

A relevância dos metadados: das infraestruturas à monitorização do Acesso Aberto

Ana Inácio

Pós-Graduada em Ciências Documentais pela Universidade Autónoma de Lisboa (UAL) - Lisboa, Portugal.
Bibliotecária atuando nos Serviços de Informação e Documentação do Instituto Univesitário de Lisboa (ISCTE) – Portugal.

<https://ciencia.iscte-iul.pt/authors/author-public-page-1796/cv>

<http://orcid.org/0000-0002-6993-9089>

E-mail: ana.inacio@iscte-iul.pt

João Dias

Licenciatura em Antropologia pela Universidade Nova de Lisboa (UNL) - Lisboa, Portugal
Gestor de Ciência na Universidade Católica Portuguesa - Biblioteca João Paulo II (UCP) - Universidade Católica Portuguesa - Portugal

<http://orcid.org/0000-0002-0751-873X>

E-mail: joao.dias@ucp.pt

Maria João Amante

Doutora em Documentação pela Universidade de Alcalá (UAH), Alcalá de Henares - Espanha
Diretora do Instituto Univesitário de Lisboa (ISCTE-IUL) – Portugal.

<http://www.degois.pt/visualizador/curriculum.jsp?key=9072561651658632>

<http://orcid.org/0000-0001-8891-9094>

E-mail: maria.amante@iscte-iul.pt

Data de submissão: 09/09/2019. Data de aprovação no ConfOA: 11/06/2019. Data de publicação:

RESUMO

O estudo testa as potencialidades oferecidas pelas plataformas de metadados abertos na monitorização do estado do Acesso Aberto (AA), caracterizando parte da produção científica do ISCTE-IUL, entre 2014 e 2018, no que se refere às dimensões selecionadas. Possibilita igualmente o levantamento das dificuldades e/ou limitações de interoperabilidade entre diferentes sistemas de informação científica. Utiliza-se como fonte de informação o sistema CRIS (Current Research Information System) Ciência-IUL, isolando o conjunto de artigos científicos produzidos no período referido (n=5991) e trabalhando as seguintes características: tipo de localização, tipo de licença associada, tipo de versão depositada e APC (Article Processing Charges) envolvidos. Recorre-se a duas plataformas: o agregador Unpaywall e o diretório DOAJ (Directory of Open Access Journals). No primeiro, a ligação das bases faz-se através do DOI (Digital Object Identifier), tendo sido recuperados 14,1% dos artigos científicos, aos quais se aplica tratamento estatístico descritivo. No caso do diretório DOAJ, considera-se o ISSN (International Standard Serial Number) e/ou e-ISSN para a verificação do tipo de licença e apuramento de valor de APC, correspondendo a informação a 10,2% do total. Os resultados do estudo reforçam a relevância das plataformas de metadados abertos, mas aconselham prudência no uso exclusivo delas para monitorizar o estado do AA. Apesar de permitirem a utilização de metadados a todos os que para elas contribuem, de serem desnecessários conhecimentos técnicos para a exportação de dados e poderem servir como pontos de partida para o desenvolvimento de novas metodologias de monitorização, a diferença de campos que apresentam dificulta a comparação e o seu cariz agregador faz depender a qualidade da informação da qualidade dos metadados das fontes (e.g. Unpaywall). Por outro lado, a informação de retorno é parcial, quer porque nem sempre existe no CRIS a informação que permite a ligação a essas bases de dados (e.g. DOI), quer porque também nessas plataformas não se encontra a totalidade dos registos pretendidos (e.g. DOAJ). A inexistência de outras ferramentas com metadados idênticos restringe a possibilidade de corroborar alguns dos resultados obtidos (e.g. APC).

Palavras-chave: Acesso Aberto. Interoperabilidade. CRIS. Plataformas de Metadados Abertos.

Metadata relevance: from infrastructures to Open Access monitoring

ABSTRACT

This study tests the potential of open metadata platforms to monitor Open Access (OA) at ISCTE-IUL, considering part of the institution's scientific production (2004-2018) and selected dimensions. It allows also identifying some of the difficulties and/or limitations related with the interoperability between diverse information management systems. ISCTE-IUL CRIS (Current Research Information System) Ciência-IUL is the information source for this study. All the scientific articles produced in the period under analysis were selected ((n=5991) and the following characteristics were analyzed: localization type, associated license type, deposited version type and related APC (article Processing Charges). To facilitate the analysis, Unpaywall aggregator and DOAJ (Directory of Open Access Journals) are the most suitable tools. In the first one, the databases connection is established through the DOI (Digital Object Identifier). A total of 14,1% scientific articles was retrieved. These articles were subject to a descriptive statistical analysis. Regarding the DOAJ, ISSN and/or EISSN are used in order to verify the license type and the cost of APC. The information retrieved represents 10,2% of the total number of articles. The results obtained reinforce the importance and relevance of open metadata platforms but some caution must be observed when using these tools to monitor Open Access practices and policies. Despite allowing the use of metadata, not requiring technical knowledge to export data, and allowing its use as a starting point to develop new tracking methodologies there is a major problem that needs to be considered. The existing difference on the fields shown is an obstacle to any comparison and acting as an aggregator, the retrieved information quality depends on the sources metadata quality (e.g. Unpaywall). On the other hand, the obtained feedback is partial, either because not always the information that allows the connection with the databases exists in the CRIS System (e.g. DOI) or because on those platforms the overall intended articles does not exist (e.g. DOAJ). The inexistence of other tools with identical metadata limits the possibility to corroborate some of the obtained results (e.g. APC).

Keywords: Open Access, Interoperability, CRIS, Open Metadata Platforms

La relevancia de los metadatos: de las infraestructuras a la monitorización del Acceso Abierto

RESUMEN

En este trabajo son testadas las potencialidades ofrecidas por las plataformas de metadatos abiertos para la monitorización del Acceso Abierto (AA), para caracterizar parte de la producción científica del año 2014 al 2018, tomando en consideración un conjunto de dimensiones previamente seleccionadas. Igualmente permite la identificación de las dificultades y/o limitaciones de interoperabilidad entre distintos sistemas de información científica. Como fuente de información hemos utilizado el sistema CRIS (Current Research Information System) Ciência-IUL, a partir del cual seleccionamos un conjunto de artículos científicos producidos en el período de tiempo anteriormente señalado (n=5991) y analizando las siguientes características: Tipo de localización, tipo de licencia asociada, tipo de versión depositada e APC (Article Processing Charges) asociados. Para tal, utilizamos dos plataformas: el integrador Unpaywall y el directorio DOAJ. En el primero, la conexión entre las bases de datos se hace a través el DOI (Digital Object Identifier). Fueron recuperados 14,1% de los artículos científicos, a los cuales fue aplicado un tratamiento estadístico descriptivo. En el caso del directorio DOAJ, se considera el ISSN y/o EISSN para confirmación del tipo de licencia y ultimación del coste del APC, correspondiendo la información al 10,2% del total de artículos. Los resultados del estudio refuerzan la relevancia de las plataformas de metadatos abiertos. Sin embargo, los mismos resultados nos indican la necesidad de alguna prudencia en el uso exclusivo de las mismas con el objetivo de monitorizar el estado del Acceso Abierto. Mientras permitieren la utilización de metadatos a todos que para ellas contribuyan, de no existir la necesidad conocimientos para la exportación de los datos e de ser posible su utilización como puntos de partida para el desarrollo de nuevas metodologías de monitorización, la diferencia de campos que presentan dificulta la comparación. Además, su dimensión integradora haz depender cualquier la calidad de la información de la calidad de los metadatos de las fuentes (e.g. Unpaywall). Por otro lado, la información de feedback obtenida es parcial: en el CRIIS ni siempre existe la información que permite la conexión a las bases de datos (e.g. DOI), en las plataformas de metadatos abiertos ni siempre existe na totalidad de los registros deseados (e.g. DOAJ). La inexistencia de otras herramientas con metadatos idénticos limita la posibilidad de confirmar algunos de los resultados obtenidos (e.g. APC).

Palabras-clave: Acceso Abierto, CRIS, Plataformas de Datos Abiertos