

DOS SISTEMAS ISOLADOS AOS PARTILHADOS: O REPOSITÓRIO INSTITUCIONAL DA UNIVERSIDADE DE COIMBRA

Maria Manuel Borges

Universidade de Coimbra (Portugal)

Resumo

O acesso à informação científica é fundamental para os investigadores poderem progredir nos seus trabalhos de investigação. Contudo, existem barreiras para o acesso a esta informação que devem ser ultrapassadas. Não se trata de um problema que afecte apenas países em vias de desenvolvimento, uma vez que aqueles mais activos na constituição de vias complementares de acesso à informação científica são o Reino Unido, os Estados Unidos da América, o Canadá e a Austrália, entre tantos outros. Uma das reacções à dificuldade em obter quer o acesso à informação, quer de proceder à sua difusão de modo a obter prioridade e a potenciar o impacto de citação, tem passado pela constituição de repositórios institucionais e temáticos. Esta proposta tem por objectivo demonstrar a abordagem e metodologia utilizadas no estudo de viabilidade para a constituição de um repositório institucional da Universidade de Coimbra.

Abstract

For researchers it is crucial to have access to scientific information. However, there are barriers to access this information that must be overcome. This is not a problem that affects only developing countries. The United Kingdom, United States, Canada and Australia are, among others, the most active in setting up institutional or thematic repositories as a complementary way to have access and disseminate the scientific information, which is essential to potentiate its impact. This study pretends to show the approach and methodology used on the case study for the implementation of an institutional repository at the University of Coimbra.

Introdução

As universidades são, simultaneamente, as grandes produtoras e consumidoras do conhecimento. Os seus investigadores produzem e validam – através do processo de revisão pelos pares e pela participação em conselhos editoriais –, a informação publicada em revistas científicas, as mesmas que são posteriormente adquiridas pelas suas bibliotecas a preços significativamente elevados. As repercussões da tecnologia digital são visíveis não apenas em novas formas de colaboração, mas também em novas possibilidades que permitem aos autores recuperar o controlo da difusão da investigação que produzem. Se os ‘sabores’ que podem ser produzidos (Willinsky, 2006) estão a ser discutidos desde a *Budapest Open Access Initiative* (BOAI) e suas sucessoras, não

é menos significativo o conjunto de estudos que procura estabelecer uma correlação positiva entre a acessibilidade e o impacto de citação.

Procurando caracterizar a comunidade de investigadores da Universidade de Coimbra (UC), e usando um conjunto de unidades de investigação seleccionadas como amostra procurou-se perceber a forma como os investigadores estão a alterar o seu comportamento face a novos ambientes trazidos pelas TIC e a adaptar os modos de acesso, produção e difusão da informação científica.

Metodologia

O estudo, basicamente exploratório e qualitativo, seguiu duas grandes linhas (Cf. Fig. 1): (a) capturar a produção científica para poder proceder à identificação das políticas de copyright dos editores, usando as bases de dados do ISI (*Arts & Humanities Citation Index*, *Social Sciences Citation Index* e *Science Citation Index*); a base de dados do Observatório da Ciência, Tecnologia e Ensino Superior (OCES) para a identificação de outros títulos das Ciências Sociais e Humanidades; a b-on [Biblioteca do Conhecimento online] e a base de dados Sherpa; (b) caracterizar os inquiridos e as unidades de investigação através de inquérito, da base de dados da Fundação para a Ciência e Tecnologia (FCT)¹ e dos sítios Web das unidades de investigação.

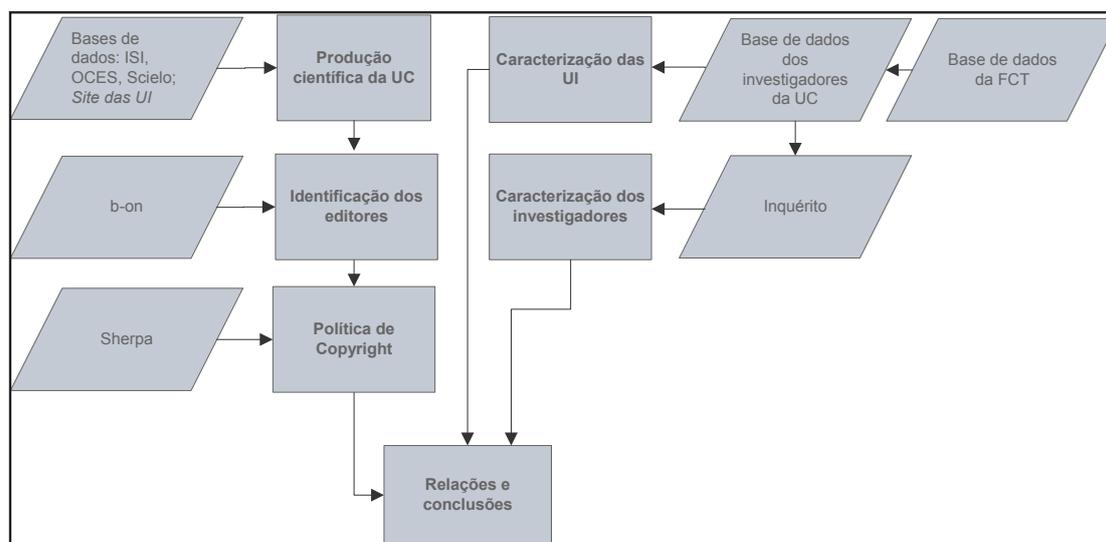


Fig. 1 – Abordagem e metodologia

Discussão

Os dados obtidos pela análise do inquérito, bem como aqueles que resultaram da análise das bases de dados citadas permitiram obter uma imagem clara sobre a percepção

¹ O OCES e a FCT são organismos dependentes do Ministério da Ciência, Tecnologia e Ensino Superior.

e comportamento dos investigadores face ao auto-arquivo das publicações, bem como identificar as principais fontes usadas e eventuais restrições aplicadas à difusão da sua investigação através de um repositório institucional.

O contexto organizacional e cultural no qual os investigadores trabalham, apenas aparentemente eliminado pela tecnologia digital, é essencial para a compreensão do seu comportamento, uma vez que existem outros constrangimentos associados à investigação que devem ser tidos em conta como, por exemplo, o acesso a equipamentos, a encontros face-a-face com investigadores de topo e a selecção dos temas de investigação.²

A Universidade de Coimbra tem seis Faculdades (Letras, Direito, Medicina, Ciência e Tecnologia, Farmácia, Economia, Psicologia e Ciências da Educação e Desporto). Os dados obtidos são relativos a trinta e nove unidades de investigação, todas elas membros do Instituto de Investigação Interdisciplinar (III) da Universidade de Coimbra e avaliadas pela Fundação para a Ciência e Tecnologia (FTC), a agência de financiamento da investigação em Portugal. O inquérito envolveu cerca de 1823 investigadores e uma percentagem de respostas válidas na ordem dos 10%. Uma das variáveis que procurámos testar era a da correlação entre a área científica e possíveis obstáculos no acesso à informação. Apesar de os resultados obtidos por Swan e Brown (2005) mostrarem que não existem diferenças significativas entre as disciplinas no que concerne ao acesso aos recursos, na UC, e para as Artes e Humanidades, as dificuldades são muito claras. Tais diferenças podem estar relacionadas com a tipologia de documentos usada na investigação, tendo em conta que, ao menos em princípios, algumas das digitais são menos restritivas no que respeita ao acesso.³

A oferta progressiva de informação disponível nas ‘pontas dos dedos’ levou-nos a inquirir sobre o uso da informação de acordo com o seu formato, analógico ou digital. Na categoria ‘essencial’⁴ o formato digital é mais utilizado em todas as áreas com excepção das Artes e Humanidades e Ciências Sociais, obtendo a mesma percentagem de uso, independentemente do formato, nas Ciências Naturais e do Ambiente.

Tabela 1 – Área científica/Uso das revistas (formatos papel e digital)

Áreas científicas	Revistas arbitradas (papel) (%)	Revistas arbitradas (digital) (%)	Revistas não arbitradas (papel) (%)	Revistas não arbitradas (digital) (%)
Artes e Humanidades	57	50	57	36
Ciências Sociais	86	82	28	28
Ciências Puras	83	96	4	13
Ciências Naturais e do Ambiente	91	91	13	26
Ciências da Vida	89	100	6	11
Engenharia e Tecnologia	73	88	13	25

² Em países como Portugal, a escolha dos temas de investigação depende fortemente da escolha europeia.

³ Apenas algumas, uma vez que a grande maioria implica pagamento. Stevan Harnad denomina-as de fontes sujeitas a portagem (*toll-gate sources*).

⁴ As categorias usadas foram: irrelevante, importante e essencial.

Um dos problemas relativos à informação prende-se com a identificação das características mais importantes das fontes utilizadas. Os resultados apresentados na tabela seguinte mostram diferenças significativas para as diferentes áreas científicas no que respeita à actualidade – Ciências Puras e Engenharia e Tecnologias são aquelas que consideram mais relevante este aspecto, o que expressa uma pressão para a publicação de resultados maior do que outras áreas como as Artes e Humanidades que detêm o valor mais baixo. A credibilidade, importante para todas as áreas, atinge valores mais elevados nas Ciências da Vida e Ciências Puras. Considerando as versões impressa e electrónica, a primeira é mais importante para as Artes e Humanidades, enquanto a última é mais importante para as Ciências Naturais e do Ambiente. O acesso ao texto integral é mais importante para as Ciências da Vida e para as Ciências Puras, enquanto o acesso por computador pessoal é relevante para as Artes e Humanidades.⁵ Este aspecto é reafirmado na resposta seguinte onde esta área indica uma preferência no acesso que se traduz por 24 [horas]/7[dias por semana]. A ligação a outros trabalhos é igualmente importante para esta área, enquanto a inclusão de elementos multimédia é preferido pelas Ciências da Vida.

Tabela 2 – Área científica: factores relevantes na selecção de fontes de informação
(1=Irrelevante; 2=Importante; 3=Muito importante)

		Média	Desvio padrão
Actualidade	Ciências Exactas	2,8	0,410
	Engenharia e Tecnologia	2,79	0,409
	Ciências Naturais e do Ambiente	2,78	0,421
	Ciências da Vida	2,76	0,437
	Ciências Sociais	2,72	0,455
	Artes e Humanidades	2,63	0,504
Credibilidade	Ciências da Vida	3	0,000
	Ciências Exactas	3	0,000
	Artes e Humanidades	2,9	0,301
	Ciências Sociais	2,9	0,364
	Engenharia e Tecnologia	2,88	0,375
	Ciências Naturais e do Ambiente	2,86	0,344
Material certificado	Ciências da Vida	2,94	0,242
	Ciências Exactas	2,8	0,410
	Ciências Naturais e do Ambiente	2,69	0,558
	Engenharia e Tecnologia	2,58	0,534
	Ciências Sociais	2,52	0,646
	Artes e Humanidades	2,09	0,831
Versão impressa	Artes e Humanidades	2	0,774
	Ciências da Vida	1,94	0,658
	Ciências Sociais	1,84	0,841
	Ciências Naturais e do Ambiente	1,82	0,7168
	Ciências Exactas	1,65	0,670
	Engenharia e Tecnologia	1,58	0,745

⁵ Apesar da preferência pela versão impressa.

Versão electrónica	Ciências Naturais e do Ambiente	2,56	0,589
	Ciências da Vida	2,52	0,624
	Ciências Exactas	2,5	0,606
	Engenharia e Tecnologia	2,43	0,6656
	Ciências Sociais	2,28	0,729
	Artes e Humanidades	2,18	0,6039
Texto integral	Ciências da Vida	2,76	0,562
	Ciências Exactas	2,75	0,444
	Ciências Naturais e do Ambiente	2,73	0,540
	Artes e Humanidades	2,72	0,646
	Ciências Sociais	2,72	0,453
	Engenharia e Tecnologia	2,69	0,463
Acesso por computador	Artes e Humanidades	2,72	0,467
	Ciências da Vida	2,64	0,701
	Engenharia e Tecnologia	2,56	0,604
	Ciências Exactas	2,55	0,604
	Ciências Sociais	2,48	0,677
	Ciências Naturais e do Ambiente	2,43	0,727
Acesso contínuo (24/7)	Artes e Humanidades	2,72	0,467
	Ciências da Vida	2,70	0,587
	Engenharia e Tecnologia	2,43	0,693
	Ciências Exactas	2,4	0,680
	Ciências Naturais e do Ambiente	2,34	0,775
	Ciências Sociais	2,3	0,762
Ligação a outros trabalhos	Artes e Humanidades	2,27	0,646
	Ciências da Vida	2,23	0,664
	Ciências Naturais e do Ambiente	2,21	0,795
	Engenharia e Tecnologia	2,20	0,660
	Ciências Exactas	2,2	0,523
	Ciências Sociais	2,06	0,739
Elementos multimédia	Ciências da Vida	1,88	0,696
	Artes e Humanidades	1,81	0,750
	Ciências Naturais e do Ambiente	1,73	0,864
	Engenharia e Tecnologia	1,41	0,633
	Ciências Sociais	1,36	0,597
	Ciências Exactas	1,35	0,587

A relação entre a área científica e a sua produção ilustra as diferentes práticas entre aquelas que reflectem tendências geográficas e linguísticas como as Artes e Humanidades e as Ciências Sociais, daquelas [área de STM] onde estas tendências não são visíveis. Assim, procurámos apurar dados que mostrassem esta diferença em revistas científicas nacionais e estrangeiras que se verificou na amostra analisada.

Para apurar a sensibilidade dos autores ao acesso livre, confrontámo-los com duas afirmações acerca da difusão e impacto dos seus trabalhos. Os resultados confirmaram aqueles obtidos no trabalho de Swan e Brown (2005), tal como se pode concluir da tabela seguinte.

Tabela 3 – Área científica: opinião acerca dos títulos em AL
(1=discordo completamente...7=concordo completamente)

Área científica		O AL é mais eficaz para a difusão	O AL aumenta o impacto de citação
Artes e Humanidades	Média	5,7	4,7
	Desvio padrão	1,414	1,380
Ciências Sociais	Média	5,5	4,2
	Desvio padrão	1,619	1,961
Ciências Puras	Média	4,9	3,1
	Desvio padrão	1,586	1,124
Ciências Naturais e do Ambiente	Média	5,6	4,7
	Desvio padrão	1,199	1,572
Ciências da Vida	Média	6,1	4,0
	Desvio padrão	0,857	2,160
Engenharia e Tecnologia	Média	5,3	4,4
	Desvio padrão	1,404	1,246
Total	Média	5,5	4,2
	Desvio padrão	1,418	1,660

Muitos dos autores da amostra já tinham procedido ao auto-arquivo usando a sua página pessoal e/ou institucional ou através da submissão a repositórios institucionais ou temáticos. Questionados acerca destas abordagem, a maioria dos autores de todas as áreas, usam sobretudo os últimos para difundir o seu trabalho e fazem-no utilizando a versão pdf do editor. Este resultado é diferente do que foi obtido por Swan e Brown (2005), no qual a população de respondentes usou a página pessoal ou institucional para difundir o seu trabalho e utilizou a versão post-print.

Tabela 4 – Área científica: auto-arquivo na pagina pessoal e/ou institucional

Áreas científicas	Preprint (%)	Postprint (%)	Pre+postprint (%)	PDF do editor (%)	Nenhuma (%)
Artes e Humanidades	7	29	7	7	50
Ciências Sociais	8	14	14	6	54
Ciências Exactas	13	9	13	17	43
Ciências Naturais e do Ambiente	4	9	9	17	61
Ciências da Vida	0	0	0	22	78
Engenharia e Tecnologia	7	45	0	4	45

Tabela 5 – Área científica: auto-arquivo em repositório institucional e/ou temático

Áreas científicas	Preprint (%)	Postprint (%)	Pre+postprint (%)	PDF do editor (%)	Nenhuma (%)
Artes e Humanidades	7	14	7	57	14
Ciências Sociais	10	12	6	54	14
Ciências Exactas	13	13	17	48	9
Ciências Naturais e do Ambiente	4	4	0	78	13
Ciências da Vida	0	11	6	78	6
Engenharia e Tecnologia	2	18	0	57	21

A abertura dos autores à participação num repositório institucional se o depósito fosse obrigatório está na linha dos resultados obtidos por Swan e Brown (2005) no qual “a maioria dos autores (81%) depositaria o seu trabalho num repositório institucional ou temático”.

Tabela 6 – área científica: participação num repositório da UC se fosse obrigatório

Área científica	Aceitação (%)
Artes e Humanidades	86
Ciências Sociais	72
Ciências Puras	78
Ciências Naturais e do Ambiente	70
Ciências da Vida	78
Engenharia e Tecnologia	82

Uma das limitações deste estudo foi o de ter sido impossível ter em conta as especificidades disciplinares. Permitted, contudo, concluir que existe uma atitude relativa ao auto-arquivo positiva na amostra analisada. Outra conclusão importante revelada por esta amostra é que as áreas tradicionalmente menos inclinadas para o uso das TIC, como as Artes e Humanidades e as Ciências Sociais, as utilizam efectivamente.

Para poder identificar os títulos mais relevantes para os autores da universidade, foram analisadas todas as publicações das bases de dados do ISI por um período de quatro quinquénios⁶. A presença mais expressiva é das áreas de STM (SCI), com uma presença significativamente menor nas Ciências Sociais (SSCI) e apenas simbólica nas Artes e Humanidades (AHCI), o que explica o recurso a outras fontes para estas duas últimas grandes áreas do conhecimento.

A identificação dos títulos bem como das suas políticas de copyright foram estabelecidas quer pela pesquisa na base de dados SHERPA/RoMEO quer pela pesquisa

⁶ O que significa que para os autores da UC, as publicações tidas em conta são maioritariamente aquelas cobertas pelas bases de dados do ISI. O objectivo foi o de identificar as políticas de copyright dos títulos usados pelos autores da universidade para poder concluir sobre a viabilidade do auto-arquivo a partir apenas de tais políticas.

directa nas páginas Web dos títulos em causa. Para os títulos do SCI, os resultados permitiram concluir que 77% dos editores envolvidos, de acordo com a classificação do SHERPA/RoMEO, são editores verdes, isto é, permitem que o autor tenha as versões pré e post-print do seu trabalho na sua página pessoal e/ou institucional e num repositório institucional ou temático.

Tabela 7 – SCI: Posição dos editores face ao auto-arquivo (1983-2003)

Política editorial	(%)
Editor verde (pre+postprint)	77
Editor azul (postprint)	5
Editor amarelo (preprint)	2
Editor branco	7
Acesso livre (com ou sem período de embargo)	1
Outros	5
Sem informação	3

Se as áreas de STM têm esta representação, podemos facilmente concluir a enorme vantagem que a Universidade de Coimbra pode obter com um repositório. A análise é particularmente importante para estas áreas por se tratar de informação cujo acesso é crucial para a prossecução da investigação. Apesar de a mesma análise ser aplicável a outras áreas como as Artes e Humanidades e as Ciências Sociais, é também verdade que, regra geral, são menos relevante em virtude dos diferentes padrões de publicação.

As Ciências Sociais são uma área que abrange uma vasta gama de disciplinas, as quais, em maior ou menor grau, não têm a revista científica como meio privilegiado de comunicação entre pares. Apesar disso, os resultados da análise são expressivos: 63% são editores verdes, 3% editores de títulos em OA e 31% são editores amarelos.

Tabela 8 – SSCI: Posição dos editores face ao auto-arquivo (1983-2003)

Política editorial	(%)
Editor verde (pre+postprint)	63
Editor amarelo (preprint)	31
Acesso livre (com ou sem período de embargo)	3
Sem informação	3

A mesma análise aplicada aos títulos indexados pela A&HCI revela que apenas 9% são editores verdes, ignorando-se a política de copyright praticada pela esmagadora maioria, 91%. Contudo, uma análise aos títulos do núcleo permitem um aumento ligeiro dos editores verdes que passam para 13%, decrescendo a percentagem daqueles de que se desconhece a política de copyright praticada para 87%. É claro o papel secundário que este tipo de publicações têm nesta grande área do conhecimento, bem como o enviesamento geográfico e linguístico da base de dados do ISI.

Tabela 9 – AHCI: Posição dos editores face ao auto-arquivo (1983-2003)

Política editorial	(%)
Editor verde (pre+postprint)	9
Sem informação	91

Outro aspecto importante a ter em conta nesta análise respeita ao número de trabalhos publicados pelos autores em títulos da UC, sobretudo nas áreas de Artes e Humanidades e Ciências Sociais. Se estes títulos fossem publicados em acesso livre, a produção científica usufruiria de maior visibilidade do que aquela que obtém em papel. A existência destas publicações justifica-se pelo aumento da reputação dos seus autores que poderiam beneficiar enormemente com a difusão electrónica, como meio principal ou suplementar de publicação, pelo que faz todo o sentido a articulação entre estas duas faces da mesma moeda, o repositório institucional e a migração para os títulos também em formato electrónico e idealmente em acesso livre.

Conclusão

O cruzamento dos dados obtidos permitiram concluir não apenas da viabilidade da constituição do repositório por inexistência de barreiras de propriedade intelectual dos títulos utilizados, bem como apurar que os investigadores estariam dispostos a aderir a tal repositório. A desmaterialização da academia, a inclusão em redes de investigadores, incluindo as redes sociais, estão a transformar a comunicação da ciência. Novas formas de apresentação e conexão do conhecimento estão a aproximar e a transformar o modo como as pessoas absorvem e se adaptam à tecnologia.

O problema para a Europa é o de aumentar o impacto da pesquisa aqui realizada, cuja visibilidade é ameaçada por iniciativas internacionais um pouco por todo o mundo. O modo de publicação tradicional requer o acesso contínuo a resultados para recuperar o investimento, mas é possível que novos modelos surjam como resposta a novos requisitos da ciência e da sociedade relativos ao acesso aos resultados da investigação, particularmente aqueles financiados com dinheiros públicos. A integração de ferramentas para acesso e criação da informação numa única plataforma transforma a comunicação. A interligação entre bases de dados com o uso de ferramentas como o Crossref⁷, o Citeseer⁸, CiteULike⁹ ou Mendeley¹⁰, etc. Estas novas ferramentas permitem evoluir para um nível de partilha e de análise impossível de conseguir

⁷ O CrossRef é uma iniciativa de uma organização não lucrativa gerida pelo consórcio *International STM Journal Publisher PILA (Publishers' International Linking Association)* com o objectivo de interligar a literatura mundial publicada em revistas científicas.

⁸ Disponível na WWW: <http://citeseer.ist.psu.edu/citeseer.html>. O código fonte é gratuito para fins não comerciais. Suporta o protocol OAI-PMH.

⁹ Um serviço com o patrocínio da Springer. Disponível na WWW: <http://www.citeulike.org/>

¹⁰ Disponível na WWW: <http://www.mendeley.com/>

com idêntica amplitude em ambiente impresso. Ao contrário das bases de dados do ISI que usam os 'core journals' para evidenciar a interligação entre a literatura, estes novos serviços usam o texto integral de artigos em acesso livre e disponíveis na Web proporcionando um serviço novo e mais vantajoso (Harnad, 2001).

Todos estes problemas são relevantes para a discussão do impacto da tecnologia digital nas universidades, as quais se verão cada vez mais envolvidas em actividades de ensino à distância. Nem todas as universidades seguirão este caminho porque os seus percursos dependem muito do contexto e do perfil das comunidades científicas, mas a tecnologia digital estará fortemente presente na academia.

Referências bibliográficas

- BORGES, Maria Manuel (2006) – *A Esfera: Comunicação Académica e Novos Media*. Coimbra: Universidade. Tese de doutoramento.
- HARNAD, S. (2001) – The Self-Archiving Initiative. *Nature*, (41). Disponível na WWW: <http://www.nature.com/nature/debates/e-access/Articles/harnad.html>.
- HARNAD, S. [et al.] (2004) - The Access/Impact Problem and the Green and Gold Roads to Open Access. *Serials Review*, 30 (4). Disponível na WWW: <http://www.ecs.soton.ac.uk/~harnad/Temp/impact.html>.
- HARNAD, S. & BRODY, Tim (2004) - Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals. *D-Lib Magazine*, 10 (6). Disponível na WWW: <http://www.dlib.org/dlib/june04/harnad/06harnad.html>
- SWAN, A. & BROWN, S. N. (2005) - *Open access self-archiving: An author study*. London: JISC. Disponível na WWW: <http://cogprints.org/4385/>.
- TERRY, RobertL; KILEY, Robert (2006) – Open access to the research literature: a funder's perspective. In JACOBS, Neil, ed. (2006) – *Open Access: Key Strategic, Technical and Economic Aspects*. Oxford: Chandos Publishing. P. 101-109.
- WILLINSKY, J. (2003) - The Nine Flavours of Open Access Scholarly Publishing . *J. of Postgraduate Medicine*, 49, (3). Disponível na WWW: <http://www.jpgmonline.com/article.asp?issn=0022-3859;year=2003;volume=49; issue =3;spage=263;epage=267;aulast=Willinsky>.
- WILLINSKY, J. (2006) – *The Access Principle: The Case for Open Access to Research and Scholarship*. Cambridge, MA; London: The MIT Press. Também disponível na WWW:<http://mitpress.mit.edu/catalog/item/default.asp?tid=10611&ttype=2>.