



Modèle de rapport de traitement des données

Titre du projet	Processus de protection des données
------------------------	-------------------------------------

Version	Auteur	Description	Date
V1	Wescott	Rapport	19/11/2024

Introduction

Dev'Immédiat est une Entreprise spécialisée en courtage d'assurances automobiles, dont le métier consiste à faire de la prospection commerciale et à réaliser une marge sur la vente de contrats. Cependant, elle rencontre des problèmes concernant la protection de ses données. Elle a été sanctionnée par la CNIL pour avoir conservé les données personnelles d'un client qui n'était plus assuré depuis plus d'un an.

Les données dans leur CRM n'ont pas été supprimées depuis sept ans. Il n'y a aucun processus défini concernant la durée de conservation des données dans le CRM, ni personne en charge de gérer le cheminement des données. De plus, il n'existe aucune documentation sur la manière de gérer ces données.

Clara Daucour, la directrice, souhaite qu'on les aide à résoudre cette situation vis-à-vis de CNIL, à anonymiser correctement leurs données, à continuer à exercer leur activité dans le respect des règles, et à mettre en place des processus internes concernant la collecte, le traitement et la documentation des données clients.



Définition:

L'anonymisation des données consiste à rendre les informations personnelles non identifiable, assurant ainsi la protection de la vie privée. Dans le secteur des assurances automobiles, elle est essentielle pour se conformer aux régulations, sécuriser les données et renforcer la confiance des clients. Elle permet aussi d'analyser les données tout en minimisant les risques liés aux violations de la confidentialité.



Contenu du rapport :

I. Les difficultés liées aux données actuelles :

- L'identification du type de données (Il existe 3 grandes catégories : données à caractère personnel, les données publiques, les données privées)

Le type de données que la société Dev'Immédiat possède dans son CRM est constitué de données à caractère personnel. Même si le consentement n'est pas obligatoire lorsque les données sont collectées pour l'exécution d'un contrat, il est cependant conseillé par la CNIL ([article 13](#) du RGPD) de les informer et de leur fournir le maximum d'informations sur le traitement de leurs données.

Exemple de donnée personnel :

metier	employeur	num_ss	groupe_sanguin	adresse	formation
entraîneur de chevaux	Besnard Michaud SARL	239023523804378	A+	6 Rue Anatole France 46100 Figeac	PhD
orthoptiste	Costa Prévost S.A.R.L.	296106544039387	B-	22 Place Saint-Pierre 56130 Marzan	z_High Scho
technicien en optique de précision	Neveu Courtois S.A.	267079402854977	AB+	75 Rue de Lyon 42130 Boën-sur-Lignon	Bachelors

- Le traitement de donnée :

Une fois que vous avez trouvé les données de base (données majoritaires) qui définissent le cœur de votre entreprise (type de véhicule, usage du véhicule, formule choisie etc.).

Il faut hiérarchiser vos données et créer un dictionnaire de données, c'est-à-dire un répertoire qui va définir ces différentes catégories et les lister concrètement. C'est une boussole qui permettra à tous de comprendre votre classification interne et ainsi de mieux se repérer dans vos données.

- La durée de vie des données :

La durée de vie de vos données fait partie du processus de tri et de classification nécessaire à une bonne gouvernance. Il existe plusieurs phases pour mettre fin à la vie de vos données : elles commencent dans la base active, puis passent par une phase d'archivage intermédiaire, suivie de l'archivage définitif, et enfin de la suppression.

La phase de suppression ne doit pas nécessairement être brutale ; vous pouvez passer par les processus d'anonymisation et de pseudonymisation de vos données.

La pseudonymisation est l'une des mesures recommandées par le RGPD pour limiter les risques liés au traitement des données personnelles.

Bien qu'il n'y ait pas de durée fixée pour la conservation des données personnelles par le RGPD, il est cependant recommandé de fixer des limites en fonction de l'atteinte de vos objectifs.

- La personne en charge de la confidentialité des données :

Dans une organisation, différents interlocuteurs peuvent être responsables ou impliqués dans la confidentialité et la protection des données.

Notamment :

- **DPO** pour les questions de conformité et de droits des individus.
- **RSSI (Responsable de la sécurité des systèmes d'information)** pour les questions de sécurité des données.
- **Équipe juridique** pour des conseils sur les aspects légaux.
- **Équipe technique/administrateurs** de bases de données pour les aspects techniques de sécurité.



- **Direction générale** pour des questions stratégiques.
- **RH** pour des questions concernant les données des employés.

Ces différentes personnes peuvent être en charge de la documentation des processus, du suivi et de la mise à jour des procédures, ainsi que de la formation et de la sensibilisation de leurs collègues.

II. Nettoyage des données

- Récupération et sélection des données :

Étape 1 : L'extraction des données et le filtrage des données nécessaires pour être conformes au RGPD.

Dans SQLite, j'ai pu filtrer sur les données selon la demande du client (l'année et les dossier complet).

Exemple de Requête SQL :

The screenshot shows a SQLite database interface with a query window and a results table. The query is:

```
1 select*
2 from base_client
3 where etat_dossier = "complet" and date_demande >= 2022;
```

The results table displays 13 rows of client data. The columns include: id, sexe, email, date_naiss, id_client, enfant_o, nombre_revenus, valeur_re, formation, usage_veh, type_veh, est_rout, points_p, age_veh, type_conduite, date_demande, etat_dossier, and forme.

	id	sexe	email	date_naiss	id_client	enfant_o	nombre_revenus	valeur_re	formation	usage_veh	type_veh	est_rout	points_p	age_veh	type_conduite	date_demande	etat_dossier	forme	
1		F	thibaultbourdon@club-internet.fr	1999-09-25	871024631	0	1	123501	0		z_High School	Private	Minivan	no	2	1.0	Highly Urban/ Urban	2022-02-25 16:13:38	complet dev_s
2		F	pascalabellie@france.com	1932-01-28	7945239	1	1	50815	0		z_High School	Private	Pickup	no	0	4.0	z_Highly Rural/ Rural	2022-01-24 08:17:21	complet dev_s
3		F	christinedias@hotmail.fr	1978-09-15	713692930	0	2	21204	0		z_High School	Private	z_SUV	no	3	5.0	Highly Urban/ Urban	2022-04-20 03:46:00	complet dev_s
4		M	monnedouard@wanadoo.fr	2021-02-17	58888550	0	0	0	106859		z_High School	Private	z_SUV	no	2	12.0	Highly Urban/ Urban	2022-03-23 01:32:40	complet dev_s
5		M	hibonnet@free.fr	1956-06-14	69165559	0	0	102161	207519	Bachelors	Commercial	Panel Truck	yes	1	10.0	Highly Urban/ Urban	2022-03-23 01:32:40	complet dev_s	
6	yes	F	charlespoly@tiscali.fr	1992-04-11	573751437	0	1	79606	225962	Bachelors	Private	z_SUV	no	2	15.0	z_Highly Rural/ Rural	2022-03-23 01:32:40	complet dev_s	
7	Nouvel	F	chantal5@bouygies.fr	1941-03-18	294816298	2	2	130540	344195	Bachelors	Commercial	Minivan	no	0	9.0	z_Highly Rural/ Rural	2022-03-23 01:32:40	complet dev_s	
8		M	ygauffier@live.com	1956-02-06	686791659	0	0	51278	230340	Bachelors	Private	z_SUV	no	1	10.0	Highly Urban/ Urban	2022-03-23 01:32:40	complet dev_s	
9	u	M	wgermain@live.com	1956-02-18	866093378	0	0	115044	311980	PhD	Private	Minivan	no	0	14.0	Highly Urban/ Urban	2022-01-29 10:17:49	complet dev_s	
10		F	dbonnet@noos.fr	1974-03-07	401200262	0	0	102161	252896	Bachelors	Commercial	Panel Truck	yes	1	10.0	Highly Urban/ Urban	2022-03-26 07:37:33	complet dev_s	
11		F	agilbert@orange.fr	1967-02-13	926073051	1	2	24090	0		Bachelors	Private	Minivan	no	0	10.0	Highly Urban/ Urban	2022-05-26 05:51:37	complet dev_s
12		M	julien72@tele2.fr	2017-05-14	45830738	0	4	20822	76022	z_High School	Private	z_SUV	no	0	1.0	Highly Urban/ Urban	2022-01-27 09:33:29	complet dev_s	
13	pentier	F	jsagner@wanadoo.fr	1950-12-20	429962959	0	0	74426	0	PhD	Private	z_SUV	no	0	18.0	Highly Urban/ Urban	2022-05-08 13:32:05	complet dev_s	

Étape 2 : Préparation des données dans Power Query (ETL).

Après avoir uploadé les données extraites avec la requête SQL (au forma CSV).

Dans Power Query j'ai effectué les modifications suivantes :

- Renommer le fichier et l'enregistrer. Supprimer les données personnelles qui peuvent être utilisées pour identifier les clients (métier, employeur, numéro de sécurité sociale, groupe_sanguin, nom, email, date de naissance, id_client, formation, adresse).



- Remplacer les sigles qui peuvent être non incompréhensibles pour certains ou dans un graphique (exemple : Féminin et Masculin).
- Dupliquer la colonne date/heure afin de les séparer pour avoir deux colonnes distinctes. Pour le filtrage.
- Modification des types pour les différentes colonnes (Entier, texte, date, heure, devis, Vrai/Faux etc.). Certain type faisait apparaître des erreurs dans les données (remplacer les points par des virgules afin de passer la colonne en nombre entier.
- Renommer et déplacer certaines colonnes.
- Ajouter une colonne personnalisée pour la génération de numéros Aléatoires afin de remplacer les id-clients provenant du CRM du client.

ÉTAPES APPLIQUÉES

Source	
Type modifié	
Données personnel supprimées	
Valeur remplacée_F par Féminin	✱
Valeur remplacée1-M par Masculin	✱
Colonnes supprimées_date_naissance	
Duplication de la colonne-séparé heure/date	
Type modifié1_heures	
Colonnes Heure_renommées	
Colonnes permutées_colonne_heure	
Type modifié_colonne_date	
Colonnes permutées_tous les colonnes	
Valeur remplacée_remplacer les points par un virgule	✱
Type modifié_tarif_devis	
Valeur remplacée_remplacer les points par des virg...	✱
Type modifié1_age_vehicule et nombre_enfants	
Lignes filtrées	
Index ajouté	✱
Personnalisée ajoutée-Random ID	✱
Index ajouté1	✱
Colonnes supprimées	
Colonne multipliée	✱
Type modifié1	
Valeur remplacée	✱
Lignes filtrées1-Remplacer les NULL par 0	
✱ Colonne conditionnelle ajoutée-crétion des interv...	✱

- Ajout d'une colonne conditionnelle pour crée les intervalles des revenus. Pour avoir des données non personnelles, tout en permettant une analyse des tendances de revenus par tranches.

Ajouter une colonne conditionnelle

Ajoutez une colonne conditionnelle calculée en fonction des autres colonnes ou valeurs.

Nouveau nom de colonne :

Personnalisé

Si	Nom de la colonne	Opérateur	Valeur	Alors	Sortie
Si	revenus	est inférieur ou é...	60000	Alors	[0-60000]
Autre...	revenus	est inférieur ou é...	60001	Alors	[60001-120000]
Autre...	revenus	est inférieur ou é...	120001	Alors	[120001-180000]
Autre...	revenus	est inférieur ou é...	180001	Alors	[180001-240000]
Autre...	revenus	est inférieur ou é...	240001	Alors	[240001-300000]
Autre...	revenus	est inférieur ou é...	300001	Alors	[300001-360000]

Ajouter une clause

Autre :

OK Annuler

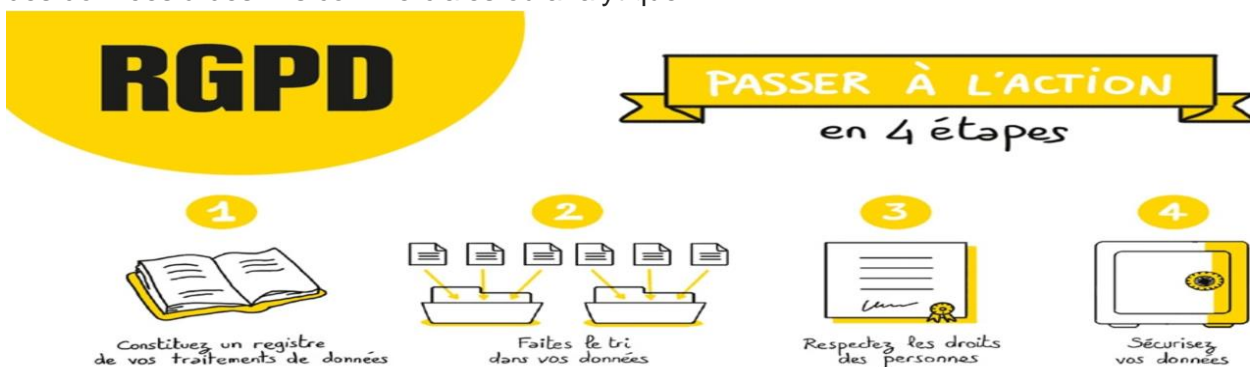
III. Solution (résultat obtenu) / Préconisations

- Résultat

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	
1	random.ID	date_demande	heure_demande	etat_dossier	formule	tarif_devis	Revenue.Interval	type_vehicule	usage_vehicule	type_conduite	age_vehicule	points_perdus	sexe	est_rouge	valeur_residence	prir_enfant	conduite_accompagn	nombre_enfants
2	282167	25/02/2022	16:13:38	complet	dev_integral	323,39	[180001-240000]	Sports Car	Commercial	Highly Urban/ Urban	7	0	Féminin	no	0	0	0	VRAI
3	226650	24/05/2022	21:07:34	complet	dev_express	257,56	[0-60000]	Minivan	Private	Highly Urban/ Urban	1	2	Féminin	no	0	1	0	VRAI
4	493204	24/01/2022	08:17:21	complet	dev_express	250,68	[0-60000]	Pickup	Private	z_Highly Rural/ Rura	4	0	Féminin	no	0	0	0	VRAI
5	571712	20/04/2022	03:46:00	complet	dev_integral	300,3	[0-60000]	z_SUV	Private	Highly Urban/ Urban	5	3	Féminin	no	106859	0	0	FAUX
6	47202	23/03/2022	01:32:40	complet	dev_vip	648,13	[0-60000]	Panel Truck	Commercial	Highly Urban/ Urban	12	2	Masculin	no	207519	0	0	FAUX
7	590326	23/03/2022	22:20:50	complet	dev_premium	459,83	[120001-180000]	z_SUV	Private	z_Highly Rural/ Rura	15	2	Féminin	no	225962	0	0	VRAI
8	598615	13/03/2022	01:10:59	complet	dev_express	261,28	[180001-240000]	Minivan	Commercial	z_Highly Rural/ Rura	9	0	Féminin	no	344195	2	0	VRAI
9	337677	09/03/2022	03:43:53	complet	dev_premium	343,17	[0-60000]	z_SUV	Private	Highly Urban/ Urban	10	1	Masculin	no	230340	0	0	FAUX
10	108241	29/01/2022	10:17:49	complet	dev_premium	389,24	[120001-180000]	Minivan	Private	Highly Urban/ Urban	14	0	Masculin	no	311980	0	0	FAUX
11	370782	26/03/2022	07:37:33	complet	dev_integral	310,66	[0-60000]	Panel Truck	Commercial	Highly Urban/ Urban	10	1	Féminin	yes	252896	0	0	FAUX
12	101625	26/05/2022	05:51:37	complet	dev_express	261,6	[0-60000]	Minivan	Private	Highly Urban/ Urban	10	0	Féminin	no	0	1	0	VRAI
13	569196	27/01/2022	09:33:29	complet	dev_premium	380,99	[0-60000]	z_SUV	Private	Highly Urban/ Urban	1	0	Masculin	no	76022	0	0	VRAI
14	84345	08/05/2022	13:32:05	complet	dev_express	251,44	[120001-180000]	z_SUV	Private	Highly Urban/ Urban	18	0	Féminin	no	0	0	0	FAUX
15	204216	02/05/2022	08:44:46	complet	dev_integral	307,66	[0-60000]	Van	Commercial	z_Highly Rural/ Rura	6	5	Masculin	no	0	0	0	FAUX
16	506517	15/04/2022	13:09:59	complet	dev_integral	302,89	[0-60000]	Minivan	Private	Highly Urban/ Urban	1	0	Masculin	no	202967	0	0	VRAI
17	112581	11/05/2022	02:24:06	complet	dev_express	279,3	[0-60000]	Minivan	Private	Highly Urban/ Urban	13	0	Masculin	no	173196	0	0	FAUX

- Préconisation

Voici d'autres préconisations pour l'anonymisation des données, afin de continuer à garantir la sécurité et la confidentialité des informations sensibles des clients tout en permettant une exploitation des données à des fins commerciales ou analytiques.





- Vous identifiez et listez de la manière la plus exhaustive possible toutes les données que vous avez dans votre entreprise et d'où viennent-ils (afin de savoir quelle source va poser problème)
- Remplacer les données spécifiques par des catégories plus larges pour éviter toute identification précise (exemple : vous pouvez ne pas conserver l'adresse exacte d'un client, mais vous pouvez le remplacer par une zone géographique).
- Supprimer les informations qui ne sont pas indispensables à l'objectif de traitement ou d'analyse (exemple : Si l'agence n'a pas besoin de la date de naissance exacte d'un client pour calculer le montant de sa prime, on peut anonymiser cette information en la regroupant sous une tranche d'âge (par exemple, "25-35 ans"))
- Anonymiser les données dès le moment où elles sont collectées, pour éviter qu'elles ne soient stockées sous une forme identifiable.
- Mettez en place un répertoire afin de documenter l'ensemble des données retenus (exemple : un dictionnaire de donnée).
- Auditer et contrôler en internes régulièrement pour s'assurer que les pratique d'anonymisation sont respectées et que les données sensibles sont bien protégées.
- Crypter les données sensibles de manière à ce qu'elles ne puissent être accessibles qu'avec une clé de décryptage. (**Exemple** : Les données de paiement ou d'autres informations sensibles sur les clients doivent être cryptées, même lorsqu'elles sont pseudonymisées, afin d'ajouter une couche supplémentaire de sécurité).
- Définir une durée de conservation de vos données (vous ne pouvez pas conserver éternellement vos données).
- Mettez en place une charte de gouvernance des données (La charte de gouvernance est le document interne de référence pour tous vos salariés. Il est essentiel car il va résumer l'ensemble des actions à mener, les objectifs poursuivis et les raisons de ces actions, mais aussi la répartition des rôles de chacun).
- Nommez un DPO (Délégué à la protection des données). C'est une personne qui s'occupe de la protection des données dans une entreprise.



Conclusion

Le processus d'anonymisation des données doit être systématique, bien documenté, et conforme aux exigences légales. En suivant ces étapes et en appliquant les préconisations ci-dessus pourra garantir la confidentialité et la sécurité des données de ses clients, tout en respectant les normes de protection des données personnelles.