



HAL
open science

Développement d'une base de données et d'un outil de visualisation des actions de renaturation, dans le cadre de la préservation de la biodiversité urbaine

Morgane Spinner

► To cite this version:

Morgane Spinner. Développement d'une base de données et d'un outil de visualisation des actions de renaturation, dans le cadre de la préservation de la biodiversité urbaine. Informatique [cs]. 2024. dumas-04790702

HAL Id: dumas-04790702

<https://dumas.ccsd.cnrs.fr/dumas-04790702v1>

Submitted on 19 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MINISTERE DE L'AGRICULTURE

ECOLE NATIONALE SUPERIEURE des SCIENCES AGRONOMIQUES de BORDEAUX
AQUITAINE
1, cours du Général de Gaulle - CS 40201 – 33175 GRADIGNAN cedex



MEMOIRE de fin d'études

pour l'obtention du titre

d'Ingénieur de Bordeaux Sciences Agro

Développement d'une base de données et d'un outil de
visualisation des actions de renaturation, dans le cadre
de la préservation de la biodiversité urbaine

SPINNER, Morgane

Spécialisation : Agrotic – Technologies de l'Information et de la communication pour
l'Agriculture et l'Environnement

Étude réalisée à : Natural Solutions
68 Rue Sainte, 13001 Marseille



- 2024 -

THE FRENCH MINISTRY OF AGRICULTURE AND FOOD

NATIONAL SCHOOL of AGRICULTURAL SCIENCES and ENGINEERING,
BORDEAUX AQUITAINE
1, cours du Général de Gaulle - CS 40201 – 33175 GRADIGNAN cedex



MASTER THESIS

Submitted in fulfillment of the requirements for the degree of
Agricultural Engineer, Bordeaux Sciences Agro

Development of a database and a visualization tool for
renaturation actions within the framework of urban
biodiversity preservation.

SPINNER, Morgane

Specialisation: AgroTIC – Technologies de l'Information et de la communication pour
l'Agriculture et l'Environnement

Study completed at: Natural Solutions
68 Rue Sainte, 13001 Marseille



- 2024 -

Résumé :

Dans un contexte d'érosion massive de la biodiversité et d'urgence de protection et de restauration des écosystèmes, les actions de renaturation émergent comme une réponse prometteuse. Natural Solutions, une entreprise spécialisée dans la technologie au service de la biodiversité, a initié un projet visant à recenser l'ensemble des actions de renaturation pour créer un outil d'aide à la décision. Ce projet s'inscrit dans le cadre de la loi Climat et Résilience de 2021, qui introduit le principe du zéro artificialisation nette (ZAN) et formalise la renaturation en France.

Le document présente le contexte du projet, expliquant le concept de renaturation et identifiant les principaux acteurs impliqués. Il détaille ensuite la méthodologie employée pour l'identification des actions de renaturation et la construction d'une base de données. L'implémentation de cette base de données est décrite, incluant le choix des outils et technologies utilisés. Enfin, le document aborde la visualisation des données, présentant l'architecture du frontend et le résultat final de l'outil développé.

Ce projet vise à fournir une ressource centralisée sur les actions de renaturation, contribuant ainsi aux efforts de préservation de la biodiversité et de restauration des écosystèmes. L'outil développé permettra une meilleure compréhension et planification des initiatives de renaturation, soutenant ainsi les décideurs et les acteurs impliqués dans ces projets cruciaux pour l'environnement.

Mots-clés :

Renaturation ; Biodiversité ; Zéro Artificialisation Nette ; Écosystèmes ; Base de données ; Visualisation ; Aide à la décision ; Restauration écologique ; Loi Climat et Résilience ; Désartificialisation

Abstract:

In the context of massive biodiversity loss and the urgent need to protect and restore ecosystems, renaturation efforts are emerging as a promising response. Natural Solutions, a company specializing in technology for biodiversity, has initiated a project aimed at cataloging all renaturation actions to create a decision-support tool. This project is part of the 2021 Climate and Resilience Law, which introduces the principle of zero net land take and formalizes renaturation in France.

The document outlines the project's background, explaining the concept of renaturation and identifying the key stakeholders involved. It then details the methodology used for identifying renaturation actions and building a database. The implementation of this database is described, including the choice of tools and technologies employed. Finally, the document discusses data visualization, presenting the frontend architecture and the final outcome of the developed tool.

This project aims to provide a centralized resource on renaturation actions, thus contributing to efforts to preserve biodiversity and restore ecosystems. The developed tool will enable a better understanding and planning of renaturation initiatives, thereby supporting decision-makers and stakeholders involved in these crucial environmental projects.

Keywords:

Renaturation; Biodiversity; Zero Net Land Take; Ecosystems; Database; Visualization; Decision Support; Ecological Restoration; Climate and Resilience Law; Land De-sealing.

Remerciements

Je tiens tout d'abord à exprimer ma profonde gratitude à Gaëtan DUHAMEL, mon maître de stage, pour m'avoir offert l'opportunité de réaliser ce stage. Je lui suis reconnaissante non seulement pour les missions et les responsabilités qu'il m'a confiées, mais aussi pour la grande autonomie qu'il m'a accordée.

Je remercie également Nathalie TOULON, ma tutrice école, pour son accompagnement précieux tout au long de la rédaction de ce mémoire. Sa disponibilité, ses conseils avisés et son engagement constant ont été des soutiens inestimables.

Je souhaite aussi exprimer ma reconnaissance à Naomi FISCHER, dont l'expertise et les échanges enrichissants ont grandement facilité la phase de design. Ses réponses à mes questions et nos discussions approfondies ont été déterminantes dans la résolution des défis techniques rencontrés.

Un grand merci à l'équipe Renaturation pour l'ambiance motivante et le soutien constant. Je tiens à remercier tout particulièrement Mickael BELLUAU pour les moments instructifs que nous avons partagés.

Enfin, je remercie chaleureusement ma famille pour leur soutien indéfectible tout au long de ce stage. Leur encouragement a été un moteur essentiel dans cette étape de ma vie professionnelle et leur aide précieuse dans la rédaction de ce mémoire a joué un rôle capital.

Sommaire

Introduction.....	1
1. Contexte du projet.....	2
1.1. Les actions de renaturation.....	2
1.1.1. Concept de renaturation.....	2
1.1.2. Acteurs de la renaturation.....	3
1.1.2.1. Acteurs mettant en place des actions de renaturation.....	3
1.1.2.2. Acteurs impliqués dans les projets de renaturation.....	3
1.1.3. État des lieux et défis des actions de renaturation.....	4
1.2. Natural Solutions, une agence au service de la biodiversité.....	5
1.2.1. Présentation de l'entreprise.....	5
1.2.2. Organisation de l'entreprise.....	6
1.2.3. Cartographier et préserver la biodiversité urbaine avec <i>Ecoteka</i>	7
1.2.3.1. <i>Ecoteka</i> , une application pour la renaturation urbaine.....	7
1.2.3.2. Modèle économique d' <i>Ecoteka</i>	9
1.3. Missions et déroulement du stage.....	9
1.3.1. Objectif et problématique.....	9
1.3.2. Organisation du stage.....	10
2. Identification des actions de renaturation et construction de la base de données.....	11
2.1. Méthodologie.....	12
2.2. Identification des sources d'informations.....	14
2.2.1. Choix des sources d'information.....	14
2.2.2. Listing des actions.....	16
2.3. Modélisation de la base de données.....	18
3. Implémentation de la base de données.....	20
3.1. Méthodologie.....	20
3.2. Architecture et solution de l'environnement du code.....	24
3.2.1. Système de gestion de base de données.....	24
3.2.1.1. Outils existants.....	24
3.2.1.2. Comparaison des outils existants.....	24
3.2.2. Outil d'extraction.....	25
3.2.2.1. Page web.....	25
3.2.2.1.1. Outils existants.....	25
3.2.2.1.2. Comparaison des outils existants.....	26
3.2.2.2. PDF.....	27
3.2.2.2.1. Outils existants.....	27
3.2.2.2.2. Comparaison des outils existants.....	27
3.2.3. Framework Web Python.....	28
3.2.3.1. Outils existants.....	28
3.2.3.2. Comparaison des outils existants.....	28
3.2.4. Système d'orchestration de conteneurs.....	29
3.2.4.1. Outils existants.....	29
3.2.4.2. Comparaison des outils existants.....	30
3.2.5. Synthèse de l'architecture.....	31
3.3. Remplissage de la base.....	32

3.3.1. Table dont le contenu est statique	32
3.3.2. Couche géographique.....	32
3.3.3. Site web.....	33
4. Visualisation	36
4.1. Méthodologie	36
4.2. Architecture et solution du frontend	36
4.2.1. Les frameworks JavaScript	36
4.2.1.1. Outils existants	36
4.2.1.2. Comparaison des outils existants	37
4.2.2. Gestionnaire de carte web	38
4.2.2.1. Outil existant	38
4.2.2.2. Comparaison des outils existants	38
4.2.3. Architecture.....	39
4.3. Le résultat final	41
4.3.1. Page cartographique.....	42
4.3.2. Liste d'actions	43
4.3.3. Filtre.....	44
4.3.4. Fiche action.....	46
4.4. Bilan et perspectives	48
4.4.1. Bilan du travail réalisé	48
4.4.2. Perspectives du projet	48
Conclusion	51
Bibliographie	53

Liste des figures

Figure 1 : Organisation de Natural Solutions	6
Figure 2 : Les fonctionnalités d'Ecoteka	8
Figure 3 : Diagramme de Gantt du projet	11
Figure 4 : Exemple de fiche de lecture	17
Figure 5 : Schéma simplifié du modèle de données.....	18
Figure 6 : Architecture de l'environnement de code avec SGBD (cette figure se complétera progressivement au fur et à mesure de l'avancement de cette section)	21
Figure 7 : Architecture de l'environnement du code avec outils d'extraction	22
Figure 8 : Architecture de l'environnement du code avec framework web Python.....	23
Figure 9 : Architecture de l'environnement du code avec plateformes de conteneurs	24
Figure 10 : Architecture détaillé de l'environnement de code du backend.....	31
Figure 11 : Schéma de répartition des informations issue des fichiers Shapefile dans les différentes tables	33
Figure 12 : Exemple d'extraction de date	34
Figure 13 : Exemple d'extraction de partenaires financiers	35
Figure 14 : Exemple d'extraction de budget.....	35
Figure 15 : Schéma des informations présentes dans le CSV et leurs répartitions dans la base de données	35
Figure 16 : Architecture de l'environnement du frontend	40
Figure 17 : Architecture détaillé de l'environnement docker	40
Figure 18 : Architecture du site web.....	41
Figure 19 : Header permettant la navigation sur les différentes pages (activé sur la page accueil).....	41
Figure 20 : Page cartographique	42
Figure 21 : Page "Liste d'actions"	43
Figure 22 : Filtre	45
Figure 23 : Fiche action	47

Liste des tableaux

Tableau 1 : Recensement de l'ensemble des sources identifiées (Légende : Les cases vertes situées dans la quatrième colonne du tableau indiquent que l'information est présente dans le document.)....	15
Tableau 2 : Comparaison des principaux SGBD à l'aide des principaux critères d'évaluation	25
Tableau 3 : Comparaison des principales bibliothèques Python pour le scraping de page web	26
Tableau 4 : Comparaison des principales bibliothèques Python pour le scraping de PDF.....	28
Tableau 5 : Comparaison des principales bibliothèques Python pour les frameworks web Python	29
Tableau 6 : Comparaison des principales bibliothèques Python pour les systèmes d'orchestration de conteneurs	31
Tableau 7 : Comparaison des principaux frameworks JavaScript.....	38
Tableau 8 : Comparaison des principaux gestionnaires de carte web	39

Liste des annexes

Annexe i : Modèle de données	60
Annexe ii : Dictionnaire de données	61

Liste des sigles et abréviations

AFNAV	Association Française de Renaturation et Aménagement des Vallées
API	Interface de Programmation d'Applications
CDC	Caisse des Dépôts et Consignations
CEO	Chief Executive Officer
CIO	Chief Information Officer
COO	Chief Operating Officer
CPO	Chief Product Officer
CSV	Comma-Separated Values
ETL	Extract, Transform, Load
GPS	Système de Positionnement Global
HTML	HyperText Markup Language
HTTPS	HyperText Transfer Protocol Secure
LiDAR	Light Detection and Ranging
NS	Natural Solutions
OFB	Office Français de la Biodiversité
ONG	Organisation Non Gouvernementale
PDF	Portable Document Format
PLU	Plans Locaux d'Urbanisme
POO	Programmation Orientée Objets
RAG	Retrieval Augmented Generation
RGPD	Règlement Général sur la Protection des Données
SCoT	Schémas de Cohérence Territoriale
SGBD	Systèmes de Gestion de Bases de Données
SIG	Système d'Information Géoréférencé
SQL	Structured Query Language
SRADDET	Schémas Régionaux d'Aménagement, de Développement Durable et d'Égalité des Territoires
SRCE	Schémas Régionaux de Cohérence Écologique
UI	User Interface
URL	Uniform Resource Locator
UX	User Experience
WWF	World Wildlife Fund
ZAN	Zéro Artificialisation Nette

Glossaire

Anthropisation : Processus par lequel un environnement naturel est transformé ou altéré par l'action humaine, notamment par l'urbanisation, l'agriculture ou l'industrie.

Back-end : Partie d'une application ou d'un site web qui gère la logique, la base de données, et la communication avec les serveurs. Le back-end est responsable du traitement des données, de la gestion des utilisateurs, de la sécurité, et de l'interaction avec les bases de données. Il fonctionne en arrière-plan et n'est pas visible par l'utilisateur final, contrairement au front-end.

Biodiversité : Variété des formes de vie sur Terre, y compris la diversité des espèces, des écosystèmes et des gènes, qui est essentielle au maintien des équilibres écologiques.

Désimperméabilisation : Processus visant à rendre perméable un sol précédemment imperméabilisé (par du béton, de l'asphalte, etc.), afin de restaurer ses capacités naturelles d'absorption des eaux de pluie et de favoriser la biodiversité.

Écosystème : Système formé par une communauté d'êtres vivants (plantes, animaux, micro-organismes) en interaction avec leur environnement physique (eau, sol, air).

Front-end : Partie visible d'une application ou d'un site web avec laquelle les utilisateurs interagissent directement. Le front-end concerne tout ce qui est lié à l'interface utilisateur (UI), incluant les éléments graphiques, les mises en page, les animations et la navigation. Il est développé en utilisant des technologies comme HTML, CSS, et JavaScript pour assurer une expérience utilisateur intuitive et responsive.

Métadonnées : Données fournissant des informations sur d'autres données, comme la date de création, l'auteur, ou la localisation géographique des données, utilisées pour faciliter leur gestion et leur recherche.

Open-source : Désigne un type de logiciel dont le code source est librement accessible, modifiable et distribuable par quiconque. Les logiciels open-source permettent aux utilisateurs de collaborer, d'améliorer le programme, et de le redistribuer, souvent dans le cadre de licences spécifiques qui garantissent la pérennité de cette ouverture.

Scraping : Technique utilisée pour extraire automatiquement des données à partir de sites web. Le scraping, fréquemment réalisé par des scripts ou des logiciels spécialisés, permet de récupérer des informations contenues dans des pages web en simulant la navigation d'un utilisateur, puis en analysant le contenu pour en extraire les données pertinentes.

Systèmes d'information géographique (SIG) : Logiciels et outils permettant la visualisation, l'analyse et l'interprétation de données spatiales pour comprendre des relations, des modèles et des tendances géographiques.

Urbanisation : Processus par lequel une zone rurale se transforme en zone urbaine, avec l'augmentation de la densité de population, la construction d'infrastructures, et la réduction des espaces naturels.

ChatGPT a été utilisé pour corriger la syntaxe et l'orthographe, traduire et reformuler des textes.

Introduction

La biodiversité, résultat de millions d'années d'évolution, connaît aujourd'hui un déclin alarmant à l'échelle mondiale. Selon un communiqué de presse de la World Wildlife Fund (WWF) (2020), les populations de vertébrés (mammifères, oiseaux, poissons, reptiles et amphibiens) ont chuté de 68 % en moyenne depuis 1970, ce que les scientifiques qualifient de 6^e extinction de masse, la première causée entièrement par les activités humaines. Certaines régions sont particulièrement touchées, comme l'Amérique latine et les Caraïbes qui ont perdu 94 % de leurs populations animales.

En France métropolitaine, le constat est tout aussi préoccupant. D'après l'évaluation de l'Office Français de la Biodiversité (OFB) en 2023, seuls 20 % des habitats naturels d'intérêt communautaire sont jugés en bon état de conservation. Les milieux humides et côtiers sont les plus dégradés, avec à peine 6 % dans un état favorable. Sur le plan des espèces, l'OFB (2020) souligne que 24 % des oiseaux communs spécialistes ont disparu depuis 1989, et que 19 % des espèces sont aujourd'hui éteintes ou menacées d'extinction, dont 16 % des plantes à fleurs.

Les principales pressions pesant sur la biodiversité sont multiples : les changements d'usage des terres, l'exploitation directe des ressources naturelles, les pollutions, le changement climatique et la prolifération d'espèces exotiques envahissantes (OFB, 2023). Face à cette érosion massive de la biodiversité, des efforts urgents de protection et de restauration des écosystèmes sont indispensables pour préserver les services écosystémiques qui contribuent au bien-être humain.

Bien que la situation de la biodiversité soit préoccupante, des initiatives prometteuses de renaturation sont mises en œuvre pour tenter d'enrayer son déclin.

La renaturation des milieux constitue une réponse à la crise environnementale mondiale en restaurant les écosystèmes dégradés. Face au changement climatique et à la perte de biodiversité, cette pratique englobe une variété d'actions d'aménagement visant à réduire l'impact humain sur un espace donné. Cela peut inclure la revégétalisation de zones déforestées, la réhabilitation de terres agricoles abandonnées, la désimperméabilisation des surfaces, ou encore la restauration de cours d'eau endommagés. Cette approche contribue à la capture du carbone (Ariluoma *et al.*, 2021), à la régulation des températures locales (Fu *et al.*, 2022), et à la préservation de la biodiversité (Latawiec *et al.*, 2016), offrant ainsi une lueur d'espoir pour l'avenir de notre planète.

C'est dans ce contexte que Natural Solutions, une entreprise spécialisée dans la technologie au service de la biodiversité, souhaite recenser l'ensemble des actions de renaturation. Cela permettrait de pouvoir les regrouper dans une base de données utilisée pour la création d'un outil d'aide à la décision.

Ce document a pour objet de présenter le projet réalisé ainsi que l'outil développé à la suite de celui-ci. Il débute par l'exposition du contexte dans lequel s'inscrit le projet, avec une brève explication de l'existant, de l'entreprise d'accueil, du déroulement du stage et de la méthodologie employée. Ensuite, il est structuré en plusieurs parties, chacune correspondant à une grande étape clé du projet. Chaque partie comprend une section détaillant la méthodologie mise en œuvre pour réaliser l'étape et une section mettant en lumière les résultats concrets obtenus, valorisant ainsi le travail accompli.

1. Contexte du projet

1.1. Les actions de renaturation

1.1.1. Concept de renaturation

Le concept de renaturation, également appelé renaturalisation, est associé à la restauration écologique et à la réparation des écosystèmes dégradés par les activités humaines. Bien que le terme soit relativement récent, il s'inspire de pratiques et de concepts développés au fil des décennies par des écologues et des spécialistes de l'environnement.

Le terme "renaturation" a été popularisé en France par l'Association Française de Renaturation et Aménagement des Vallées (AFNAV) dans les années 2000. La renaturation désigne les processus par lesquels les espèces vivantes recolonisent spontanément un milieu ayant subi des perturbations écologiques, ainsi que les opérations d'aménagement et de gestion visant à restaurer l'état écologique et paysager de sites dégradés (Castaing, Monod and Noreve, 2022).

En France, la renaturation a été formalisée dans le cadre de la loi Climat et Résilience de 2021, qui introduit le principe du zéro artificialisation nette (ZAN). Cette loi définit la renaturation comme l'action de restauration ou d'amélioration de la fonctionnalité d'un sol, transformant un sol artificialisé en un sol non artificialisé (Joinet, Laruelle and Delaville, 2023). Ce cadre législatif a renforcé l'importance de la renaturation dans les politiques publiques, notamment en milieu urbain, où elle est devenue un outil clé pour lutter contre les effets du changement climatique et améliorer la biodiversité (L'Institut Région Paris, 2023 ; Joinet, Laruelle and Delaville, 2023).

Une action de renaturation désigne un ensemble d'interventions visant à réduire le degré d'anthropisation d'un espace et à restaurer son état naturel. Ces actions peuvent inclure la revégétalisation, la remise en eau, la désimperméabilisation des sols et la reconstitution d'habitats naturels (Bouron, 2023). L'objectif est de transformer un sol artificialisé en un sol non artificialisé, favorisant ainsi la biodiversité et la résilience écologique.

La renaturation peut prendre diverses formes, allant de la plantation de haies bocagères à la création de jardins partagés et de parcs urbains. Elle peut également impliquer des travaux plus lourds comme la dépoldérisation d'un littoral ou la destruction de barrages pour restaurer des cours d'eau (Bouron, 2023 ; Caillavet and Cabrejas, 2020). Ces actions visent à recréer des écosystèmes fonctionnels et à améliorer la qualité des sols et des eaux.

Après avoir exploré le concept de renaturation et ses implications pour l'environnement et la biodiversité, nous allons examiner les forces vives qui animent ces projets. La renaturation n'est pas simplement une notion abstraite, mais une démarche concrète nécessitant l'intervention et la collaboration de multiples acteurs. Ces intervenants, issus de divers horizons, jouent chacun un rôle déterminant dans la conception, la réalisation et le suivi des initiatives de restauration écologique. Il convient donc d'analyser les principaux acteurs impliqués dans le processus de restauration écologique et leurs contributions respectives. Ce processus implique plusieurs étapes clés : l'évaluation initiale du site et de ses dommages, la planification détaillée des interventions, la suppression des facteurs de dégradation, la réintroduction d'espèces natives, la reconstruction des habitats et le suivi à long terme.

1.1.2. Acteurs de la renaturation

1.1.2.1. Acteurs mettant en place des actions de renaturation

Les acteurs mettant en place des actions de renaturation sont variés et complémentaires, chacun jouant un rôle spécifique dans ce processus complexe. Les collectivités territoriales sont en première ligne, particulièrement dans le cadre de l'objectif de zéro artificialisation nette (ZAN). Elles intègrent les continuités écologiques dans leurs documents d'urbanisme, mènent des projets de restauration écologique sur les friches et les cours d'eau, et valorisent la nature en ville à travers la gestion différenciée des espaces verts (L'Institut Region Paris, 2023). Les communes et leurs groupements, ainsi que les syndicats mixtes et établissements publics, peuvent bénéficier de subventions départementales pour des projets de désartificialisation et de renaturation des sols imperméabilisés (Jeanneau, 2024).

Les acteurs privés, tels que les entreprises et les propriétaires fonciers, participent également. Ils gèrent leurs sites pour favoriser la biodiversité, adoptent une gestion différenciée de leurs espaces verts, et végétalisent leurs espaces avec des espèces locales. Les organisations environnementales, comme le Conservatoire du Littoral, jouent un rôle en menant des actions concrètes telles que la démolition de bâtiments, le décompactage des sols et le rétablissement du fonctionnement hydraulique des zones humides remblayées (L'Institut Region Paris, 2023).

L'État et ses établissements publics, comme l'Office français de la biodiversité, les Parcs nationaux et l'Office national des forêts, contribuent également. Ils gèrent écologiquement leurs propriétés, supervisent certains espaces protégés, et accompagnent d'autres acteurs dans la mise en place de la Trame Verte et Bleue. À l'échelle régionale, des outils tels que les Schémas Régionaux de Cohérence Écologique (SRCE), désormais intégrés dans les Schémas Régionaux d'Aménagement, de Développement Durable et d'Égalité des Territoires (SRADDET), sont mobilisés. Les documents d'urbanisme, comme les Schémas de Cohérence Territoriale (SCoT) et les Plans Locaux d'Urbanisme (PLU), permettent d'identifier des zones préférentielles de renaturation, assurant ainsi une planification plus cohérente et efficace des opérations de renaturation (Marchangelo, 2023 ; France nation verte, 2023).

Après avoir exploré les différents acteurs qui initient et mettent en œuvre des actions concrètes de renaturation, il est important de reconnaître que ces projets ne se réalisent pas en vase clos. La réussite de ces initiatives repose souvent sur un réseau plus large de parties prenantes, chacune apportant son expertise, ses ressources et son soutien. Cette collaboration étendue permet de garantir le succès et la pérennité des efforts de renaturation. Il convient maintenant d'examiner plus en détail les divers intervenants qui jouent un rôle dans la planification, l'exécution et le suivi de ces projets ambitieux.

1.1.2.2. Acteurs impliqués dans les projets de renaturation

Les projets de renaturation impliquent une variété d'acteurs, chacun apportant des compétences et des perspectives uniques pour assurer la réussite des initiatives. Les élus locaux et les collectivités territoriales jouent un rôle central en intégrant les continuités écologiques dans leurs documents de planification et d'urbanisme, et en menant des actions de sensibilisation et de valorisation de la nature en ville (Site de la trame verte et bleue, 2023). Les services publics, tels que l'Office français de la biodiversité, les Parcs nationaux, ainsi que les Parcs Naturels Régionaux, qui couvrent une plus grande surface à l'échelle nationale, apportent leur expertise en gestion écologique et en supervision des espaces

protégés. Ils accompagnent également les autres acteurs dans la mise en place de la Trame Verte et Bleue (Site de la trame verte et bleue, 2023).

Les opérateurs privés, y compris les entreprises et les gestionnaires d'infrastructures, contribuent en adoptant des pratiques favorisant la biodiversité, comme la gestion différenciée des espaces verts et la végétalisation des sites avec des espèces locales (Site de la trame verte et bleue, 2023). Les experts en écologie, quant à eux, jouent un rôle en évaluant les potentialités écologiques des sites et en proposant des solutions adaptées aux écosystèmes locaux (Barra and Grandin, 2023).

La participation des citoyens est aussi importante, car les espaces renaturés sont destinés à être utilisés par les habitants pour des activités récréatives, sociales et éducatives. Leur implication dès les phases de conception permet de s'assurer que les aménagements répondent aux besoins et aux aspirations de la communauté, privilégiant ainsi l'appropriation des espaces par les usagers (Cerema, 2021). Des initiatives comme celles de Natura City montrent l'importance de la collaboration entre les différentes parties prenantes pour diagnostiquer les terrains, planifier les projets et donner une identité et une fonction aux espaces renaturés (Nexity, 2021).

Cette approche collaborative permet de transformer des espaces dégradés ou sous-utilisés en véritables atouts pour la biodiversité urbaine et le bien-être des habitants. La renaturation devient ainsi un processus de co-création où chaque acteur, qu'il soit public, privé ou citoyen, apporte sa pierre à l'édifice d'un environnement urbain plus vert et plus vivant. Les projets de renaturation, soutenus par des subventions et des initiatives gouvernementales, contribuent à la création de villes plus résilientes et durables, capables de faire face aux défis environnementaux actuels et futurs (L'Institut Region Paris, 2023 ; France nation verte, 2023).

L'exploration des différents acteurs impliqués dans les efforts de renaturation met en lumière l'importance de comprendre le contexte plus large dans lequel ces initiatives s'inscrivent. Les défis rencontrés par ces acteurs et les progrès accomplis jusqu'à présent définissent l'état actuel de la renaturation. Un examen plus approfondi de la situation sur le terrain et des obstacles à surmonter permet de mieux apprécier l'ampleur de la tâche et l'importance des collaborations entre ces acteurs clés pour atteindre les objectifs de restauration écologique. Cette compréhension globale contribue à évaluer l'efficacité des stratégies actuelles et identifier les domaines nécessitant une attention particulière dans le processus de renaturation.

1.1.3. État des lieux et défis des actions de renaturation

Les bénéfices de la renaturation sont multiples. Elle permet de lutter contre les îlots de chaleur urbains, d'améliorer la qualité de l'air et de l'eau, et de privilégier la recharge des nappes phréatiques. De plus, elle contribue à la création d'habitats diversifiés pour la faune et la flore, et à la protection des berges contre l'érosion (Caillavet and Cabrejas, 2020 ; Savi, 2020). En réintroduisant la nature en ville, la renaturation améliore aussi le bien-être des habitants et réduit les dépenses de santé publique liées au stress et à l'anxiété.

Les actions de renaturation jouent un rôle dans la lutte contre l'érosion de la biodiversité. En restaurant les écosystèmes dégradés et en recréant des habitats naturels, la renaturation permet de reconstituer des environnements favorables au développement de la faune et de la flore (Debœuf De Los Rios Serrano, Barra and Grandin, 2022 ; Rouen.fr, 2023).

De plus, la loi Climat et Résilience de 2021 a renforcé le cadre réglementaire autour de la renaturation, en l'inscrivant dans l'objectif de zéro artificialisation nette d'ici à 2050 (Légifrance, 2024 ; CDC Biodiversité, 2023). Cette dynamique législative vise à encourager les opérations de renaturation à grande échelle, pour enrayer le déclin de la biodiversité et préserver les services écosystémiques indispensables au bien-être humain (Légifrance, 2024 ; Légifrance, 2023).

Cependant, la mise en œuvre de projets de renaturation peut rencontrer des obstacles, notamment financiers et réglementaires. Les coûts de déconstruction, de désimperméabilisation et de dépollution peuvent être élevés, et les délais pour retrouver une bonne fertilité du sol peuvent être longs (Caillavet and Cabrejas, 2020). De plus, les plans locaux d'urbanisme ne sont pas toujours conçus avec des objectifs bioclimatiques, ce qui peut freiner l'intégration de la renaturation dans les projets urbains.

Actuellement, il n'existe pas de base de données centralisée regroupant toutes les actions de renaturation. Ces actions sont morcelées et dispersées à travers diverses sources de données, se présentant sous différentes formes. Certaines sont disponibles sous forme de couches géographiques ou de cartographies web permettant de géolocaliser les actions. D'autres sont compilées dans des recueils hétérogènes, tant sur la forme que sur le contenu, pouvant contenir de 5 à 100 actions. Les informations fournies varient considérablement, allant du simple nom de l'action à des détails plus complets tels que les parties prenantes impliquées et le budget alloué à la réalisation de l'action.

Après avoir exploré les diverses actions de renaturation et leurs bénéfices pour l'adaptation au changement climatique et l'amélioration du cadre de vie urbain, il est pertinent d'examiner comment ces initiatives peuvent être mises en œuvre concrètement. C'est là qu'interviennent des acteurs spécialisés, capables d'accompagner les collectivités et les entreprises dans leurs démarches en faveur de la biodiversité et de la renaturation des espaces. L'entreprise Natural Solutions est l'un d'eux.

1.2. Natural Solutions, une agence au service de la biodiversité

1.2.1. Présentation de l'entreprise

Le cœur de métier de Natural Solutions réside dans son engagement en faveur de la biodiversité à travers une palette de services spécialisés. L'édition de logiciels sur mesure constitue le pivot central de son activité, offrant des solutions adaptées aux besoins spécifiques de ses clients. En parallèle, la conception graphique et le développement de sites web et d'applications mobiles représentent des outils pour véhiculer leur message et sensibiliser le public à l'importance de la préservation de la nature. Le conseil ainsi que la recherche et développement complètent cet ensemble, permettant à Natural Solutions d'innover constamment pour mieux servir la cause de la biodiversité. (Natural Solutions, 2023)

Natural Solutions utilise des outils mobiles tels que les smartphones et les GPS pour collecter et gérer efficacement des données environnementales sur le terrain, notamment celles concernant la faune, la flore et la géographie. L'entreprise se spécialise également dans la visualisation et l'édition de ces données à travers des interfaces graphiques et des outils SIG. De plus, elle développe des supports numériques tels que des sites web et des applications mobiles afin de promouvoir les sciences participatives et la préservation de la biodiversité. (Natural Solutions, 2023)

1.2.2. Organisation de l'entreprise

Natural Solutions est dirigée par un CEO (*Chief Executive Officer*) responsable de la stratégie, de la vision d'entreprise et des partenariats. L'entreprise est structurée en quatre unités principales : deux unités de production, l'unité "Connaissance des individus et des écosystèmes" et l'unité "Santé et agronomie", et deux unités de support, l'unité "Platform" et l'unité "RH et comptabilité" (**Figure 1**).

Chaque unité de production est gérée par un CPO (*Chief Product Officer*) chargé d'assurer l'uniformisation des pratiques et des technologies au sein de son unité. Le CPO communique également la vision et les priorités à l'ensemble des parties prenantes impliquées dans les projets de son unité.

En ce qui concerne les unités de support, la première est dirigée par un CIO (*Chief Information Officer*). Son rôle consiste à analyser les besoins technologiques de l'entreprise, à mettre en place l'infrastructure IT nécessaire et à en assurer le suivi. La seconde unité de support est quant à elle gérée par un COO (*Chief Operating Officer*) responsable de la stratégie d'entreprise, des finances, de la comptabilité et du recrutement.

Chacune des quatre unités comprend un designer UX/UI, un chargé de développement et un portefeuille de produits et projets. Au sein même de ces unités, on retrouve des équipes composées d'un *Product Owner*, d'un *Lead Developer* et de développeurs dédiés aux différents projets en cours.

"Données supprimées : non conformes RGPD"

Figure 1 : Organisation de Natural Solutions

Natural Solutions opère selon une méthodologie agile, mettant en avant la collaboration, la souplesse et une distribution efficace des produits et services grâce à des itérations rapides et une communication transparente. Cette méthode se concentre sur la livraison rapide de parties fonctionnelles du produit, avec une adaptation continue aux changements de priorités et aux retours des clients. Les équipes agiles

travaillent en cycles courts appelés "sprints" et impliquent étroitement les parties prenantes tout au long du processus.

Faisant partie de l'unité "Connaissance des individus et des écosystèmes", nous nous concentrerons plus particulièrement sur les produits de cette unité :

- **GeoNature** est un logiciel *open-source* développé par l'équipe Observation, offrant une solution complète pour la gestion des données naturalistes en faune et flore. Ce système d'information dédié à la biodiversité comprend une suite d'applications web et mobiles disponibles sous licence libre, permettant aux acteurs de la biodiversité de saisir, gérer, consulter, analyser, exporter et diffuser efficacement leurs données. Les composants principaux de GeoNature incluent l'application web principale pour la saisie, la consultation et l'export des données, ainsi que OccTax mobile pour la collecte sur le terrain. GeoNature comprend également GeoNature Citizen, un outil de sciences participatives permettant aux citoyens de collecter des données naturalistes via des programmes dédiés, et GeoNature-atlas pour la diffusion publique des informations. Le système offre des fonctionnalités avancées telles que la gestion des référentiels taxonomiques et utilisateurs, la saisie de données selon divers protocoles et la gestion des métadonnées. Il permet également l'intégration de données partenaires et l'exportation des données dans différents formats adaptés aux besoins des partenaires.
- L'équipe Océan a développé **Wings Map**, un produit phare conçu pour l'ONG française *Wings of the Ocean*. Cette organisation, qui lutte contre la pollution marine, en particulier les déchets plastiques, organise des opérations de nettoyage sur les plages françaises. *Wings Map* répond à un double besoin : gérer les aspects opérationnels sur le terrain et traiter les données pour produire des rapports destinés au grand public et aux actions de *lobbying*. L'application comprend une page d'accueil avec une carte interactive montrant les actions de nettoyage, ainsi qu'un *dashboard* présentant diverses métriques. Les utilisateurs peuvent saisir des informations via des formulaires personnalisables par l'administrateur. Ce projet marque la première utilisation complète de la méthode agile, avec des révisions régulières du cahier des charges toutes les unes à deux semaines. Cette approche a permis d'optimiser le temps de développement en se concentrant sur les fonctionnalités essentielles.
- **Ecoteka**, développée par l'équipe Renaturation, est une application web et mobile *open-source* qui révolutionne la gestion du patrimoine arboré urbain. Conçue pour optimiser les processus de gestion et soutenir activement les efforts de renaturation des villes, cette solution innovante joue un rôle central dans l'amélioration de l'environnement urbain. Étant donné que la solution développée durant le stage doit s'intégrer à *Ecoteka*, nous examinerons plus en détail les caractéristiques et les fonctionnalités de ce produit dans la section suivante.

1.2.3. Cartographier et préserver la biodiversité urbaine avec *Ecoteka*

1.2.3.1. *Ecoteka*, une application pour la renaturation urbaine

Ecoteka est une application web et mobile *open-source* conçue pour une planification intelligente de la renaturation urbaine.

Ecoteka offre une cartographie arborée tout-en-un, incluant inventaire, diagnostics, suivi budgétaire et exports SIG. Face aux enjeux du changement climatique, l'application propose également des modules de calculs automatiques de services écosystémiques tels que le stockage de CO₂, les îlots de fraîcheur, la préservation de la biodiversité, les corridors écologiques et un inventaire automatisé.

L'approche d'*Ecoteka* repose sur trois piliers principaux :

- 1) La méthode UX qui consiste en une démarche de conception de services prenant en compte les besoins et les caractéristiques propres à chaque utilisateur.
- 2) L'*open-source* qui permet de co-construire le projet et de garantir la pérennité des différentes réalisations.
- 3) La méthode agile par incrémentation permettant de réadapter le produit à chaque étape de développement pour correspondre au mieux aux attentes du client.

Les outils utilisés par *Ecoteka* associent les technologies web, mobile, le développement logiciel et les outils cartographiques. L'utilisation de données ouvertes permet l'analyse et la visualisation sous forme de *Dashboard*. L'innovation réside dans la co-construction de l'outil pour utiliser la juste technologie pour la nature.

Les avantages d'*Ecoteka* incluent son caractère *open-source*, projet libre, ouvert et collaboratif, son équipe pluridisciplinaire avec une double compétence en écologie urbaine et numérique, ainsi que son hébergement sur des serveurs dédiés et sécurisés.

Ecoteka propose plusieurs fonctionnalités, telles que la cartographie, un tableau de bord, la gestion de l'inventaire, la détection LiDAR et l'illustration des trames vertes. Chacune de ces fonctionnalités permet l'accomplissement de diverses actions (**Figure 2**).

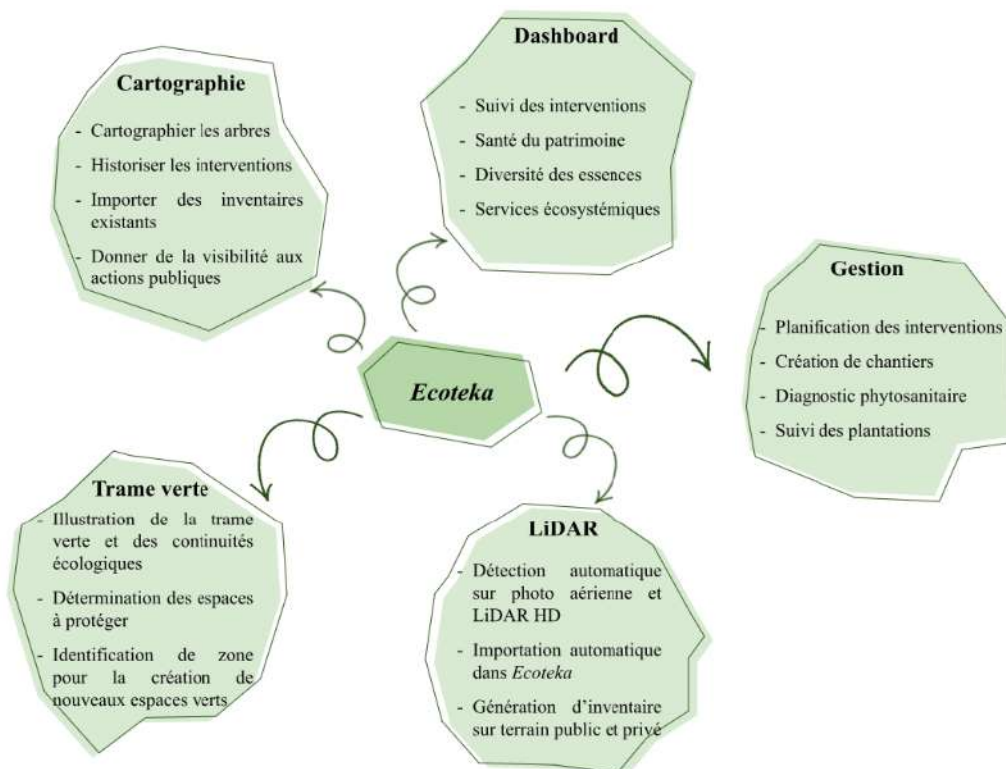


Figure 2 : Les fonctionnalités d'*Ecoteka*

Après avoir exploré les fondements et les objectifs *d'Ecoteka*, il s'agit de comprendre comment cette plateforme fonctionne concrètement et quelles sont les possibilités qu'elle offre aux collectivités et aux professionnels de l'environnement urbain. En examinant de plus près son modèle économique et ses fonctionnalités clés, il sera possible de mieux apprécier la valeur ajoutée *qu'Ecoteka* apporte dans la gestion et la réinvention des écosystèmes urbains.

1.2.3.2. Modèle économique *d'Ecoteka*

Ecoteka est un outil *open-source* basé sur un modèle de double licence, ou dual licensing. Ce modèle permet aux utilisateurs de choisir entre une licence *open-source*, gratuite, mais imposant certaines contraintes, et une licence commerciale, offrant plus de flexibilité. Les petits utilisateurs peuvent ainsi utiliser le logiciel gratuitement, tandis que les entreprises peuvent opter pour une licence commerciale pour éviter les contraintes du copyleft. Les utilisateurs paient pour l'installation du logiciel, un forfait de maintenance pour les mises à jour trimestrielles, et des ajustements de développement spécifiques via des appels à projets ou des offres sur mesure. Ces nouvelles fonctionnalités, une fois développées, sont ensuite intégrées à la version *open-source*, bénéficiant ainsi à l'ensemble de la communauté.

L'outil *open-source* inclut des fonctionnalités de base telles que la cartographie, des tableaux de bord pour le suivi des arbres (nombre d'arbres, stockage de carbone, etc., sans les corridors écologiques), et la gestion des chantiers. Des modules commerciaux supplémentaires, tels que l'intégration de données LiDAR terrestre et aérien, ainsi que la gestion de l'arrosage, sont également disponibles.

Ecoteka cible principalement trois types d'utilisateurs : les collectivités, les professionnels de terrain, et les élus et citoyens. Pour les collectivités, l'outil offre une meilleure connaissance et gestion de leur patrimoine arboré grâce à des fonctionnalités de cartographie détaillée. Les professionnels sur le terrain, tels que les élagueurs, les diagnostiqueurs et les experts arboricoles, bénéficient d'un logiciel de gestion de chantier adapté à leurs besoins. Enfin, pour les élus et les citoyens, souvent peu informés sur les arbres, mais directement concernés, *Ecoteka* fournit un accès simplifié à l'information via un tableau de bord intuitif. Cette approche favorise ainsi la sensibilisation et l'engagement citoyen dans la gestion du patrimoine arboré urbain.

Natural Solutions se positionne ainsi comme un acteur majeur de l'innovation numérique dédiée à la biodiversité. L'agence accompagne de nombreux acteurs dans leurs projets environnementaux. C'est dans ce contexte stimulant que j'ai eu l'opportunité d'effectuer mon stage, me permettant, après avoir présenté l'idée de départ, de contribuer concrètement à cette mission tout en développant mes compétences.

1.3. Missions et déroulement du stage

1.3.1. Objectif et problématique

La dégradation croissante de la nature par les activités humaines perturbe les cycles géo-bio-chimiques, rendant cruciales les initiatives pour la protéger. À la suite de discussions entre le directeur de Natural Solutions et des prospects tels que la CDC biodiversité, une idée novatrice a émergé : centraliser les informations sur la renaturation. Cette initiative ne répond pas à une demande client spécifique, mais découle de signaux faibles, c'est-à-dire de manifestations récurrentes d'une problématique révélatrice d'une difficulté ressentie par des utilisateurs potentiels.

Face à cette constatation, le directeur de Natural Solutions a décidé de lancer un projet visant à regrouper un maximum d'informations sur la renaturation. L'ambition de ce projet est de diffuser et de faciliter l'accès à ces informations pour soutenir la prise de décision en matière de mise en œuvre d'actions de renaturation. Concrètement, il s'agit de concevoir une base de données la plus complète possible, capable d'éclairer le choix des futurs porteurs de projets de renaturation, et de développer un outil permettant d'interagir avec cette base de données.

Le projet aboutirait à la création d'un bot conversationnel destiné à rendre les informations sur les actions de renaturation plus accessibles aux acteurs concernés. Ce bot, alimenté par une base de connaissances riche en documents et informations, serait intégré à *Ecoteka* comme un plugin supplémentaire, disponible pour les utilisateurs de cette plateforme.

Étant donné que le projet ne répond pas directement à un besoin client, il est possible que l'outil développé ne soit pas utilisé, et ce, pour trois raisons. La première est qu'il ne présente pas d'intérêt réel. Néanmoins, ce point n'est pas trop risqué, car le projet est issu de discussions soulevant l'utilité d'un tel outil. Le deuxième risque est un risque de temporalité : il est possible que l'outil soit intéressant, mais ne tombe pas au bon moment. Enfin, le troisième risque est un manque de communication autour de l'outil. Pour pallier ces risques, l'idée est de développer un module à *Ecoteka*, ce qui permettrait d'augmenter les chances d'utilisation de l'outil en le rendant plus pertinent, opportun et visible pour les utilisateurs potentiels.

Le stage a pour objectif d'explorer la création d'une base de connaissances dédiée à la renaturation, en centralisant des informations actuellement dispersées. Ce travail exploratoire vise à affiner l'idée initiale et à préparer le terrain pour le développement d'un prototype fonctionnel. Bien que ce prototype ne soit pas destiné à être une version finalisée, il permettra de démontrer la faisabilité du concept et de tester son utilisation. En parallèle, un outil de visualisation sera développé pour illustrer les informations regroupées. Étant donné que la collecte des données est soumise au règlement général sur la protection des données (RGPD), cet outil sera exclusivement développé en interne pendant la durée du stage.

L'objectif de ce stage consistera à répondre à la problématique suivante :

Comment centraliser et mettre à disposition les données relatives aux actions de renaturation passées afin de pouvoir, à terme, aider à la mise en place et au pilotage de nouvelles actions de renaturation ?

Après avoir clarifié les objectifs de ce projet et défini la problématique centrale, nous allons nous pencher sur les aspects pratiques de sa réalisation. La mise en œuvre efficace de ces ambitions nécessite une planification minutieuse et une structure bien définie. C'est pourquoi il convient à présent d'examiner en détail comment ce stage a été organisé pour répondre au mieux aux enjeux identifiés et atteindre les résultats escomptés.

1.3.2. Organisation du stage

Le stage est découpé en quatre grandes étapes. Les principales étapes du stage comprennent tout d'abord l'**identification** des sources pertinentes de renaturation, puis la **création** d'une base de données pour stocker ces informations. Ensuite, il s'agira d'**intégrer** les données collectées dans cette base, en veillant

à leur qualité et leur cohérence. Enfin, une phase de **visualisation** des données sera entreprise, afin de les rendre accessibles et compréhensibles en fonction des besoins spécifiques des utilisateurs. Chaque grande étape est découpée en plus petite étape dans le but de définir des tâches plus précises pour lesquelles il est possible d'estimer le temps de travail nécessaire. Ces étapes sont réparties sur l'ensemble de la durée du stage (**Figure 3**).

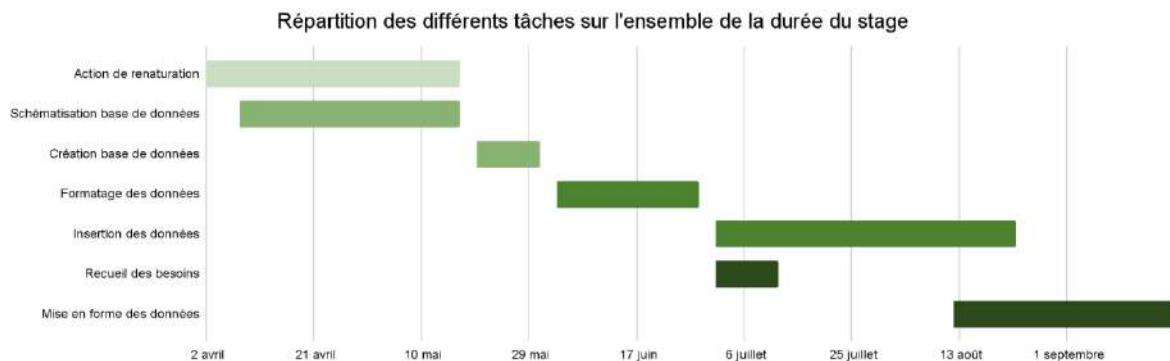


Figure 3 : Diagramme de Gantt du projet

■ Identification ; ■ Création ; ■ Intégration ; ■ Visualisation

Les risques identifiés sont variés. Tout d'abord, il y a un manque d'information sur les actions de renaturation et un manque de recul sur leurs impacts, souvent dû à la nouveauté ou à une documentation insuffisante. Un risque majeur est le manque de fiabilité des données. Une structuration inadéquate, résultant de contraintes de temps, pourrait compliquer leur importation. Il est également possible de viser des objectifs trop larges ou trop ambitieux, diluant ainsi les efforts et les ressources. Enfin, la base de données pourrait devenir inutilisable en raison d'un manque d'informations pertinentes ou de l'incapacité à calculer l'impact des actions de renaturation.

Pour atténuer ces risques, plusieurs pistes peuvent être explorées. D'abord, il est envisageable de contacter directement les acteurs impliqués dans les actions de renaturation afin de recueillir les informations nécessaires à la source. Une communication régulière avec les tuteurs en entreprise peut par ailleurs aider à détecter rapidement les points de blocage potentiels et à y remédier efficacement. De plus, il faut avoir le courage d'abandonner une voie sans issue pour éviter de gaspiller des ressources précieuses dans des efforts vains.

2. Identification des actions de renaturation et construction de la base de données

L'étude sur la renaturation des espaces naturels commence par l'identification des actions menées. Cette démarche implique une méthodologie pour sélectionner des sources d'informations pertinentes, recenser les actions existantes et construire une base de données structurée.

La méthodologie inclut la collecte et le traitement des informations, suivis de l'identification des sources fiables. Un inventaire des actions de renaturation sera réalisé, et la base de données sera modélisée pour répondre aux objectifs de l'étude.

2.1. Méthodologie

Dans cette partie, trois objectifs ont été identifiés. Le premier correspond à l'identification des sources de données. Le deuxième concerne le listing de l'ensemble des informations pertinentes trouvées. Le troisième objectif était de schématiser le modèle de la base de données.

- Identification des sources

Notre objectif était d'identifier le plus grand nombre possible de sources décrivant des actions de renaturation. Pour cela, nous avons posé plusieurs questions, réparties en deux catégories : celles portant sur la méthode de recherche et celles concernant les éléments recherchés.

Pour la méthode de recherche, nous devons déterminer les lieux de recherche, à savoir s'il était préférable de nous orienter vers des recherches sur Internet ou de prendre contact avec des acteurs potentiels. Il était également nécessaire de préciser les types de fichiers d'intérêt, tels que des petites bases de données ou des fichiers CSV.

Concernant les éléments recherchés, nous devons d'abord cibler la zone géographique de nos recherches, que ce soit le monde entier, l'Europe, la France, ou une échelle encore plus restreinte. Nous avons également besoin d'évaluer la pertinence des sources, en distinguant celles qui décrivaient des actions menées par les auteurs eux-mêmes de celles qui étaient de simples listes de sources compilées par différents acteurs. Enfin, il nous fallait clarifier la forme sous laquelle se présentaient les informations et le nombre d'actions décrites par document.

Une fois tous ces points identifiés, nous avons pu réellement commencer la recherche. Elle s'est déroulée sur Internet, en utilisant différents mots-clés tels que "action de renaturation", "restauration écologique", "expérience de renaturation" et "renaturation d'espace". Nous avons également établi des contacts avec des parcs nationaux et des parcs naturels régionaux, mais ceux-ci n'ont pas répondu.

Aucune restriction n'avait été imposée quant au type de fichier : toutes les formes, y compris les pages web, PDF, CSV et Shapefile, ont été prises en compte. La zone géographique a joué un rôle dans le tri des différentes sources identifiées. En effet, nous avons décidé de nous concentrer uniquement sur les actions ayant lieu en France métropolitaine. Néanmoins, l'échelle de ces actions, qu'elle soit communale, départementale, régionale ou nationale, n'avait pas d'importance.

Nous avons choisi de nous concentrer principalement sur des sources ayant elles-mêmes mis en place des actions de renaturation. Ce choix s'expliquait par la précision plus fine des informations sur les actions et par une meilleure justification de certains choix techniques. En revanche, aucune restriction n'avait été appliquée quant au domaine d'activité de la source, qui pouvait aussi bien provenir du secteur public que du secteur privé, et opérer à une échelle régionale, départementale ou nationale.

Pour la forme des actions, nous avons décidé de nous concentrer principalement sur les recueils d'actions pour deux raisons principales :

- Ils permettaient de regrouper plusieurs actions dans une même source, réduisant ainsi le nombre de sources à traiter.
- Ils offraient des descriptions plus concises, facilitant l'accès aux informations recherchées.

Pendant la recherche, nous avons conservé certaines sources même si, au premier abord, elles ne présentaient pas d'actions de renaturation. Ce choix s'expliquait par le fait que la plupart des recueils dépassaient une centaine de pages. Sachant qu'un second tri serait effectué et pour gagner du temps dans l'identification des sources, nous avons décidé de ne pas lire l'intégralité des recueils dans un premier temps. Ainsi, pour éviter d'écartier des sources potentiellement intéressantes lors d'une lecture plus approfondie, nous avons décidé de les garder.

- Listing des actions

Notre objectif était de créer un fichier clair et synthétique permettant de centraliser toutes les sources et informations. Pour atteindre cet objectif, nous avons posé plusieurs questions, réparties en deux catégories : celles concernant la forme du fichier et celles relatives au contenu.

Pour la mise en forme générale du fichier, il s'agissait de déterminer quel type de fichier créer (tableau, fiches de lecture, etc.) et de réfléchir à l'organisation des informations afin que le document soit facilement compréhensible par des personnes extérieures au projet.

Concernant le contenu du fichier, nous devons identifier les actions à conserver, ainsi que les informations pertinentes pour chaque action et celles qu'il était important de mettre en avant. Il fallait également décider comment retranscrire les détails de l'information en comparaison avec d'autres documents.

Une fois tous ces points identifiés, nous avons pu commencer le listing des informations. Nous avons décidé de centraliser toutes les informations dans un même fichier en répertoriant les sources. Nous avons choisi un document Word, avec chaque page représentant une source. Chaque page était conçue comme une fiche de lecture, permettant d'avoir à la fois des tableaux pour identifier rapidement les informations les plus importantes, et des paragraphes pour décrire l'organisation du document, facilitant ainsi la navigation.

Les recueils contenaient parfois un grand nombre d'actions diverses, telles que des actions de renaturation, d'accompagnement financier ou de communication. Étant donné que notre objectif principal était de créer une base de données recensant des actions de renaturation, nous avons décidé de nous concentrer uniquement sur la première catégorie. Ainsi, seules les actions concrètes sur le terrain ont été retenues.

La recherche et l'identification des informations se sont déroulées en deux phases. Tout d'abord, nous avons effectué une lecture complète du document pour identifier les informations présentes dans toutes les actions, ce qui nous a permis d'isoler seulement les informations les plus pertinentes. En effet, certaines informations présentes dans certaines actions étaient soit absentes dans d'autres documents, soit non nécessaires pour le modèle de données. Une fois toutes les informations du document identifiées, nous avons déterminé lesquelles étaient vraiment importantes et donc nécessaires à isoler.

Nous avons décidé de conserver les métadonnées, c'est-à-dire les informations décrivant le document lui-même, telles que le nom de la source, la date de publication et le type de document. La date de publication sert principalement à dater les actions et à évaluer s'il existe ou non d'autres documents sur un compte rendu de l'efficacité des actions. Le type de document sera utile plus tard lors de l'extraction d'informations. De plus, pour chaque document, un petit paragraphe descriptif est rédigé. Celui-ci facilite la compréhension et la navigation dans le document en identifiant les différentes sections

présentes ainsi que les numéros de page correspondants. Pour les pages web, le paragraphe explique toutes les actions possibles pour obtenir le maximum d'informations concernant une fiche de renaturation. En ce qui concerne les informations décrivant l'action, nous avons choisi de conserver les principales : le lieu de réalisation, les partenaires impliqués, la date de réalisation de l'action et le budget alloué.

Nous avons décidé de présenter certaines informations clés dans un tableau coloré. Ce tableau met en évidence les éléments les plus importants, à savoir le lieu de réalisation, les partenaires impliqués, la date de réalisation de l'action et le budget alloué. Le code couleur est composé de rouge, jaune et vert.

- Le rouge indique que l'information n'est pas présente dans le document.
- Le vert indique que l'information est présente pour l'ensemble des fiches actions.
- Le jaune signifie deux choses selon l'information concernée :
 - Pour la localisation, cela signifie qu'elle est présente, mais imprécise, par exemple, uniquement le nom de la commune sans le point géographique précis.
 - Pour les trois autres points, cela signifie que l'information est présente dans le document, mais pas pour l'ensemble des actions.

Pour faciliter le comptage du nombre total d'actions, nous avons également décidé de noter le nombre d'actions de renaturation présentes dans le document, et si les actions présentes étaient ou non accompagnées de documents annexes.

● Modélisation du schéma de la base de données

Notre objectif était de réaliser le schéma de modélisation de la base de données. Pour atteindre cet objectif, nous nous sommes posés plusieurs questions. Tout d'abord, il fallait déterminer à quel moment réaliser le premier jet du schéma de modélisation, c'est-à-dire avant ou après la recherche des différentes sources. Nous devons également réfléchir à l'agencement des différentes tables et identifier, parmi les informations relevées à l'étape précédente, celles qui étaient utiles.

La modélisation s'est déroulée en deux étapes principales. Dans un premier temps, avant la lecture des recueils d'actions, nous avons jugé opportun de lister les informations qui nous semblaient importantes, en proposant une première ébauche de regroupement. L'objectif était d'éviter toute influence des informations trouvées ultérieurement. Ensuite, après avoir pris connaissance des différents recueils, nous avons revisité notre modèle pour y intégrer des éléments supplémentaires. Par exemple, nous avons pris en compte les espaces protégés qui justifiaient certaines actions et ajouté des fichiers complémentaires pour mieux les décrire. Enfin, des ajustements ont été apportés au schéma pour le rendre plus exhaustif, tout en restant prudent dans l'intégration de nouvelles informations.

2.2. Identification des sources d'informations

2.2.1. Choix des sources d'information

À l'issue de la première phase de recherche, nous avons retenu 18 sources potentiellement intéressantes (voir **Tableau 1**). Ces sources couvrent l'ensemble de la France et, pour certaines d'entre elles, les départements d'outre-mer. Parmi ces sources, plus de la moitié traitent de la continuité écologique des cours d'eau, définie comme la libre circulation des espèces animales et végétales ainsi que le transport naturel des sédiments dans un écosystème aquatique.

La plupart des sources sont au format PDF. Nous examinerons dans la suite de ce mémoire les implications de ce type de document sur la récupération d'information. En ce qui concerne les informations pertinentes telles que le lieu de réalisation, les partenaires impliqués, la date de réalisation de l'action et le budget alloué, près de la moitié des sources fournissent ces quatre informations.

Tableau 1 : Recensement de l'ensemble des sources identifiées (Légende : Les cases vertes situées dans la quatrième colonne du tableau indiquent que l'information est présente dans le document.)

Source	Type de document	Type d'action	Information pertinente			
			Où	Qui	Quand	Comb.
Cerema	Shapefile	Mesures compensatoires	Où	Qui	Quand	Comb.
Capitales Françaises de la Biodiversité	Page Web	Actions de renaturation	Où	Qui	Quand	Comb.
France nation verte	PDF	Axe permettant la classification des actions	Où	Qui	Quand	Comb.
OFB	PDF	Renaturations des sols	Où	Qui	Quand	Comb.
Syndicat mixte d'étude et d'aménagement de la Garonne	PDF	Actions réalisées sur la Garonne	Où	Qui	Quand	Comb.
Syndicat du Haut-Rhône	PDF	Actions réalisées sur le Haut Rhône	Où	Qui	Quand	Comb.
Pays Beauce Gâtinais en Pithiverais	PDF	Actions en lien avec la trame verte et bleue	Où	Qui	Quand	Comb.
Agence Régionale Pour la Biodiversité	PDF	Initiatives régionales	Où	Qui	Quand	Comb.
Symvahem	PDF	Continuité écologique sur la rivière de la Hem	Où	Qui	Quand	Comb.
Gest'eau	PDF	Actions réalisées sur le bassin versant des 4 Vallées	Où	Qui	Quand	Comb.
Métropole du Grand Paris	PDF	Projets d'actions de renaturation	Où	Qui	Quand	Comb.
Life tourbière du Jura	PDF	Restauration des tourbières	Où	Qui	Quand	Comb.

ZABR graie	PDF	Actions réalisées sur le bassin Rhône méditerranée	Où	Qui	Quand	Comb.
Site de la trame verte et bleue	Page Web	Retour d'expérience trame verte et bleue	Où	Qui	Quand	Comb.
Site de la trame verte et bleue	PDF	Restauration de la continuité écologique pour les poissons migrateurs	Où	Qui	Quand	Comb.
Agence de l'eau	PDF	Actions réalisées sur le bassin Artois-Picardie	Où	Qui	Quand	Comb.
ZonesHumide.org	Page Web	Retours d'expériences cours d'eau et zones humides	Où	Qui	Quand	Comb.
INPN	Shapefile	Délimitation des différents espaces protégés en France	Où	Qui	Quand	Comb.

2.2.2. Listing des actions

Après avoir identifié toutes les sources, celles-ci ont été compilées dans des fiches de lecture. La **Figure 4** en présente un exemple. Chaque fiche de lecture est constituée de dix sections, numérotées de ① à ⑩.

5. Fiches actions : Nouvelle approche du territoire de la Garonne débordante entre Toulouse et Saint-Nicolas-de-la-Grave ①

https://www.lagaronne.com/sites/default/files/upload/1_plan_actions_tfe_sept2014.pdf ②

Source : Syndicat mixte d'étude et d'aménagement de la Garonne ③

Zone géographique : Garonne ④

Date : 2014 ⑤

Type de données : Fiche actions ⑥

Description :

Recueil des actions de restauration sur la Garonne (p.11). Chaque fiche actions est un tableau de maximum deux colonnes permettant de séparer clairement les différentes rubriques. Action au sens large : travaux, communication, recherche. ⑦

Informations présentes : ⑧Ⓐ

OÙ	QUI	QUAND	COMBIEN
----	-----	-------	---------

Axe ; objectif ; bénéfique ; enjeu partagé ; indicateurs de suivi/réussite ⑨

Autre document : oui, document technique de référence ⑩

Nombre d'actions : ⑩

Action au total __ 32

Figure 4 : Exemple de fiche de lecture

Les sections ①, ②, ③ et ⑤ fournissent des informations générales sur le recueil. Elles incluent le titre ①, l'URL pour accéder au document ②, le nom de l'organisme ayant publié le document ③ et la date de parution ⑤.

La "zone géographique" ④ précise la région de France où se déroulent les différentes actions. Le "type de données" ⑥ décrit la forme sous laquelle les actions sont présentées, par exemple : fiches actions, couches géographiques ou cartographie web. Les fiches actions, généralement d'une à deux pages chacune, détaillent individuellement toutes les actions menées. Les couches géographiques sont des fichiers Shapefile, où chaque ligne correspond à une action. La cartographie web permet de localiser les actions directement depuis le site source.

La "description" ⑦, comme son nom l'indique, présente le document et sa structure. Elle facilite la navigation de l'utilisateur à l'intérieur du document. Indiquer le numéro de page de la première action permet d'identifier rapidement le début. Une description générale de la forme des fiches actions aide le lecteur à savoir où trouver les informations importantes.

Les "informations présentes" ⑧ peuvent être divisées en deux :

- Un tableau Ⓐ
- Une section de liste d'informations supplémentaires ⑨.

Comme expliqué dans la section 2.1.2, une case du tableau Ⓐ peut être de trois couleurs : rouge, jaune ou vert. Dans cet exemple, aucune information sur la date de réalisation des actions n'est disponible, car la case est rouge. En revanche, toutes les fiches actions précisent le montant du budget attribué, puisque la case est verte. Concernant la localisation, la couleur jaune indique que celle-ci est mentionnée, mais de manière imprécise. En effet, le document décrit la localisation sans entrer dans les détails, en se basant sur des cartographies d'actions et des caractéristiques générales du lit de la Garonne, sans fournir de coordonnées géographiques spécifiques. Pour ce qui est des partenaires, la couleur jaune signifie qu'ils ne sont pas mentionnés pour toutes les actions et, dans ce document, les informations

restent vagues. Par exemple, pour les partenaires de communication, il est simplement noté “partenaires classiques”, ce qui manque de clarté.

La deuxième section de la partie “informations présentes” ⑧ est une liste d’informations complémentaires ⑧ décrites dans le document. Dans cet exemple, en plus des quatre informations principales, le document inclut les différents objectifs, les bénéfices de l’action ainsi que les indicateurs de suivi et de réussite.

La section “autre document” ⑨ indique si les actions sont liées à d’autres documents. Dans cet exemple, il s’agit d’un document technique de référence, mais cela peut également inclure des liens vers des vidéos explicatives ou des notes de synthèse.

Le nombre d’actions ⑩ est précisé pour faciliter le comptage total des actions recensées.

Une fois les sources clairement définies, on peut procéder à la structuration des données recueillies de manière cohérente et exploitable. La modélisation de la base de données est une étape qui permet de transformer les informations brutes en un ensemble organisé et facilement accessible.

2.3. Modélisation de la base de données

Douze tables principales ont été créées avec leurs identifiants et les données associées, permettant de regrouper toutes les informations isolées précédemment. La **Figure 5** représente un schéma simplifié de ce modèle.

Dictionnaire de données et le modèle réel sont présents en **Annexe i** et **Annexe ii**.

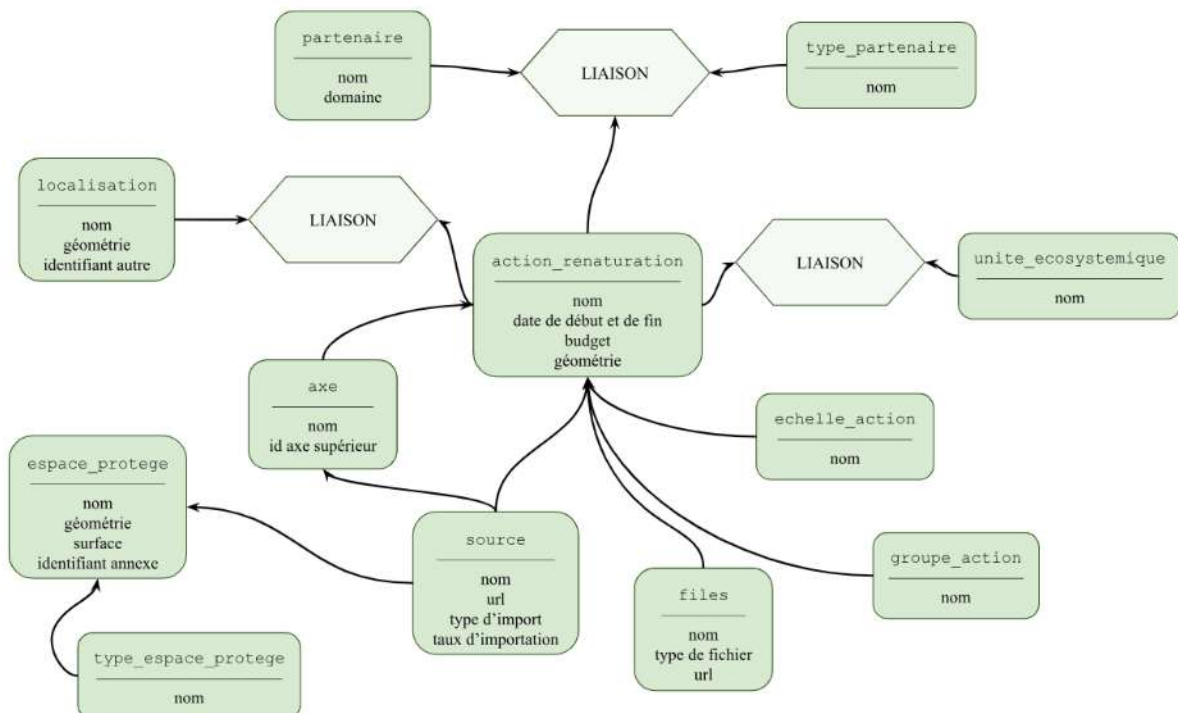


Figure 5 : Schéma simplifié du modèle de données

La table centrale de ce schéma est la table `action_renaturation`, qui comprend le nom, les dates de début et de fin du projet, le budget alloué à l’action et sa géométrie si elle en contient une. Une action

peut avoir un ou plusieurs documents annexes, stockés dans la table `files` avec leur URL, le nom et le type de fichier.

Lors de certains aménagements, plusieurs travaux peuvent être réalisés pour un même projet. Tous les travaux sont enregistrés dans la table `action_renaturation`, et pour indiquer qu'ils font partie d'un même groupe, la table `groupe_action` permet de les lier.

Pour faciliter les filtres de recherche, l'échelle de l'action est placée dans une table distincte. Une unité écosystémique, correspondant au domaine d'action comme l'agriculture, l'économie, la forêt ou le littoral, est également incluse.

Certaines actions sont catégorisées dans des axes, renseignés dans la table homonyme. Un champ `id_parent` a été ajouté pour indiquer l'identifiant de l'axe supérieur, évitant ainsi une chaîne de tables successives.

Concernant la localisation, qui peut être imprécise (comme métropole ou communauté de communes), le nom global donné dans la fiche action est utilisé, et non uniquement des noms de villes. La géométrie de la localisation et d'autres identifiants sont également ajoutés, permettant de faire des liens avec des bases de données telles qu'OpenStreetMap ou Wikidata.

Les partenaires des actions de renaturation sont renseignés dans une table séparée, incluant leur domaine pour simplifier la saisie des données. Le type de partenaire est dans une autre table, car selon les actions de renaturation, un partenaire peut avoir différents rôles (financier, technique, maître d'œuvre, services de collectivité, etc.).

Il a été identifié que la protection de l'environnement influe sur les actions, d'où l'ajout d'une table pour les espaces protégés, permettant des filtres. La table contenant les espaces protégés a une structure similaire à celle des localisations, incluant géométrie, surface et identifiants d'OpenStreetMap et Wikidata pour créer des liens entre les espaces et autres localisations.

Enfin, une table recensant toutes les sources trouvées sur internet et les reliant aux tables `action_renaturation`, `axe` et `espace_protege` a été ajoutée. Outre le nom de la source et l'URL, le type d'import (manuel ou automatique) et le taux d'import de la source sont inclus.

Trois tables de liaison ont été créées pour faciliter les liens entre la table `action_renaturation` et les tables `partenaire`, `type_partenaire`, `localisation` et `unite_ecosystemique`.

Bilan _ Partie 2

Dans cette section, nous avons présenté le point de départ du projet, qui a débuté sans aucune source de données préexistante sur les actions de renaturation. La première étape a donc été de réaliser un inventaire des sources disponibles. Nous avons décrit le processus d'identification, de sélection et de regroupement de ces sources sous forme de fiches de lecture. Nous avons également détaillé la manière dont le modèle de données a été conceptualisé à partir des informations recueillies.

Au total, **dix-huit sources** ont été répertoriées, représentant **8 250 actions**. Cependant, en raison de contraintes de temps, seules deux sources ont été retenues pour l'intégration dans la base de données : les couches géographiques fournies par le Cerema et le site web Capitales Françaises de la Biodiversité, représentant ensemble 7 717 actions. Ces deux sources ont été choisies en raison de leur exhaustivité et de la facilité d'extraction des informations qu'elles offrent.

3. Implémentation de la base de données

Cette section aborde l'implémentation de la base de données, en se concentrant sur les différentes étapes nécessaires à sa conception et à sa mise en œuvre. Nous commencerons par décrire la méthodologie adoptée pour structurer le processus d'implémentation. Ensuite, nous examinerons l'architecture et les solutions technologiques choisies pour l'environnement du code. Enfin, nous détaillerons le processus de remplissage de la base, qui inclut l'importation et la structuration des données initiales.

3.1. Méthodologie

Afin de mieux comprendre comment chaque composant de l'environnement de code s'intègre dans l'architecture globale, les **Figure 6** à **Figure 9** illustrent progressivement cette intégration. À chaque étape, le schéma se complète, offrant une vision claire et détaillée de l'assemblage des différentes briques. Le nouveau composant, représenté en vert, apparaît dans chaque figure pour souligner son introduction et son rôle au sein de l'architecture. Cette progression visuelle permet d'appréhender l'évolution du système au fil des sections, facilitant ainsi la compréhension des interconnexions et des dépendances entre les différents éléments.

- **Système de gestion de base de données**

Dans le cadre du projet, nous nous sommes posés plusieurs questions pour comprendre comment manipuler, optimiser, sécuriser et analyser les informations recueillies précédemment. D'abord, nous nous sommes interrogés sur les outils et techniques à utiliser pour gérer de vastes ensembles de données. Ensuite, nous avons considéré la question de l'optimisation de l'accès et du partage des informations. Nous avons également examiné les mesures nécessaires pour assurer la sécurité et la protection des données. Enfin, nous avons cherché à identifier les outils et techniques les plus efficaces pour l'analyse des données.

Pour répondre à ces questions, les Systèmes de Gestion de Bases de Données (SGBD) se sont révélés être des solutions potentielles. D'autres options incluent les bases de données NoSQL, qui offrent une plus grande flexibilité avec les données non structurées pour centraliser et analyser efficacement de grandes quantités de données. Les SGBD permettent de stocker, organiser et gérer les données de manière centralisée, améliorant ainsi la cohérence et l'intégrité des informations (Opendatasoft, 2023 ;

Mamou, 2024). Ils optimisent l'accès et le partage des données, permettant à plusieurs utilisateurs de collaborer efficacement tout en assurant la sécurité et le contrôle des accès (Garnier, 2023). En termes de sécurité, les SGBD offrent des mécanismes pour contrôler les accès, chiffrer les informations sensibles et réaliser des sauvegardes régulières, protégeant ainsi les données contre les accès non autorisés et les pertes accidentelles (Garnier, 2023 ; Mamou, 2024). Enfin, pour l'analyse des données, les SGBD fournissent des outils puissants pour interroger les bases de données, générer des rapports et effectuer des analyses complexes, aidant ainsi à la prise de décision (Opendatasoft, 2023 ; Eyrolles, 2010).

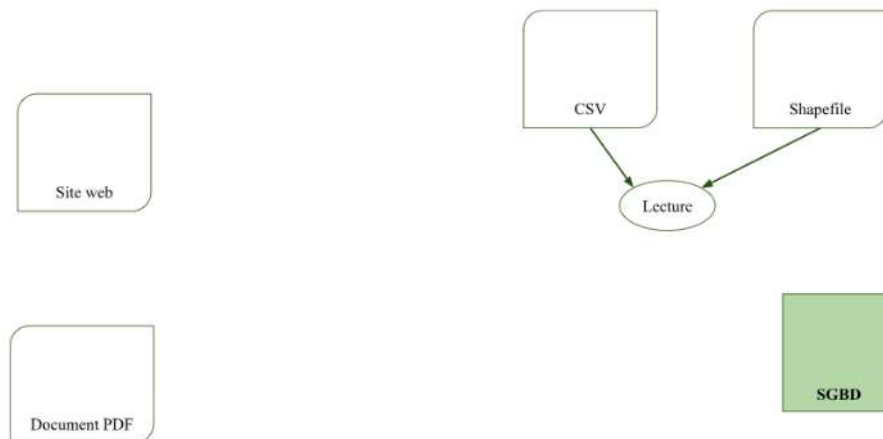


Figure 6 : Architecture de l'environnement de code avec SGBD (cette figure se complétera progressivement au fur et à mesure de l'avancement de cette section)

- **Outil d'extraction**

Lors de la collecte des informations à grande échelle, plusieurs questions se sont posées pour optimiser le processus. Premièrement, nous avons dû déterminer les méthodes les plus efficaces pour économiser du temps lors de la collecte d'informations. Ensuite, il nous a été important de structurer et d'organiser efficacement les données collectées pour optimiser leur exploitation dans une base de données. Cette structuration automatique des données en différents types (noms, adresses, prix, etc.) a facilité leur classement et leur utilisation ultérieure. Une autre question concernait la mise à jour régulière et efficace des informations, étant donné que de nombreux sites web étaient fréquemment mis à jour. Enfin, nous nous sommes interrogés sur les meilleures pratiques pour la collecte et l'analyse d'informations à grande échelle, cherchant des moyens de traiter des volumes importants de données.

Les réponses à ces questions ont mis en lumière les avantages des outils d'extraction. D'autres options incluent l'utilisation de solutions ETL (Extract, Transform, Load) pour automatiser davantage le traitement des données ou encore l'externalisation du processus de collecte à des services spécialisés, bien que cette approche puisse être moins flexible et plus coûteuse. Cependant, les solutions ETL, bien qu'efficaces, sont complexes à installer et nécessitent une courbe d'apprentissage importante, ce qui les rend moins adaptées à ce stage. Les outils d'extraction permettent d'automatiser la collecte de données, économisant ainsi un temps considérable et réduisant les erreurs humaines (Persechini, 2024 ; Joomla, 2024). De plus, ces outils structurent et organisent les données extraites, les rendant facilement exploitables dans des bases de données (Scrapster, 2021). Ils identifient et catégorisent automatiquement différents types d'informations, garantissant ainsi une organisation cohérente et prête à l'utilisation (Joomla, 2024). Ces outils peuvent également être programmés pour des extractions périodiques, assurant que la base de données contient toujours les informations les plus récentes (Persechini, 2024 ; Kinsta, 2023). Enfin, ils permettent de traiter des volumes de données bien

supérieurs à ce que nous pourrions gérer manuellement, facilitant l'enrichissement de la base de données avec des informations externes pertinentes (Persechini, 2024 ; Joomla, 2024).

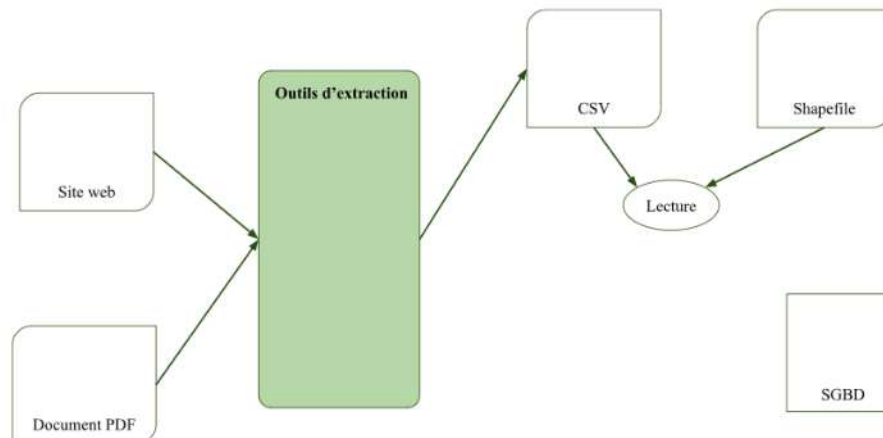


Figure 7 : Architecture de l'environnement du code avec outils d'extraction

- **Framework web Python, lien entre information et base de données**

Pour faciliter les interactions entre les informations récupérées et la base de données, plusieurs questions se sont posées. Nous avons dû déterminer les meilleures pratiques pour sécuriser ces interactions, en garantissant que les données ne soient accessibles que par des opérations autorisées. L'optimisation des interactions avec la base de données a été une autre notion importante pour assurer une performance et une efficacité maximales. Enfin, l'intégration et l'échange de données entre différents systèmes informatiques étaient des points clés, car ils permettaient à diverses applications de fonctionner de manière harmonieuse et de partager des informations de manière fluide.

Pour sécuriser les interactions entre les applications et les bases de données, nous utilisons des frameworks web Python comme intermédiaire sécurisé. Une autre option consiste à utiliser des protocoles de sécurité comme HTTPS pour chiffrer les données en transit. De plus, l'utilisation de pare-feu d'application web aide à détecter et à bloquer les tentatives d'accès non autorisées. Les frameworks web Python agissent comme un tampon, contrôlant l'accès aux données et garantissant que seules les opérations autorisées sont effectuées, ce qui protège l'intégrité des données en évitant les accès directs non contrôlés (Lempereur, 2023 ; RedHat, 2024). En ce qui concerne l'optimisation des interactions, les frameworks web Python standardisent et simplifient ces processus. Ils fournissent des méthodes cohérentes pour récupérer, créer, mettre à jour et supprimer des données, indépendamment de la structure sous-jacente de la base de données (Software AG, 2023). En optant pour ces choix, nous avons pu travailler efficacement sans avoir à maîtriser les détails spécifiques de l'implémentation de la base de données. Enfin, les frameworks web Python facilitent l'intégration et l'échange de données entre différents systèmes et applications. Ils permettent à diverses applications de communiquer et de partager des informations de manière structurée, favorisant ainsi l'interopérabilité et la création de systèmes interconnectés plus complexes (Vi, 2021 ; Harguindeguy, 2022).

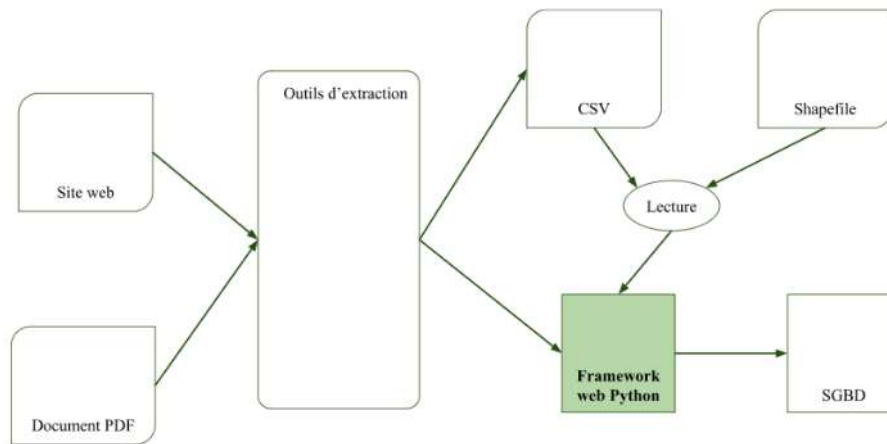


Figure 8 : Architecture de l'environnement du code avec framework web Python

- **Système d'orchestration de conteneurs**

Lors de la gestion et du déploiement du projet, plusieurs questions se sont posées pour assurer une transition fluide entre les environnements de développement. Nous nous sommes interrogés sur la manière d'assurer la cohérence entre les environnements, ainsi que sur les outils et méthodes à utiliser pour gérer les ressources et les performances. La portabilité des applications et la flexibilité dans le choix des environnements d'hébergement ont également été des préoccupations importantes, tout comme la gestion de la flexibilité et de la maintenabilité.

Pour répondre à ces questions, les plateformes de conteneurs se révèlent efficaces. Cependant, d'autres options existent, telles que le déploiement bleu-vert, qui permet de basculer entre deux environnements identiques pour minimiser les interruptions, et le déploiement progressif, qui introduit les nouvelles versions de manière incrémentale pour détecter les problèmes potentiels. Les plateformes de conteneurs nous permettent d'empaqueter une application avec toutes ses dépendances dans un conteneur, assurant ainsi une cohérence parfaite entre les différents environnements et simplifiant le processus de déploiement (RedHat, 2023 ; IBM, 2023). En termes de gestion des ressources, les conteneurs, partageant le même système d'exploitation hôte, sont plus légers et rapides à démarrer, ce qui permet une utilisation plus efficace des ressources matérielles (Veritas, 2022 ; IBM, 2023). Concernant la portabilité, les conteneurs peuvent être facilement déplacés entre divers environnements, facilitant ainsi la mise en œuvre de stratégies multi-cloud (Scafe, 2023 ; Robert, 2024). Enfin, les architectures de microservices soutenues par les conteneurs améliorent la flexibilité et la maintenabilité du projet, nous permettant de gérer indépendamment les différentes parties de l'application (RedHat, 2023 ; Scafe, 2023).

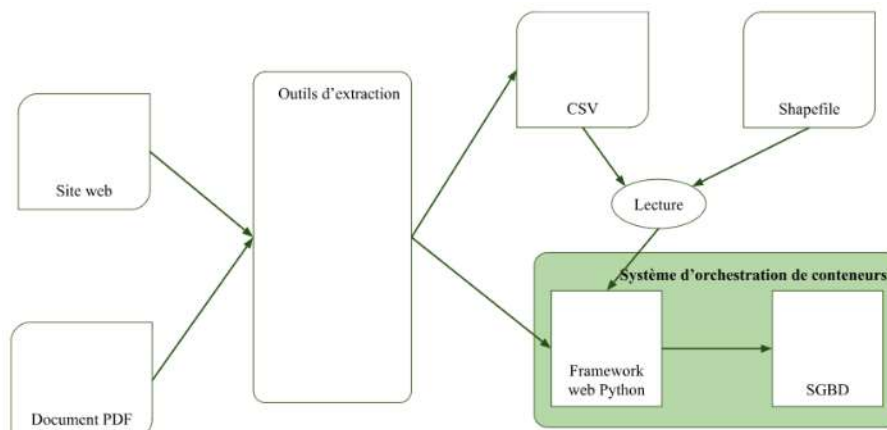


Figure 9 : Architecture de l'environnement du code avec plateformes de conteneurs

3.2. Architecture et solution de l'environnement du code

3.2.1. Système de gestion de base de données

3.2.1.1. Outils existants

Les systèmes de gestion de base de données (SGBD) sont des logiciels pour la gestion, le stockage et la manipulation des données. Il existe plusieurs types de SGBD, chacun adapté à des besoins spécifiques. Les SGBD relationnels, comme MySQL, PostgreSQL et Oracle, sont les plus courants et utilisent des tables pour organiser les données, avec des relations entre elles (Oracle, 2024 ; sql.sh, 2020). Les SGBD NoSQL, tels que MongoDB, Cassandra et Redis, sont conçus pour des applications nécessitant une grande flexibilité et des performances élevées, notamment dans le traitement de grandes quantités de données non structurées ou semi-structurées (sql.sh, 2020 ; Ambient IT, 2021). Les SGBD orientés objets, comme db4o, permettent de stocker des objets directement, facilitant la gestion des données dans des environnements de programmation orientée objet (sql.sh, 2020). Enfin, les SGBD en mémoire, tels que Redis et Memcached, offrent des performances extrêmement rapides en stockant les données directement dans la mémoire vive (sql.sh, 2020 ; Ambient IT, 2021). Chaque type de SGBD présente des avantages et des inconvénients spécifiques, rendant le choix du système dépendant des exigences particulières de chaque projet (Ambient IT, 2021).

Pour notre projet, nous utilisons un SGBD relationnel, car il offre une structure de données robuste et fiable, idéale pour gérer des relations complexes entre différentes entités et assurer l'intégrité des données.

3.2.1.2. Comparaison des outils existants

Le **Tableau 2** présente une comparaison multidimensionnelle de plusieurs critères pour évaluer et choisir un SGBD relationnel. Les principaux aspects à considérer sont :

- **Open-source** : Ce critère est crucial pour garantir une flexibilité et une transparence optimales. Il permet de déterminer si le SGBD est librement accessible et modifiable, offrant ainsi aux utilisateurs la possibilité de personnaliser le système selon leurs besoins spécifiques.
- **Moteur géographique** : La présence d'un moteur géographique est essentielle pour les applications nécessitant la gestion et le traitement de données spatiales. Il permet une gestion précise et efficace des données géospatiales, répondant ainsi aux besoins spécifiques des projets impliquant des informations de localisation.

- **Performances** : Évaluant la rapidité et l'efficacité du SGBD dans le traitement des requêtes, ce critère est indispensable pour les applications nécessitant un traitement rapide des données.
- **Scalabilité** : Mesurant la capacité du système à s'adapter à une charge croissante, ce critère assure la pérennité du système face à l'évolution des besoins de l'entreprise.
- **Sécurité** : Concernant les fonctionnalités de protection des données et de contrôle d'accès, ce critère est essentiel pour garantir l'intégrité et la confidentialité des informations gérées par le SGBD.
- **Utilisation en entreprise** : Ce critère assure que la solution est adaptée aux besoins opérationnels et stratégiques de l'entreprise, garantissant ainsi une intégration harmonieuse dans les processus existants.

Ces critères sont essentiels pour aider les utilisateurs à choisir la solution la plus adaptée à leurs besoins spécifiques en termes de fonctionnalités, de performances et de contexte d'utilisation. L'importance accordée à l'open-source, au moteur géographique et à l'utilisation en entreprise permet de garantir une solution flexible, capable de gérer efficacement les données spatiales et intégrées aux stratégies opérationnelles de l'entreprise.

Tableau 2 : Comparaison des principaux SGBD à l'aide des principaux critères d'évaluation

Critère	MySQL	PostgreSQL	Oracle Database	Microsoft SQL Server	SQLite
Open-source	Oui [1 ; 2]	Oui [2 ; 4]	Non [1 ; 3]	Non [1 ; 3]	Oui [4 ; 5]
Moteur géographique	Oui (via extension) [2]	Oui (PostGIS) [2 ; 4]	Oui (Oracle Spatial) [3]	Oui (intégré) [3]	Oui (Spatialite) [4 ; 5]
Performances	Bonnes [1 ; 2]	Très bonnes [2 ; 4]	Excellentes [1 ; 3]	Très bonnes [1 ; 3]	Bonnes (embarqué) [4 ; 5]
Scalabilité	Moyenne [2 ; 3]	Bonne [2 ; 4]	Excellente [1 ; 3]	Bonne [1 ; 3]	Limitée [4 ; 5]
Sécurité	Bonne [1 ; 2]	Excellente [2 ; 4]	Excellente [1 ; 3]	Très bonne [1 ; 3]	Basique [4 ; 5]
Utilisé dans l'entreprise	Non	Oui	Non	Non	Non

[1] (sql.sh, 2020)

[2] (Pomelo, 2021)

[3] (Ambient IT, 2021)

[4] (AltexSoft, 2023)

[5] (Durantay, Marlé and Ollier, 2024)

Le système de gestion de base de données recommandé pour la solution est **PostgreSQL**.

3.2.2. Outil d'extraction

3.2.2.1. Page web

3.2.2.1.1. Outils existants

Il existe plusieurs outils d'extraction web qui facilitent la collecte de données en ligne. Parmi les plus populaires, on trouve Beautiful Soup, Selenium et Scrapy pour Python, qui permettent de scraper des sites web en analysant le HTML (Vasilis, 2024 ; Peters, 2024). Octoparse et ParseHub sont des solutions sans code qui offrent des interfaces intuitives pour extraire des données sans nécessiter de compétences en programmation (Bhagat, 2024). Pour les utilisateurs d'Excel, Import.io et Web Scraper sont des

extensions pratiques qui permettent d'extraire directement des données dans des feuilles de calcul (Pavlovskyté, 2023). Enfin, Diffbot utilise l'intelligence artificielle pour transformer des pages web en données structurées, simplifiant ainsi le processus d'extraction (Robot, 2022). Chaque outil a ses spécificités, permettant aux utilisateurs de choisir celui qui correspond le mieux à leurs besoins.

Pour ce projet, nous nous concentrons principalement sur les bibliothèques Python, qui offrent des outils puissants et flexibles pour l'extraction et l'analyse des données, et une utilisation en interne.

3.2.2.1.2. Comparaison des outils existants

Le **Tableau 3** présente une comparaison des outils d'extraction web en Python, en se basant sur plusieurs critères définis comme suit :

- **Facilité d'utilisation** : Évalue la simplicité de prise en main et d'implémentation de l'outil. Ce critère est essentiel pour garantir qu'à la suite du stage, les repreneurs du projet puissent facilement le prendre en main.
- **Performance** : Mesure la vitesse et l'efficacité de l'extraction des données, ce qui est crucial pour optimiser le temps et les ressources nécessaires à l'extraction.
- **Flexibilité** : Indique la capacité de l'outil à s'adapter à différents types de sites web et de structures de données. Une flexibilité élevée facilite la mise à jour de la base de données et l'adaptation aux changements des sites web.
- **Documentation** : Évaluée sur sa clarté, sa complétude et son accessibilité pour les utilisateurs, elle est essentielle pour une utilisation efficace et une résolution rapide des problèmes.
- **Communauté** : Reflète la taille et l'activité du groupe d'utilisateurs et de développeurs autour de l'outil. Une communauté active peut offrir un meilleur support et garantir des mises à jour régulières.
- **Gestion JavaScript** : Évalue la possibilité de l'outil à extraire des données de pages web dynamiques utilisant JavaScript. Ce critère est primordial pour de nombreux sites web modernes, car il assure une performance optimale et une interaction fluide avec le site web.

Les critères principaux de notre évaluation incluent la gestion JavaScript, la facilité d'utilisation et la flexibilité. Ces éléments garantissent une interaction fluide avec les sites web, une prise en main aisée par les nouveaux utilisateurs et une adaptation rapide aux diverses structures de données des sites web.

Tableau 3 : Comparaison des principales bibliothèques Python pour le scraping de page web

Critère	Beautiful Soup	Scrapy	Selenium
Facilité d'utilisation	Simple [1 ; 2]	Complexe [1 ; 3]	Modérée [1 ; 3]
Performance	Modérée [2 ; 4]	Élevée [2 ; 3 ; 4]	Lente [3 ; 4]
Flexibilité	Limitée [2 ; 4]	Étendue [2 ; 3 ; 4]	Très étendue [1 ; 3 ; 4]
Documentation	Excellente [2]	Bonne [3]	Bonne [1]
Communauté	Large [2]	Active [3]	Très large [1 ; 3]
Gestion JavaScript	Basique [2 ; 4]	Avancée [3 ; 4]	Excellente [1 ; 3 ; 4]

[1] (Swapnil, 2024)

[2] (Pavlovskyté, 2023)

[3] (Daniel, 2021)

[4] (Dubois, 2022)

L'outil d'extraction web recommandé pour la solution est **Selenium**.

3.2.2.2. PDF

3.2.2.2.1. Outils existants

Il existe divers outils d'extraction de PDF, chacun offrant des fonctionnalités spécifiques pour répondre à différents besoins. Des logiciels comme UPDF permettent d'extraire efficacement des images de PDF, ce qui est particulièrement utile pour les chercheurs et les professionnels du marketing (Deschamps, 2023). Pour l'extraction de données tabulaires, des outils spécialisés tels que Tabula, PdfTables, Docparser, Camelot et Excalibur facilitent le transfert de données de tableaux PDF vers des formats plus utilisables comme Excel ou CSV (Nieters, 2022). Des solutions en ligne comme Smallpdf et ILovePDF offrent une extraction rapide et facile directement via le navigateur, sans nécessiter d'installation. Pour les développeurs, des bibliothèques comme PyPDF2 pour Python ou PDFBox pour Java fournissent des solutions programmatiques robustes pour manipuler et extraire des données de fichiers PDF. Enfin, pour une automatisation plus poussée, des outils d'extraction de données PDF utilisant l'intelligence artificielle, comme ceux mentionnés par Nanonets, offrent des solutions globales pour extraire des données de PDF complexes (Prithiv, 2024).

Pour ce projet, nous nous concentrons principalement sur les bibliothèques Python, qui offrent des outils puissants et flexibles pour l'extraction et l'analyse des données.

3.2.2.2.2. Comparaison des outils existants

Le **Tableau 4** de comparaison des bibliothèques Python pour l'extraction de données PDF se base sur plusieurs critères essentiels :

- **Facilité d'utilisation :** Ce critère évalue la simplicité de mise en œuvre et la courbe d'apprentissage de chaque bibliothèque, garantissant une bonne reprise du projet sans courbe d'apprentissage abrupte.
- **Extraction de texte :** Mesure la capacité à extraire efficacement le contenu textuel des PDF. Une extraction de texte efficace et fiable est indispensable pour obtenir des informations pertinentes à partir de documents textuels.
- **Extraction de tableaux :** Évalue la performance dans la reconnaissance et l'extraction de données tabulaires. L'habileté à extraire des tableaux de manière précise est essentielle pour traiter des données structurées.
- **Gestion des PDF scannés :** Indique la possibilité de traiter les documents numérisés, souvent plus complexes à analyser. Cette capacité est importante pour les documents qui ne sont pas directement accessibles en format texte.
- **Performance :** Se réfère à la vitesse et à l'efficacité globale du traitement des fichiers PDF, un aspect important pour des opérations à grande échelle.
- **Fonctionnalités avancées :** Englobent les capacités supplémentaires telles que l'extraction d'images, la manipulation de métadonnées, ou d'autres fonctions spécialisées offertes par chaque bibliothèque.

Ces critères permettent une évaluation complète et nuancée des différentes options disponibles pour les développeurs Python travaillant avec des fichiers PDF. Les critères principaux à considérer sont l'extraction de tableaux, l'extraction de texte, et la facilité d'utilisation, car ils sont déterminants pour la précision, l'efficacité, et la clarté dans le traitement des documents PDF.

Tableau 4 : Comparaison des principales bibliothèques Python pour le scraping de PDF

Critère	PyPDF2	PDFMiner	Tabula-py	Konfuzio SDK
Facilité d'utilisation	Simple [1]	Complexe [1]	Moyenne [1 ; 2]	Complexe [1]
Extraction de texte	Bonne [1]	Excellente [1]	Limitée [1 ; 2]	Excellente [1]
Extraction de tableaux	Limitée [1]	Bonne [1]	Excellente [1 ; 2]	Excellente [1]
Gestion des PDF scannés	Non [1]	Limitée [1]	Non [2]	Bonne [1]
Performance	Moyenne [1]	Bonne [1]	Bonne [1 ; 2]	Excellente [1]
Fonctionnalités avancées	Basiques [1]	Avancées [1]	Spécialisées (tableaux) [1 ; 2]	Très avancées [1]

[1] (Zyprian, 2023)

[2] (Ariga, 2019)

L'outil d'extraction web recommandé pour la solution est soit **Konfuzio SDK** pour l'extraction des tableaux et du texte, soit **PyPDF2** pour sa prise en main aisée.

3.2.3. Framework Web Python

3.2.3.1. Outils existants

Les frameworks web Python offrent des outils et des bibliothèques pour faciliter le développement de sites et d'applications web. Parmi les plus populaires, on trouve Django, un framework complet et robuste qui suit le principe "batteries included" pour simplifier le développement rapide de sites complexes (Bajaj, 2024 ; Svitla, 2023). Flask, en revanche, est un micro-framework minimaliste qui offre une grande flexibilité et est idéal pour les projets plus petits ou lorsque l'on souhaite une architecture plus modulaire (Bajaj, 2024 ; Tomaszuk, 2024). Pyramid, quant à lui, se situe entre les deux, offrant la flexibilité de Flask avec plus de fonctionnalités intégrées (Tomaszuk, 2024 ; Svitla, 2023). D'autres frameworks notables incluent Tornado, conçu pour des applications nécessitant une gestion concurrente et des performances élevées (Tomaszuk, 2024 ; Svitla, 2023), et FastAPI, reconnu pour sa rapidité et son efficacité dans la création d'API modernes avec une gestion asynchrone des requêtes (Bajaj, 2024). Chacun de ces frameworks répond à des besoins spécifiques et offre des avantages distincts en fonction des exigences du projet (Bajaj, 2024 ; Tomaszuk, 2024 ; Svitla, 2023).

3.2.3.2. Comparaison des outils existants

Le **Tableau 5** comparatif des frameworks web Python repose sur plusieurs critères essentiels pour les développeurs :

- **Écosystème** : Représente l'étendue des bibliothèques, plugins et outils disponibles pour le framework, ce qui peut grandement faciliter le développement et l'extension des projets.
- **Performance** : Évalue la rapidité et l'efficacité du framework dans le traitement des requêtes. Une performance optimale garantit que le système fonctionne de manière efficace et rapide, répondant ainsi aux besoins des utilisateurs.

- **Facilité d'utilisation** : Mesure la simplicité d'apprentissage et d'implémentation du framework pour les développeurs. Une interface et des fonctionnalités intuitives et accessibles minimisent le besoin de formation et augmentent la satisfaction des utilisateurs.
- **Critère asynchrone** : Évalue le support natif ou la facilité d'implémentation de la programmation asynchrone, permettant une meilleure gestion des opérations d'entrée/sortie et une meilleure performance dans certains scénarios.
- **Documentation** : Évalue la qualité, la clarté et l'exhaustivité des ressources officielles et communautaires, assurant une bonne assistance aux développeurs à toutes les étapes de leur projet.
- **Validation des données** : Bien que non explicitement mentionné dans les critères initiaux, ce critère est crucial pour garantir l'exactitude et la fiabilité des informations traitées, réduisant ainsi les risques d'erreurs et de mauvaises décisions.

Ces critères fournissent une évaluation complète et nuancée des frameworks web Python, aidant les développeurs à choisir la solution la plus adaptée à leurs besoins. Les critères principaux à considérer sont la performance, la facilité d'utilisation et la validation des données, car ils garantissent un système efficace, intuitif et fiable.

Tableau 5 : Comparaison des principales bibliothèques Python pour les frameworks web Python

Critère	Django	Flask	FastAPI	Pyramid
Écosystème	Framework complet [2]	Micro-framework [2 ; 5]	Framework API moderne [1 ; 2]	Framework flexible [1]
Performance	Bonne	Très bonne	Excellente [1 ; 2]	Bonne
Facilité d'utilisation	Modérée	Facile [4]	Facile [2]	Modérée
Gestion asynchrone	Limitée	Limitée	Native [1 ; 3]	Limitée
Documentation	Basique	Basique	Avancée (OpenAPI/Swagger) [1 ; 2]	Basique
Validation des données	Manuelle	Manuelle	Automatique (Pydantic) [2 ; 4]	Manuelle

[1] (Tiangolo, 2024)

[2] (Adebord, 2023)

[3] (Baca, 2023)

[4] (Deleuze, 2021)

[5] (Kamunya, 2024)

Le framework web Python recommandé pour la solution est **FastAPI**.

3.2.4. Système d'orchestration de conteneurs

3.2.4.1. Outils existants

Les systèmes d'orchestration de conteneurs sont des outils pour gérer et automatiser le déploiement, la mise à l'échelle et les opérations de conteneurs d'applications. Parmi les plus populaires, Kubernetes se distingue par sa robustesse et sa polyvalence, permettant une orchestration à grande échelle avec des fonctionnalités avancées comme la gestion de la mise à l'échelle automatique et le déploiement sans interruption (RedHat, 2019 ; Kubernetes, 2024). Docker Swarm, bien que plus simple et intégré

nativement à Docker, est apprécié pour sa facilité de configuration et sa rapidité de déploiement (DevUniversity, 2023). Red Hat OpenShift, basé sur Kubernetes, offre une solution clé en main avec des fonctionnalités supplémentaires pour la sécurité et la gestion des applications en entreprise (RedHat, 2019). Enfin, des solutions comme Nomad de HashiCorp se démarquent par leur simplicité et leur capacité à orchestrer non seulement des conteneurs, mais aussi des applications non conteneurisées, offrant ainsi une grande flexibilité d'utilisation (L, 2024).

3.2.4.2. Comparaison des outils existants

Le **Tableau 6** de comparaison des systèmes d'orchestration de conteneurs repose sur plusieurs critères clés :

- **Complexité** : Évalue la facilité d'utilisation et la courbe d'apprentissage du système, ainsi que le niveau de difficulté associé à l'utilisation et à la maintenance de la solution. Une faible complexité facilite l'adoption et l'efficacité des opérations.
- **Scalabilité** : Mesure la compétence du système à gérer efficacement un grand nombre de conteneurs et à s'adapter à la croissance, assurant ainsi la pérennité et l'adaptabilité à l'augmentation des charges de travail.
- **Écosystème** : Représente l'étendue des outils, plugins et intégrations disponibles pour le système, ce qui peut grandement enrichir les fonctionnalités et l'interopérabilité avec d'autres technologies.
- **Déploiement** : Indique la rapidité et la facilité avec lesquelles le système peut être mis en place et rendu opérationnel. La facilité de déploiement est cruciale pour une mise en œuvre rapide et efficace des solutions.
- **Flexibilité** : Évalue la capacité du système à s'adapter à différents environnements et cas d'utilisation, garantissant une utilisation polyvalente et personnalisée selon les besoins spécifiques de l'entreprise.
- **Intégration** : Fait référence à la façon dont le système s'intègre avec d'autres technologies et plateformes, ainsi qu'à sa compatibilité avec différents types de conteneurs et d'applications, facilitant une harmonisation avec l'infrastructure existante.
- **Utilisation en entreprise** : Assure que la solution répond aux besoins opérationnels et stratégiques de l'entreprise, en garantissant que le système est conçu pour les déploiements à grande échelle et les environnements complexes.

Ces critères fournissent une évaluation complète et nuancée des systèmes d'orchestration de conteneurs, aidant les entreprises à choisir la solution la plus adaptée à leurs besoins spécifiques. Les critères principaux à prendre en compte sont l'utilisation en entreprise, le déploiement et la complexité, car ils garantissent une solution qui est à la fois efficace, rapide à mettre en œuvre et facile à gérer.

Tableau 6 : Comparaison des principales bibliothèques Python pour les systèmes d'orchestration de conteneurs

Critère	Docker Swarm	Kubernetes	OpenShift	Nomad
Complexité	Simple [2]	Complexe [2]	Modérée [2]	Simple [2]
Scalabilité	Limitée [1]	Excellente [1]	Excellente [1]	Bonne [2]
Écosystème	Restreint [2]	Vaste [3]	Étendu [3]	Modéré [2]
Déploiement	Rapide [2]	Lent [3]	Modéré [3]	Rapide [2]
Flexibilité	Modérée [2]	Élevée [3]	Élevée [3]	Élevée [2]
Intégration	Native Docker [2]	Polyvalente [3]	Enterprise [4]	Polyvalente [2]
Utilisation dans l'entreprise	Non	Oui	Non	Non

[1] (Laarschot, 2023)

[2] (Ohaba, 2023)

[3] (Fernandes and Ziemiecki, 2023)

[4] (Slingerland, 2023)

Les outils présentés permettent de gérer des clusters de serveurs, mais comme l'application sera déployée sur un serveur unique, les problématiques de mise à l'échelle et de gestion de la capacité de charge ne se posent pas. Dans ce contexte, Docker et Docker Compose sont suffisants pour nos besoins actuels. L'architecture du Docker est décrite dans la section 4.2.3. Toutefois, si à l'avenir, nous devons mettre en place des clusters, **Kubernetes** sera alors l'outil recommandé à NS pour l'orchestration de conteneurs.

3.2.5. Synthèse de l'architecture

Après avoir examiné en détail chacune des solutions techniques employées, il est possible de visualiser leur interconnexion grâce à la **Figure 10**. Cette figure schématise clairement comment chaque choix technique s'intègre au sein du système global du backend, qui est la partie du système informatique responsable du traitement des données et de la logique métier, mettant en évidence les points de connexion et les interactions entre les différentes composantes.

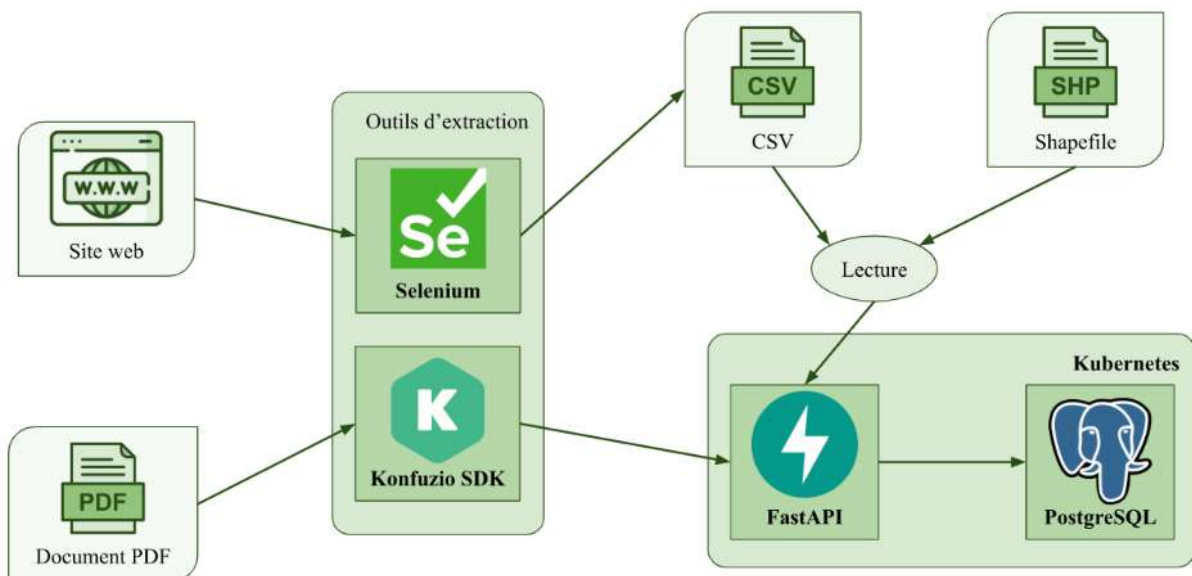


Figure 10 : Architecture détaillée de l'environnement de code du backend

3.3. Remplissage de la base

Le remplissage de la base de données fait partie intégrante du processus de gestion des informations. Cette section se penche sur les différentes méthodes et les outils utilisés pour alimenter la base de données de manière optimale. Nous commencerons par aborder la création et l'importation des tables dont le contenu est statique, souvent sous forme de fichiers CSV, qui permettent une structuration initiale des données. Ensuite, nous explorerons l'intégration de couches géographiques, qui ajoutent une dimension spatiale aux données. Enfin, nous discuterons de l'intégration des données dans un site web, facilitant ainsi l'accès et la manipulation des informations.

3.3.1. Table dont le contenu est statique

L'ajout d'informations à l'aide de fichiers CSV s'effectue dans deux cas de figure spécifiques :

- Tout d'abord, il y a les tables qui ont un nombre prédéfini d'informations. C'est notamment le cas des tables telles que `echelle_action`, `source`, `type_espace_protege`, et `type_partenaire`. Ces tables possèdent une structure de données stable, avec des entrées bien définies et limitées, facilitant ainsi le processus de mise à jour via des fichiers CSV.
- Ensuite, il y a les tables où les informations sont réparties dans de nombreux documents et sites web. Les tables `axe` et `unique_ecosystemique` en sont des exemples typiques. Dans ces cas, les données sont plus dispersées et variées, nécessitant une collecte et une centralisation minutieuses et manuelles pour une insertion correcte dans les bases de données.

Pour procéder à l'ajout des informations, le fichier CSV doit contenir l'identifiant (ID) et les informations nécessaires pour remplir chaque table. Ce fichier est d'abord lu, ligne par ligne, pour extraire les données. Chaque ligne du CSV correspond à une entrée de table, et les informations sont ensuite insérées une à une dans la table respective. Ce processus, bien que répétitif, permet une mise à jour précise et méthodique de la base de données, assurant ainsi l'intégrité et la cohérence des informations stockées.

3.3.2. Couche géographique

L'ajout de fichiers Shapefile, issue de Cerema, est plus complexe que celui des fichiers CSV, car un Shapefile contient bien plus d'informations, lesquelles doivent être réparties dans plusieurs tables (**Figure 11**). Contrairement aux CSV, il n'est pas possible d'ajouter un fichier Shapefile directement. À la place, on utilise un fichier ZIP contenant l'ensemble des fichiers associés à un Shapefile.

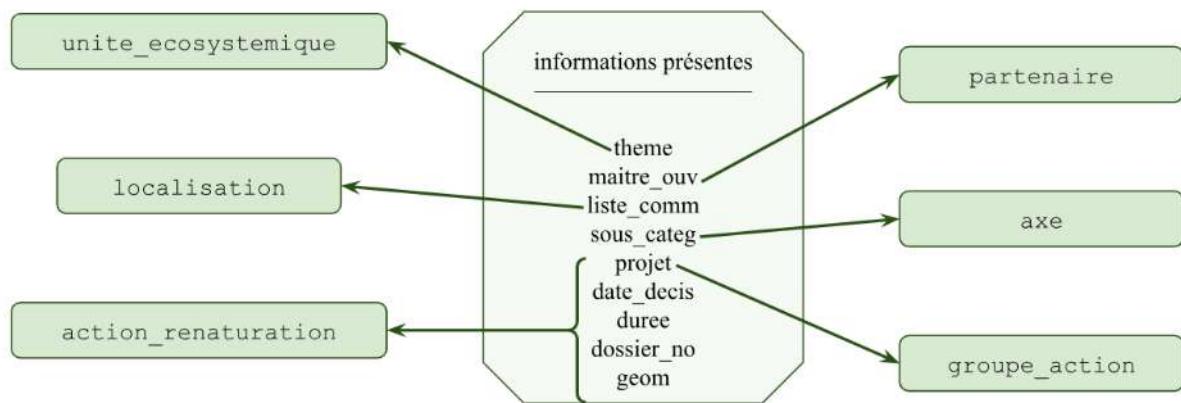


Figure 11 : Schéma de répartition des informations issue des fichiers Shapefile dans les différentes tables

Le processus commence par la récupération du fichier Shapefile à partir de l'archive ZIP et la lecture des données ligne par ligne. L'ajout des informations se déroule en trois étapes distinctes.

La première étape consiste à ajouter l'ensemble des informations dans les tables, à l'exception de la table `action_renaturation`. Pendant cette étape, les données sont traitées si nécessaire, notamment pour les champs concernant les maîtres d'œuvre et les communes. Ces informations sont souvent sous forme de texte séparé par des virgules, et il est nécessaire de les séparer pour pouvoir les intégrer individuellement dans la table. Ensuite, une vérification est effectuée pour déterminer si les entrées existent déjà dans la table. Si elles n'existent pas, elles sont ajoutées, et les identifiants (ID) sont récupérés pour être utilisés dans les tables de liaison.

La deuxième étape concerne l'ajout des actions dans la table `action_renaturation`. Les dates sont converties au format approprié (jj/mm/aaaa). Pour la date de fin du projet, si la date de début et la durée sont spécifiées, un calcul est effectué pour déterminer la date de fin. Une fois les informations correctement formatées, elles sont ajoutées à la table, et l'ID de l'action nouvellement intégrée est récupéré.

Enfin, la troisième étape consiste à ajouter les liens dans les tables de liaison en utilisant les ID récupérés précédemment. Cette étape permet de maintenir l'intégrité référentielle entre les différentes tables et assure que toutes les relations sont correctement établies. Grâce à cette approche méthodique, le processus d'intégration des Shapefile dans la base de données est réalisé de manière structurée et cohérente.

3.3.3. Site web

Le site web de la Capitale Française de la Biodiversité contient 127 pages, chacune présentant 5 actions, soit un total de 635 actions. Cliquer manuellement sur chaque action pour en extraire les informations aurait nécessité un temps de travail considérable, limitant ainsi la possibilité de mener d'autres tâches en parallèle. Pour optimiser ce processus et gagner du temps, nous avons décidé de procéder au scraping du site web. Il est important de noter qu'aucune demande de récupération des données n'a été adressée aux administrateurs du site. Par respect du droit d'auteur, ces informations seront utilisées uniquement en interne, sans aucune monétisation ni partage sans l'accord préalable du site.

L'extraction d'information se fait en plusieurs temps :

- **Analyse de la structure HTML du site :**

L'analyse de la structure HTML du site permet de comprendre la construction des pages et d'identifier les éléments à extraire. La bibliothèque Python utilisée pour le scraping repose sur cette structure. En spécifiant un type de division HTML, il devient possible de localiser et d'extraire les informations contenues dans cette division.

- **Choix du style de programmation :**

Nous avons opté pour la programmation orientée objets (POO). Cette méthode organise le code en "objets", qui regroupent des informations et des méthodes associées. Par exemple, un objet "Voiture" pourrait contenir des caractéristiques telles que la couleur et la marque, ainsi que des méthodes comme "démarrer" ou "freiner". Cette approche facilite la compréhension et la gestion du code, tout en permettant la réutilisation de certaines parties dans différents projets.

- **Extraction des informations :**

Le programme d'extraction est rédigé autour de l'objet "site web", comprenant des méthodes de navigation et d'extraction (changement de page, scroll, etc.). Ces méthodes sont conçues pour être aussi génériques que possible afin de pouvoir être réutilisées sur d'autres sites sans modification majeure. Le programme procède alors à l'extraction en cliquant successivement sur chaque action, récupérant les informations, revenant en arrière, et répétant ce processus pour l'ensemble des pages et actions. Il a fallu adapter le programme pour prendre en compte les variations dans la structure HTML des fiches actions au fil de la navigation.

- **Mise en forme des informations extraites :**

Les données extraites nécessitent une mise en forme pour pouvoir être intégrées dans la base de données. Il était crucial de gérer les caractères spéciaux qui pourraient entraver la recherche ou l'insertion de ces informations.

- Les dates étaient exprimées sous divers formats (mois en toutes lettres, semestres, saisons, etc.). Un dictionnaire a été créé pour associer chaque format à sa valeur numérique correspondante. Si une date sans jour ou sans mois est rencontrée, il a été décidé d'utiliser par défaut le 1^{er} janvier. Il a également été nécessaire de distinguer entre la date de début et la date de fin du projet. Cette distinction a été effectuée en recherchant des mots-clés spécifiques en début de phrase. (**Figure 12**)

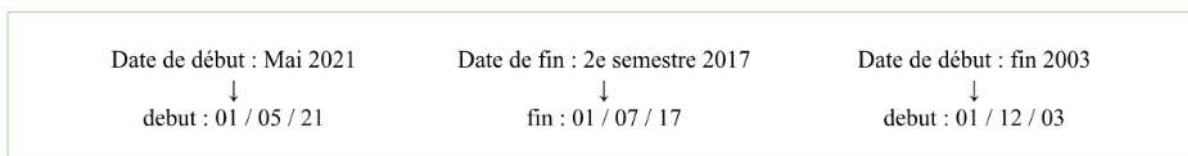


Figure 12 : Exemple d'extraction de date

- La gestion des partenaires est complexe, car il fallait identifier et séparer les partenaires, puis les classer par type (technique, financier, etc.). Une liste de mots clés a été utilisée pour identifier le type de partenaire dans la phrase extraite, ensuite les partenaires étaient isolés à l'aide de séparateurs (virgules, points-virgules, etc.). (**Figure 13**)

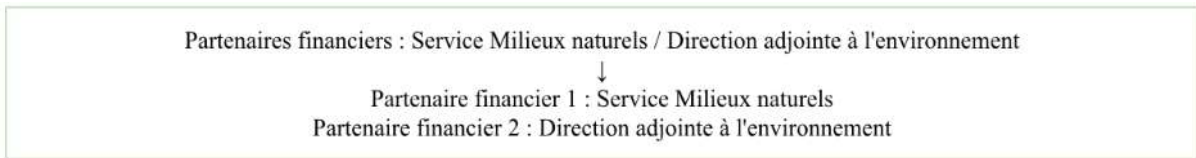


Figure 13 : Exemple d'extraction de partenaires financiers

- Le coût, présenté sous forme de phrase sur le site web, a été extrait en recherchant les parties numériques de la phrase, et en vérifiant si elles étaient suivies de "€" ou "euro". Les montants trouvés ont ensuite été additionnés pour obtenir le coût total de l'action. (Figure 14)

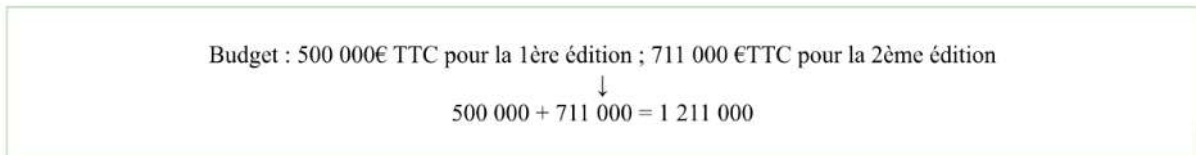


Figure 14 : Exemple d'extraction de budget

- **Enregistrement des informations dans un fichier CSV :**

Une fois extraites, les informations ont été enregistrées dans un fichier CSV (Figure 15). Il n'était pas possible de les intégrer directement à cause des limitations imposées par l'environnement de travail (comme l'utilisation de Docker). Ainsi, seules les données de la première page pouvaient être extraites dans ce procédé.

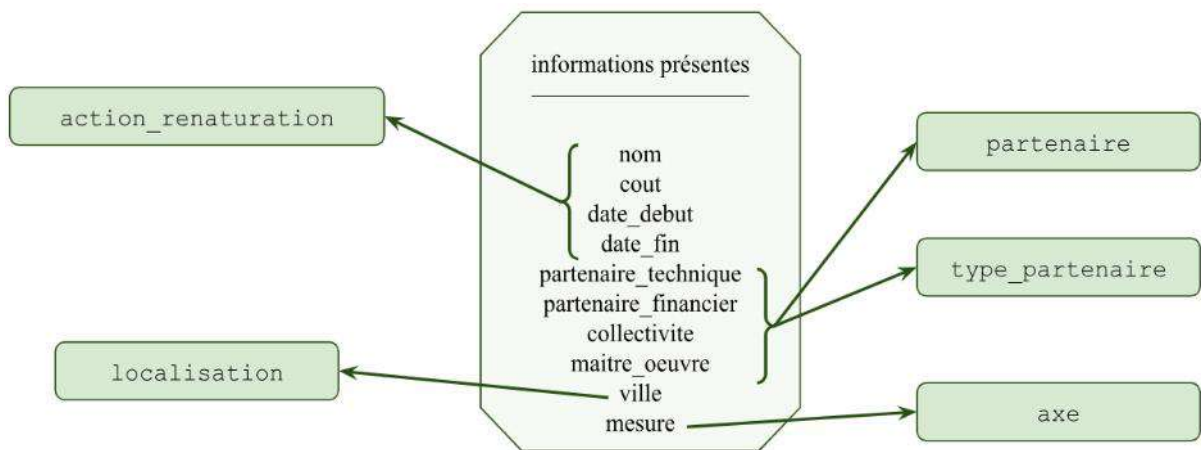


Figure 15 : Schéma des informations présentes dans le CSV et leurs répartitions dans la base de données

- **Insertion des données dans les différentes tables :**

L'intégration des données du fichier CSV dans les bases de données s'est faite en plusieurs étapes identiques à celle expliquée dans la partie précédente. D'abord, les informations hors table `action_renaturation` ont été intégrées, puis celles destinées à la table `action_renaturation` et enfin les tables de liaisons.

Bilan _ Partie 3

Dans cette section, nous avons détaillé la construction du backend de notre application, en expliquant les choix techniques qui ont conduit à son architecture actuelle. Nous avons également illustré comment la base de données a été alimentée à partir de trois types de documents différents : des fichiers CSV remplis manuellement, des fichiers Shapefile, et des CSV obtenus par le scraping de pages web.

La base de données contient désormais **7 717 actions**. Ces actions seront mises en avant dans l'outil de visualisation, qui permettra d'identifier celles qui sont les mieux adaptées à un projet de renaturation.

4. Visualisation

Dans cette section, nous détaillons le processus de création de notre outil de visualisation. Nous abordons d'abord la méthodologie employée, puis l'architecture utilisée, avant de présenter le résultat final. Cette vue d'ensemble permet de comprendre les choix techniques et les étapes clés qui ont conduit à la réalisation de l'outil.

4.1. Méthodologie

Notre objectif était de développer un outil destiné à mettre en avant les informations pertinentes concernant les actions de renaturation, afin de faciliter au mieux la prise de décision lors de la mise en place de nouveaux projets. Pour atteindre cet objectif, nous avons d'abord identifié les informations les plus intéressantes à représenter dans l'outil, et déterminé la forme sous laquelle elles devaient être présentées. Nous avons ensuite réfléchi au nombre de pages web nécessaires, ainsi qu'à la manière de les agencer pour garantir une navigation fluide et intuitive. De plus, nous nous sommes assurés de vérifier s'il existait une charte graphique ou des contraintes visuelles imposées à respecter dans la conception de l'outil.

La réalisation du projet s'est déroulée en plusieurs étapes clés. Tout d'abord, nous avons créé des maquettes pour visualiser l'ensemble des pages nécessaires à l'outil de visualisation, ainsi que l'organisation des différents éléments sur ces pages. Ensuite, nous avons pris rendez-vous avec les principaux acteurs concernés, notamment la designer UX/UI et la *Product Owner*, afin de recueillir leurs avis et leurs demandes de modifications sur les maquettes. Ces retours nous ont permis d'ajuster les maquettes pour mieux répondre à leurs besoins, d'autant plus que l'outil était initialement développé en interne. Enfin, nous avons veillé à respecter la charte graphique *d'Ecoteka* lors de l'intégration future de l'outil, afin d'assurer une cohérence visuelle et une intégration harmonieuse avec les autres éléments de la marque.

4.2. Architecture et solution du frontend

4.2.1. Les frameworks JavaScript

4.2.1.1. Outils existants

Il existe plusieurs frameworks JavaScript populaires, chacun offrant des caractéristiques et des avantages spécifiques pour le développement web. Parmi les plus utilisés, on trouve React, Angular, et

Vue.js. React, développé par Facebook, est une bibliothèque orientée vers la création d'interfaces utilisateur dynamiques et réactives grâce à son architecture basée sur les composants (ACSEO, 2024). Angular, maintenu par Google, est un framework complet qui offre une solution tout-en-un avec des outils intégrés pour le développement d'applications web complexes. Il est basé sur TypeScript et propose des fonctionnalités comme l'injection de dépendances et la gestion des états. Vue.js est un framework progressif qui permet une intégration facile dans des projets existants et offre une courbe d'apprentissage douce grâce à sa syntaxe intuitive (Martin, 2023). D'autres frameworks notables incluent Svelte, qui compile le code à l'avance pour éliminer le runtime, et Ember.js, connu pour sa convention sur la configuration et son écosystème riche. Chacun de ces frameworks a ses propres atouts, rendant le choix dépendant des besoins spécifiques du projet et des préférences des développeurs (Martin, 2023).

4.2.1.2. Comparaison des outils existants

Le **Tableau 7** de comparaison des frameworks JavaScript fournit un aperçu complet de leurs caractéristiques essentielles :

- **Architecture** : Décrit la structure fondamentale du framework, qu'elle soit basée sur des composants, MVC, ou autre. L'architecture adoptée affecte directement la manière dont les applications sont structurées, maintenues et évoluent, impactant la facilité de développement et de scalabilité.
- **Langage** : Spécifie le langage de programmation principal utilisé par le framework, déterminant ainsi la syntaxe et les fonctionnalités disponibles. Le choix du langage a un impact sur la productivité des développeurs, la compatibilité avec leurs compétences existantes et les besoins du projet.
- **Performance** : Met en lumière les mécanismes d'optimisation employés par chaque framework, un facteur crucial pour la réactivité et l'efficacité des applications, assurant une expérience utilisateur fluide et rapide.
- **Courbe d'apprentissage** : Évalue la facilité avec laquelle les développeurs peuvent maîtriser le framework. La courbe d'apprentissage est essentielle pour déterminer la rapidité avec laquelle une équipe de développeurs peut prendre en main la solution, influençant la formation et l'adoption.
- **Communauté et écosystème** : Reflètent le soutien disponible, la richesse des ressources et la vitalité de l'environnement entourant chaque framework. Une communauté active et un écosystème riche sont des éléments clés pour le développement à long terme et la résolution de problèmes.

Les critères principaux à considérer lors du choix d'un framework JavaScript sont le langage utilisé, l'architecture adoptée et la courbe d'apprentissage associée. Ces critères déterminent la compatibilité avec les compétences des développeurs, la manière dont les applications sont conçues et la facilité avec laquelle une équipe peut adopter le framework, assurant ainsi un développement efficace et évolutif.

Tableau 7 : Comparaison des principaux frameworks JavaScript

Critère	React	Angular	Vue.js	Svelte	Ember.js
Architecture	Basée sur les composants [1]	MVC (Model-View-Controller) [1]	Basée sur les composants [1]	Basée sur les composants [2]	Convention over Configuration [3]
Langage	JavaScript [1]	TypeScript [1]	JavaScript [1]	JavaScript [2]	JavaScript [3]
Performance	Virtual DOM [1]	Change Detection [1]	Virtual DOM [1]	Compilation à l'avance [2]	Virtual DOM [3]
Courbe d'apprentissage	Modérée [1]	Abrupte [1]	Douce [1]	Modérée [2]	Abrupte [3]
Communauté et Écosystème	Très large et active [1]	Très large et active [1]	Large et active [1]	En croissance rapide [2]	Large, mais plus restreinte [3]

[1] (Sharma, 2018)

[2] (MacInnis, 2024)

[3] (Adam, 2021)

Le Framework JavaScript recommandé pour le projet est **React**.

4.2.2. Gestionnaire de carte web

4.2.2.1. Outil existant

Les gestionnaires de cartes web sont des outils importants pour intégrer des cartes interactives dans des applications web. Parmi les plus connus, on retrouve Leaflet, une bibliothèque open-source légère et facile à utiliser, idéale pour des projets simples (Agafonkin, 2023). OpenLayers est une autre option populaire, offrant plus de fonctionnalités avancées pour la gestion des données géospatiales et l'intégration de multiples sources de cartes (OpenLayer, 2024). Mapbox GL JS, bien que puissant, est devenu payant pour certains usages, ce qui a conduit à la création de MapLibre GL (Mapbox, 2024). MapLibre est une bifurcation open-source de Mapbox GL JS, qui conserve toutes les fonctionnalités de l'original tout en restant libre et communautaire (MapLibre, 2024). Ce gestionnaire permet de créer des cartes vectorielles haute performance avec une grande flexibilité, le rendant particulièrement adapté aux projets nécessitant des cartes personnalisées et interactives sans dépendre de services propriétaires (Pridal, 2024).

4.2.2.2. Comparaison des outils existants

Le **Tableau 8** de comparaison des gestionnaires de carte web fournit un aperçu complet de leurs caractéristiques essentielles :

- **Performance** : évalue la rapidité et l'efficacité du rendu des cartes et des interactions.
- **Personnalisation** : indique le degré de flexibilité offert pour adapter l'apparence et les fonctionnalités des cartes.
- **Facilité d'utilisation** : mesure l'accessibilité et la simplicité de mise en œuvre pour les développeurs.
- **Open Source** : indique si le gestionnaire est librement accessible et modifiable par la communauté.

- **Fonctionnalités avancées** : englobent la richesse des options disponibles pour la manipulation des données géospatiales et la création de cartes complexes.
- **Compatibilité mobile** : évalue l'adaptabilité et les performances sur les appareils mobiles.

Parmi ces critères, il est important de souligner que l'open source, la personnalisation et les fonctionnalités avancées sont considérés comme les critères principaux. Ces aspects offrent une plus grande flexibilité, un contrôle accru sur le code source et la possibilité de créer des applications cartographiques sophistiquées et sur mesure.

Tableau 8 : Comparaison des principaux gestionnaires de carte web

Critères	Leaflet	OpenLayers	Mapbox GL JS	MapLibre GL
Performance	Rapide [1]	Moyenne [2]	Excellente [3]	Excellente [4]
Personnalisation	Modérée [1]	Élevée [2]	Très élevée [3]	Très élevée [4]
Facilité d'utilisation	Très simple [1]	Complexe [2]	Modérée [3]	Modérée [4]
Open Source	Oui [1]	Oui [2]	Non [3]	Oui [4]
Fonctionnalités avancées	Limitées [1]	Nombreuses [2]	Très nombreuses [3]	Très nombreuses [4]
Compatibilité mobile	Excellente [1]	Bonne [2]	Excellente [3]	Excellente [4]

[1] (Agafonkin, 2023) [2] (OpenLayer, 2024) [3] (Mapbox, 2024) [4] (MapLibre, 2024)

Le gestionnaire de carte web recommandé est **MapLibre GL**. En plus de MapLibre GL, il est important que nous ajoutons le service **Martin** à notre application. Martin est un serveur de cartographie, lié à MapLibre GL, qui génère des tuiles vectorielles à la volée, permettant un affichage cartographique rapide et efficace. Il transforme les données géospatiales en tuiles vectorielles, facilitant ainsi le rendu dynamique et interactif des cartes. En effet, notre application front se connecte à une API pour récupérer des données au format GeoJSON, mais elle se connecte également à Martin pour récupérer un flux de tuiles vectorielles.

4.2.3. Architecture

La **Figure 16** illustre de manière détaillée comment les différentes briques du frontend interagissent pour créer une interface utilisateur à la fois fonctionnelle et esthétique. Le frontend, qui constitue la partie visible et interactive d'une application, regroupe l'ensemble des éléments avec lesquels l'utilisateur interagit. Cette figure met en lumière les connexions et les flux d'information entre ces composants, montrant comment ils s'articulent pour offrir une expérience utilisateur fluide et intuitive.

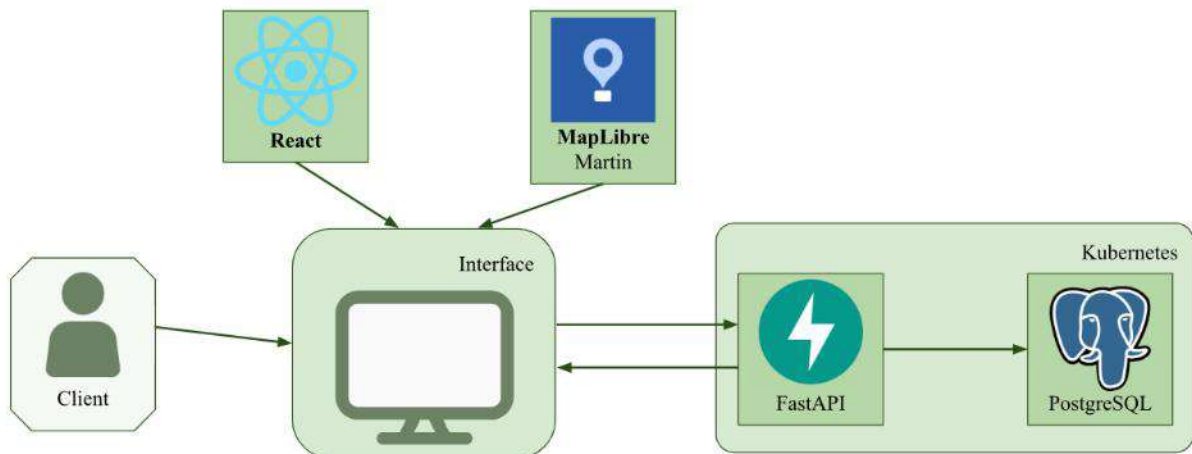


Figure 16 : Architecture de l'environnement du frontend

La **Figure 17** illustre l'architecture des conteneurs utilisés dans notre application à l'aide de Docker Compose. Docker Compose est un outil qui permet de définir et de gérer des applications multi-conteneurs. Dans notre cas, chaque conteneur représente un service distinct de l'application, tels que la base de données, le serveur web, et les différents composants backend. Grâce à Docker Compose, ces services sont définis dans un fichier YAML, où chaque conteneur est configuré avec ses paramètres spécifiques, y compris les images Docker à utiliser, les ports exposés, les volumes montés, et les réseaux auxquels ils sont connectés. Cette architecture modulaire facilite le déploiement, la mise à jour et la gestion des différents composants de l'application, assurant une isolation des services tout en permettant une communication efficace entre eux via des réseaux Docker dédiés. En somme, la **Figure 17** offre une vue d'ensemble claire de la manière dont les différents conteneurs interagissent et se complètent au sein de notre application Docker Compose.

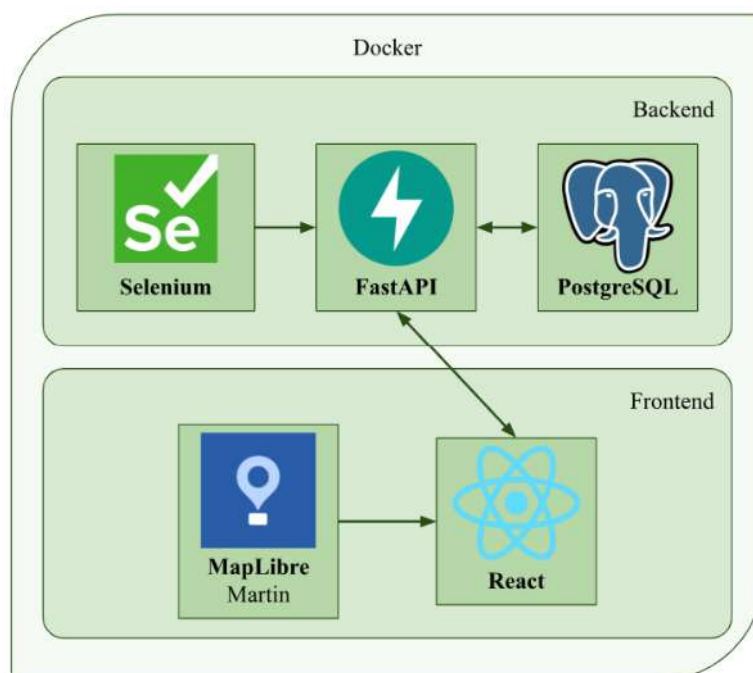


Figure 17 : Architecture détaillée de l'environnement docker

4.3. Le résultat final

Le site se structure autour de quatre pages principales (**Figure 18**), chacune jouant un rôle spécifique dans la présentation et la navigation des informations relatives aux actions de renaturation.

- La première, la **page d'accueil**, sert d'introduction en présentant le principe clé de la renaturation ainsi que le projet qui a rendu possible la création de ce site web. Cette page pose les bases pour les utilisateurs en leur offrant une vue d'ensemble sur les objectifs et l'importance des actions de renaturation.
- La deuxième page, la **page cartographique**, est conçue pour permettre aux utilisateurs de visualiser les actions de renaturation sur une carte interactive. Seules certaines actions sont géoréférencées, offrant ainsi une représentation visuelle claire et intuitive de la répartition partielle des projets sur le territoire.
- La troisième page, intitulée "**Liste d'actions**", propose une vue différente en affichant toutes les actions présentes dans la base de données sous forme de liste. Cette page permet aux utilisateurs de parcourir les actions de manière structurée et d'accéder facilement à des informations spécifiques sans passer par la carte.
- Enfin, la quatrième page est dédiée aux **détails d'une action spécifique**. Accessible uniquement depuis la page cartographique et la page liste d'actions, cette page permet aux utilisateurs de visualiser l'ensemble des informations disponibles pour décrire une action particulière, fournissant ainsi un niveau de détail plus approfondi sur chaque projet.

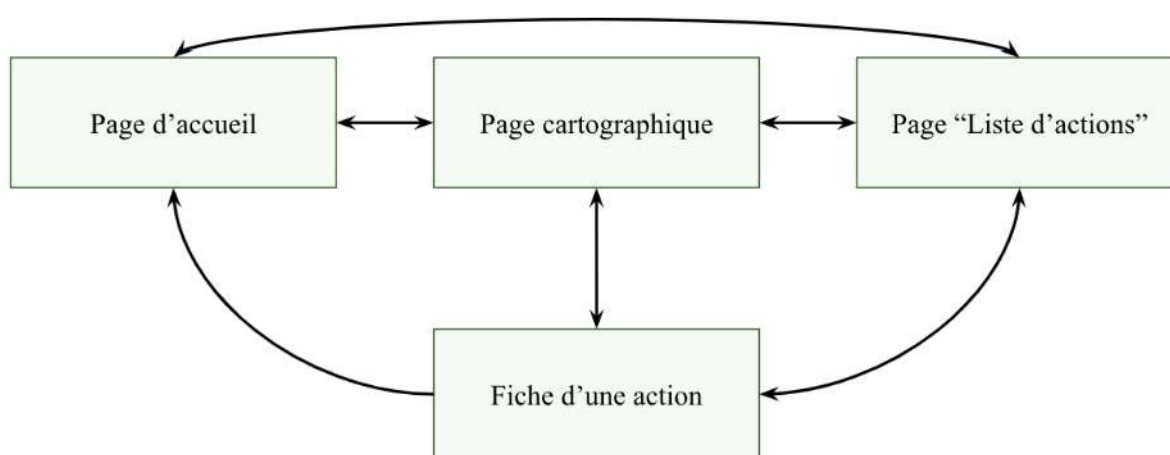


Figure 18 : Architecture du site web

En plus de ces pages, le site intègre un bandeau d'entête (header) (**Figure 19**) présent sur chaque page, conçu pour faciliter la navigation. Ce bandeau offre un accès rapide et direct aux différentes sections du site, garantissant une expérience utilisateur fluide et intuitive.



Figure 19 : Header permettant la navigation sur les différentes pages (activé sur la page accueil)

4.3.1. Page cartographique

La page cartographique (**Figure 20**) est un outil interactif conçu pour offrir une visualisation claire et facile à comprendre de l'ensemble des actions géoréférencées. Elle intègre plusieurs fonctionnalités essentielles pour permettre une navigation fluide et une compréhension approfondie des données affichées.

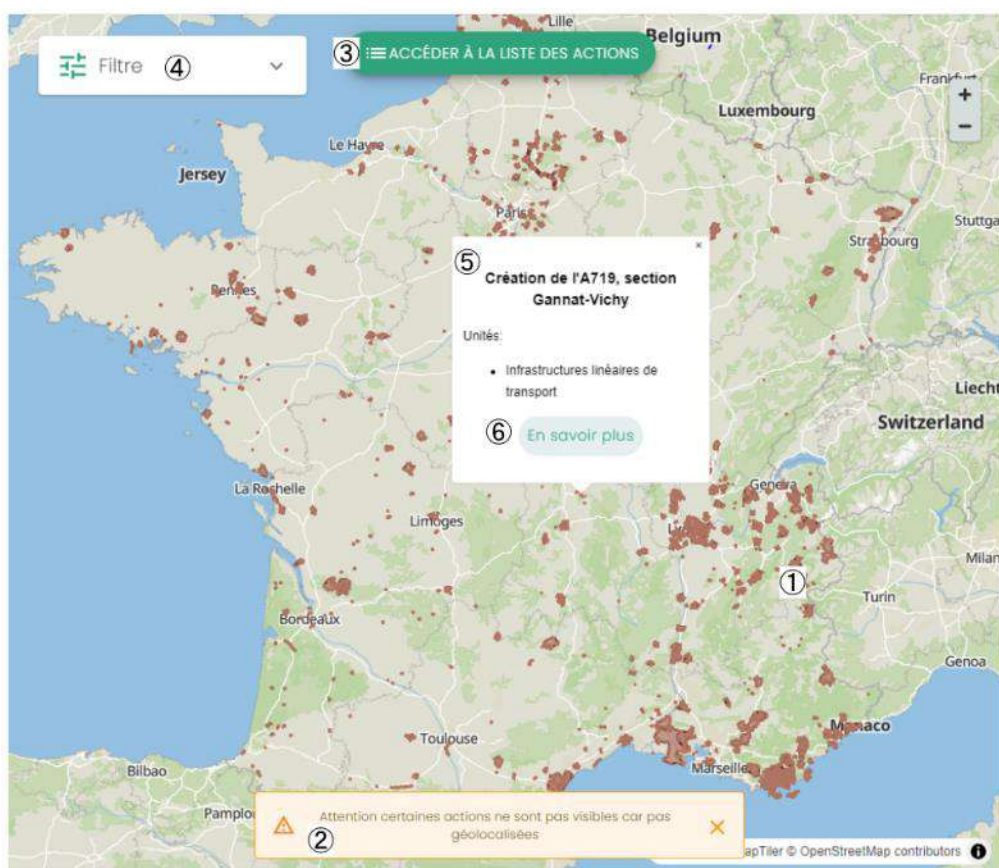


Figure 20 : Page cartographique

La carte présente l'ensemble des actions géoréférencées (1), représentées par des points, des lignes ou des polygones selon leur nature. La couleur de chaque action correspond à une palette définie par la charte graphique d'Ecoteka. Pour améliorer l'expérience utilisateur, la taille des points, l'épaisseur des lignes, ou la visibilité des polygones augmente dynamiquement en fonction du niveau de zoom, facilitant ainsi la visualisation des actions sur des zones spécifiques. La carte est entièrement interactive : l'utilisateur peut se déplacer librement, zoomer et dézoomer pour explorer les différentes actions à travers le territoire.

Un bandeau de couleur orange (2) est visible en bas de la carte pour informer l'utilisateur que certaines actions ne sont pas représentées sur la carte, faute de coordonnées géographiques disponibles. Cette indication vise à garantir une transparence totale sur les données affichées, en précisant que la carte ne contient que les actions disposant d'une localisation précise.

Un bouton (3), placé de manière accessible, permet à l'utilisateur de consulter la page "Liste d'actions". Cette page regroupe l'ensemble des actions disponibles, qu'elles soient géoréférencées ou non, offrant ainsi une vision exhaustive des initiatives répertoriées dans la base de données d'Ecoteka.

Un filtre avancé ④ est intégré à la carte pour permettre à l'utilisateur de restreindre l'affichage des actions en fonction de critères spécifiques. Ce filtre est un outil puissant qui sera détaillé plus en avant dans la section 4.3.3, afin de mieux comprendre ses capacités et son utilisation optimale.

Lorsqu'un utilisateur clique sur une action représentée sur la carte (qu'il s'agisse d'un point, d'une ligne ou d'un polygone), une fenêtre pop-up ⑤ s'ouvre pour fournir des informations supplémentaires. Cette fenêtre affiche le nom de l'action sélectionnée, l'unité écosystémique à laquelle elle se rattache, ainsi que le coût associé si ces données sont présentes dans la base de données. Cette fonctionnalité enrichit l'expérience utilisateur en offrant un accès immédiat à des détails pertinents sur chaque action.

Un bouton "En savoir plus" ⑥ est disponible dans la pop-up, permettant à l'utilisateur de naviguer vers une fiche détaillée de l'action. Cette fiche sera détaillée dans la section 4.3.4.

4.3.2. Liste d'actions

La page "Liste d'actions" (Figure 21) offre un accès complet à l'ensemble des actions répertoriées dans la base de données. Cette page est conçue pour permettre aux utilisateurs de naviguer facilement à travers les différentes initiatives, tout en offrant un aperçu rapide et détaillé de chaque action.

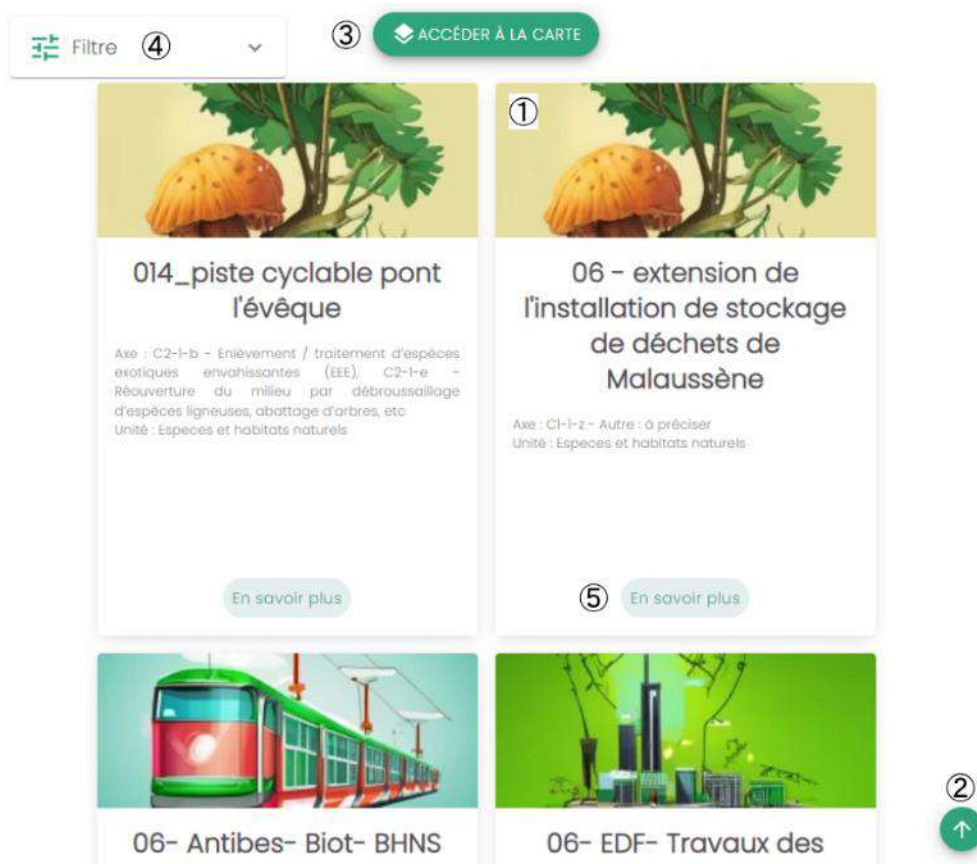


Figure 21 : Page "Liste d'actions"

Chaque action est présentée sous forme de carte individuelle ①, conçue de manière uniforme pour garantir une présentation cohérente et facile à lire. En haut de chaque carte, une image est affichée, correspondant à l'unité écosystémique de l'action, permettant ainsi une identification visuelle rapide.

Juste en dessous, le nom de l'action est clairement indiqué. Au centre de la carte, diverses informations clés sont présentées, telles que l'axe de l'action, l'unité écosystémique, l'échelle d'intervention, et le coût associé, lorsque ces données sont disponibles dans la base de données. La taille des cartes étant fixe, si le texte dépasse l'espace alloué, une barre de défilement horizontale est automatiquement ajoutée pour permettre à l'utilisateur de visualiser l'intégralité du contenu sans perdre en lisibilité.

Un bouton situé en bas à droite (2) de la page permet à l'utilisateur de revenir instantanément en haut de la page en un seul clic. Cette fonctionnalité est particulièrement utile pour faciliter la navigation sur des longues pages, où de nombreuses actions sont listées.

Un autre bouton (3), également placé de manière stratégique, permet d'accéder directement à la page cartographique. Cette option est idéale pour les utilisateurs souhaitant basculer entre la vue d'ensemble des actions sur une carte et leur liste détaillée.

Un filtre (4) avancé est intégré à la page pour permettre aux utilisateurs de trier les actions en fonction de critères spécifiques. Ce filtre, puissant et flexible, sera expliqué en détail dans la section 4.3.3, où ses différentes options et fonctionnalités seront décrites pour maximiser son utilité.

Enfin, un bouton "En savoir plus" (5) est disponible sur chaque carte d'action. Ce bouton donne accès à une fiche détaillée de l'action, où les utilisateurs peuvent consulter des informations plus complètes sur l'initiative, incluant son contexte, ses objectifs, et d'autres détails pertinents. Cette fiche sera expliquée en profondeur dans la section 4.3.4.

4.3.3. Filtre

Comme mentionné précédemment, un filtre (**Figure 22**) est présent sur la page cartographique ainsi que sur la page "Liste d'actions", permettant à l'utilisateur d'afficher uniquement les actions qui correspondent à ses intérêts spécifiques. Ce filtre est composé de cinq sections permettant de trier les actions (1 à 5) et de trois boutons pour interagir avec le filtre (6 à 8).

The image shows a web-based filter interface titled "Filtre". It consists of several sections:

- 1 Localisation:** A dropdown menu for selecting a location.
- 2 Partenaire:** A dropdown menu for selecting a partner.
- 3 Unité écosystémique:** A grid of 16 buttons representing different ecosystem units. The units are:

AGRICULTURE	AMÉNAGEMENT ET DOCUMENTS D'URBANISME	ECONOMIE	EDUCATION À LA NATURE, À L'ENVIRONNEMENT ET AU DÉVELOPPEMENT DURABLE
ESPECES ET HABITATS NATURELS	FORÊT	INFRASTRUCTURES LINÉAIRES DE TRANSPORT	INFRASTRUCTURES ÉNERGÉTIQUES
LITTORAL ET MER	MILIEUX AQUATIQUES ET HUMIDES	MILIEUX URBAIN	PAYSAGE
POLLUTION LUMINEUSE / TRAME NOIRE	PÊCHE ET ACTIVITÉ AQUICOLES	SANTÉ	TOURISME
NULL			
- 4 Echelle:** A row of buttons for selecting the scale of the action: COMMUNALE, DÉPARTEMENTALE, EUROPÉENNE, MONDIALE, NATIONALE, RÉGIONALE, and NULL.
- 5 Budget:** A text input field for specifying a budget.
- 6 Réinitialiser:** A button to reset the filters.
- 7 Annuler:** A button to cancel the current filter selection.
- 8 Filtrer:** A button to apply the selected filters.

Figure 22 : Filtre

- **Sections du Filtre**

La première section du filtre concerne la localisation (1). Étant donné le grand nombre de localisations différentes répertoriées dans la base de données, la sélection se fait via un système d'auto-complétion. L'utilisateur commence à taper les premières lettres de la localisation souhaitée et une liste de suggestions apparaît, basée sur les correspondances trouvées dans la base de données. Ce système simplifie la recherche, surtout lorsque le nombre de localisations possibles est important.

La deuxième section du filtre est dédiée aux partenaires (2) impliqués dans les actions. Comme pour la localisation, le choix du partenaire se fait à l'aide d'une auto-complétion en raison du nombre potentiellement élevé de partenaires dans la base de données. L'utilisateur commence à taper le nom du partenaire, et les suggestions correspondantes s'affichent, facilitant ainsi une sélection rapide et précise.

Le troisième élément du filtre concerne l'unité écosystémique (3). Comme le nombre d'unités écosystémiques est fixe et relativement limité (16 au total), la sélection se fait via des boutons. Chaque unité écosystémique possède son propre bouton, et un bouton supplémentaire est présent pour sélectionner les actions qui n'ont pas d'unité écosystémique renseignée. Cette approche par boutons permet une sélection rapide et visuelle.

La quatrième section du filtre est consacrée à l'échelle (4) de l'action. Comme pour les unités écosystémiques, le nombre d'échelles est fixe (6), ce qui a motivé le choix de boutons pour la sélection. La section contient donc sept boutons : un pour chaque échelle et un supplémentaire pour les actions

sans échelle renseignée. Cette configuration permet à l'utilisateur de sélectionner rapidement l'échelle d'intervention souhaitée.

Le cinquième et dernier élément du filtre est le coût ⑤ de l'action. L'utilisateur peut entrer un montant spécifique, et le filtre affichera toutes les actions dont le coût est inférieur ou égal à ce montant. Cela permet de cibler les actions en fonction de leurs coûts, ce qui peut être particulièrement utile pour gérer les budgets ou évaluer la faisabilité de certaines actions.

- **Boutons d'Interaction**

Le premier bouton, intitulé "Réinitialiser" ⑥, permet à l'utilisateur de supprimer tous les filtres appliqués, revenant ainsi à un état de départ où aucune action n'est filtrée. Cette fonctionnalité est particulièrement utile pour recommencer la recherche à zéro sans avoir à désélectionner manuellement chaque critère.

Le deuxième bouton, "Annuler" ⑦, offre la possibilité de revenir à la configuration de filtre précédente, annulant ainsi la sélection qui vient d'être effectuée. Cela permet à l'utilisateur de corriger rapidement une sélection sans avoir à tout recommencer, en conservant les filtres appliqués précédemment.

Le dernier bouton, intitulé "Filtrer" ⑧, permet de valider et d'appliquer les critères de filtre sélectionnés. Ce bouton a une fonction double : il applique les filtres non seulement à la page sur laquelle l'utilisateur se trouve (cartographique ou "Liste d'actions"), mais également sur l'autre page en parallèle. Ainsi, l'utilisateur peut naviguer entre la page cartographique et la page "Liste d'actions" sans avoir à reconfigurer les filtres, offrant une continuité et une fluidité dans l'exploration des données.

4.3.4. Fiche action

La page "Fiche action" (**Figure 23**) est conçue pour offrir une vue détaillée et complète de chaque action, en rassemblant et en illustrant toutes les informations disponibles dans la base de données. Elle est structurée de manière à faciliter la lecture et la compréhension des détails de l'action.

⑤ ← Retour ① **67 - Déviation de Châtenois - RN59**

③

Nom de l'action	Coût
67 - Déviation de Chât...	Aucune information
Date de début	Date de fin
14/08/2019	Aucune information
Unité écosystémique	Échelle
Infrastructures linéaires...	Aucune information
Axe	
C1-1-a - Création ou renaturation d'habitats et d'habita...	

④

Financier	Maître d'œuvre	Services de collectivité associé	Technique
	DREAL GRAND EST/SERVICE TRANSPORT, Communauté Européenne d'Alsace		

Figure 23 : Fiche action

Le nom de l'action sert de titre principal ① à la page, occupant une position centrale en haut de celle-ci. Ce placement met en valeur l'action et assure que l'utilisateur sait immédiatement sur quelle initiative il a cliqué. Le titre est affiché en grand et en gras pour le distinguer clairement du reste du contenu.

Une carte interactive ② est intégrée pour situer l'action géographiquement, à condition que des coordonnées soient disponibles. Dès l'ouverture de la page, la carte effectue un zoom automatique sur la zone spécifique où se trouve l'action, offrant ainsi une visualisation immédiate de sa localisation. Si l'action n'a pas de coordonnées géographiques associées, la carte reste dézoomée, affichant une vue d'ensemble de l'Europe, permettant à l'utilisateur de conserver une orientation générale.

À droite de la carte se trouve un espace dédié à l'affichage de toutes les informations pertinentes sur l'action ③, telles qu'elles sont consignées dans la base de données. Cet espace présente les éléments suivants : le nom de l'action, le coût associé, la date de début, la date de fin, l'unité écosystémique, l'échelle de l'action et l'axe stratégique. Si une information n'est pas disponible dans la base de données, une mention "Aucune information" est affichée à la place. Cela garantit que l'utilisateur est informé de l'absence de certaines données, plutôt que de supposer qu'elles ont été omises.

Sous la carte et les informations générales, un tableau ④ résume les différents partenaires impliqués dans l'action. Le tableau est organisé en colonnes, chaque colonne représentant un type de partenaire spécifique tel qu'il est entré dans la base de données. Chaque ligne à l'intérieur d'une colonne correspond à un partenaire distinct. Cette disposition permet à l'utilisateur de voir d'un coup d'œil quels partenaires sont engagés dans l'action et de comprendre leur rôle respectif dans sa mise en œuvre.

En haut à gauche de la page, un bouton "Retour" ⑤ permet à l'utilisateur de revenir à la page précédemment consultée, qu'il s'agisse de la page cartographique ou de la page "Liste d'actions". Lors du retour, les filtres appliqués précédemment restent en mémoire et sont réappliqués automatiquement. Cela assure une continuité dans la navigation, permettant à l'utilisateur de poursuivre son exploration des actions sans perdre les critères de sélection qu'il a déjà définis.

4.4. Bilan et perspectives

4.4.1. Bilan du travail réalisé

Le travail réalisé s'est structuré en plusieurs étapes clés, chacune visant à répondre aux objectifs du projet et à démontrer la faisabilité de la solution proposée.

Tout d'abord, une recherche bibliographique approfondie a été effectuée pour recenser un maximum d'actions de renaturation déjà entreprises en France métropolitaine. Cette phase a permis de rassembler une quantité importante de données, contribuant ainsi à un premier regroupement d'informations sur les initiatives de renaturation.

La deuxième étape a consisté en la conception et la réalisation technique d'une base de données robuste et flexible. Celle-ci a été alimentée à partir de deux sources complémentaires : un site web et un fichier géographique. Ces sources ont été sélectionnées en raison de la richesse des informations qu'elles offraient, bien que la forte hétérogénéité des données ait constitué une limite significative dans la perspective de créer une base de connaissances complète sur les projets de renaturation en France. Ainsi, un réel effort a été demandé pour les flux d'extraction et de transformation, notamment avec l'utilisation de Selenium pour le scraping des données. Cette complexité technique a limité le développement de l'orchestration et l'automatisation de flux ETL normalisés. Cet aspect mériterait d'être approfondi après la fin du stage.

L'étape suivante a impliqué l'extraction d'informations spécifiques du site "Capitales Françaises de la Biodiversité", ce qui a permis d'enrichir la base de données avec des exemples diversifiés d'actions de renaturation. Les API développées pour ce projet assurent un échange fluide et sécurisé des informations entre l'application et la base de données. Il se présente cependant ici une autre limite essentielle qui concerne les droits d'utilisation des données. Pour finaliser ce projet, que ce soit en le rendant public ou en le valorisant en interne comme une application RAG (Retrieval Augmented Generation) thématique, il sera nécessaire d'obtenir des accords de coopération avec les auteurs des contenus ciblés lors du stage. Faute de temps, ces accords n'ont pas pu être réalisés au cours du stage.

Enfin, un outil de visualisation a été mis en place pour exploiter les données recueillies. Cet outil comprend une page d'accueil, une carte interactive pour localiser les actions, une liste exhaustive des initiatives répertoriées, et des fiches détaillées pour chaque action. Ce premier outil de visualisation, bien que perfectible, constitue une base solide qui pourra être modifiée et améliorée par la suite.

Un test utilisateur a été réalisé en interne lors d'une démonstration, validant ainsi la fonctionnalité de l'outil et les premières implémentations. Globalement, cette contribution a permis d'établir une première base de données et un outil de visualisation, qui serviront de fondement aux développements futurs du projet.

Concernant les aspects non entièrement explorés, le traitement des fichiers PDF reste à approfondir ainsi que la diminution du temps de chargement de la page "Liste d'actions".

4.4.2. Perspectives du projet

Le projet présente plusieurs perspectives intéressantes pour son développement futur. L'une des premières étapes envisagées est l'importation d'actions supplémentaires à partir des autres sources déjà

identifiées. Cela permettra d'enrichir la base de données existante et d'offrir une vision encore plus exhaustive des initiatives de renaturation en France métropolitaine.

Un autre axe de travail consistera à développer des scripts automatisés pour la mise à jour régulière de la base de données. Ces scripts permettront de maintenir l'outil à jour avec les dernières informations disponibles, tout en réduisant le temps nécessaire pour effectuer ces mises à jour manuellement.

Parallèlement, il sera crucial que les auteurs des sites web utilisés pour recueillir les données soient contactés afin d'obtenir leur autorisation pour un usage plus large que celui prévu en interne. Ce point est d'autant plus important que les droits d'utilisation des données doivent être distingués selon leur nature : le droit d'auteur, qui protège l'ensemble structuré des données, doit être différencié du droit *sui generis*, qui s'applique aux contenus rédigés.

Dans la structure de la base de données construite, une table des sources a été ajoutée pour permettre la connexion des futurs accords de partenariat relatifs à l'usage des données. Sans ces accords, la publication des données sur un site public ne sera pas possible.

Une autre perspective de valorisation des sources identifiées pourrait consister en leur intégration dans un agent conversationnel basé sur des techniques émergentes telles que le RAG (Retrieval Augmented Generation). La base de données construite permettra d'identifier les corpus de textes et de documents pouvant être indexés. Seules pourront être indexées les sources ouvertes ou celles pour lesquelles des accords de partenariat auront été obtenus.

À l'instar de Climate Q&A (<https://climateqa.com/>) développé par Ekimetrics, il sera nécessaire que cet agent conversationnel permette aux utilisateurs d'accéder à des sources d'information pertinentes. Cette démarche est essentielle pour garantir le respect des droits d'auteur et pour établir des partenariats durables avec les sources d'information.

L'implémentation des espaces protégés sur la carte constitue également une perspective clé. En intégrant ces zones, la carte interactive sera davantage enrichie, offrant ainsi une représentation plus complète des territoires concernés par les actions de renaturation.

Le projet prévoit aussi d'élargir son champ d'action en recherchant des initiatives de renaturation au-delà de la France métropolitaine. Cette extension géographique permettrait de comparer les pratiques et les résultats obtenus dans différents contextes, apportant ainsi une dimension internationale à l'outil.

Au regard de la diversité des données disponibles, il serait pertinent de proposer un schéma de données et des métadonnées spécifiques aux actions de renaturation. Cela faciliterait leur agrégation et pourrait s'avérer indispensable pour évaluer leur impact et améliorer leur suivi à différentes échelles spatiales et administratives.

Enfin, à plus long terme, il serait intéressant d'envisager l'intégration de ce projet à Ecoteka. Une telle intégration offrirait une plus grande visibilité au projet tout en le reliant à d'autres initiatives similaires, ce qui encouragerait les échanges et les collaborations dans le domaine de la renaturation.

Bilan _ Partie 4

Dans cette section, nous avons exploré la conception, la structure de l'outil de visualisation et les perspectives du stage. Le site, organisé en **quatre pages** principales, offre une vue d'ensemble cohérente et détaillée des actions répertoriées dans la base de données.

Le projet a débuté par une recherche bibliographique et la création d'une base de données, suivie du développement d'un outil de visualisation incluant une carte interactive. Malgré des défis techniques et des limitations d'usage actuelles, les fonctionnalités ont été validées par un test utilisateur.

Les prochaines étapes comprennent l'enrichissement de la base de données, l'automatisation des mises à jour, et l'obtention des autorisations nécessaires pour une diffusion plus large. Des développements futurs, tels que l'intégration d'un agent conversationnel et l'extension internationale du projet, sont également envisagés, avec la possibilité d'une intégration à Ecoteka pour renforcer sa visibilité et encourager les collaborations.

Conclusion

La problématique initiale de ce mémoire portait sur la centralisation et la mise à disposition des données concernant les actions de renaturation passées, avec pour objectif de faciliter la planification et la gestion de nouvelles initiatives. Le défi était de concevoir un outil permettant aux porteurs de projets futurs de sélectionner les actions les plus adaptées à leurs besoins spécifiques.

Dans le cadre de ce travail, nous avons identifié et regroupé un nombre significatif d'actions de renaturation menées en France métropolitaine. Ces initiatives, auparavant dispersées et difficilement accessibles, ont été centralisées dans une base de données unique. Cette centralisation a pris la forme d'un outil interactif que nous avons développé, dépassant la simple forme d'un répertoire pour devenir un véritable instrument d'analyse. Cet outil permet aux utilisateurs de sélectionner les interventions les plus pertinentes en fonction de critères spécifiques, tels que la localisation, l'échelle de l'action ou encore les partenaires, apportant ainsi un éclairage précieux pour les décisions futures en matière de renaturation et optimisant les stratégies à adopter.

Le point de départ du projet était caractérisé par l'absence de sources de données préexistantes sur les actions de renaturation. La première étape a donc consisté à inventorier les sources disponibles, à partir desquelles le modèle de données a été conceptualisé. Bien que 18 sources aient été répertoriées, seules deux ont été retenues, représentant 7 717 actions, en raison de leur exhaustivité et de la facilité d'extraction des informations.

La construction du backend de l'application a impliqué des choix techniques qui ont conduit à son architecture actuelle. Nous avons opté pour PostgreSQL comme système de gestion de base de données (SGBD) en raison de sa robustesse et de ses capacités de gestion de données géospatiales. Pour alimenter la base de données, nous avons utilisé plusieurs sources : des fichiers CSV manuels, des fichiers Shapefile, ainsi que des données extraites par web scraping à l'aide de Selenium et du Konfuzio SDK. L'application web a été développée en utilisant le framework Python FastAPI, reconnu pour sa performance et sa simplicité. Enfin, le déploiement et la gestion de l'application seraient orchestrés via Kubernetes, assurant ainsi une scalabilité et une gestion optimisée des conteneurs. Grâce à ces choix techniques, nous avons pu constituer une base de 7 717 actions mises en avant dans l'outil de visualisation.

L'outil de visualisation a été conçu et structuré en quatre sections principales : une page d'accueil, une carte interactive permettant de localiser les actions, une liste détaillée des actions réalisées et des fiches spécifiques pour chaque initiative. Chaque composant de l'outil contribue à offrir une vue d'ensemble cohérente et détaillée des actions répertoriées, facilitant ainsi la planification de futures initiatives de renaturation.

À l'avenir, cet outil pourrait être enrichi par l'intégration de nouvelles sources de données et l'ajout de fonctionnalités supplémentaires, comme l'évaluation des impacts des actions de renaturation ou la simulation de scénarios pour de futurs projets. L'élargissement de la base de données à d'autres régions ou pays constituerait également une étape intéressante pour créer un référentiel international des actions de renaturation.

Ce projet s'inscrit dans un cadre global de réflexion sur la renaturation des écosystèmes, un enjeu crucial face aux défis environnementaux actuels. En améliorant la gestion et la conservation des milieux

naturels grâce à une meilleure utilisation des données disponibles, ce travail contribue à la préservation de la biodiversité et à la lutte contre le changement climatique. Il souligne l'importance de la centralisation et de l'accessibilité des données pour encourager des actions concertées et efficaces à grande échelle.

Bibliographie

Mémoire :

- ACSEO (2024) ‘Les technologies ReactJS et React Native par Facebook’. Available at: <https://www.acseo.fr/techno/react/> (Accessed: 22 July 2024).
- Adam, J. (2021) ‘Strengths, weaknesses and use cases of the most popular JS frameworks’, 8 January. Available at: <https://kruschecompany.com/ember-jquery-angular-react-vue-what-to-choose/> (Accessed: 22 July 2024).
- Adebord (2023) *FastAPI, le framework python alternatif à Django et Flask ?*, Créateur de site internet - agence web WordPress. Available at: <https://www.agencedebord.com/blog/fastapi-le-framework-python-alternatif-a-django-et-flask> (Accessed: 18 July 2024).
- Agafonkin, V. (2023) *Leaflet — an open-source JavaScript library for interactive maps*. Available at: <https://leafletjs.com/> (Accessed: 28 August 2024).
- AltexSoft (2023) *Database Management Systems (DBMS) Comparison: MySQL, Postgr, AltexSoft*. Available at: <https://www.altexsoft.com/blog/comparing-database-management-systems-mysql-postgresql-mssql-server-mongodb-elasticsearch-and-others/> (Accessed: 17 July 2024).
- Ambient IT (2021) ‘Top 10 des meilleures bases de données à choisir en 2024’, *Ambient IT*, 24 September. Available at: <https://www.ambient-it.net/meilleures-base-de-donnees/> (Accessed: 17 July 2024).
- Ariga (2019) *tabula-py: Read tables in a PDF into DataFrame — tabula-py documentation*. Available at: <https://tabula-py.readthedocs.io/en/latest/> (Accessed: 18 July 2024).
- Ariluoma, M. et al. (2021) ‘Carbon sequestration and storage potential of urban green in residential yards: A case study from Helsinki’, *Urban Forestry & Urban Greening*, 57, p. 126939. Available at: <https://doi.org/10.1016/j.ufug.2020.126939>.
- Baca, J. (2023) *Le meilleur framework Python : Django vs FastAPI vs Flask*, BacaSoftware. Available at: <https://www.bacasoftware.com/fr/le-meilleur-framework-python-django-vs-fastapi-vs-flask/> (Accessed: 18 July 2024).
- Bajaj, N. (2024) *Comparing Django, Flask, and FastAPI as Python Web Frameworks - Search My Expert Blog*. Available at: <https://blog.searchmyexpert.com/python-web-frameworks-comparison/> (Accessed: 18 July 2024).
- Barra, M. and Grandin, G. (2023) *Étapes essentielles d’un projet de renaturation, Techniques de l’Ingénieur*. Available at: <https://www.techniques-ingenieur.fr/base-documentaire/environnement-securite-th5/genie-ecologique-en-milieu-urbain-42703210/renaturer-les-villes-ge1076/etapes-essentielles-d-un-projet-de-renaturation-ge1076niv10004.html> (Accessed: 28 June 2024).
- Bhagat, R. (2024) *Beautiful Soup or Scrapy or Selenium - Best tool for Python Web scraping? | LinkedIn*. Available at: <https://www.linkedin.com/pulse/beautiful-soup-scrapy-selenium-best-tool-python-web-rutam-bhagat-m2jhf/> (Accessed: 18 July 2024).
- Bouron, J.-B. (2023) *Renaturation, Géoconfluences*. École normale supérieure de Lyon. Available at: <http://geoconfluences.ens-lyon.fr/glossaire/renaturation> (Accessed: 3 June 2024).
- Caillavet, V. and Cabrejas, F. (2020) *Renaturation : la solution pour soutenir la résilience de la ville ?*, *Envies de ville*. Available at: <https://www.enviesdeville.fr/transition-ecologique/renaturation-la-solution-pour-soutenir-la-resilience-de-la-ville/> (Accessed: 3 June 2024).
- Castaing, J., Monod, K. and Noreve, V. (2022) *Renaturer les sols, des solutions pour des territoires durables*. Available at:

- <https://www.ofb.gouv.fr/sites/default/files/Fichiers/Plaquettes%20et%20rapports%20instit/renature-r-les-sols.pdf> (Accessed: 22 April 2024).
- CDC Biodiversité (2023) ‘Renaturer | CDC Biodiversité’. Available at: <https://www.cdc-biodiversite.fr/renaturer-nature2050-compensation/> (Accessed: 14 June 2024).
 - Cerema (2021) *Construire la nature en ville avec les citoyens : 3 questions aux auteures du P'Tit Essentiel | Cerema*. Available at: <http://www.cerema.fr/fr/actualites/construire-nature-ville-citoyens-3-questions-aux-auteurs-du> (Accessed: 28 June 2024).
 - Daniel (2021) *Scrapy Vs Selenium Vs BeautifulSoup – Which Is Best For Web Scraping?* Available at: <https://www.proxyrack.com/blog/scrapy-vs-selenium-vs-beautiful-soup-which-is-best-for-web-scraping> (Accessed: 18 July 2024).
 - Debœuf De Los Rios Serrano, G., Barra, M. and Grandin, G. (2022) *Renaturer les villes, Agence Régionale de la Biodiversité en Île-de-France*. Available at: <https://www.arb-idf.fr/nos-travaux/publications/renaturer-les-villes/> (Accessed: 14 June 2024).
 - Deleuze, C. (2021) ‘Les différences entre les Frameworks Flask et FastAPI’, *pythoniste.fr*, 16 March. Available at: <https://www.pythoniste.fr/python/fastapi/les-differences-entre-les-frameworks-flask-et-fastapi/> (Accessed: 18 July 2024).
 - Deschamps, A. (2023) ‘Les 5 meilleurs outils d’extraction d’images PDF’, *UPDF*, 20 October. Available at: <https://updf.com/fr/edit-pdf/pdf-image-extractor/> (Accessed: 18 July 2024).
 - DevUniversity (2023) *Orchestration Docker : Guide complet sur Swarm et Kubernetes*. Available at: <https://www.devuniversity.com/blog/orchestration-docker-guide-complet-sur-swarm-et-kubernetes> (Accessed: 18 July 2024).
 - Dubois, A. (2022) ‘Scrapy vs BeautifulSoup vs Selenium – Which One to Use?’, *Proxyway*, 17 January. Available at: <https://proxyway.com/guides/scrapy-vs-beautiful-soup-vs-selenium> (Accessed: 18 July 2024).
 - Durantay, Q., Marlé, B. and Ollier, J. (2024) *Découvrez le système de gestion de base de données (SGBD), OpenClassrooms*. Available at: <https://openclassrooms.com/fr/courses/6971126-implementez-vos-bases-de-donnees-relationnelles-avec-sql/7139618-decouvrez-le-systeme-de-gestion-de-base-de-donnees-sgbd> (Accessed: 17 July 2024).
 - Eyrolles (2010) *Tout apprendre sur les systèmes de gestion de bases de données, Developpez.com*. Available at: <https://sgbd.developpez.com/tutoriels/cours-complet-bases-de-donnees/> (Accessed: 17 July 2024).
 - Fernandes, C. and Ziemiecki, G. (2023) *Kubernetes vs Mesos vs Swarm, Sumo Logic*. Available at: <https://www.sumologic.com/blog/kubernetes-vs-mesos-vs-swarm/> (Accessed: 18 July 2024).
 - France nation verte (2023) *Cahier d’accompagnement des porteurs de projet et des services instructeurs*. Available at: https://agence-cohesion-territoires.gouv.fr/sites/default/files/2023-01/Cahier%20accompagnement_Axe2_Renaturation.pdf (Accessed: 28 June 2024).
 - Fu, J. *et al.* (2022) ‘Optimized greenery configuration to mitigate urban heat: A decade systematic review’, *Frontiers of Architectural Research*, 11(3), pp. 466–491. Available at: <https://doi.org/10.1016/j.foar.2021.12.005>.
 - Garnier, A. (2023) *SGBD : qu’est-ce qu’un système de gestion de base de données ?* Available at: <https://blog.hubspot.fr/marketing/sgbd> (Accessed: 17 July 2024).
 - Harguindeguy, B. (2022) *Sécurité des API : le guide complet pour 2022*. Available at: <https://www.pingidentity.com/fr/resources/blog/post/complète-guide-to-api-security.html> (Accessed: 17 July 2024).
 - IBM (2023) *Qu’est-ce que la conteneurisation ?* Available at: <https://www.ibm.com/fr-fr/topics/containerization> (Accessed: 17 July 2024).

- Jeanneau, A. (2024) *Bénéficiaire d'une subvention départementale pour la renaturation de sols imperméabilisés - Mon Espace Collectivité*. Available at: <https://monespacecollectivite.incubateur.anct.gouv.fr/ressource/236/> (Accessed: 28 June 2024).
- Joinet, H., Laruelle, N. and Delaville, D. (2023) *La renaturation des quartiers en politique de la ville, une dynamique à intensifier*, L'Institut Paris Region. Available at: <https://www.institutparisregion.fr/nos-travaux/publications/la-renaturation-des-quartiers-en-politique-de-la-ville-une-dynamique-a-intensifier/> (Accessed: 13 June 2024).
- Joomla (2024) *Web scraping : quelle utilité pour votre entreprise ?* -. Available at: <https://www.wmagence.com/web-scraping-quelle-utilite-pour-votre-entreprise.html> (Accessed: 17 July 2024).
- Kamunya, T. (2024) *FastAPI vs. Flask : lequel des deux vous convient le mieux ?* | Geekflare, Geekflare France. Available at: <https://geekflare.com/fr/fastapi-vs-flask/> (Accessed: 18 July 2024).
- Kinsta (2023) *Qu'est-ce que le Web Scraping ? Comment extraire légalement du contenu web*, Kinsta®. Available at: <https://kinsta.com/fr/base-de-connaissances/web-scraping/> (Accessed: 17 July 2024).
- Kubernetes (2024) *Solution professionnelle d'orchestration de conteneurs*. Available at: <https://kubernetes.io/fr/> (Accessed: 18 July 2024).
- L, B. (2024) *Kubernetes – Tout savoir sur la plateforme d'orchestration de containers*, LEBIGDATA.FR. Available at: <https://www.lebigdata.fr/kubernetes-definition> (Accessed: 18 July 2024).
- Laarschot, L.L.V. de (2023) 'Docker Swarm VS Kubernetes VS Openshift', *Next2IT*, 26 January. Available at: <https://www.next2it.co.uk/docker-swarm-vs-kubernetes-vs-openshift/> (Accessed: 18 July 2024).
- Latawicz, A. *et al.* (2016) 'Natural regeneration and biodiversity: a global meta-analysis and implications for spatial planning', *Biotropica*, 48, pp. 844–855. Available at: <https://doi.org/10.1111/btp.12386>.
- Légifrance (2023) *Section I : Sites naturels de compensation, de restauration et de renaturation, Article L163-I-A*. Available at: <https://www.legifrance.gouv.fr/codes/id/LEGISCTA000048245737> (Accessed: 14 June 2024).
- Légifrance (2024) *Chapitre III : Restauration de la biodiversité, renaturation et compensation des atteintes à la biodiversité, Articles L163-I-A à L163-5*. Available at: <https://www.legifrance.gouv.fr/codes/id/LEGISCTA000048248812> (Accessed: 14 June 2024).
- Lempereur, K. (2023) *11 bonnes pratiques de sécurité API : nos conseils pour protéger vos actifs numériques*, DataDome. Available at: <https://datadome.co/fr/menaces/11-meilleures-pratiques-de-securite-api-conseils-pour-protoger-vos-actifs-numeriques/> (Accessed: 17 July 2024).
- L'Institut Region Paris (2023) *Cycle de webinaires : La renaturation en Île-de-France*, L'Institut Paris Region. Available at: <https://www.institutparisregion.fr/environnement/la-renaturation-en-ile-de-france/> (Accessed: 13 June 2024).
- MacInnis, J. (2024) *React vs Vue.js vs Angular vs Svelte vs Next.js vs Ember.js vs Gatsby vs Astro*, Medium. Available at: <https://javascript.plainenglish.io/react-vs-vue-js-vs-angular-vs-svelte-vs-next-js-vs-ember-js-vs-gatsby-vs-astro-33dc4e30a31e> (Accessed: 22 July 2024).
- Mamou, M. (2024) 'Système de Gestion de Bases de Données (SGBD) : Définition', *Formation Data Science | DataScientest.com*, 17 January. Available at: <https://datascientest.com/sghd-tout-savoir> (Accessed: 17 July 2024).
- Mapbox (2024) *Mapbox GL JS*, Mapbox. Available at: <https://docs.mapbox.com/mapbox-gl-js/guides/> (Accessed: 28 August 2024).
- MapLibre (2024) *MapLibre*. Available at: <https://maplibre.org/> (Accessed: 28 August 2024).

- Marchangelo, P. (2023) *Zones préférentielles de renaturation : le décret est paru*. Available at: <https://www.banquedesterritoires.fr/zones-preferentielles-de-renaturation-le-decret-est-paru> (Accessed: 28 June 2024).
- Martin, S. (2023) ‘Top Reasons to Choose ReactJS for Your Next Project’, *ValueCoders / Unlocking the Power of Technology: Discover the Latest Insights and Trends*. Available at: <https://www.valuecoders.com/blog/technology-and-apps/5-reasons-choose-facebooks-reactjs/> (Accessed: 22 July 2024).
- Natural Solutions (2023) *Notre agence*. Available at: <https://www.natural-solutions.eu/lagence> (Accessed: 6 June 2024).
- Nexity (2021) *Présentation Natura City.pdf*. Available at: <https://www.odeys.fr/sites/default/files/2021-05/Pr%C3%A9sentation%20Natura%20City.pdf> (Accessed: 28 June 2024).
- Nieters, R. (2022) *Extraire efficacement les données de fichier PDF*, *Klippa*. Available at: <https://www.klippa.com/fr/blog/informations/extraire-donnee-pdf/> (Accessed: 18 July 2024).
- OFB (2020) *Six questions sur la biodiversité en France*. Available at: https://naturefrance.fr/sites/default/files/2021-01/PublicationONB_2020_V14_comprese_0.pdf (Accessed: 17 May 2024).
- OFB, A. (2023) *La biodiversité française en déclin 10 ans de chiffres-clés par l’Observatoire national de la biodiversité*. Available at: https://naturefrance.fr/sites/default/files/2023-06/Publication_10ansONB_web.pdf (Accessed: 17 May 2024).
- Ohaba, E. (2023) *Container Orchestration Smackdown: Kubernetes vs. Docker Swarm vs. Nomad / HackerNoon*. Available at: <https://hackernoon.com/container-orchestration-smackdown-kubernetes-vs-docker-swarm-vs-nomad> (Accessed: 18 July 2024).
- Opendatasoft (2023) ‘Système de gestion de base de données (SGBD)’, *Opendatasoft*. Available at: <https://www.opendatasoft.com/fr/glossaire/systemes-de-gestion-de-base-de-donnees/> (Accessed: 17 July 2024).
- OpenLayer (2024) *OpenLayers - Welcome*. Available at: <https://openlayers.org/> (Accessed: 28 August 2024).
- Oracle (2024) *PostgreSQL vs MySQL*. Available at: <https://www.oracle.com/fr/database/postgresql-versus-mysql/> (Accessed: 17 July 2024).
- Pavlovskytè, E. (2023) *Scrapy vs. BeautifulSoup: A Comparison of Web Scraping Tools*. Available at: <https://oxylabs.io/blog/scrapy-vs-beautifulsoup> (Accessed: 18 July 2024).
- Persechini, B. (2024) ‘Les avantages de l’automatisation du scrapping pour votre entreprise’, *Agence No Code à Mulhouse (Alsace)*. bruno persechini, 2 February. Available at: <https://www.seinsights.com/les-avantages-de-lautomatisation-du-scrapping-pour-votre-entreprise/> (Accessed: 17 July 2024).
- Peters, R. (2024) *Scrapy Vs. BeautifulSoup: Which is the best choice for your business?*, *Bright Data*. Available at: <https://brightdata.com/blog/web-data/scrapy-vs-beautiful-soup> (Accessed: 18 July 2024).
- Pomelo, H. (2021) ‘Quel système de gestion de bases de données choisir ?’, *Hello Pomelo*, 17 March. Available at: <https://hello-pomelo.com/articles/quel-systeme-de-gestion-de-bases-de-donnees-choisir/> (Accessed: 17 July 2024).
- Pridal, P. (2024) ‘maplibre/maplibre-gl-js’. MapLibre. Available at: <https://github.com/maplibre/maplibre-gl-js> (Accessed: 28 August 2024).
- Prithiv, S. (2024) *Extraire des données d’un PDF : 6 méthodes d’extraction*. Available at: <https://nanonets.com/blog/fr/extraire-des-donnees-de-fichiers-pdf/> (Accessed: 18 July 2024).

- RedHat (2019) *L'orchestration des conteneurs, qu'est-ce que c'est ?* Available at: <https://www.redhat.com/fr/topics/containers/what-is-container-orchestration> (Accessed: 18 July 2024).
- RedHat (2023) *La conteneurisation, qu'est-ce que c'est ?* Available at: <https://www.redhat.com/fr/topics/cloud-native-apps/what-is-containerization> (Accessed: 17 July 2024).
- RedHat (2024) *Sécurité des API [DO NOT OVERWRITE FR PAGE]*. Available at: <https://www.redhat.com/fr/topics/security/api-security> (Accessed: 17 July 2024).
- Robert, S. (2024) *Construire et Orchestrer des conteneurs*. Available at: <https://blog.stephane-robert.info/docs/conteneurs/introduction/> (Accessed: 17 July 2024).
- Robot, S. (2022) *BeautifulSoup vs. Scrapy: Which Is Better For Web Scraping?*, *Scraping Robot*. Available at: <https://scrapingrobot.com/blog/beautifulsoup-vs-scrapy/> (Accessed: 18 July 2024).
- Rouen.fr (2023) *Qu'est-ce que la renaturation ?* Available at: <https://rouen.fr/definition-renaturation> (Accessed: 14 June 2024).
- Savi (2020) *Fiche action renaturation.pdf*. Available at: https://www.savi37.fr/images/telechargements/Fiches_action/Fiche%20action%20renaturation.pdf (Accessed: 3 June 2024).
- Scafe (2023) 'Comprendre la conteneurisation informatique', *Scafe*, 25 November. Available at: <https://www.scafe.io/comprendre-la-conteneurisation-informatique/> (Accessed: 17 July 2024).
- Scrapster (2021) *Quels outils de Web Scraping utiliser ?* Available at: <https://www.scrapster.io//quels-outils> (Accessed: 17 July 2024).
- Sharma, N. (2018) *React vs Angular vs Ember vs Vue: Compare JavaScript Framework*, *Apptunix Blog*. Available at: <https://www.apptunix.com/blog/react-vs-angular-vs-ember-vs-vue-best-javascript-framework/> (Accessed: 22 July 2024).
- Site de la trame verte et bleue (2023) *Quels sont les acteurs de la préservation et de la remise en bon état des continuités écologiques ? | Trame verte et bleue*. Available at: <https://www.trameverteetbleue.fr/presentation-tvb/foire-aux-questions/quels-sont-acteurs-preservation-remise-bon-etat-continuites> (Accessed: 28 June 2024).
- Slingerland, C. (2023) *Kubernetes Vs. Docker Vs. OpenShift, CloudZero*. Available at: <https://www.cloudzero.com/blog/kubernetes-vs-docker/> (Accessed: 18 July 2024).
- Software AG (2023) *Qu'est-ce que la sécurité des API? | Solutions de sécurité API et bonnes pratiques*, *Software AG*. Available at: https://www.softwareag.com/fr_fr/resources/api/article/api-security.html (Accessed: 17 July 2024).
- sql.sh (2020) *Système de Gestion de Base de Données (SGBD), SQL*. Available at: <https://sql.sh/sghbd> (Accessed: 17 July 2024).
- Svitla (2023) *Python Web Frameworks Overview by Svitla Systems*. Available at: <https://svitla.com/blog/python-web-frameworks-django-flask-bottle-pyramid-tornado-web2py-falkon> (Accessed: 18 July 2024).
- Swapnil (2024) *Beautiful Soup vs Scrapy vs Selenium - Which one to use? Python Web sc.* Available at: <https://www.getodata.com/blog/beautiful-soup-vs-scrapy-vs-selenium-which-one-to-use-python-web-scraping> (Accessed: 18 July 2024).
- Tiangolo (2024) *Alternatives, inspiration et comparaisons - FastAPI*. Available at: <https://fastapi.tiangolo.com/fr/alternatives/> (Accessed: 18 July 2024).
- Tomaszuk, D. (2024) *Python's Frameworks Comparison: Django, Pyramid, Flask, Sanic, Tornado, BottlePy and More*. Available at: <https://www.netguru.com/blog/python-frameworks-comparison> (Accessed: 18 July 2024).

- Vasilis, T. (2024) *Web scraping: Beautiful Soup or Scrapy?*, Apify Blog. Available at: <https://blog.apify.com/beautiful-soup-vs-scrapy-web-scraping/> (Accessed: 18 July 2024).
- Veritas (2022) *Qu'est-ce que la conteneurisation ? Quels sont les avantages ?* Available at: <https://www.veritas.com/fr/fr/information-center/containerization> (Accessed: 17 July 2024).
- Vi, L. (2021) *7 best practices pour sécuriser vos API*, Astrakhan Innovation Management. Available at: <https://astrakhan.consulting/fr/blog/7-best-practices-pour-securiser-vos-api/> (Accessed: 17 July 2024).
- WWF (2020) *Communiqué de presse: La biodiversité s'est effondrée de 68 % depuis 1970, selon un nouveau rapport du WWF*. Available at: <https://wwf.be/fr/communiqués-de-presse/communiqué-de-presse-la-biodiversité-s'est-effondrée-de-68-depuis-1970-selon> (Accessed: 17 May 2024).
- Zyprian, F. (2023) *Analyser des PDF et d'autres formats de données en Python*, Konfuzio. Available at: <https://konfuzio.com/fr/pdf-parsing-python/> (Accessed: 18 July 2024).

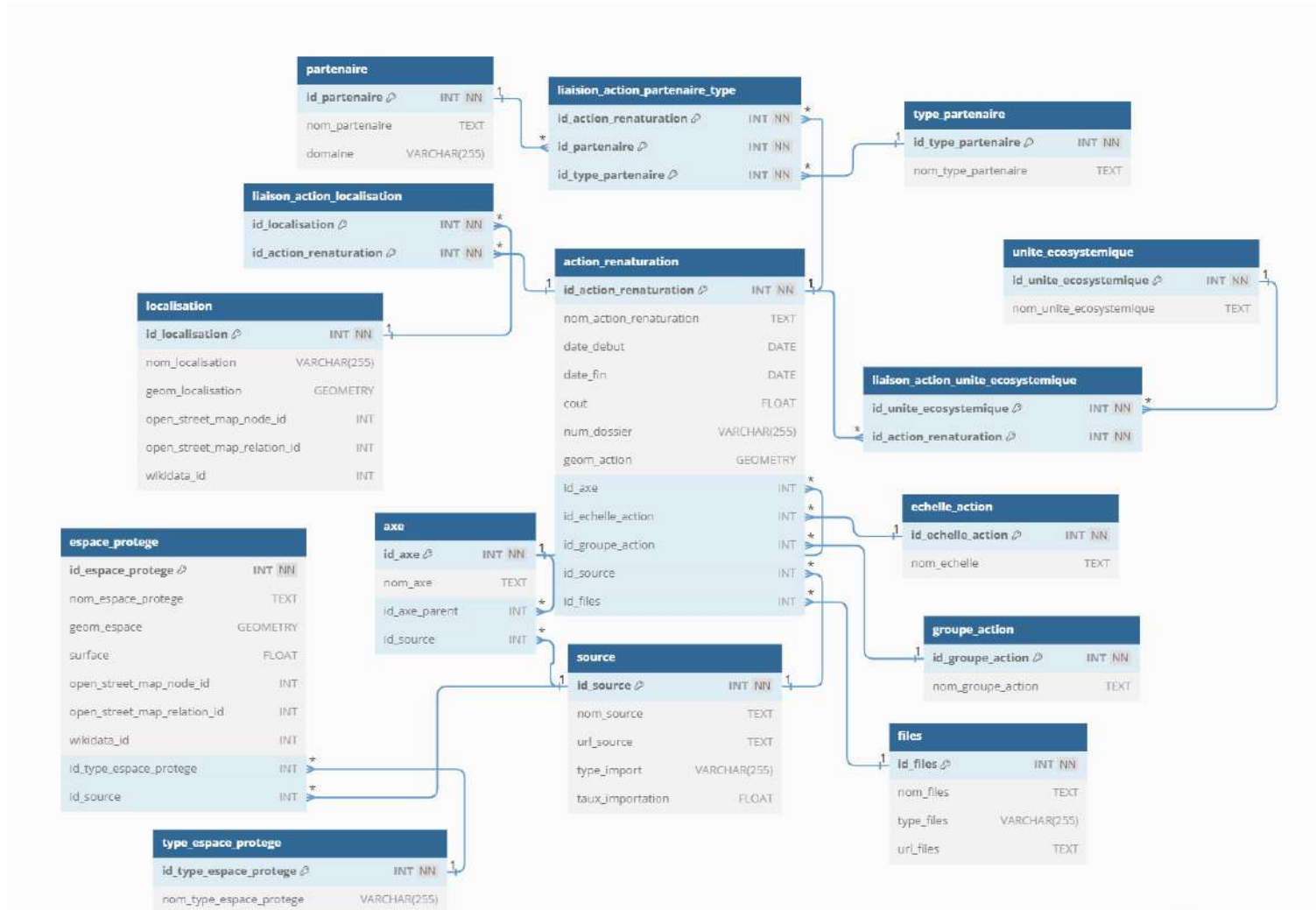
Source d'actions de renaturation :

- Agence Régionale pour la Biodiversité en Aquitaine (2017) *Recueil d'initiatives régionales en faveur de la biodiversité*. Available at: <https://www.biodiversite-nouvelle-aquitaine.fr/wp-content/uploads/2021/02/Recueil-initiatives-r%C3%A9gionales-biodiversit%C3%A9-2017-Version-Web.pdf> (Accessed: 24 April 2024).
- Calvar, E. et al. (2021) *Recueil d'expériences : Restauration fonctionnelle de tourbières dans le massif du Jura*. Available at: http://www.life-tourbieres-jura.fr/server/php/download4/Recueil-d-experiences_Life-tourbieres-du-Jura_FR_WEB.pdf (Accessed: 30 April 2024).
- Capitales Françaises de la Biodiversité (no date) *Recueil d'expériences | Capitales Françaises de la Biodiversité*. Available at: <https://www.capitale-biodiversite.fr/carte-experiences> (Accessed: 22 April 2024).
- Castaing, J., Monod, K. and Noreve, V. (2022) *Renaturer les sols, des solutions pour des territoires durables*. Available at: <https://www.ofb.gouv.fr/sites/default/files/Fichiers/Plaquettes%20et%20rapports%20institut/renaturer-les-sols.pdf> (Accessed: 22 April 2024).
- Cerema (2024) *Téléchargement de données*. Available at: <https://catalogue.cdata.cerema.fr/geosource/panierDownloadFrontalParametrage?LAYERIDTS=8497148> (Accessed: 19 April 2024).
- Conservatoire d'espaces naturels Centre (2015) *Etude Trame Verte et Bleue du Pays Beauce Gâtinais en Pithiverais*. Available at: https://www.pays-du-pithiverais.fr/fic_bdd/pdf_fr_fichier/tvb-pbpg-fiches-actions_1424950445.pdf (Accessed: 23 April 2024).
- DREAL Centre-Val de Loire (2023) *Programme de priorisation des actions de restauration de la continuité écologique du bassin Loire-Bretagne*, DREAL Centre-Val de Loire. Available at: <https://www.centre-val-de-loire.developpement-durable.gouv.fr/programme-de-priorisation-des-actions-de-a3735.html> (Accessed: 23 April 2024).
- France nation verte (2024) *Plans nationaux d'actions en faveur des espèces menacées*. Available at: <https://biodiversite.gouv.fr/projet-pna/> (Accessed: 22 April 2024).
- GRAIE, AFB and FCEN (2018) *Fleuves et grandes rivières du bassin Rhône méditerranée et de la région Auvergne-Rhône-Alpes : Recueil de retours d'expériences exemplaires*. Available at:

- <https://www.graie.org/zabr/RecueilREX/Recueil-REXFleuvesRiviereRMC.pdf> (Accessed: 30 April 2024).
- Grandin, G. and Lecuir, G. (2019) *Capitale française de la biodiversité ; climat : la nature source de solutions*. Available at: <https://www.capitale-biodiversite.fr/sites/default/files/experience/documents/cfb2020-recueil-actions.pdf>.
 - NatureFrance (2022) *Catalogue de données sur la biodiversité*. Available at: <https://data.naturefrance.fr/geonetwork/srv/fr/catalog.search#/metadata/cc1bfe07-625e-4df6-86a2-49fd07d193d7> (Accessed: 24 April 2024).
 - OFB (2021) *Le recueil d'expériences sur l'hydromorphologie, Le portail technique de l'OFB*. Available at: <https://professionnels.ofb.fr/fr/node/217> (Accessed: 2 May 2024).
 - OFB (2023) *Retours d'expériences cours d'eau et zones humides | Zones Humides*. Available at: <https://www.zones-humides.org/agir/retours-d-experiences-cours-d-eau-et-zones-humides> (Accessed: 2 May 2024).
 - SDAGE (2023) *Données techniques de référence du SDAGE 2022-2027 | L'eau dans le bassin Rhône-Méditerranée*. Available at: <https://www.rhone-mediterranee.eaufrance.fr/gestion-de-leau/sdage-2022-2027-elaboration/donnees-techniques-de-reference-du-sdage-2022-2027> (Accessed: 26 April 2024).
 - Stratégie nationale biodiversité (2023) *Classeur des fiches mesures*. Available at: <https://www.ecologie.gouv.fr/sites/default/files/Cahier-des-fiches-mesures-SNB2030.pdf> (Accessed: 22 April 2024).
 - Symvahem (2019) *Recueil d'expérience : La reconquête de la continuité écologique de la Hem*. Available at: <https://symvahem.fr/wp-content/uploads/2019/12/recueil-d-experience-continuite-ecologique-hem-bd-double-page.pdf> (Accessed: 24 April 2024).
 - Syndicat Mixte d'Etudes et d'Aménagement de la Garonne (2014) *Nouvelle approche du territoire de la Garonne débordante entre Toulouse et Saint-Nicolas-de-la-Grave : Fiche action*. Available at: https://www.lagaronne.com/sites/default/files/upload/1_plan_actions_tfe_sept2014.pdf (Accessed: 22 April 2024).
 - Syndicat Rivières des 4 Vallées (2015) *Contrat de rivière des 4 Vallées ; Recueil de fiches actions*. Available at: https://www.gesteau.fr/sites/default/files/gesteau/content_files/document/15-2-pdf-20161102221535_compressed.pdf (Accessed: 25 April 2024).
 - Thibault, O. (2010) 'Recueil d'opérations localisation des sites'.
 - Trame verte et bleue (no date) *Carte dynamique des retours d'expériences TVB | Trame verte et bleue*. Available at: <https://www.trameverteetbleue.fr/retours-experiences> (Accessed: 30 April 2024).
 - Zones Humides (2024) *Les sites Ramsar en France | Zones Humides*. Available at: <https://www.zones-humides.org/les-sites-ramsar-en-france> (Accessed: 23 April 2024).

Liste des annexes

Annexe i : Modèle de données



Annexe ii : Dictionnaire de données

Table	Colonne	Type de donnée	Commentaire
action_renaturation	id_action_renaturation	int	
	nom_action_renaturation	text	
	date_debut	date	
	date_fin	date	
	cout	float	Budget nécessaire pour la réalisation de l'action
	num_dossier	varchar(255)	
	geom_action	geom	
axe	id_axe	int	
	nom_axe	text	Axe dans lequel s'inscrit l'action
	id_axe_parent	int	
echelle_action	id_echelle_action	int	
	nom_echelle	text	Départementale, régionale, européenne, ...
espace_protege	id_espace_protege	int	
	nom_espace_protege	text	
	geom_espace	geom	
	surface	float	
	open_street_map_node_id	int	id d'Open Street Map pour favoriser la recherche d'information

	open_street_map_relation_id	int	id d'Open Street Map pour favoriser la recherche d'information
	wikidata_id	varchar(50)	id de Wikidata pour favoriser la recherche d'information
files	id_files	int	
	nom_files	text	Nom du fichier pouvant être rattaché à l'action
	type_files	varchar(255)	
	url_files	text	
groupe_action	id_groupe_action	int	
	nom_groupe_action	text	
localisation	id_localisation	int	
	nom_localisation	varchar(255)	
	geom_localisation	geom	
	open_street_map_node_id	int	id d'Open Street Map pour favoriser la recherche d'information
	open_street_map_relation_id	int	id d'Open Street Map pour favoriser la recherche d'information
	wikidata_id	int	id de Wikidata pour favoriser la recherche d'information
partenaire	id_partenaire	int	
	nom_partenaire	text	
	domaine	varchar(255)	Public, privé, ...
source	id_source	int	

	nom_source	text	
	url_source	text	
	type_import	varchar(255)	Manuel ou automatisé
	taux_importation	float	
type_espace_protege	id_type_espace_protege	int	
	nom_type_espace_protege	varchar(255)	Parc national, aire marine protégé, ...
type_partenaire	id_type_partenaire	int	
	nom_type_partenaire	text	Partenaire financier, technique, maître d'ouvrage, ...
unite_ecosystemique	id_unite_ecosystemique	int	
	nom_unite_ecosystemique	text	Trame bleue, trame verte, forêt, ...