

ESTUDO DE TÉCNICA PARA PREPARAÇÃO DE METADADOS DE REPOSITÓRIOS DIGITAIS¹

Leticia Ignácio Ribeiro¹, Divino Ignácio Ribeiro Junior²

¹ Bolsista vinculada ao projeto “Indexação Automática: Desenvolvimento De Uma Plataforma De Indexação De Acervos Textuais Digitais Para Repositórios Institucionais”

² Orientador, Departamento de Biblioteconomia – FAED – divino.ribeiro@udesc.br.

A pesquisa desenvolvida nesse último ano, se classifica como exploratória de abordagem bibliográfica, tendo por **objetivo geral** investigar as tecnologias conhecidas aplicáveis à indexação automática de documentos oriundos de repositórios institucionais (artigos científicos, teses, dissertações etc.) e os resultados de suas aplicações em plataforma experimental de indexação automática. inicial levantar e analisar leituras e documentos sobre tecnologias e softwares usados na realização da indexação automática (em Repositórios Digitais), assim como estudar essas referências e filtrá-las. E a partir dessa fundamentação, realizar testes de coleta de dados com a finalidade de desenvolver uma base de dados de testes para um repositório digital baseado na plataforma DSpace.

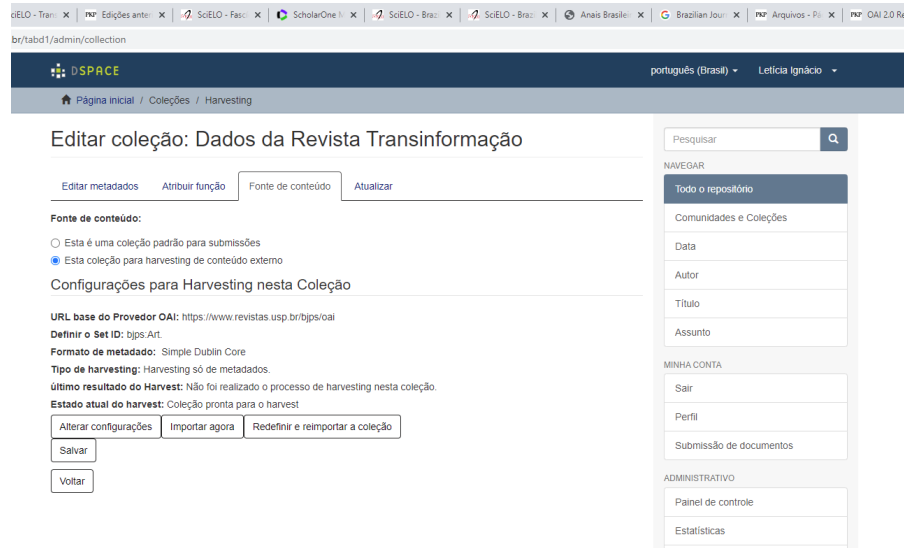
Procedimentos metodológicos: trata-se de pesquisa aplicada, com abordagem qualitativa, com procedimentos de laboratório, de natureza experimental.

Resultados: O processo de preparação da base de dados consistiu na coleta de metadados de periódicos científicos baseados na plataforma OJS, através do protocolo OAI-PMH, para posterior tratamento e montagem da base de dados. O processo se divide em três partes: a) na primeira, foi desenvolvido um processo para coleta de dados através do protocolo OAI/PMH para obter metadados de periódicos feitos com OJS para futura importação no Dspace. De forma resumida os passos dessa etapa de coleta são:

- Encontrar um periódico feito na plataforma OJS.
- Ao lado da URL do periódico, acrescentar “/OAI” e verificar se aparece a interface do provedor de dados OAI-PMH.
- No Dspace, criar uma coleção nova, colocar nome, pequena descrição e salvar.
- Ainda na coleção, ir na aba “fonte de conteúdo”, selecionar a opção “esta coleção é para harvesting de conteúdo externo” e salvar.
- Voltar ao provedor de dados e copiar o endereço “base URL” (clique no verbo “identify”)
- Copiar o endereço do provedor de dados e colar no campo “URL base do provedor OAI”
- Voltar ao provedor de dados, clique no verbo “list sets” e escolher um ID de set spec. Copiar o nome do se spec.
- Salvar os dados da coleção.

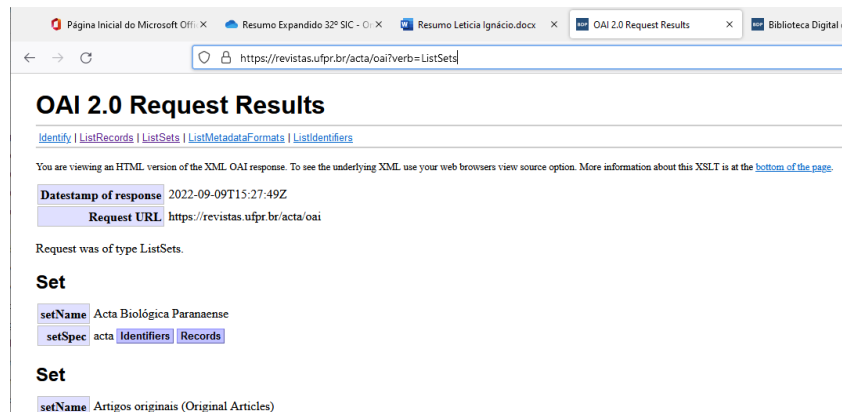
Abaixo a imagem da preparação de uma coleção no repositório experimental DSpace do LabTecGC. Para cada periódico foi montada uma coleção:

Figura 1: *Configuração da coleção para coleta de dados de artigos de periódicos baseados em OJS*



Para cada revista é criada uma coleção e importados apenas a seção de artigos de cada uma. Os dados do provedor OAI-PMH são informados no site de cada revista, conforme descrito na primeira etapa. A imagem a seguir ilustra uma dessas interfaces do provedor de dados:

Figura 2: *Interface do provedor de dados OAI-PMH*



Na segunda parte foi quando a coleta dos dados foi iniciada. Os dados são de revistas da área de saúde baseadas na plataforma OJS e carregadas para o Dspace do LabTecGC. E na terceira parte houve a preparação e correção dos dados obtidos, processo realizado com a exportação dos metadados coletados em formato CSV para produzir as correções em uma planilha Excel, para posterior importação para a base de testes.

Palavras-chave: OAI-PMH. Tratamento de Dados. Repositório Digital. Indexação automática.