

Produção científica da UP em Acesso Aberto: retrato atual

Cristina Sousa Lopes

Faculdade de Engenharia da Universidade do Porto

cslopes@fe.up.pt

Luís Miguel Costa

Faculdade de Engenharia da Universidade do Porto

lmcosta@fe.up.pt

Resumo

Este estudo tem por objetivo avaliar a cobertura da produção científica da Universidade do Porto em Acesso Aberto pelo Repositório Aberto (enfoque na via verde) para o período de 2000 a 2013. Baseia-se numa análise do levantamento de artigos publicados em revistas científicas a partir de diversas fontes internas (SiGARRA e Repositório Aberto) e externas (SCOPUS e WoS). De entre os resultados obtidos verifica-se uma taxa de cobertura do Repositório Aberto de 16% relativamente ao potencial a disponibilizar. Ainda que para esse período o número de artigos depositados tenha aumentado, verifica-se que há uma margem grande de crescimento dos artigos a disponibilizar no Repositório Aberto, como veículo promotor do acesso e divulgação da investigação.

Palavras-chave: Acesso Aberto, Produção científica, Universidade do Porto, Repositório Aberto da Universidade do Porto, SCOPUS, Web of Science

Abstract

This study aims to assess the Open Access coverage of scientific production at University of Porto by “Repositório Aberto” (focus on green route) between 2000–2013. It starts from the analysis of published articles in scientific journals from several internal (SiGARRA and “Repositório Aberto”) and external sources (SCOPUS and WoS). From the attained results, one can see that the coverage of “Repositório Aberto” represents 16% of all the publications which could potentially be made available. The number of deposited articles had increased for this period. However, there is still room for an increase in the number of articles that can be made available at “Repositório Aberto” which is as a vehicle for the promotion and release of research.

Key-words: Open Access, Scientific publications, University of Porto, Open Repository of University of Porto, SCOPUS, Web of Science

Introdução

O Acesso Aberto (AA), ao proclamar uma alteração de princípios e práticas na partilha e acesso à literatura científica que as tecnologias de informação e comunicação vieram tornar possíveis, originou mudanças que tiveram um efeito de “pedra no charco”, com repercussões económicas, políticas e também culturais para os diferentes atores envolvidos no sistema de comunicação científica. Havendo uma identificação plena entre a designação e os seus princípios, o AA advoga um acesso à literatura científica livre, gratuito e sem restrições. Para a sua concretização, dois modelos ou vias têm sido adotados: a via dourada, associada à publicação em revistas de AA, e a via verde, associada ao depósito das publicações em repositórios (institucionais ou temáticos); para além de outras variantes, como modelos híbridos e novos modelos de publicação que têm surgido.

A Universidade do Porto (U.Porto), desde há alguns anos que tem vindo a apostar na criação de condições para a concretização do AA à produção científica e intelectual da sua comunidade. Enquadrada pelas declarações de Budapeste, Bethesda e Berlim e na declaração de "Acesso Livre à literatura científica" do Conselho de Reitores das Universidades Portuguesas em 2008, a U.Porto estabelece nesse mesmo ano de 2008 uma política para o AA à produção científica, materializada sob a forma de um Regulamento (Universidade do Porto 2008), que se concretizou na adoção de uma Política de Acesso Livre (Open Access)". Daí resultou o desenvolvimento do seu repositório institucional, que em 2007 arrancou com a criação do Repositório Aberto (RA) para as publicações científicas, tendo integrado posteriormente outras vertentes, como o Repositório Temático para bibliotecas e arquivos digitais e mais recentemente o Repositório de Dados para dados científicos. O RA desde a sua criação tem registado um franco crescimento, contando atualmente com mais de 33.000 publicações, colocando-o como o maior repositório a nível nacional.

Tendo por objetivo avaliar a cobertura da produção científica da U.Porto em AA pelo RA, este estudo, com particular enfoque na via verde, faz uma análise comparativa entre o número de artigos publicados em revistas científicas pela comunidade da U.Porto atualmente disponibilizado através do RA e o potencial que seria possível disponibilizar de acordo com as políticas de arquivo das revistas. Este estudo pretende traçar o panorama evolutivo na última década do AA da produção científica da U.Porto (abrangendo o período entre 2000 e 2013), com dados desagregados por faculdade. Procura-se assim avaliar a taxa de adesão ao AA nos processos de publicação da comunidade U.Porto e refletir sobre as diferenças entre as áreas em que a comunidade desenvolve a sua investigação.

O Repositório Aberto da Universidade do Porto

O repositório institucional da U.Porto foi disponibilizado na Internet em 2007, usando a plataforma DSpace. De entre os marcos históricos que caracterizam a sua evolução, para além dos já referidos, destaca-se em 2008 a sua certificação face às diretivas DRIVER, em 2009 ter integrado o repositório nacional RCAAP e em 2013 repositórios internacionais, como o DART Europe e o NDLTD.

Desde a sua criação que o RA tem tido um forte crescimento, principalmente a partir de 2011. É alimentado por diversas unidades orgânicas da U.Porto, respeitantes às 14 faculdades, à Porto Business School e às unidades de I&RD. De entre as unidades orgânicas que registam um maior número de publicações, destacam-se faculdades como a FEUP (cerca de 11.000 documentos), a FLUP (cerca de 6.000 documentos) e a FPCEUP (cerca de 4.500 documentos).

Relativamente aos tipos de publicações disponibilizadas, para este estudo foram agregadas em seis categorias principais. O tipo de documentos com maior representatividade corresponde às dissertações, teses e trabalhos académicos, com 60%, a seguir estão representados os artigos em revistas com 19% e os artigos em conferências com 9%.

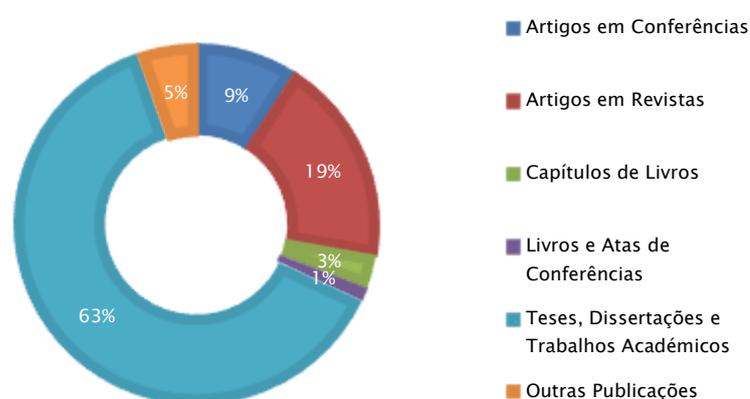


Figura 1:
Repositório Aberto: tipo de publicações

Para alimentar o RA a U.Porto beneficia de uma infraestrutura e sistemas de informação integrados, a partir dos quais são transferidas as publicações (metadados e texto integral), como o SiGARRA (Sistema de Gestão Agregado de Registos e Recursos Académicos), que se assume como um CRIS (*Current Research Information System*) de suporte a múltiplas valências e serviços, no âmbito do qual se destaca o módulo de publicações e o módulo de teses. Outras fontes como o DigiTool da FEUP, o ALEPH da FLUP e

ainda a introdução manual de documentos alimentam também o RA (as contribuições de cada uma destas fontes de dados encontram-se indicadas na Figura 2).

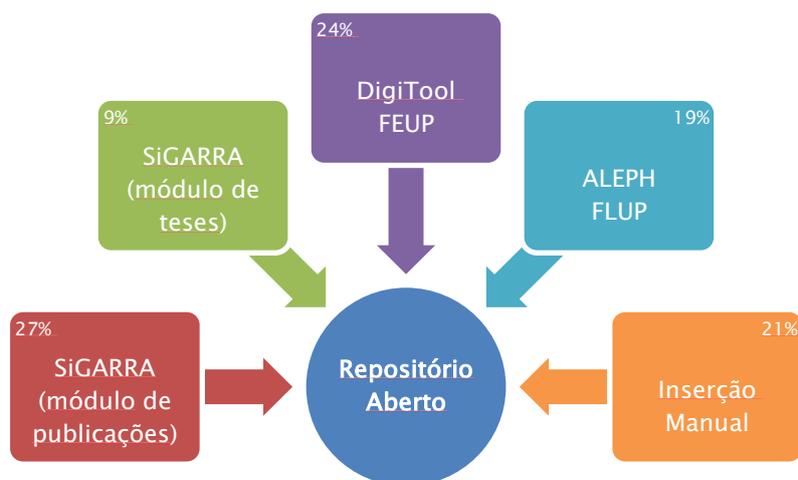


Figura 2:
Repositório Aberto: fontes de dados [total de documentos:32.558]

Importa realçar a importância do módulo de publicações do SiGARRA como a principal fonte de dados para o RA. Importância acrescida pelo facto de ser o sistema em que a comunidade académica deve registar toda a sua produção científica para efeitos de avaliação de desempenho dos professores. O processo de comunicação e transferência de dados do SiGARRA para o RA caracteriza-se por ser um processo automatizado, sendo de referir que apenas são transferidas publicações que tenham associado o respetivo ficheiro de texto integral e indicado como de AA.

Adicionalmente, neste estudo foi ainda apurada a evolução do número de artigos depositados no RA por data de depósito entre 2007 e 2013, verificando-se um notório aumento do número de depósitos, sobretudo a partir de 2010, com um crescimento muito significativo até à atualidade. Evolução que acompanha o crescimento global do número de publicações no RA. Para este crescimento contribui sobretudo o processo de depósito mediado implementado pela Biblioteca da FLUP, contribuindo com 29,8% do total de artigos depositados.

Revisão da literatura

A nível internacional e nacional, diversos estudos têm sido elaborados com o objetivo de fazer o retrato da situação e evolução do AA, no que toca a estudos generalizados e longitudinais (Archambault et al., 2013, Bjork et al., 2014, Bjork et al., 2010, Gargouri et al., 2012, Harnad et al., 2008, Hajjem, Harnad e Gingras, 2005, Laakso, 2014, Miguel, Chinchilla-Rodríguez e Moya-Anegón, 2011, Saraiva e Rodrigues, 2010). Na sua maioria são estudos quantitativos que procuram avaliar a situação real e potencial do AA da literatura científica, quer ao nível das revistas quer dos artigos, tendo em conta as políticas de acesso

das revistas. Estudos mais focados no AA verde procuram saber a proporção que os artigos de revistas de AA verde representam relativamente ao total de artigos publicados num determinado ano. Deve-se referir que análise do AA verde se torna um pouco mais complexa por contemplar variantes como os possíveis sítios de depósito e os diferentes tipos de versões disponibilizadas. Tal como refere Laakso (2014), a diversidade de conceitos e metodologias adotadas nesses estudos têm dificultado desenvolvimentos longitudinais e comparações entre os estudos, apontando por isso para resultados um pouco diversos.

Destaca-se a nível nacional um estudo sobre o acesso aberto à literatura científica e a implementação de repositórios institucionais em Portugal, em que Saraiva e Rodrigues (2010), concluíram que a produção científica depositada em repositórios era pouco significativa, menos de 10%.

Bjork (2014), partindo de estudos anteriores e de uma definição lata de AA verde ao considerar todas as cópias de artigos livremente acessíveis na web, incluindo diferentes versões dos artigos que existem em diferentes sítios para além do sítio do editor, estima uma taxa de 12%. Indica ainda haver diferenças significativas do AA verde consoante as disciplinas, devido às diferentes características das culturas científicas. Miguel (2011), focado ao nível do título da revista, estimou que o número de revistas que têm políticas explícitas de AA verde é de 32%, tendo verificado também existirem diferentes políticas consoante a região geográfica. Laakso (2014), num estudo elaborado para analisar as políticas de AA verde praticadas por editores, concluiu que 80,4% das versões aceites ou publicadas podem ser disponibilizadas num repositório institucional ou temático após um ano da publicação, havendo maior permissividade dos editores quanto à disponibilização da versão aceite em repositórios institucionais (79,9%) ou em páginas web pessoais (78,1%), quando comparado com a disponibilização em repositórios temáticos (32,8%). Relativamente a estudos prévios que apontam para a existência de percentagens inferiores de artigos disponibilizados em AA verde, sugere que esses resultados indiciam um franco subaproveitamento do AA verde.

Nesta mesma linha, um outro estudo longitudinal da Science Metrix (Archambault et al., 2013) avalia a disponibilidade das publicações académicas para o período entre 2004 e 2011 no âmbito de 22 áreas do conhecimento, estimando que o número de artigos em AA já ultrapassou o ponto de viragem dos 50% em alguns países e algumas áreas do conhecimento.

Metodologia

O estudo apresentado nasce da reflexão sobre o desenvolvimento das políticas de AA na U.Porto e a realidade da sua implementação, condicionada pela multiplicidade de culturas e práticas científicas da comunidade e a disparidade de sistemas com que os autores têm

que interagir para divulgar a sua investigação. Procura-se perceber em que medida o RA contribui de facto para a promoção e divulgação da investigação desenvolvida.

Iniciada a recolha das publicações de autores ligados à U.Porto, rapidamente se percebe a impossibilidade de se abranger toda a investigação, surgindo imediatamente o limite de se analisar as publicações em revistas científicas (tipicamente as publicações mais relevantes na avaliação dos repositórios) e ano 2000 como ponto de partida (pelo simbolismo da data). Foram definidas como questões a procurar responder:

1. Qual o número de artigos em cada uma das fontes analisadas (Repositório Aberto, SCOPUS e WoS) comparativamente ao SiGARRA?
2. Qual o número de artigos indexados no SCOPUS e WoS em Acesso Aberto comparativamente ao Repositório Aberto?
3. Qual o número de artigos em Acesso Aberto (dourado e verde) nas diversas fontes (SiGARRA, SCOPUS e WoS)?
4. Qual o número de artigos disponibilizados no Repositório Aberto face ao potencial a disponibilizar, distribuído por ano de publicação?
5. Qual a percentagem de artigos atualmente disponibilizados no Repositório Aberto face ao potencial a disponibilizar, distribuído por unidade orgânica?

Como ponto de partida, é efetuado o levantamento dos artigos em revistas científicas da U.Porto registados no SiGARRA e dos artigos depositados no RA. Este levantamento é complementado com a recolha dos artigos indexados nas duas bases de dados de referência, SCOPUS e WoS. Procura-se assim identificar toda a produção científica publicada em revistas científicas por autores associados à U.Porto, consolidando as diferentes listagens obtidas numa única lista com a identificação das coincidências e a eliminação de duplicados (resultando um total de 34.918 artigos). Para cada um destes artigos procurou-se ainda identificar a faculdade em que os autores estão afiliados, recorrendo à informação do SiGARRA, às coleções do RA e à afiliação identificada pelos serviços da Reitoria da U.Porto sobre os artigos da WoS, não tendo sido possível identificar a afiliação para 6.001 artigos.

A lista final consolidada de artigos é numa segunda fase cruzada com os dados das publicações divulgadas pelo DOAJ para identificação das revistas AA dourado e com a informação disponibilizada pelo SHERPA/RoMEO, para verificação das políticas de AA praticadas por cada uma das revistas. Obtém-se assim para cada um dos artigos a indicação de constar de uma revista em AA dourado (3.140 publicações) e também de constar de uma revista que permite o depósito em repositórios (24.721 publicações). Dos restantes 10.197 artigos há um conjunto que aparece identificado no SHERPA/RoMEO como não sendo de AA verde (4.490 publicações) e um outro (5.707 publicações) para qual não foi possível identificar uma política AA, seja porque ainda não dispõem de avaliação ou simplesmente não existem nesta ferramenta.

As fases finais do processo consistiram na aplicação de fórmulas à listagem consolidada para obter os dados necessários à leitura e análise e assim construir a informação de resposta a cada uma das questões de análise colocadas.

Resultados

A partir do universo de publicações da lista consolidada e dos múltiplos cruzamentos efetuados, foram obtidos os seguintes resultados para cada uma das questões.

Qual o número de artigos em cada uma das fontes analisadas (Repositório Aberto, SCOPUS e WoS) comparativamente ao SiGARRA?

Para a primeira questão foi apurado, após consolidação dos dados, o número de artigos existente em cada uma das fontes com a afiliação associada à U.Porto: SiGARRA 19.489, RA 4.014, SCOPUS 21.653 e WoS 20.144. Fazendo-se depois a comparação dos artigos destas fontes coincidentes com o SiGARRA, verifica-se que o RA é a que apresenta uma menor percentagem de sobreposição, com 3.104 (21%), apresentando a WoS: 8.297 (43%) e o SCOPUS: 8.544 (44%), tal como apresentado na Figura 3.

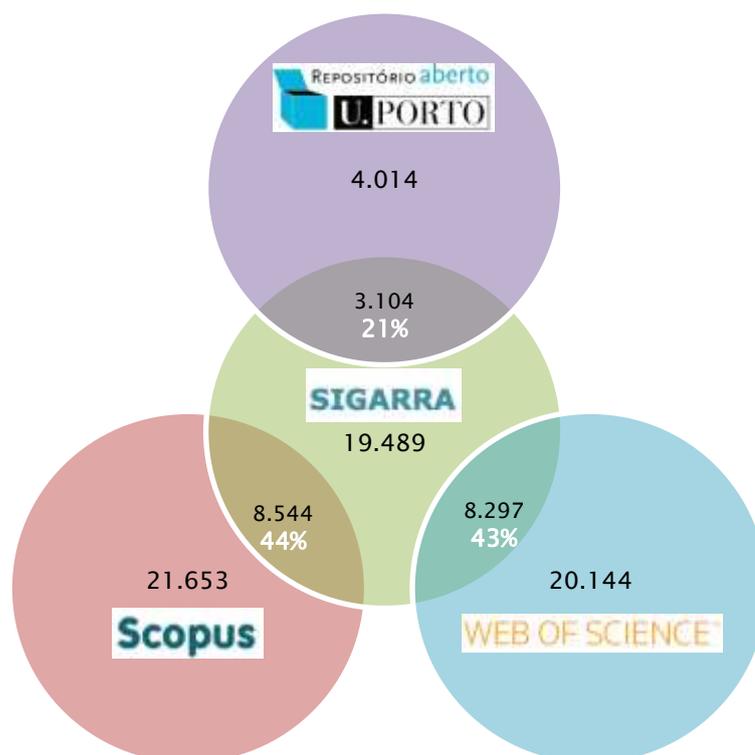


Figura 3:

Número de artigos nas diferentes fontes [consolidados em 34.918 artigos]

Qual o número de artigos indexados no SCOPUS e WoS em Acesso Aberto comparativamente ao Repositório Aberto?

Em relação à segunda questão, para o universo de revistas analisadas e tendo por base as políticas de acesso das revistas divulgadas pelo SHERPA/RoMEO, foram identificados 24.721 artigos de AA que podem depositados em repositórios, 2.773 sem AA que não permitem o depósito e 7.424 para os quais não foi possível identificar uma política de acesso. Comparando os artigos publicados nas revistas de AA identificados com os existentes no RA, verifica-se que o número de artigos coincidentes é relativamente baixo, 605 para a SCOPUS e 654 para a WoS, o que representa para ambas as fontes uma percentagem equivalente a cerca de 4%, como é mostrado na Figura 4.

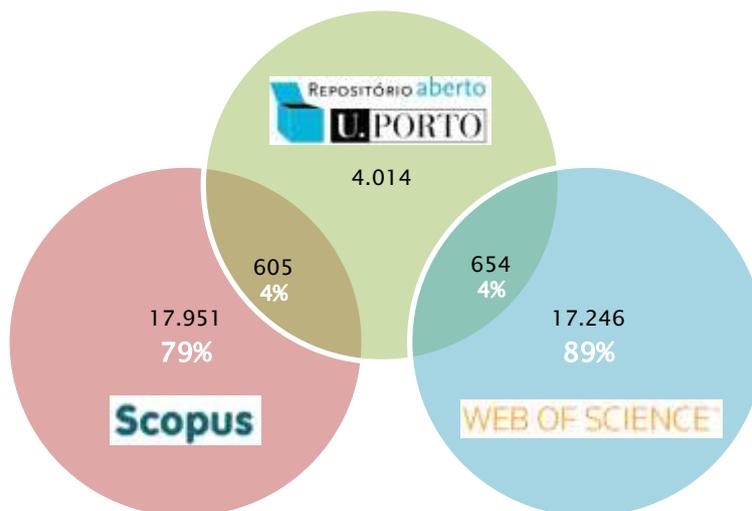


Figura 4:
Número de artigos em AA [consolidados em 24.721 artigos AA]

Qual o número de artigos em Acesso Aberto (dourado e verde) nas diversas fontes (SiGARRA, SCOPUS e WoS)?

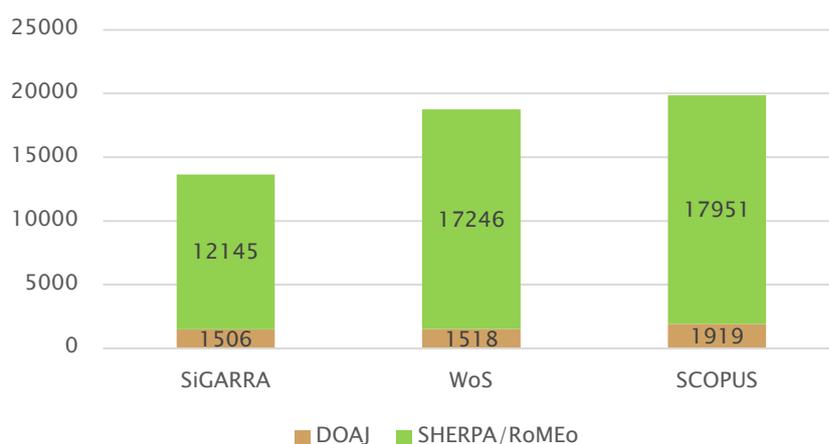


Figura 5:
Número de artigos AA (dourado e verde) em cada fonte

Relativamente à terceira questão, tomando por referência o SERPA/RoMEO e o DOAJ, verifica-se que para as três fontes analisadas o AA verde é preponderante relativamente ao

AA dourado. No SiGARRA o número de artigos em AA verde é de 12.145 face a 1.506 em AA dourado, na WoS é de 17.246 face a 1.518 e na SCOPUS é de 17.951 face a 1.919, como pode ser visto na Figura 5.

Qual o número de artigos disponibilizados no Repositório Aberto face ao potencial a disponibilizar, distribuído por ano de publicação?

No que toca à quarta questão, verifica-se que a relação entre o número total de artigos existente no RA foi de 4.014 face a um potencial de 24.721, tendo em conta o conjunto das fontes já referidas. Isto significa que a percentagem de cobertura do RA face aos artigos potencialmente disponibilizáveis em AA é de 16,2%. Ainda que esta percentagem não seja muito expressiva, pode dizer-se que não está desfasada dos valores que a bibliografia aponta, relativamente a estudos que comparam o AA real com o potencial.

Distribuindo por ano de publicação, verifica-se que o número de artigos disponibilizáveis em AA cresce significativamente a partir de 2007 enquanto no RA a evolução do número de artigos de facto disponibilizados se mantém estável, denotando apenas uma ligeira subida a partir de 2011. O aumento do número potencial deve-se sobretudo ao aumento em valor absoluto do número de artigos publicados em revistas e, por outro lado, ao aumento das revistas que possibilitam o AA verde.

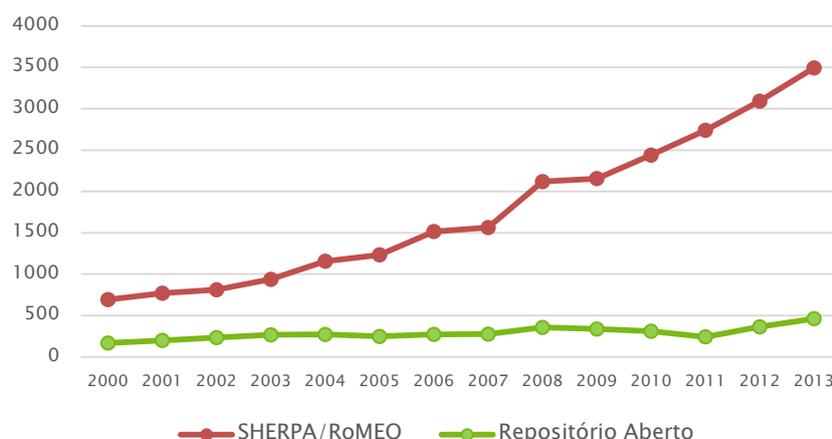


Figura 6:
Evolução do número real de artigos face ao potencial por ano de publicação

Qual a percentagem de artigos atualmente disponibilizados no Repositório Aberto face ao potencial a disponibilizar, distribuído por unidade orgânica?

Relativamente à quinta e última questão, verifica-se que diferença entre o real e potencial número de artigos disponibilizados em AA através de RA se manifesta de forma muito diferenciada consoante as unidades orgânicas. Assim, as Faculdades de Direito, Belas Artes, Arquitetura, Psicologia e Ciências da Educação são as que apresentam uma maior percentagem de artigos depositados no RA, enquanto o ICBAS e as Faculdades de Ciências,

Medicina e Engenharia são as que apresentam uma menor percentagem de artigos depositados no RA face ao potencial em AA. Isto demonstra, em certa medida, diferentes comportamentos e práticas de publicação e depósito por parte das diversas comunidades científicas. Mas nota-se particularmente que nas unidades orgânicas em que as práticas de depósito mediado pela Biblioteca (FLUP e FPCEUP) são a prática corrente as percentagens de documentos de facto disponibilizados aumenta significativamente. É ainda de realçar que nas comunidades onde o número potencial de artigos aumenta, diminui significativamente a percentagem de disponibilização por não ser uma prática corrente o depósito dos artigos.

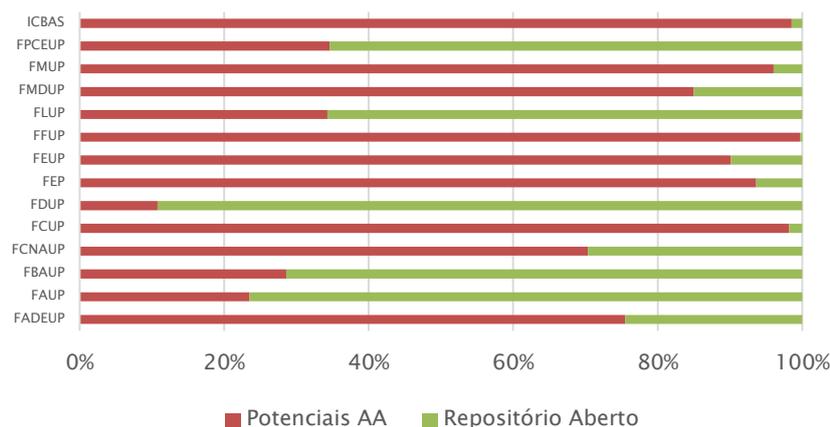


Figura 7:

Distribuição da percentagem real de artigos face ao potencial por unidade orgânica

Conclusões

Sendo o RA da U.Porto o maior repositório a nacional em número de publicações, no que respeita aos artigos publicados em revistas a sua representatividade é de 19% face ao total de publicações. Comparando o número real de artigos depositados no RA relativamente ao número potencial passível de ser depositado num repositório institucional tendo em conta as políticas de acesso das revistas, verifica-se que à data atual a percentagem de artigos depositados é de cerca de 16%. Ainda assim, o número de artigos no RA tem aumentado, beneficiando da infraestrutura e sistemas de informação integrados na U.Porto. Embora o registo de publicações no SiGARRA seja cada vez maior e mais facilitado, verifica-se que ainda há uma margem grande de crescimento dos artigos a disponibilizar no RA. Concomitantemente, as realidades muito diferentes nas práticas de publicação e de depósito das várias unidades orgânicas contribuem para explicar a ainda pouca cobertura de artigos em revistas científicas pelo RA.

No que respeita às dificuldades detetadas e áreas de intervenção futuras destacam-se a necessidade de fazer uma análise aprofundada sobre a qualidade dos metadados do RA (bem como do SiGARRA) e posterior correção/validação, em virtude das inconsistências detetadas. Outra necessidade prende-se com a verificação da conformidade das versões dos

ficheiros depositados no RA consoante as políticas de AA autorizadas, para proteger a U.Porto de eventuais inconformidades.

Como desenvolvimentos futuros e métodos mais efetivos para incrementar a taxa de depósito, espera-se que os requisitos de AA impostos pelo Horizonte 2020 venha a dar os seus frutos, o que poderá ser reforçado com a criação de uma política *de facto* mandatária por parte da U.Porto, bem como o reforço das políticas de incentivo e de campanhas de sensibilização para o aumento da taxa de depósito por parte dos autores. Alguns desenvolvimentos simples nos sistemas informáticos, como a integração com a API do SHERPA/RoMEO, poderão contribuir para a clarificação das permissões efetivas das revistas no momento do depósito e assim facilitar as opções dos autores.

Sendo este um estudo que procura construir um retrato do estado atual da produção científica da U.Porto em AA, espera-se que desta análise se mostrem evidências relativamente às práticas da comunidade académica da U.Porto, bem como a cobertura e representatividade de cada um dos sistemas analisados face às publicações passíveis de serem disponibilizadas. Assim e para além dos resultados que se conta virem a ser úteis e relevantes para a U.Porto, espera-se igualmente que possam contribuir para o enriquecimento e construção de um conhecimento mais alargado e aprofundado sobre a situação do AA a nível nacional.

Referências bibliográficas

ARCHAMBAULT, Eric [et al.] (2013) – *Proportion of Open Access Peer-Reviewed Papers at the European and World Level: 2004–2011* [Em linha]. Montréal : Science Metrix. [Consult. 19 set. 2014]. Disponível na Internet: <URL: http://www.science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf>.

BJORK, Bo-Christer [et al.] (2014) – Anatomy of green open access. *Journal of the American Society for Information Science and Technology* [Em linha]. Vol 65, Nº 2. 237–250. [Consult. 19 set. 2014]. Disponível na Internet: <URL: <http://dx.doi.org/10.1002/asi.22963>>. ISSN 1532-2882.

BJORK, Bo-Christer [et al.] (2010) – Open Access to the Scientific Journal Literature: Situation 2009. *Plos One* [Em linha]. Vol 5, Nº 6 .e11273. [Consult. 19 set. 2014]. Disponível na Internet: <URL: <http://dx.doi.org/10.1371/journal.pone.0011273>>. ISSN 1932-6203.

GARGOURI, Yassine [et al.] (2012) – Green and Gold Open Access percentages and growth, by discipline. In *7th International Conference on Science and Technology Indicators (STI)*. Montreal, Canada: 2012. [Consult. 19 set. 2014]. Disponível na Internet: <URL: <http://eprints.soton.ac.uk/340294>>.

HAJJEM, Chawki; HARNAD, Stevan; GINGRAS, Yves (2005) – Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact. *IEEE Data Engineering Bulletin* [Em linha]. Vol 28, Nº 4. 39–47. [Consult. 19 set. 2014]. Disponível na Internet: <URL: <http://eprints.soton.ac.uk/261688/>>.

HARNAD, Stevan [et al.] (2008) – The Access/Impact Problem and the Green and Gold Roads to Open Access: An Update. *Serials Review* [Em linha]. Vol 34, Nº 1. 36–40. [Consult. 19 set. 2014]. Disponível na Internet: <URL: <http://dx.doi.org/10.1016/j.serrev.2007.12.005>>. ISSN 0098–7913.

LAAKSO, Mikael (2014) – Green open access policies of scholarly journal publishers: a study of what, when, and where self-archiving is allowed. *Scientometrics* [Em linha]. Vol 99, Nº 2. 475–494. [Consult. 19 set. 2014]. Disponível na Internet: <URL: <http://dx.doi.org/10.1007/s11192-013-1205-3>>. ISSN 0138–9130.

MIGUEL, Sandra; CHINCHILLA-RODRIGUEZ, Zaida; DE MOYA-ANEGÓN, Félix (2011) – Open access and Scopus: A new approach to scientific visibility from the standpoint of access. *Journal of the American Society for Information Science and Technology* [Em linha]. Vol. 62, Nº 6. 1130–1145. [Consult. 19 set. 2014]. Disponível na Internet: <URL: <http://dx.doi.org/10.1002/asi.21532>>. ISSN 1532–2890.

SARAIVA, Ricardo; RODRIGUES, Eloy (2010) – O Acesso Livre à literatura científica em Portugal: a situação actual e as perspectivas futuras. In *Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas – “Políticas de informação na sociedade em rede : actas”*. Guimarães: 2010. APBAD. [Consult. 19 set. 2014]. Disponível na Internet: <URL: <http://repositorium.sdum.uminho.pt/handle/1822/10528>>.

UNIVERSIDADE DO PORTO (2008) – *Regulamentos de Política de Acesso Livre (Open Access) da U.PORTO* [Em linha]. [Consult. 19 set. 2014]. Disponível na Internet: <URL: https://sigarra.up.pt/up/pt/web_gessi_docs.download_file?p_name=F87706059/Regulamentos_Open_Access.pdf>.