

PROPOSTA DE ALGORITMO PARA BUSCA DE TERMOS GEOLÓGICOS EM RELATÓRIOS DE POÇOS DO SISTEMA SIAGAS

Paula Josyane dos Santos Francisco¹; Juliana Pertille da Silva²

RESUMO

Os recursos hídricos subterrâneos são utilizados como fonte de abastecimento em diversas cidades brasileiras. Por isso estudos que viabilizam a compreensão desses recursos são relevantes para a correta gestão destes mesmos mananciais. Esse trabalho apresenta um algoritmo de busca que permite realizar pesquisas por formação geológica nos relatórios de poços inseridos no banco de dados SIAGAS (Sistema de Informações de Águas Subterrâneas). A principal função desta ferramenta é otimizar o tempo de retorno ao usuário de informações sobre as profundidades e distribuições de determinadas camadas estratigráficas. Além disso, este algoritmo é uma ferramenta para a construção de modelos conceituais que descrevam as condições geológicas locais de distintos sistemas aquíferos. O resultado deste algoritmo é a saída das respectivas profundidades inicial e final nos dados de poços, acompanhadas do termo de busca.

ABSTRACT

The groundwater resources are used for water supply in several brazilian cities. Therefore, studies to contribute the understanding of these resources are relevant for the best management of these sources. This work presents a search algorithm that allows searching by the geological formation in the well reports inserted in the SIAGAS (Sistema de Informações de Águas Subterrâneas) database. The main function of this tool is to optimize the return time to the user of information about the depths and distributions of certain stratigraphic layers. Moreover, this algorithm is a tool for the construction of conceptual models that describe local geological conditions of different aquifer systems. The result of this algorithm is the output of the respective initial and final depths in the well data, accompanied by the search term.

Palavras-chave: Algoritmo de busca, formação geológica, SIAGAS

¹ Universidade Federal de Pelotas; Rua Gomes Carneiro, 1; e-mail: paula.josyane@uol.com.br

² Universidade Federal de Pelotas; Rua Gomes Carneiro; e-mail: juliana.pertill@gmail.com

1 INTRODUÇÃO

Grande parte da água doce no planeta é encontrada no meio subterrâneo, por esta razão estudos relacionados com o fluxo subterrâneo, tornam-se ferramentas imprescindíveis para análise, através de modelos conceituais, que subsidiam modelos matemáticos [1]. Com o aumento da base de dados disponíveis, a compilação de dados para elaboração de modelos conceituais, torna-se cada vez mais complicada e onerosa. Desta forma, ferramentas que auxiliem a busca e organização dos dados tornam-se cada vez mais importantes para o gerenciamento de águas subterrâneas. Este trabalho apresenta uma proposta de script de busca para termos geológicos nacionais na mais completa plataforma de registro de informações de águas subterrâneas do país, o SIAGAS da CPRM.

2 OBJETIVOS

O objetivo principal deste estudo é propor um algoritmo de busca de termos geológicos específicos nos relatórios dos poços tubulares profundos cadastrados no SIAGAS. Os objetivos específicos incluem a otimização do tempo de busca e o auxílio na confecção de modelos conceituais de sistemas aquíferos. Esta ferramenta é particularmente útil para sistemas aquíferos multicamadas.

3 MATERIAIS E MÉTODOS

Nesse trabalho foram desenvolvidas as seguintes etapas metodológicas: 1) Seleção da área de interesse no município de Alegrete, RS; 2) Seleção e compilação dos dados de poços do município através de 17 relatórios; 3) Extração das informações geológicas dos relatórios através da linguagem de programação Python e a IDE (*Integrated Development Environment*) PyCharm, que será detalhada a seguir.

A parte inicial do *script* realiza a importação dos módulos utilizados na conversão, leitura e listagem dos arquivos no diretório escolhido pelo usuário. Neste *script*, o termo de busca é tratado como uma variável e o tipo de dado que é aceito, denomina-se *string*. O

termo *string* designa um texto, que é essencialmente a forma de apresentação de uma informação [2]. A Figura 1 exibe nas linhas 8, 15 e 16 do *script*, os *inputs* a serem realizados pelo usuário. Os *inputs* referem-se respectivamente ao diretório onde está localizado o banco de dados, o termo de busca e se o usuário deseja buscar a formação em um relatório específico ou em todos os relatórios. A segunda parte do código, ocupa-se em extrair a segunda página do arquivo com extensão pdf, com informações sobre o perfil geológico dos poços. Todas as operações contidas neste *script* são realizadas, em modo lote, isso significa que a mesma operação é realizada para todos os arquivos no banco de dados. As partes seguintes do *script* convertem o formato do arquivo para extensão xlsx e realizam uma varredura nestes arquivos em busca do termo desejado. A partir da linha 18 (Figura 1), os arquivos passam por um processo de leitura, conversão e um resultado é impresso.

```
1 import os
2 import xlrd
3 import PyPDF2
4 import pathlib2 as p
5 import camelot
6 print(os.getcwd())
7 print('Escolha o diretório onde estão localizados os relatórios de poço')
8 os.chdir(r'{}'.format(input('Insira agora diretório da banco de dados:')))
9 print('O diretório utilizado para esta busca será: {}'.format(os.getcwd()))
10 rel_dir = [i for i in os.listdir(os.getcwd()) if i.endswith('.pdf')]
11 print('Os relatórios neste diretório estão listados abaixo:')
12 r = sorted(rel_dir, key=lambda item: (int(item.partition(' ')[0]) if item[0].isdigit() else float('inf'), item))
13 print('\n'.join(r))
14 print('18 - Todos os relatórios.')
15 busca = input('Busque uma formação geologica:')
16 escolha = int(input('Escolha um relatório para analisar:'))
17 if escolha == 18:
18     for h in rel_dir:
19         print(os.path.abspath(h))
```

Figura 1. Linhas 1 a 21 do *script* de busca.

4 RESULTADOS E DISCUSSÕES

O principal resultado deste trabalho é um algoritmo que possui uma relação interativa com o usuário, permitindo que seja realizada a busca por termos geológicos específicos em bancos de dados. A primeira parte do *script* apresenta a lista de arquivos no diretório, informações a serem preenchidas pelo usuário e o diretório absoluto em que o arquivo se encontra (Figura 2a). A segunda parte do *script* exibe o resultado principal do algoritmo de busca (Figura 2b). Ao longo de todo *script* foi utilizada a função *print*, como forma de identificar em qual ponto o algoritmo encontrava-se. Todo o *script* é executado em 158,79 segundos, considerando o tempo de inserção do diretório da base de dados, a formação geológica e opção de executar a busca para todos os arquivos.

```

Prompt de Comando - python "arquivo final.py"
Escolha o diretório onde estão localizados os relatório de poço
Insira agora diretório da banco de dados:C:\Users\Paula2017\Desktop\Secao1
0 diretório utilizado para esta busca será: C:\Users\Paula2017\Desktop\Secao1
0s relatórios neste diretório estão listados abaixo:
1 - 4300009360.pdf
2 - 4300009283.pdf
3 - 4300024184.pdf
4 - 4300001346.pdf
5 - 4300024854.pdf
6 - 4300007972.pdf
7 - 4300001353.pdf
8 - 4300001354.pdf
9 - 4300001338.pdf
10 - 4300001350.pdf
11 - 4300001643.pdf
12 - 4300007965.pdf
13 - 4300007967.pdf
14 - 4300007966.pdf
15 - 4300001257.pdf
16 - 4300001256.pdf
17 - 4300001255.pdf
18 - Todos os relatórios.
Busque uma formação geologica:basalto
Escolha um relatório para analisar:18
C:\Users\Paula2017\Desktop\Secao1\1 - 4300009360.pdf
C:\Users\Paula2017\Desktop\Secao1\10 - 4300001350.pdf
C:\Users\Paula2017\Desktop\Secao1\11 - 4300001643.pdf
C:\Users\Paula2017\Desktop\Secao1\12 - 4300007965.pdf
C:\Users\Paula2017\Desktop\Secao1\13 - 4300007967.pdf
C:\Users\Paula2017\Desktop\Secao1\14 - 4300007966.pdf
C:\Users\Paula2017\Desktop\Secao1\15 - 4300001257.pdf
C:\Users\Paula2017\Desktop\Secao1\16 - 4300001256.pdf
C:\Users\Paula2017\Desktop\Secao1\17 - 4300001255.pdf
C:\Users\Paula2017\Desktop\Secao1\2 - 4300009283.pdf
(a)
Prompt de Comando
C:\Users\Paula2017\Desktop\Secao1\17 - 4300001255.pdf
C:\Users\Paula2017\Desktop\Secao1\2 - 4300009283.pdf
C:\Users\Paula2017\Desktop\Secao1\3 - 4300024184.pdf
C:\Users\Paula2017\Desktop\Secao1\4 - 4300001346.pdf
C:\Users\Paula2017\Desktop\Secao1\5 - 4300024854.pdf
C:\Users\Paula2017\Desktop\Secao1\6 - 4300007972.pdf
C:\Users\Paula2017\Desktop\Secao1\7 - 4300001353.pdf
C:\Users\Paula2017\Desktop\Secao1\8 - 4300001354.pdf
C:\Users\Paula2017\Desktop\Secao1\9 - 4300001338.pdf
-----
Poço:4300009360
De(m): 0 Até(m): 30 Litologia: Basalto
De(m): 30 Até(m): 40 Litologia: Basalto
De(m): 40 Até(m): 50 Litologia: Basalto
De(m): 50 Até(m): 64 Litologia: Basalto
De(m): 64 Até(m): 72 Litologia: Basalto
De(m): 72 Até(m): 100 Litologia: Basalto
De(m): 100 Até(m): 120 Litologia: Basalto
De(m): 120 Até(m): 154 Litologia: Basalto
De(m): 154 Até(m): 215 Litologia: Basalto
-----
Poço:4300001350
De(m): 0 Até(m): 20 Litologia: Basalto
De(m): 20 Até(m): 32 Litologia: Basalto
De(m): 32 Até(m): 64 Litologia: Basalto
-----
Poço:4300007965
De(m): 3 Até(m): 26 Litologia: Basalto
De(m): 26 Até(m): 32 Litologia: Basalto
De(m): 32 Até(m): 44 Litologia: Basalto
-----
Poço:4300007966
De(m): 1 Até(m): 38 Litologia: Basalto
De(m): 38 Até(m): 43 Litologia: Basalto
De(m): 43 Até(m): 73 Litologia: Basalto
(b)

```

Figura 2. a) Primeira parte do algoritmo de busca; b) Execução do algoritmo, com resultado da busca para o termo basalto.

5 CONSIDERAÇÕES FINAIS

O algoritmo aplicou-se bem ao número de poços testados. O tempo de execução do script pode variar de acordo com o número de relatórios inseridos na base de dados e com a capacidade de processamento do computador. O resultado obtido neste script encontra-se apto a ser inseridos em um banco de dados mais robusto como *MySQL* ou *SQLite*. Este algoritmo é parte integrante do modelo conceitual a ser construído para o perímetro urbano do município de Alegre/RS.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- [1] FEITOSA, F. A. C.; MANUEL FILHO, J. Hidrogeologia: Conceitos e Aplicações. CPRM, Rio de Janeiro, 2008.
- [2] SUMMERFIELD, Mark. Programação em Python 3: uma introdução completa à linguagem Python. Alta Books, Rio de Janeiro, 2012.