

Desafio FIAP />

IdWall />

GRUPO DataTales

idwall

FIAP

Desafio

FBI /> INTERPOL

Criar um sistema de gerenciamento sobre as pessoas mais procuradas pelo FBI e Interpol, para a IdWall.

Resumo Definições Técnicas

Construção da API

- Linguagem de Programação: Java
- Framework: Spring Boot
- Bibliotecas/Frameworks Adicionais: Spring Security, Spring Data JPA, Hibernate, Spring Doc
- Gerenciador de Dependências: Maven
- Testes: Postman
- Documentação da API: Swagger

Banco de Dados

- Modelagem de dados: Sql Developer Data Modeler
- Banco de Dados: Oracle SQL

Aquisição de Dados

- API FBI ([link](#))
- API Interpol ([link](#))
- Linguagem de script: Python
- Bibliotecas: requests, json, pandas, cx_Oracle, pycountry, np, sys

Scripts Python para aquisição de dados das APIs

Os scripts são destinados para execução periódica em ferramenta de automação (sugestão: semanal, Jenkins ou Airflow). É importante instalar o módulo `python cx_Oracle` e o `client Oracle` em sua máquina para conseguir executar os scripts, assim como alterar o caminho para sua pasta descompactada diretamente no código (indicação no próprio código).

Extração

Script ***extract-fbi-api.py***: Realiza extração da base de dados completa do FBI e os insere na tabela `FBI_CRIMINALS_DATABASE` dedicada apenas para dados do FBI. Algumas colunas sem utilidade identificada são abandonadas.

Script ***extract-interpol-api.py***: Realiza extração da base de dados completa da Interpol e os insere na tabela `INTERPOL_CRIMINALS_DATABASE` dedicada apenas para dados do FBI. Algumas colunas sem utilidade identificada são abandonadas. É comum que dados aninhados em alguns atributos sejam "explodidos" para colunas. A API da Interpol concedeu apenas 160 registros distintos no retorno de sua API>

Transformação e Carregamento

Script ***transform-load.py***: Recoleta os dados de ambas tabelas `FBI_CRIMINALS_DATABASE` e `INTERPOL_CRIMINALS_DATABASE` em dataframes pandas para realizar transformações, como renomear suas colunas para estabelecer as equivalências identificadas entre ambas ou converter listas para strings de itens separados por ponto e vírgula. Faz a união de ambas e a salva em 2 tabelas diferentes: uma para manter o histórico de criminosos, outra para manter a lista mais atualizada da última execução em uma tabela de produção, utilizada para consumo pela API.

Banco de Dados

Os scripts executados no slide anterior alimentam um **banco de dados relacional Oracle SQL 19c remoto da faculdade FIAP**.

Os dados de acesso completo estão disponíveis como variáveis no início de cada script, podendo ser alterados facilmente para tornar o script utilizável no mesmo tipo de banco para outro usuário. Conforme a biblioteca `cx_Oracle` foi utilizada em todos os scripts, o código está fortemente vinculado a este tipo de banco relacional, precisando ser refatorado para uso em outro tipo de banco.

As tabelas geradas contêm todas as colunas com o tipo de dado `VARCHAR2(4000)`.

FBI_CRIMINALS_DATABASE	
age_max	VARCHAR2 (4000 BYTE)
age_min	VARCHAR2 (4000 BYTE)
nationality	VARCHAR2 (4000 BYTE)
dates_of_birth_used	VARCHAR2 (4000 BYTE)
images	VARCHAR2 (4000 BYTE)
languages	VARCHAR2 (4000 BYTE)
place_of_birth	VARCHAR2 (4000 BYTE)
sex	VARCHAR2 (4000 BYTE)
weight	VARCHAR2 (4000 BYTE)
age_range	VARCHAR2 (4000 BYTE)
aliases	VARCHAR2 (4000 BYTE)
build	VARCHAR2 (4000 BYTE)
details	VARCHAR2 (4000 BYTE)
eyes_raw	VARCHAR2 (4000 BYTE)
field_offices	VARCHAR2 (4000 BYTE)
hair_raw	VARCHAR2 (4000 BYTE)
height_max	VARCHAR2 (4000 BYTE)
height_min	VARCHAR2 (4000 BYTE)
modified	VARCHAR2 (4000 BYTE)
ncic	VARCHAR2 (4000 BYTE)
occupations	VARCHAR2 (4000 BYTE)
person_classification	VARCHAR2 (4000 BYTE)
possible_countries	VARCHAR2 (4000 BYTE)
possible_states	VARCHAR2 (4000 BYTE)
poster_classification	VARCHAR2 (4000 BYTE)
publication	VARCHAR2 (4000 BYTE)
race	VARCHAR2 (4000 BYTE)
race_raw	VARCHAR2 (4000 BYTE)
remarks	VARCHAR2 (4000 BYTE)
reward_text	VARCHAR2 (4000 BYTE)
status	VARCHAR2 (4000 BYTE)
subjects	VARCHAR2 (4000 BYTE)
suspects	VARCHAR2 (4000 BYTE)
title	VARCHAR2 (4000 BYTE)
url	VARCHAR2 (4000 BYTE)
warning_message	VARCHAR2 (4000 BYTE)
weight_max	VARCHAR2 (4000 BYTE)
weight_min	VARCHAR2 (4000 BYTE)
analyzed_at	VARCHAR2 (4000 BYTE)

INTERPOL_CRIMINALS_DATABASE	
charge	VARCHAR2 (4000 BYTE)
nationalities	VARCHAR2 (4000 BYTE)
date_of_birth	VARCHAR2 (4000 BYTE)
distinguishing_marks	VARCHAR2 (4000 BYTE)
eyes_colors_id	VARCHAR2 (4000 BYTE)
forename	VARCHAR2 (4000 BYTE)
height	VARCHAR2 (4000 BYTE)
images	VARCHAR2 (4000 BYTE)
issuing_country_id	VARCHAR2 (4000 BYTE)
languages_spoken_ids	VARCHAR2 (4000 BYTE)
name	VARCHAR2 (4000 BYTE)
sex_id	VARCHAR2 (4000 BYTE)
weight	VARCHAR2 (4000 BYTE)
age_range	VARCHAR2 (4000 BYTE)
aliases	VARCHAR2 (4000 BYTE)
hairs_id	VARCHAR2 (4000 BYTE)
country_of_birth_id	VARCHAR2 (4000 BYTE)
analyzed_at	VARCHAR2 (4000 BYTE)
entity_id	VARCHAR2 (4000 BYTE)

FBI_INTERPOL_WANTED_CRIMINALS	
wanted_origin_id	VARCHAR2 (4000 BYTE)
charges	VARCHAR2 (4000 BYTE)
nationality	VARCHAR2 (4000 BYTE)
dates_of_birth_used	VARCHAR2 (4000 BYTE)
distinguishing_marks	VARCHAR2 (4000 BYTE)
eyes_color	VARCHAR2 (4000 BYTE)
forename	VARCHAR2 (4000 BYTE)
hair_color	VARCHAR2 (4000 BYTE)
height	VARCHAR2 (4000 BYTE)
images	VARCHAR2 (4000 BYTE)
issuing_country_id	VARCHAR2 (4000 BYTE)
languages	VARCHAR2 (4000 BYTE)
name	VARCHAR2 (4000 BYTE)
place_of_birth	VARCHAR2 (4000 BYTE)
sex	VARCHAR2 (4000 BYTE)
weight	VARCHAR2 (4000 BYTE)
wanted_origin	VARCHAR2 (4000 BYTE)
age_range	VARCHAR2 (4000 BYTE)
aliases	VARCHAR2 (4000 BYTE)
age_max	VARCHAR2 (4000 BYTE)
age_min	VARCHAR2 (4000 BYTE)
build	VARCHAR2 (4000 BYTE)
complexion	VARCHAR2 (4000 BYTE)
details	VARCHAR2 (4000 BYTE)
eyes_raw	VARCHAR2 (4000 BYTE)
field_offices	VARCHAR2 (4000 BYTE)
hair_raw	VARCHAR2 (4000 BYTE)
height_max	VARCHAR2 (4000 BYTE)
height_min	VARCHAR2 (4000 BYTE)
modified	VARCHAR2 (4000 BYTE)
ncic	VARCHAR2 (4000 BYTE)
occupations	VARCHAR2 (4000 BYTE)
person_classification	VARCHAR2 (4000 BYTE)
possible_countries	VARCHAR2 (4000 BYTE)
possible_states	VARCHAR2 (4000 BYTE)
poster_classification	VARCHAR2 (4000 BYTE)
publication	VARCHAR2 (4000 BYTE)
race	VARCHAR2 (4000 BYTE)
race_raw	VARCHAR2 (4000 BYTE)
remarks	VARCHAR2 (4000 BYTE)
reward_text	VARCHAR2 (4000 BYTE)
status	VARCHAR2 (4000 BYTE)
subjects	VARCHAR2 (4000 BYTE)
suspects	VARCHAR2 (4000 BYTE)
title	VARCHAR2 (4000 BYTE)
url	VARCHAR2 (4000 BYTE)
warning_message	VARCHAR2 (4000 BYTE)
weight_max	VARCHAR2 (4000 BYTE)
weight_min	VARCHAR2 (4000 BYTE)
analyzed_at	VARCHAR2 (4000 BYTE)

FBI_INTERPOL_WANTED_CRIMINALS_HISTORY	
wanted_origin_id	VARCHAR2 (4000 BYTE)
charges	VARCHAR2 (4000 BYTE)
nationality	VARCHAR2 (4000 BYTE)
dates_of_birth_used	VARCHAR2 (4000 BYTE)
distinguishing_marks	VARCHAR2 (4000 BYTE)
eyes_color	VARCHAR2 (4000 BYTE)
forename	VARCHAR2 (4000 BYTE)
hair_color	VARCHAR2 (4000 BYTE)
height	VARCHAR2 (4000 BYTE)
images	VARCHAR2 (4000 BYTE)
issuing_country_id	VARCHAR2 (4000 BYTE)
languages	VARCHAR2 (4000 BYTE)
name	VARCHAR2 (4000 BYTE)
place_of_birth	VARCHAR2 (4000 BYTE)
sex	VARCHAR2 (4000 BYTE)
weight	VARCHAR2 (4000 BYTE)
wanted_origin	VARCHAR2 (4000 BYTE)
age_range	VARCHAR2 (4000 BYTE)
aliases	VARCHAR2 (4000 BYTE)
age_max	VARCHAR2 (4000 BYTE)
age_min	VARCHAR2 (4000 BYTE)
build	VARCHAR2 (4000 BYTE)
complexion	VARCHAR2 (4000 BYTE)
details	VARCHAR2 (4000 BYTE)
eyes_raw	VARCHAR2 (4000 BYTE)
field_offices	VARCHAR2 (4000 BYTE)
hair_raw	VARCHAR2 (4000 BYTE)
height_max	VARCHAR2 (4000 BYTE)
height_min	VARCHAR2 (4000 BYTE)
modified	VARCHAR2 (4000 BYTE)
ncic	VARCHAR2 (4000 BYTE)
occupations	VARCHAR2 (4000 BYTE)
person_classification	VARCHAR2 (4000 BYTE)
possible_countries	VARCHAR2 (4000 BYTE)
possible_states	VARCHAR2 (4000 BYTE)
poster_classification	VARCHAR2 (4000 BYTE)
publication	VARCHAR2 (4000 BYTE)
race	VARCHAR2 (4000 BYTE)
race_raw	VARCHAR2 (4000 BYTE)
remarks	VARCHAR2 (4000 BYTE)
reward_text	VARCHAR2 (4000 BYTE)
status	VARCHAR2 (4000 BYTE)
subjects	VARCHAR2 (4000 BYTE)
suspects	VARCHAR2 (4000 BYTE)
title	VARCHAR2 (4000 BYTE)
url	VARCHAR2 (4000 BYTE)
warning_message	VARCHAR2 (4000 BYTE)
weight_max	VARCHAR2 (4000 BYTE)
weight_min	VARCHAR2 (4000 BYTE)
analyzed_at	VARCHAR2 (4000 BYTE)

Modelagem de Dados (MER)

Documentação API

[challenge-idwall](#) | [1.0.0](#) | [cgo devs](#) | [SwaggerHub](#)

SMARTBEAR
SwaggerHub

Explore APIs for free

challenge-idwall 1.0.0

InfoTagsSearchParâmetros de buscaModels

AaSAVESync

```
1 # Link para nossa documentação: https://app.swaggerhub.com/apis/cgo devs/challenge-idwall/1.0.0
2
3 # Caso o link para a documentação acima não esteja disponível, basta colar o conteúdo desse arquivo
  .yaml no https://editor.swagger.io/ para visualizar nossa doc em Swagger.
4
5 swagger: '2.0'
6 info:
7   version: "1.0.0"
8   title: Idwall Background Check
9   description: A API opera sobre diversos parâmetros de busca para realizar o background check,
    trazendo uma lista de possíveis matches em relação aos dados informados, utilizando de um
    conjunto para gerar maior assertividade. O filtro de busca para todos os parâmetros, exceto
    full_name que é definido por exact match, procura por dados que contenham a palavra da busca,
    não sendo equivalência exata para retornar um criminoso correspondente.
10 paths:
11   /wanted:
12     get:
13       tags:
14         - Parâmetros de busca
15       description: Retorna lista de possíveis criminosos através de busca por conjunto de
        parâmetros.
16       operationId: getCriminalByDescription
17       parameters:
18         - in: query
19           name: wanted_origin_id
20           type: string
21           description: ID do procurado na respectiva organização do registro (FBI ou pela Interpol
            ); Este é o id do criminoso em nosso banco.
22         - in: query
23           name: alias
24           type: string
25           description: "Retorna lista de possíveis criminosos para uma busca por nome. (Coluna do
            banco de dados relacional: \"aliases\")"
26         - in: query
27           name: exact_match
28           type: boolean
29           description: Parâmetro opcional que deve ser informado junto do parâmetro "full_name". A
            busca é feita para nomes estritamente iguais ao parâmetro full_name caso exact_match
            =true, caso seja igual a false, o nome é usado para trazer correspondências parciais,
            considerando o parâmetro uma substring dos registros correspondentes retornados.
30         - in: query
31           name: birthdate
32           type: string
33           description: "Adiciona o critério de data de nascimento na busca do indivíduo. (Coluna do
            banco de dados relacional: \"dates_of_birth_used\")"
34         - in: query
35           name: distinguishing_marks
36           type: string
37           description: Adiciona o critério de características físicas mais marcantes na busca do
            indivíduo. Separar itens com ponto e vírgula.
38         - in: query
39           name: nationality
40           type: string
41           description: Adiciona o critério de nacionalidades na busca do indivíduo. Separar itens
            com ponto e vírgula.
42         - in: query
43           name: languages
44           type: string
45           description: Adiciona o critério do idioma falado na busca do indivíduo.
46         - in: query
```

Last Saved: 10:53:36 pm - Sep 30, 2023

Idwall Background Check

1.0.0

A API opera sobre diversos parâmetros de busca para realizar o background check, trazendo uma lista de possíveis matches em relação aos dados informados, utilizando de um conjunto para gerar maior assertividade. O filtro de busca para todos os parâmetros, exceto full_name que é definido por exact match, procura por dados que contenham a palavra da busca, não sendo equivalência exata para retornar um criminoso correspondente.

Note: The interactive forms are disabled, as no 'host' property was found in the specification. Please see: [OAI 2.0/#swagger-object](#)

Parâmetros de busca


GET /wanted

Retorna lista de possíveis criminosos através de busca por conjunto de parâmetros.

Name	Description
wanted_origin_id string (query)	ID do procurado na respectiva organização do registro (FBI ou pela Interpol); Este é o id do criminoso em nosso banco. <input type="text" value="wanted_origin_id"/>
alias string (query)	Retorna lista de possíveis criminosos para uma busca por nome. (Coluna do banco de dados relacional: "aliases") <input type="text" value="alias"/>
exact_match boolean (query)	Parâmetro opcional que deve ser informado junto do parâmetro "full_name". A busca é feita para nomes estritamente iguais ao parâmetro full_name caso exact_match=true, caso seja igual a false, o nome é usado para trazer correspondências parciais, considerando o parâmetro uma substring dos registros correspondentes retornados. <input type="text" value="--"/>
birthdate string	Adiciona o critério de data de nascimento na busca do indivíduo. (Coluna do banco de dados relacional: "dates_of_birth_used") <input type="text" value="birthdate"/>

Modelo Site

127.0.0.1:5500/frontend/components/listarID.html




Id Wall

Search by ID

List all

Most wanted



Search

Obrigado!

GRUPO DataTales

Kennedy Cesar Pimentel (RM95249)
kennedywork06@gmail.com

Lucas Ferreira Venturini (RM94027)
flucas.venturini@hotmail.com

Caroline Gonçalves de Oliveira (RM95511)
kharol.go@gmail.com

Pedro Henrique Lucena Brandão (RM95051)
pedrohlucena413@gmail.com