Table des matières

[Description du commun 1](#_Toc47428560)

[Les questions à explorer 1](#_Toc47428561)

[Les données à disposition en Open Data 1](#_Toc47428562)

[Livrables attendus 2](#_Toc47428563)

[Solution pour la captation pour les flux cita.lu et telraam 2](#_Toc47428564)

[Analyse et prédiction 2](#_Toc47428565)

[Description des données 2](#_Toc47428566)

[Données du trafic des autoroutes au Luxembourg 2](#_Toc47428567)

[Données du traffic des autoroutes 2](#_Toc47428568)

[Données des événements trafic 3](#_Toc47428569)

[Données météo 3](#_Toc47428570)

[Données du trafic place général de gaulle (mouans sartoux, France) 4](#_Toc47428571)

# Description du commun

L'objet de ce commun est de produire de la connaissance pour identifier quelles sont les conditions minimales pour alimenter une IA capable de prédire les flux de circulation à un certain horizon de temps.

## Les questions à explorer

* Valider l’hypothèse qu'il est possible d'avoir une IA qui prédise les flux de circulation avec 5, 10, 15 minutes en avance
* Identifier le faisceau des données nécessaires pour produire une IA pertinente
* Clarifier les besoins en infrastructure de captation et distribution des données pour alimenter l'IA

## Les données à disposition en Open Data

* Les données du trafic des autoroutes du Luxembourg à travers 186 caméras via [www.cita.lu](http://www.cita.lu)
* Les données de la caméra place général de gaulle (Mouans-Sartoux) via [www.telraam.net](http://www.telraam.net)

## Livrables attendus

### Solution pour la captation pour les flux [cita.lu](http://cita.lu/) et telraam

* Batch d'alimentation : batch générique Open Source en java à l’aide du Framework Spring permettant de récolter les données des caméras ([spring.io/](https://spring.io/))
* Intégration et alimentation d’une base de données de type « Time Series » (timescaledb de Postgresql [www.timescale.com/](https://www.timescale.com/))
* Visualisation des données à l’aide d’outil dédié (Graphana [grafana.com/grafana/](https://grafana.com/grafana/))

Les livrables pour la solution d’alimentation sont le code source dans un dépôt Git  (java + sql), la documentation de la mise en œuvre et des exemples de dashboard de visualisation.

### Analyse et prédiction

Rapport d'analyse sur les résultats de l'algorithme prédictif et description de l’infrastructure nécessaire à son cycle de vie. Les outils utilisés sont les librairies python de « machine learning »  [scikit-learn.org/stable/](https://scikit-learn.org/stable/) [www.tensorflow.org/](https://www.tensorflow.org/) et [keras.io/](https://keras.io/)

Le rapport d'analyse décrit notamment :

* Quelles sont les données nécessaires pour la mise en œuvre d'un algorithme prédictif ?
* Quelles prédictions (flow, vitesse), quel horizon de temps?
* Décrire les besoins pour mettre cet algorithme en production

# Description des données

## Données du trafic des autoroutes au Luxembourg

L’ administration des Ponts et Chaussées du Luxembourg met à disposition un ensemble de données ouvertes via la plate-forme de données luxembourgeoise [data.public.lu](http://data.public.lu/)  
( <https://data.public.lu/fr/organizations/administration-des-ponts-et-chaussees>  )

Notamment les données suivantes peuvent être interrogées en temps réels au format DATEX II, format de données de trafic routier (norme européenne)  
( Datex II : <http://trafic-routier.data.cerema.fr/la-norme-europeenne-datex-ii-a58.html> )

* Données du trafic des autoroutes
* Données des événements trafic
* Données météo

Données du traffic des autoroutes  
Pour chacune des 180 caméras sur les autoroutes (A1, A3, A4, A6, A7, A13, B40) les données suivantes sont exposées toutes les 5 minutes:  
Données:

* Identifiant de la caméra
* Timestamp de la mesure
* Sens de prise de vue
* Latitude et longitude de la caméra
* Distance de la caméra au début de l’autoroute (caméra au km x )
* Vitesse moyenne en km/h
* Concentration du trafic en pourcentage
* Flux du trafic en nombre de voitures par heure

Format : XML DATEX II  
URLs A1 : <https://www.cita.lu/info_trafic/datex/trafficstatus_a1>  URL pour xx : <https://www.cita.lu/info_trafic/datex/trafficstatus_xx>   avec xx dans (A1, A3, A4, A6, A7, A13, B40)

Données des événements trafic  
L’API donne la liste des événements (situation) en cours.

On peut trouver les données suivantes à dispositions pour chaque situation, identifiant de la situation, timestamp de la création de la situation, position, description de l'événement.  
La description complète des données disponible est décrite ici : <https://diffusion-numerique.info-routiere.gouv.fr/IMG/pdf/interface-datexii-avectipi-2017-171124.pdf>

Le format est en XML DATEX II  
URL : <https://www.cita.lu/info_trafic/datex/situationrecord>

Données météo  
L’API donne des informations météos pour un ensemble de point de mesure.

* Identifiant du point de mesure
* Timestamp de la mesure
* Pourcentage d’humidité
* Précipitation en mm par heure
* Température de la surface de route
* Température de l'air
* Température de rosée
* Vitesse et direction du vent

Le format est en XML DATEX II  
URL : <https://www.cita.lu/info_trafic/datex/weather_dynamic.xml>

## Données du trafic place général de gaulle (mouans sartoux, France)

<https://www.telraam.net/fr/location/9000000411>

Cette caméra fait parti du dispositif Telraam qui met à disposition des citoyens intéressés des caméras de comptage. (<https://www.telraam.net/fr/what-is-telraam>)

Chaque heure la caméra relève les données suivantes :

* Pourcentage d'activité de la caméra
* Comptage des piétons, voitures, vélos, camions (total, gauche et droite de la rue)
* Histogramme des vitesses des voitures pour les intervalles [0-10[  [10-20[ [20- 30[     .....  [70 et plus [

Le format est en JSON  
URLs

* Récupération des données : <https://telraam-api.net/v0/reports/9000000411>
* Documentation de l’API : <https://telraam.zendesk.com/hc/en-us/articles/360027325572-Want-more-data-Telraam-API>

# Mise en œuvre de la DB

NOTES

Exemple

"segment\_id": "9000000411",

 "date": "2020-07-01T22:00:00.000Z",

 "pct\_up": 0.448981481481482,

"timezone": "Europe/Paris",

"pedestrian": 604.391643145026,

"bike": 1553.10258573979,

"car": 9226.5770122573,

"lorry": 1063.81217958995,

"pedestrian\_lft": 284.86737748684,

 "bike\_lft": 894.621476307705,

"car\_lft": 5193.30653377209,

"lorry\_lft": 364.777851777449,

"pedestrian\_rgt": 319.524265658186,

 "bike\_rgt": 658.481109432086,

 "car\_rgt": 4033.27047848521,

"lorry\_rgt": 699.034327812504,

"car\_speed\_histogram": [2289.38747172499,3104.1790056431,2722.70916279142,725.617402833619,143.54595136182,65.1015948034983,58.1040594288343,117.932363670011],

            "car\_speed\_bucket": [0,1,2,3,4,5,6,7]

Livrables attendus ( repris de la fiche commun)  
Connecteur Spring : Connecteur batch générique Open Source en java à l'aide du framework Spring permettant de récolter les données des caméras

Base de données « Time Series » : Intégration intégration et alimentation d'une base de données « Time Series »

Outil de visualisation : oOutil de visualisation de données  
Test IA de prédiction à l’aide d’outil de « machine learning »  
Rapport d'analyse sur les résultats de l'algorithme prédictif et description de l’infrastructures  
nécessaires au bon fonctionnement de l'IA à son cycle de vie.

Yann :

Les technologies choisies répondent à ces critères :

+ Open Source

+ Très bien documentées

+ Très utilisées et la connaissance est très répandue chez les développeurs

**Solution pour la captation pour les flux**[**cita.lu**](http://cita.lu/)**et telraam**

Batch d'alimentation : java à l'aide du framework Spring (<https://spring.io/>)

DB « Time Series »  : timescale de Postgresql <https://www.timescale.com/>

Visualisation : Graphana <https://grafana.com/grafana/>

(timescaledb et graphana tournent dans un container Docker)

Livrables attendus

+ Code sources dans un repo Git  (java + sql)

+ Documentation de mise en oeuvre sur un poste local

+ Deux exemples de dashboard Graphana : 1 cita + 1 telraam

**Analyse / prédiction**

Outils : python à l'aide des librairies <https://scikit-learn.org/stable/> <https://www.tensorflow.org/> et <https://keras.io/>

Code sources dans un repo Git

Documentation de mise en oeuvre

Rapport d'analyse pour la mise en oeuvre d'un algorithme prédictif :

+ Quelles sont les données nécessaires ?

+ Quelles prédictions ( flow, vitesse ) , quel horizon de temps?

+ Décrire les besoins pour mettre cet algorithme en production