

INDEXAÇÃO E RECUPERAÇÃO SEMÂNTICA DA INFORMAÇÃO EM REPOSITÓRIOS INSTITUCIONAIS

RAIMUNDA FERNANDA DOS SANTOS - UNIVERSIDADE FEDERAL DO RIO DE JANEIRO/UNIVERSIDADE ESTADUAL DE LONDRINA, E-MAIL: FERNANDA@FACC.UFRJ.BR, ORCID: 0000-0002-7750-3269

THAMIRES NASCIMENTO DE OLIVEIRA - UNIVERSIDADE FEDERAL DO RIO DE JANEIRO, E-MAIL: THAMIRESOLI91@GMAIL.COM, ORCID: 0000-0002-3791-5875

Introdução: Os objetos informacionais devem ser representados a partir de um conjunto de operações que contemplam a descrição física e a representação do seu conteúdo para fins de recuperação, acesso e uso. No atual contexto dos repositórios institucionais, os metadados temáticos dos objetos informacionais (teses, dissertações, livros, artigos de periódico etc.) referem-se aos campos específicos para a descrição dos seus assuntos, que podem ser representados através da linguagem natural (palavras-chave) ou de instrumentos de controle terminológico (listas de cabeçalhos de assunto, tesouros, taxonomias, por exemplo). Esses assuntos podem ser fornecidos pelo próprio autor do documento, pelo catalogador/indexador ou ainda podem ser provenientes de um processo de indexação colaborativa integrando as duas modalidades supracitadas a partir de um sistema híbrido de representação da informação.

Objetivo: delinear recomendações para a melhoria da indexação e recuperação semântica da informação em Repositórios Institucionais a partir das práticas já realizadas nesses ambientes.

Metodologia: pesquisas bibliográfica, documental, exploratória, descritiva com abordagem quali-quantitativa. Foi realizado também um levantamento das Universidades Federais brasileiras através dos dados disponibilizados no site do MEC (Ministério da Educação). Posteriormente, foi averiguada a existência dos Repositórios Institucionais destas Universidades, mediante busca nos sites institucionais das Universidades, no site do IBICT e no diretório global OpenDoar. Além disso, foi feita a navegação nos menus e submenus dos Repositórios Institucionais, objetivando recuperar documentos como Política de Indexação, manuais, tutoriais de busca, entre outros. A coleta desses dados foi realizada entre os meses de março e junho de 2021, com revisão de março a junho de 2022.

Resultados: Dos 53 (cinquenta e três) Repositórios Institucionais existentes e ativos recuperados, apenas 10 (dez) possuem orientações relacionadas à indexação registradas no sistema, o que equivale a 18,86% dos sistemas analisados. Como sugestões para melhorias das práticas de representação nos Repositórios Institucionais analisados, recomenda-se: a) elaborar (caso não exista) uma política de indexação para o sistema com diretrizes e normas que contemplem as características das coleções da instituição mantenedora, seus usuários, seus recursos e todo o contexto relacionado ao Sistema; b) traduzir e tornar pública a política de indexação em questão em formato adequado à linguagem dos usuários; c) o não uso de termos gerais ligados por hífen a termos específicos para subdividir assuntos, para fins de adaptação à linguagem adotada pelo software; d) criação de uma nuvem de tags na tela inicial de cada repositório, com vistas a melhorar a experiência de navegação dos usuários; e) acompanhamento das atualizações do software e, de maneira padronizada, implementar a última versão do DSpace; f) uso de Machine Learning (aprendizagem de máquina) para a criação, implementação e atualização de vocabulários controlados (como ontologias, taxonomias e tesouros) em repositórios institucionais.

Referências bibliográficas

- Dias, E. W., & Naves, M. M. L. (2007). *Análise de assunto: teoria e prática*. Thesaurus.
- Oliveira, T. N., & Santos, R. F. dos. (2022). Políticas e diretrizes de indexação em Repositórios Institucionais das Universidades Federais brasileiras. *Revista Informação Na Sociedade Contemporânea*, 6(1), e29444.
- Santos, R. F. dos. (2017). Indexação em repositórios digitais: uma abordagem sobre o metadado assunto da Biblioteca Digital de Monografias da UFRN. *Revista Informação Na Sociedade Contemporânea*, 1, 1-22. <https://doi.org/10.21680/2447-0198.2017v1n01D12279>.