



Octopize : #avatars #anonymisation sans compromis #innovation de rupture

Problème

Le volume des données personnelles augmente de manière exponentielle et offre un gisement de valorisation pour les acteurs économiques. Néanmoins, à ce jour, le traitement des données personnelles implique souvent de faire un mauvais compromis entre confidentialité et partage de données exploitables.

De façon légitime, le RGPD (Règlement sur la Protection des Données Personnelles) est venu renforcer la confidentialité. Les acteurs peuvent exploiter les données personnelles à condition de mettre en place des mesures pour préserver la vie privée des individus¹.

Cependant, la difficulté à exploiter des données de qualité demeure. L'anonymisation apparaît comme la solution.

État de l'art

Il réside une incompréhension sur ce que l'on considère comme des données « anonymes » versus « pseudonymes » au sens de la CNIL. « *La pseudonymisation consiste à supprimer toutes données directement identifiantes (nom, prénom, etc.) d'un jeu de données, par des données indirectement identifiantes afin de réduire leur sensibilité* ». De nombreux écueils démontrent que les données pseudonymisées ne sont pas une garantie de confidentialité. En effet, les individus peuvent être ré-identifiés par croisement².

Par opposition, l'anonymisation est « *un traitement qui consiste à utiliser un ensemble de techniques de manière à rendre impossible, en pratique, toute identification de la personne par quelque moyen que ce soit et de manière irréversible* ». Les autorités de protection des données européennes (ex G29), ont déterminé les trois critères permettant de s'assurer qu'un jeu de données est véritablement anonyme : individualisation, corrélation et inférence. Elles n'ont identifié aucune méthode en capacité de garantir une protection totale sur les trois critères³. D'autre part, ces méthodes impliquent une perte de qualité des données.

Solution Octopize

Les avatars, évalués avec succès par la CNIL, garantissent la confidentialité, sans faire de compromis sur la qualité des données.

Ce sont des données de synthèse qui permettent de partager sans dévoiler et exploiter tout le potentiel de valeur de la donnée, de façon éthique. Ils permettent d'affirmer : « *qu'il n'y plus aucune justification de faire prendre des risques de réidentification à des individus* » et que nous devons « *réserver les données personnelles pour les usages personnels* ».

Les avatars sont une innovation de rupture : ils garantissent sécurité et reproductibilité. N'étant plus considérés comme des données personnelles, ils sortent du RGPD.

La solution est déjà utilisée dans le secteur de la santé, par l'industrie pharmaceutique, le CHU de Nantes, ou encore dans le cadre du projet européen HAP2⁴. Elle est commercialisée sous forme de logiciel (on premise) ou d'avatars as a service.

Proposition de valeur unique

Les avatars sont la solution pour tous les usages ultérieurs sur les données personnelles pour lesquels il n'y a pas un besoin de ré-identification des individus : analyse, recherche médicale, partage avec des tiers hors UE, open data, ou encore conservation des données sans durée de détention. La méthode est agnostique et se déploie sur toutes les verticales d'activité.

Avec les avatars, exploitons les données au service de tous mais dans le respect de chacun.

¹ Le non-respect expose à un risque de sanctions financières jusqu'à 20 000 000€ ou 4% du CA annuel mondial

² <https://www.nature.com/articles/s41467-019-10933-3>

³ https://www.cnil.fr/sites/default/files/atoms/files/wp216_fr.pdf

⁴ www.hap2-project.com