

Ivan Marin
Vivo Data Labs
ivan.smarin@telefonica.com
2016-02-18

Gerador de Hashes

O script Python gera hashes usando o algoritmo [SHA-256](#) e faz a codificação destes hashes em [Base64](#).

Algoritmo

O script é capaz de executar as seguintes operações:

- gerar hash em SHA 256 e fazer encoding em Base64
- gerar hash em SHA 256

O algoritmo básico é descrito por

```
para linha em arquivo:  
    hash = SHA-256(linha)  
    hash_base = base64(hash)
```

Execução

O script requer Python versão 2.7. Para executar, em um terminal utilize o comando

```
python generate_hash.py -i test.csv -o out.csv -n cpf -f hashed_cpf -b yes
```

As opções são:

- **-i**: Arquivo de entrada com os dados a serem gerados os hashes
- **-o**: Arquivo de saída, com os dados originais e os dados transformados em hash
- **-f**: Nome do campo do arquivo de entrada que contém os dados a serem gerados hash
- **-f**: Nome do campo para os hashes gerados
- **-b**: **yes** se o hash SHA 256 deve ser encodado em Base64 e **no** se o hash deve ser somente em SHA 256. O padrão é 'no'

O arquivo `test.csv` representa o arquivo de entrada com o nome do tipo de dado e os dados a serem gerados hashes e `out.csv` é o arquivo de saída com os valores originais de entrada e uma coluna adicional dada por `-f`.

O executável do Python deve estar no caminho do terminal.