

MONAEROPORT.FR

OUTIL DE DATA-VISUALISATION DES INFORMATIONS RELATIVES AU TRANSPORT AERIEN,
CONÇU ET REALISE PAR SIMON HEGE

RAPPORT DE PHASE 1

DEMARCHE UTILISEE

L'application interactive MonAeroport.fr a été développée dans le but d'offrir un accès facilité aux informations relatives au transport aérien, pour un aéroport donné. Il s'agit de fournir des représentations aisément compréhensibles pour les différents internautes visitant le site, ainsi qu'une navigation interactive. Les différentes vues doivent permettre de facilement comprendre les enjeux de l'aéroport qu'il appelle « mon aéroport ».

D'un point de vue technique, l'objectif a également été de mettre en œuvre une solution ne nécessitant pas de serveur de contenu dynamique (python, php, ...) afin de simplifier la mise en production et la maintenance applicative.

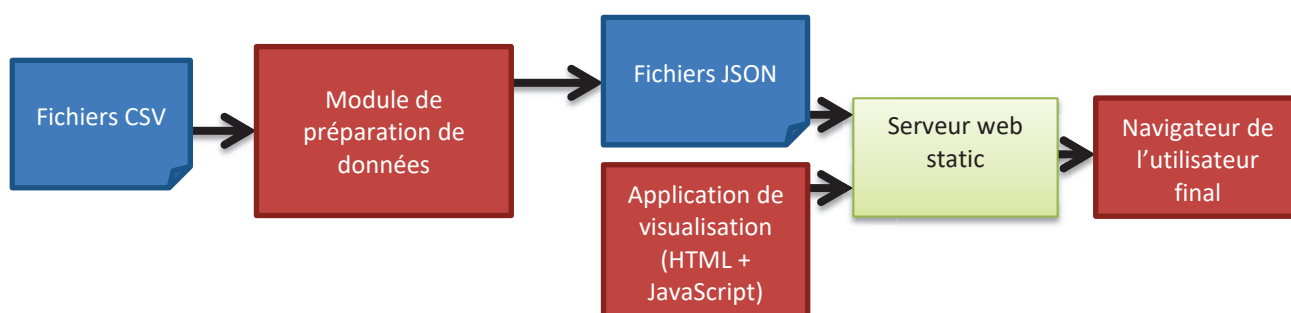
Il est à noter que l'ensemble des fonctionnalités prévues initialement n'ont pu être réalisées dans cette phase 1, mais pourront ultérieurement, en particulier lors de la prochaine phase.

PRINCIPES DE LA SOLUTION

La solution est réalisée au travers d'une exécution en deux étapes.

La première est une préparation des données hors ligne. Elle consiste à agréger les données fournies en CSV et les valider. Cette étape se réalise au travers d'un programme en ligne de commande écrit en langage Go¹. Les données sont exportées dans des fichiers JSON. Le temps de traitement requis est de l'ordre de la minute.

La seconde partie de la solution est l'affichage des données transformées. Elle est réalisée au travers d'une application web monopage écrite en JavaScript au travers de la bibliothèque Angular. Trois modules sont disponibles, permettant l'affichage des informations de trafic, de retards et d'émissions. Une page d'accueil comprend une vue cartographique, puis des pages de détails existe pour chacun des aéroports français métropolitain ou d'outre-mer.



La visualisation a été testée sur les navigateurs Chrome, Firefox et Microsoft Edge.

L'ensemble du code source est publié sous licence MIT, et se base sur les dernières versions des différentes bibliothèques open-sources utilisées, en date de début de projet.

¹ <https://www.golang.org>

TRAITEMENTS DES DONNEES

Dans les deux phases, les données sont manipulées au travers d'hyper-cubes OLAP² afin d'avoir toujours une source centrale de données et de construire des vues agrégées à la demande. Ces cubes sont définies par leurs dimensions (par exemple la période temporelle, l'aéroport concerné, ...) ainsi que par leurs champs de valeurs (par exemple le nombre de vols). En effectuant des manipulations élémentaires sur ces cubes de données (extraction, rotation, agrégation, ...), on peut extraire les informations pertinentes à afficher à l'utilisateur.

PHASE DE PRETRAITEMENT (HORS-LIGNE)

Trois hyper-cubes sont principalement utilisés lors de la phase de prétraitement des données.

Un premier hyper-cube contient les informations relatives au trafic des liaisons aéroportuaires. Ses dimensions sont les suivantes : Année, Mois, aéroport de départ, aéroport de destination. Les valeurs associées sont celles issues des fichiers de trafic (nombre de vols et nombre de passagers). L'agrégation des valeurs se fait par somme simple.

Un second hyper-cube contient les informations relatives au retard. Il est construit à partir du premier en y incluant les données disponibles dans les fichiers de retard. Les dimensions du cube sont les mêmes que pour le précédent. Les champs de valeurs sont issus des fichiers de trafic et de retard. Lorsqu'une agrégation de données est requise, les moyennes et les pourcentages sont ainsi calculés en se basant sur les informations de trafic.

Ces deux cubes sont ensuite transformés pour remplacer le couple origine/destination par un triplé aéroport/direction/escale. Cela permet, d'accéder ainsi par une unique dimension aux informations combinées des vols au départ et à l'arrivée.

Le troisième hyper-cube contient les informations relatives aux émissions LTO. Les dimensions et les champs de valeurs correspondent aux informations présentes dans les fichiers correspondant.

Chacun de ces hyper-cubes est ensuite éclaté par aéroport et exporté en fichier JSON, directement compréhensibles par l'appliquatif frontal de visualisation. La taille des fichiers varie ainsi de 1ko à quelques Mo dans le cas de l'aéroport de Paris Charles-de-Gaulle.

Des fichiers généraux (liste et position des aéroports) sont également exportés lors de cette phase.

PHASE DE VISUALISATION (EN-LIGNE)

Lors de la phase de visualisation, l'hyper-cube correspondant à l'aéroport sélectionné et au type d'information souhaité est chargé dynamiquement dans le navigateur. Les cartes, tableaux et graphiques se mettent à jour dynamiquement on fonction des paramètres de recherche. Les agrégations de données se font sur le même principe que lors de la phase précédente (somme, recalcule des pourcentages et moyennes, ...)

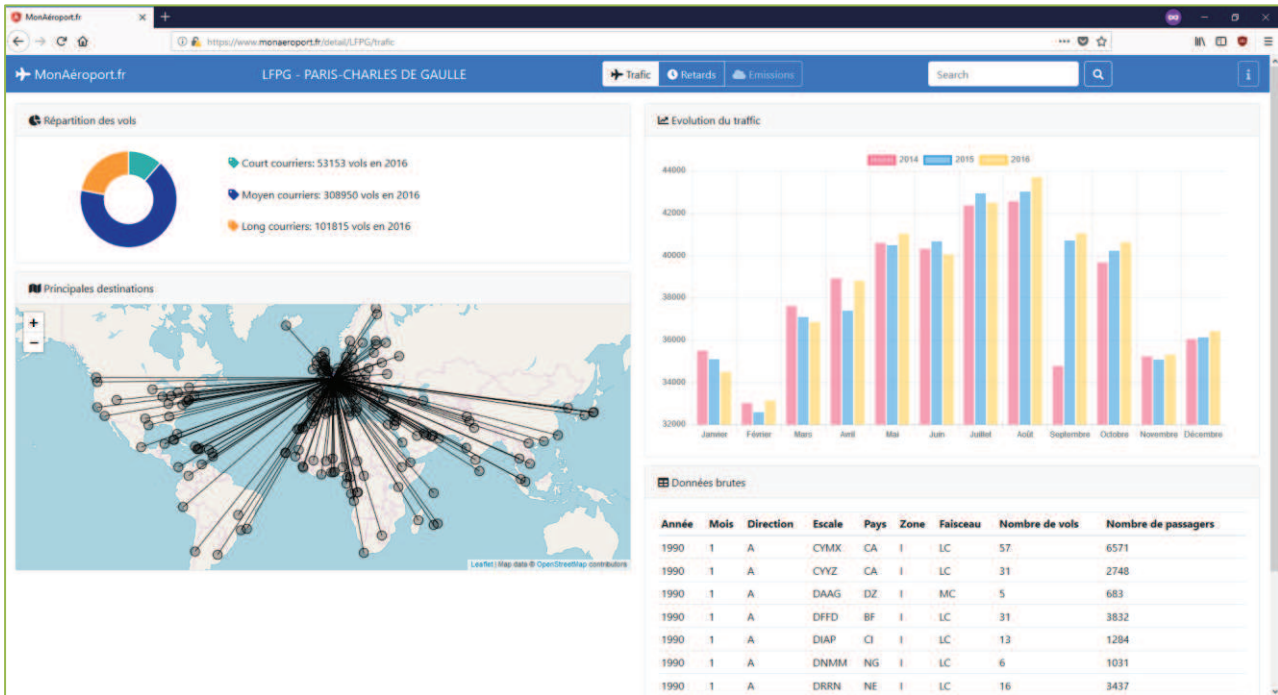
PERSPECTIVES

Grâce à la disponibilité des données brutes au travers d'un moteur d'analyse OLAP dans l'appliquatif de visualisation, il est relativement rapide d'intégrer de nouveaux graphiques ou tableaux à l'application. Il serait possible par exemple d'ajouter la liste des 10 destinations ayant la plus forte moyenne de retard à l'arrivée. On pourrait également imaginer une interactivité laissée à l'utilisateur pour lui permettre de créer ses propres éléments d'analyse (choix des champs affichés : nombre de vols ou de passagers, année de l'analyse, ...).

La page d'accueil pourrait également être complétée par des outils de recherche et de filtrage avancés (aéroports répondant à certains critères).

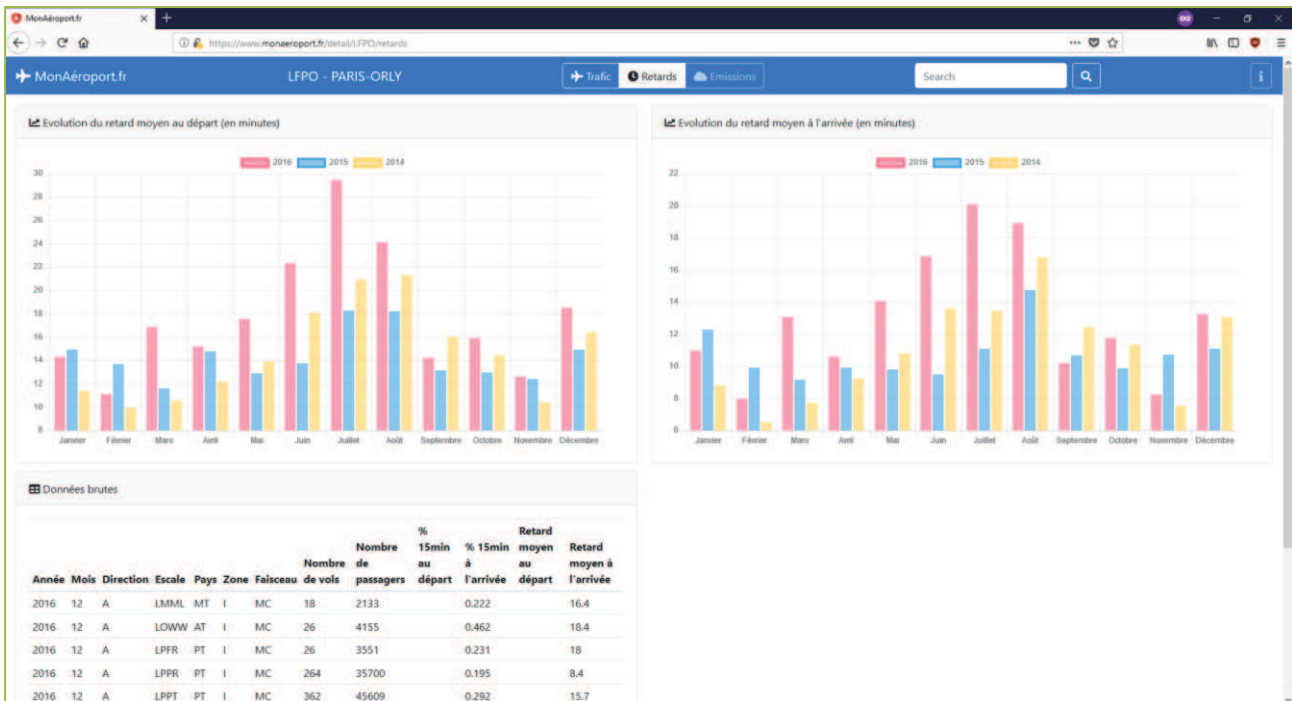
² https://en.wikipedia.org/wiki/OLAP_cube

TRAFFIC DE L'AEROPORT LFPG – PARIS-CHARLES DE GAULLE



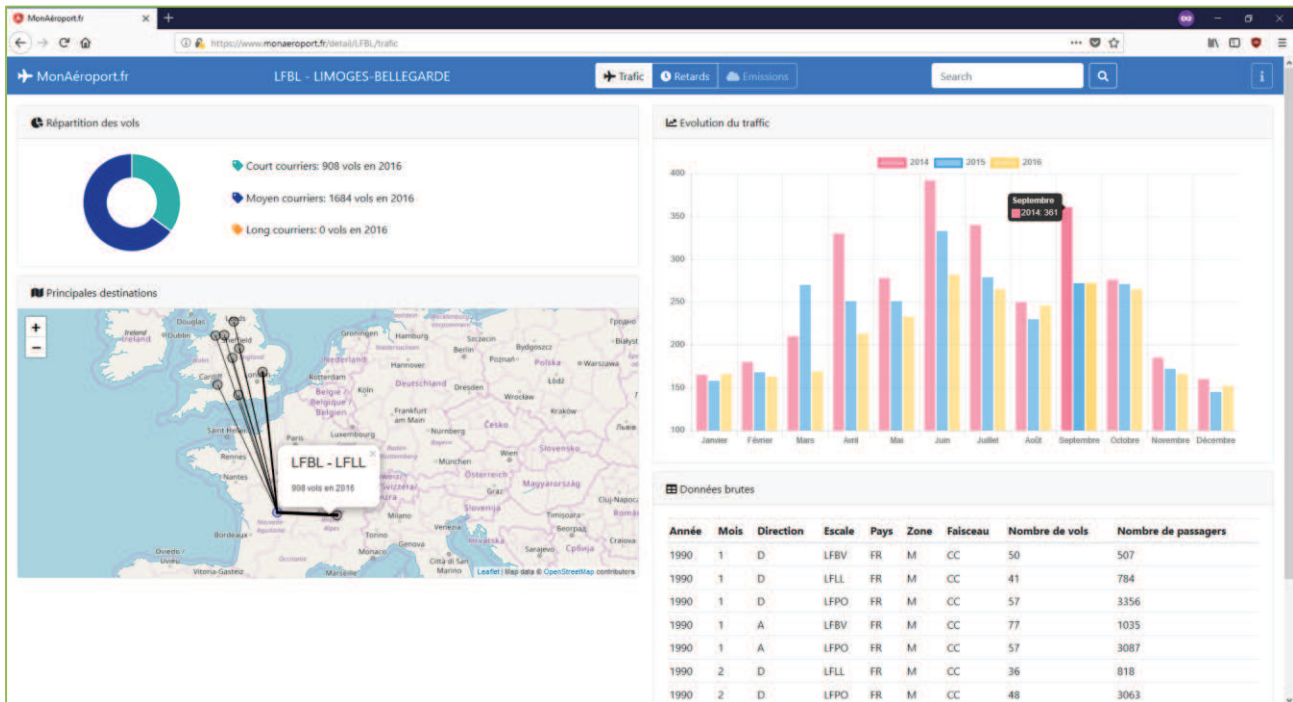
La visualisation cartographique permet de découvrir l'étendue du territoire couvert au départ de l'aéroport de Roissy, de Santiago du Chili à Tokyo en passant par Vancouver. Seule l'Océanie reste hors de portée.

RETARDS A L'AEROPORT LFPO – PARIS ORLY



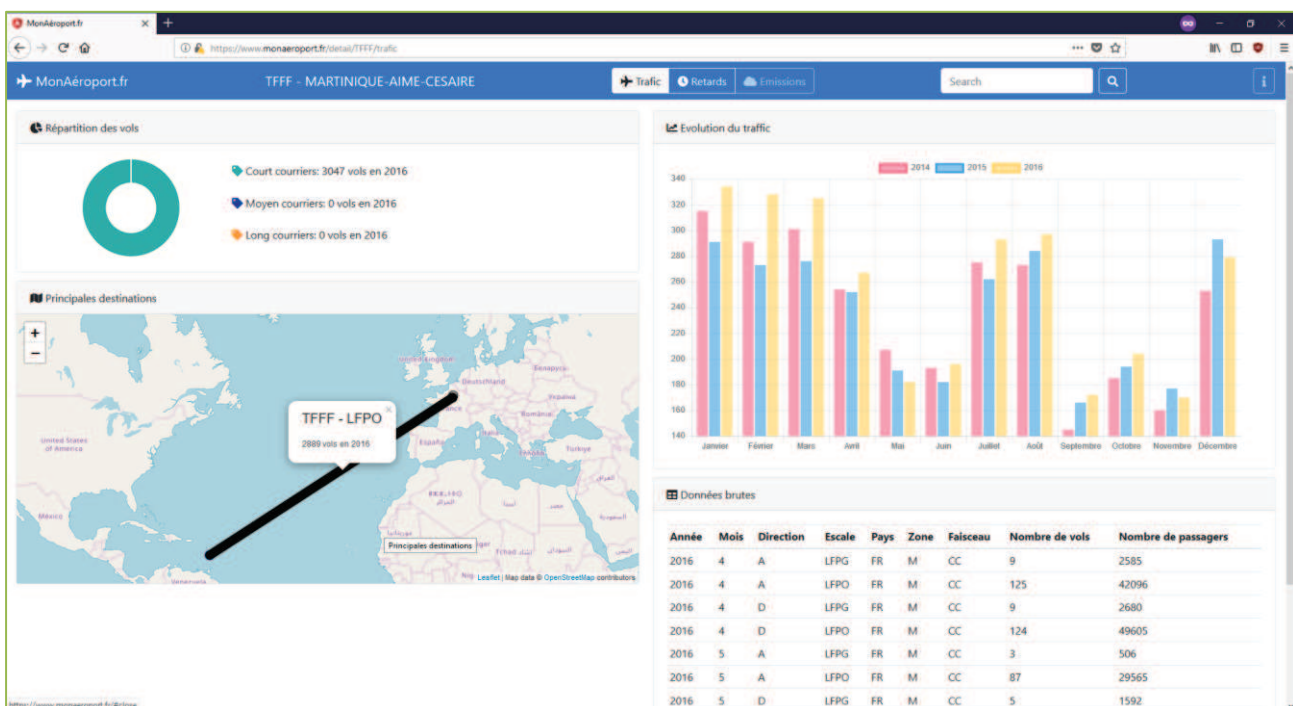
On peut au travers de cette représentation des données une hausse de la durée moyenne des retards au départ (avec un maximum à plus de 28 minutes de moyennes en juillet 2016).

TRAFFIC DE L'AEROPORT LFBL - LIMOGES-BELLEGARDE



La page concernant l'aéroport de Limoges permet de découvrir la répartition géographique particulière du trafic.

TRAFFIC DE L'AEROPORT TFFF - MARTINIQUE-AIME-CESAIRE



La saisonnalité des vols de cet aéroport d'outre-mer est particulièrement marquée (juillet/août puis de décembre à mars). On peut remarquer également que 2889 des 3047 vols long courrier de 2016 étaient à destination d'Orly.

LIEN VERS LA VIDEO DE DEMONSTRATION

Une vidéo de démonstration (au format GIF) est disponible au lien suivant :

<https://drive.google.com/file/d/1U6HB61X8J6G8LDhKCAQbUeOC9Lf8bJBp/view?usp=sharing>