

# Trabalho 1 - Técnicas Anonimização

Javam Machado

Abril 2022

## 1 Objetivo:

- Implementar três técnicas de anonimização de um conjunto de dados.

## 2 Especificação:

- Carregue o conjunto de dados “covid.csv” que lhe será fornecido. Este data set já estará pseudoanonimizado. Você deve implementar um programa para fazer as seguintes operações:
  - **Supressão:** A interface vai permitir a escolha de um ou mais atributos cujos valores serão substituídos por um valor curinga (\*) por exemplo) em todos os registros do dataset;
  - **Generalização:** defina uma árvore de generalização para o atributo “Município” considerando as regiões do Estado do Ceará e uma outra árvore para data de nascimento com 3 níveis (dado original, data sem dia, data sem dia e mês). Forneça uma interface para o usuário escolher a generalização que ele deseja aplicar sobre os dois atributos.
  - **Perturbação:** Defina uma função de perturbação sobre um atributo numérico e sobre um atributo alfanumérico com domínio finito (considere os valores presentes no dataset) e forneça uma interface para o usuário aplicar a perturbação.
- A saída deve ser um novo data set, dito “covid\_publico.csv”. Programe uma interface para visualização do data set gerado.

## 3 Requisitos

- Linguagem: C++, Java, Python
- Meio de entrega: Empacotar os artefatos de software (programa, arquivos de entrada e de saída) e fazer o *upload* pelo classroom.

- Equipes de até 2 pessoas.
- Somente um um aluno por equipe precisa fazer o *upload* pelo classroom.

## 4 Entrega

- *Upload* pelo classroom até as 13:59 do dia 20/04, quarta-feira.
- Demo com arguição em sala de aula na quarta-feira, 20/04, em ordem aleatória. A equipe deve usar seu próprio equipamento que será ligado ao projetor (HDMI) da sala.