Título:	**Automação da extração e estruturação em FHIR de dados de Fichas de Atendimento Multiprofissional para simulação do preenchimento de requisições de mamografia no SISCAN**		
Autores:	Leonardo Antonio Patatt Murilo Antonio Wollmann Daniela Duarte da Silva Bagatini Priscila Schmidt Lora Gabriela de Andrade Monteiro Ana Paula Müller Adriano de Oliveira Martins Débora Oliveira da Silva Cristiane Drebes Pedron		
Área	[ ] Humanas [ ] Sociais Aplicadas [ ] Biológicas e da Saúde [x] Exatas, da Terra e Engenharias	Dimensão:	[ ] Ensino [x] Pesquisa [ ] Extensão [ ] Inovação

Resumo:
A Ficha de Atendimento Multiprofissional (FAM) serve de insumo para o preenchimento da Requisição de Mamografia de Rastreamento e Diagnóstica no SISCAN. Entretanto, o processo manual é demorado, repetitivo e suscetível a erros, dado o volume de documentos gerados na rotina clínica. A ausência de padronização das fichas dificulta a automação desse preenchimento, exigindo soluções capazes de reconhecer variações na escrita e adaptar-se a diferentes formatos de entrada. O objetivo deste trabalho foi desenvolver um sistema de automação capaz de extrair os dados da FAM, estruturá-los segundo o padrão FHIR e utilizá-los na integração e no preenchimento automatizado de um formulário que simula a requisição de exames de mamografia no SISCAN. O método envolveu: (1) uso da biblioteca PyMuPDF (módulo Fitz) para leitura e extração do conteúdo textual dos PDFs da FAM; (2) desenvolvimento de expressões regulares (Regex) para reconhecer os campos mesmo com pequenas diferenças nos rótulos (ex.: "CPF:" ou "CPF"), acentuação ou espaçamento; (3) extração, organização e limpeza dos valores encontrados, removendo inconsistências; (4) criação de uma tabela CSV relacionando cada campo da FAM ao campo correspondente na Requisição do SISCAN, classificados em mapeados (presentes nos dois documentos), não mapeados (presentes em apenas um) e calculados automaticamente pelo SISCAN; (5) estudo da estrutura da Requisição de Mamografia com base na documentação pública do SISCAN; (6) criação de um formulário simulado no Google Forms para representar e testar o preenchimento; (7) conversão dos dados extraídos para um Bundle FHIR R5, serializado em JSON; e (8) automação do preenchimento do formulário simulado com Playwright. Foram processadas 33 FAMs, com 100% de sucesso na extração, parsing e geração dos Bundles FHIR R5. A acurácia foi calculada apenas sobre os campos efetivamente presentes na FAM, distinguindo campo ausente ou vazio no

documento original, o que não representa erro do extrator, e sim indisponibilidade da informação na fonte, de campo presente na FAM mas não localizado pelo sistema, o que caracteriza falha de extração. Dos 528 campos avaliados, observou-se 100% de acurácia, sem omissões ou incorreções, e os 462 campos originalmente vazios foram preservados sem preenchimentos indevidos. Como validação adicional, foi realizada conferência manual dos valores e dos campos faltantes nas 33 fichas. O tempo médio de processamento foi de 0,05 s por documento (P95 = 0,07 s). O sistema demonstrou capacidade de extrair corretamente dados de exames requisitados e prontuários, além de metadados como usuário, data de impressão e responsável pelo registro e atendimento, mantendo os campos vazios sem preenchimento. Regras específicas foram aplicadas para separar campos próximos ou situados na mesma linha, evitando que a extração associasse o valor de um campo ao seguinte, especialmente em áreas organizadas em colunas. A organização do sistema em módulos independentes facilita sua manutenção e evolução. Conclui-se que o sistema apresenta uma solução viável para automatizar a extração e a estruturação de dados clínicos, contribuindo para o ODS 3 – Saúde e Bem-Estar. Para trabalhos futuros, pretende-se validar o sistema no formulário oficial do SISCAN, mediante acesso institucional, e ampliar o mapeamento para outros tipos de requisição e exame. Trabalho financiado pelo Edital Chamada Nº 21/2023 – Estudos Transdisciplinares em Saúde Coletiva.

Link do Vídeo: Mostra_BolsaCnpq_Unisc_VII_2026.mp4