

UNIVERSITE DE YAOUNDE I

\*\*\*\*\*

ECOLE NATIONALE SUPERIEURE  
POLYTECHNIQUE

\*\*\*\*\*

DEPARTEMENT DE GENIE  
INFORMATIQUE

UNIVERSITY OF YAOUNDE I

\*\*\*\*\*

NATIONAL ADVANCED SCHOOL  
OF ENGINEERING

\*\*\*\*\*

DEPARTMENT OF COMPUTER  
ENGINEERING



prooftag™  
CATIS S.A

---

---

## CONCEPTION D'UN PORTAIL OPEN DATA DES ACCIDENTS DE LA CIRCULATION

---

---

### Rapport de stage PRÉ-INGÉNIEUR

Présenté par :

**KAMGA YOUTH ROEL MATHURIN**

Sous la supervision de :

**ING. SERGE TAMPOLLA**

Année Académique 2023-2024



.5



.5

---

## CONCEPTION D'UN PORTAIL OPEN DATA DES ACCIDENTS DE LA CIRCULATION

---

### **Rapport de stage PRÉ-INGÉNIEUR**

Présenté par :

**KAMGA YOUTH ROEL MATHURIN**

Sous la supervision de :

**ING. SERGE TAMPOLLA**

**Année Académique 2023-2024**

## REMERCIEMENTS

**E**n premier lieu, je tiens à remercier M. SERGE TAMPOLLA, En tant que maître de stage, il m'a beaucoup appris et a partagé ses connaissances dans le domaine d'informatique. Je saisi saisir cette occasion pour adresser mes profonds remerciements aux responsables et aux personnels de l'École Nationale Supérieure Polytechnique de Yaoundé. Je désire aussi remercier les professeurs de l'ENSPY, qui m'ont fourni les outils nécessaires au bon déroulement de mon stage. Je tiens à remercier spécialement Pr KOUAMOU, qui fut le premier à me soutenir dans ma démarche de stage. Un grand merci à ma Famille , mes amis et le Personnel de Prooftag-Catis pour leurs conseils ainsi que leur soutien.

Je remercie également mes parents, mes frères et sœurs pour toute leur aide, leur soutien et leur amour.

Merci à tous mes camarades et à tous ceux que je n'ai pas cité.

## ABRÉVIATIONS

<b>Acronymes</b>	<b>Significations</b>
<b>CKAN</b>	Comprehensive Knowledge Archive Network
<b>DMS</b>	Data management system
<b>CMS</b>	Content management system
<b>API</b>	Application programming interface

## ABSTRACT

The project aims to create an Open Data portal dedicated to accidents, providing free and transparent access to crucial information. The main goal is to centralize data on accidents, making it easily accessible for researchers, public authorities, and the general public. The portal will provide detailed information on the location, severity, causes, and contributing factors of accidents, allowing for in-depth analysis and the implementation of preventive measures.

By fostering collaboration among various stakeholders, the portal seeks to enhance understanding of accident trends, facilitating the development of public policies and targeted solutions. The use of Open Data standards will ensure data interoperability, promoting reuse and combination with other datasets for a more holistic perspective.

This project aligns with a commitment to transparency and accountability, enabling everyone to contribute to road safety. By providing accurate and up-to-date data, the Open Data portal on accidents aims to play a key role in accident prevention and reduction.

## RÉSUMÉ

**L**e projet vise à créer un portail Open Data dédié aux accidents, offrant un accès libre et transparent à des données cruciales. L'objectif principal est de centraliser les informations sur les accidents, les rendant facilement accessibles pour les chercheurs, les autorités publiques et le grand public. Le portail fournira des données détaillées sur la localisation, la gravité, les causes et les facteurs contributifs des accidents, permettant une analyse approfondie et la mise en place de mesures préventives.

En favorisant la collaboration entre les différentes parties prenantes, le portail cherche à améliorer la compréhension des tendances d'accidents, facilitant ainsi le développement de politiques publiques et de solutions ciblées. L'utilisation de normes Open Data garantira l'interopérabilité des données, favorisant ainsi la réutilisation et la combinaison avec d'autres ensembles de données pour une perspective plus holistique.

Ce projet s'inscrit dans une démarche de transparence et de responsabilité, permettant à chacun de contribuer à la sécurité routière. En mettant à disposition des données précises et actualisées, le portail Open Data sur les accidents aspire à jouer un rôle clé dans la prévention et la réduction des accidents de la route.

## LIST OF TABLES

TABLE	Page
-------	------

## LIST OF FIGURES

FIGURE	Page
3.1 CKAN . . . . .	23
3.2 Diagramme . . . . .	25
3.3 Diagramme . . . . .	26
3.4 Diagramme . . . . .	26
3.5 Postgresql . . . . .	29
3.6 ckan installation . . . . .	31
3.7 . . . . .	32
3.8 . . . . .	32
3.9 . . . . .	33
3.10 . . . . .	33
3.11 . . . . .	34
3.12 . . . . .	34
3.13 . . . . .	35
3.14 . . . . .	35
3.15 . . . . .	36
3.16 . . . . .	36
3.17 . . . . .	37
3.18 . . . . .	37
3.19 . . . . .	37
3.20 . . . . .	38
3.21 . . . . .	38
3.22 . . . . .	38



## SOMMAIRE

<b>Dedication</b>	<b>i</b>
<b>Acknowledgements</b>	<b>ii</b>
<b>Abstract</b>	<b>vi</b>
<b>Résumé</b>	<b>vii</b>
<b>List of Tables</b>	<b>viii</b>
<b>List of Figures</b>	<b>ix</b>
<b>Table of Contents</b>	<b>xi</b>
<b>Abréviations</b>	<b>ii</b>
<b>Introduction</b>	<b>1</b>
<b>1 Présentation de l'entreprise</b>	<b>5</b>
1.1 PRÉSENTATION . . . . .	6
1.1.1 C'est Quoi Prooftag-Catis ? . . . . .	6
1.1.2 Quels Solutions apporte-elle? . . . . .	6
1.2 LES ACTIVITÉS MENÉS À PROOFTAG-CATIS . . . . .	6
1.2.1 LA CELLULE RECHERCHE . . . . .	6
1.2.2 LE CENTRE DE CONTRÔLE TECHNIQUE . . . . .	6
1.3 PRÉSENTATION DE CERTIDOCs CT . . . . .	8
1.4 QUELQUES PARTENAIRES DE PROOFTAG AU CAMEROUN . . . . .	8
1.4.1 Le Ministère des Transports . . . . .	8
1.4.2 Les différents centres de contrôle technique du Cameroun . . . . .	9
<b>2 État de l'art</b>	<b>11</b>
2.1 Introduction . . . . .	12

## TABLE OF CONTENTS



2.1.1	Contexte . . . . .	12
2.1.2	Justification du Projet . . . . .	12
2.2	OBJECTIFS DU PROJET . . . . .	12
2.2.1	Objectifs Généraux . . . . .	12
2.2.2	Objectifs Spécifiques . . . . .	13
2.3	Portée du Projet . . . . .	14
2.3.1	Description Du Périmètre du Projet . . . . .	14
2.3.2	Les Parties Prenantes. . . . .	14
2.4	MÉTHODOLOGIE ET APPROCHE . . . . .	14
2.4.1	Méthodologie Proposée . . . . .	14
2.4.2	Outils à Utilisés . . . . .	15
2.4.3	Calendrier des activités . . . . .	16
2.5	LIVRABLES ATTENDUS . . . . .	16
2.6	RESSOURCES NÉCESSAIRES . . . . .	17
2.6.1	Ressources Humaines . . . . .	17
2.6.2	Ressources Matérielles . . . . .	17
2.6.3	Ressources financière . . . . .	18
<b>3</b>	<b>Méthodologie, implémentation et résultats</b>	<b>20</b>
3.1	Choix de l'Outil et technologies utilisés . . . . .	21
3.2	OUTILS ET TECHNOLOGIES UTILISÉS . . . . .	22
3.2.1	CKAN . . . . .	22
3.2.2	Pourquoi choisir CKAN ? . . . . .	23
3.2.3	Types de données qui peuvent être importer dans une plateforme CKAN .	24
3.3	Analyse . . . . .	24
3.3.1	description détaillée . . . . .	25
3.4	Installation de CKAN . . . . .	27
3.4.1	Installation de CKAN 2.9 Sur Ubuntu 20 .04 . . . . .	28
3.4.2	Configuration . . . . .	31
3.4.3	Résultats Obtenus . . . . .	32
3.4.4	Perspectives . . . . .	38
	<b>Conclusion</b>	<b>40</b>
	<b>Bibliographie</b>	<b>41</b>



## INTRODUCTION

### CONTEXTE

**A** l'ère numérique en constante évolution, l'accès libre aux données publiques est devenu un pilier fondamental de la transparence, de l'efficacité et de la collaboration dans tous les domaines. Ces données offrent de nombreuses opportunités pour étendre le savoir humain et créer de nouveaux produits et services de qualité. Fort de cette conviction, ProofTag-Catis souhaite s'engager activement dans l'ouverture de ses données pour servir le bien public, stimuler l'innovation et renforcer les capacités de prise de décision basées sur des éléments tangibles.



## HISTORIQUE

L'entreprise prooftag-catis voir le jour au Cameroun en 2017 par le biais d'un accord entre l'entreprise française Prooftag SAS et le ministère des transport du Cameroun. L'application Cer-tiDocs CT produit de l'entreprise Prooftag est déployé dans tout les centres de visites techniques d'automobile jusqu'au 16 mars 2020 .

## OBJECTIFS

Notre projet vise a mettre sur pied **un système d'information des accidents ou une base de données sur les accidents** utilisant Open DATA CKAN. ce système est capable de :

1. Permettre la collecte de données sur les accidents de la route, y compris des informations telles que la date, l'heure, le lieu, les types de véhicules impliqués, les conditions météorologiques, les blessures et les décès.
2. Fournir des outils d'analyse pour examiner les tendances, les facteurs de risque et les schémas des accidents de la route, afin d'identifier les problèmes et de prendre des mesures préventives appropriées.
3. Offrir des fonctionnalités de visualisation des données, telles que des graphiques et des cartes interactives, pour faciliter la compréhension et la communication des informations sur les accidents de la route.
4. Permettre la création de profils d'utilisateurs avec des autorisations spécifiques pour accéder, modifier ou télécharger des données sur les accidents de la route, en fonction de leur rôle et de leurs responsabilités.
5. Faciliter le partage des données sur les accidents de la route avec les parties prenantes concernées, telles que les décideurs politiques, les chercheurs, les professionnels de la sécurité routière et le grand public.
6. Générer des rapports automatiques sur les statistiques et les indicateurs clés liés aux accidents de la route, pour faciliter la prise de décision basée sur des données probantes.
7. Assurer l'interopérabilité des données sur les accidents de la route avec d'autres systèmes d'information pertinents, tels que les services d'urgence médicale ou les bases de données de véhicules, pour une meilleure coordination et une meilleure compréhension globale de la sécurité routière.
8. D'être disponible dans plusieurs langues pour atteindre un public mondial.





9. Permettre aux utilisateurs de laisser des commentaires ou des réactions sur les données ou les analyses.
10. Permettre aux développeurs d'accéder aux données à l'aide d'API pour créer des applications ou des services personnalisés.
11. Les données peuvent inclure des informations démographiques sur les personnes impliquées dans les accidents.
12. Les utilisateurs peuvent filtrer les données en fonction de critères tels que la localisation, la date, le type d'accident, etc.



## FEUILLE DE ROUTE

La suite de ce document est organisée comme suit :



## PRÉSENTATION DE L'ENTREPRISE

**D**ans ce chapitre, nous allons faire une présentation de l'entreprise dans laquelle nous avons effectué notre stage pré-ingénieur : Prooftag-Catis



## 1.1 PRÉSENTATION

### 1.1.1 C'est Quoi Prooftag-Catis ?

Prooftag-Catis est le nouvel opérateur chargé de la sécurisation des vignettes des visites techniques automobile au Cameroun. Prooftag-Catis est le fournisseur exclusif du logiciel d'enregistrement et d'un ensemble d'équipements nécessaires à la mise en place de cette nouvelle solution. Le processus de sécurisation et d'authentification est basé sur Plusieurs points. Le siège social est situé à Douala au quartier Bonapriso à l'immeuble le carat.

### 1.1.2 Quels Solutions apporte-elle?

Cet Opérateur est chargé de la sécurisation des vignettes des visites techniques automobile sur l'étendue du territoire. Les solutions apportés par cet Organisme sont multiple ,on peut noter :

- la réduction du taux d'accidents de circulation due aux mauvais états des véhicules. Puisque le logiciel CertiDocs suit le processus de la visite technique de l'automobile de l'enregistrement du véhicule jusqu'à la délivrance du PV.
- Vérification de l'effectivité de la dite Visite.
- Gain d'argent dans les Caisses de l'État .
- Permet au ministère des Transport de Superviser tout les activités dans chaque centre ,en interconnectant tout les centres de visites techniques au ministère des Transport.
- Retrouver les Véhicules Volés.
- Identifier une Voiture Ainsi que son propriétaire.
- Vérifie l'authenticité de la vignette automobile.

## 1.2 LES ACTIVITÉS MENÉS À PROOFTAG-CATIS

L'organisme prooftag-catis a deux sites d'activités dans la ville de douala.

### 1.2.1 LA CELLULE RECHERCHE

Situé à la zone portuaire de bonaberie , Dans cette cellule on retrouve l'équipe des développeurs, qui ont pour principale rôle de développé les applications proposé par l'organisme. Cette cellule travaille sur une application de collecte de données que nous présenterons un peu plus tard.

### 1.2.2 LE CENTRE DE CONTRÔLE TECHNIQUE

Situé à Bonapriso , dans ce centre au retrouve trois services.







#### **1.2.2.1 LE SERVICE DE SUPPORT TECHNIQUE**

Ce service a pour but principal d'assurer le bon fonctionnement des serveurs installés dans chaque centre. Les employés de ce service utilisent un serveur central (Zabbix) qui est connecté à tous les serveurs de tous les centres de visites techniques grâce à un logiciel appelé AnyDesk.

Ce serveur central affiche les activités et les états de tous les serveurs qui lui sont connectés. L'équipe de ce service résout les problèmes (liés à un serveur d'un centre) affichés par ce serveur central.

En parallèle, les centres peuvent être appelés pour exposer un problème lié à l'utilisation de l'application ou un bug du serveur par exemple ce qu'on appelle la supervision active. Dans ce cas, l'équipe en place va trouver une solution au problème posé ; Le serveur central Zabbix détecte également automatiquement les problèmes sur les autres serveurs qui lui sont liés et l'envoie à un logiciel qui lui est associé GLPI qui à son tour va notifier un membre du service pour lui faire part du problème rencontré c'est la supervision proactive.

Il est à noter qu'avant de résoudre un problème, les employés de ce service utilisent le logiciel GLPI, pour créer ce problème avant de le résoudre. Ceci dans le but de sauvegarder la solution au problème résolu, et servant aussi de preuve de travail. On dénombre 5 Employés dans cette cellule.

#### **1.2.2.2 LA CELLULE DU SURVIE**

Les rôles majeurs de cette cellule c'est d'éviter les fraudes dans tous les centres de visites techniques, vérifier l'authenticité des PV fournis et faire des études statistiques permettant de prendre de meilleures décisions plus tard. Les employés de ce service se rassurent qu'effectivement le rapport fourni par le banc de test (rapport machine) est corrélé à 100

#### **1.2.2.3 LA CELLULE COMPTABLE**

C'est dans cette cellule qu'est géré tout le mouvement financier de l'organisme. Comme toutes les autres entreprises, Prooftag a besoin de services comptables pour éclaircir les différentes activités financières au sein de l'organisation. On dénombre 1 Employé dans cette cellule.

#### **1.2.2.4 LA CELLULE RH**

Cette cellule a pour mission de guider les employés dans leur parcours dans l'entreprise, accompagner dans leur projet professionnel, et gérer les congés et les absences. On dénombre 2 Employés dans cette cellule.





## 1.3 PRÉSENTATION DE CERTIDOCs CT

CertiDocs est l'application de délivrance de la vignette de visite technique ; Elle constitué par une partie web et une partie Mobile. Les étapes du processus sont les suivants: • Enregistrement du véhicule, la voiture passe à l'état 'à traiter'. • Mise à jour des données du certificat d'immatriculation , la voiture passe à l'état 'Inspecter'. • Le Véhicule est encours d'inspection par le contrôleur du centre technique, qui relève les informations suivantes : le kilométrage ,le nombre d'essieux et numéro de chassie du véhicule. Puis valide • Le Véhicule passe sous le banc de test (ripage,suspension et freinage), qui au terme renvoi les données d'analyse sur le serveur mysql du centre et sur dans la base de donnée du centre du support qui par la suite aiderons pour la vérification des données qui seront inscrit sur le PV. • Le contrôleur effectue un contrôle visuel et repleie d'autres données sur la tablette, prend les photos du véhicule et signe. • Et Impression du PV.

## 1.4 QUELQUES PARTENAIRES DE PROOFTAG AU CAMEROUN

### 1.4.1 Le Ministère des Transports

Le siège social et le bureau commercial de Megasoft se situent à Yaoundé au lieu-dit Rue Biyem-Assi Lac (L'entrée juste après le lac).





Suite à de nombreuses irrégularités observées dans le processus de délivrance de la vignette de visite technique automobile, le ministère des transports a opté pour une nouvelle solution numérique. Dans le but d'atteindre les objectifs de la décennie d'actions pour la sécurité routière 2011-2020 et réduire l'impact négatif des accidents sur nos routes dû au mauvais état technique des véhicules. Cette solution est l'application Certidocs CT proposée par PROOFTAG CATIS qui permettra au ministère des Transports, à la Gendarmerie et à la Sûreté nationale d'avoir des informations à temps réel sur la santé technique réelle du véhicule et de sécuriser et authentifier les vignettes. Cette sécurisation de ladite vignette permettra d'éradiquer le phénomène des fausses vignettes et à l'État de rattraper le gap de 1 milliard de francs CFA que cela cause dans ses caisses chaque année.

#### **1.4.2 Les différents centres de contrôle technique du Cameroun**

Les centres de contrôle technique sont des installations spécialisées chargées de vérifier la conformité des véhicules automobiles aux normes de sécurité et d'émissions en vigueur. Ils effectuent des inspections techniques périodiques visant à s'assurer que les véhicules sont en bon état de fonctionnement, conformes aux réglementations en matière de sécurité routière, et respectent les normes environnementales. Ces centres réalisent des contrôles approfondis sur différents aspects des véhicules, notamment les freins, les phares, les émissions de gaz d'échappement,



la direction, les pneus, et d'autres composants essentiels. Leur mission est de garantir que les véhicules circulant sur les routes sont sûrs pour les conducteurs, les passagers et les autres usagers de la route, tout en contribuant à la réduction de l'impact environnemental des véhicules à moteur. Le Cameroun compte à ce jour plus d'une cinquantaine centres de contrôle qui sont repartis sur l'ensemble du territoire.

## ÉTAT DE L'ART

**D**ans ce chapitre nous allons parler de ce qui en ai d'une plateforme open data . Nous allons tout d'abord présenté le projet, ensuite les Objectifs les projets, La portée et la Methodologie Proposée



## 2.1 Introduction

### 2.1.1 Contexte

A l'ère numérique en constante évolution, l'accès libre aux données publiques est devenu un pilier fondamental de la transparence, de l'efficacité et de la collaboration dans tous les domaines. Ces données offrent de nombreuses opportunités pour étendre le savoir humain et créer de nouveaux produits et services de qualité. Fort de cette conviction, ProofTag-Catis souhaite s'engager activement dans l'ouverture de ses données pour servir le bien public, stimuler l'innovation et renforcer les capacités de prise de décision basées sur des éléments tangibles.

### 2.1.2 Justification du Projet

Dans notre projet de création d'un portail Open Data, il serait essentiel d'identifier un certain nombre de besoins spécifiques pour s'assurer que le portail réponde aux attentes et aux objectifs fixés. Voici certains des besoins clés à identifier : • Besoin d'accès aux données publiques sur les Accidents: Le Portail doit Permettre Aux Consommateurs d'accéder facilement aux données.

- Besoin de transparence et de responsabilité : Les données publiques doivent être facilement accessibles et compréhensibles par tous.
- Besoin de standardisation des données : Pour faciliter la réutilisation des données.
- Besoin de mise à jour régulière : Il est important de prévoir la possibilité de mettre à jour régulièrement les données publiées sur le portail pour garantir leur pertinence et leur actualité.
- Besoin de mise à jour régulière : Il est important de prévoir la possibilité de mettre à jour régulièrement les données publiées sur le portail pour garantir leur pertinence et leur actualité .
- Besoin de formation et de sensibilisation : Identifier le besoin de fournir des formations aux parties prenantes sur l'utilisation du portail et de mener des actions de sensibilisation pour encourager l'adoption de l'Open Data

## 2.2 OBJECTIFS DU PROJET

### 2.2.1 Objectifs Généraux

Le principal objectif de ce projet est de concevoir, développer et déployer un Portail Open Data robuste et convivial . Parmi les objectifs spécifiques du projet, on peut citer : • Fournir un accès facile, rapide et sécurisé aux données publiques du Portail dans un format ouvert, réutilisable et exploitable.

- Faciliter la recherche, la navigation et le téléchargement des jeux de données publiques par les parties prenantes, les chercheurs, les journalistes et le grand public.
- Promouvoir la transparence et la responsabilité en permettant aux citoyens d'accéder aisément aux informations publiques pertinentes.





- Favoriser l'innovation en ouvrant de nouvelles opportunités pour les développeurs, les startups et les entreprises qui souhaitent utiliser les données publiques pour créer des applications et des solutions novatrices.
- Améliorer l'efficacité de l'organisation en facilitant le partage de données à l'interne et en réduisant les coûts liés à la diffusion d'informations publiques.

### 2.2.2 Objectifs Spécifiques

Pour atteindre notre objectif principal, le portail Open Data s'attachera à réaliser les objectifs spécifiques suivants :

- **Faciliter l'accès aux données publiques** : Le portail permettra aux citoyens, chercheurs, médias, entreprises et autres parties prenantes d'accéder facilement aux données sur le portail. Il offrira une interface conviviale et intuitive pour une recherche aisée et une navigation fluide dans les différentes catégories de données.

- **Améliorer la prise de décision basée sur les données** : Le portail permettra aux décideurs, aux chercheurs et aux analystes d'accéder facilement aux données pertinentes et actualisées. Cela facilitera une prise de décision informée et basée sur des éléments tangibles, contribuant ainsi à l'amélioration de l'efficacité et de la pertinence des politiques et des actions au sein de l'Organisation.

- **Favoriser la collaboration et l'engagement citoyen** : En donnant accès aux données publiques, le portail encouragera la participation citoyenne et la collaboration avec les parties prenantes externes. Les données ouvertes pourront servir de base pour des projets de recherche, des initiatives de société civile et des débats publics éclairés.

- **Assurer la mise à jour régulière des données** : Le portail sera conçu pour permettre la mise à jour régulière des données publiques afin de garantir leur pertinence et leur actualité pour les utilisateurs.

- **Promouvoir la transparence et la responsabilité** : En mettant à disposition les données publiques de manière ouverte et accessible, le portail favorisera la transparence des activités et des décisions de l'Organisation. Il renforcera également la responsabilité en permettant aux citoyens de mieux comprendre les actions de l'organisation.

- **Encourager l'innovation et la réutilisation des données** : Le portail Open Data servira de catalyseur pour stimuler l'innovation au sein de la société civile, des entreprises et de la communauté académique. En ouvrant de nouvelles opportunités de réutilisation des données, il encouragera le développement d'applications, d'outils et de solutions novatrices basées sur les données publiques

- **Renforcer la communication et la sensibilisation** : Le portail jouera un rôle actif dans la communication et la sensibilisation concernant les activités de l'Organisation. Il permettra de mieux informer le public sur les sujets clés, les réalisations, les indicateurs de performance ,etc.





## 2.3 Portée du Projet

### 2.3.1 Description Du Périmètre du Projet

- Le Portail Open Data conçue doit être accessible dans tout l'étendu du territoire camerounais dans le secteur des Transports et de l'Infrastructure .
- Les données publiées sur le portail seront les Données Actuelles , Données Historiques et Données Spécifiques (données publiques spécifiques à un événement ou à une période donnée).

### 2.3.2 Les Parties Prenantes.

Parties prenantes	
<b>Utilisateurs du Portail</b>	Les utilisateurs du portail sont les principales parties prenantes bénéficiaires.
<b>Responsable Technique du Portail</b>	Cette entité est en charge de la gestion et de la maintenance technique.
<b>Société Civile et Groupes d'Intérêt</b>	Ces parties prenantes peuvent utiliser les données publiques pour mener des activités.
<b>Développeurs et Entreprises</b>	Les développeurs et les entreprises peuvent utiliser les données publiques pour développer des applications.
<b>Public et Médias</b>	Le grand public et les médias jouent un rôle important dans l'utilisation des données.
<b>Gestionnaires de Projet</b>	Les gestionnaires de projet sont responsables de la planification, de l'exécution et du suivi du projet.
<b>Organismes Publics Décentralisés</b>	chaque organisme public décentralisé est responsable de la publication des données.
<b>ProofTag-Catis</b>	Il est le commanditaire du projet et définit les objectifs globaux du portail.

## 2.4 MÉTHODOLOGIE ET APPROCHE

### 2.4.1 Méthodologie Proposée

#### 2.4.1.1 Approches et Méthodes

Pour atteindre les objectifs de notre Projet de création d'un portail Open Data, plusieurs approches et Méthodes peuvent être mises en œuvre. Voici quelques-unes des approches clés qui peuvent être utilisées pour atteindre les objectifs spécifiques du projet :

- **Collecte Systématique des Données** : Mettre en place un processus systématique de collecte des données publiques auprès des organismes gouvernementaux et d'autres sources pertinentes. Assurer que les données sont actualisées régulièrement pour garantir leur pertinence.
- **Normalisation et Validation des Données** : Appliquer des normes de normalisation et de validation des données pour assurer leur cohérence, leur qualité et leur fiabilité avant leur publication sur le portail Open Data.
- **Transparence et Engagements Des Organismes Publics** : Établir des relations de collaboration avec les organismes publics responsables de la publication des données. Sensibiliser ces organismes à l'importance de l'Open Data, à la transparence et à la réutilisation des données.







publiques.

- Consultations Publiques : Organiser des consultations publiques pour identifier les priorités en matière de données, recueillir les besoins des utilisateurs potentiels et s'assurer que le portail répond aux attentes du public.
- Formation et Sensibilisation : Fournir des formations et des séances de sensibilisation pour les parties prenantes concernées, y compris les organismes publics, les développeurs, la société civile et les utilisateurs du portail Open Data.
- Méthode de Collecte des Données : Utiliser des méthodes de collecte de données efficaces, telles que les formulaires en ligne, les API de données, les extractions de bases de données, etc., pour obtenir les données publiques auprès des organismes gouvernementaux et d'autres sources.
- Méthode de Formation et de Sensibilisation : Fournir des sessions de formation et de sensibilisation pour les employés des organismes gouvernementaux, les développeurs et les parties prenantes sur l'importance de l'Open Data et les avantages de la réutilisation des données publiques.
- Méthode de Création de Visualisations de Données : Utiliser des outils de visualisation de données pour créer des graphiques, des cartes et des tableaux interactifs qui permettent aux utilisateurs de mieux comprendre les informations fournies par les données publiques.
- Méthode de Suivi et d'Évaluation : Mettre en place un système de suivi et d'évaluation pour mesurer les progrès du projet, évaluer l'utilisation du portail par les utilisateurs, et identifier les domaines d'amélioration.
- Méthode de Gestion de Projet : Utiliser des méthodes de gestion de projet efficaces pour planifier, suivre et coordonner toutes les activités liées à la création et à la gestion du portail Open Data.

#### 2.4.2 Outils à Utilisés

- CKAN : CKAN est une plateforme open-source spécialement conçue pour gérer des portails Open Data. Il offre des fonctionnalités avancées pour la publication, la recherche, et la réutilisation des données.
- Liferay : Liferay est une plateforme open-source de gestion de contenu d'entreprise (CMS) qui offre une gamme complète de fonctionnalités pour créer des sites web, des portails d'entreprise, des applications web et des expériences numériques.
- DKAN et JKAN qui sont des adaptations de CKAN pour d'autres CMS
- Socrata : Socrata est une autre plateforme populaire qui permet aux organisations de publier et de partager des données ouvertes. Elle offre des fonctionnalités avancées pour gérer et visualiser les ensembles de données.
- WordPress : WordPress peut être utilisé pour créer un portail Open Data personnalisé. On peut utiliser des plugins spécifiques tels que DKAN pour ajouter des fonctionnalités Open Data au portail





### 2.4.3 Calendrier des activités

Date	Activités
<b>Du 10/07/2023 au 17/07/2023</b>	Analyse du Problème
<b>Du 17/07/2023 au 24/07/2023</b>	Recherche Solutions
<b>Du 24/07/2023 au 31/07/2023</b>	Prototypage
<b>Du 31/07/2023 au 29/08/2023</b>	Implémentation
<b>Du 29/08/2023 au 11/09/2023</b>	Test et déploiement

## 2.5 LIVRABLES ATTENDUS

Le contenu des livrables peut varier en fonction des spécificités du projet, des objectifs, des besoins et des attentes des parties prenantes. Les Modèles de Livrables Pour le Projet de Portail Open Data sont les suivant :

- Document de Spécifications des Besoins : Un document détaillant les besoins du projet, les objectifs, les parties prenantes, les exigences fonctionnelles et techniques, et les attentes en matière de données à publier sur le portail Open Data.
- Plan de Projet : Un plan de projet décrivant la portée, les étapes, les ressources, les délais et les responsabilités pour la création et la mise en œuvre du portail Open Data.
- Rapport d'État Initial : Un rapport initial qui décrit l'état actuel des données publiques, les défis et les opportunités identifiés, ainsi que les domaines qui bénéficieront le plus de la mise en œuvre du portail Open Data.
- Design et Wireframes du Portail : Des maquettes et des wireframes ou schémas du portail Open Data, montrant la structure, la mise en page, la navigation et l'organisation des données.
- Code Source du Plugin (si applicable) : Le code source du plugin personnalisé créé pour le portail Open Data, accompagné d'une documentation pour faciliter la maintenance et les mises à jour ultérieures.
- Rapport de Collecte et de Validation des Données : Un rapport décrivant les méthodes de collecte des données publiques, les sources, les processus de normalisation, de nettoyage et de validation des données, ainsi que des statistiques sur la qualité des données publiées.
- Interface Utilisateur du Portail : Une interface utilisateur fonctionnelle du portail Open Data, comprenant les fonctionnalités de recherche, de filtrage, de visualisation des données, etc.
- Documentation de l'API (si applicable) : Une documentation détaillée de l'API du portail Open Data, si celle-ci est mise à disposition pour les développeurs tiers.
- Manuel d'Utilisation : Un manuel d'utilisation expliquant comment naviguer et utiliser le portail Open Data, comment accéder aux données, et comment utiliser les fonctionnalités disponibles.
- Rapport Final : Un rapport final résumant l'ensemble du projet, les réalisations, les défis relevés, les leçons apprises et les prochaines étapes pour la pérennisation du portail Open Data.





## 2.6 RESSOURCES NÉCESSAIRES

### 2.6.1 Ressources Humaines

Pour la réalisation d'un portail Open Data avec un CMS , on aura besoin d'une équipe de ressources humaines avec des compétences spécifiques adaptées à l'utilisation du CMS choisi. Voici une liste des rôles clés pour la création du portail Open Data avec un DMS :

- **Chef de Projet** : Responsable de la planification, de la coordination et de la supervision générale du projet, en veillant à ce que les objectifs soient atteints dans les délais fixés.
- **Développeurs Web** : Des développeurs spécialisés dans le CMS spécifique que vous allez utiliser pour personnaliser le portail Open Data, créer des thèmes, des plugins et implémenter des fonctionnalités spécifiques.
- **Designer Web/UI-UX** : Chargé de la conception de l'interface utilisateur du portail, en s'assurant qu'elle est intuitive, conviviale et attrayant.
- **Spécialiste en Base de Données** : Pour gérer la base de données du CMS et s'assurer que les données publiées sur le portail sont stockées et récupérées de manière efficace.
- **Spécialiste en Open Data** : Un expert ayant une connaissance approfondie des principes de l'Open Data, des normes, des licences et des meilleures pratiques pour assurer la qualité et la réutilisabilité des données publiées.
- **Spécialiste en Sécurité** : Pour garantir que le portail Open Data est sécurisé contre les vulnérabilités et les attaques potentielles, et pour mettre en œuvre des mesures de sécurité appropriées.
- **Spécialiste en Collecte de Données** : Pour faciliter la collecte des données publiques auprès des organismes gouvernementaux et d'autres sources pertinentes, et s'assurer de leur qualité.
- **Spécialiste en Visualisation de Données** : Pour créer des visualisations de données interactives et informatives afin de rendre les données plus compréhensibles pour les utilisateurs.
- **Testeurs**: Des testeurs pour évaluer le portail, s'assurer de son bon fonctionnement, que les fonctionnalités sont conformes aux spécifications et qu'il n'y a pas de bugs ou d'erreurs. Chaque membre de l'équipe jouera un rôle important dans la réussite du projet de portail Open Data.

### 2.6.2 Ressources Matérielles

Pour la réalisation d'un portail Open Data , vous aurez besoin de certaines ressources matérielles pour héberger le DMS, stocker les données et assurer le bon fonctionnement du portail. Voici une liste des principales ressources matérielles nécessaires.

- **Serveur Web** : Vous aurez besoin d'un serveur web pour héberger le CMS et le portail Open Data. Vous pouvez opter pour un serveur dédié, un serveur virtuel privé (VPS) ou un hébergement cloud en fonction de vos besoins et de vos ressources disponibles.
- **Stockage** : Un espace de stockage suffisant pour héberger les fichiers du CMS, les données du portail Open Data et les éventuels fichiers médias, images, vidéos, etc
- **Base de Données** : Un système de gestion de base de données (par exemple, MySQL, Post-





greSQL) pour stocker les données du portail Open Data.

- Sauvegardes : Une solution de sauvegarde régulière pour assurer la sécurité des données et la possibilité de restaurer le site en cas de problème.
- Connexion Internet : Une connexion Internet rapide et fiable pour permettre l'accès au portail et faciliter les mises à jour du CMS et des données.
- Ordinateurs et Périphériques : Des ordinateurs et des périphériques pour les membres de l'équipe travaillant sur le développement, la conception et la gestion du portail Open Data.
- Logiciels : Les logiciels nécessaires pour le développement, la gestion de projet, la conception graphique, etc.
- Équipement de Test : Des appareils et des outils pour tester le portail sur différentes plates-formes et navigateurs pour assurer une compatibilité optimale.
- Services Cloud : Selon les besoins et les ressources disponibles, vous pouvez utiliser des services cloud pour l'hébergement, la sauvegarde ou la gestion de certaines fonctionnalités.

Se Assurer de prendre en compte les exigences techniques du CMS (par exemple, les ressources minimales requises, les extensions recommandées) lors du choix des ressources matérielles. Il est également possible d'utiliser des solutions d'hébergement gérées qui offrent des services spécifiquement optimisés pour l'exécution du CMS. En fonction de la taille du projet et de la charge de trafic attendue, on peut adapter les ressources matérielles pour répondre aux besoins spécifiques de votre portail Open Data.

### 2.6.3 Ressources financière

Les ressources financières jouent un rôle crucial dans la réalisation d'un projet de portail Open Data avec un DMS. Elles couvrent les dépenses liées à l'ensemble du projet, y compris le développement du portail, les coûts d'hébergement, les frais de personnel, les licences de logiciels, les coûts de formation, les dépenses de communication, etc. Voici quelques éléments importants à prendre en compte concernant les ressources financières :

- Budget de Projet : Établissez un budget réaliste pour le projet, en tenant compte des coûts associés à chaque étape, du développement du portail jusqu'à son lancement et sa maintenance.
- Dépenses liées au CMS : Si on utilise un CMS open-source les coûts liés au logiciel en lui-même seront généralement limités aux frais de personnalisation, de plugins, de thèmes et d'éventuelles extensions payantes.
- Hébergement et Infrastructure : Les coûts d'hébergement dépendront du type de serveur choisi (dédié, VPS ou cloud) et des spécifications matérielles requises pour répondre aux besoins du portail Open Data.
- Ressources Humaines : Le coût du personnel dépendra de la taille de l'équipe et des compétences nécessaires. Assurez-vous de prévoir le coût des développeurs, des concepteurs, des experts en base de données, etc





- **Formation** : Si des formations sont nécessaires pour le personnel ou les parties prenantes impliquées dans le projet, incluez également ces coûts dans le budget.
- **Maintenance et Mises à Jour** : Prévoyez également des ressources financières pour la maintenance continue du portail, les mises à jour du CMS, des plugins et des thèmes, ainsi que pour répondre aux besoins futurs du projet.
- **Sécurité et Sauvegardes** : Investissez dans des solutions de sécurité pour protéger le portail et les données, ainsi que dans des sauvegardes régulières pour éviter la perte de données en cas de problème.
- **Contingences** : Il est toujours conseillé de prévoir un montant pour les imprévus ou les situations d'urgence qui pourraient survenir pendant la réalisation du projet.

Il est essentiel de planifier soigneusement les ressources financières dès le début du projet et de les allouer de manière responsable pour assurer la réussite du portail Open Data.



## MÉTHODOLOGIE, IMPLÉMENTATION ET RÉSULTATS

Dans ce chapitre, nous allons présenter les résultats obtenus dans notre travail



Lors des Premières semaines à l'entreprise, il nous a été demandé d'automatiser certaines tâches notamment :

- De Produire un script qui permet naviguer dans un dossier et de compter le nombre de fichiers présent dans ce dossier par date puis rangés le résultat dans un autre fichier.
- D'écrire un programme qui permet de compter le nombre de caractères du numéro de chassie des véhicules (dans un fichier excel ) ceci dans le but de connaître les numéros de chassies qui ont été mal saisie.
- D'écrire un programme qui permet de filtrer les numéros de chassies comportant les lettre 'o', 'q' et 'i'.
- etc

Ensuite m'a été assigné la tâches de concevoir un système d'information des accidents utilisant Open DATA CKAN.

### 3.1 Choix de l'Outil et technologies utilisés

Avec la généralisation de l'ouverture des données dans les administrations et dans les services publics, de nombreux décideurs se posent des questions sur la stratégie à observer concernant les plateformes open data et les outils de gestion des données. En effet, une politique d'ouverture des données fait apparaître de nombreux challenges dont le choix d'une plateforme open data reste un point important. Il existe une grande disparité dans les outils pour la mise en œuvre de ces projets

- CKAN : Une solution libre avec une communauté active est la solution bien souvent préférée par les « gros » sites d'open data. Reste le problème de l'hébergement qui doit être assuré indépendamment de la plateforme. Des combinaisons avec d'autres CMS sont possibles pour assurer un hébergement et une interface plus conviviale. D'un point de vue technique, CKAN est un logiciel libre donc le code source est accessible et modifiable, il s'agit de code en Python pour le backend et de JavaScript pour la partie interface. Des technologies de gestion de données type PostgreSQL ou SQLAlchemy sont utilisés.
- Socrata : Socrata se différencie de CKAN par le fait qu'il s'agit d'une solution commerciale hébergée. Il propose aujourd'hui une solution open source pour faciliter les migration. Les produits Socrata sont des produits propriétaires basés sur des technologies open source telles que PostgreSQL, MongoDB, Elasticsearch, Angular, D3.js, et Chef. Le marketplace permet de mettre à la disposition de tous des applications utilisant les données ouvertes.
- DKAN : DKAN est une autre plateforme open source populaire qui offre des fonctionnalités similaires à CKAN. Il est également facile à utiliser et dispose d'une interface utilisateur intuitive. DKAN est apprécié pour sa flexibilité et sa capacité à gérer de grands ensembles de données. Toutefois, il peut ne pas offrir autant d'options de personnalisation que CKAN.





## 3.2 OUTILS ET TECHNOLOGIES UTILISÉS

### 3.2.1 CKAN

CKAN est un système de gestion de données open-source (DMS) utilisé pour alimenter des centres de données et des portails de données. Il facilite la publication, le partage et l'utilisation des données. CKAN est utilisé pour créer des hubs de données et des portails de données, en permettant aux utilisateurs de publier, partager et utiliser facilement les données. Une plateforme open data sur les accidents de la route présente de nombreux avantages importants pour la société, la sécurité routière, les gouvernements et d'autres parties prenantes. Voici quelques-unes des raisons pour lesquelles une telle plateforme est nécessaire :

- Amélioration de la sécurité routière : En rendant les données sur les accidents de la route accessibles au public, on sensibilise davantage aux problèmes de sécurité routière, ce qui peut contribuer à une conduite plus responsable et à une réduction des accidents.
- Prévention des accidents : L'analyse des données sur les accidents de la route peut aider à identifier les zones à haut risque et les facteurs de risque, ce qui permet de prendre des mesures préventives pour réduire le nombre d'accidents.
- Meilleure planification urbaine : Les données sur les accidents peuvent influencer la planification urbaine, l'aménagement des routes et la signalisation, ce qui peut contribuer à la création de villes plus sûres.
- Soutien à la recherche : Les chercheurs, les universitaires et les experts en sécurité routière peuvent utiliser ces données pour mener des recherches approfondies sur les causes des accidents et les tendances, ce qui peut conduire à des avancées dans la sécurité routière.
- Transparence gouvernementale : La publication des données sur les accidents de la route par les autorités gouvernementales renforce la transparence et la responsabilisation. Cela permet au public de surveiller la performance des gouvernements en matière de sécurité routière.
- Informations pour les assureurs : Les compagnies d'assurance peuvent utiliser ces données pour évaluer les risques et déterminer les primes, ce qui peut favoriser une tarification plus équitable.
- Soutien aux victimes : Les victimes d'accidents de la route, ainsi que leurs familles et avocats, peuvent utiliser ces données pour comprendre les circonstances d'un accident et éventuellement engager des poursuites judiciaires si nécessaire.
- Éducation et sensibilisation : Les données sur les accidents de la route peuvent être utilisées pour l'éducation du public et la sensibilisation à la sécurité routière, ce qui peut contribuer à une conduite plus prudente et à une réduction des accidents.
- Suivi des progrès : La collecte et la publication continues de données sur les accidents permettent de suivre les progrès dans la réduction des accidents de la route et de mesurer l'impact des initiatives de sécurité routière.
- Conduite autonome et véhicules connectés : Les données sur les accidents sont essentielles







pour le développement et le test de technologies de conduite autonome et de véhicules connectés, qui ont le potentiel de réduire considérablement les accidents de la route.

Il fournit des extensions pour émuler l'interaction de l'utilisateur avec les navigateurs, un serveur de distribution pour la mise à l'échelle de l'allocation des navigateurs, et l'infrastructure pour les implémentations de la W3C WebDriver spécification qui vous permet d'écrire un code interchangeable pour tous les principaux navigateurs Web.



Figure 3.1: CKAN

### 3.2.2 Pourquoi choisir CKAN ?

CKAN est un choix populaire pour la gestion de données ouvertes, et il présente plusieurs avantages qui expliquent pourquoi il est préféré par de nombreuses organisations par rapport à d'autres technologies open data. Voici quelques raisons pour lesquelles CKAN est utilisé :

- **Facilité d'utilisation** : CKAN est conçu pour être convivial. Il offre une interface utilisateur intuitive qui permet aux utilisateurs de cataloguer, publier et rechercher des données sans avoir besoin de compétences techniques avancées. Cela le rend accessible à un large public.
- **Flexibilité** : CKAN est hautement personnalisable et extensible. Il peut être adapté aux besoins spécifiques de votre organisation en ajoutant des extensions, en personnalisant l'apparence et en modifiant les fonctionnalités selon vos besoins.
- **Communauté active** : CKAN bénéficie d'une communauté mondiale active d'utilisateurs, de développeurs et de contributeurs. Cela signifie que vous pouvez trouver un support, des ressources et des extensions supplémentaires facilement, ce qui peut faciliter le développement et la maintenance.
- **Normes et bonnes pratiques** : CKAN suit des normes et des bonnes pratiques en matière de gestion de données ouvertes, ce qui en fait un choix solide pour les organisations qui cherchent à se conformer aux directives de l'open data, notamment celles émises par des gouvernements.
- **Intégration de données géospatiales** : CKAN prend en charge la gestion de données géospatiales, ce qui le rend approprié pour les organisations qui ont besoin de publier des données cartographiques ou liées à la localisation.





- **Évolutivité** : CKAN peut gérer de grandes quantités de données, ce qui en fait une solution adaptée pour les organisations ayant d'importants ensembles de données à gérer.
- **Sécurité** : CKAN offre des fonctionnalités de sécurité pour protéger les données sensibles et contrôler l'accès aux informations.
- **Métadonnées avancées** : CKAN permet de fournir des métadonnées détaillées pour chaque ensemble de données, améliorant la recherche et la compréhension des données.
- **API robuste** : Il offre une API (Application Programming Interface) puissante qui permet l'intégration avec d'autres systèmes et applications.
- **Support international** : CKAN est utilisé dans le monde entier, ce qui signifie qu'il est adapté à un large éventail de contextes et de langues.

### 3.2.3 Types de données qui peuvent être importer dans une plateforme CKAN

CKAN est conçu pour gérer et publier une grande variété de types de données. Voici quelques-unes des structures de données courantes que vous pouvez importer sur CKAN :

- **Fichiers tabulaires** : CKAN peut gérer des fichiers de données tabulaires au format CSV, Excel, ou d'autres formats similaires. Cela inclut les données dans des feuilles de calcul, des bases de données ou d'autres sources tabulaires.
- **Données géospatiales** : CKAN prend en charge les données géospatiales, y compris les données cartographiques au format GeoJSON, Shapefile, KML, etc. Cela permet de créer des cartes interactives et de visualiser des informations géographiques.
- **Données temporelles** : Les données temporelles, telles que les séries chronologiques, les plages de dates et les horodatages, peuvent être importées et gérées dans CKAN. Cela est utile pour la visualisation de tendances temporelles.
- **Données non structurées** : CKAN peut gérer des documents non structurés, tels que des rapports, des PDF, des images, des vidéos, des fichiers texte, etc. Vous pouvez ajouter des métadonnées décrivant ces documents.
- **Données RDF (Resource Description Framework)** : CKAN peut gérer des données RDF, qui sont couramment utilisées pour représenter des données sémantiques. Cela peut être utile pour les données liées et les ontologies.
- **Données de streaming** : CKAN prend en charge la publication de données en temps réel, ce qui permet de suivre en continu des flux de données en direct.
- **Données liées à des API externes** : CKAN peut intégrer des données provenant d'API externes. Cela permet d'agréger des données en provenance de sources variées.

## 3.3 Analyse

### 3.3.0.1 Diagramme de Cas d'utilisation



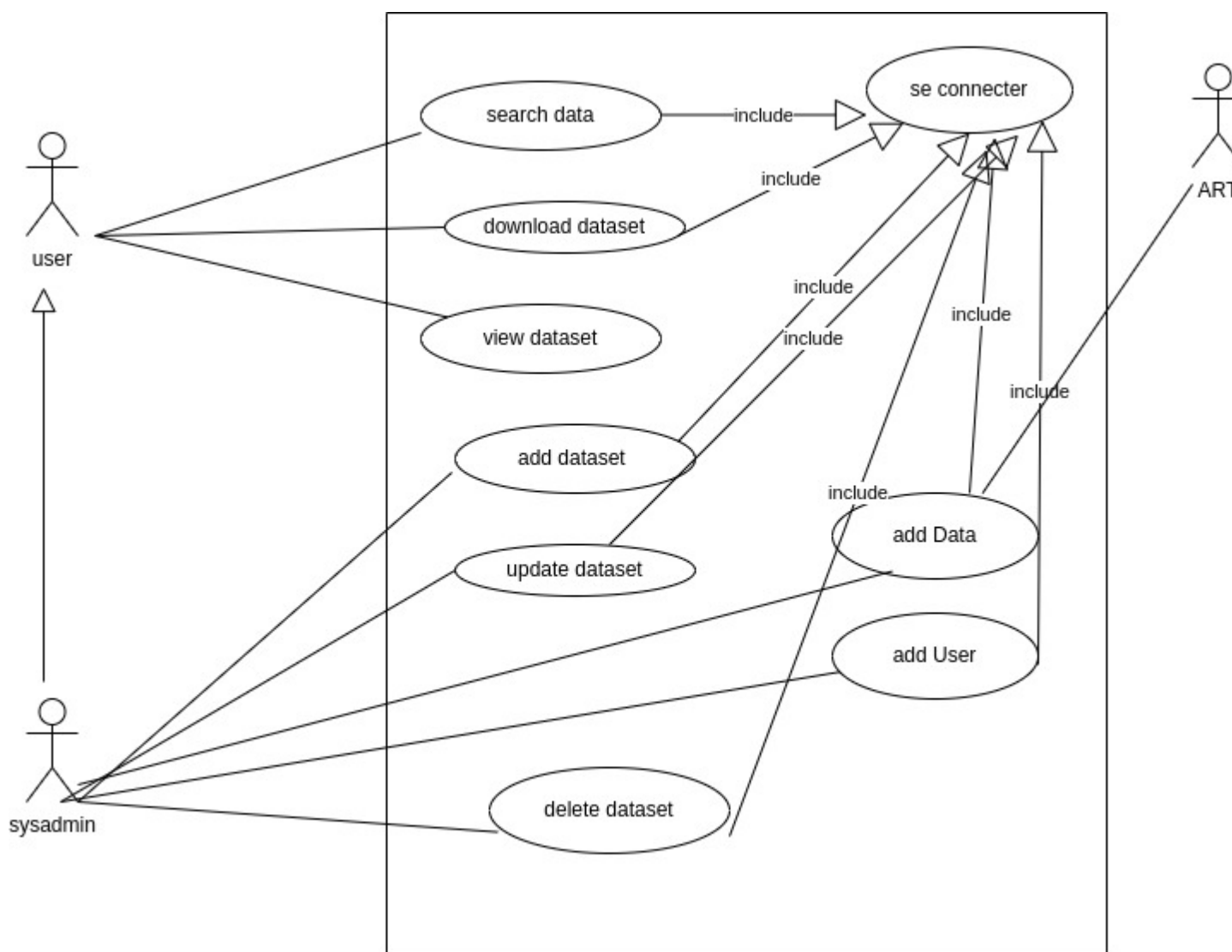


Figure 3.2: Diagramme

### 3.3.1 description détaillée

#### 3.3.1.1 sysadmin

. L'administrateur système (sysadmin) du portail Open Data assume les exigences suivantes:

- Installation et configuration du portail Open Data : il est responsable de l'installation du portail Open Data sur les serveurs, ainsi que de sa configuration initiale en fonction des besoins spécifiques de l'organisation.

- Gestion des utilisateurs : Vous devrez créer et gérer les comptes d'utilisateurs sur le portail, en veillant à ce que les autorisations et les rôles soient correctement attribués. il mets en place



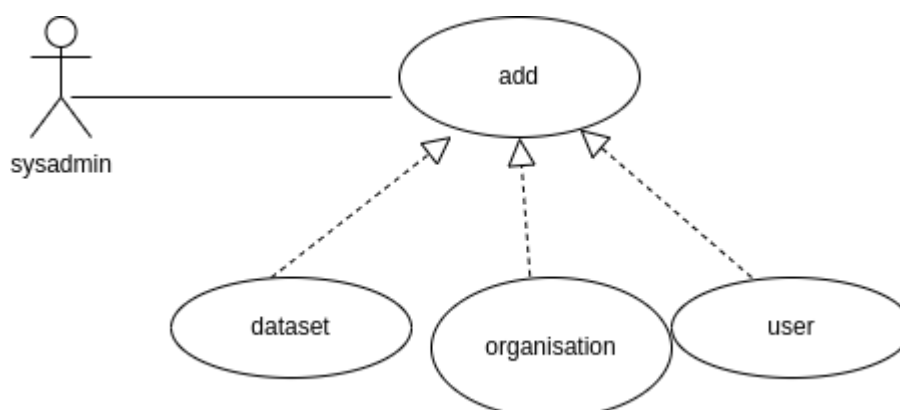


Figure 3.3: Diagramme

des mécanismes d'authentification pour sécuriser l'accès aux données.

• **Gestion des données** : il est responsable de la gestion des ensembles de données sur le portail. Cela peut inclure l'importation de nouvelles données, la vérification de leur qualité, la suppression des données obsolètes, ainsi que l'organisation et le classement des ensembles de données pour faciliter leur recherche et leur utilisation.

### 3.3.1.2 Utilisateur

. L'utilisateur du portail Open Data assume les exigences suivantes:

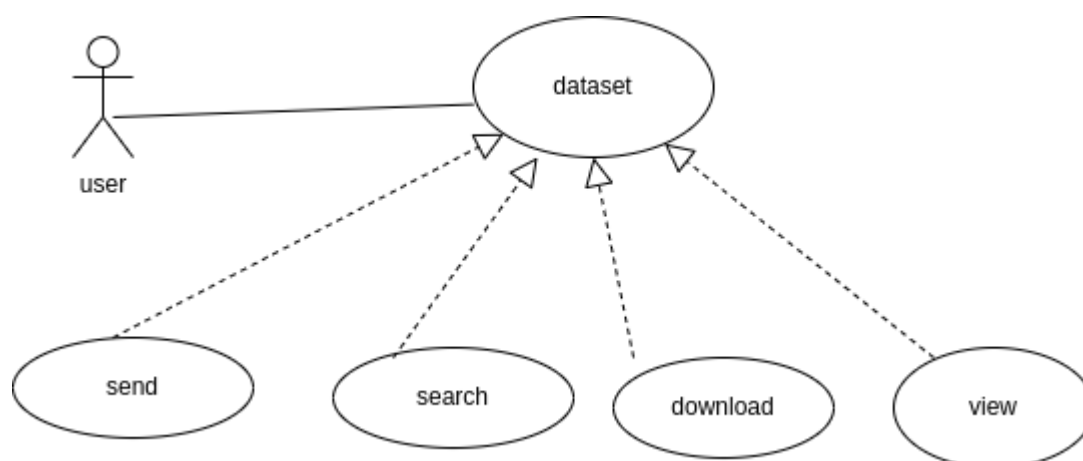


Figure 3.4: Diagramme

• **Recherche de données** : il effectue des recherches spécifiques à l'aide de mots-clés, de catégories ou de filtres pour trouver les ensembles de données qui vous intéressent.



- **Exploration des ensembles de données** : il navigue à travers les différentes collections ou catégories de données disponibles sur le portail pour découvrir de nouvelles informations.

- **Téléchargement de données** : il a la possibilité de télécharger les ensembles de données en divers formats (CSV, JSON, XML, etc.) afin de les utiliser pour vos propres analyses et projets.

- **Visualisation des données** : le portail Open Data proposent des outils de visualisation qui permet d'explorer les ensembles de données de manière interactive, en utilisant des graphiques, des cartes ou d'autres représentations visuelles.

## 3.4 Installation de CKAN

CKAN est conçu pour fonctionner sur le framework .NET/Mono 4.5, donc il est compatible avec les systèmes d'exploitation Windows, Mac et Linux. Dans notre projet nous avons installer CKAN 2.9.

### 3.4.0.1 Prérequis et Exigences du serveur

Pour déployer CKAN sur un serveur, on doit prendre en compte plusieurs exigences en termes de matériel, de système d'exploitation, de logiciels et de configuration.

- **Système d'exploitation :**

CKAN est conçu pour fonctionner sur des systèmes d'exploitation basés sur Linux. Les distributions Linux couramment utilisées pour CKAN sont Ubuntu(20.04 ou 22.04 pour CKAN 2.9), CentOS, Debian, et Red Hat Enterprise Linux (RHEL).

- **Serveur Web :**

On a besoin d'un serveur web pour héberger CKAN. Apache est généralement recommandé, bien que Nginx puisse également être utilisé.

- **Base de données :**

CKAN nécessite une base de données pour stocker ses métadonnées et ses informations internes. PostgreSQL est fortement recommandé. On doit également installer l'extension PostGIS si nous avons besoin de prendre en charge des données géospatiales.

- **Langage de programmation** : est écrit en Python, donc vous aurez besoin de Python 3.8 installé sur votre serveur pour déployer CKAN 2.9.

- **Solr :**

CKAN utilise Solr pour la recherche textuelle. On doit installer et configurer Solr pour que CKAN fonctionne correctement.





- **Espace disque :**

On a besoin d'un espace disque suffisant pour stocker les données que vous publierez via CKAN, ainsi que pour les fichiers système et les journaux.

- **RAM et CPU :**

Les exigences en matière de RAM et de CPU dépendront de la charge de travail prévue. Pour un usage de base, un serveur avec au moins 2 Go de RAM et un CPU dual-core devrait suffire. Cependant, pour un déploiement plus important, nous devons augmenter ces ressources.

- **SSL (facultatif) :**

OSi on prévoit d'utiliser des connexions sécurisées, on doit configurer un certificat SSL pour votre serveur.

- **Certificat de domaine (facultatif) :**

Si on prévoit d'utiliser un domaine personnalisé pour l'instance CKAN, on doit disposer d'un certificat de domaine valide.

- **Configuration du pare-feu et des port :**

Se Rassurer que le ports nécessaire (5000) est ouvert sur le serveur et configurés dans le pare-feu.

### 3.4.1 Installation de CKAN 2.9 Sur Ubuntu 20 .04

1. Se Connecter en tant qu'utilisateur root, entrez le mot de passe lorsque vous y êtes invité  
`Sudo -i`
2. Mettre à jour les paquets
3. Mettre à jour les paquets  
`Sudo apt update upgrade`
4. Installez les packages requis par Ubuntu pour CKAN  
`sudo apt install -y libpq5 redis-server nginx supervisor`
5. Télécharger le package CKAN. Dans ce guide, nous installons CKAN pour Python 3  
`wget https://packaging.ckan.org/python-ckan2.9-py3-focalamd64.deb`

6. Installer les packages supplémentaires nécessaires

`sudo apt-get install python3.8-distutils`

7. Installer le paquet CKAN

`sudo dpkg -i python-ckan2.9-py3-focalamd64.deb`





## 8. Installer PostgreSQL

```
sudo apt install -y postgresql
```

## 9. Vérifier qu'il est correctement installé en exécutant la commande suivante et en vous assurant que l'encodage des bases de données est UTF8

```
sudo -u postgres psql -l
```

Name	Owner	Encoding	Collate	Ctype	
ckan_default	ckanuser	UTF8	fr_FR.UTF-8	fr_FR.UTF-8	=
datastore_default	ckanuser	UTF8	fr_FR.UTF-8	fr_FR.UTF-8	=
postgres	postgres	UTF8	fr_FR.UTF-8	fr_FR.UTF-8	=
template0	postgres	UTF8	fr_FR.UTF-8	fr_FR.UTF-8	=
template1	postgres	UTF8	fr_FR.UTF-8	fr_FR.UTF-8	=

Figure 3.5: Postgresql

## 10. Créer un utilisateur de base de données et créez un mot de passe pour le nouvel utilisateur lorsque vous y êtes invité.

```
sudo -u postgres createuser -S -D -R -P username
```

11. Créer une nouvelle base de données PostgreSQL appartenant au nouvel utilisateur, remplacer `ckan_default` par un nom de base de données de votre choix et `username` par le nom d'utilisateur qui

```
sudo -u postgres createdb -O username ckan_default -E utf-8
```

## 12. Si vim n'est pas installé, vous devrez l'installer avec ce qui suit

```
apt install vim
```

## 13. Modifier le fichier de configuration CKAN et remplir le mot de passe, la base de données et l'utilisateur de la base de données pour la base de données que vous avez créée.

```
vim /etc/ckan/default/ckan.ini
```

## 14. Installer Solr

```
sudo apt install -y solr-tomcat
```

## 15. Remplacer le port par défaut sur lequel Tomcat s'exécute par celui attendu par CKAN

```
vim /etc/tomcat9/server.xml
```

Editez la ligne suivante `<Port du connecteur="8080" protocole="HTTP/1.1"` À `<Port du connecteur="8983" protocole="HTTP/1.1"`

16. Remplacer le fichier `schema.xml` par défaut par un lien symbolique vers le fichier de schéma CKAN.



```
sudo mv /etc/solr/conf/schema.xml /etc/solr/conf/schema.xml.bak  
sudo ln -s /usr/lib/ckan/default/src/ckan/ckan/config/solr/schema.xml /etc/solr/conf/schema.xml
```

#### 17. Configurer Solr

```
vim /etc/ckan/default/ckan.ini  
solr_url = http://51.210.48.159:8983/solr/ckanRedémarrerSolr  
sudoservicetomcat9restart
```

#### 18. Vérifier que Solr est en cours d'exécution en saisissant ce qui suit dans votre navigateur

<http://51.210.48.153:8983/solr/>

#### 19. Configuration de base du site CKAN

```
vim /etc/ckan/default/ckan.ini  
ckan.devserver.host = 51.*.*.* (adresse ip de la machine)  
ckan.devserver.port = 5000  
ckan.site_url = http://51.*.*.*:5000
```

#### 20. Initialiser la base de données CKAN en exécutant la commande suivante

```
sudo ckan db init
```

#### 21. Recharger le Supervisor pour que les nouveaux processus soient récupérés

```
sudo supervisorctl reload
```

#### 22. Vérifier l'état des processus

```
sudo supervisorctl status  
Ce qui suit devrait apparaître sans erreur  
ckan-datapusher:ckan-datapusher-00 RUNNING pid 1963, uptime 0:00:12  
ckan-uwsgi:ckan-uwsgi-00 RUNNING pid 1964, uptime 0:00:12  
ckan-worker:ckan-worker-00 RUNNING pid 1965, uptime 0:00:12
```

#### 23. Redémarrer Nginx

```
sudo service nginx restart
```

Entrez ce qui suit dans votre navigateur Web et la page d'accueil de CKAN devrait apparaître  
[http://51.\\*.\\*.\\*:5000](http://51.*.*.*:5000)





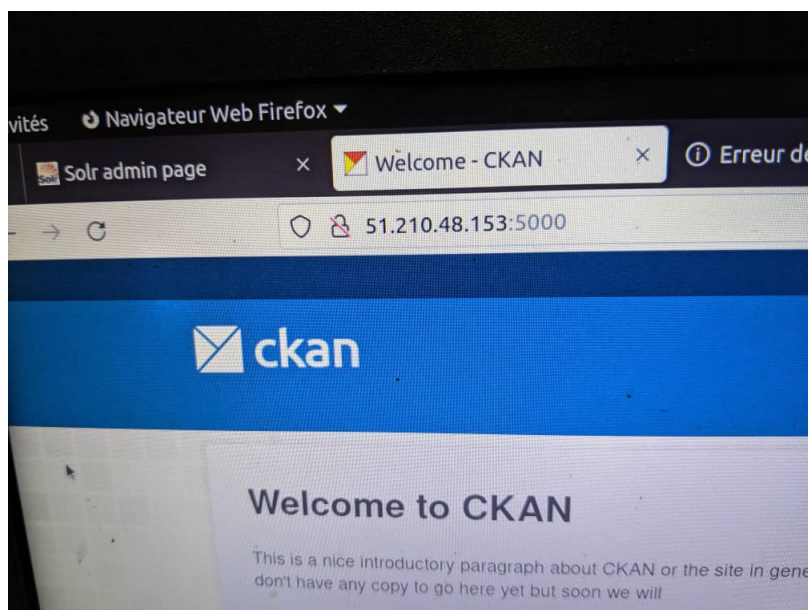


Figure 3.6: ckan installation

### 3.4.2 Configuration

nous présentons quelques configuration à faire pour répondre aux objectifs cités plus haut.

- Configuration du datastore : c'est un outil puissant qui facilite la gestion, la publication, l'accès et l'analyse de données tabulaires au sein de la plateforme CKAN. Il permet aux organisations, aux gouvernements et aux particuliers de mieux exploiter et de partager leurs données, contribuant ainsi à la transparence, à la recherche, à la prise de décision éclairée et à la collaboration dans divers domaines. Pour l'installer il faut suivre la procédure suivant la documentation officielle <https://docs.ckan.org/en/2.9/maintaining/datastore.html>.

- Configuration du Filestore : le FileStore de CKAN est un composant essentiel pour la gestion et le stockage des fichiers de données associés à des jeux de données dans la plateforme. Pour l'installer il faut suivre la démarche suivant la documentation officielle <https://docs.ckan.org/en/2.9/maintaining/filestore.html>

- Configuration du datapusher : le Datapusher est un composant essentiel de CKAN qui simplifie la gestion et la préparation des données tabulaires pour une utilisation facile dans la plateforme. Il facilite la conversion, l'intégration, l'automatisation et l'analyse des données, ce qui contribue à une meilleure utilisation des données ouvertes et à une plus grande transparence dans diverses organisations et domaines d'application. Suivre le lien <https://github.com/ckan/datapusher> pour l'installation

- Créer le super administrateur : `./usr/lib/ckan/default/bin/activate`

```
cd /usr/lib/ckan/default/src/ckan
```

```
ckan -c /etc/ckan/default/ckan.ini sysadmin add seanh email=seanh@localhost name=nom_admin
```





La configuration nécessaire dans le fichier de configuration : type de fichier chargé dans CKAN

```
ckan.preview.loadable = html htm rdf+xml owl+xml xml n3 n-triples turtle plain atom csv tsv txt json pdf text/plain knl geotjson
ckan.display_timezone = server
```

Figure 3.7

• Connexion à une entité externe (comme mysql par exemple) : après avoir installé le datastore, ouvrir le fichier de configuration et ajouter la commande 'ckan.datastore.write\_url = mysql : //utilisateur : mot\_de\_passe@nom\_hôte/nom\_base\_de données'.

↳ Suivre la documentation officielle de CKAN 2.9 Pour toutes les configurations : [//docs.ckan.org/en/2.9/](http://docs.ckan.org/en/2.9/)

### 3.4.3 Résultats Obtenus

il à noter que CKAN n'est pas configuré pour être exécuté automatiquement en tant que service système. On doit le lancer manuellement dans n'importe quel répertoire via la commande 'sudo ckan -c /etc/ckan/default/ckan.ini run' puis entrer l'url de notre plateforme 'http://51.\*.\*.\*:5000' dans un navigateur.

- Interface D'accueil

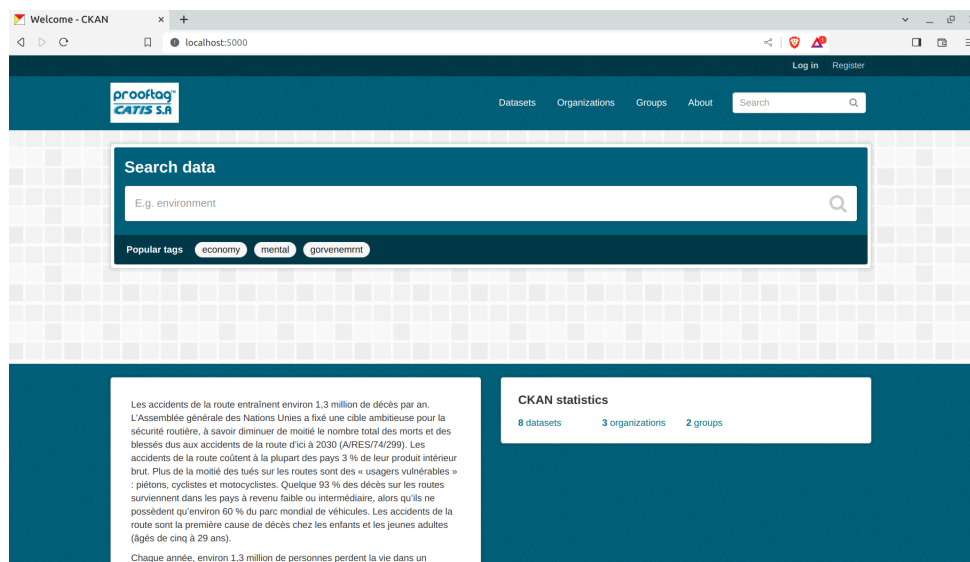


Figure 3.8

- Interface de connexion
- username : k\*\*\*\*
- password : \*\*\*\*\*



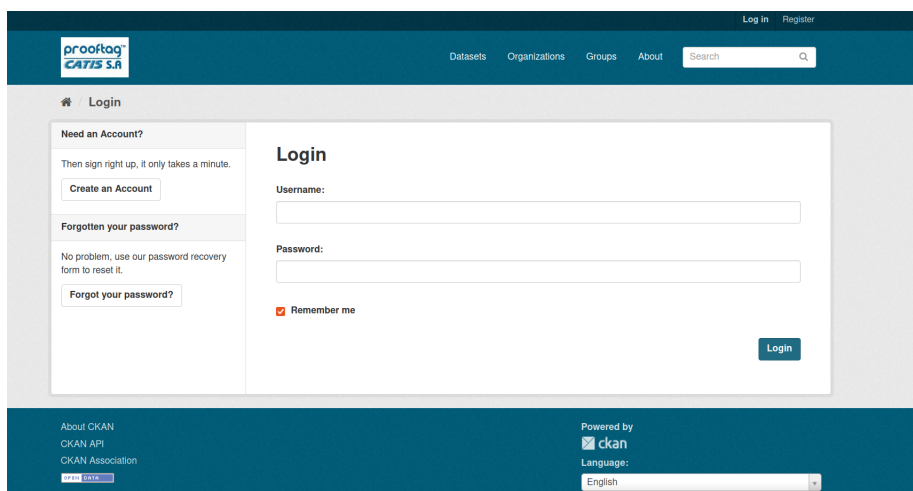


Figure 3.9

Les jeux de données qui seront présents sur notre portail proviendront de la base de données de l'application DOSER et seront transmis au portail via un outil de reporting. Ceci n'étant pas fait nous avons utilisés les jeux de données téléchargées en local pour effectuer des tests.

- Dashbord administrateur

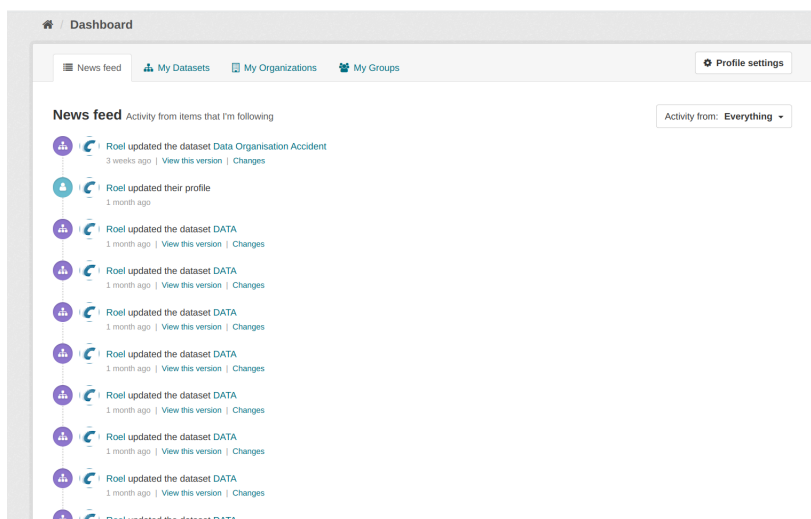


Figure 3.10

- Dashbord Datasets



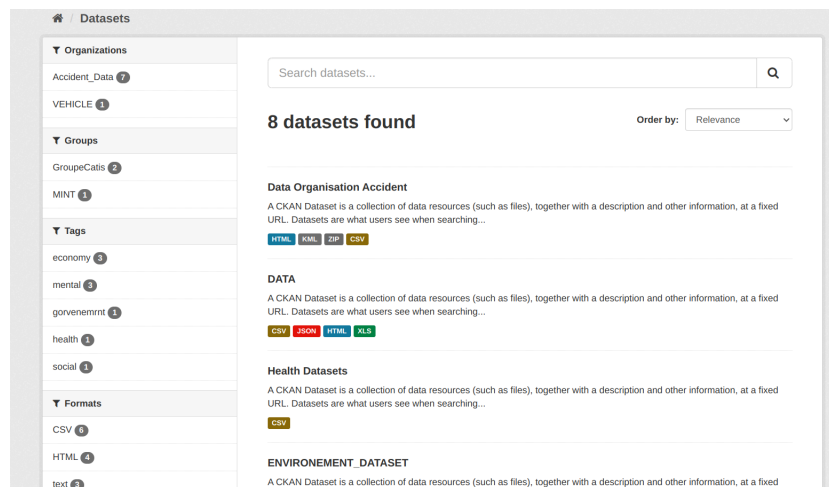


Figure 3.11

- Organisations

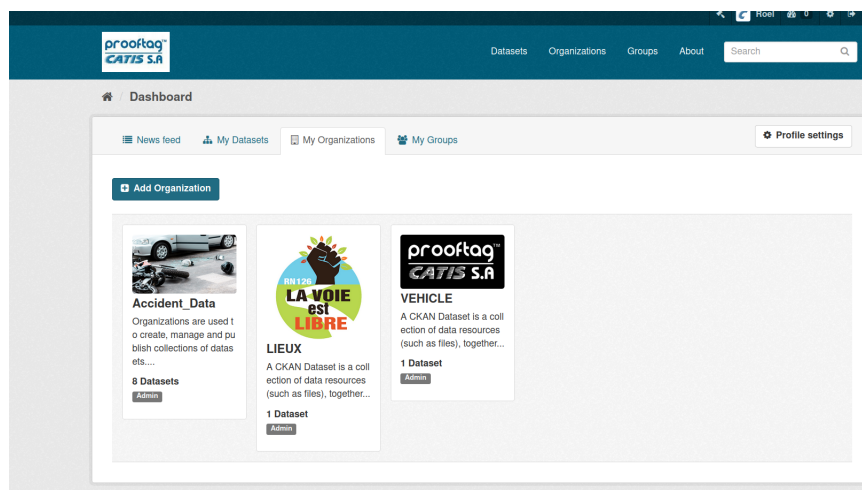


Figure 3.12

- Collecte de données



**What's a resource?**

A resource can be any file or link to a file containing useful data.

☒ New resource ← All resources

**Data:**

**Name:**

**Description:**  
  
You can use Markdown formatting here

**Format:**

This will be guessed automatically. Leave blank if you wish

Figure 3.13

- Visualisation des données(datastore grid)

**Lottery\_Mega\_Millions\_Winning\_Numbers\_\_Beginnin...**

URL: [http://localhost:5000/dataset/998530-1356-4488-9119-38781ba64d1/resource/209e1d90-ec0b-4063-9f6f-06108735e9/download/lottery\\_mega\\_millions\\_winning\\_numbe...](http://localhost:5000/dataset/998530-1356-4488-9119-38781ba64d1/resource/209e1d90-ec0b-4063-9f6f-06108735e9/download/lottery_mega_millions_winning_numbe...)

A CKAN Dataset is a collection of data resources (such as files), together with a description and other information, at a fixed URL. Datasets are what users see when searching for data.

Grid Graph Map about 2000 records

Id	Draw Date	Winning...	Mega Ball	Multiplier
1	2020-09-...	20 36 37...	16	2
2	2020-09-...	14 39 43...	19	3
3	2020-10-...	09 38 47...	25	2
4	2020-10-...	15 16 18...	17	3
5	2020-10-...	05 11 25...	13	2
6	2020-10-...	11 44 45...	25	2

Figure 3.14

- Visualisation des données(datastore graph)

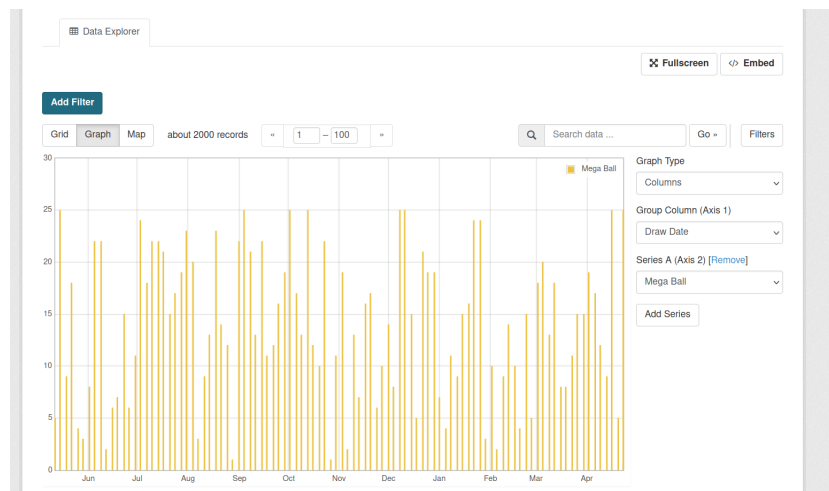


Figure 3.15

- Visualisation des données(datastore Map)

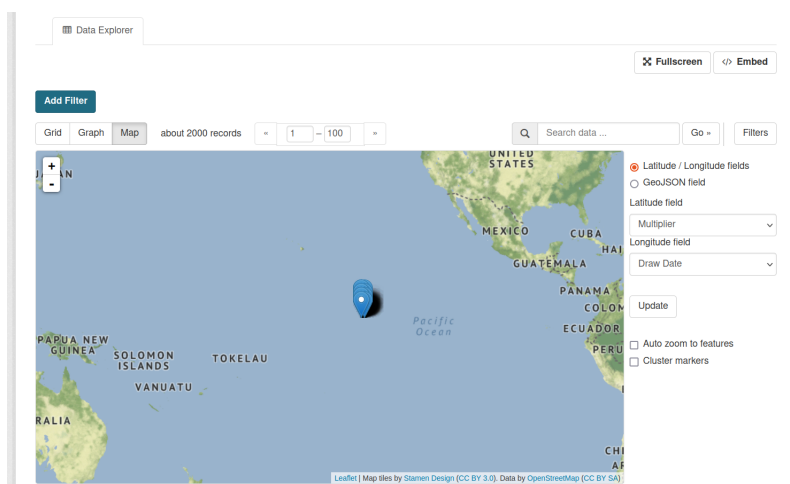


Figure 3.16

- Visualisation des données(Json View)

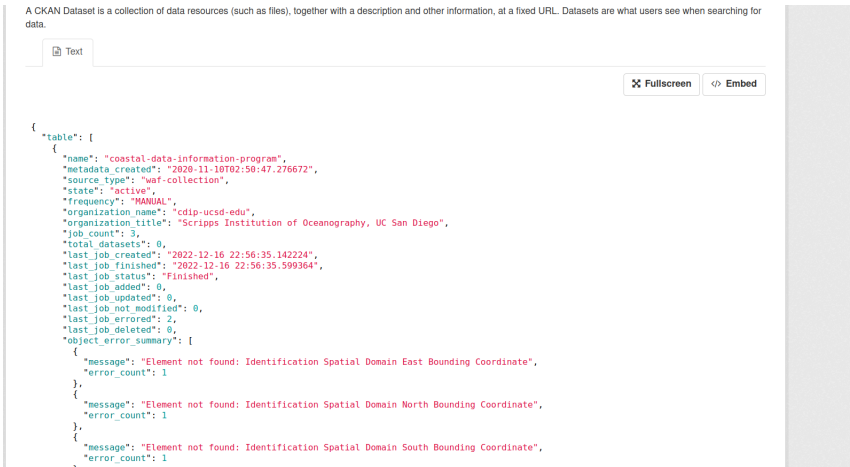


Figure 3.17

- Gestion des utilisateurs

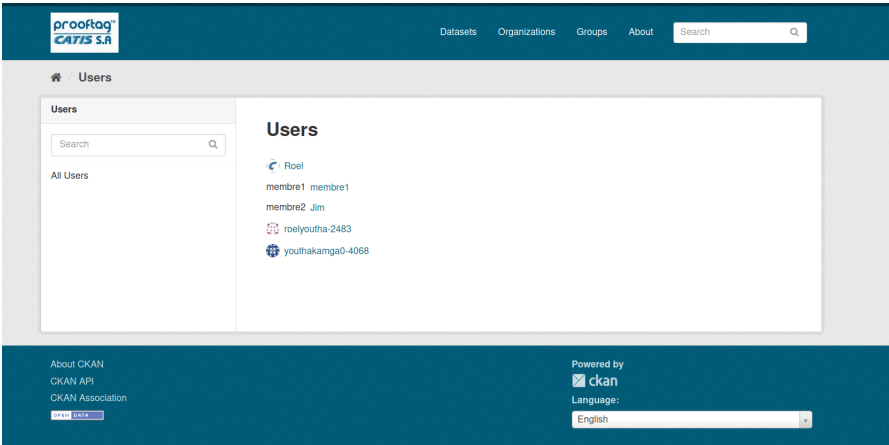


Figure 3.18

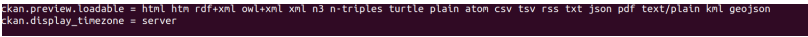


Figure 3.19

- 
- Partage des données



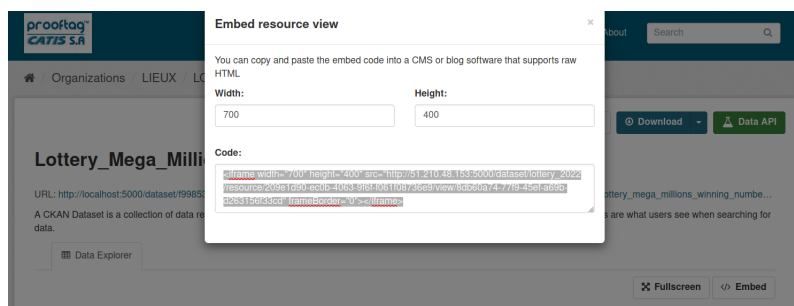


Figure 3.20

- Rapport Automatique sur les statistiques

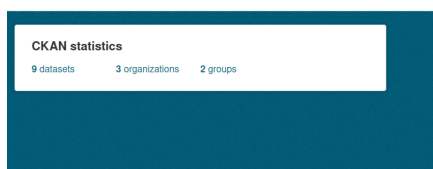


Figure 3.21

- Internationalisation

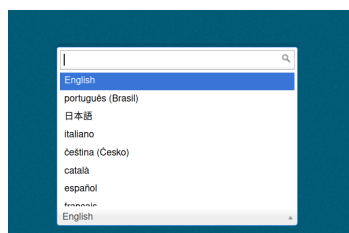


Figure 3.22

etc..

### 3.4.4 Perspectives

Pour rendre nos portails complets et respectant les besoins de l'organisation, il faut :

- Connecter la plateforme CKAN à un outil de reporting : ceci facilite la collecte des données sur la plateforme et assure également une mise à jour régulière des données sur la plateforme.
- Intégrer la plateforme CKAN à un CMS comme liferay : Pour mieux personnaliser la plateforme et l'intégrer facilement à l'application DOSER.

- Personnalisation des alertes : Permettre aux utilisateurs de s'abonner à des alertes pour être informés des mises à jour et des événements liés à la sécurité routière dans leur région.

- Récolte de retours : Solliciter les commentaires des utilisateurs pour améliorer en permanence







le portail et les données qu'il contient.

- intégrer les Outils d'analyse : Intégrer des outils d'analyse de données pour permettre aux utilisateurs de mener des analyses plus poussées sur les données d'accidents de la route.
- Analyse prédictive des accidents : Utiliser l'IA pour analyser les données historiques des accidents de la route et développer des modèles prédictifs qui identifient les zones à risque élevé. Les prédictions peuvent aider les autorités à prendre des mesures préventives.
- Détection automatique d'accidents : Intégrer des systèmes de vision par ordinateur pour détecter automatiquement les accidents à partir de caméras de surveillance ou de vidéos de tableau de bord. Cela peut aider à signaler rapidement les accidents et à améliorer les temps de réponse des secours.
- Analyse des causes d'accidents : Utiliser l'IA pour analyser les données d'accidents et identifier les causes sous-jacentes. Cela peut aider les autorités à mettre en œuvre des politiques ciblées pour réduire les types spécifiques d'accidents.
- Recommandations pour l'amélioration de la sécurité routière : Utiliser l'IA pour générer des recommandations basées sur les données, telles que des améliorations de l'infrastructure routière, des modifications de la réglementation ou des campagnes de sensibilisation.
- Systèmes d'assistance à la conduite : Intégrer des systèmes d'assistance à la conduite basés sur l'IA pour aider les conducteurs à éviter les situations dangereuses, comme le freinage d'urgence automatique.



## CONCLUSION

### Rappel du problème et difficultés rencontrées

Rendu au terme de notre travail , il était question pour nous de concevoir et de déployer une plateforme open data sur les accidents de la routes . Cette plateforme open data sur les accidents de la route est un outil essentiel pour améliorer la sécurité routière, prévenir les accidents, éduquer le public et soutenir la recherche. Elle permet également une plus grande transparence gouvernementale et contribue à la création de villes plus sûres et de routes plus sécurisées. Tout au long de mon stage j'ai toujours du doigt les projets qui visent à résoudre les problèmes de la vie réelle et améliorer mes connaissances.

Tout au long de notre travail, nous avons rencontré de nombreuses difficultés :

- La première difficulté que nous avons rencontrée est la prise en main de CKAN qui est un outil utilisé pour la mise sur pied des plateformes open data.
- Ensuite, l'instabilité de la connexion internet
- L'instabilité du serveur sur lequel est déployée cette plateforme

## BIBLIOGRAPHIE

1. CKAN 1.9 documentation <https://orihoch-ckan.readthedocs.io/en/latest/maintaining/installing/install-from-source.html>
2. List of Extensions <https://github.com/ckan/ckan/wiki/List-of-extensions>
3. Open Data tools <https://www.stat4decision.com/fr/plateformes-open-data/>
4. CKAN datastore <https://orihoch-ckan.readthedocs.io/en/latest/maintaining/datastore.html>
5. CKAN filestore <https://orihoch-ckan.readthedocs.io/en/latest/maintaining/datastore.html#relationship-to-filestore>
6. <https://ckanext-importer.readthedocs.io/en/stable>