

Proposta para a obtenção automática da base de metadados de citantes e citados da biblioteca SciELO

Raphael de Freitas Saldanha¹, Mônica S. Auler², Rosângela Cordeiro Assef³, Cicera Henrique da Silva⁴, Maria Cristina Soares Guimarães⁵, Rosane Abdala Lins⁶

¹0000-0003-0652-8466 – Fiocruz, Rio de Janeiro, Brasil. raphael.saldanha@icict.fiocruz.br

² 0000-0003-2937-0067 – Fiocruz, Instituto de Informação Científica e Tecnológica em Saúde, Rio de Janeiro, Brasil; e Fiocruz, Casa de Oswaldo Cruz, Rio de Janeiro, Brasil. monica.auler@gmail.com

³ 0000-0003-2296-0689 – Fiocruz, Instituto de Informação Científica e Tecnológica em Saúde, Rio de Janeiro, Brasil. rosangela.cordeiro@icict.fiocruz.br

⁴ 0000-0002-3298-795X – Fiocruz, Instituto de Informação Científica e Tecnológica em Saúde, Rio de Janeiro, Brasil. cicera.henrique@icict.fiocruz.br

⁵ 0000-0003-2717-381X – Fiocruz, Instituto de Informação Científica e Tecnológica em Saúde, Rio de Janeiro, Brasil. cristina.guimaraes@icict.fiocruz.br

⁶ 000-0002-7518-3265 – Fiocruz, Instituto de Informação Científica e Tecnológica em Saúde, Rio de Janeiro, Brasil. rosane.abdala@icict.fiocruz.br

RESUMO

Criada em 1996 com o objetivo de dar visibilidade e permitir a produção de indicadores da produção nacional, a Scientific Electronic Library (SciElo) se constitui como uma fonte de informação importante da produção científica no cenário científico nacional e internacional. Entretanto, a coleta dos metadados dos artigos citantes e respectivas referências citadas é uma tarefa trabalhosa, visto que não há como obtê-los com facilidade, de forma automática. Diante dessa realidade, a equipe do grupo de pesquisa autora deste estudo propôs-se a desenvolver, em acesso aberto, uma ferramenta que pudesse agilizar o desenvolvimento de estudos de citação. O objetivo deste trabalho é apresentar a metodologia utilizada no desenvolvimento dessa ferramenta, utilizando como protótipo um estudo exploratório do campo da história das ciências da saúde, como parte de uma dissertação de mestrado em andamento que busca conhecer o perfil da literatura citada nos artigos do periódico História, Ciências, Saúde: Manguinhos. Utilizando a linguagem R, foram criados *scripts* para varredura da API disponibilizada para acesso ao banco de dados existe na plataforma SciElo. A partir dessa varredura, foram recuperados os metadados dos artigos citantes em formato XML, contendo em cada arquivo XML um bloco dos dados com as referências citadas, ponto de partida para construção do banco de dados disponibilizado para o usuário. Por meio da tecnologia utilizada, foi possível a construção de um banco de dados contendo os metadados das coleções SciElo Saúde Pública e SciElo Brasil, com a indexação de mais de 370 periódicos. A recuperação e

filtragem dos dados por meio do sistema criado utilizando as tecnologias R + SQLite + Shiny apresenta resultados fidedignos aos encontrados manualmente nos periódicos.