



ESTUDO DE TÉCNICAS PARA AUTOMATIZAR A COLETA DE DADOS VIA WEB

Bruna Pinto Pereira, Annabell D. R. Tamariz

As tecnologias digitais de comunicação e informação estão possibilitando muitas mudanças. As redes não só de máquinas e de informação, mas principalmente de pessoas e de comunidades, estão permitindo configurar novos espaços de interação e de aprendizagem. Os serviços de localização nos dispositivos móveis surgiram no ano de 1996, quando o governo americano, numa tentativa de agilizar os atendimentos, determinou que as operadoras de telefonia móvel deveriam fornecer às centrais públicas de atendimento e segurança a localização de um cliente que realizasse um chamado de emergência ao número 911. Nosso foco nesta pesquisa de tecnologias digitais de comunicação e informação se concentra na telefonia móvel, e baseada nestas informações pretendesse criar um aplicativo inteligente customizado para localização de unidades de saúde usando dispositivos móveis. Para realizar esta tarefa, são necessários os endereços detalhados de todas as unidades de saúde. Este projeto é responsável por automatizar a coleta de endereços e coordenadas, especificamente dos hospitais, a partir do site do Cadastro Nacional de Estabelecimentos de Saúde (CNES), para o desenvolvimento desta fase utilizaremos a linguagem Python e a biblioteca *BeautifulSoup*. O objetivo é coletar o endereço completo de todos os hospitais do Brasil, assim como suas coordenadas geográficas para, posteriormente, a implementação de um mapa contendo todos esses endereços. Para isso, está em construção um *script* escrito em Python e mais especificamente utilizando uma biblioteca chamada *BeautifulSoup*, que é responsável pela coleta de dados em um arquivo HTML de acordo com o padrão seguido pelo código da página. Os endereços são exportados para um arquivo JSON em forma de lista, que será usada como banco de dados para o aplicativo. O site do CNES é atualizado frequentemente, por isso o *script* tem que estar hábil para ser executado pelo servidor periodicamente para que o banco de dados se mantenha sempre atualizado.

Palavras-chave: Automatização, Coleta, Endereços.

Instituição de fomento: UENF