

Repositórios de acesso aberto brasileiros: características, crescimento e possibilidades futuras

Raphael Figueiredo Xavier¹

¹ Bibliotecário. Fundação Getúlio Vargas – FGV. Email: raphael.xavier@fgv.br

Resumo

Este trabalho, adequado ao eixo temático "Repositórios Institucionais e Temáticos", busca avaliar a evolução de repositórios brasileiros entre 2011 e início de 2019 e com isso proporcionar uma visão global da infraestrutura destas plataformas. O surgimento de repositórios está diretamente ligado a popularização de tecnologias de informação e comunicação e seu impacto no processo da comunicação científica. Com o avanço do movimento de políticas de acesso aberto em todo o mundo, os repositórios tiveram um desenvolvimento relevante e significativo na última década. Através do diretório OpenDOAR, pretende-se verificar a quantidade de repositórios brasileiros, o idioma do conteúdo, visão geral dos softwares utilizados nas plataformas, assunto e conteúdo das coleções e o crescimento no período de 2011 a 2019. Com isso, é possível considerarmos trajetórias futuras no desenvolvimento de repositórios como a possibilidade de infraestrutura distribuída globalmente e serviços de valor agregado, transformando o sistema mais centrado em pesquisa, aberto e apoiando a inovação, enquanto, ao mesmo tempo, é coletivamente gerenciado pela comunidade acadêmica.

Palavras-chave: Repositórios Digitais; Open-Access; Movimento de acesso livre a informação; OpenDOAR

Abstract

This work, appropriate to the thematic axis "Institutional and Thematic Repositories", seeks to evaluate the evolution of Brazilian repositories between 2011 and early 2019 and thus provide a global view of the infrastructure of these platforms. The emergence of repositories is directly linked to the popularization of information and communication technologies and their impact on the process of scientific communication. With the advancement of open access policies worldwide, repositories have developed relevant and significantly over the last decade. Through the OpenDoar directory, we intend to check the number of Brazilian repositories, content language, an overview of the software used in these platforms, subject and

content of the collections and growth in the period from 2011 to 2019. With this, it is possible to consider possible future paths in the development of repositories such as the possibility of globally distributed infrastructure and value-added services, transforming the system more focused on research, open and supported innovation, while at the same time collectively managed by the academic community.

Keywords: Digital repositories; Open-Access; OpenDOAR

Introdução

Repositórios institucionais podem ser descritos como coleções digitais que coletam e preservam a produção intelectual de uma comunidade específica (PINFIELD, 2009). Estes repositórios armazenam e tornam acessíveis o patrimônio educacional, de pesquisa e associado a uma instituição. O surgimento destas coleções está diretamente atrelado ao desenvolvimento de tecnologias de informação e comunicação e seu impacto no processo de comunicação científica. Com o avanço do movimento de políticas de acesso aberto em todo mundo (Baseada na declaração de Budapeste, em 2001), os repositórios tiveram um desenvolvimento relevante e significativo na última década.

Diante deste fato, esta pesquisa busca avaliar a evolução de Repositórios brasileiros entre 2011 e início de 2019 e com isso proporcionar uma visão holística da infraestrutura destas plataformas. Os dados para esta pesquisa são de origem do OpenDOAR, um diretório que reúne iniciativas de acesso aberto e monitoram o crescimento do movimento de acesso aberto à informação científica no mundo através do registro de repositórios acadêmicos Open-Access. Nesse sentido, implica, por um lado, abordagem teórica sobre a temática acerca dos conceitos basilares e, de outro, abordagem exploratória e comparativa no que tange o objeto empírico: os repositórios brasileiros registrados no OpenDOAR.

Com isto, além de delinear uma progressão histórica e as características atuais da situação dos repositórios, a pesquisa traz discussões acerca de possíveis trajetórias futuras no desenvolvimento de repositórios como a possibilidade de infraestrutura distribuída globalmente e serviços de valor agregado, transformando o sistema mais centrado em pesquisa, aberto e apoiando a inovação e, ao mesmo tempo, coletivamente gerenciado pela comunidade acadêmica.

Definições

Inicialmente, é importante estabelecer algumas definições de conceitos-chave, particularmente 'Acesso aberto a Informação' e 'Repositórios', por serem fundamentos bases para esta pesquisa.

O Acesso aberto a informação é respectivamente o resultado:

(1) de uma reação dos pesquisadores ao modelo de negócios de editoras comerciais de revistas científicas (e seus preços cada vez mais altos de

assinatura); e da (2) crescente conscientização do aumento de impacto provocado pela disponibilização de documentos científicos livres de barreiras ao acesso. O mote do movimento mundial em favor do Acesso Livre a resultados de pesquisa, portanto, é a disseminação ampla e irrestrita dos resultados de pesquisas financiadas com recursos públicos (BAPTISTA et al, 2007, p.2).

No início do século XXI a Open Society Institute (OSI) promoveu uma reunião em Budapeste que resultou, provavelmente em um dos mais relevantes documentos no marco das iniciativas do movimento de Acesso Livre. O documento ficou conhecido como Budapest Open Access Initiative (BOAI)¹. A BOAI estabeleceu ali duas estratégias práticas para iniciativa Open-Access: uma seria através de revistas de acesso livre, sem restrição de acesso e sem cobrança de assinatura ou taxas de acesso; outra seria pelo auto arquivamento, através dos autores, de artigos publicados nas revistas científicas em repositórios, disciplinares ou institucionais. A primeira foi denominada de via dourada e a segunda, via verde. (HOUGHTON; SWAN, 2013)

Ambos mecanismos possuem larga literatura discutindo suas especificidades e efeitos. Neste ano, o Directory of Open Access Journals – DOAJ registrou 12.932 revistas de Acesso livre que cumprem diretrizes da via dourada², enquanto o OpenDOAR tem inscritos em março de 2019 cerca de 4.000 Repositórios Institucionais operando na via verde³. Pinfield (2009) nos recorda que esses dois caminhos são vias paralelas distintas, alternativas e não complementares e que defensores de ambos os lados tem apresentado discussões na literatura da área.

O que é perceptível é que a discussão acerca de repositórios está vinculada diretamente ao seu papel de suporte à Via Verde (PINFIELD, 2009; BJÖRK ET AL., 2014; HOUGHTON e SWAN, 2013) e os Repositórios incluídos neste estudo estão neste escopo. Pinfield (2009, p.165, tradução nossa) define os Repositórios como:

Um conjunto de sistemas e serviços que facilita a ingestão, o armazenamento, o gerenciamento, a recuperação, a disponibilização e a reutilização de objetos digitais. Repositórios podem ser criados por instituições, comunidades, financiadores de pesquisa ou outros grupos. Podem fornecer acesso a uma variedade de objetos digitais, incluindo artigos de periódicos avaliado por pares, capítulos de livros, teses, conjuntos de dados, objetos de aprendizado ou arquivos de mídia. (TRADUÇÃO NOSSA)⁴

¹ A declaração original pode ser lida em: <https://www.budapestopenaccessinitiative.org/read>

² Acessar a página oficial do DOAJ. Disponível em <https://doaj.org/>. Acesso em: 01/03/2019

³ Acessar a página oficial do OpenDOAR. Disponível em <https://v2.sherpa.ac.uk/pendoar/>. Acesso em: 01/03/2019

⁴ As a set of systems and services that facilitates the ingest, storage, management, retrieval, display, and reuse of digital objects. Repositories may be set up by institutions, subject communities, research funders, or other groups. They may provide access to a variety of digital objects, including peer-reviewed journal articles, book chapters, theses, datasets, learning objects, or rich media files.

Distinções, tipologias e características são muito bem colocadas em Heery e Anderson (2005). Aqui, cabe destacar que o presente artigo analisa toda população de repositórios brasileiros de acesso aberto através do OpenDOAR, um diretório de repositórios de acesso aberto baseado na Universidade de Nottingham, Reino Unido, realizado inicialmente em parceria com a Universidade de Lund, na Suécia. O OpenDOAR foi criado em 2005 para criar “um diretório oficial de repositórios acadêmicos de acesso aberto” em todo o mundo (OPENDOAR). Seus registros são periodicamente revisados e atualizados em seu banco de dados oferecendo um escopo atualizado do cenário de repositórios em todo mundo.

O serviço OpenDOAR é, atualmente, usado como fonte de dados para pesquisas de acesso aberto e diferentes projetos de monitoramento dos avanços do Open Access (projetos como o BASE, CORE, Open Access Tracker, OpenAIRE, OpenDepot, Repository 66 e ROAR).

Resultados

Crescimento dos repositórios

No geral, o número total de repositórios brasileiros no OpenDOAR cresceu de 28 em janeiro de 2011 para 109 em abril de 2019 (conforme mostra a figura 1). Isso representa um aumento de 289% durante o período. No entanto, a taxa de crescimento não foi inteiramente consistente e equilibrada durante esse período: houve crescimento muito acentuado em 2011 (de 28 para 56 repositórios entre janeiro e dezembro) e um período estagnado de crescimento em todo o ano 2018.



Figura 1: Evolução de repositórios brasileiros cadastrados no OpenDOAR

Repositórios Disciplinares tinham 28,6% (8 repositórios), seguidos por 3 Governamentais (10,7%) e 1 Agregador (3,6%).

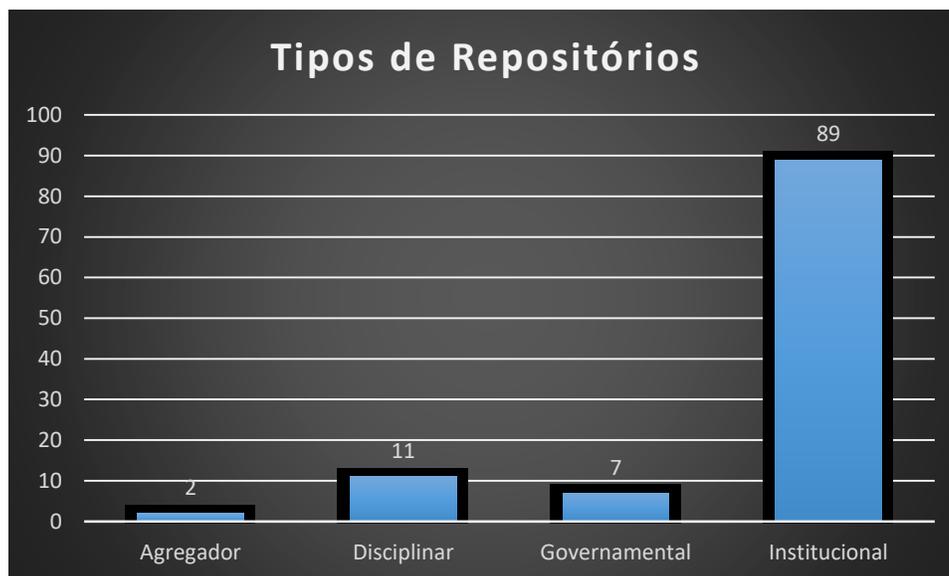


Figura 4: Classificação de tipos de Repositórios brasileiros pelo OpenDOAR. 2011-2019

Em se tratando de cobertura de assunto, o registro do mês de abril de 2019 coloca praticamente metade dos repositórios como multidisciplinares, o que não é surpresa, visto o domínio em quantidade dos repositórios institucionais. A Tabela 1 mostra que Repositórios multidisciplinares representam 69 (47%) dos repositórios brasileiros.

Onde os repositórios têm uma cobertura de assunto identificável específica (em vez de ser multidisciplinar), os assuntos são registrados no OpenDOAR. Há uma grande variedade de assuntos listados sem que qualquer área temática principal esteja dominando: todos estão abaixo de 10% do número total de repositórios.

Assunto	Frequência	Porcentagem
Multidisciplinar	69	47%
Saúde e Medicina	8	5%
Direito e Política	8	5%
Ciências Gerais	7	5%
Educação	6	4%
Computação e TI	5	3%
Biblioteconomia e Ciência da Informação	4	3%
Administração e Planejamento	4	3%

Física e Astronomia	4	3%
Tecnologia	3	2%
Negócios e Economia	3	2%
Agricultura, Alimentos e Veterinária	3	2%
Química e Tecnologia Química	3	2%
Artes e Humanidades	2	1%
Geografia e Estudos Regionais	2	1%
História e Arqueologia	2	1%
Ciências Sociais	2	1%
Psicologia	2	1%
Geologia e Ciências Planetárias	2	1%
Arquitetura	1	1%
Engenharia Civil	1	1%
Engenharia Elétrica e Eletrônica	1	1%
Engenharia Mecânica e de Materiais	1	1%
Belas Artes e Artes Cênicas	1	1%
Linguagem e literatura	1	1%
Biologia e Bioquímica	1	1%
Ecologia e Meio Ambiente	1	1%
Matemática e Estatística	1	1%

Tabela 1: Cobertura de assunto em Repositórios Brasileiros

No que se refere a tipologia de materiais em repositórios, a figura 5 nos mostra que os dois tipos de conteúdo mais comum para os repositórios brasileiros listados no OpenDOAR são teses e dissertações e artigos científicos: cerca de 63% dos repositórios incluem estes dois tipos de conteúdo. Outros tipos de conteúdo que se destacam são capítulos de livros (presente em 41% dos repositórios) e anais de conferências (29%). Foram identificados outros tipos, como formatos específicos, incluindo material didático, bibliografias, materiais audiovisuais, dentre outros, em menor número de repositórios. A maioria dos repositórios inclui vários tipos de conteúdo que apoiam claramente o processo de comunicação acadêmica e institucional de várias maneiras. Essas proporções não mudaram desde 2011.

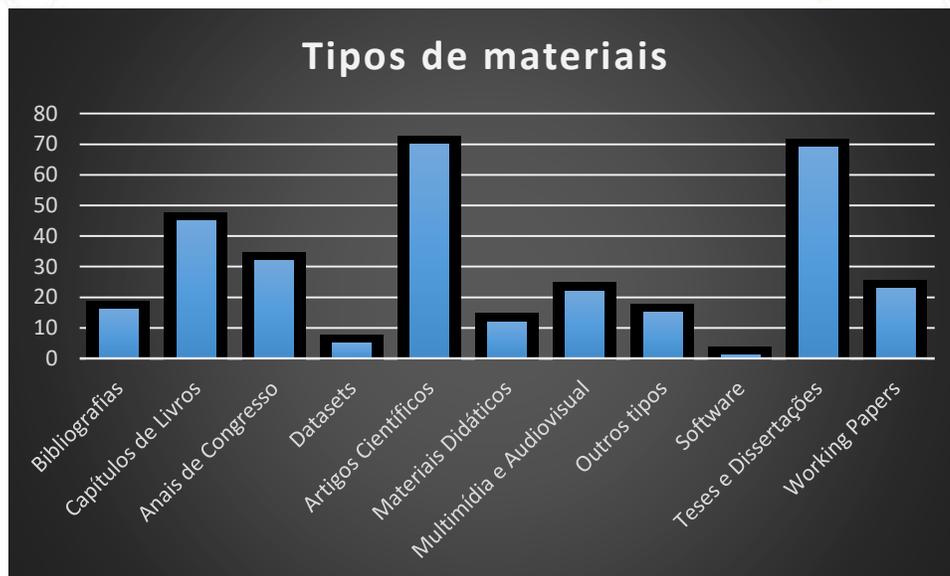


Figura 5: Tipos de materiais hospedados nos Repositórios Digitais Brasileiros

O idioma também é uma característica importante dos repositórios identificados no OpenDOAR. Em abril de 2019, os repositórios brasileiros listados continham recursos em três idiomas. A língua portuguesa é majoritária. 108 repositórios (99%) incluíam material em língua portuguesa (Tabela 2), naturalmente o predomínio do idioma se explica devido a língua nativa do país. O segundo idioma com mais ocorrência é o inglês (32; 29%), que reflete a posição da língua como a globalmente aceita em estudos, particularmente nos domínios da ciência, tecnologia e medicina (Tardy, 2004). Por fim, 16 repositórios (15%) foram listados com ocorrência do idioma espanhol.

Idioma	Total	Porcentagem
Inglês	32	29%
Espanhol	16	15%
Português	108	99%

Tabela 2: Idiomas presentes em Repositórios Brasileiros

Uma questão que merece atenção associada a repositórios é seu tamanho. Grandes repositórios contêm milhões de itens. Por exemplo, o inglês 'Europe PubMed Central' possui, de acordo com o OpenDOAR, 420.000.000 de registros. No entanto a grande maioria dos repositórios é relativamente pequena e isso está de acordo com o contexto brasileiro. O maior repositório brasileiro registrou 275.112 registros. Embora a média de registros em repositórios brasileiros seja 20.451, a Figura 6 nos mostra que a incidência de repositórios entre 100 e 1.000 registros são bem maiores do que em outros grupos. Os dados mostram, em termos simplificados, que há um número muito pequeno de repositórios grandes (ou seja,

com mais de 150.000) e um grande número de pequenos repositórios (com entre 100 e 25.000 itens) que compõem o cenário brasileiro de repositórios.

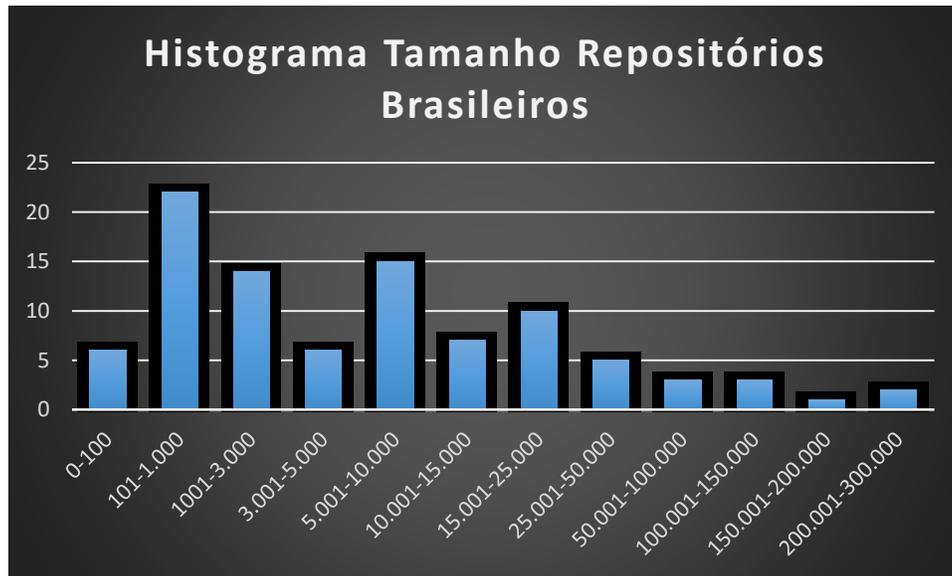


Figura 6: Histograma de frequência de quantidade de itens nos Repositórios Digitais Brasileiros

Ao analisar as plataformas de software utilizadas pelos repositórios, percebemos que o DSpace tem domínio majoritário do mercado e conseqüentemente permite aos repositórios a utilização de serviços de interoperabilidade com o protocolo OAI. Podemos verificar ainda uma quantidade considerável de soluções desenvolvidas localmente, que dificulta a identificação de seus recursos.

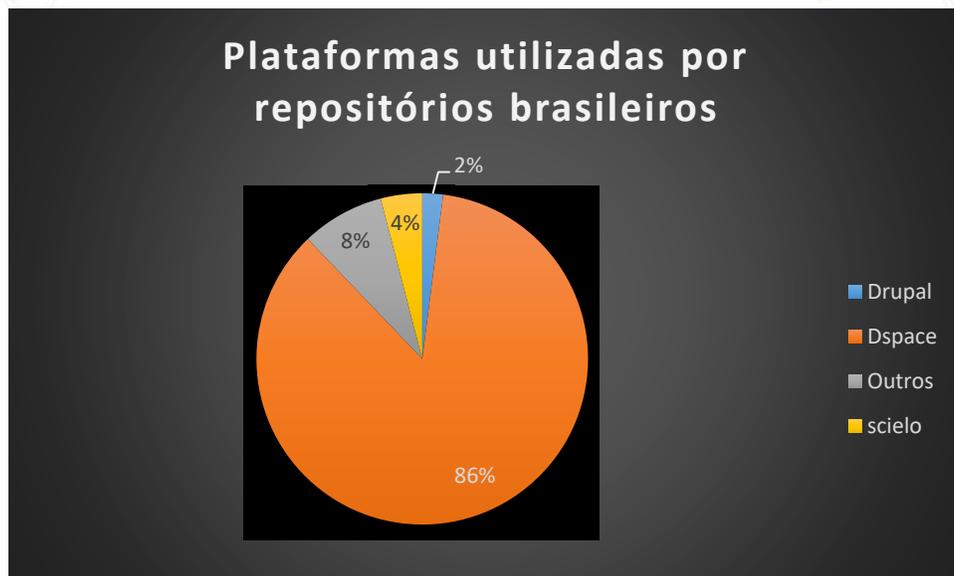


Figura 7: Plataformas de *softwares* utilizados por Repositórios Digitais Brasileiros

Considerações

Os dados apresentados mostram crescimento em repositórios brasileiros entre 2011 e início de 2019. O crescimento foi concentrado na região Sudeste, seguido pelas regiões Centro-Oeste (devido ao forte desempenho do Distrito Federal) e Sul. Somente mais recentemente, houve surgimento de repositórios na região Norte, que possui níveis comparativamente baixos de desenvolvimento de repositórios. Os maiores números de repositórios estão em São Paulo, Distrito Federal e Rio de Janeiro.

Repositórios são predominantemente institucionais em seu foco e multidisciplinar em sua cobertura. Eles ainda têm forte papel como biblioteca de teses e dissertações, porém já têm estabelecido guarda de artigos científicos e, são majoritariamente em português. Embora o tamanho dos repositórios seja difícil de avaliar com precisão, os dados disponíveis indicam que há um pequeno número de grandes repositórios e um grande número de pequenos repositórios. A configuração técnica mais comum para um repositório envolve o uso do software DSPACE (compatível com OAI de código aberto), mas, significativamente, muitas vezes não inclui o licenciamento formal de conteúdo.

Em contexto global, Xia (2012) traz considerações sobre tendências e características da difusão de Repositórios Digitais e Periódicos de acesso aberto. Ele identifica um número de facilitadores e barreiras à adoção que ajudam a explicar muitas das principais tendências apresentadas pelos repositórios brasileiros. Questões de infraestrutura tecnológica podem ser razoavelmente assumidas como barreiras significativas para o desenvolvimento de repositórios na região Norte, por exemplo.

Aparentemente ainda existe um potencial considerável para a continuação da expansão da infraestrutura do repositório no Brasil mesmo que o número de repositórios em determinadas regiões esteja mais próximo de um ponto de saturação. Por exemplo, em São Paulo, todas as grandes instituições de pesquisa agora têm repositórios e, portanto, é provável que o crescimento no número de repositórios futuro não seja o mesmo - embora os dados aqui apresentados ainda não sugiram que isso pode ter começado a acontecer. Isso, é claro, se aplica ao crescimento do número de repositórios em vez de crescimento em seus conteúdos. Como já foi demonstrado, para a grande maioria dos repositórios, permanece um potencial considerável para expandir o conteúdo.

É muito provável que as diferentes regiões continuem a avançar em velocidades diferentes, dependendo do complexo conjunto de fatores técnicos, culturais, gerenciais e econômicos. É importante que tais desenvolvimentos continuem a ser rastreados e monitorados.

Os dados do serviço OpenDOAR apresentados aqui fornecem uma perspectiva importante sobre o desenvolvimento da infraestrutura de repositórios brasileiros na última década. A partir destes resultados, a percepção é de que são necessárias mais pesquisas de análise e comparação com outras regiões do planeta para monitorar tendências de desenvolvimento e avaliar a importância de fatores que os influenciam. O trabalho evidencia a necessidade de fortalecimento de serviços de repositório, particularmente em torno da criação e disponibilização de políticas de reutilização de recursos.

Referências

BAPTISTA, Ana Alice et al. Comunicação científica: o papel da open archives initiative no contexto do acesso livre 10.5007/1518-2924.2007v12nesp1p1. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, p. 1-17, dez. 2007.

BJÖRK, Bo-christer et al. Anatomy of green open access. **Journal Of The Association For Information Science And Technology**, [s.l.], v. 65, n. 2, p.237-250, 6 nov. 2013.

HOUGHTON, John; SWAN, Alma. Planting the Green Seeds for a Golden Harvest: Comments and Clarifications on. **D-lib Magazine**, [s.l.], v. 19, n. 1/2, jan. 2013.

OPENDOAR. **About OpenDOAR**. Disponível em: <http://v2.sherpa.ac.uk/opensoar/information.html>. Acesso em: 26 de março de 2018

PINFIELD, Stephen. Journals and repositories: an evolving relationship?. **Learned Publishing**, [s.l.], v. 22, n. 3, p.165-175, jul. 2009.



TARDY, C. The role of English in scientific communication: lingua franca or Tyrannosaurus rex?. **Journal Of English For Academic Purposes**, [s.l.], v. 3, n. 3, p.247-269, jul. 2004. Elsevier BV. <http://dx.doi.org/10.1016/j.jeap.2003.10.001>.

XIA, Jingfeng. Diffusionism and open access. **Journal Of Documentation**, [s.l.], v. 68, n. 1, p.72-99, 2012. Emerald.