TRABALHO COMPLETO
FULL PAPER
XXVIII SEPE
Simpósio de Ensino, Pesquisa e Extensão
Teaching, Research and Extension Symposium
ISSN: 2316-9745
24 A 26 DE SETEMBRO DE 2024
SEPTEMBER 24th to 26th, 2024
UFN
https://doi.org/10.48195/sepe2024.28773

VISUALIZADOR WEB PARA ANÁLISE DE PRODUÇÃO CIENTÍFICA EM CURRÍCULO LATTES UTILIZANDO PYTHON E MACHINE LEARNING

Paulo André Machado Maracci; Gabriel Azenha Fachim; Mirkos Ortiz Martins

RESUMO

A Plataforma Currículo Lattes é amplamente usada no Brasil para registrar produções e atividades de pesquisadores, funcionando como um sistema de avaliação baseado em pontuação. No entanto, analisar currículos Lattes tornou-se demorado devido à dificuldade de acessar informações no formato XML disponibilizado. Para agilizar esse processo, foi desenvolvido um software em Python que utiliza técnicas de machine learning para coletar e analisar dados dos currículos. Esse software permite baixar e processar arquivos XML de múltiplos currículos simultaneamente, facilitando a seleção de pesquisadores por universidades, empresas e outras instituições. A criação da aplicação web foi viabilizada pelo framework Flask, que permite a rápida implementação e extensibilidade de funcionalidades, proporcionando uma solução eficiente para analisar dados de produção científica de docentes cadastrados no Currículo Lattes.

Palavras-chave: Framework, aplicação, plataforma, software.

ABSTRACT

The Currículo Lattes Platform is widely used in Brazil to record researchers' productions and activities, functioning as a scoring-based evaluation system. However, analyzing Lattes curricula has become time-consuming due to the difficulty of accessing information in the provided XML format. To streamline this process, a Python software was developed that uses machine learning techniques to collect and analyze curriculum data. This software allows for the simultaneous downloading and processing of XML files from multiple curricula, facilitating the selection of researchers by universities, companies, and other institutions. The development of the web application was made possible by the Flask framework, which enables rapid implementation and extensibility of features, providing an efficient solution for analyzing scientific production data of faculty registered in the Currículo Lattes

Keywords: Framework, application, platform, software.

Eixo Temático: Tecnologia, Inovação e Desenvolvimento Sustentável (TIDS).

1. INTRODUÇÃO

A plataforma de currículo lattes é utilizada em todo o país para registro de produções e atividades desenvolvidas por pesquisadores, a Plataforma tornou-se um Sistema de avaliação, inclusive por pontuação, da carreira dos pesquisadores baseado em seus projetos.

Entretanto, desenvolver uma análise produtiva de currículos dessa Plataforma tornou-se um processo moroso devido a dificuldade de acesso a informação de forma facilitada devido ao formato XML disponível para download na Plataforma governamental, o qual limita a possibilidade de gestão dos dados contidos no currículo lattes.

Com isso, utilizando o Flask, um framework web que se destacou como uma escolha eficaz (Pallets Project, 2010), foi possível a rápida criação de uma aplicação web, oferecendo extensibilidade para implementar funcionalidades conforme necessário, resultando em uma aplicação web que realiza uma análise de dados com informações de produção científica de docentes cadastrados no Currículo Lattes utilizando Python e Machine Learning.

2. METODOLOGIA

O trabalho foi desenvolvido seguindo a metodologia ágil Feature Driven Development (FDD), que se destaca por sua abordagem incremental na construção de sistemas, implementando funcionalidades de forma gradual (Feature Driven Development, 2021). Essa abordagem proporcionou um desenvolvimento mais controlado e direcionado, garantindo que cada componente do sistema fosse cuidadosamente desenvolvido antes de prosseguir para a etapa seguinte.

No início do projeto, foi escolhido o framework FastAPI, da linguagem Python, para o desenvolvimento da aplicação que funcionaria como servidor back-end do sistema (Triangolo, 2018). Essa escolha foi feita com o objetivo de facilitar a integração dos dados com o servidor front-end, desenvolvido com Python Flask, que além de se comunicar diretamente com o usuário final, é responsável por realizar o upload dos currículos Lattes ao back-end.

Para testar a aplicação, foram utilizados os currículos dos pesquisadores do Programa de Pós-Graduação em Nanociências da Universidade Franciscana, bem como de alunos de Ciência da Computação.

3. RESULTADOS E DISCUSSÕES

Como resultado deste trabalho, foi desenvolvido um software em Python que permite a coleta e análise de dados por meio de técnicas de machine learning. O software possibilita o download dos dados em formato XML e permite a análise simultânea de múltiplos currículos, facilitando o processo de seleção de pesquisadores por universidades, empresas e outras instituições.

Além disso, com o uso do Flask, um framework web escolhido por sua eficácia, foi possível criar rapidamente uma aplicação web extensível. Essa aplicação realiza a análise de dados referentes à produção científica de docentes cadastrados na Plataforma Lattes, utilizando Python e Machine Learning.

Na Figura 1 é mostrado a parte principal do software, onde os principais dados do pesquisador são exibidos, juntamente com as abas para acessar sua formação acadêmica, artigos publicados, trabalhos/eventos, livros e capítulos.

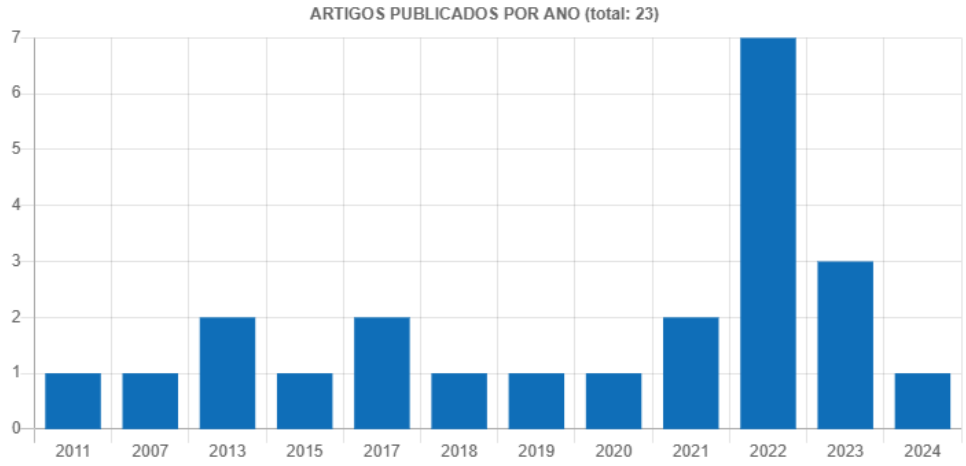
Figura 1 – Tela de dados do pesquisador.



Fonte: elaborado pelo autor (2024).

Na Figura 2 é exibido um gráfico em barras mostrando a quantidade de artigos publicados pelo pesquisador por ano e o total de artigos publicados.

Figura 2 – Gráfico de artigos publicados por ano.



Fonte: elaborado pelo autor (2024).

Na Figura 3 é exibido a lista dos pesquisadores que tiveram seus currículos carregados no software.

Figura 3 – Lista de pesquisadores carregados no sistema.



Fonte: elaborado pelo autor (2024).

4. CONCLUSÃO

A criação de um software utilizando Python e machine learning, em conjunto com o framework Flask, representa uma solução eficaz para a análise de currículos na Plataforma Lattes. Ao possibilitar a coleta e a análise de dados de forma simultânea em múltiplos currículos, a aplicação desenvolvida facilita o processo de seleção de pesquisadores para universidades, empresas e outras instituições.

Essa ferramenta não só simplifica o acesso às informações contidas nos currículos Lattes, como também potencializa a gestão de dados e o processo de avaliação da produção científica de docentes, tornando-se um recurso valioso para a comunidade acadêmica e científica no Brasil.

AGRADECIMENTOS

Este projeto é financiado pela Universidade Franciscana (UFN) de várias formas, incluindo apoio financeiro por meio de bolsas de iniciação científica e suporte estrutural com o acesso aos laboratórios para o desenvolvimento e testes da aplicação. Os autores expressam seu agradecimento à UFN pelo apoio integral recebido.

REFERÊNCIAS

PALLETS PROJECTS. Flask Documentation (Version 3.0.x). Disponível em: https://flask.palletsprojects.com/en/3.0.x/#user-s-guide. Acesso em: 17 ago. 2024.

PYTHON SOFTWARE FOUNDATION. Python.org. Disponível em: https://www.python.org/. Acesso em: 17 de ago. 2024.

FEATURE DRIVEN DEVELOPMENT. Feature Driven Development Official Site. Disponível em: http://www.featuredrivendevelopment.com. Acesso em: 19 ago. 2024.

TIANGOLO. FastAPI Documentation. Disponível em: https://fastapi.tiangolo.com. Acesso em: 19 ago. 2024.