promouvoir l'intérêt de produire, entretenir et partager une donnée publique de qualité, proposer des outils pour l'exploiter et pour la croiser avec des données tierces, mettre en place les dispositifs de protection de la donnée et de conseil aux services, veiller harmoniser les attentes des acteurs internes et externes du pôle ministériel, organiser la montée en compétence collective.

**La circulaire du Premier ministre du 27 avril 2021 à destination des ministres et des préfets de région relative à une politique de la donnée, des algorithmes et des codes sources** pose de nouvelles ambitions.

Cette circulaire fait suite au Comité interministériel de la transformation publique (CITP) du 5 février 2021, qui a également prescrit l'ouverture de nouveaux jeux de données publics, et aux préconisations du rapport du député Eric Bothorel « Pour une nouvelle ère de la politique publique de la donnée » remis au gouvernement le 23 décembre 2020.

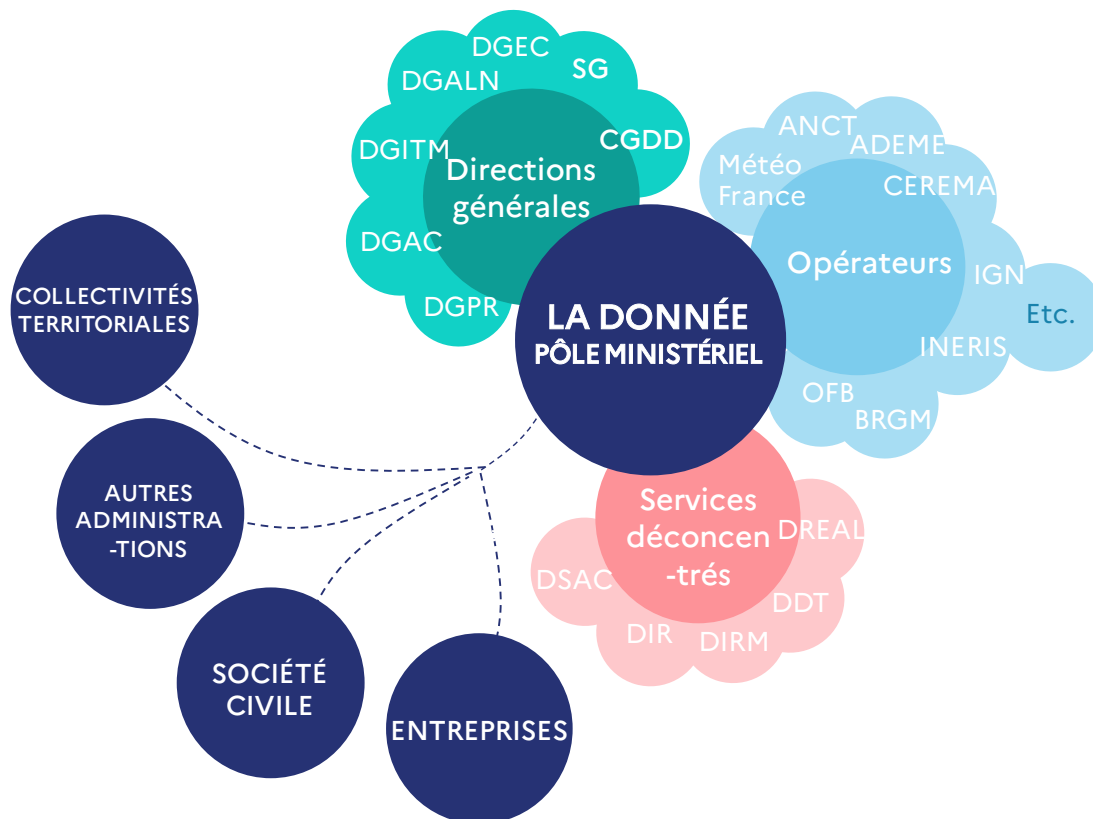
Ces fortes impulsions politiques et institutionnelles reposent sur le constat de l'importance stratégique de l'ouverture, de l'exploitation et du partage de la donnée, des algorithmes et des codes sources publics, et de leur meilleur croisement avec les données des acteurs privés. Elles resituent ces enjeux dans le contexte ambitieux des initiatives lancées au niveau européen sur la transition numérique, la donnée et l'intelligence artificielle. Elles sont également un élément de la participation de la France au Partenariat mondial pour un gouvernement ouvert (PGO), auquel elle a adhéré en 2015 et qui ambitionne un meilleur dialogue des administrations avec la société civile et les citoyens pour contribuer à la diffusion d'une culture d'ouverture de l'action publique.

La feuille de route constitue un élément clé de la mise en œuvre par le pôle ministériel de la circulaire du 21 avril 2021 « Politique de la donnée, des algorithmes et des codes sources ». Elle fait leur place aux enjeux plus récents des algorithmes et codes sources, qui sont des données publiques d'un type particulier, en renforçant leur ouverture sous pilotage du service du numérique du secrétariat général.

La feuille de route est un repère et un appui pour les stratégies et plans d'actions des services et opérateurs du pôle ministériel. Une **communauté de la donnée, des algorithmes et des codes sources du pôle ministériel** est constituée au printemps 2021 comme lieu de partage et de mise en synergie des usages de la donnée.

La feuille de route s'adresse à **l'ensemble des agents et équipes du pôle ministériel**, quel que soit leur niveau d'expertise lié à la donnée : décideurs, praticiens métiers ou acteurs techniques de la donnée. Les actions à déployer bénéficieront à tous : directions centrales ; opérateurs ; services déconcentrés.

L'écosystème de la donnée représenté dans le schéma ci-dessous rappelle que les acteurs de la donnée au sein du pôle ministériel ne peuvent agir dans le déploiement actif de cette politique sans liens étroits avec leurs partenaires. La publication de cette feuille de route y participera, notamment dans le cadre des objectifs 3, 4 et 5 qui y sont dédiés. Leur périmètre de responsabilité quant à la politique de la donnée est précisé et détaillé en annexe.



**Notre pôle ministériel est en avance, tiré par les obligations réglementaires et législatives des 20 dernières années, qui très tôt ont contraint la sphère environnementale à agir**

Au titre des obligations générales qui pèsent sur les administrations en matière de politique de la donnée, la transparence et le partage de l'information environnementale sont des acquis issus de deux textes prépondérants : la convention d'Aarhus - accès à l'information, participation du public au processus décisionnel, accès à la justice en matière d'environnement (1998) - et la directive européenne INSPIRE – qui vise à établir une infrastructure d'information géographique dans l'Union européenne pour favoriser la protection de l'environnement (2007).

Ce contexte a conduit le pôle ministériel à être précurseur en matière de dispositifs techniques et institutionnels de mise à disposition de ses données. Le périmètre des données relatives à la transition écologique (ex. transports, énergie, prévention des risques, logement, mer) s'étend bien entendu au-delà des seules données environnementales.

La politique de la donnée du pôle ministériel s'est renforcée ensuite sous l'impulsion des dispositions nées en grande partie de la loi du 28 décembre 2015 relative à la gratuité et aux modalités de la réutilisation des informations du secteur public et de la loi du 7 octobre 2016 pour une République numérique. Ces textes ont largement renforcé les dispositifs d'ouverture, de réutilisation et de partage des données publiques (*open data*). D'autres dynamiques s'y ajoutent, telles que les stratégies de l'Union européenne sur la transition numérique, les données ou l'intelligence artificielle.



## 2. Répondre à trois enjeux essentiels

La feuille de route de la donnée, des algorithmes et des codes sources pour la transition écologique sert trois enjeux essentiels :

- **Exploiter les données pour mener des politiques publiques plus pertinentes et efficaces**
- **Partager les données, les algorithmes et les codes sources pour renforcer l'impact écologique des politiques publiques**
- **S'inscrire dans la démocratisation de la gestion de la donnée, des algorithmes et des codes sources pour une plus grande transparence de l'action publique et davantage d'innovation**

### 2.1. Exploiter les données pour mener des politiques publiques plus pertinentes et efficaces

«[...] Ne pas anticiper, se contenter de réagir sans prévoir, sans prévenir, ne pas construire aujourd'hui les conditions de notre résilience demain : ce serait une faute impardonnable. » (la ministre de la transition écologique Barbara Pompili le 10 novembre 2020 à l'Assemblée nationale lors de la présentation du projet de loi de Finances pour 2021 pour le pôle ministériel).

En maîtrisant des données de qualité (= savoir), des décisions pertinentes peuvent être prises (= comprendre) afin d'améliorer l'efficacité et l'efficience d'une politique publique (= agir). Cet enchaînement d'actions constitue la trame type de toute fonction de pilotage.

**Le tableau ci-dessous illustre de façon simplifiée chaque étape de mobilisation de la donnée par un usage de la transition écologique :**

Etapes de mobilisation de la donnée	Exemples d'usage de la donnée pour la transition écologique et les territoires	Politiques du pôle ministériel concernées par les exemples
<b>Mieux savoir</b> ce qui se passe sur le terrain		
Mesurer	Quantifier les émissions de gaz à effet de serre	Climat
Agréger	Recueillir les comptages d'espèce marine des chercheurs, pêcheurs, plongeurs, associations, citoyens, plaisanciers...	Biodiversité / Mer
Qualifier	Evaluer la disponibilité de matières secondaires (recyclées)	Economie circulaire
Construire en partenariat	Construire des « communs » de l'information géographique	Information géographique
Assurer une veille	Détecter des algorithmes et codes sources dont le partage peut améliorer l'efficience de nouveaux services	Tous
<b>Comprendre et anticiper</b> pour prendre des décisions pertinentes		
Croiser	Analyser les effets des polluants urbains sur la santé des insectes et sur les interactions plantes-pollinisateurs	Santé-environnement
Modéliser	Prévoir la diffusion d'un événement météorologique	Risques
Simuler	Construire le « jumeau numérique » d'un quartier pour tester les impacts de nouveaux aménagements	Aménagement
Evaluer	Apprécier l'application de la nouvelle réglementation sur les drones afin d'en anticiper l'éventuelle adaptation	Aviation civile

Adapter son action pour améliorer l'efficacité et l'efficience d'une politique		
Informer	Mieux détecter les jeux de données afin de mieux appliquer la réglementation relative à l'information des citoyens en matière d'environnement	Environnement
Diffuser	Diffuser les données à jour relatives à l'emplacement des bornes de recharge de véhicules électriques	Mobilités
Contrôler	Vérifier que les incitations fiscales sont bien affectées à des travaux livrés par des entreprises qualifiées RGE	Bâtiment
Détecter	Détecter les fuites de méthane depuis des images satellitaires	Energie
Optimiser	Optimiser la stratégie de contrôle des ouvrages d'art afin de réduire les risques et d'optimiser le plan de charge des contrôleurs	Infrastructures

*Savoir* suppose de décrire précisément par les DONNEES ce qui passe sur le terrain.

- Mesurer depuis le terrain physique et humain,
- Agréger méthodiquement les mesures depuis plusieurs sources,
- Evaluer et certifier la qualité des données agrégées (ex : base de données).

*Comprendre* et anticiper suppose de décrire par la CONNAISSANCE ce qui devrait se passer

- Croiser : créer la connaissance par le croisement de données de sources hétérogènes, par exemple en s'aidant de l'intelligence artificielle (IA),
- Cibler : décider du meilleur scénario en fonction d'objectifs de politique publique,
- Modéliser : proposer une explication des causes et conséquences du phénomène mesuré, les représenter visuellement (*data visualisation*),
- Simuler : construire des scénarios d'action sur les causes (du modèle) et anticiper ses conséquences (attendues selon le modèle),

*Agir* suppose d'adapter son action en fonction de la compréhension du présent et d'identifier ce qu'il faut faire pour qu'à l'avenir, le réel se conforme à la décision

- Diffuser : compléter un processus par des données en temps réel pour qu'il fonctionne et/ou qu'il gagne en précision et prévisibilité,
- Contrôler : comparer les DONNEES actualisées à la CONNAISSANCE initiale,
- Détecter : identifier les écarts par rapport au modèle normatif, et identifier si c'est le réel, sa mesure ou la norme qui sont déviants,
- Corriger : lancer des actions correctives, soit dans la mise en œuvre de la politique publique, soit en enrichissant la connaissance et les modèles,
- Suivre/auditer : quantifier les écarts entre la politique publique et son exécution sur le terrain.

L'usage de la donnée à chaque étape, tant sur les aspects métier (i.e. les données de la transition écologique) que sur les fonctions support (finances, RH, etc.), nous permet de gagner en efficacité et en pertinence. Par la même, l'autorité, la souveraineté et la puissance de notre action publique en sont renforcées, y compris vis-à-vis du secteur concurrentiel et en partenariat avec lui.

## 2.2. Partager les données, les algorithmes et les codes sources pour renforcer l'impact écologique des politiques publiques

Les phénomènes environnementaux – climatiques - d'économie circulaire - énergétiques, sont fondés sur des processus complexes. Ils nécessitent le croisement de données variées et multiples. Les



politiques d'ouverture, de partage et d'usage des données constituent un levier pour prendre en compte et exploiter cette complexité, au service de la transition écologique.

Les données utiles à la transition écologique sont par nature **territorialisées**. A ce titre, la bonne gouvernance de la donnée géographique doit apporter un bénéfice conséquent pour constituer et valoriser le patchwork des données et travailler sur toutes les échelles de territoire. Pour cela, le conseil national de l'information géographique va voir son instance plénière de gouvernance relancée.

Les phénomènes environnementaux sont sujets à des effets en cascade ou de seuil. Un haut degré d'analyse est dès lors requis et repose sur la disponibilité de données diversifiées et précises. La mesure des impacts d'un aménagement sur la biodiversité, l'analyse du cycle de vie d'un produit ou l'évaluation des émissions de gaz à effet de serre d'un mode de transport sont rendus difficiles par l'insuffisance des données disponibles. L'usage des données utiles aux analyses de la transition écologique se prête donc par nature aux outils de modélisation, de datavisualisation et d'intelligence artificielle. Ces usages sont adaptés au caractère « écosystémique » des politiques publiques de la transition écologique qui impliquent et sont dépendantes d'une grande variété d'acteurs presque systématiquement.

L'urgence d'intervenir en cas de risques de dommages irréparables, mais aussi les perspectives de la ville intelligente ou du véhicule autonome, rendent nécessaire un accès plus systématique à certaines **données en temps réel** (ex. recueil et analyses automatiques de toutes sortes de captages, données satellitaires).

Les données utiles à la transition écologique sont par nature **éparses**. Il faut organiser et réussir le partage de jeux de données avec les collectivités territoriales, les opérateurs publics, les entreprises, les nombreuses associations qui œuvrent dans le domaine de la transition écologique, les communs numériques (ex. OpenStreetMap, Open Food Fact) et même les particuliers qui déjà participent à cet enrichissement (ex. Observation de la faune/flore).

**A titre illustratif, les données ci-dessous témoignent de cette hétérogénéité :**

Exemple de données	Cause de la fragmentation
Emissions de Gaz à Effet de Serre	De multiples sources d'émission : nationales (ex. bâtiment, transport, industrie, agriculture) voire internationales (ex. déforestation)
Effectif de population animale en déclin	Comptage humain et manuel à faire sur le tout territoire, en s'appuyant sur des bénévoles plus ou moins qualifiés, étude des habitats naturels pour combler l'absence de données de comptage, ...
Déchets ménagers	De multiples systèmes communaux de tri des ordures ménagères, et des dépôts sauvages par définition non comptabilisés

### 2.3. S'inscrire dans la démocratisation de la gestion de la donnée, des algorithmes et des codes sources pour une plus grande transparence de l'action publique et davantage d'innovation

L'accès à la donnée est l'un des piliers de la transparence démocratique.

Les sources de données prolifèrent, aussi bien dans le domaine privé que dans les administrations et chez les acteurs territoriaux. De même, les citoyens sont désormais dotés d'un accès constant à toute sorte d'informations. Ils ont ainsi développé une appétence accrue pour l'accès aux données et demandent plus de transparence sur l'action publique.

Au-delà de l'accès, le traitement de la donnée et sa visualisation (« *data viz* ») se démocratisent. Il ne s'agit plus d'une affaire exclusive d'experts. Les formations initiales en matière technique, économique et social intègrent désormais ces dimensions. De plus, les nouvelles technologies telles que l'intelligence artificielle (IA) accélèrent l'exploitation et l'interprétation des données.

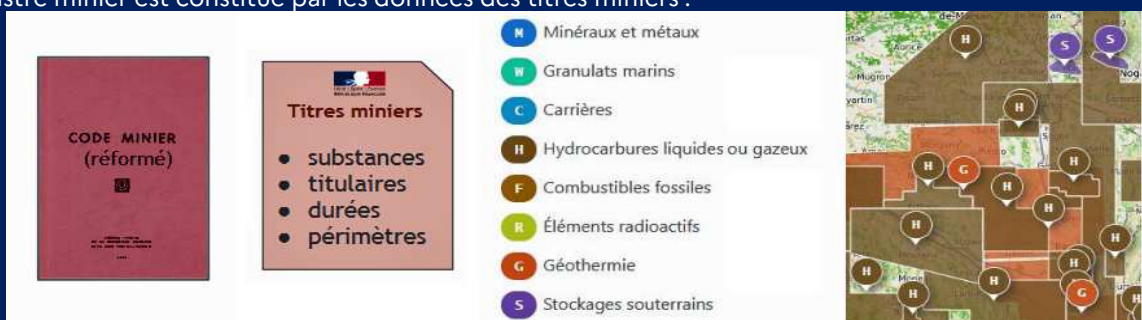
La production et le stockage de données se massifient à moindre coût. La donnée devient ainsi une commodité sur un marché international en voie de régulation, avec les ambitions de l'Union européenne sur le numérique, la donnée et l'intelligence artificielle.

Ces opportunités constituent un tremplin pour l'administration française pour accélérer sa transformation et son action et répondre à l'exigence de lisibilité et transparence de la société civile.

Cette transparence s'est aussi déjà affirmée comme un levier puissant pour accélérer l'innovation. L'accès à ces données crée des opportunités pour les chercheurs, les citoyens, les entrepreneurs innovants et les acteurs de la transformation publique. D'ores et déjà, de nombreuses start-up ont construit leur solution innovante grâce à l'accès et au partage de ces données. Il s'agit donc de poursuivre cet effort de transparence et d'accès aux données pour que celles-ci forment un terrain d'expérimentation fécond : à la fois producteur d'innovations publiques et privées, et outil clé pour les chercheurs.

**Exemple de produit numérique innovant répondant aux 3 enjeux : CAMINO, piloté par son impact en matière de politique publique avec des effets induits positifs sur les données.**  
**La propositions de valeur de CAMINO est d'ouvrir le cadastre minier pour mieux les gérer les projets.**  
**La valeur d'usage de ce service et sa stratégie de mise en marché a permis de démultiplier l'acquisition et la fiabilisation des données entraînant la croissance du partage et des usages.**

Le cadastre minier est constitué par les données des titres miniers :



2018 : de nombreux irritants pour l'administration des mines et les utilisateurs :

- Grandes difficultés à connaître l'état du domaine minier sur un territoire
- Déposer un dossier était un parcours du combattant et il était très difficile pour le demandeur et les citoyens de suivre une démarche d'instruction
- Des complications de gestion au sein de l'administration et des délais d'instruction trop longs
- Impossible ou presque pour un citoyen de connaître les projets miniers près de chez lui
- Accès aux données quasi impossible.

La réponse : CAMINO (« Cadastre minier ouvert »), un projet de la DGALN incubé par la Fabrique numérique du ministère.

Le projet est mené selon une démarche incrémentale centrée sur les utilisateurs où se nourrissent mutuellement de nouveaux services et l'amélioration des données (contenu et qualité). L'ouverture stratégique d'un premier lot de données incomplet et de qualité discutable a permis de créer rapidement un premier service numérique capable de capter la contribution naturelle des utilisateurs. L'augmentation des données disponibles et l'amélioration progressive de leur qualité permet de proposer des services toujours plus complets adoptés par l'ensemble des parties prenantes au-delà des attentes initiales. La connaissance de meilleure qualité a été étendue progressivement à l'ensemble du



territoire et partagée en tenant compte des prérogatives d'accès et exigence des confidentialité liées au secret des affaires.

Les données sont disponibles via une API sous différents formats ouverts et réutilisable avec des niveau d'accès régulés sur les données couvertes par le secret des affaires.

<https://docs.camino.beta.gouv.fr/> et <https://api.camino.beta.gouv.fr/>

2021 : CAMINO, c'est :



La Fabrique numérique met à disposition le code de tous les services numériques qu'elle réalise. Par exemple, tout le processus de l'application CAMINO est ouvert : l'ensemble des règles de collecte et de redistribution des taxes et redevance minières est donc à disposition de tous, sur la plateforme GitHub.

Au final, une source de vérité qui équilibre les débats et éclaire sur les impacts d'une politique publiques avec des statistiques ouvertes : <https://camino.beta.gouv.fr/statistiques/globales>



Les informations en temps réel et graphiques concernant la Guyane - sont consultable à <https://camino.beta.gouv.fr/statistiques/guyane>

## PARTIE 2

# Quels sont les objectifs de la feuille de route?

Pour répondre aux trois enjeux Exploiter, Partager et Démocratiser les données, algorithmes et codes sources, deux ambitions :

- **Moderniser l'action du pôle ministériel concernant la donnée** : avoir les données, les outils, les méthodes, les règles, les réflexes, les compétences et les ressources humaines pour produire et conserver la donnée, rendre la donnée mobilisable, consommer la donnée produite par d'autres et en percevoir tout le potentiel.
- **Accélérer la transition écologique grâce à la donnée** : permettre à tous les acteurs- Etat, collectivités, entreprises, citoyens - d'agir en faveur de la transition écologique en accédant aux données utiles et en les valorisant au travers d'usages pertinents.

### Accélérer la transition écologique grâce à la donnée : les GreenTech, un exemple d'usage des données au service de la transition écologique

Nam'R a développé des algorithmes et outils d'intelligence artificielle à partir de plusieurs sources de données (DPE – diagnostic de performance énergétique notamment) pour faire des propositions de rénovation énergétique des bâtiments dont la pertinence augmente avec les usages.

La start-up Ecofilae partage les données qu'elle a collectées dans le domaine des usages de l'eau en créant, en collaboration avec l'OFB et le BRGM, la plateforme HotspotReuse. Cette ouverture de données est destinée à faciliter le passage à l'acte en matière de réutilisation des eaux usées.

Pour relever le défi posé par ces deux ambitions, 6 objectifs de travail sont poursuivis. Deux leviers soutiendront les actions menées : une gouvernance (prévoyant entre autres la création d'une communauté des données, des algorithmes et des codes source) et un support technique :



## 1. Objectif : Rendre la donnée mobilisable

	RESPONSABLE	CONTRIBUTEUR	JALON PREVISIONNEL *
<b>ACTION PHARE : Augmenter la qualité des données en vue de leur réutilisation</b>			
>> Diffuser des principes directeurs communs au pôle ministériel en matière de qualité de la donnée à destination des producteurs de données et des porteurs de projet(s) numérique(s)	ECOLAB	SNUM ; Communauté de la donnée	CT (V0) MT (V1)
>> Intégrer la qualité de la donnée le plus tôt possible dans la gouvernance et la conception des projets et des politiques publiques afin de prévoir les conditions de son réemploi	PORTEUR PROJET		MT
>> Réaliser un inventaire des données de référence / données pivot pour alimenter entre autres le futur guichet d'accès à la donnée de la transition écologique et de la cohésion des territoires	ECOLAB	Communauté de la donnée	MT
>> Créer et déployer un label de qualité de la donnée de la transition écologique pour garantir à l'utilisateur le niveau de qualité atteint	ECOLAB		LT
>> Effectuer des diagnostics qualité (notamment sur les données de référence) et réaliser des plans d'actions de montée en qualité de la donnée	ECOLAB	Communauté de la donnée	MT
>> Signaler des défauts majeurs sur la donnée via un processus souple	ECOLAB	SNUM - Communauté de la donnée	CT
<b>ACTION PHARE : Promouvoir et amplifier l'APIsation de la donnée de la transition écologique</b>			
>> Elaborer un programme d'Apisation des nouveaux projets lancés pour amplifier l'utilisation des APIs au sein du pôle ministériel	SNUM		MT
>> Promouvoir l'offre API-management du SNUM	SNUM		CT
<b>ACTION PHARE : Mettre en place un guichet d'accès à la donnée de la transition écologique et de la cohésion des territoires</b>			
>> Cartographier, via une démarche incrémentale, les données et les référentiels pour alimenter au fil de l'eau le guichet	ECOLAB	SNUM	MT

	RESPONSABLE	CONTRIBUTEUR	JALON PREVISIONNEL *
>> Systématiser la traçabilité et la qualification des données dans le guichet (utilisation et affichage systématiques des métadonnées)	ECOLAB	SNUM	CT
>> Définir les processus de maintenance pour éviter l'obsolescence précoce du guichet	ECOLAB	SNUM	MT
<b>ACTION PHARE : Généraliser le partage des codes sources et algorithmes présentant un intérêt général en terme de connaissance et de transparence de l'action publique</b>	SNum		LT
Autre action			
Mettre en place des services de conseil « sur-mesure » pour accompagner les porteurs de projet dans la montée en qualité de leurs données et de les aider à appliquer les principes FAIR	ECOLAB	EXPERTS A IDENTIFIER	CT

CT : 2021-1er semestre 2022 ; MT : 2022-2023 ; LT : > 2023

\*Premières réalisations tangibles

### 1.1. Augmenter la qualité des données en vue de leur réutilisation

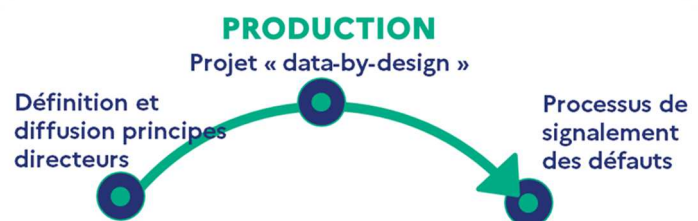
La **qualité intrinsèque** d'une donnée se juge, pour son producteur comme pour ses utilisateurs, autant à des critères absolus (tels que précision, exhaustivité, homogénéité, fraîcheur ou pérennité) qu'à l'adéquation aux **usages souhaités** de la donnée.

A titre illustratif, la précision de certaines sections du plan cadastral peut être jugée faible en soi pour alimenter un plan local d'urbanisme, mais elle demeure suffisante vis-à-vis de son principal usage : la perception des taxes.

La montée en qualité des données est assurée à trois étapes clés (cf. schéma ci-contre).

Afin de favoriser la réutilisation des données, il est important :

- de définir des principes directeurs de la qualité de la donnée
- d'envisager le futur usage des données dès leur conception (« data by design »)
- d'organiser un processus de signalement des défauts des données après leur diffusion.



L'enjeu est de **sensibiliser le concepteur** de projets numériques au respect de principes directeurs de qualité de la donnée tels que la standardisation, la précision sémantique et l'interopérabilité. La nécessité de définir et de partager des principes directeurs de qualité de la donnée dans l'ensemble du pôle ministériel a été largement mise en avant par les parties prenantes au cours de la concertation.



Concevoir des projets intégrant la donnée « by-design » envisage toutes les étapes d'un projet :

- prévoir des modalités d'ouverture ou de partage des données dès la préparation des textes législatifs ou réglementaires relatifs à une politique publique ;
- intégrer les exigences de qualité de la donnée dès le démarrage des projets produisant ou récoltant des données ouvertes ;
- organiser des séquences de revue de projet dédiées à la qualité de la donnée ;
- mettre à disposition des utilisateurs des modalités de signalement facile des défauts majeurs de la donnée et mettre en place les mesures correctrices dans un délai rapide ;
- faire dialoguer sur la durée producteurs et utilisateurs de la donnée.

Cela suppose des moyens humains et une organisation agile (capacité à ne pas être dépendant d'une ressource ou d'une expertise) pour maintenir dans le temps la qualité, gage de la réutilisation des données.

Ainsi, si le projet ambitionne in fine la constitution de jeux de données participant au « Once-only principle »<sup>3</sup> en cours de développement au sein de l'Union européenne ou de son équivalent français (« Dites-le nous une fois »), cela nécessite de prévoir dès l'amont les modalités de recueillement (dont le consentement éclairé des personnes au titre du RGPD) ainsi que les paramètres qualité de la donnée au regard des usages qui seront identifiés, y compris hors de la sphère du projet.

#### Constituer et promouvoir l'usage des données pivot (ou de référence)<sup>4</sup>

La mise en place de standards, de formats de stockage et de règles d'interopérabilité partagés entre utilisateurs facilite considérablement l'exploitation des données. Les **données « pivot » ou « de référence »** contribuent également au caractère interopérable des données (exemples : référentiels géographiques de l'IGN, dictionnaires sur l'eau du portail Sandre, limites terre-mer, géostandards), ce qui suppose de les identifier et de les promouvoir.

L'augmentation de la qualité et l'usage des données « pivot » ou « de référence » est l'affaire de tous. Les acteurs nationaux et territoriaux doivent converger dans cette direction, par exemple pour la production des Bases Adresse Locale et de la Base Adresse Nationale.

Lorsque la montée en qualité des données de référence sera assurée, elle pourra donner lieu à l'octroi d'un **label**. Le label fixe un objectif d'excellence à atteindre par le producteur de la donnée. En conférant une légitimité à des jeux de données, le label permet de communiquer autour des données de référence et de réduire les éventuels doublons. C'est aussi un gage de pérennité (mise à jour, qualité et diffusion) des services numériques développés à partir de ces données labellisées.

Ce label pourra concerner les jeux de données participant au Once-Only Principle (exposé ci-avant). Les travaux se baseront sur les travaux interministériels menés au sein du Ministère de la transformation et de la fonction publique par la DINUM pour créer un « label des services producteurs de données » conformément à la circulaire du Premier ministre du 27 avril 2021.

---

<sup>3</sup> <https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/Once+Only+Principle>

<sup>4</sup> Données pivot ou de référence = données qui soutiennent l'activité courante d'une organisation ; elles sont en général partagées entre plusieurs applications numériques de l'organisation et leur permettent de « communiquer » (même nom, même adresse, même code SIRET...). Cas emblématique de « données de référence » : le service public de la donnée créé par l'article 14 de la loi pour une République numérique vise à mettre à disposition, en vue de faciliter leur réutilisation, les jeux de données de référence qui présentent le plus fort impact économique et social ; ceux-ci sont actuellement au nombre de neuf (base adresses nationale, référentiel de l'organisation administrative de l'Etat, etc).

## 1.2. Promouvoir et amplifier l'APIsation de la donnée de la transition écologique

L'accès aux données est un irritant manifesté lors de la préparation de la feuille de route au sein du pôle ministériel. Plusieurs facteurs freinent l'accès aux données :

- une qualité insuffisante qui nécessite un travail avant exploitation ;
- un partage insuffisant ou des plateformes difficiles à identifier ;
- des problèmes techniques.

La feuille de route prévoit donc de rendre les données accessibles via une infrastructure prévue pour les partager et les télécharger grâce au recours à des API<sup>5</sup> et des webservices, avec des niveaux de services différenciés selon les besoins : fréquence régulière ou élevée et vitesse rapide pour les données dynamiques (ex. mobilité, météo, crue) ou au gré des rythmes de mise à jour pour les données plus statiques. L'archivage prévu dans l'infrastructure améliorera la conservation des données.

La conception d'API au sein des systèmes d'information nationaux dès leur création ("API First") devient la norme. Pour y parvenir, deux actions sont à engager : l'élaboration du programme et du calendrier d'Apisation des nouveaux projets de systèmes d'information du pôle ministériel ; la promotion de l'offre d'accompagnement des services en API-management du service du numérique (SNum). Le futur guichet d'accès à la donnée de la transition écologique et de la cohésion des territoires intégrera ces services d'API.

## 1.3. Mettre en place un guichet d'accès à la donnée de la transition écologique et de la cohésion des territoires

Trouver des données de la transition écologique et des territoires au sein de la multiplicité des portails existants et en préparation (internationaux, européens, nationaux, territoriaux, thématiques...) est un enjeu clé, qui doit en outre être adapté au public le plus large et le moins expert des domaines d'activité concernés.

La création d'un guichet d'accès à la donnée de la transition écologique et de la cohésion des territoires est donc l'une des actions phares de la feuille de route et a été retenue au printemps 2021 au guichet du plan de relance. Elle facilitera la recherche des données relatives aux politiques du pôle ministériel. La construction du guichet doit s'opérer en deux temps, ciblant d'abord les données produites par le pôle ministériel y compris ses opérateurs et ses services déconcentrés, puis celles des autres acteurs des politiques publiques de la transition écologique et des territoires (collectivités territoriales, associations, grand public...).

L'inventaire (cartographie) des données du pôle ministériel réalisé à l'occasion de la construction du guichet offrira un outil supplémentaire pour améliorer la cohérence des données entre elles et leur interopérabilité, et mieux gérer leur cycle de vie (collecte, traitement, documentation, stockage, conservation, partage). Il facilitera également la création d'Espaces Communs de Données (v. plus bas).

La préfiguration du projet de guichet d'accès à la donnée de la transition écologique et des territoires prévoit de s'appuyer sur les démarches existantes de catalogage des données afin de limiter le travail des équipes (ex : FEC Fichier des Entrées au Catalogue du CGDD/SDES, "WikiSI" de la DGALN, "SIdSI" de la DREAL Hauts-de-France, Géocatalogue du BRGM).

---

<sup>5</sup> API est un acronyme anglais qui signifie « Application Programming Interface », que l'on traduit par interface de programmation d'application (≈ dialogue de machine à machine, d'application à application).

Le guichet a vocation à être moissonné par data.gouv.fr et data.europa.eu. In fine cela a aussi vocation à améliorer la qualité des données de la transition écologique dans ces deux portails. Les API du guichet seront déclarées dans API.gouv.fr.

### Nettoyer les données et enrichir les métadonnées<sup>6</sup> associées pour les réutiliser

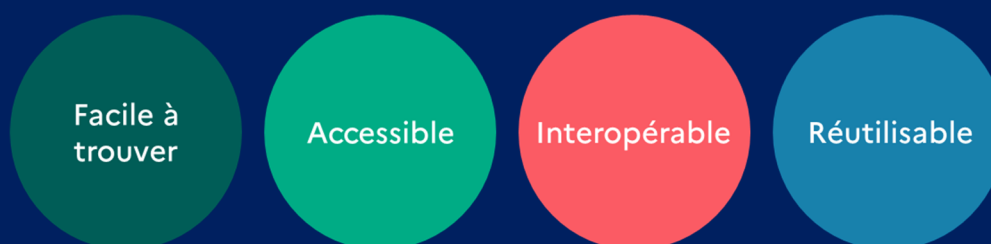
Les données doivent être « nettoyées » en vue de leur réutilisation :

- protection des données personnelles et des données non communicables par filtrage, agrégation, anonymisation, pseudonymisation ;
- accompagnement des jeux de données par des métadonnées qui placent les données dans leur contexte (service producteur, droits et licence assortis, date et fréquence de mise à jour, critères de qualité, mode d'accès, « traçabilité » des traitements qu'elles ont subis, etc.). Les métadonnées permettent de cataloguer les données, elles sont nécessaires pour apparaître dans le guichet d'accès décrit dans ce paragraphe.
- accompagnement éventuel par des informations de mise en perspective (ex : articles destinés au grand public).

#### Principes F.A.I.R.

Les principes exposés dans cet objectif se regroupent dans la notion de « Données FAIR » (ou FAIR Data) pour « Facile à trouver, Accessible, Interopérable, Réutilisable ».

La feuille de route vise à développer au sein du pôle ministériel la mise en œuvre de ces principes, déjà déployés largement notamment dans le monde de la recherche qui les a définis.



Les agents du pôle ministériel doivent pouvoir bénéficier d'un accompagnement pour les aider à évaluer la maturité de leurs données au regard des principes de qualité et d'accessibilité de la donnée. Les experts mobilisables assurant le rôle de consultants internes seront ainsi identifiés au sein du pôle ministériel.

#### 1.4. Généraliser le partage des codes sources et algorithmes présentant un intérêt général en termes de connaissance et de transparence de l'action publique

##### Ouverture des algorithmes et des codes sources

La circulaire du Premier ministre du 27 avril 2021 fixe un objectif concret : « les Administrations placées sous votre autorité devront rechercher en permanence la meilleure circulation de la donnée, des algorithmes et des codes, dans des formats exploitables par des tiers ».

<sup>6</sup> Métadonnée = donnée servant à caractériser / décrire les autres données

## Qu'entend-on par algorithmes et codes sources ?

Le rapport pour une politique publique de la donnée, porté par la Mission Bothorel, définit un code source comme « un ensemble d'instructions exécutables par un ordinateur. » Un algorithme est défini par la CNIL comme « la description d'une suite finie et non ambiguë d'étapes (ou d'instructions) permettant d'obtenir un résultat à partir d'éléments fournis en entrée » ; l'algorithme n'est pas nécessairement informatisé. Le rapport pour une politique publique de la donnée ajoute « De manière simplifiée, l'algorithme est une recette de cuisine, et le code sa réalisation concrète : ouvrir l'algorithme est un premier acte de transparence, mais rien ne garantit que la recette a été réellement suivie. ».

## Pourquoi ouvrir les algorithmes et codes sources ?

L'ouverture des algorithmes et des codes sources partage des enjeux communs avec l'ouverture des données au sens large. Concrètement, on retrouve deux principaux enjeux :

- **La transparence de l'action publique.** Outre le partage des données, partager les algorithmes et codes sources permet d'apporter davantage de lisibilité et de transparence sur des modes de calcul et l'application des règles. Par exemple, en exposant les règles de calcul de certaines taxes ou d'attribution des aides à la pierre.
- **L'amélioration de la puissance et de la sécurité des applications.** Partager les algorithmes et codes sources sur des plateformes équipées d'une communauté large d'utilisateurs actifs, c'est permettre à d'autres de contribuer aux applications de toutes les manières possibles et de collecter leurs retours. Cela se traduit par une meilleure qualité de code pour les applications et par une sécurité des données renforcée.

Au-delà des enjeux de fond, le partage répond aux obligations réglementaires posées entre autres par la Loi Lemaire de 2016. L'administration doit continuer à être précurseur en la matière.

## Quelles actions mettre en place pour poursuivre l'ouverture des algorithmes et codes sources ?

Le pôle ministériel se situe en avant-garde en matière d'ouverture de ses données, algorithmes et codes sources. La Fabrique numérique met à disposition le code de tous les services numériques qu'elle réalise. Par exemple, tout le processus de l'application CAMINO est ouvert : l'ensemble de la redistribution de la taxe sur l'or en Guyane est donc à disposition de tous, sur la plateforme [GitHub](#). Le service du numérique met à disposition des communautés les développements réalisés sur les logiciels libres utilisés (ex. : développement d'une suite d'outils améliorant la prise en main du logiciel PostgreSQL – système de gestion de bases de données - et de deux plug-ins associés au logiciel libre Q-Gis - système d'information géographique).

Pour poursuivre son action et dans le cadre de sa stratégie, le pôle ministériel se fixe les actions suivantes :

- **Porter spécifiquement le sujet de l'ouverture des algorithmes et des codes sources dans la gouvernance de la donnée** (cf. objectif 7 de la présente feuille de route) en nommant un correspondant en charge d'animer et de faire vivre cet objectif. Ce correspondant sera nommé par le comité stratégique de la donnée
- **Généraliser le partage, dès que cela représente un intérêt général, des algorithmes et codes sources des services et applications futurs.** Pour y parvenir, mettre en place un processus léger de vérification avant ouverture (ex. vérifier la qualité du code, la protection des données sensibles, l'intérêt général en matière de connaissance et de transparence de l'action publique). La publication sera assurée, lorsque la communauté d'utilisateurs sera présente, sur la plateforme interministérielle [code.gouv.fr](#).

## 2. Objectif : Développer la culture et l'usage de la donnée

	RESPONSABLE	CONTRIBUTEUR	JALON PREVISIONNEL *
<b>ACTION PHARE : Diffuser la culture et développer les compétences de la donnée</b>			
>> Construire et déployer le plan de sensibilisation et de formation aux sujets relatifs à la donnée pour assurer une montée en compétence collective du pôle ministériel sur la donnée à tous niveaux	DRH	ECOLAB - SNUM (expression de besoin)	CT
>> Rendre progressivement accessibles aux agents des outils de récolte, de nettoyage, d'analyse et de visualisation des données	SNUM	ECOLAB	MT
>> Définir et mettre en œuvre un plan de recrutement de compétences en « sciences des données »	DRH	ECOLAB - SNUM (expression de besoin)	CT
>> Concevoir et/ou diffuser les guides pour acculturer les praticiens métier à la donnée et favoriser les bonnes pratiques (principes FAIR, règles juridiques d'ouverture et d'archivage)	DRH	ECOLAB – SNUM - DAF	MT
<b>ACTION PHARE : Incuber des services répondant aux usages de la donnée et déployer la feuille de route de l'intelligence artificielle</b>			
>> Cartographier les usages par direction métier pour identifier les leviers de valeur de la donnée	ECOLAB	DG - SNUM	CT
>> Déployer la feuille de route de l'intelligence artificielle pour maximiser la valeur de l'usage des données	ECOLAB		Jalons FDR IA
>> Incuber des services répondant aux usages de la donnée	ECOLAB	SNUM (FABRIQUE NUMERIQUE)	CT
<b>Autres actions</b>			
Mettre en place un comité de pilotage du site Applisat pour contribuer à améliorer et enrichir son contenu au bénéfice de la communauté des utilisateurs du satellitaire	CGDD	CNES – CEREMA - IGN	CT

CT : 2021-1<sup>er</sup> semestre 2022 ; MT : 2022-2023 ; LT : > 2023

\*Premières réalisations tangibles

### 2.1. Faire prendre aux décideurs le virage de la donnée

Considérer la donnée comme un patrimoine commun à faire fructifier se fera à chaque niveau hiérarchique et avec l'appui décisif des dirigeants. Il est nécessaire de diffuser la valeur que peut générer la donnée et d'y allouer les moyens adéquats.

Cette acculturation est à mener dans l'ensemble du pôle ministériel, y compris au sein des territoires. Elle est réalisée au travers de séminaires de sensibilisation aux enjeux de la donnée.

Il est possible de s'insérer dans les outils de management déjà mis en œuvre auprès des décideurs pour intégrer la thématique de la donnée (ex. culture de la donnée, enjeux du partage, connaissances sur les enjeux de conformité et de sécurité, opportunités de financement, ingénierie financière). Par exemple, un plan de formation dédié à la transformation numérique et comportant un volet relatif à la donnée est en cours de construction au SNum et pourra servir de base aux travaux. En amont, cette acculturation peut être intégrée au sein des cycles de formation initiale des futurs cadres du pôle ministériel.

## 2.2. Faire monter en compétence les praticiens métiers (*makers*)

Au-delà des décideurs, il s'agit de rendre accessible aux agents du pôle ministériel des savoir-faire concrets pour en faire de futurs praticiens de la donnée (*makers*) par :

- **Des outils** : de récolte (*webscraping*), d'analyse de la donnée (*Dataiku*), de nettoyage de la donnée (anonymisation), de visualisation (*Dataviz*) ;
- **Des communautés d'utilisateurs**, des outils dédiés (*Osmose*) et des partages d'expérience entre administrations du pôle ministériel (voir au 7. Faire vivre la feuille de route de la donnée autour de la communauté de la donnée) ;
- **Des tutoriels et des guides** en vue de doter les agents d'un vocabulaire et de pratiques communs : principes FAIR, règles juridiques d'ouverture, de restitution graphique orientée utilisateur (*dataviz et UX design*), d'archivage ;
- **Des formations continues**, pour lesquelles une expression de besoins est à définir. A ce titre, une réflexion est à initier sur l'instauration d'une formation obligatoire à la sensibilité (secret) des données. Un catalogue de formation unique pourra être mis à disposition des acteurs du pôle ministériel.

### **Applisat : un site dédié à la Communauté du satellitaire**

Lancé en novembre 2019, le site internet Applisat offre un espace commun à l'ensemble des utilisateurs d'applications satellitaires et permet de rassembler les différents acteurs, favoriser l'appropriation des savoirs et partager les bonnes pratiques.

L'observation de la Terre par satellites représente en effet une source d'informations précieuse pour le suivi environnemental et pour l'élaboration et le suivi des politiques publiques en faveur de l'environnement. De plus en plus d'offres de services et d'outils sont disponibles. Pourtant, leur potentiel est encore mal connu et leur utilisation reste à développer. Il y a donc un réel enjeu à faire connaître les données, les produits et services disponibles ainsi que des cas d'usages des données satellitaires.

## 2.3. Se doter de ressources humaines et d'ingénierie financière

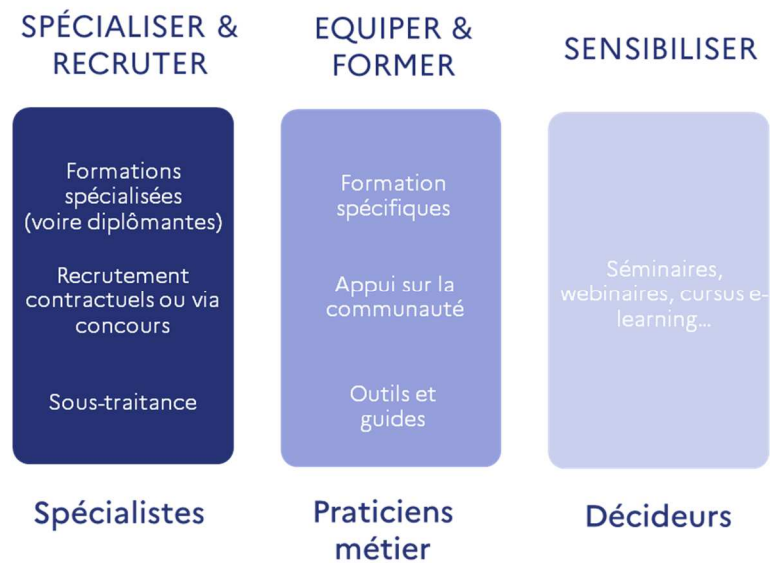
Pour lancer et généraliser des initiatives, des projets, le pôle ministériel doit pouvoir s'appuyer sur une ingénierie financière (montage de dossier de financement).

En ce qui concerne les ressources humaines, le pôle ministériel a besoin des experts de la gestion de données (*data handler*), des experts de l'analyse de données (*data scientist*), des experts de la restitution des données formés à l'UX design (*UX designers*). Ce sont l'ensemble des directions générales, des territoires et des opérateurs qui pourront renforcer leurs équipes en ce sens. Pour ce faire, plusieurs leviers sont à activer par ordre croissant de pérennité des compétences :

- Appel à un agent contractuel de l'Etat, à un sous-traitant
- Formation continue spécialisée (éventuellement diplômante) pour les agents de formation initiale scientifique, pour qui l'évolution vers les métiers de la *data* peut-être un souhait
- Recrutement par concours de *data scientist*, nécessitant de rénover les parcours de carrière et donner la possibilité à ces spécialistes d'exercer leur métier au long de leur carrière dans l'administration.



Schématiquement, les trois volets d'action peuvent être représentés de la manière suivante :



Sur ces trois buts, un travail d'identification des cibles et de documentation des besoins doit être mené en collaboration avec les acteurs du pôle RH en vue de la conception et de la mise en œuvre d'un plan de sensibilisation, d'acculturation, de formation et de recrutement.

Les formations aujourd'hui disponibles au sein du pôle ministériel autour de la donnée couvrent essentiellement son exploitation (ex. statistiques, géomatique).

#### L'INITIATIVE « DATA AU FEMININ » Comment féminiser les métiers de la data science

"Data au féminin" est un projet initié par le commissariat général au développement durable (CGDD), avec le soutien de la Haute fonctionnaire à l'égalité du secrétariat général, lauréat du fond pour l'égalité professionnelle de la DGAFP, soutenu par de nombreux partenaires publics et privés. L'objectif est de présenter des expériences, outils et rôles modèles aux agentes publiques pour leur permettre de prendre conscience de la place de la donnée dans leur quotidien personnel et professionnel et des parcours pour acquérir de nouvelles compétences dans ce domaine. Il inscrit son action dans les objectifs suivants :

- Anticiper l'émergence de nouveaux métiers
- Développer le management des talents : encourager les femmes à s'approprier les compétences numériques et favoriser leur montée en expertise
- Faire face à la pénurie de profils compétents en numérique et big data au sein de l'administration et favoriser la mixité et les transitions professionnelles
- Les métiers RH et les big data : sensibiliser les personnels et adapter les outils pour le traitement de données nombreuses, mouvantes et disparates
- Déconstruire les biais de genre.

#### 2.4. Innover pour entretenir la compétence « donnée » dans le temps et fidéliser les ressources

La donnée est un domaine en pleine évolution, porté par des évolutions technologiques fréquentes. La compétence doit être constamment entretenue pour rester pertinente. Par ailleurs, la fidélisation des ressources expertes en données est un défi pour l'administration. Outre la rénovation des parcours de carrière rénovés, évoqués dans le paragraphe précédent, maintenir le niveau de compétence des agents au plus proche de la transformation numérique est un prérequis. Ainsi, pour

favoriser des apprentissages en continu, le recours à des modules de formation plus « innovants » est souhaitable. Ils seront par ailleurs soutenus par des guides et outils dotés de format facilement appropriables et évolutifs (ex. accessibilité via un site web, formats synthétiques).

#### Formats illustratifs de modules « innovants »

- MOOC « Massive Open Online Course » : des modules de cours à distance pouvant accueillir de nombreux participants ;
- D'autres formats de e-learning : Serious Game ; micro-learning (contenus très courts diffusés au sein de la communauté) ;
- Des cursus de formation fédérant des communautés, type « Data Academy », proposant notamment des exercices de mise en situation
- Des modules de formation délivrés par un expert favorisant le retour d'expérience, type « Master class Data »
- Des modules de formation réalisés par des ambassadeurs, qui répliquent leur propre formation, type «Trainer-Trainee » ou « mentorat inversé »
- Des échanges de bonnes pratiques entre directions, de type « vis-ma-vie »

Pour y parvenir, le pôle ministériel pourra à la fois s'appuyer sur ses propres bonnes pratiques concernant d'autres domaines de compétence, sur les bonnes pratiques de formation et de fidélisation des compétences des autres ministères et d'autres acteurs publics et privés, et sur des modules de formation existant au sein d'autres structures ou plateformes.

#### 2.5. Mobiliser des méthodes pragmatiques pour montrer la valeur de la donnée

La culture de la donnée se diffuse également par la preuve. La mise en place de projets concrets, réalisés au travers de **méthodes pragmatiques** (ex. cycles itératifs courts de prototypage / tests) permet de partager des **résultats tangibles et rapides** aux métiers et aux usagers, internes ou externes au pôle ministériel.

Une approche croisée métier / spécialistes de la donnée produit des résultats : identification de cas métier à prototyper par les métiers eux-mêmes ou à partir de l'analyse des données.

A titre illustratif, la Direction générale de l'Aménagement, du Logement et de la Nature (DGALN) a déjà engagé une démarche de « **chantiers dynamiques** » en se positionnant aux côtés des métiers. La démarche les fait collaborer aux travaux pour identifier des offres pertinentes et pour s'appuyer sur des cas d'usage.

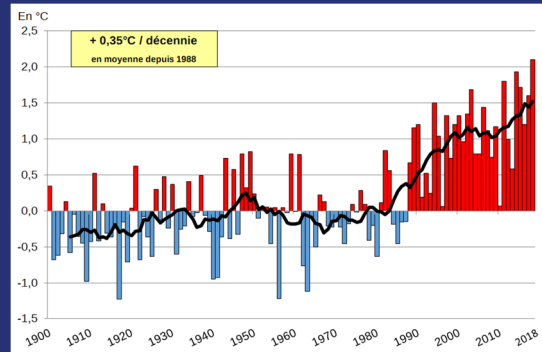
On retrouve trois principales familles de cas d'usage.

- **Des services internes** répondant à un besoin de pilotage de l'efficacité et de l'efficience des politiques publiques ;
- **Des services innovants tournés vers les politiques publiques** de la Transition écologique à fort impact ;
- **Des services aux usagers**, dans le cadre des obligations réglementaires ou non, en matière de dématérialisation ou favorisant la nécessaire transparence de l'action publique via la *data viz*. Ex. La start-up d'Etat Wikicarbone développe un outil qui informe les consommateurs sur l'impact environnemental des produits et des services. Les consommateurs contribuent à l'alimentation de l'outil en données

**Les services à concevoir pourront être soutenus par les technologies d'intelligence artificielle. Cette thématique est traitée de manière complète dans la *Feuille de route de l'intelligence artificielle* du pôle ministériel, une des actions phares de cette feuille de route.**

## LA DATAVIZ, A QUOI ÇA SERT ?

La visualisation de la donnée (Dataviz) constitue un outil puissant pour mettre en valeur les données, notamment à des fins de pilotage et de transparence. Elle facilite la découverte, l'exploration et l'analyse des données, la « recherche de sens » (e.g filtre, comparaisons, rapprochements) et la communication de l'information dont elle est l'un des principaux vecteurs. Le "Data journalism", journalisme basé sur l'exploitation de données, utilise naturellement et fréquemment la Dataviz.



Évolution des températures moyennes annuelles en France métropolitaine (site : notre-environnement.gouv.fr, source Météo-France)

Afin d'entretenir la dynamique autour des usages de la donnée, un **processus d'incubation** vise à :

- **Capter et prioriser les cas d'usage** en mettant en place les conditions nécessaires (grille de notation ; mise en place d'un processus et de l'instance dédiés) ;
- **Incuber les cas d'usage** priorités pour accompagner les porteurs de projet dans la création de services autour de la donnée ;
- **Accompagner les porteurs de projet pour passer à l'échelle** (conception sécurisée, organisation cible et ressources qui soutiennent le service identifiées et mobilisables).

En amont, les métiers peuvent cartographier leurs usages pour identifier les leviers de valeur qui se dégagent du traitement de leurs données.

### 2.6. Poursuivre les actions des start-up d'Etat

Dans le cadre de l'intrapreneuriat public, les start-up d'Etat permettent à un agent public de travailler avec une petite équipe à la conception d'un service numérique, généralement piloté par l'usage et offrant une solution à un irritant de politique publique. Les données sont souvent au cœur de ces services numériques.

Au sein du pôle ministériel, la Fabrique numérique du Service du numérique (SNum), qui fait partie de la communauté beta.gouv pilotée par le département Etalab de la direction interministérielle du numérique, est le lieu d'incubation de services numériques innovants de la transition écologique sur le modèle des start-up d'Etat. Par exemple, [trackdéchets](#) vise à apporter une solution numérique au fonctionnement parfois défectueux du suivi par des bordereaux physiques de déchets dangereux (12 millions de tonnes par an) et crée de fait les bases de données nécessaires.

### 3. Objectif : Partager les données entre administrations

	RESPONSABLE	CONTRIBUTEUR	JALON PREVISIONNEL *
<b>ACTION PHARE : Poursuivre l'ouverture des données, des algorithmes et des codes sources</b>			
>> Créer les « rendez-vous de la donnée publique de la transition écologique et des territoires » proposant masterclass et ateliers autour du partage des données entre administrations pour créer une dynamique commune	ECOLAB		MT
>> Mettre en place des schémas territoriaux de l'usage de la donnée notamment pour faciliter les arbitrages et la planification en matière d'ouverture de données	ECOLAB	DREAL	CT
>> Tenir à jour et diffuser le tableau de bord de l'ouverture des données aux acteurs du pôle ministériel	ECOLAB	POLE MINISTRERIEL	CT
<b>ACTION PHARE : Concevoir un dispositif national de la connaissance pour la transition écologique, à partir des administrations territoriales et centrales</b>			
	ECOLAB		MT
Autres actions			
Renforcer l'animation des plateformes régionales et des communautés de logiciels libres	CGDD	SNUM - DREAL	MT

CT : 2021-1<sup>er</sup> semestre 2022 ; MT : 2022-2023 ; LT : > 2023

\*Premières réalisations tangibles

Le principe du partage des données entre administrations est inscrit dans la loi (article 113-13 du code des relations entre le public et l'administration [CRPA]). Ce principe a vocation non seulement à aider les différents services publics à étayer leurs politiques avec les données nécessaires, mais aussi, de plus en plus, à éviter de demander aux administrés des documents déjà détenus par une autre administration (principe « Dites-le nous une fois »).

En outre, mieux partager la donnée de la transition écologique et des territoires permet de mieux diffuser les enjeux des politiques publiques qui en sont les garantes.

Or, dans les faits, l'hétérogénéité des pratiques, la réticence au partage de données sensibles ou complexes, la crainte de diffuser des données erronées, voire l'existence de redevances de mise à disposition freinent ce partage.

La feuille de route de la donnée fixe pour objectif de faciliter ces échanges :

- entre les niveaux national et territorial du pôle ministériel, en comprenant les opérateurs ;
- au niveau interministériel.

#### 3.1. Animer le partage de données entre les administrations

Avec :

- la création des « rendez-vous de la donnée publique de la transition écologique et des territoires » pour créer une dynamique (masterclass de partage de la donnée publique, ateliers / openlab thématiques, etc.)
- la poursuite d'actions d'innovation autour de la donnée, avec le cas échéant l'appui de la direction interministérielle de la transformation publique.
- la mise en place de la communauté de la donnée, des algorithmes et des codes sources (v. objectif 7)
- la mise en place d'un accompagnement juridique permettant de lever les freins liés à la complexité de la réglementation de l'ouverture, du partage et de la protection des données (v. objectif 6).

### 3.2. Réussir les initiatives de coopération entre administrations

#### Mettre en place l'Espace Commun de Données Santé-Environnement

En cohérence avec la stratégie européenne de la donnée, plusieurs Espaces Communs de Données (ECD) sont en cours de définition ou de préfiguration au sein du pôle ministériel.

L'initiative Green Data for Health (GD4H) portée par l'Ecolab du commissariat général au développement durable en application du Plan national Santé – Environnement 4, premier projet d'ECD au sein du pôle ministériel, vise à créer un Espace Commun de Données Santé-Environnement qui :

- permette de croiser les données pertinentes et disponibles pour caractériser les expositions aux facteurs environnementaux susceptibles d'affecter la santé,
- facilite l'accès aux données environnementales, y compris au niveau des territoires,
- stimule un appariement rigoureux des données environnementales avec les données de santé

Le GD4H est dans un premier temps développé à l'usage des chercheurs et donc notamment de la sphère publique.

La mise en place des premiers Espaces Communs de Données fait l'objet d'une action phare décrite dans l'objectif 4.

#### Concevoir un dispositif national de la connaissance pour la transition écologique

Les missions « connaissance » des administrations déconcentrées de l'Etat ont un rôle prépondérant dans la production et l'analyse de la donnée. Leurs analyses territoriales sont un outil de premier plan pour la définition, la mise en œuvre et l'évaluation des politiques publiques au service des citoyens et de la transition écologique.

A la suite du rapport d'inspection « Optimisation des missions de connaissance » du CGEDD (2021), l'objectif est de capitaliser sur les atouts des missions « connaissance » pour remettre au premier plan la donnée territoriale, rendre plus impactante l'action publique et améliorer le service rendu aux citoyens.

#### Renforcer l'animation des plateformes régionales et des communautés de logiciels libres

L'Ecolab du Commissariat général au développement durable anime les plateformes territoriales de la donnée géographique et à ce titre renforce son rôle de coordination de la communauté du logiciel libre Prodiges. Ce logiciel, initialisé en Région Rhône-Alpes, est utilisé aujourd'hui par plus de la moitié des plateformes territoriales, mais aussi par le ministère chargé de la santé pour sa plateforme atlaSanté. Prodiges constitue *de facto* un commun numérique interministériel partagé également avec des collectivités.

## PRODIGE : un commun numérique ?

Prodige est une solution Web open source destinée à un grand nombre d'acteurs, et notamment aux acteurs territoriaux. Il valorise et diffuse les données d'un territoire au bénéfice des acteurs du territoire et du grand public, et de respecter les obligations de la directive INSPIRE.

Prodige dispose d'un générateur d'applications métier permettant de répondre très rapidement à de nouveaux besoins tout en mutualisant les ressources. Par exemple, le ministère de la santé l'a utilisé pour lister et localiser les centres de vaccination pour diffusion au grand public sur le site dédié.

Le SNum organise une animation de communauté de logiciel libre autour de QGIS et PostGIS qui réunit les acteurs de la géomatique au sein des directions générales, services déconcentrés et notamment les DDT(M) mais aussi au-delà car il y a des membres issus des établissements publics, de collectivités territoriales et d'organismes publics.

Elle donne principalement lieu à 2 forums sous forme de liste à abonnement (QGIS, PostGIS), à des interventions dans les réunions régulières du domaine et à des contributions sous forme de plugins QGIS ou d'extensions PostgreSQL dédiées aux métiers du pôle ministériel, déposés sur le site de dépôt du ministères et/ou dans le GitHub du pôle ministériel.

Ces actions contribuent à l'animation et à la promotion interministérielles en matière de logiciel libre et de communs numériques demandées par la circulaire du Premier ministre du 27 avril 2021.

### 3.3. Poursuivre l'ouverture des données publiques, des algorithmes et des codes sources, en priorisant les usages

#### Poursuivre l'ouverture et les échanges de données entre administrations

Le pôle ministériel poursuit sa politique d'ouverture des jeux de données et référentiels clés en open data ou en accès restreint. Le comité interministériel de la transformation publique du 5 février 2021 lui a prescrit l'ouverture prochaine de 14 nouveaux jeux de données, API et codes sources. A titre illustratif, le site [transport.data.gouv.fr](http://transport.data.gouv.fr) rend accessible les données de mobilité aux autres administrations et ré-utilisateurs de données.

[transport.data.gouv.fr](http://transport.data.gouv.fr) est le Point d'Accès National français aux données de mobilité. C'est une plateforme publique destinée à rassembler toutes les données de mobilité utiles aux développeurs et éditeurs d'application qui facilitent le calcul d'itinéraire, au bénéfice des voyageurs.

#### Lancer une réflexion sur la mise en place de schémas territoriaux de l'usage de la donnée

Pour maximiser la réutilisation des données ouvertes, des schémas territoriaux de l'usage de la donnée permettront à l'échelon local de prioriser plus facilement les données à ouvrir et le calendrier associé. Par exemple, la DREAL Bretagne s'est dotée d'une feuille de route de la donnée sur 10 politiques publiques prioritaires.

#### Diffuser le tableau de bord « ouverture des données de la transition écologique »

Pour mieux accompagner les acteurs du pôle ministériel dans leurs choix, le tableau de bord « ouverture des données de la transition écologique », tenu par le CGDD au titre de son rôle d'administrateur de la donnée, sera diffusé régulièrement.

Le tableau de bord mentionnera notamment les bases de données dont la réutilisation deviendra gratuite en application de la circulaire du 1er ministre "relative à la politique publique de la donnée, des algorithmes et des codes sources" du 27/4/2018.



## S'appuyer sur l'existant pour la poursuite de la diffusion des codes sources

Que ce soit les travaux sur la suite Asgard (traitement de jeux de données volumineux et réalisation d'analyse spatiale complexes) ou bien les productions de la Fabrique numérique (Camino, Trackdéchet, Zero logement vacant, ..), de nombreux codes sources du pôle ministériel sont déjà conçus dans une logique de partage de la connaissance et de transparence. Ils sont d'ores et déjà accessible sur les plateformes dédiées (Github). Ces partages vont se poursuivre et permettre, en s'appuyant sur ces expériences, un élargissement de cette politique.

## 4. Objectif : Partager des données par la coopération public-privé

	RESPONSABLE / Contributeur	CONTRIBUTEUR	JALON PREVISIONNEL *
<b>ACTION PHARE :</b> Mettre en place les premiers Espaces Communs de Données en matière de Transition écologique	ECOLAB	AUTRES ENTITES DU MINISTERE	<b>MT</b>
<b>Autres actions</b>			
Intensifier les actions de sensibilisation et de promotion du partage de données (ex : évènement Données d'intérêt général, Hackathons)	ECOLAB	OPERATEURS	<b>CT</b>
Mettre en place des correspondants au sein du pôle ministériel et un point d'accès pour faciliter le recueil et l'instruction des demandes provenant du secteur privé et de la recherche pour l'accès aux données	COMITE STRATEGIQUE		<b>CT</b>
Organiser une équipe d'appui pour appuyer les négociations d'accès aux données d'intérêt général	ECOLAB		<b>CT</b>
S'appuyer sur la communauté pour partager les expériences d'accès aux données privées au sein du pôle ministériel	ECOLAB		<b>CT</b>

**CT** : 2021-1<sup>er</sup> semestre 2022 ; **MT** : 2022-2023 ; **LT** : > 2023

\*Premières réalisations tangibles

De nombreux acteurs privés voient un intérêt à accéder aux données publiques de la transition écologique (ex. données géographiques de l'IGN, données partagées par Météo France, données de biodiversité, sur l'eau). Cet intérêt est partagé par la puissance publique. Car l'utilisation des données par le secteur privé constitue un vecteur de valeur pour les citoyens ou la collectivité dans son ensemble et améliorer l'efficacité de l'action publique.

### *Du secteur public vers le privé*

Une meilleure circulation des données du secteur public vers le secteur privé bénéficiera à chacune des parties :

- Pour les *acteurs du secteur privé et la société civile* : Le Géoportail de l'urbanisme, par exemple, agrège des données issues de documents d'urbanisme différents pour déployer des services d'aide à l'aménagement du territoire. De même, la mise à disposition des données relatives aux zonages d'évitement (base GéoMCE) permet de limiter les atteintes à la

biodiversité en informant les maîtres d’ouvrage et les bureaux d’études des zones porteuses de mesures compensatoires.

- Pour les *producteurs publics* : leurs données seront de plus en plus utilisées, et en retour les utilisateurs pourront ainsi contribuer à enrichir la donnée publique (*crowdsourcing*), qui maintiendra son avantage qualitatif (niveau d’actualisation, de précision) et avec un coût maîtrisé

### ***Du secteur privé vers le public : les données d’intérêt général***

Inversement, de nombreuses données sont récoltées, produites ou détenues par le secteur privé : entreprises, associations, citoyens, etc. Pour certains domaines d’action du pôle ministériel, ces données privées sont essentielles, l’acteur privé occupant une position dans la chaîne de production de la donnée qui le rend incontournable.

Voici pour illustration quelques données d’intérêt public détenues par des acteurs du secteur privé :

Données d’intérêt public	Producteur privé	Position stratégique
Forages	Entreprises pétrolières	Exclusivité à la source
Compteurs Linky	Opérateur du réseau de distribution électrique	Exclusivité à la source
Agriculture numérique	Fabricants de matériel	Exclusivité à la source
Mobilité	Entreprises de transport des personnes, constructeurs automobiles, services d’aide à la circulation, services de location de vélos	Exclusivité à la source

Certains développements anticipent déjà une gestion et mise à disposition sécurisée de données entre tous types d’acteurs, comme par exemple l’application GéoDAE, réalisée avec le logiciel libre Prodigé décrit dans l’objectif 3 et dont la communauté est animée par le pôle ministériel. Cette application a permis de répondre à la loi du 28/06/2018 relative au défibrillateur cardiaque, et de créer en crowdsourcing une base de données recensant la localisation des défibrillateurs automatisés externes et leur accessibilité (partenaires : Direction générale de la Santé, ARLoD, propriétaires de défibrillateurs, services de secours, applications web citoyenne, sociétés de maintenance des défibrillateurs, fédération française de cardiologie).

La Feuille de route de la donnée concoure à positionner le Pôle Ministériel en plaque-tournante (*Hub*) des données pour la transition écologique.

#### **4.1. Promouvoir le partage de données entre le secteur public et le secteur privé**

Le pôle ministériel organise différents types d’actions de sensibilisation et de promotion du partage de données associant acteurs privés et publics. La promotion du partage de la Donnée capitalise sur la démarche d’innovation ouverte GreenTech, avec par exemple :

- Des actions d’innovation et d’idéation communes autour de l’usage de la donnée au service de la Transition écologique. A titre illustratif, un Hackathon RenovAction a été mené en 2020 autour de la rénovation énergétique ;
- Des appels à manifestation d’intérêt portant par exemple sur les Start up Green tech en 2020 ou sur les Communs de données / numériques (à venir).

Ces actions ont vocation à s'intensifier dans les prochaines années. En 2021, un évènement « *Données d'intérêt Général* » sera ainsi organisé.

De son côté, l'IGN s'investit dans un projet de « géo-communs » en lien avec son projet de Géoplateforme décrit dans l'objectif 8.

## 4.2. Réussir des premières initiatives concrètes, notamment les Espaces Communs de Données (ECD)

### Mettre en place les premiers Espaces Communs de Données

En cohérence avec l'ambition européenne en matière de stratégie de données pour les domaines d'intérêt public, les premiers Espaces Communs de Données seront mis en place pour faciliter l'accès, la mobilisation voire la production et la montée en qualité des données ayant trait aux politiques publiques du ministère. Le premier Espace commun de donnée à déployer est celui dédié à Santé environnement, ou Green Data for Health (décrit dans l'objectif 3) qui est aussi une action majeure du quatrième plan national santé environnement (PNSE4). D'autres espaces communs de données sont à préfigurer sur des secteurs à définir parmi lesquels l'Économie Circulaire (en lien avec le MEFR/DGE et le comité stratégique de filière), la Prévention des risques naturels et la Rénovation énergétique des bâtiments. Ils ont vocation à associer des acteurs du pôle ministériel (centrales et territoires) et des acteurs externes, parfois publics (ex : la communauté des chercheurs en santé environnement) et souvent privés. Ils devront prendre la forme de dispositifs adaptés aux enjeux métiers sectoriels en termes d'outil, d'acteurs et de gouvernance.

Ces initiatives permettent d'animer – par une structuration sectorielle de la donnée orientée vers les usages - des communautés d'utilisateurs et de producteurs de la donnée intégrant des acteurs privés. Une doctrine est à construire qui aidera aussi à porter auprès de la Commission Européenne la vision française des principes qui doivent présider à la mise en œuvre de ces espaces dans le champ de la transition écologique. Le pôle ministériel pourra s'appuyer sur ces premières initiatives variées pour faire vivre la donnée publique et privée, l'enrichir, et favoriser ses usages.

#### Qu'est-ce qu'un Espace Commun de données ?

En février 2020, la Commission Européenne a présenté sa stratégie pour les données dans laquelle elle ambitionne de créer un marché unique européen de la donnée. Elle introduit la notion d'espace commun de données. Ils ont pour vocation de mettre en commun sectoriellement les données européennes, publiques et privées (en particulier les données industrielles), inter- et intra-Etats Membres, et avec un objectif d'interopérabilité. Le pacte vert, la mobilité, la santé, l'énergie font par exemple partie des grandes thématiques identifiées.

Des premières précisions ont été apportées à la notion d'Espace commun de données en novembre 2020, avec la parution d'un projet de règlement sur la gouvernance des données : la Commission envisage un futur où les acteurs du domaine privé ou public entreposeraient leurs données chez un prestataire de partage de données, qui « organiserait la neutralité des données pour augmenter la confiance », et jouerait un rôle d'intermédiaire de confiance.

### Produire et entretenir en commun des données dans le cadre de la Géoplateforme

La Géoplateforme est décrite dans le paragraphe [Réussir la Géoplateforme : l'outil collaboratif de production et d'entretien de la donnée](#) de l'objectif 8.

### 4.3. Faciliter l'accès aux données publiques de la transition écologique via un guichet d'accès et des correspondants clairement identifiés

Comme détaillé dans l'objectif 1, le pôle ministériel met en place un guichet d'accès à la donnée de transition écologique et de la cohésion des territoires. Grâce à ce service, les acteurs du secteur privé et de la société civile accéderont également plus facilement aux données.

Le pôle ministériel poursuit sa politique d'ouverture des jeux de données et référentiels clé en *open data* ou en accès restreint. A titre illustratif, en 2021, de nombreuses données seront rendues accessibles comme des données sur les sites et sols pollués, sur l'artificialisation des sols ou encore les données 3D issues des campagnes de LIDAR de l'IGN.

Pour faciliter le recueil et l'instruction des demandes d'accès aux données provenant du secteur privé (entreprises ou société civile) et de la recherche, des correspondants au sein du pôle ministériel et un point d'accès (Ecolab) seront mis en place. Cette organisation s'articule avec la promotion de l'accès aux données d'intérêt général et la désignation d'un médiateur national dédié.

### 4.4. Faciliter l'accès aux données privées d'intérêt général pour la Transition écologique

#### Organiser une équipe d'appui aux négociations d'accès aux données d'intérêt général

Au-delà de l'organisation mise en place par le pôle ministériel pour faciliter le dialogue avec les acteurs privés (point d'accès pour les demandes et correspondants internes), une équipe d'appui sera mise en place. Elle mènera les négociations d'accès aux données d'intérêt général aux côtés des directions et opérateurs, auprès des acteurs ayant des missions de service public ou produisant des données environnementales. Cela s'ajoute à l'accompagnement juridique proposé pour faciliter le partage de données (cf. 6.2 Offrir un accompagnement juridique centralisé).

#### Capitaliser sur les premières initiatives de partage de données

Le pôle ministériel tirera profit des premiers projets concrets développés avec le secteur privé – en particulier les Espaces Communs de Données – pour définir des lignes directrices en matière d'accès, d'usage et de réemploi des données d'intérêt général. Ces lignes directrices ont vocation à servir de guide et enseignements pour les projets lancés par les différentes entités internes mais également à faciliter le partage de données par le secteur privé. En parallèle, l'échelon interministériel (ex : CNIL) lance une réflexion sur la clarification du cadre juridique lié au partage des données d'intérêt général.

#### Partager les expériences d'accès aux données privées au sein du pôle ministériel

Les directions et opérateurs du ministère ont différentes expériences de partenariats public-privé. La Communauté servira à partager des retours d'expérience sur ces projets concrets mais également d'avoir des échanges sur les réflexions et les résultats en matière d'accès à des données d'intérêt général auprès du secteur privé, sur le bilan et les bonnes pratiques en matière d'acquisition de données (ex. optimisation des règles, licences, modèle économique).

## 5. Objectif : Positionner la France dans la stratégie de donnée de l'Union européenne

	RESPONSABLE	CONTRIBUTEUR	JALON PREVISIONNEL *
Identifier les pays actifs et précurseurs en matière de politique de la donnée de la transition écologique et mettre en place un club avec eux pour nourrir en continu les travaux sur la donnée	ECOLAB		MT
Etablir une relation renforcée avec la commission européenne en intégrant les instances de travail des textes européens relatifs à la donnée afin de démontrer l'identité de la France en la matière	A IDENTIFIER – AUX COTES DU DINUM		MT

CT : 2021-1<sup>er</sup> semestre 2022 ; MT : 2022-2023 ; LT : > 2023

\*Premières réalisations tangibles

En février 2020, la Commission européenne a publié une stratégie européenne pour les données ayant pour but de créer un marché unique de la donnée : « *La Commission entend soutenir l'élaboration de systèmes technologiques et d'infrastructures de nouvelle génération, qui permettront à l'UE et à tous les acteurs d'exploiter le potentiel de l'économie des données. Elle cofinancera des projets européens à forte incidence, portant sur des espaces européens des données et des infrastructures en nuage économes en énergie et fiables. Elle lancera des mesures sectorielles spécifiques pour le déploiement d'espaces européens communs des données* ».

Cela s'est traduit par une proposition de règlement sur la gouvernance européenne des données (**Data Governance Act**) en novembre 2020 qui «  *vise à favoriser la disponibilité de données en vue de leur utilisation, en augmentant la confiance dans les intermédiaires de données et en renforçant les mécanismes de partage de données dans l'ensemble de l'UE* », elle définit un cadre pour la mise à disposition de données du secteur public soumises à droits d'autrui, pour le partage de données entre entreprise contre rémunération, pour l'intermédiation de partage de données à caractère personnel et enfin pour permettre l'utilisation de données pour des motifs altruistes.

En 2021, la commission poursuit ses travaux en vue d'un acte législatif : « loi sur les données » (Data Act). Elle «  *a pour objectif de proposer des mesures visant à créer une économie des données juste en garantissant l'accès aux données et leur utilisation, notamment dans des situations interentreprises et entre les entreprises et les administrations publiques.* » Elle pourra contenir une révision de la directive 96/9/CE concernant la protection juridique des bases de données afin qu'elle reste pertinente pour l'économie des données.

Les actions déclenchées dans le cadre de cette feuille de route seront nourries par les évolutions européennes. La vocation évolutive du document s'adaptera aux avancées avec une volonté de participation, d'anticipation et de relai :

### 5.1. Mettre en place un dispositif de coopération européenne

Pour poser les bases de son action dans le cadre des instances européennes, un premier travail de veille et d'analyse est à engager pour suivre l'actualité de la donnée au sein des instances communautaires mais également au sein des pays membres précurseurs en la matière de données au service de la transition écologique. Ce travail de fond permettra d'identifier des partenaires européens éligibles à la construction d'une coopération Cette coopération peut prendre la forme d'un club où les pays enrichissent leurs pratiques, identifient des axes de développement et trouvent de nouveaux relais de croissance communs. Ce club pourrait se rencontrer une à deux fois par an.

La France devant présider le Conseil de l'Union européenne au premier semestre 2022, ce calendrier constitue pour elle une étape utile pour lancer de telles initiatives et coopérations.

## 5.2. Faire de la France une force de proposition dans les négociations

Dans le cadre de son activité européenne, le pôle ministériel souhaite faire de la France une force de proposition dans les négociations. D'une manière générale, l'objectif est de positionner le pôle ministériel en amont de la production législative et réglementaire européenne pour partager ses positions et démontrer son identité. Le pôle ministériel, et notamment le BRGM (Bureau de recherches géologiques et minières) et l'IGN, a participé activement aux travaux relatifs à la directive INSPIRE. L'ensemble des travaux successifs à la publication de la *stratégie européenne pour les données* et des projets de textes qui la mettent en œuvre (projets de *data governance act*, de *data act*, ...) constituent pour la France une opportunité de pousser sa vision, mettre en avant ses travaux et s'inscrire en interface avec la Commission européenne.

Dans sa stratégie pour la donnée, la Commission européenne promeut le développement des espaces européens communs des données et l'interconnexion d'infrastructures en nuage (« *cloud* »). Plusieurs projets d'espaces communs des données interopérables à l'échelle de l'Union européenne concordent avec le périmètre d'action du pôle ministériel, comme pacte vert, mobilité ou énergie. A terme, l'utilisation de bases de données européennes sera un vecteur fort de valeur. Par exemple, la gestion des déchets à la frontière nécessite une vision transfrontalière.

(Voir les initiatives de la feuille de route sur les espaces communs de données dans les objectifs 3 et 4).

## 5.3. Relayer auprès des services et des partenaires les initiatives européennes et favoriser les recherches de partenariat et de financement européens

Le partage d'expérience autour de la recherche de financements européens est à encourager. La Communauté de la donnée, des algorithmes et des codes sources du pôle ministériel peut servir cet objectif.

# 6. Objectif : Utiliser les données de manière conforme et sécurisée

	RESPONSABLE	CONTRIBUTEUR	JALON PREVISIONNEL *
<b>ACTION PHARE : Compléter l'accompagnement juridique</b>			
>> Réaliser et diffuser une cartographie de l'existant relatives aux rôles et responsabilités des acteurs en matière d'accompagnement juridique	ECOLAB	SG - CGDD	<b>CT</b>
>> Définir une offre de services en matière d'accompagnement juridique des projets / initiatives du pôle ministériel, orientée utilisateurs, sur la base de la cartographie de l'existant	ECOLAB	SG - CGDD	<b>CT</b>

**CT** : 2021-1<sup>er</sup> semestre 2022 ; **MT** : 2022-2023 ; **LT** : > 2023

\*Premières réalisations tangibles



L'ouverture des données est une notion encadrée juridiquement. Des normes de plusieurs natures indiquent, avec la complexité que renferment leurs recoupements, si une donnée doit être ouverte ou pas. La notion de partage est plus vaste, il s'agit d'aller au-delà du cadre législatif ou réglementaire pour faire circuler une donnée.

## 6.1. Cerner le périmètre juridique concerné

L'ouverture d'une donnée requiert d'examiner des exigences législatives et réglementaires complexes et qui interagissent entre elles. Ainsi :

- La loi pour une République numérique du 7 octobre 2016 prévoit que toutes les données administratives sont sujettes par défaut à l'ouverture, sauf exceptions liées à leur régime de protection particulier
- la Commission d'accès aux documents administratifs (CADA) exerce un contrôle et émet des avis relatifs à l'accès aux données publiques
- les articles L124-1 à L124-8 du code de l'environnement organisent le droit d'accès aux informations environnementales (en application de la Convention d'Aarhus - Accès à l'information, participation du public et accès à la justice en matière d'environnement) et les articles L127-1 à L127-9 de ce même code contiennent les dispositions transposées de la directive INSPIRE (Infrastructure d'information géographique dans la Communauté européenne)
- l'article L127-10 du code de l'environnement définit un cadre de publication des données géographiques, en particulier l'adresse et les parcelles cadastrales qui servent de socle de publication de certaines données
- l'article L125-5 du code de l'environnement permet de publier des données d'émissions de substance dans l'environnement, en vertu de la convention d'Aarhus, sans les limitations habituelles relevant du secret des affaires
- des réglementations sectorielles (ex : données énergétiques des réseaux publics de distribution, données de la mobilité) posent des règles spécifiques en matière d'ouverture de données.

En matière de protection des données entraînant des restrictions à leur ouverture, et encore sans prétendre à l'exhaustivité, on trouve :

- la protection des données à caractère personnel, en vertu du règlement général européen sur la protection des données (RGPD) et de la loi Informatique et Libertés
- le secret de la défense, qui protège la sécurité nationale
- le secret des affaires, qui protège la libre-concurrence
- l'article L124-4 du code de l'environnement (par exemple non-divulgateion de l'emplacement d'une biodiversité menacée car convoitée)
- la souveraineté économique
- la prise en compte des contrats et de leurs clauses.

Enfin, la perspective de l'organisation d'échanges accrus de données avec les entreprises ou les ONG, qui passeront notamment par des espaces communs de données, nécessite une expertise juridique particulière.

Au-delà de la dichotomie ouverture/fermeture, et avec les possibilités de partage ou d'accès sécurisés à la donnée sensible, il est indispensable que la donnée soit exposée, qu'elle circule et qu'elle soit consommée autant que possible.

## 6.2. Vers un accompagnement juridique centralisé

Face à ce foisonnement normatif, une offre de services en matière juridique doit être mise en œuvre pour accompagner les entités du pôle ministériel.

Pour les lever, un premier travail de cartographie devra être mené pour clarifier les rôles et responsabilités en matière de conseil juridique. Dans un deuxième temps, une offre de service orientée « utilisateurs » pourra être pensée en s'appuyant sur les services en place.

Ce service pourrait être incarné par un groupe d'experts métiers et/ou des acteurs de la protection des données. Lorsque l'administration a besoin d'accéder à des données ou qu'une demande d'accès à ses données est réalisée, ce groupe d'experts fournirait un conseil juridique sous forme d'avis et recommandations documentant les risques, opportunités et obligations en matière de partage de données. Parmi ces experts, il est entre autres envisagé de mobiliser le Département Sécurité et gestion de crise du Service du numérique (DSGC), les représentants du Délégué à la Protection des Données et la Personne responsable de l'accès aux documents administratifs et des questions relatives à la réutilisation des informations publiques (PRADA – *transfert en cours de la DAF à la DAJ*) de la Direction des affaires juridiques ou le service du Haut-Fonctionnaire de Défense et de Sécurité (HFDS).

## 7. Faire vivre la feuille de route via une gouvernance dédiée

	RESPONSABLE	CONTRIBUTEUR	JALON PREVISIONNEL *
<b>ACTION PHARE : Déployer et animer la communauté des utilisateurs et producteurs de la donnée</b>	ECOLAB		CT
<b>Autres actions</b>			
Déployer les comités permettant de faire vivre dans la durée la feuille de route de la donnée	ECOLAB	Membres gouvernance	CT
Mettre en place une animation spécifique au niveau territorial qui fédère les dispositifs existants	ECOLAB	SNUM-DDD	CT

CT : 2021-1<sup>er</sup> semestre 2022 ; MT : 2022-2023 ; LT : > 2023

\*Premières réalisations tangibles

### 7.1. Elaborer et mettre en œuvre la feuille de route de la donnée, des algorithmes et des codes sources

Au sein du pôle ministériel, le commissariat général au développement durable « est chargé, en lien avec le secrétariat général et en appui aux directions régionales du ministère, de la supervision générale des données sur le logement, l'énergie, les transports, l'environnement ainsi que de l'évaluation et de la mobilisation des moyens nécessaires pour le développement de la mise à disposition de ces données à des fins de connaissance, d'étude, de recherche et d'évaluation »<sup>7</sup>.

<sup>7</sup> Décret n°019-1528 du 30 décembre 2019, article 2, modifiant le décret n°2008-680 du 9 juillet 2008.

Dans son rôle d'administrateur ministériel des données et en lien avec les entités du pôle ministériel, il coordonne la stratégie (feuilles de route, gouvernance des données, circulation des données) et joue un rôle d'animation (communautés, appui, expérimentations).

Au sein du CGDD, l'Ecolab « prépare et coordonne les actions du superviseur général des données »<sup>8</sup> et est ainsi missionné pour élaborer et animer la feuille de route de la donnée.

Pour les aspects techniques et les outils de la donnée, il se coordonne étroitement avec le service du numérique (SNum) du secrétariat général qui porte la mission générale de définition et de mise en œuvre de la politique numérique du ministère.

Les rôles et responsabilités des deux entités au titre de la donnée sont décrits en annexe.

Depuis mai 2021, le ministère de la cohésion des territoires et des relations avec les collectivités territoriales dispose de son propre administrateur ministériel des données, au sein de l'agence nationale de la cohésion des territoires [ANCT]).

## 7.2. Piloter dans le temps la mise en œuvre de la feuille de route de la donnée, des algorithmes et des codes sources

La feuille de route est un document évolutif, à actualiser régulièrement. Pour s'adapter en continu aux besoins du pôle ministériel et des métiers et produire des effets concrets, la feuille de route nécessite un pilotage impliquant toutes les parties prenantes de la donnée, des algorithmes et des codes sources à des degrés divers.

Son pilotage politique sera assuré par un comité de pilotage se réunissant 2 fois par an et réunissant le niveau des directeurs du pôle ministériel.

Son animation et son suivi plus régulier sera assuré par une Communauté de la donnée, des algorithmes et des codes sources.

---

<sup>8</sup> Arrêté du 30 décembre 2019 modifiant l'arrêté du 9 juillet 2008 portant organisation de l'administration centrale du ministère de l'énergie, de l'électricité, du développement-durable et de l'aménagement du territoire. Equivalent de « Administrateur général des données ».

## INSTANCE DECISIONNELLE



QUI ?

### Comité de pilotage

(2 fois / an)



QUOI ?

- Les directeurs généraux
- L'administrateur ministériel des données
- Représentant(s) DREAL / DDT

Au lancement du pilotage de la feuille de route, l'ouverture du comité de pilotage à des représentants élargis sera étudiée : opérateurs ; représentants d'entreprises ; représentants de collectivités

- Assurer l'adéquation de la feuille de route avec la stratégie ministérielle
- Piloter l'atteinte des objectifs de la feuille de route
- Arbitrer les points organisationnels, techniques stratégiques, etc.
- Infuser les objectifs de la feuille de route dans le pôle ministériel

### 7.3. Faire vivre la feuille de route autour de la Communauté de la donnée, des algorithmes et des codes sources

La Communauté de la donnée, des algorithmes et des codes sources, rassemblant les producteurs et les utilisateurs, est l'instance de base de pilotage de la feuille de route. Son noyau comprend le pôle ministériel, au niveau central et territorial, et son ambition est d'associer des acteurs privés, collectivités locales, associations relevant des missions du pôle ministériel. Elle doit être préfigurée à la mi-2021 (la première réunion a eu lieu le 21 juin).

La Communauté rassemble autour d'intérêts communs :

- partager les besoins, lever les blocages
- monter en compétences
- partager les bonnes pratiques, les innovations, les retours d'expérience (ex. partenariats public-privé fructueux ; pratiques en matière de partage de données ; ingénierie financière des projets ; points durs juridiques)
- constituer une base de connaissances en accès continu
- faire circuler l'information, les outils et les méthodes (ex. évolutions du cadre législatif ; circulation d'une nouvelle norme européenne ; partage d'outils)
- s'entraider
- porter des initiatives, des événements
- promouvoir la feuille de route.

La mise en place de la communauté sera soutenue par un outil dédié permettant à la fois d'échanger en direct, de publier des contenus dans des fils de discussion et de manière pérenne (stockage documentaire) et de gérer différents cercles d'acteurs.

## INSTANCE DE PARTAGE

### Communauté de la donnée, des algorithmes et des codes sources

(1 fois / trimestre)

- Les responsables des chantiers de la feuille de route : représentants décisionnaires
  - Les autres membres de la communauté  
Représentants des DAC, opérateurs, DREAL, DDT, entreprises, collectivités...
- 
- Suivre l'exécution opérationnelle de la feuille de route
  - Partager les besoins, lever les blocages
  - Monter en compétence collectivement
  - Partager les bonnes pratiques
  - Faire circuler l'information, les outils et les méthodes

20

### 7.4. Animer la donnée au niveau territorial dans un lieu d'échange unique

L'animation de la politique de la donnée au sein des territoires est assurée dans le pôle ministériel par :

- La communauté des chefs de services connaissance et développement durable, portée par la délégation au développement durable (DDD) ;
- Les Journées Nationales des Services Régionaux – Information géographique (JNSR), rassemblant les administrateurs de données - coorganisées par Ecolab et le SNum ;
- Les Conseillers en Management de Système d'Information Géographique (CMSIG), agents du service du numérique pour l'infrastructure en information géographique.

La feuille de route préconise la création d'un lieu unique de dialogue sur la donnée pour assurer l'animation territoriale des sujets autour de la donnée. Son animation territoriale devrait intégrer le responsable de l'administration de la donnée qui doit être nommé par chaque préfet de région, et associer les problématiques des collectivités.

### 7.5. Actionner la feuille de route de manière coordonnée avec les gouvernances extérieures

Au niveau interministériel, le dispositif de gouvernance de la feuille de route profitera de la participation de l'AMD à la gouvernance interministérielle et a vocation à y contribuer également (Comité interministériel des administrateurs de données, animation des AMD).

Hors du pôle ministériel, d'autres instances œuvrent à la création, au partage, à la circulation et à un meilleur usage de la donnée de la transition écologique. Parmi celles-ci, deux d'entre elles se positionnent en lien étroit avec le pôle ministériel : le conseil national de l'information géographique

(CNIL) et l'association française pour l'information géographique (AFIGEO). Elles seront associées à des actions de la feuille de route. Il en est de même des gouvernances pluridisciplinaires intégrant la transition écologique notamment au travers des espaces communs de données.

## 8. Équiper le pôle ministériel avec des outils adaptés en matière de données

	RESPONSABLE	CONTRIBUTEUR	JALON PREVISIONNEL *
<b>ACTION PHARE</b> : Promouvoir l'offre de services de la donnée : plateforme de données	SNUM		<b>MT</b> S1 2022
<b>ACTION PHARE</b> : Concevoir et déployer la Géoplateforme	IGN		<b>MT</b> SOCLE TECHNIQUE ET OFFRE ACTUELLE DU GEOPORTAIL <b>LT</b> SERVICES GEOGRAPHIQUES ET AUTRES APPLICATIFS
<b>Autres actions</b>			
Mettre à disposition des services les outils permettant la réalisation d'une cartographie technique et applicative sur l'outillage de la donnée au sein du pôle ministériel	SNUM		<b>CT</b>
Recenser les besoins métier en matière de donnée et proposer des solutions de mise en œuvre en regard (« trajectoire SI »)	SNUM		<b>MT</b> S1 2022
Mettre en place un outil communautaire pour faire vivre la communauté de la donnée, des algorithmes et des codes-sources	ECOLAB		<b>CT</b>

**CT** : 2021-1<sup>er</sup> semestre 2022 ; **MT** : 2022-2023 ; **LT** : > 2023

\*Premières réalisations tangibles

La notion d'outils développée ici porte sur tous les supports technologiques (SI et infrastructures) soutenant la mise en œuvre de la feuille de route, c'est-à-dire permettant de capter, stocker, partager et valoriser les données.



Schématiquement, voici les principaux types de technologies concernées, selon les objectifs de la feuille de route :

Obj. 1	Obj. 2	Obj. 3	Obj. 4	Obj. 5	Obj. 6	Gouvernance
Outil collaboratif de production et d'entretien des données	E-learning / MOOC API / webservices					
Outil de catalogage (guichet d'accès)	Outils et technologies d'exploitation de données : dataviz, Dataiku, ...	Hub de données (socle ou thématiques)	-		Outil d'anonymisation	Outil collaboratif
Outil d'archivage électronique (SAE)						

La définition et la mise en place d'un guichet d'accès à la donnée de la transition écologique et territoriale est décrit dans l'objectif 1 et fait l'objet d'une action phare de cette feuille de route.

### 8.1. S'inscrire au maximum dans l'écosystème technique et applicatif existant ou planifié

Le pôle ministériel est doté d'un écosystème technique et applicatif dense :

- Dans le cadre du Plan de transformation numérique, le service du numérique (SNum) a défini une offre de services sur le volet des données. Le plan de transformation numérique du pôle ministériel consacre une fiche action au volet donnée : fiche n° 6 "Mettre la donnée au profit de l'action publique". En s'appuyant sur la trajectoire de la présente feuille de route, les actions majeures prévues pour 2021 y sont précisées. Un engagement fort est pris pour 2021 concernant le déploiement des outils de gestion de l'information basés sur l'intelligence artificielle au bénéfice des services instructeurs de l'autorité environnementale par la coopération entre l'Écolab du CGDD et le service connaissance de la DREAL Bretagne."
- Les directions générales, dotées de leurs propres budgets SI, ont enrichi leurs outils en matière de données ;  
Ex. La plateforme Internet Géorisques, développée et opérée par le BRGM et mise en service en 2014, est progressivement devenue le canal principal de diffusion des données sur les risques naturels et technologiques de la DGPR ;
- Les territoires, au travers des DREAL et des DDT, déploient des projets variés parfois en collaboration avec la Région ;  
Ex. La DREAL Grand-Est étudie la mise en place d'une solution technique facilitant l'accès aux données utiles à ses missions et de mieux diffuser ses données ;  
Ex. Les plateformes territoriales, reposant sur l'utilisation des logiciels libres : Prodige, geOrchestra ;
- Les opérateurs, selon leurs besoins, ont des niveaux de maturité très différents sur la valorisation de la donnée.

De fait, pour **répondre au besoin métier** décrit dans la feuille de route, il est nécessaire de s'inscrire dans le paysage technique et applicatif existant et de capitaliser au maximum sur les ressources disponibles. Pour y parvenir, il convient tout d'abord de mettre à plat la vision existante autour d'une **cartographie technique et applicative, partagée**, à réaliser sur le périmètre décrit. Ce travail pourra être mené par les services métier avec l'appui du SNum et de manière incrémentale pour obtenir des résultats probants à court terme.

Ensuite, une trajectoire précise pourra être définie en fonction : reprise de l'existant, évolutions de l'existant, lancement de nouveau projet, etc.

Les principales briques applicatives existantes face aux besoins définis sont décrites dans les paragraphes suivants.

## 8.2. Réussir la Géoplateforme : l'outil collaboratif de production et d'entretien de la donnée

Le projet de **Géoplateforme** conçu par l'Institut national de l'information géographique et forestière (IGN) a pour objectif de construire, avec des partenaires publics et privés, un **espace public de l'information géographique**, sous forme d'une plateforme de données géographiques ouverte et mutualisée dotée d'une gouvernance partagée.

### CONCRETEMENT, QUE PERMET LA GÉOPLATEFORME ?

- Administrer les données géographiques par les communautés ;
- Rendre accessible un panel très large de données, avec à la clé la garantie de l'interopérabilité avec des plateformes et des systèmes d'informations référencés ;
- Produire et entretenir des bases de données de manière collaborative ;
- Déployer des solutions à grande échelle par la mutualisation d'infrastructures sans se soucier de la capacité, de la fiabilité ou des performances ;
- Développer des nouveaux services en offrant des capacités de traitement sur les données et des services à haute valeur ajoutée ;
- Gérer des communautés autour de la donnée géographique.

La Géoplateforme sera construite selon un enjeu double de disposer d'un socle technique performant d'ici fin 2022 tout en apportant rapidement de la valeur par le début des développements de services et des interfaces utilisateurs dès 2022.

Ce programme, lauréat du fonds de transformation de l'action publique (FTAP), s'étale jusqu'en 2024.

## 8.3. Mettre à disposition une plateforme de services de la donnée et les capacités d'hébergement associées

Au-delà de la Géoplateforme, destinée à la donnée géographique et réalisée par l'IGN, le Service du numérique lance une offre de services en matière de valorisation des données auprès des Directions métiers au travers d'une **plateforme de services autour de la donnée**.

Le développement de la plateforme est réalisé de manière incrémentale à partir de cas d'usage en fournissant les briques technologiques nécessaires à leur bon fonctionnement et à leur passage à l'échelle. Début 2022, celle-ci devrait couvrir un cas d'usage : la création d'un outil de pilotage des contrôles et des zones contrôlées en mer.

Le projet vise aussi à renforcer la capacité de la plateforme à fournir les briques technologiques nécessaires au bon fonctionnement et au passage à l'échelle des cas d'usage.

Cette offre de services<sup>9</sup> fera l'objet d'une communication étendue pour gagner en visibilité au sein du pôle ministériel.

---

<sup>9</sup> Annexe n°4

## CONCRETEMENT, QUE PERMET LA PLATEFORME ?

- Collecter et croiser des données internes et externes
- Rendre autonomes les métiers pour la création de data-visualisation, de tableaux de bord et d'algorithmes métier de data science
- Accélérer la création et l'industrialisation des algorithmes d'IA via des outils standards (R et Python) et prêts à l'emploi pour les datascientists
- Fournir des capacités de stockage de volumes de données importants
- Disposer des capacités de modernisations des applications décisionnelles historiques

### 8.4. Déployer des outils d'aide à la décision (visualisation de la donnée ou solutions décisionnelles)

Plusieurs solutions de visualisation de la donnée existent au sein du pôle ministériel :

- La solution privée de plateforme de data science Dataiku, déployé auprès de quelques directions<sup>10</sup> depuis 2020, intègre une brique dédiée ;
- Les directions générales utilisent des outils tels que Tableau™ ;
- La future plateforme de données, décrite au paragraphe précédent, prévoit une fonctionnalité de visualisation de la donnée.

Ces solutions, ou d'autres équivalentes en fonction de l'offre du marché, apporteront une réponse aux premiers cas d'usage et aux utilisateurs potentiels et s'enrichiront pour répondre aux besoins complémentaires qui apparaîtront.

### 8.5. Mettre en place un outil collaboratif communautaire

Plusieurs outils collaboratifs, ayant des usages proches mais différenciés, sont d'ores et déjà mis en œuvre au sein du pôle ministériel : des outils d'échange communautaire tels qu'Osiose – mis à disposition par la DINUM - ; des outils de gestion électronique de documents (GED) collaboratifs (ex. Alfresco ou MS Sharepoint, dont les fonctionnalités vont au-delà).

En s'appuyant sur l'écosystème existant, une solution destinée à la Communauté de la donnée, des algorithmes et des codes sources couvrira les principaux besoins :

- Gestion de cercles thématiques de la Communauté ;
- Partage documentaire et travail collaboratif ;
- Visio-conférence, chat ; etc.

---

<sup>10</sup>ECOLAB, DGITM/DAM, SDES/BACS, DGALN

# Annexe 1 - Méthode mise en œuvre et participants

Cette feuille de route est le fruit d'un travail collectif mené sous la direction d'Ecolab, en charge de la mise en œuvre de l'administration ministérielle des données au nom du commissaire général au développement durable, administrateur ministériel des données.

Pour ce travail, Ecolab s'est coordonné avec le Service du numérique du secrétariat général. Il a été accompagné par le cabinet conseil Accenture.

La conception de cette feuille de route s'est appuyée sur un ensemble de réunions de discussions avec les directions générales du pôle ministériel, des opérateurs et des services déconcentrés qui ont permis de mettre en avant freins et irritants et de faire émerger les axes de progression et les actions proposés.

- ADEME : Noam Leandri, Laurent Morice, Marc Viot
- ANCT : Pierre de Saint Ferjeux, Sylviane Le Guyader, Pierre-Louis Rolle
- BRGM : Arnaud Garnier, Vincent Labbé, Jean-Marc Trouillard
- CEREMA : Eric Lajarge
- CGDD/DDD : Edwige Duclay
- CGDD/SDES : Béatrice Sédillot, Frédéric Vey
- CGDD/SDSED : Jérôme Baret, Nicolas Chucho, Lionel Janin
- CGDD/SRI : Hélène Bégon, Hugo Berthelé, Solène Cadiou, Hélène Costa de Beauregard, Thomas Cottinet, Thierry Courtine, Benoît David, Olivier Dissard, Magali Domergue, Oriane Gauffre, Gwenaëlle Hello, Clément Jaquemet, Marc Léobet, Gabrielle Pastré, Sylvain Pradelle, Claire Sallenave
- DINUM/Etalab : Romain Talès
- DGAC : Stéphane Chatty, Thibaud Figueroa, Nicolas Naudin-Roy, Pascal Planchon, Ivonig Le Ruyet
- DGALN : Guillaume Levieux
- DGEC : Julien Barreteau, Alice Vieillefosse, Christophe Kassiotis
- DGITM : Victor Alonso, Michel Deboscher, Mathieu Fernandez-Vandewalle, Hervé Philippe, Rémi Pochez, Benoit Queyron, Natacha Roger
- DGPR : Guillaume Leseq
- DREAL : Chantal Adjriou, Christine Bordier, Yann Deffin, Emmanuel Fillon, Philippe Gazeau, Isabelle Kauffman, Enrique Portola, Alexandrine Sens, Marie-Alexandrine Valere
- IGN : Claude Pénicand
- INERIS : Jean-Yves Chatelier
- Météo France : Christophe Morel
- OFB : François Hissel, Alexandre Liccardi
- SGD/DAF : Camille Tatger
- SG/DAJ : Fabrice Breteche, Alexandre Mouton, Etienne Orleanski, Sophie Guyen
- SG/DRH : Valéry Lemaitre
- SG/SHFDS : Juliette Brezillon, Eric Tetelin
- SG/SNum : Delphine Auger, Jean-François Boutier, Christophe Boutonnet, Pierre-Louis Chamelot, Fabienne Chevalier, Olivier Cordelier, Frédéric Damiens, Serge Doba, Hervé Domalain, Frederic Greggio, Gilles Guillouet, Corinne Lafont, JP Lang, Ghizlane Lebel, Sylvie Mompert, Erwan Salmon, Gilles Toquin, Aurélie Vieillefosse, Alain Wermeister
- Université Gustave Eiffel : Jean-Bernard Kovarik

## Annexe 2 - Panorama des acteurs de la donnée du pôle ministériel

Service	Entité	Responsabilités principales	Contenus
Commissariat général au développement durable (CGDD)	Ecolab	Administrateur ministériel des données pour les ministères de la transition écologique et de la mer	Co-élaboration et supervision de la feuille de route de la donnée, des algorithmes et des codes sources du pôle ministériel (ministère de la Transition écologique - ministère de la Cohésion des territoires et des Relations avec les collectivités territoriales - ministère de la Mer) avec le Service du numérique Garantie du respect des obligations d'ouverture des données pour la transparence de l'action ministérielle et l'information des citoyens sur l'environnement Travaux sur le catalogage, les référentiels communs Animation de l'usage de la donnée au sein du pôle ministériel, recherche de cas d'usage de la donnée Promotion de la création d'espaces communs de données sur les territoires et dans des écosystèmes économiques Animation de la montée en compétences sur la donnée au sein du pôle ministériel
		Promotion de l'usage de la donnée	Organisation d'événements « données » (ex. journée de la donnée, hackathons, data camps, data sessions, ...) Conseil en matière de financement et de marchés publics en matière de politique de la donnée
		Expertise en intelligence artificielle	Coordination de la feuille de route de l'intelligence artificielle du pôle ministériel Conseil sur des projets en intelligence artificielle Elaboration de preuves de concept (POC – Proof of concept) en intelligence artificielle au service des politiques ministérielles Animation du réseau métier « Le Garage IA » des data scientists du pôle ministériel
		Gouvernance de la donnée géographique	Vice-présidence et secrétariat général du conseil national de l'information géographique Application de la directive Inspire Pilotage de l'outil Prodige pour les plateformes régionales de données géographiques
		Accompagnement de l'écosystème français des start-up et incubateurs GreenTech	Labellisation et promotion de start-up et PME vertes et innovantes, dont en matière d'usage de la donnée

Service	Entité	Responsabilités principales	Contenus
	Sous-direction de l'animation scientifique et technique (AST)	Animation du réseau scientifique et technique ministériel	Réflexions et travaux sur la production et l'usage de données au sein du réseau Tutelle de Météo-France et de l'Institut national de l'information géographique et forestière (IGN) ; par ex : travaux sur le modèle économique de Météo France et IGN face aux textes relatifs à la gratuité de la réutilisation des données publiques
	Sous-direction de la recherche (SDR)	Données satellitaires	Elaboration et suivi du plan d'application satellitaire du pôle ministériel
	Sous-direction de la stratégie éditoriale et de la diffusion (SDSED)	Gestion du portail d'information environnementale Diffusion de la connaissance	<a href="http://notre-environnement.gouv.fr">notre-environnement.gouv.fr</a> diffuse la connaissance en matière environnementale et du développement durable
	Service des données et des statistiques (SDES)	Service statistique du pôle ministériel	Production et mise à disposition des jeux de données statistiques du pôle ministériel Production d'indicateurs et d'analyses statistiques, datavisualisation
Secrétariat général	Service du numérique (SNum)	Contributeur majeur de la feuille de route de la donnée du pôle ministériel	Participation à l'élaboration de la feuille de route de la donnée, des algorithmes et des codes sources du pôle ministériel comme contributeur majeur Portage des actions d'ouverture des algorithmes et des codes sources du pôle ministériel Architectures et solutions techniques, urbanisation, infrastructures, sécurité Accompagnement des services dans la mise en œuvre de la sécurité de leurs données Accompagnement des services pour mettre en œuvre la politique du Dites-le nous une fois, le règlement général sur la protection des données (RGPD), l'archivage
		Fabrique numérique	Elaboration de preuves de concept (POC – Proof of concept) au service des politiques ministérielles
		Donnée géographique	Pilotage de l'outil Géo-IDE pour les plateformes régionales de données géographiques
	Direction des ressources humaines (DRH)	GPEEC, recrutement, gestion de carrière	Recruter et gérer la carrière des experts de la donnée Etablir le lien avec la DGAFP et la DINUM sur la politique interministérielle de ressources humaines pour la donnée
		Formation [FORCQ, CMVRH]	Faire monter en compétence des spécialistes de la donnée Former et acculturer à la politique et aux usages de la donnée

Service	Entité	Responsabilités principales	Contenus
		Se doter d'une stratégie et d'outils en matière de production et d'usage de la donnée	Préparer la stratégie de l'usage des données des ressources humaines au sein du pôle ministériel
	Direction des affaires financières (DAF)	Mission des archives et de la gestion de l'information électronique (MAGIE)	Politique de l'archivage des données Préparer la stratégie de l'usage des données financières au sein du pôle ministériel
		Se doter d'une stratégie et d'outils en matière de production et d'usage de la donnée	Préparer la stratégie de l'usage des données financières au sein du pôle ministériel
	Service du Haut fonctionnaire de défense et de sécurité (HFDS)	Sécurité et intelligence économique Gestion de crise	Dispositions relatives à la sécurité de défense, à la protection du secret et à la protection du patrimoine scientifique et technique Dispositions relatives à la sécurité des systèmes d'information.
	Direction des affaires juridiques (DAJ)	Correspondant de la commission nationale de l'informatique et des libertés (CNIL)	Garantir l'application du règlement général sur la protection des données (RGPD) au sein du pôle ministériel
		Personne responsable de l'accès aux documents administratifs (PRADA) ( <i>en cours</i> )	Accès aux données, correspondant ministériel de la Commission d'accès aux documents administratifs (CADA) ( <i>en cours de transfert de la DAF à la DAJ</i> )
		Conseil aux services en matière de droit de la donnée	Protection des données personnelles Droit de la propriété intellectuelle Licences de réutilisation Conventions relatives à l'échange de données
		Se doter d'une stratégie et d'outils en matière de production et d'usage de la donnée	Préparer la stratégie de l'usage des données juridiques au sein du pôle ministériel
	Service du pilotage et de l'évolution des services (SPES)	Structuration des services en matière de Connaissance	Participer au pilotage du projet « Connaissance » du pôle ministériel (positionnement de la mission « Connaissance » de la transition écologique au sein des services)
		Tutelle des écoles du pôle ministériel [PPSTE]	Formation et acculturation à la donnée
		Sous-direction de l'innovation pédagogique et collaborative [IPEC]	Formation et acculturation à la donnée
Conseil général de l'environnement et du développement durable CGEDD)			Participer au pilotage de la mission « Connaissance » du pôle ministériel (positionnement de la mission « Connaissance » de la transition écologique au sein des services) Contrôler et évaluer la politique de la donnée du pôle ministériel



Service	Entité	Responsabilités principales	Contenus
			Se doter d'une stratégie et d'outils en matière de production et d'usage de la donnée produite par le CGEDD
Directions générales (DG)		Se doter d'une stratégie et d'outils en matière de production et d'usage de la donnée	Appliquer la politique d'ouverture de la donnée publique pour la transparence de l'action ministérielle et l'information des citoyens sur l'environnement Se doter en propre ou en lien avec des partenaires de solutions d'usage de la donnée au service des politiques publiques portées par la direction générale Co-porter la promotion de la création d'espaces communs de données entrant dans le champ de compétences de la direction générale Se doter le cas échéant d'un laboratoire d'innovation par la donnée
Services déconcentrés du pôle ministériel		Se doter d'une stratégie et d'outils en matière de production et d'usage de la donnée	Appliquer la politique d'ouverture de la donnée publique pour la transparence de l'action ministérielle et l'information des citoyens sur l'environnement Se doter en propre ou en lien avec des partenaires de solutions d'usage de la donnée au service des politiques publiques portées par le service déconcentré Entrer dans la préfiguration du projet « Connaissance » du pôle ministériel (positionnement de la mission « Connaissance » de la transition écologique au sein des services) Etablir des liens avec les collectivités territoriales porteuses de projets d'utilisation de données ou d'espaces communs de données
Opérateurs du pôle ministériel <u>Non exhaustif</u>	Agence nationale de la cohésion des territoires (ANCT)	Administrateur ministériel des données du ministère de la Cohésion des territoires et des Relations avec les collectivités territoriales (MCTRCT). Rôle de pilote, en appui au MCTRCT, en matière de lien entre l'Etat-et les collectivités territoriales pour les politiques de la donnée	Pilotage stratégique, animation d'acteurs en matière de politique générale de la donnée Production, récupération, diffusion et utilisation de nombreux jeux de données
		Production, récupération, diffusion et utilisation de nombreux jeux de données et notamment production de données considérées comme des données de référence pour leur écosystème  Le cas échéant démarches pilotes d'intelligence artificielle	<ul style="list-style-type: none"> <li>• Agence de la transition écologique - ADEME : économie circulaire, sols pollués et friches, énergie et climat, air et bruit...</li> <li>• Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement (CEREMA) : environnement, mobilité, foncier...</li> <li>• Bureau de recherches géologiques et minières (BRGM) : sciences de la terre</li> <li>• Office français de la biodiversité (OFB) : biodiversité, eau</li> <li>• Météo-France : données météorologiques et climatiques</li> <li>• Institut national de l'information géographique et forestière (IGN) : données géographiques et forestières</li> <li>• Institut national de l'environnement industriel et des risques (INERIS) : risques industriels et environnementaux</li> </ul>

Service	Entité	Responsabilités principales	Contenus
			<ul style="list-style-type: none"> <li>• Service hydrographique et océanographique de la Marine (SHOM) : information géographique maritime et littorale</li> <li>• Museum national d'histoire naturelle (MNHM) : biodiversité, espaces naturels (<u>Etc.</u>)</li> </ul>

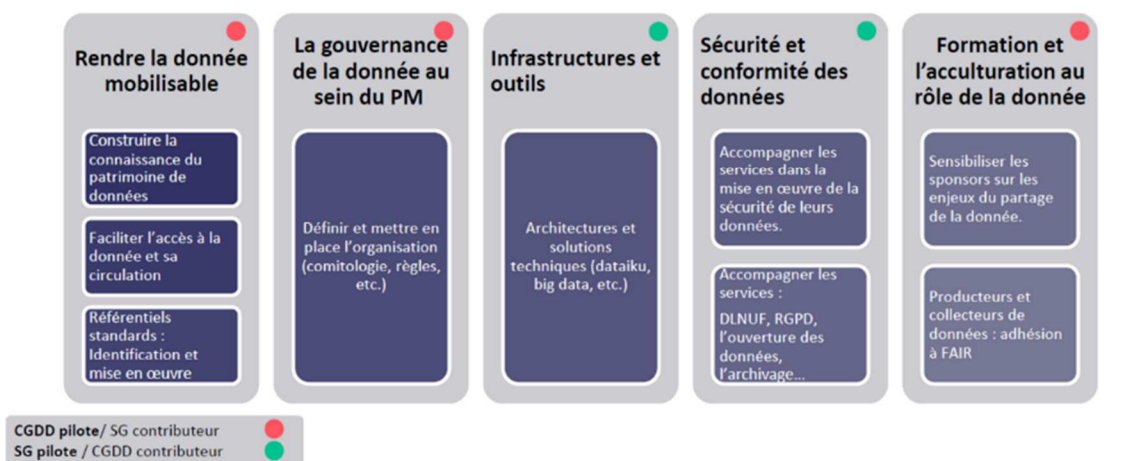
# Annexe 3 – Répartition des rôles CGDD-SNum sur le volet data du Plan de Transformation Numérique



Partenariat Snum- CGDD  
 Point d'entrée Snum : Dptn  
 Point d'entrée CGDD : Ecolab

## Répartition des rôles SG-CGDD sur le volet data du PTN

(SG : snum, daj, magie, shfds)



SG/SNUM/DPTN

1

17/12/2020

# Annexe 4 – Extrait de l'offre de services du Service du Numérique (SNum)

Secrétariat général

## EXTRAIT DE L'OFFRE DE SERVICE SNUM (ANNEXE 1) SYNTHÈSE DES OFFRES

Présentation

**HÉBERGEMENT & EXPLOITATION CLOUD**

Un service d'hébergement moderne basé sur le cloud.

- En fonction des besoins (disponibilité, sécurité, élasticité...), différentes solutions sont disponibles Cloud Interne ou Cloud Externe.
- Ce service d'hébergement est complété par l'exploitation applicative selon différents niveaux de services

**● VALORISATION DES DONNÉES** *Offre en cours de construction*

Un socle technologique Big Data pour collecter, stocker, traiter et valoriser les données du patrimoine applicatifs et les croiser avec des données exogènes. Ce service offre l'ensemble des outils nécessaires pour :

- Moderniser les solutions décisionnelles historiques
- Créer des tableaux de bord en autonomie
- Créer des algorithmes d'intelligence artificielle avec les datascientists

**● INTEROPÉRABILITÉ ET DÉMATÉRIALISATION**

Un bouquet de services pour :

- Faciliter la dématérialisation des démarches et des formulaires et répondre aux enjeux de simplification des relations avec les usagers
- Accélérer et maîtriser le partage et la publication des données avec les partenaires via des API (interface standard à l'état de l'art)

MTE - SNUM - Offre Socle technologique

19

Secrétariat général

## EXTRAIT DE L'OFFRE DE SERVICE DU SNUM (ANNEXE 1) OFFRE DE VALORISATION DES DONNÉES

Présentation de l'offre

*Offre en cours de construction* **01101  
01101  
01101**

Les besoins couverts	PRINCIPES DIRECTEURS	DÉTAIL
<ul style="list-style-type: none"> <li>Croiser des données de différents silos applicatifs</li> <li>Collecter et croiser des données internes et externes</li> <li>Rendre autonome les métiers pour la création des tableaux de bord et d'algorithmes métier de datascience</li> <li>Accélérer la création et l'industrialisation des algorithmes d'IA via des outils standards (R &amp; Python) et prêts à l'emploi pour les datascientists</li> <li>Fournir des capacités de stockage de volume de données important</li> <li>Disposer des capacités de modernisation des décisionnels historiques</li> </ul>	<p><b>DESCRIPTION ET BÉNÉFICES</b></p> <p>Mise à disposition d'un service de collecte, de stockage, de traitement et de valorisation des données sur la base des technologies Big Data. Les principales fonctionnalités offertes sont :</p> <ul style="list-style-type: none"> <li>Modernisation les SI décisionnels existants</li> <li>Créer des indicateurs et tableaux de bord en autonomie (dataiku ou lumira)</li> <li>Créer des algorithmes d'intelligence artificielle (R &amp; Python, dataiku)</li> </ul> <p>Cette plateforme est capable de traiter des données structurées et non structurées, interne ou externe en facilitant le croisement et l'exploitation.</p>	
	<p><b>SUPPORT ET ENGAGEMENTS</b></p> <ul style="list-style-type: none"> <li>Support aux équipes utilisatrices à la prise en main des outils mis à disposition</li> <li>Fourniture d'un cadre méthodologique à l'évaluation de l'éligibilité des usages à la plateforme</li> <li>Garantie de disponibilité de la plateforme mise à disposition</li> </ul>	
	<p><b>PRÉ-REQUIS ET EXIGENCES</b></p> <ul style="list-style-type: none"> <li>L'offre est ouverte à toute entité du pôle ministériel</li> <li>L'utilisation de l'offre nécessite de mobiliser des compétences - data - pour développer les cas d'usage (ingénieur data, datascientist...)</li> <li>La création de projet sur les plateformes data du pôle ministériel passe par un entretien sur l'éligibilité avec MSP/DAM dépendant à la fois du projet (faisabilité, disponibilité attendue, budget, planning) et de l'équipe</li> </ul>	
	<p><b>COÛT (RUN ET INVESTISSEMENT)</b></p> <ul style="list-style-type: none"> <li>Sur devis</li> </ul>	

11/2020

MTE - SNUM - Offre Socle technologique

20