

2 a 4
MARÇO
2021

Evento Virtual



Uma década de conhecimento em conexão

#CampusOsório10anos



INSTITUTO FEDERAL
Rio Grande do Sul
Campus Osório

ISSN: 2526-3250

Automatização da aquisição de dados de Covid-19 por web scraping e atualização de mapas do ArcGis Online utilizando Python.

Autor(es):

- Eliseu José Weber
- Ricardo de Sampaio Dagnino
- Martim Kowalczyk Presser

Nível de Ensino: Ensino Superior

Área do Conhecimento: Pesquisa - Ciências Exatas e da Terra

Resumo:

Em 2020, o novo coronavírus ocasionou uma pandemia da doença Covid-19, acompanhada de uma onda de desinformação. Na tentativa de esclarecer e fornecer dados atualizados para a população, várias iniciativas de monitoramento e análise de dados surgiram, dentre elas uma plataforma com mapas e estatísticas atualizadas diariamente da doença nos municípios do Rio Grande do Sul. O objetivo deste trabalho é apresentar o processo de construção dessa plataforma e as melhorias implementadas ao longo do processo visando a diminuição dos erros de entrada de dados e a economia de tempo nas atualizações diárias. Inicialmente, a coleta e o processamento dos dados eram feitos manualmente: consultados diariamente na página HTML da Secretaria Estadual de Saúde e inseridos no Sistema de Informação Geográfica baseado em ArcGis Online para serem disponibilizados publicamente. A partir do crescimento dos casos e a demanda maior de tempo para as atualizações, foi implementada em agosto de 2020 a automatização desse processo de aquisição de dados e atualização dos mapas através de um programa em python utilizando a técnica de web scraping (raspagem de dados). O programa, depois de iniciado, automaticamente analisa os dados na página HTML e CSV hospedado, coleta, processa, cria arquivos CSV secundários e envia os dados para a base de dados do Sistema de Informação Geográfica baseado em ArcGis Online. O programa, que levou aproximadamente 100 horas de trabalho para ser criado, realiza em apenas 10 minutos o trabalho que anteriormente poderia levar aproximadamente 1 hora, ou até passar disso caso alguma inconsistência ou erro de digitação fossem detectados. Assim, estima-se que desde sua implementação em agosto de 2020, o uso do programa tenha gerado uma economia de no mínimo 127 horas de trabalho. Uma economia que varia entre 83%, se continuássemos levando 60 minutos para atualizar os dados, a 97%, caso o tempo decorrente da atualização de 2500 casos fosse aumentando proporcionalmente. Conclui-se que o programa desenvolvido otimizou bem o tempo de trabalho, facilitando-o e também mostra boa modularidade para projetos futuros, ou seja, grandes partes de seu código podem ser reaproveitadas, visto que têm pouca dependência entre si.

Disponível em <https://moexp.osorio.ifrs.edu.br/uploads/anai/2020/Anais MoExP 2020 Etapa II.1805.pdf>

Anais da Mostra de Ensino, Extensão e Pesquisa do Campus Osório - MoExP.
<https://moexp.osorio.ifrs.edu.br/anais>