**Plateforme de données Terreze**

**pour un territoire zéro carbone**

****

**Description du projet**

Il s’agit de déployer **une infrastructure de gestion des données du territoire** et d’un un portail web pour les exploiter. Cet outil nous aidera à **construire la trajectoire pour atteindre l’objectif de neutralité carbone** à l’horizon 2040. Cette plateforme logicielle appelée **Terreze** permettra de collecter, traiter et analyser des informations toujours plus utiles aux acteurs territoriaux. Elle pourra grandir et évoluer pour intégrer de nouvelles données et fonctionnalités.

**Le partage et la mise en circulation des données des acteurs du territoire** permet aux administrations, aux entreprises privées, aux associations et aux chercheurs d’accéder facilement à une grande diversité de données afin de produire des services performants, respectueux de la vie privée et de l’intérêt général.

Avec la plateforme, les membres du consortium LRTZC et les acteurs territoriaux pourront :

* **Piloter et évaluer** leurs actions par des indicateurs alimentés avec des données.

La valeur ajoutée de la plateforme réside dans la centralisation des données et référentiels nécessaires à la production des tableaux de bord, dans l’automatisation progressive de la collecte des données et l’actualisation continue des indicateurs.

* **Produire des services numériques** utiles, utilisables, utilisés qui soient véritablement vecteurs de progrès, basés sur des données publiques ou privées au plus près des besoins du territoire et des utilisateurs.

Des cas d’usage spécifiques à LRTZC seront réalisés (1 par axe, soit 10 en tout). Ils sont l’expression des besoins des membres du consortium et des différents axes du projet. Il s’agit de répondre à des problématiques concrètes par la collecte, le croisement et l’analyse de données. L’intérêt de la plateforme réside dans la fédération d'informations disparates liées à la multitude d'acteurs disposant chacun d'une partie des éléments et dans sa capacité à mettre en évidence les liens de causalité.

Les principaux éléments qui composent cette structure sont les suivants :

* + Une **infrastructure dédiée**, qui puisse stocker de manière sécurisée, performante et éco-responsable de grandes quantité de données. Il est prévu que cette infrastructure soit hébergée par l’université de La Rochelle dans un premier temps, puis au centre de données durable de Rochefort dans un second temps.
	+ Une **pile logicielle scalable** destinée à structurer les grands processus d'acquisition, de traitement, de mise en forme, de stockage et de sécurisation des données.
	+ Des outils pour **instrumenter l'analyse des données**, dans l’objectif de leur donner du sens (des analyses basiques jusqu’au bureau virtuel du data scientist).
	+ Un système de gestion temps réel des consentements et de protection des anonymats (RGPD *by design*).
	+ Un système de ***smart contracts*** destiné à faciliter le conventionnement entre producteur et réutilisateur de données : gestion du contrôle d’accès apposé aux jeux de données en *shared. data* et gestion des licences de propriété intellectuelle.
	+ Les garde-fous techniques garantissant **l’audibilité et l’explicabilité** du système (transparence algorithmique) selon les préconisations de l’ACPR.
	+ Un site internet "frontal" décliné sur trois profils d'utilisation, offrant chacun des fonctionnalités propres :
		- "Citoyen engagé" : destiné au particulier, il lui permet de s'informer sur l'avancement du projet de réduction des émissions carbone via diverses analyses et de contribuer activement en confiant ses données personnelles (en lien avec le projet de Self Data).
		- "Organisation engagée" : destiné aux personnes morales (entreprises, universités, collectivités, associations etc.), il leur permet de s'informer et de partager leurs données pour créer de nouveaux usages.
		- "Expert engagé" : destiné aux experts en science des données, il leur est permis d'accéder directement à un bureau virtuel directement hébergé sur la plate-forme pour leur permettre de travailler sur les données de manière sécurisée.
1. La deuxième composante, portant sur "le contenu", consiste à donner forme aux 11 cas d'usages demandés par les autres Axes du projet LRTZC et perçus comme créateurs de valeur. L’enjeu est de pouvoir mesurer, suivre et modéliser les indicateurs des actions de chaque axe de LRTZC pour une aide à la décision, et globalement de piloter l’atteinte des objectifs de consommation carbone sur le territoire.

La plateforme de données est un système permettant de **partager des données**, avec pour objectif de :

* Construire une **méta-plateforme** consolidant toutes les données publiques et privées du territoire ;
* Mettre en visibilité et communiquer sur les **moyens et les résultats quantitatifs** des efforts portés par les acteurs des différents chantiers de la démarche « [La Rochelle Territoire Zéro Carbone](https://www.agglo-larochelle.fr/projet-de-territoire/territoire-zero-carbone)» (LRTZC) au travers d’analyses basées sur des croisements de données qui soient éclairantes et communicatives ;
* Fédérer toutes les institutions publiques et privées du territoire dans une logique de gestion partagée des données à travers une gouvernance idoine et un **service public territorial des données** ;
* **Inviter les citoyens à s’engager** dans les projets en enrichissant des jeux de données avec leurs données personnelles tout en protégeant leur vie privée ;
* Favoriser l’émergence de **nouveaux services** (hors LRTZC) fondés sur des données du territoire ;
* Créer des **modèles descriptifs et prédictifs** pour éclairer la prise de décision, faciliter la coordination entre plusieurs parties prenantes ou encore optimiser des automatismes ;
* Minimiser et mesurer l’empreinte carbone de la plateforme par l’application de la **norme ISO 14062** relative à l’**éco-conception** d’un service numérique.

L’ambition du projet est également de pouvoir **capitaliser** sur les différents éléments développés pour se mettre en capacité de les dupliquer dans d’autres contextes territoriaux et à l’international, en s’appuyant sur les principes des Communs et de l’open source.

**L’action est construite sur 2 niveaux :**

A/ Ingénierie technique :

La structure élémentaire de la plateforme qui se veut avant tout :

* + **Générique,** pour pouvoir être répliquée et intégrer pour d'autres usages (voire à d’autres problématiques citoyennes) ;
	+ **Evolutive**, pour pouvoir s'enrichir à terme de nouvelles fonctionnalités.

B/ Ingénierie organisationnelle : la mise en place de la gouvernance des données doit permettre :

* De faire **l’inventaire des données territoriales** et créer un espace de débat et de décision pour identifier, accompagner et promouvoir des réutilisations à forte valeur ajoutée ;
* De penser la **mise en circulation des données** comme un nouveau service public d’intérêt général, accessible à tous les acteurs du territoire (individu, association, entreprise, institution) souhaitant, via ce guichet, obtenir avis, conseils et accompagnement techniques, juridiques, pratiques pour accéder ou diffuser des données : un service public territorial des données (SPTD) ;
* De se doter de **règles communes de régulation** fondées sur l’éthique et le respect de la vie privée.

La **gouvernance** est un prérequis à la réussite de la plateforme, notamment pour créer les conditions de la confiance pour le partage des données des uns et des autres.

Chaque cas d’usage fera l’objet d’une **analyse d’impact sur les données** préalable conformément à la méthodologie proposée par l’INRIA. Des métas-indicateurs seront également construits en sus et d’après des indicateurs sectoriels par un datascientist et un data-analyst.

Par ailleurs, l’infrastructure de données a vocation à **collecter des données** via de l’IoT, du crowdsourcing de données personnelles (Self Data territorial), des sources de données ouvertes, des données issues des SI Métiers des partenaires. Pour respecter les standards de l’**éco-conception** et de **sécurité**, il n’est pas attendu de la plateforme qu’elle duplique, pour autant qu’elle y donne accès, l’ensemble des informations du territoire. L’analyse préparatoire des besoins permettra de ne collecter que les données nécessaires. L’alimentation de la plateforme implique un nettoyage des données et des traitements intermédiaires notamment d’**anonymisation** ou de **standardisation** avant stockage. Une attention particulière sera portée à l’évolutivité du système à travers l’étude et l’implémentation d’ontologies.



**Objectifs du projet**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|

|  |
| --- |
| * Mettre en place d’une **infrastructure de collecte et de stockage** de données administrée et sécurisée (entreprises, institutionnels, universitaires, particuliers, associations ... ) et concevoir une **plateforme transversale, open source et réplicable**
 |
| * Mesurer, suivre et modéliser des **indicateurs** des actions de chaque levier de LRTZC pour une **aide à la décision** et permettre un **pilotage prédictif** grâce aux *data sciences* et au *data mining*
 |
| * Prouver que l’**éco-conception** et le **privacy by design** sont possibles pour une plateforme de données de cette ampleur
 |
| * Créer un **écosystème** en fédérant des acteurs variés et créer une **communauté** favorisant la diffusion de la connaissance et des bonnes pratiques en autonomie
 |
| * Impliquer les usagers dans les politiques publiques via le **partage** de leurs données (programme Self Data) et renforcer la confiance entre le citoyen et les acteurs institutionnels.
 |
| * Démocratiser la **culture de la donnée**
 |
| * Penser la circulation des données sur le territoire comme un nouveau service public à part entière faisant l’objet de politiques publiques spécifiques. La création d’un guichet accessible à tous, le **service public territorial des données**, sera une des matérialisations de la gouvernance territoriale des données
 |
| * Recenser l’ensemble des **données disponibles** sur le territoire, les normaliser et les structurer
 |
| * Intégrer la dimension des **Communs** comme gage de pérennité de la métaplateforme
 |
| * Développer des fonctionnalités inédites, telles que l’utilisation des **smartcontract** pour la contractualisation du partage de données directement de réutilisateur à producteur.
 |
| * Développer des CGU selon une approche de **legal design**
 |
| * Développer des **ontologies** pour favoriser l’**accessibilité** pérenne et économe en énergie des jeux de données
 |

 |