

Avaliando a Integridade de um Repositório

David Marques
Emely Rossim
Jean-Rémi Bourguet
Elias de Oliveira

A identificação de objetos espúrios em um repositório é uma tarefa árdua. Dependendo do volume de itens no acervo da instituição, essa atividade pode demandar muito tempo dos especialistas na execução de um *procedimento do tipo censo*. Diante dessa dificuldade, pode-se adotar uma *metodologia de amostragem*, desde que o erro associado esteja dentro de um intervalo aceitável para os padrões de qualidade almejados pela instituição ou exigidos por normas específicas. Neste trabalho, propomos uma estratégia baseada em Estatística Bayesiana para reduzir o esforço do especialista no treinamento do algoritmo de inteligência artificial na identificação de objetos espúrios. Nossos experimentos mostram que é possível reduzir para cerca de 3.96% do tamanho do acervo, o esforço humano necessário para fornecer exemplos ao treinamento do algoritmo na identificação do restante do acervo.

Palavras-chave: Integridade de Repositórios; Estatística Bayesiana; Metodologia de Amostragem.

Linked Open Archives (LOA): Um hub para aplicar a Ontologia RiC e conectar acervos arquivísticos na web semântica

Jair Martins de Miranda
Alison Carlos Figueiras
Eusébia de Fátima Diniz

O modelo conceitual Records in Contexts Conceptual Model (RiC-CM) e a ontologia Record-in-Context (RiC-O) foram lançados em novembro de 2023 pelo *Experts Group on Archival Description do International Council on Archives (EGAD/ICA)*. Tal feito apresenta mudanças paradigmáticas que influenciam diretamente os processos de descrição, recuperação e gestão dos arquivos digitais, uma vez que extrapolam os limites estáticos impostos dos padrões estabelecidos (ISAD, ISAAR, ISDF e ISDIAH). Esse novo paradigma, segundo o EGAD/ICA, oferece melhorias consideráveis na descrição e no gerenciamento de arquivos e registros em ambientes digitais, além de facilitar a interoperabilidade e aumentar as capacidades de recuperação da informação por meio de consultas SPARQL. O presente artigo visa apresentar o protótipo da plataforma digital Linked Open Archives (LOA), cujo principal objetivo é aplicar experimentalmente a nova ontologia Records in Contexts (RiC-O 1.0), conectando acervos do Projeto Diáspora Africana no Brasil. Conclui que as Ontologias, a Websemântica e as tecnologias Linked Open Data (LOD) trarão grande impacto nos princípios, nas práticas arquivísticas e no acesso aos arquivos.

Palavras-chave: Records in Contexts; Ontologia; Linked Open Data (LOD); Websemântica; Diáspora Africana no Brasil