

Modèle de rapport de traitement des données

Version	Auteur	Description	Date
V _I	Loïc Stéphane BAMENOU	<i>Rapport synthèse des différentes manipulations réalisées sur la base des clients (CRM), afin de fournir un fichier de travail après les sanctions de CNIL au service de la performance de DEV'IMMEDIAT</i>	27.12.2023

Introduction

Ce rapport synthèse retrace l'ensemble des traitements qui doivent être apporté aux fichiers sources de gestion de la clientèle (CRM) pour respecter les normes du RGPD. Dans le but de lever les sanctions de la CNIL, nous avons décidé de mettre pratiques l'ensemble de recommandations déjà souhaitées pour la continuité des activités de la société.

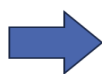
Pour ce traitement, nous allons procéder dans un premier temps à un état des lieux des fichiers retraçant des informations personnelles sensibles ou non. Puis dans un second temps, nous mettrons les traitements à réaliser pour respecter anonymisation des DCP. Ce traitement permettra ainsi aux services de performances de commercialisation de ne pas rester pénalisé par les sanctions de l'agence de contrôle des données informatiques.

Contenu du rapport

I. Etat des lieux d'une extraction limitée des données du logiciel de gestion des clients (CRM)

Pour nos travaux, nous utiliserons des données de nos clients les assurés ayant des dossiers complets datant de l'année 2022. Ces données ont été obtenues avec la requête ci-dessous écrite dans SQLite studio :

```
SELECT *  
FROM base_client  
WHERE etat_dossier = 'complet'  
AND LOWER (date_demande) LIKE '2022%';
```



Soit **1158** clients contrats ou pré-contractualisation au cours de l'année 2022 dont la majorité proviennent des hommes (51 %).

Voici un tableau récapitulatif du glossaire des variables et du devenir de celles-ci dans le cadre de la mise en place des nouvelles réglementations du RGPD. Ce tri a permis de ne garder que les informations jugées indispensables dans le cadre de notre activité.

Tableau 1 : Synthèse de l'état des lieux des données brutes mise à notre disposition

Variables de la table CRM_Complet_2022	Types de données	Description	Typologie de la donnée au sens du RGPD	Devenir des variables
metier	Texte	Profession de l'assuré(e) dans la vie courante	DCP Courante	A supprimer
employeur	Texte	Entreprise dans laquelle l'assuré(e) travaillée	DCP Courante	A traiter
num_ss	Alphanumérique	Numéro de sécurité sociale de l'assuré(e)	DCP Très sensible	A supprimer
groupe_sanguin	Texte	Groupe sanguin de l'assuré(e)	DCP sensible	Facultative
id_site_web	Texte	Pseudonyme électronique de l'assuré(e)	DCP courante	A supprimer
nom	Texte	Nom et prénom de l'assuré(e)	DCP sensible	A supprimer
sexe	Texte	Sexe de l'assuré(e)	DNCP	A garder
email	Texte	Adresse électronique de l'assuré(e)	DCP sensible	A supprimer
date_naissance	Alphanumérique	Date de naissance de l'assuré(e)	DNCP	A traiter
id_client	Alphanumérique	Identifiant secret lié au code du contrat de l'assuré(e)	DNCP	A revoir/Facultative
enfant_conduite_accompagne	Numérique	Nombre d'enfants de l'assuré(e) en conduite accompagnée	DNCP	A traiter
nombre_enfants	Numérique	Nombre d'enfants de l'assuré(e)	DNCP	A traiter
revenus	Numérique	Revenus de l'assuré(e)	DNCP	A traiter
valeur_residence_prin	Numérique	Valeur de résidence principale de l'assuré(e)	DNCP	A traiter
formation	Texte	Niveau d'instruction de l'assuré(e)	DNCP	A garder/Traiter
usage_vehicule	Texte	Usage véhicule du l'assuré(e)	DNCP	A garder
type_vehicule	Texte	Modèle de véhicule de l'assuré(e)	DNCP	A garder
est_rouge	Binaire/Booléan	Est Rouge de l'assuré(e)	DNCP	Facultative
points_perdus	Numérique	Nombre de points perdus de l'assuré(e) lors des infractions commissent	DNCP	A traiter
age_vehicule	Numérique	Durée de vie du véhicule de l'assuré(e)	DNCP	A traiter
type_conduite	Binaire/Booléan	Zone de conduite de l'assuré(e)	DNCP	A traiter/Corriger
date_demande	Alphanumérique	Date de réalisation de la demande par de l'assuré(e)	DNCP	D revoir/Facultative
etat_dossier	Texte	Etat de la Dossier de l'assuré(e)	DNCP	A garder
formule	Texte	Formule de la prestation de l'assuré(e)	DNCP	A garder
tarif_devis	Numérique	Devis du tarif de l'assurance proposé à l'assuré(e)	DNCP	A traiter
adresse	Alphanumérique	Adresse de l'assuré(e)	DCP Très sensible	A traiter
lat	Numérique	Latitude de l'assuré(e)	DCP Très sensible	A supprimer
lon	Numérique	Longitude de l'assuré(e)	DCP Très sensible	A supprimer

II. Traitements accordés aux variables retenues pour le service performance

A la suite de ce tri, nous avons débuté l'importation du fichier dans Power Query en lançant l'éditeur de cette extension d'Excel qui gère les données de masses.

- Utilisation du connecteur « Texte/csv » pour la récupération des données puis charger la table de données pour procéder à l'apurement de cette dernière (Diapo n°5);
- Apurement du fichier (*suppression des colonnes inutilisables*) dans le but de ne garder que les données brutes dont nous avons besoin pour la suite des événements (*revoir le devenir des variables Tableau n°1*) ;
- Traitement ou transformation de la donnée brute restante en fichier exploitable qui reste les normes et règles du RGPD. A ce niveau, nous avons procédé à trois (03) types de traitements à savoir :
 - Correction des variables : rendre les modalités d'une variable plus simple et compréhensible (*présence d'enfant, type de conduite, points perdus sur le permis et la date de la demande*) ;

- Création d'intervalle pour les variables sensibles : ce type de transformation consiste à empêcher la reconnaissance des individus par rapport à un ou plusieurs critères, c'est-à-dire la randomisation (*Tranche de revenus, Tranche d'âge du véhicule, Tranche de valeur de la résidence principale et intervalle des tarifs de devis*) ;
- Création de nouvelles variables : ceux des variables qui selon des critères ou des transformations particulières donne lieu à d'autres informations (*Typologie d'employeur, Tranche d'âge, Index personnalisé et Code département*) ;

Le détail sur les différentes fonctionnalités et les opérations utilisées pour faire ressortir ses nouvelles variables sont présent dans l'interface de Power Query et glossaire des transformations qui est associé (*Voir feuille « Glossaire » du classeur Excel*).

- Vérification ou contrôle des nouvelles variables quelles respectent les règles de l'anonymisation à l'aide des tableaux croisés dynamiques (TCD) ;
- Cloisonnement des données sur un serveur avec un mot de passe complexe dont seulement les membres du service performance peuvent y avoir accès ;
- Diffuser les résultats de cette nouvelle réflexion et solliciter la mise en place d'atelier d'amélioration pour une continuité vers les nouvelles aspirations de l'entreprise.

Conclusion

Voilà en quelques lignes un résumé rapide de l'ensemble des traitements ayant été appliqué sur une petite extraction des données de DEV'IMMEDIAT pour s'assurer de l'application des nouvelles règles du RGPD. Elle met en application différents pré-requis pour s'assurer d'une anonymisation efficace (randomisation et généralisation).

Cette nouvelle orientation permettra ainsi de construire un guide qui servira de registre de traitement des données dans l'atteinte des objectifs visés par l'entreprise dans ces activités.