

RESUMO EXPANDIDO  
XXVI Congresso de Iniciação Científica

## MÉTODO DE CONVERSÃO DE ARQUIVOS PDF PARA PLATAFORMA ON-LINE

Maria Fernanda Lima<sup>1</sup>

Silvia Cristina Martini<sup>2</sup>

1. Discente do curso de ADS; e-mail: [fmarialima225@gmail.com](mailto:fmarialima225@gmail.com)
2. Docente na Universidade de Mogi das Cruzes; e-mail: [silviac@umc.br](mailto:silviac@umc.br).

**Área de Conhecimento:** Tecnologia, Saúde, Imaginologia.

**Palavras-Chave:** PDF, Python, Imaginologia, Questões, Radiografia, Interpretação, Questionário, Extrair.

### Como citar:

Lima MF, Martini SC. Método de conversão de arquivos PDF para plataforma on-line. Revista Científica UMC [Internet]. 27º de outubro de 2023;8(2):e080200047.

Disponível em: <https://revista.umc.br/index.php/revistaumc/article/view/1904>

**Fluxo de revisão:** o presente resumo expandido foi revisado por pares pela comissão do evento.

Recebido em: 11/09/2023

Aprovado em: 26/10/2023

ID publicação: e080200047

DOI:

Licença CC BY 4.0 DEED

## INTRODUÇÃO

A As novas tecnologias computacionais desenvolveram maiores possibilidades de produção e uso de imagens, permitindo uma hipermediação que consiste na combinação da informação em suas múltiplas dimensões o que potencializou o emprego da fotografia nas diversas áreas do conhecimento (RODRIGUES, 2007).

Segundo Giger (2017) essas fotografias conhecidas na área da saúde como imagens médicas tem uma importância muito grande além do conhecimento da quantidade que os especialistas analisam para o diagnóstico. Isso mostra a importância de um diagnóstico correto utilizando imagens médicas.

No contexto educacional vem se falando constantemente no uso de "metodologias ativas", onde o aluno passa a ser o principal responsável por sua aprendizagem. O uso das Tecnologias da Informação e Comunicação (TIC's), especificamente na área da saúde, tem evidenciado um método significativo e inovador que, quando inserido no ensino, mostra-se como uma ferramenta facilitadora e fomentadora do conhecimento (MORGADO et al., AMES e SILVESTRE, 2019), tanto que tecnologias antes utilizadas apenas em cursos EAD estão sendo frequentemente usadas em cursos presenciais como o MOODLE, WebAula, TeleEduc, entre outros (CALIARI, ZILBER e PEREZ, 2017), (PEZZUOL, Jorge L.; ABREU, 2017).

O questionário é uma técnica com alta eficiência no processo de ensino aprendizagem e seu uso corriqueiro em sala de aula contribui com o aprendizado dos alunos e desenvolvimento de suas habilidades na matéria ministrada (DUNLOSKY et al., 2013).

Sendo assim o uso de recursos tecnológicos nesse processo é constantemente utilizado, e essa metodologia aplicada ao ensino da saúde no Brasil já vem sendo discutida (ROMAN, et al., 2017). Com o uso desse recurso o professor tem a possibilidade de direcionar mais as aulas dentro de sala de aula, e torná-las mais atrativas e interativas aos alunos, além da possibilidade de transpor as barreiras de sala de aula, onde o aluno pode ter acesso ao conteúdo das aulas de qualquer lugar e revê-los sempre que achar necessário ou for solicitado pelo professor.

## OBJETIVO

O objetivo geral deste trabalho foi desenvolver um conversor que possa auxiliar o professor a utilizar materiais de concurso ou prova de proficiência que normalmente estão no formato pdf para sua aula convencional alunos/residentes na interpretação de imagens médicas.

## METODOLOGIA

A Este projeto foi desenvolvido seguindo as etapas estabelecidas em seu cronograma, sendo elas resumidamente: estudar a linguagem de programação que melhor se encaixasse para a manipulação de PDF's, no caso a linguagem Python, que é uma linguagem que possui em sua distribuição padrão uma vasta coleção de módulos prontos para uso (Borges, 2015); pesquisar as principais bibliotecas para extração de texto e imagens de PDF e por meio de testes, escolher as bibliotecas que melhor atendessem as necessidades da aplicação; desenvolver a aplicação em si e testar sua funcionalidade em atender os objetivos da aplicação.

Visando atender essas etapas, foram utilizadas como ferramentas de estudo livros – como por exemplo Python Eficaz: 59 Maneiras de Programar Melhor em Python, de Brett Slatkin -, grandes fóruns de programação como Stack Overflow e comunidades de Python nas redes sociais Reddit e Telegram e as próprias documentações oficiais, tanto das bibliotecas de extração de texto e imagens, quanto do próprio Python.

Para possibilitar uma melhor interação entre o usuário e o conversor, foi utilizado o Framework de desenvolvimento web em Python, Django; as linguagens de marcação de hipertexto HTML e de estilização de páginas CSS para a criação de interfaces, disponibilizando assim o conversor em forma de website.

## RESULTADOS E DISCUSSÃO

E A biblioteca escolhida para extração dos textos nos PDF's foi a PyPDF2, uma biblioteca gratuita e de código aberto, que contém um catálogo robusto de ferramentas para manipulação de PDF's e uma comunidade ativa nos fóruns de programação.

Por meio da utilização do Framework Django e as linguagens HTML e CSS, foi montada uma aplicação simples e intuitiva, na qual o usuário realiza tudo através de etapas. Na primeira etapa é realizado o upload do arquivo PDF, que fica armazenado temporariamente; na segunda, o usuário digita o intervalo das páginas em que estão as questões a serem extraídas; na terceira, o usuário indica qual é a estrutura de formatação utilizada para caracterizar as questões, por exemplo 1-, 1) ou 1. e então o texto do intervalo indicado é extraído em forma de questões, que são separadas e salvas na memória; a quarta etapa é a listagem, deleção e edição de pequenos erros que estão sujeitos a ocorrer durante a extração.

## CONSIDERAÇÕES FINAIS

A Em meio à toda informação adquirida por meio de pesquisa e testes, o maior desafio se encontra em manter o alinhamento dos textos originais após a extração, assim como também conservar as imagens, que são extraídas separadamente dos textos.

Com a realização deste estudo, espera-se contribuir com futuras pesquisas na área de manipulação de PDF's, assim como no desenvolvimento e/ou aprimoramento de ferramentas com esta mesma finalidade ou similares.

Já com o desenvolvimento desta aplicação, almeja-se atender a necessidade dos professores em possuir uma forma mais simples e dinâmica de criar questionários e extrair questões das mais diversas fontes e em cascata, contribuir com a problemática do treinamento de especialistas em exames de imagem.

## REFERÊNCIAS BIBLIOGRÁFICAS

- RODRIGUES, Ricardo C. Análise e tematização da imagem fotográfica. *CI.Inf.*, Brasília, v. 36, n. 3, p. 67-76. 2007.
- GIGER, Maryellen L. Machine Learning in Medical Imaging. *Journal of the American College of Radiology* Volume 15, Edição 3, Parte B, Páginas 512-520. Março de 2018.
- MORGADO, Manoel do Vale; AMES, Romeu F. Maia; SILVESTRE, Larissa Jácome Barros. Aplicativos móveis na medicina: um estudo das tecnologias e sua relevância no processo de aprendizagem do aluno. *Revista de Teorias e Práticas Educacionais - RTPE* vol.25, n.1,pp.10-15 (Out - Dez 2019).
- CALIARI, Ketter V. Z.; ZILBER, Moisés A. e PEREZ, Gilberto. Information and Communication Technologies in higher education (face-to-face teaching): An analysis of the variables that influence its adoption. *REGE - Revista de Gestão*. Universidade Presbiteriana Mackenzie, São Paulo, SP, Brasil. v. 24, n.1, p. 247-255. 2017.
- PEZZUOL, Jorge L.; ABREU, Francisco D. L.; SILVA, Simone M.; TENDOLINI, Aline; BISSACO, Márcia A. e RODRIGUES, Silvia C. M. Virtual setting for training in interpreting mammography images. In. *SPIE Medical Imaging, 10138, 2017, Londres. Imaging Informatics for Healthcare, Research, and Applications, Orlando, Flórida - Estados Unidos: SPIE. 2017. Disponível em: < <http://dx.doi.org/10.1117/12.2254677>> Acesso em: 2017.*
- DUNLOSKY, John; RAWSON, Katherine A; MARSH, Elizabeth J; NATHAN, Mitchell J e WILLINGHAM, Daniel T. Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest*, 2013.

ROMAN, Cassiela; ELLWANGER, Juliana; BECKER, Gabriela C.; SILVEIRA, Anderson D.; MACHADO, Carmen L. B. e MANFROI, Waldomiro C. Active Teaching-Learning Methodologies in the Teaching Health process in Brazil: A Narrative Review. Clin Biomed Res. v. 37, n. 4, p. 349-357. 2017.

BORGES, Luiz Eduardo. Python para Desenvolvedores. 3 ed. São Paulo: Novatec, 2015.