**le projet :le Big Data pour optimiser**

**le client :**

**La Postale (www.laposte.fr)**

La Poste est un grand groupe de services organisé en cinq branches d’activités : La branche Services - Courrier - Colis, La Banque Postale, Geopost (colis-express), Le réseau La Poste (les bureaux de poste), et Le numérique. Le Groupe La Poste réunit plus de 251 000 collaboratrices et collaborateurs animés par des valeurs citoyennes qui sont depuis toujours au cœur de son ADN et de l’identité postale. Avec un réseau de plus de 17 000 points de contact sur tout le territoire, Le Groupe La Poste est le 1er réseau commercial de proximité en France. À l’international, il renforce ses positions, avec une présence dans plus de 40 pays sur cinq continents. Depuis janvier 2011, 100% du chiffre d’affaires du Groupe La Poste est réalisé sur des marchés en totale concurrence. En dix ans, Le Groupe La Poste s’est modernisé en profondeur en intégrant les dimensions numérique et écologique sur un marché concurrentiel.

Le Groupe affirme aujourd’hui son ambition de devenir le leader de la distribution de services de proximité, la banque citoyenne de référence et un acteur international du colis-express.

**La Banque Postale, c’est :**

- 24,1milliards d’euros de CA

- 22,4% à l'international

-

C*hiffres 2017*

**La date du projet : 27/10/2015 au 31/05/2016**

**L’adresse : Siège social du Groupe La Poste Principal 9** rue du colonel Avia, Paris, 75015

**Le projet :**

La Poste a consolidé les milliards d’événements issus de ses plates-formes de tri du courrier dans Hadoop pour les analyser dans Tableau. Objectif : améliorer l’efficacité opérationnelle de cet outil industriel.

**Le contexte**

Utiliser la donnée pour induire un changement des mentalités, Un usage massivement répandu au sein de la société, Des usines à trier le courrier, qui gèrent chaque année 12 milliards de lettres et 300 millions de colis, générant 30 milliards de traces, des données issus soit du suivi des objets de bout en bout, soit de leur traitement sur des robots de tri.

Les données n’étaient pas valorisées à part au sein d’études ponctuelles d'où la mise sur pied d’un cluster Hadoop appelé « Butia » pour digérer toutes ces données, et leur mise à disposition dans la dataviz de Tableau, pour une valorisation à J+1.

Des indicateurs de performance envoyés à chaque plate-forme industrielle du courrier (PIC), au nombre de 80 en France,Ce qui permet par exemple de visualiser la part de trafic local, permettant d’adapter le plan de transport, ou le taux de lettres vertes traitées en prioritaire, soit avec une qualité de service supérieure à l'engagement de la poste.

**Le positionnement de la solution :**

* Réduire les coûts
* Optimisation de la tournée des facteurs.
* Satisfaction de la Clientèle : facilité et rapidité des services proposés par la poste tel que le service de sont aussi achetées tel que service Montimbrenligne qui permet d’imprimer des timbres depuis chez soi
* L’améliorer les services existants ou d’en créer de nouveaux

**Les parties prenantes et utilisateurs :**

**L'utilisateur final:** la Banque Postale

**L’Équipe : 05 Collaborateur (Chef de projet, AMOA,Expert Métier,2 Développeurs)**

**Le responsable Mr**: Laurent Dabbagh, directeur de l’innovation et du pilotage opérationnel de la performance au sein de la branche Courrier et colis de La Poste

**L'architecture de projet :**

Pour les plates-formes de tri du courrier, l’analyse des tableaux de bord peut déboucher sur de vrais gains d’efficacité. Par exemple en mettant en évidence les facteurs empêchant les courriers d’être correctement triés en fonction des tournées des facteurs. Ou en faisant ressortir une baisse du taux de reconnaissance de la **reconnaissance optique de caractère**, signe d’une obsolescence IT grandissante.

Grâce à cette analyse des données, on met le doigt sur les bons indicateurs, ce qui évite les débats byzantins sur le diagnostic d’un problème, pour l’instant, seuls les utilisateurs du siège peuvent manipuler les données dans Tableau, les opérationnels dans les territoires recevant des tableaux de bord statiques, les opérationnels pourraient s’éparpiller, se perdre dans des outils permettant de manipuler les données dans tous les sens.

Au-delà des plates-formes de tri, Tableau est également exploité pour piloter certaines nouvelles activités du groupe. Car si l’activité Courrier pèse environ la moitié du chiffre d’affaires de La Poste (23 milliards d’euros à l’année), le recul du volume de courriers traités pousse la société nationale à développer des relais de croissance, en s’appuyant sur son maillage local et son réseau de facteurs. C’est par exemple le cas de Recygo, un service de recyclage des déchets de bureau.

Sur cette activité (qui pèse moins de 10 millions d’euros en 2016, NDLR), nous disposions d’indicateurs globaux, mais sans compréhension fine des leviers d’efficacité à notre portée. Nous avons exploité Tableau pour analyser en détail l’évolution du chiffre d’affaires et redéfinir les segments client afin d’y ajuster nos offres

**L'architecture Technique du projet :**

340 K€ sont investis en serveurs pour constituer un Framework Hadoop. Ce dernier est associé au requêteur Hive pour récupérer les données travaillées sur la base d’un historique de 18 mois. Cinq licences desktop de Tableau Software, un outil de business intelligence qui met l’accent sur la représentation des données (dataviz)

**Les Taches :**

**La** **reconnaissance optique de caractère** : le tri postal doit utiliser des techniques automatisées :

* Crée un job Spark qui va s’occuper de la lecture et la reconnaissance des caractères à partir d’un fichier OCR,

Dans le fichier OCR se compose de nom, prénom et adresse de l’expéditeur et le destinataire

XML

XML

XML

KAFKA

(en temps réel )

SPARK STREAMING

ETL

**HDFS HDFS**

/

Elasticsearch

Kibana

Hive

TABLEAU

Elastic search utilise les données sous forme d(indice)

Système décisionnel

Système opérationnel

* **Le suivi de courrier en temps réels :** ce suivie se fait par le bais des codes-barres qui vont être scanner dans le chaque bureau et aussi dans le centre de distribution qui va nous aider à détecter les anomalies avant quel arrive à destinataire (courriers au mauvaise adresse)

Bureau1

Sud

Bureau2

Centre de distribution ile de France

Lille ex

nord

Bureau3

* **L’optimisation de la tournée de facteur :**

**Méthode de Travail :**

**AGIL**

# https://www.silicon.fr/wp-content/uploads/2016/02/LaPoste_BigData.jpg

<https://www.pourlascience.fr/sd/informatique/la-reconnaissance-automatique-de-lecriture-4509.php>

<https://www.silicon.fr/poste-big-data-optimiser-tri-courrier-168261.html?inf_by=5bec76dd671db865178b45a2>

<https://www.laposte.fr/entreprise/produits-et-services/sna-normalisation-des-adresses>











