

# Rapport du projet d'anonymisation de données

## 1. INTRODUCTION

Ce projet a pour objectif d'anonymiser des données de la CNAM. Le client a envoyé un ensemble de dossiers contenant des fichiers .xml avec des données d'individus. Le cahier de charge indique qu'il est nécessaire de modifier ces données sans changer la structure des fichiers .xml. Ce projet est réalisé en Java avec l'IDE IntelliJ. On retrouve dans les grandes lignes du projet des fonctions permettant de générer des noms, prénoms, adresse, RIB etc..., dans les fichiers .xml, mais également des programmes permettant de parcourir l'ensemble des dossiers.

## 2. FONCTIONNEMENT

Dans un premier temps, le programme d'anonymisation est assez simple à lancer. Après avoir téléchargé le projet, Il suffit de lancer le fichier main.java puis des messages indiquant la démarche à suivre apparaissent. Le premier message signale qu'il faut choisir un fichier .properties dans la fenêtre graphique des répertoires. Le second message annonce qu'il faut choisir un dossier avec les fichiers .tgz ou .zip correspondant. Enfin, le dernier message indique que le dossier à choisir est celui qui possède les données modifiées.

Dans un second temps, l'algorithme doit se lancer, il commence par extraire tous les fichiers compressés dans le dossier sélectionné. Cette opération dure environ 5 minutes. Une barre de progression est affichée à l'écran pour suivre l'avancement des modifications des fichiers .xml. Le temps que l'algorithme termine dépend en grande partie du nombre de fichiers .xml fourni et de leurs tailles. Néanmoins, on estime qu'un fichier .xml fourni par la CNAM met au alentour de 3 secondes à être traité.

Dès que l'algorithme a terminé sa tâche, un message s'affiche pour signaler la terminaison. Il suffit alors de récupérer les fichiers .xml dans leurs .zip correspondants.

### 3.Note à l'utilisateur

- a. Lors de la récupération du projet, il est nécessaire pour utiliser pleinement les capacités de l'algorithme d'installer une librairie spécifique pour dézipper les dossiers. Il convient donc de télécharger le .jar commons-compress-1.18.jar au projet. Des tutoriels simples sont disponibles sur internet pour l'implémenter dans le projet selon l'IDE utiliser..
- b. Le traitement de toutes les données fournies par la CNAM dure environ deux heures.
- c. Un fichier maven est donné dans le projet qui installe automatiquement toutes les dépendances nécessaires.
- d. Le programme prend en compte deux types de fichiers: DKPIE\_A,DK\_A.
- e. Attention, on ne peut pas prendre un fichier properties DKPIE\_A avec des données DK\_A, l'inverse est également impossible.