



XXXIII SIC SALÃO INICIAÇÃO CIENTÍFICA

Evento	Salão UFRGS 2021: SIC - XXXIII SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2021
Local	Virtual
Título	Automatização da aquisição de dados de Covid-19 por web scraping e atualização de mapas do ArcGis Online utilizando Python
Autor	MARTIM KOWALCZUK PRESSER
Orientador	RICARDO DE SAMPAIO DAGNINO

Automatização da coleta de dados de Covid-19 por web scrap e atualização de mapas do ArcGis Online utilizando Python. Em 2020, o novo coronavírus ocasionou a pandemia de Covid-19, acompanhada de uma onda de desinformação. Para esclarecer e fornecer dados atualizados para a população, o Projeto SIG Litoral criou uma plataforma com mapas e estatísticas da doença nos municípios do Rio Grande do Sul. Inicialmente, a coleta e o processamento dos dados eram feitos manualmente: consultados diariamente na página da Secretaria Estadual de Saúde e inseridos no Sistema de Informação Geográfica baseado em ArcGis Online. A partir do crescimento dos casos e a demanda maior de tempo para as atualizações, em agosto de 2020 implementou-se 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 analisa os dados na página HTML e CSV hospedado, realiza cálculos e transformações e em seguida envia tudo para a base de dados dentro do ArcGis Online; tudo isso, que feito manualmente poderia levar horas, agora é realizado em apenas 10 minutos. Assim, estima-se que desde sua implementação em agosto de 2020 até hoje, o uso do programa tenha gerado uma economia de centenas de horas de trabalho: economia que varia de 83%, se continuássemos levando 60 minutos para atualizar os dados, a 97%, caso o tempo decorrente da atualização fosse aumentando proporcionalmente aos casos. Conclui-se que o programa desenvolvido resultou na diminuição dos erros de entrada de dados e em economia de tempo, otimizando bem o tempo de trabalho, facilitando-o. Também mostra boa modularidade, ou seja, aproveitamento de código para a continuidade do projeto. Exemplo disso foi a alteração no programa realizada em novembro de 2020 quando o site da Secretaria Estadual de Saúde passou a utilizar um mecanismo de segurança no protocolo de transferência de hipertexto (HTTP – HyperText Transfer Protocol), utilizando HTTPS.