





# Sommaire

<b>1</b>	<b>Synthèse</b>	<b>4</b>
<b>2</b>	<b>La solution proposée</b>	<b>7</b>
2.1	Périmètre, produits, services visés par le projet	7
2.1.1	<i>Le portail</i>	7
2.1.2	<i>La gouvernance</i>	8
2.1.3	<i>Le groupe d'experts</i>	9
2.1.4	<i>Le cercle des producteurs et réutilisateurs</i>	9
2.2	Description détaillée de la solution adoptée, innovations et ruptures technologiques impliquées	9
2.2.1	<i>Innovation organisationnelle</i>	9
2.2.2	<i>Innovation technique</i>	9
2.3	Etat des lieux, enjeux et difficultés présentes	11
2.4	Attentes des acteurs concernés (usagers, administrations)	11
2.5	Conditions et facteurs clés de succès / risques principaux	11
2.6	Impact attendu en termes de modernisation de l'action publique	12
<b>3</b>	<b>Méthode retenue</b>	<b>13</b>
3.1	Présentation des entités porteuses de projets et de la pertinence du partenariat	13
3.2	Principales étapes et méthodologie retenue pour mener le projet	13
3.2.1	<i>Phase 1 – Amorçage</i>	15
3.2.2	<i>Phase 2 – Préindustrialisation et lancement</i>	16
3.2.3	<i>Phase 3 – Extension de la démarche et déploiement</i>	16
3.3	Composition de l'équipe (acteurs et compétences mobilisées)	17
3.3.1	<i>Partenaires</i>	17
3.3.2	<i>Cercle des producteurs et réutilisateurs sur la filière de l'eau</i>	19
3.4	Terrains d'expérimentation pressentis	19
<b>4</b>	<b>Actions prévues pour pérenniser la solution en cas de succès</b>	<b>20</b>
4.1	Gouvernance du projet	20
4.2	Acteurs institutionnels mobilisés pour pérenniser le projet en cas d'expérimentation réussie	20
4.3	Documentation prévue du projet en cours de réalisation	20
<b>5</b>	<b>Résultats attendus</b>	<b>21</b>
5.1	Gains potentiels estimés en termes de qualité de service	21
5.2	Impact financier et économique potentiel	21
5.3	Autres résultats obtenus par le projet	21



# 1 Synthèse

Piloté par le SGAR de Midi-Pyrénées, avec une implication forte des services régionaux de l'état et un partenariat avec de nombreuses collectivités locales, le projet **Open Datalab vise à développer des pratiques innovantes d'animation et de gouvernance autour de l'Open Data pour maximiser les réutilisations** et ainsi favoriser tout autant le développement économique, la transparence de la vie publique et la simplification de l'administration. Par ses aspects innovants, Open Datalab constitue ainsi le laboratoire de la valorisation des données publiques, dont les expérimentations et les résultats seront portés à la connaissance de toutes les autres régions dans une démarche de partage des bonnes pratiques. Cette approche répond précisément au cadre défini par le mandat d'élaboration des CPER, justifiant ainsi le soutien et le portage institutionnel apportés par le CGET.

## Favoriser et maximiser la création de valeur en levant les freins aux réutilisations

Avec l'ouverture grandissante des données publiques, il apparaît de plus en plus évident que la simple mise à disposition de données brutes ne suffit pas à déclencher les réutilisations et les créations spontanées d'applications souvent mises en avant dans le cadre des projets d'ouverture de données.

L'objectif du laboratoire de la valorisation des données publiques est de lever les freins à la réutilisation en expérimentant des solutions dans le cadre de l'ouverture des données sur la région Midi-Pyrénées. Ces freins à la réutilisation des données publiques ouvertes s'expliquent notamment par :

- La non prise en compte des besoins des utilisateurs
- Le faible intérêt des données ouvertes
- Le manque de standardisation du mode de diffusion
- L'utilisation de standards trop spécialisés ou inadaptés
- La complexité et le manque de documentation des modèles de données
- La mise à jour sporadique des jeux de données
- La trop faible qualité des données

Par ailleurs, le projet va chercher à démultiplier l'ouverture des données à toutes les échelles du territoire. Pour cela, il doit prendre également en compte la méconnaissance par les petites et moyennes collectivités territoriales de l'Open Data, notamment de ses aspects techniques et juridiques. Ce point est d'autant plus important compte tenu de la loi NOTRe, portant obligation pour les collectivités locales de plus de 3500 habitants et leurs EPCI de mettre en ligne leurs documents publics sous forme électronique.

## Constituer le laboratoire de valorisation des données publiques

Le projet Open Datalab vise à créer et animer ce laboratoire et est constitué de trois volets principaux :

- **La mise en place d'un portail de mise à disposition et de valorisation des données publiques de la région.** Ce portail permet de recenser et de présenter les données publiques de Midi-Pyrénées qui sont stockées soit sur un portail déjà existant, soit sur le portail géré par la mission Etalab : [data.gouv.fr](http://data.gouv.fr)
- **Le support aux collectivités territoriales pour l'ouverture et la valorisation de leurs données publiques.** Ce support comprend plusieurs éléments, activés selon les besoins et demandes des collectivités :
  - des conseils juridiques pour définir le périmètre des données publiques qu'il serait possible d'ouvrir,
  - des conseils techniques pour héberger les données soit sur [data.gouv.fr](http://data.gouv.fr), soit sur une plateforme propre, soit sur une plateforme déjà existante,



- des outils techniques (sous forme d'API) et les conseils associés pour permettre d'afficher et de valoriser, sur un site internet propre à la collectivité, les données publiques relatives à cette collectivité et hébergées sur d'autres plateformes
- des conseils organisationnels pour structurer les démarches Open Data de ces collectivités

➤ **La définition et l'animation d'une gouvernance avec l'ensemble des acteurs nécessaires pour assurer la meilleure production et valorisation des données** ; cette gouvernance s'attache notamment à structurer les données et les rendre plus pertinentes par rapport à des réutilisations envisagées. Cela passe notamment par l'écoute continue des besoins et des demandes que peuvent exprimer la société civile, des professionnels, mais aussi des organismes publics réutilisateurs de données ...

### Organiser l'ouverture des données par filière en mode projet pour maximiser les possibilités de réutilisation et le développement de services innovants

Afin de tester les concepts et principes proposés, **le projet sera développé de façon itérative, en mode startup d'état**. Ainsi, il est proposé de mener deux actions d'ouverture des données en parallèle :

- Une action de mise à disposition des données publiques disponibles, tant au niveau des services de l'Etat qu'au niveau des collectivités, afin de rendre accessible auprès du grand public un maximum de données, sans préjuger des réutilisations qui pourront en être faites (ouverture large, quantité de jeux de données)
- **Une action de structuration des données par filière, afin de fédérer et coordonner les acteurs d'une filière, en mode projet**, pour identifier les données les plus pertinentes, les réutilisations permettant de faire évoluer de façon significative l'activité de la filière et de faire émerger les standards et règles de gouvernance permettant d'assurer la meilleure interopérabilité possible des données (structuration et qualité des données).



**Schéma de principe de l'ouverture large et de la structuration par filière des données**

Enfin, ces actions d'ouverture seront accompagnées des communications et événements nécessaires afin de faire connaître les données, initier des réutilisations, faire se rencontrer les acteurs...

### Cofinancer le projet avec les parties prenantes pour assurer leur engagement

Le plan de financement mis en place propose d'associer les acteurs majeurs sur les deux actions :

- l'Agence de l'Eau pour participer à la structuration de la filière eau, proposée comme première filière pour expérimenter et mettre en œuvre la démarche avant reproduction sur les autres filières,
- la métropole de Toulouse, acteur en pointe sur le sujet Open Data au niveau de la région et du pays, pour apporter son expertise dans les démarches d'ouverture des collectivités territoriales et services de l'état.



Par sa structuration en mode projet, filière par filière, le projet Open Datalab assure une **ouverture des données cohérente et harmonisée** à l'échelle de la région Midi-Pyrénées et permet de maximiser les possibilités de réutilisation et de développement de services innovants.

Il permet par ailleurs de **développer une méthode de travail reproductible** qui pourra ainsi être utilisée partout en France, à commencer par les territoires de la nouvelle région.

Enfin, par l'approche technique proposée, **Open Datalab contribuera d'une part à l'alimentation de la plateforme data.gouv.fr** avec l'ensemble des données publiques de Midi-Pyrénées et d'autre part à la possibilité de **consulter ces données à n'importe quel niveau du maillage territorial**.



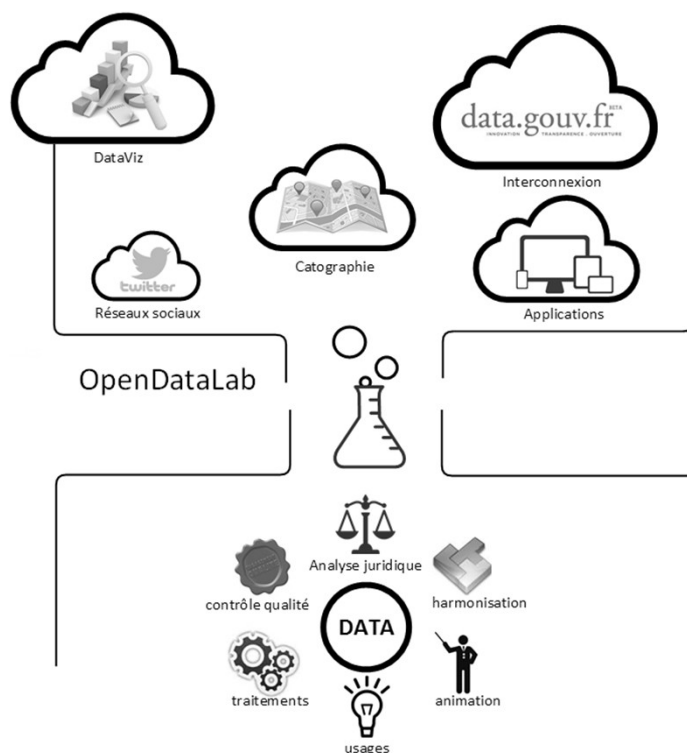
## 2 La solution proposée

### 2.1 Périmètre, produits, services visés par le projet

Ce laboratoire de la valorisation des données publiques sera constitué de quatre éléments majeurs :

- Un portail regroupant les fonctionnalités (API) de mise à disposition et de valorisation des données et des réutilisations, mais aussi assurant la capitalisation autour des expériences acquises à travers le projet (questions-réponses, études de cas, success stories, guides pratiques...)
- Une gouvernance pour cadrer et animer le projet
- Un groupe d'experts pour assurer et aider la mise en œuvre
- Un cercle des producteurs et réutilisateurs qui sera à géométrie variable en fonction des filières traitées

A terme, ce sont toutes les données publiques qui pourront être ouvertes et rendues disponibles. Le laboratoire se concentrera néanmoins dans un premier temps sur la structuration des données portant sur **le thème de l'eau**.



**Schéma de principe du laboratoire**

#### 2.1.1 Le portail

Un portail permettra de recenser, mettre à disposition et valoriser les données publiques stockées sur la plateforme nationale (data.gouv.fr).

Ce portail n'a pas vocation à héberger des données pour les collectivités locales ne disposant pas de l'infrastructure nécessaire. A la place, il leur sera proposé une aide technique pour publier leurs données sur la plateforme data.gouv.fr. En effet il nous apparaît comme primordiale d'inciter au maximum les organismes publics à utiliser la plateforme d'Etalab.



## **L'API Catalogue**

Afin de pouvoir consulter l'ensemble des données publiques disponibles au niveau de Midi-Pyrénées, le portail s'appuiera sur une API javascript, développée dans le cadre du projet et qui utilisera l'API d'Etalab. Le but est que le fonctionnement de cette API soit testé en situation réelle via le portail du projet. Ensuite cette API sera mise à disposition des collectivités territoriales (voir point 2.2.2).

## **L'API SIG**

De plus une seconde API sera développée pour s'interfacer avec la plateforme SIG régionale de l'Etat : MiPyGéo. Cette dernière API permettra d'afficher des cartes thématiques selon un thème et/ou une granularité particulière. Dans un deuxième temps, cette API sera elle aussi mise à disposition des collectivités territoriales souhaitant intégrer des données cartographiques provenant de MiPyGéo (voir point 2.2.2).

A terme, et grâce à ces API, ce portail jouera ainsi un rôle de condensateur de l'ensemble des données publiques de la région Midi-Pyrénées.

## **La capitalisation des connaissances**

Le portail capitalisera sur les réflexions et évolutions du portail data.gouv.fr en termes de fonctionnalités sociales pour renforcer l'animation autour des données et des réutilisations. Il permettra aussi de formaliser l'ensemble du savoir qui aura émergé du projet notamment à travers les retours d'expériences réalisés :

- Série de questions-réponses (FAQ)
- Étude de cas concrets
- Valorisation des succès obtenus (success stories)
- Témoignages
- Guides pratiques (dans la mesure où ils n'existent pas déjà)

Cette capitalisation des connaissances est particulièrement importante dans une optique de répliquabilité du projet et de pérennisation des actions.

## **2.1.2 La gouvernance**

La gouvernance du projet est structurée autour :

➤ d'un **comité de pilotage** constitué des parties les plus avancées et / ou les plus moteurs dans la démarche. Il sera fait en sorte que les membres de ce comité de pilotage soient d'origines diverses. Idéalement il comprendra des entités venant des services de l'État, des collectivités territoriales, de la société civile et des entreprises privées (startup et PME). Cette approche novatrice dans la constitution du comité de pilotage devrait permettre d'élargir davantage la vision open data du laboratoire créé. Le CGET, en tant qu'entité porteuse du projet, sera présent au comité de pilotage.

Le Comité de pilotage intégrera dans sa démarche à la fois les aspects organisationnels et les aspects scientifiques liés à la thématique choisie (l'eau). Pour les aspects scientifiques liés aux cycles de l'eau, il s'appuiera donc sur le Pôle de Compétitivité mondial de l'eau, présent à Montpellier. Le Pôle de Compétitivité a déjà manifesté son intérêt pour le projet (voir point 2.3.1).

➤ d'un **chef de projet** qui assure l'animation globale du projet et des acteurs et gère notamment les phases projet visant à structurer les données et réutilisations pour une filière.

➤ d'une **assistance à maîtrise d'ouvrage** qui anime et accompagne les réflexions sur la vision stratégique du projet, la gouvernance globale et les modes de fonctionnement à mettre en place.





### 2.1.3 Le groupe d'experts

Le groupe d'experts constitue l'équipe de maîtrise d'œuvre en charge d'orchestrer les actions, de faire le relais avec les producteurs de données et de garantir la livraison des réalisations ; cette équipe s'appuie sur un réseau d'experts techniques et juridiques identifiés progressivement au sein des partenaires du projet, de structures nationales, et de sous-traitants.

Ce groupe d'experts a aussi un rôle très important dans l'accompagnement des collectivités territoriales dans leur démarche d'ouverture des données. Il permettra d'informer les collectivités sur les aspects juridiques de l'open data mais aussi de les accompagner d'un point de vue technique.

### 2.1.4 Le cercle des producteurs et réutilisateurs

Un cercle des producteurs et réutilisateurs est mis en place pour chaque projet visant à organiser et structurer la mise à disposition des données publiques pour une filière. Il sera animé via des réunions régulières et via l'organisation de BarCamp dont le rôle sera de faire émerger les données à ouvrir en priorité et selon quelle méthode (fichier brut, API, Webservice).

Le rôle de ce cercle est d'alimenter le laboratoire en pistes d'actions mais aussi d'être critique envers les productions de données. Ce groupe devra nécessairement comporter des réutilisateurs d'horizons divers : services de l'Etat, collectivités territoriales, DSP, société civile, pôle de compétitivité, universités, et entreprises privées (startup, bureaux d'étude, PME)

Il sera également étudié de près la possibilité d'intégrer les citoyens à la dynamique d'amélioration continue des données, notamment via du crowdsourcing.

## 2.2 Description détaillée de la solution adoptée, innovations et ruptures technologiques impliquées

### 2.2.1 Innovation organisationnelle

Par rapport aux initiatives d'ouverture de données traditionnelles visant à mettre à disposition toutes les données publiques disponibles, le projet Open DataLab vise à ajouter une approche complémentaire : le cadrage et la structuration de l'ouverture de données par filière en mode projet.

Ce mode projet permet de fédérer et mobiliser les différents acteurs d'une filière, producteurs et réutilisateurs potentiels et conduit ainsi à :

- Identifier les besoins et demandes de la société civile, des professionnels, des experts, de l'administration en termes de données (quantité et qualité)
- Faire émerger les standards et formats de données permettant la meilleure interopérabilité entre les différents acteurs
- Co-construire les modes de fonctionnement entre les acteurs pour produire, mettre à disposition, qualifier et valoriser les données

Ce mode projet se traduit par la définition, pour chaque filière traitée, d'un plan projet précisant les acteurs concernés, la gouvernance, le planning et les livrables attendus.

Le projet s'appuie sur la conviction que cette approche complémentaire en mode projet par filière est le meilleur levier pour maximiser les réutilisations et la création de valeur à travers les données publiques.

Enfin cette méthode de travail ne peut fonctionner entre les seuls services de l'Etat et une réelle volonté sera présente pour inclure tout type d'organisation dans la démarche : services de l'Etat, collectivités territoriales, DSP, société civile et entreprises privées (startup et PME).

### 2.2.2 Innovation technique

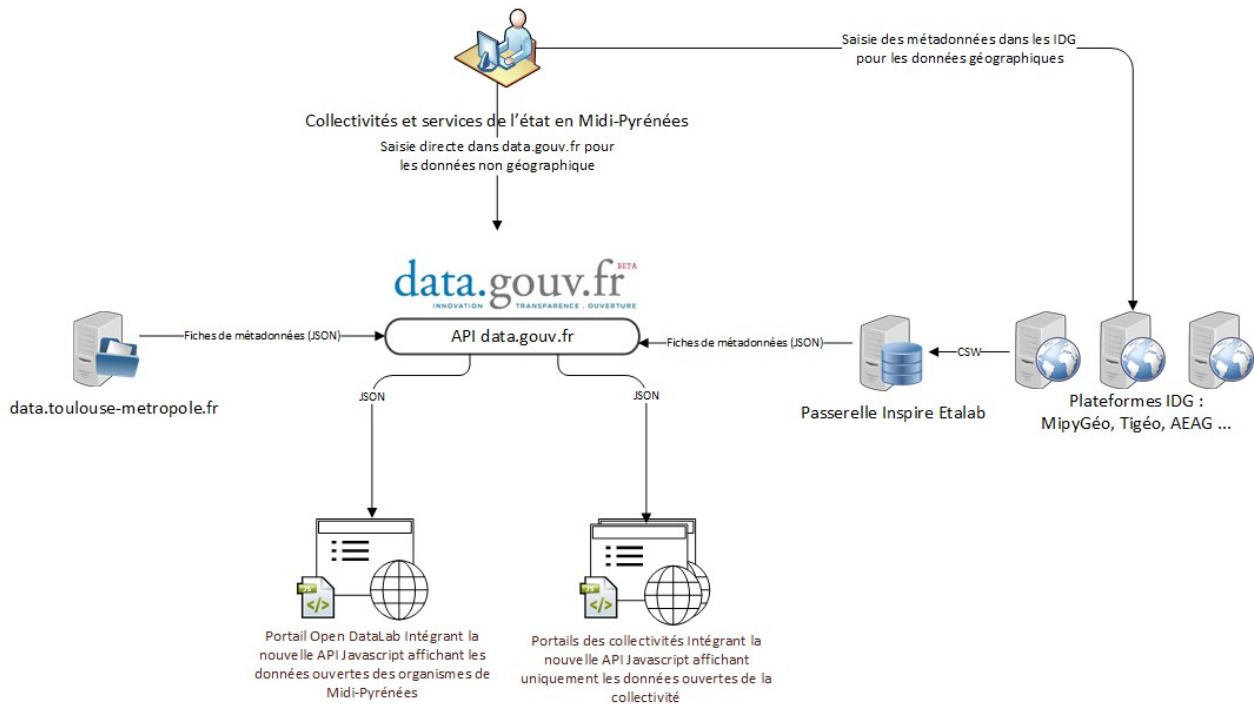
Au niveau technique, le projet propose de travailler sur plusieurs axes :

- Le développement d'une API javascript basée sur l'API de mise à disposition des données de data.gouv.fr permettant de rechercher et d'afficher les jeux de données ouverts, filtrés selon diverses métadonnées (organismes ou territoire par exemple). Cette API sera utilisée sur le portail régional mais sera également librement utilisable pour être intégrée sur d'autres sites Internet. Ce développement réalisé en Open Source suivra les évolutions des fonctionnalités notamment



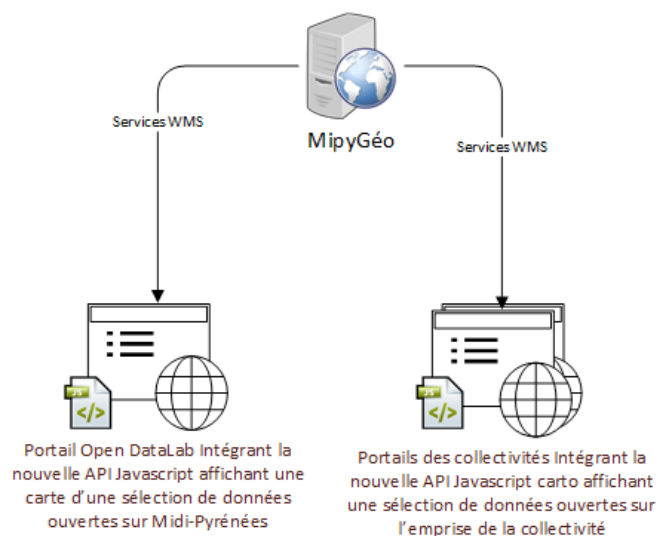
sociales apportées à data.gouv.fr au fil de l'eau. La démarche technique pour une collectivité souhaitant ouvrir ses données sans pour autant les héberger sera donc la suivante :

1. Mise à disposition des données sur data.gouv.fr
2. Intégration de l'API développée pour afficher sur le site internet de la collectivité son catalogue de données présent sur data.gouv.fr (pas de doublons des jeux de données car les données ne sont stockées qu'une fois sur data.gouv.fr)



**Schéma illustrant le fonctionnement de l'API Catalogue**

➤ Le développement d'une API javascript cartographique très simple permettant aux partenaires du projet d'afficher une carte intégrant une sélection de données ouvertes sur leurs propres sites Internet. Prenons l'exemple d'une mairie souhaitant afficher les chemins de randonnées sur son territoire. Si la mairie n'a pas les compétences ni les moyens de mettre en place des outils SIG, elle pourra fournir les données à MiPyGéo. Une fois les données intégrées sur la plateforme MiPyGéo, le webmaster du site internet de la mairie pourra très simplement afficher la carte MiPyGéo sur le site de la collectivité. Lorsque les données seront mises à jour sur MiPyGéo, la carte sur le site de la collectivité sera automatiquement à jour elle aussi.



**Schéma illustrant le fonctionnement de l'API SIG**



➤ L'harmonisation des formats de données dans l'optique de maximiser l'interopérabilité et la mise en place de solutions d'automatisation et d'industrialisation du maintien et de la mise à jour des données. Cela pourra se faire via des développements spécifiques qui seront définis en fonction des besoins émergents du projet (POC – Proof of concept).

➤ Un soutien personnalisé et local sur le plan technique et juridique pour aider les collectivités locales dans leur démarche d'ouverture des données.

La mise en place de projets par filière permet de faciliter l'émergence de formats de données interopérables et de solutions d'industrialisation des traitements sur les données.

Par ailleurs, le portail et les API seront développés, autant que possible, sur la base de solutions open source, permettant ainsi de faire progresser également ces solutions.

## 2.3 Etat des lieux, enjeux et difficultés présentes

Un tel projet peut s'appuyer sur les atouts identifiés au niveau de la région :

➤ Une expérience et des compétences acquises, notamment dans l'animation d'acteurs divers, à travers le portail de données géographiques MiPyGeo

➤ Des acteurs publics à l'échelle de la région volontaires dans la démarche d'ouverture des données

➤ Des acteurs privés en pointe dans le numérique et l'innovation comme en témoigne notamment la labellisation French Tech, l'écosystème d'incubateurs et d'associations telles que La Mêlée, TIC Valley, DigitalPlace, La Cantine...

➤ La volonté de participation au projet de Toulouse Métropole qui possède déjà une très bonne expérience Open Data (création de leur plateforme, appui à l'association Open Data France, ...)

➤ Des thèmes porteurs pour expérimenter l'ouverture des données, notamment autour de l'eau, thème pour lequel le contexte actuel (COP21) fait naître des besoins d'informations et de données et pour lequel il existe des acteurs moteurs (tels que l'Agence de l'eau)

De plus, le projet correspond aux attentes du volet numérique de la circulaire CPER : organisation de la gouvernance et des actions sur les plates-formes de données, les SIG, et l'Open Data.

## 2.4 Attentes des acteurs concernés (usagers, administrations)

Le projet doit permettre de répondre à plusieurs grandes attentes :

➤ La simplification des échanges et le partage d'information pour les administrations

➤ La demande d'information juridique et technique sur l'open data de la part des petites collectivités

➤ La visibilité accrue sur les données pour les citoyens et les administrations

➤ La transparence demandée par la société civile et les citoyens (traçabilité des données)

➤ Un interlocuteur local pour guider de manière personnalisée les administrations et les collectivités

## 2.5 Conditions et facteurs clés de succès / risques principaux

Le succès du projet réside en grande partie dans la conjonction des volontés de tous les acteurs concernés et l'envie d'avancer sur un projet en mode start-up d'état. Les modes de fonctionnement innovants proposés doivent permettre de donner un nouveau souffle à des démarches d'ouverture qui peuvent parfois s'essouffler quand elles sont menées avec des approches plus habituelles.

Les hypothèses qui permettront d'assurer un succès de l'action sont :

➤ Pas de blocage des producteurs pour les demandes d'ouverture de données des utilisateurs, dès lors que celles-ci rentrent dans le champ de l'open data

➤ Système de métadonnées de data.gouv.fr cohérent avec la démarche de structuration



- Bonne coopération entre les producteurs et utilisateurs venant de différents milieux (public, privé et associatif)
- Pas de blocage de la part des élus des collectivités
- Bonne inter-opérabilité avec le plugin data.gouv.fr
- Bonne inter-opérabilité avec le système MiPyGéo pour la partie SIG

Ces risques ont été limités entre autres par les actions suivantes :

- Organisation d'une première réunion en novembre 2014, puis de deux réunions en mars 2015 afin de vérifier la volonté des producteurs de données de participer à la démarche
- Prise de contact anticipée avec les développeurs d'Etalab via un cabinet d'étude avec qui le SGAR a un marché en cours
- Implication forte du chef de projet MiPyGéo qui a personnellement contribué à l'élaboration de ce projet

## 2.6 Impact attendu en termes de modernisation de l'action publique

L'impact en termes de modernisation de l'action publique se concrétise sur plusieurs axes :

- Développer la sensibilité Open Data pour accélérer l'ouverture des données et la réutilisation par les différents acteurs. L'administration elle-même peut et devrait être un réutilisateur des données. En effet, l'un des premiers bénéficiaires de l'ouverture des données est la simplification des échanges de données entre administrations.
- Aider les collectivités territoriales à ouvrir et publier leurs données. Ce point est d'autant plus important pour permettre aux collectivités territoriales de taille modeste, et ne bénéficiant pas nécessairement de plateformes techniques, de mettre également à disposition leurs données et ainsi pouvoir respecter l'amendement du 5 mars 2015 de la loi NOTRe leur donnant obligation de publier leurs données publiques dès lors qu'elles existent au format électronique.
- Rendre plus pertinentes les données mises à disposition des services de l'Etat en fonction des retours des usagers. L'action publique sera ainsi plus clairement comprise des citoyens.
- Améliorer la plateforme data.gouv.fr en ayant un dialogue continu permettant d'identifier des axes d'amélioration de la plateforme nationale en fonction des constatations réalisées dans le projet
- Améliorer le service public rendu aux citoyens (amélioration de la qualité de l'eau, prévention des crues, signalement de fuites...)



# 3 Méthode retenue

## 3.1 Présentation des entités porteuses de projets et de la pertinence du partenariat (le cas échéant)

L'entité porteuse du projet est le Commissariat général à l'égalité des territoires (CGET).

Le porteur opérationnel du projet est le Secrétariat général aux affaires régionales de la Préfecture de Midi-Pyrénées.

Ces deux entités étant inter-ministérielles, le projet pourra donc toucher un maximum de services de l'État et avoir ainsi un fort coefficient multiplicateur.

## 3.2 Principales étapes et méthodologie retenue pour mener le projet (calendrier prévisionnel, jalons et résultats clés...)

La méthodologie retenue se basera sur le cadre logique. Cette matrice du cadre logique permet d'identifier les rapports logiques entre le problème global, les objectifs spécifiques, les résultats anticipés et les activités. Le cadre logique facilite l'analyse de ces rapports et de ceux qui les relient à l'environnement. Le cadre logique explicite également des indicateurs qui serviront tout au long du projet à formaliser des tableaux de bords permettant d'apprécier l'avancée du projet.

Cette méthodologie a clairement comme but d'orienter le projet vers l'obtention de résultats atteignables et mesurables.

Vous trouverez à la page suivante la matrice de cadre logique du projet.



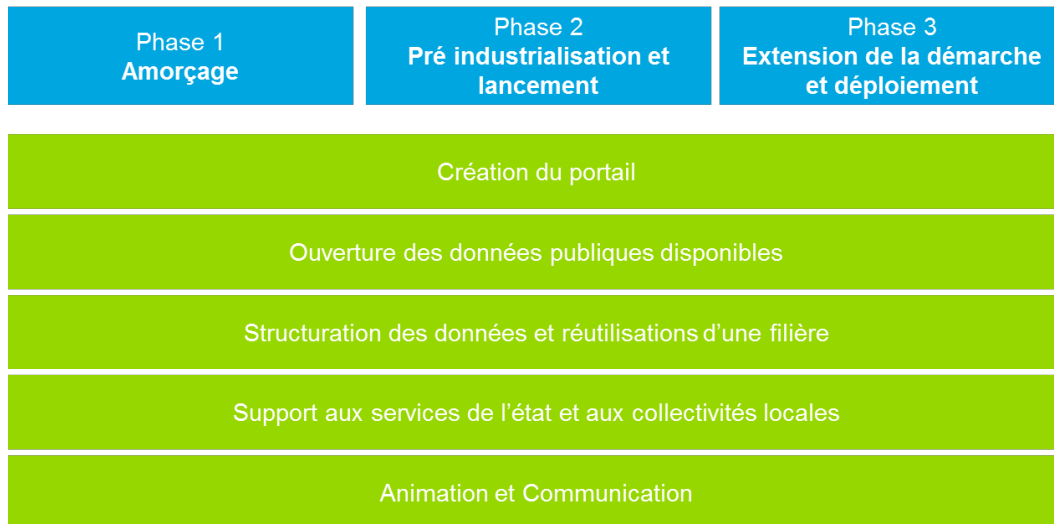
## Matrice de cadre logique

Objectif principal	Industrialiser la mise à disposition de données ouvertes		
Objectif du projet	Appuyer de manière innovante les producteurs de données sur le plan quantitatif et qualitatif		
Résultat 1 : Structurer la production des données publiques en mettant en relation les producteurs et les utilisateurs dans une démarche d'amélioration continue	Indicateurs : - nombre de nouvelles données publiques mise en ligne sur data.gouv.fr - nombre de données en ligne avec la métadonnée Midi Pyrénées - nombre de réunions organisées entre producteurs et utilisateurs de données - nombre de demandes de changement formalisées par les utilisateurs - augmentation de la pertinence des données pour les utilisateurs	Sources de vérification : - Portail data.gouv.fr - Portail du projet - Comptes-rendus de réunions - Enquêtes sur la pertinence des données réalisées en début et en fin de projet	Hypothèses : - Pas de blocage des producteurs pour les demandes d'ouverture de données des utilisateurs - Système de métadonnées de data.gouv.fr cohérent avec la démarche de structuration - Bonne coopération entre les producteurs et utilisateurs venant de différents milieux (public, privé et associatif)
Résultat 2 : Renforcer les capacités des collectivités territoriales de plus de 3500 habitants en matière d'ouverture de données publiques	Indicateurs : - nombre de collectivités aidées dans leur démarche d'ouverture de données - nombre de nouvelles collectivités de Midi Pyrénées présentes sur data.gouv.fr - satisfaction des collectivités appuyées - retours d'expériences partagés sur le portail du projet - mise en production de l'API catalogue permettant d'importer les données de data.gouv.fr sur un site de collectivité - mise en production de l'API SIG permettant d'importer des cartes thématiques sur un site de collectivité	Sources de vérification : - Portail data.gouv.fr - Portail du projet - Comptes-rendus de réunions - Enquête de satisfaction auprès des collectivités appuyées	Hypothèses : - Pas de blocage de la part des élus des petites collectivités - Bonne inter-opérabilité avec le plugin data.gouv.fr - Bonne inter-opérabilité avec le système MiPyGéo pour la partie SIG
Activité 1 : Recensement des données à ouvrir dans les services de l'Etat et aide à l'ouverture sur data.gouv.fr	Moyens sur 18 mois:  - Ingénierie de projet (1 chef de projet dédié) - Assistance maîtrise d'ouvrage / prestations intellectuelles – 1,5 j/sem pendant 12 mois - Assistance maîtrise d'œuvre / prestations techniques - Expertises pour l'appui aux collectivités et aux services de l'Etat - Communication autour du projet et événements - Equipements informatiques		
Activité 2 : Structuration des données de manière itérative en se focalisant sur une thématique choisie (l'eau) et en confrontant de manière continue les producteurs et utilisateurs			
Activité 3 : Appui aux collectivités territoriales de plus de 3500 habitants pour qu'elles deviennent productrices de données sur data.gouv.fr			
Activité 4 : Réflexion sur la pérennisation des résultats et capitalisation des connaissances autour des premières expériences avec les collectivités et avec le cercle d'utilisateurs/producteurs. Partage de ces retours d'expériences (sous forme de FAQ, d'études de cas, de success stories ou de guide pratique)			
Activité 5 : Développement et support du portail du projet et des deux API			
Activité 6 : Communication autour du projet : organisation de 2 bar camp (ateliers-événements participatifs) et d'un événement présentant les résultats obtenus en fin de projet			

Le plan de travail est découpé en quatre thèmes et un volet d'animation et de communication transverse. Les quatre thèmes sont :

- La création du portail et des API. Ce thème comprend également la capitalisation des connaissances du projet (activités 4, 5 et 6 du cadre logique)
- L'ouverture des données publiques disponibles (activité 1)
- La structuration des données et réutilisations d'une filière (activité 2)
- Le support aux services de l'État et aux collectivités locales (activité 3)

Ce plan de travail s'échelonne sur 18 mois, durée qui semble la plus appropriée pour mobiliser un maximum d'acteurs autour des objectifs fixés.



*Aperçu schématique de la méthode retenue*

### 3.2.1 Phase 1 – Amorçage

- La création du portail et des API
  - Création d'un portail Beta
- L'ouverture des données publiques disponibles
  - Recensement des données publiques disponibles dans les différents services de l'État
- La structuration des données et réutilisations d'une filière
  - Création du cercle des producteurs et réutilisateurs sur la filière de l'eau
  - Identification des données nécessaires sur la filière de l'eau
  - Réflexion autour d'une large implication des citoyens (crowdsourcing)
- Support aux services de l'État et aux collectivités locales
  - Entretiens / rencontres avec les services et collectivités pour présenter le projet et comprendre les éventuelles difficultés rencontrées par ces acteurs
  - Identification des données publiques disponibles
  - Définition détaillée des supports à mettre en place pour accompagner ces services et collectivités
- Animation et communication transverse
  - Animation régulière d'un Comité Technique regroupant les représentants, au niveau de la région, des services de l'État et des collectivités territoriales. Ce Comité



Technique est le lieu privilégié pour informer et sensibiliser sur les enjeux de l'Open Data et sur le projet Open DataLab en particulier. Il permet également d'élargir progressivement le réseau de partenaires impliqués dans le projet.

- Mise en relation avec les acteurs locaux de l'économie numérique, notamment les associations telles que La Mêlée, TIC valley, DigitalPlace, pour permettre de faire connaître le projet encore plus largement et d'impliquer progressivement le secteur privé notamment en tant que réutilisateur de données.
- Mise en relation avec la société civile

### 3.2.2 Phase 2 – Préindustrialisation et lancement

- La création du portail
  - Lancement du portail
  - Mise en place d'outils de contrôle qualité des données et des API
  - Développement de fonctionnalités sociales pour faire connaître et rayonner les données et réutilisations mises à disposition
- L'ouverture des données publiques disponibles
  - Aide à l'ouverture via data.gouv.fr
  - Lancement du portail avec une première vague de jeux de données ainsi que les données et réutilisations déjà préparées sur la filière de l'eau
- La structuration des données et réutilisations d'une filière
  - Identification des standards et formats de données pour assurer l'interopérabilité
- Support aux services de l'État et aux collectivités locales
  - Formalisation de la démarche d'appui aux collectivités (juridique et technique) et mise à disposition de ce soutien aux services de l'État et aux collectivités
  - Accompagnement de quelques services de l'État et collectivités pour leur ouverture de données. Les collectivités « pilotes » seront choisies de manière à être le plus représentatif du paysage des collectivités en France (nombre d'habitants, territoire urbain, ruraux, montagne...)
- Animation et communication transverse
  - Organisation de concours pour créer les premières réutilisations sur la filière de l'eau. Il sera envisagé à chaque fois de s'adosser à des concours déjà existants sur le territoire afin de ne pas multiplier les événements et risquer de diluer l'implication des acteurs
  - Développement de moyens pour « écouter » les besoins et demandes de la part des producteurs et utilisateurs (société civile, professionnels...)
  - Première évaluation de la démarche pour préparer son extension

### 3.2.3 Phase 3 – Extension de la démarche et déploiement

- **Réflexions autour de nouvelles stratégies pour la suite du projet :**  
Lors de cette dernière phase un groupe de travail sera constitué pour définir la stratégie post projet :
  - stratégie sur les retombées économiques induites par l'ouverture des données...
  - stratégie sur le caractère reproductible des activités du projet et mise en place d'activités tendant à démontrer cette reproductibilité
  - Formalisation d'un cadre générique de structuration des données issu de l'expérience acquise sur la thématique « eau »
  - Production d'un référentiel méthodologique documenté partageable par tous les autres SGAR et toutes les collectivités intéressées à l'échelon national.





- La création du portail
  - Capitalisation des retours d'expériences
  - Mise en place de POC (Proof of Concept)
- L'ouverture des données publiques disponibles
  - Elargissement du recensement des données publiques et mise à disposition à travers data.gouv.fr
- La structuration des données et réutilisations d'une filière
  - Identification d'une nouvelle filière à traiter
  - Support aux services de l'État et aux collectivités locales
  - Mise en place de la répliquabilité auprès d'autres collectivités
  - Accompagnement des services de l'État et des collectivités vers une autonomisation
- Animation et communication transverse
  - Développement des événements et concours pour dynamiser les réutilisations et faire se rencontrer les acteurs
  - Seconde évaluation de la démarche

### 3.3 Composition de l'équipe (acteurs et compétences mobilisées)

#### 3.3.1 Partenaires

Le projet rassemble un ensemble large de partenaires et peut également s'appuyer sur un réseau dynamique sur les sujets numériques et Open Data. En effet, la métropole de Toulouse a été à l'initiative de la création de l'association Open Data France et a obtenu le label French Tech.

Par ailleurs, de nombreux partenaires potentiels ont été approchés et se sont dits intéressés par le projet. Voici la liste de ces partenaires qui a vocation à être élargie une fois le projet lancé :

Nom du partenaire potentiel	Citation du partenaire potentiel sur le projet	Contact
Agence de l'eau Adour-Garonne	« Etant donné la thématique retenue (focus sur l'eau), le projet Open DataLab Midi-Pyrénées que vous portez, nous intéresse particulièrement pour son approche d'animation par filière permettant : - de confronter de manière continue les producteurs et les utilisateurs, - de croiser et structurer les données de différents producteurs, - de capitaliser les expériences et faciliter la réutilisation. L'Agence de l'eau s'impliquera dans la mise en œuvre de ce projet, sur la base des données sur l'eau qu'elle produit et diffuse, et de son expérience en matière de gestion et catalogage des données. »	Martine Gaëckler Directrice de département Département Connaissance et Système d'Information
Bureau de recherches Géologiques et minières (BRGM)	« Dans le cadre de ce projet, la thématique de l'eau ayant été retenue comme thématique prototype, le BRGM Midi-Pyrénées peut apporter son appui dans le recensement et la structuration des données en tant que producteurs de données sur les eaux souterraines régionales. »	Philippe Roubichou Directeur Régional
DIRECCTE	« Le projet d'Open DataLab porté par la SGAR intéresse la DIRECCTE à double titre. D'une part, le service des études de la DIRECCTE est chargé de valoriser des données que produisent nos partenaires. Qu'un projet de grande envergure mette en avant ces données, parmi d'autres, et permette de les faire connaître au public, entre en parfaite cohérence avec cette mission de la DIRECCTE. D'autre part, le service développement industriel s'occupe de soutenir les entreprises de l'écosystème régional, en particulier en incitant celles-ci à développer leur innovation. Dans cette perspective, et en lien avec l'initiative French Tech, la DIRECCTE pourra apporter son soutien au projet en mobilisant	Sébastien Gueremy Chef du service Développement Industriel, Technologique et International (DITI) Pôle Entreprises, Emploi et Economie



	les startups et les codeurs, pour qu'ils s'approprient les données qui seront ouvertes par l'Administration, et fassent ainsi naître de nouvelles applications et de nouveaux usages innovants. »	
DREAL	<p>« La DREAL cherche désormais à développer cette valorisation en accédant à un public plus large. De nombreux gains sont attendus d'une telle démarche tant en termes de transparence que de simplification dans l'accès à nos données. En outre, le projet « OpenData Lab » prévoit une écoute fine des utilisateurs qui doit permettre de rendre nos données plus facilement valorisables.</p> <p>L'initiative « OpenDataLab » a donc retenu notre attention puisqu'elle propose une méthode de travail innovante, se construit autour d'une gouvernance solide et se fonde sur un vaste partenariat impliquant aussi bien les services de l'Etat que ceux des collectivités territoriales. Cette démarche devrait permettre :</p> <ul style="list-style-type: none"> <li>de dynamiser le tissu entrepreneurial et associatif local en donnant accès à des informations de qualité ;</li> <li>de créer des boucles vertueuses d'améliorations de nos données ;</li> <li>de disposer de ré-utilisations rendant nos données plus facilement assimilables par les usagers.</li> </ul> <p>Nous souhaitons donc nous associer à cette démarche. »</p>	Hubert Ferry-Wilczek Directeur régional
INSEE	<p>« L'Insee est intéressé à plusieurs titres par un tel projet : il est en effet à la fois producteur et utilisateur de données, et le site internet www.insee.fr met gratuitement à disposition du public une grande quantité de données.</p> <p>Vous m'avez proposé d'associer l'Insee à ce projet. Je vous confirme par la présente que la direction régionale Midi-Pyrénées de l'Insee participera à la réalisation de ce projet. »</p>	Le directeur régional
ARS	<p>« L'ARS et notamment le pôle eaux du département Santé Environnement a été partenaire depuis son origine du projet Mipy-géo. Cet outil est devenu maintenant incontournable pour l'ARS, tant au niveau de la diffusion de données aux différents partenaires, qu'au niveau de son utilisation lors des contraintes.</p> <p>Ce projet Open data, dans la continuité de Mipygéo, semble particulièrement intéressant pour la valorisation de nos données publiques.</p> <p>L'Ars et notamment le pôle eaux pourrait être partenaire de ce projet en tant que fournisseur et utilisateur des données mises à disposition. »</p>	Jean-Marc VACHER Responsable du pôle Eaux Département Santé-Environnement
Société Publique Locale ARPE / Observatoire Régional Eau et Territoires en Midi-Pyrénées	<p>« Nous partageons votre analyse sur la nécessaire levée de freins et sur le travail en mode "start-up d'Etat" qu'un tel projet implique. [...] Le choix de la thématique de l'eau pour cette première consolidation des données ouvertes nous paraît donc pertinent.</p> <p>Pour ces motifs, je vous confirme notre intérêt à participer à ce projet et à pouvoir intégrer, si vous le jugez pertinent, le comité de pilotage du laboratoire de la réutilisation des données publiques »</p>	Martine Gayraud-Mathorel Directrice Générale Adjointe
Pôle de Compétitivité de l'eau (Montpellier) et Laboratoire Ecolab	<p>« Ce projet nous intéresse à double titre, d'une part pour participer avec nos bases de données déjà existantes (sous réserve de confidentialité et de délai de valorisation par la recherche avant mise à disposition), en particulier sur nos Observatoires de Recherche en Environnement sur le long terme (SO, SOERE, OHM, ZA, LTER...) à la constitution de base de données sur l'Eau et l'Environnement en Midi-Pyrénées ouvertes à un public plus large, et d'autre part pour faire remonter les besoins des laboratoires et des programmes de recherche en terme de données publiques existantes (meilleures connaissances des données existants, disponibilité et mise à disposition).</p> <p>Nous sommes prêts à participer à ce projet pour voir ensemble comment construire cet Open Data et échanger sur les besoins de données publiques et sur les possibilités de mise à disposition de nos données. Il est clair que la France a pris du retard en matière de constitution de base données et de métadonnées et de mise à disposition de ces données par rapport à ce qui existe aujourd'hui dans d'autres pays d'Europe et aux Etats-Unis. »</p>	Jean-Luc PROBST Administrateur du Pôle mondial de l'eau Directeur d'Ecolab
Toulouse Métropole	<p>« Ce projet rencontre l'intérêt de Toulouse Métropole pour être partenaire. En effet, la Métropole œuvre pour l'ouverture des données de l'ensemble de ses communes y compris les plus petites. De plus, l'ouverture des données relatives à la filière de l'eau sont également dans le spectre de la collectivité vis-à-vis de ses délégataires. Et enfin, la collectivité est attentive à faire part de ses retours d'expériences, en particulier sur ce type de projet encore assez novateurs. »</p>	André Thomas Directeur Général Délégué



Open Data France	« Les éléments du projet suscitent l'intérêt d'Open Data France dont les missions sont notamment d'œuvrer à l'ouverture des données publiques et d'aider les collectivités territoriales entamant une démarche en ce sens. En effet, les données publiques relatives aux territoires ne seront complètes que lorsque l'ensemble des acteurs auront procédé à l'ouverture de leurs données, qu'il s'agisse des collectivités de toute taille mais aussi des services déconcentrés de l'Etat. Toute action en ce sens ne peut que rencontrer l'intérêt d'Open Data France. »	Bertrand Serp Président
France Nature Environnement Midi Pyrénées	« FNE Midi-Pyrénées pourra faire émerger les problématiques que nous rencontrons en tant qu'utilisateurs de données, nos attentes et nos retours d'expérience. Parallèlement, nous ne manquerons pas d'être à l'écoute des échanges avec les producteurs de données. »	Marie-Laure Cambus Directrice

D'autres acteurs ont été impliqués dans la construction du projet, notamment via des réunions qui se sont déroulées en novembre 2014 et en mars 2015 :

- La DRAAF, qui s'est montrée intéressée par le projet et la thématique de l'eau
- Le Conseil Régional, qui s'est montré très intéressé par le renforcement des capacités des collectivités et qui émet l'hypothèse d'articuler cet appui avec sa plateforme SIP (voir le point 3.2)
- Les Conseils Généraux de la région Midi-Pyrénées
- Les DSP de l'agglomération de Toulouse (Véolia et Suez), qui par le biais de la responsable Open Data de Toulouse Métropole, ont manifesté un intérêt pour le projet
- La Communauté d'Agglomération du Sicoval
- L'IGN

### 3.3.2 Cercle des producteurs et réutilisateurs sur la filière de l'eau

Pour traiter la filière de l'eau, le cercle est constitué par les organismes particulièrement concernés : DREAL, ARS, BRGM, ARPE, agences de l'eau, DSP, mais aussi la société civile (France Nature Environnement par exemple) et universitaire (Ecolab, Pôle mondial de l'eau).

## 3.4 Terrains d'expérimentation pressentis

Les filières qui pourront être traitées en priorité en mode projet sont : l'eau, le tourisme et l'accès aux services publics.

Le projet se concentre dans un premier temps sur la mise en place et l'animation du cercle des producteurs et réutilisateurs sur la filière de l'eau.



# 4 Actions prévues pour pérenniser la solution en cas de succès

## 4.1 Gouvernance du projet

Une des activités principale du projet sera de mener une réflexion sur la pérennisation du projet et sur la gouvernance post-projet. Ce point est spécifié à l'activité 4 du cadre logique.

L'un des éléments clefs utilisé pour la pérennisation de l'action est la capitalisation de toutes les connaissances acquises au niveau régional sur le projet. Cette capitalisation est également un facteur facilitant grandement la répliquabilité du projet sur d'autres territoires.

Un autre élément permettant l'appropriation post-projet des outils créés est le fait de les développer en open source et de les documenter.

Il est à noter que le Conseil Régional fournit un service d'API permettant d'afficher sur les sites internet des collectivités des informations pratiques (plateforme SIP - Service d'Information Publique). Ces informations peuvent provenir du site internet du Conseil Régional mais aussi d'autres sites comme service-public.fr. Le Conseil Régional a donc créé un réseau d'assistance au niveau départemental via les Centres de Gestion. Il a donc été envisagé avec la Région de réutiliser ce réseau départemental pour assurer la pérennité du suivi et du support des 2 API créées dans le cadre du projet proposé.

## 4.2 Acteurs institutionnels mobilisés pour pérenniser le projet en cas d'expérimentation réussie

Le projet étant porté par le SGAR, tous les services de l'État en région seront mobilisés pour assurer la pérennité du projet.

De plus, il est à souhaiter que la DREAL, déjà fortement impliquée dans les SIG pour les services de l'État, porte ensuite l'initiative de l'Open Data en région. Ainsi d'un pôle SIG pour les services de l'État (MiPyGéo), la DREAL pourrait acquérir une compétence plus large englobant les SIG (avec données ouvertes ou pas) et l'Open Data. Cette évolution paraît naturelle au vu de la montée en puissance de la thématique Open Data.

Le Conseil Régional et sa plateforme SIP pourrait également s'impliquer dans la pérennisation du projet en intégrant le soutien aux collectivités (API, bonnes pratiques) à son réseau SIP déjà existant dans les départements de la Région.

## 4.3 Documentation prévue du projet en cours de réalisation.

Le projet mettra particulièrement l'accent sur la documentation issue du projet. Ainsi l'activité 4 du cadre logique prévoit la capitalisation des connaissances acquises sur le projet. Il sera ainsi formalisé l'ensemble du savoir qui aura émergé du projet notamment à travers les retours d'expériences réalisées :

- Série de questions-réponses (FAQ)
- Étude de cas concrets
- Valorisation des succès obtenus (success stories)
- Témoignages
- Guides pratiques (dans la mesure où ils n'existent pas déjà)



# 5 Résultats attendus

Deux résultats principaux sont attendus sur le projet (voir cadre logique à la page 13) :

## Résultat 1 :

Structurer la production des données publiques en mettant en relation les producteurs et les utilisateurs dans une démarche d'amélioration continue

## Résultat 2 :

Renforcer les capacités des collectivités territoriales de plus de 3500 habitants en matière d'ouverture de données publiques

Des indicateurs et des sources de vérification sont associés à chacun de ces résultats. Le projet étant un projet d'amarce, il semble difficile de se fixer un objectif en termes d'indicateurs en démarrage du projet. Toutefois le fait de définir d'ores et déjà ces indicateurs permettra de suivre l'évolution du projet et à la fin de celui-ci, de quantifier les résultats par rapport à la situation existante.

## 5.1 Gains potentiels estimés en termes de qualité de service

L'organisation innovante du projet OpenDataLab et des outils proposés visent à rendre les données publiques produites par les partenaires du projet :

- plus visibles et en adéquation avec les usages
- harmonisées
- qualifiées

Tous ces éléments ont comme objectif commun un accroissement de la qualité de service une multiplication des réutilisations de données publiques.

## 5.2 Impact financier et économique potentiel

Même si les impacts financiers directs seront difficilement quantifiables, l'ensemble des actions proposées dans le cadre de ce projet visent à améliorer l'efficacité des modes de production et de mise à disposition des données publiques. Si les objectifs sont atteints, certaines tâches de traitements et d'extraction de données ne seront plus nécessaires. L'idée est également que certains services travaillent plus efficacement grâce à l'accès à des données produites par un d'autres partenaires.

La mise à disposition en Open Source des API javascript de catalogue et de cartographie permettront à certaines collectivités d'éviter de faire l'acquisition d'outils, mais de disposer d'un service équivalent gratuitement.

## 5.3 Autres résultats obtenus par le projet

Au-delà des retombées économiques directes et de l'amélioration de la qualité de service de l'administration, le projet permet de développer l'acculturation des citoyens à l'utilisation des données. Ce développement permettra un effet de levier important pour démultiplier les retombées en élargissant et intensifiant l'écosystème.

Ce résultat sera atteint notamment en encourageant les réutilisations visant à proposer des services d'information, de partage des connaissances, de vulgarisation des données, de visualisation, de datajournalisme...

