



ANÁLISE DE *OUTLINKS*: uma visita aos periódicos da área de Ciências Sociais Aplicadas no Brasil (estratos A1 e A2)

Eixo temático: Informetria e Webometria

Modalidade: Apresentação oral

1 INTRODUÇÃO

Com o desenvolvimento da *Word Wide Web* um universo de novas possibilidades descortinou-se diante da comunidade científica. A *web* é marcada por uma série de evoluções que implicam em alterações tecnológicas e nas formas como seus usuários, consumidores potenciais, se relacionam com as informações e produtos disponibilizados através da rede. Essas novas relações fundamentam-se no uso de um dispositivo central para a comunicação mediada pelo computador: o *link*

[...] não é apenas um auxílio para a navegação, permitindo ao leitor, por exemplo, navegar até o fim da página, voltar à “*home-page*” etc. O criador de uma página A, ao assinalar nesta um “*link*” para a página B, está indicando que considera a página B, de alguma maneira, relacionada semanticamente a página A. (MENDONÇA *et. al.*, 2004, p. 1).

Defende-se aqui a concepção de que o *link* transcende o papel de mero nó estruturante e relacional na *web*. Interpretações acerca do *link* indicam que novas possibilidades de entrelaçamento hipertextual são estabelecidas por meio das ligações que se efetivam entre dois ou mais *links*.

2 A COMUNICAÇÃO CIENTÍFICA E A WEBOMETRIA

Com a explosão bibliográfica, fomentada pelo advento dos tipos móveis e do papel de celulose, a comunicação científica atinge outro patamar de desenvolvimento. O surgimento do periódico científico determina o início de uma nova maneira de se fazer Ciência. A

possibilidade de divulgação dos resultados de pesquisa entre a comunidade, de modo mais abrangente e rápido, representa um enorme ganho para os pesquisadores. Disseminar a informação científica por meio de periódicos resulta em uma comunicação formal mais precisa e mais econômica. A criação de um periódico com temática específica representa um grande avanço e amadurecimento para qualquer área do saber. (ZIMAN, 1979).

No que tange à Webometria, Thelwall, Vaughan e Björneborn (2003) apontam que ela trata de avaliar os aspectos quantitativos da constituição e do uso da web, abrangendo quatro áreas de investigação: análise de conteúdo das páginas web; análise da estrutura dos weblinks; análise do uso da web (exploração dos programas que registram os comportamentos de pesquisa e busca na web) e análise do desempenho dos motores de busca.

Nesse sentido concorda-se com Thelwall (2004, p. 1) de que a “[...] análise de *links* adota e adapta técnicas existentes de informação científica à meta-análise dos documentos através da investigação das interconexões.” (tradução nossa). Entende-se que o uso de técnicas webométricas, baseadas em avaliações de caráter quali-quantitativo, proporciona subsídios à análise dos conteúdos. Os *links* não cumprem apenas uma função argumentativa na construção do sentido hipertextual. A lógica do *link* transcende o contexto citado, permitindo que o leitor tenha acesso a mais informações que podem ser desveladas de acordo com a navegação no hipertexto. O *link* funciona mais como a possibilidade de acesso a informações em outro contexto, do que como uma afirmação ou negação ao texto. Sendo assim, eles podem indicar tendências e alternâncias quanto aos objetos de estudo, as composições discursivas, aos canais adotados, aos públicos atingidos.

3 METODOLOGIA

O estudo circunscreve as análises dos periódicos brasileiros da área de Ciências Sociais Aplicadas 1 nos estratos A1 e A2 determinados pelo Qualis/CAPES 2008. São eles: Ciência da Informação (A2), Dados (A1), Educação e Sociedade (A2) e Perspectivas em Ciência da Informação (A2). As informações, que resultaram no corpus do estudo, basearam-se no *crawler* feito entre os dias 15 e 16 de maio de 2012.

Trata-se de um estudo exploratório que busca obter um mapeamento inicial em relação a totalidade de ligações (*outlinks*) estabelecidas a partir dos periódicos eletrônicos mencionados. Desse modo, utilizou-se o limite relativo de 5.000 páginas por URL visando observar de que maneira as ligações semânticas feitas entre os periódicos e o exterior, através dos *outlinks*, são tecidas. A escolha por privilegiar a análise dos *outlinks* justifica-se pelo interesse em vislumbrar como os periódicos científicos se relacionam com outros tipos de documentos *web*. Compreende-se, por *outlinks*, os *links* de uma página *web* que se liga (ou está ligada) a duas ou mais páginas concomitantemente. (BJÖRNEBORN, 2004).

A pesquisa utiliza técnicas cientométricas e webométricas baseado na análise de conteúdo, especificamente a análise categorial segundo Bardin (1977). O instrumento de extração dos dados relativos ao número de *links* foi o software SocSciBot4, que permitiu recuperar listas com o número total de *links* a partir da URL do periódico eletrônico.

4 RESULTADOS

Com base no limite estabelecido de 5.000 páginas por URL obteve-se um total de 220.949 *outlinks* dispersos em 9.967 páginas relativas aos quatro periódicos investigados. Cumpre ressaltar a participação da área de Ciência da Informação (50%) dentre os periódicos brasileiros mais bem colocados no Qualis/CAPES 2008 nas CSA1. O periódico Dados, pertencente às Ciências Sociais, apresenta a melhor classificação (A1). A área de Educação foi representada pelo periódico Educação & Sociedade (A2) que apresentou o menor número de *outlinks*. (ver Tabela 1).

Tabela 1 – Periódicos e seus Outlinks

Site	Pages	Outlinks
portaldeperiodicos.eci.ufmg.br%2Findex.php	4950	187379
revista.ibict.br%2Fciinf%2Findex.php	4960	33496
www.scielo.br/dados	29	38
www.scielo.br/es	28	36

Fonte: SocSciBot4

Observou-se que os dois periódicos da área de Ciência da Informação apresentaram os maiores números de *outlinks* (220.875). Essa fato demonstra um alto grau de conectividade desses canais com informações externas. Dentre os *outlinks* com maior incidência, identificou-se a presença de informações relacionadas ao movimento de Acesso Aberto. No periódico Perspectivas em Ciência da Informação esse percentual foi de 23,46%, já na Ciência da Informação foi ainda mais elevado (46,34%). Isso significa que 27,61% do total de *outlinks* dos dois veículos acima mencionados estabeleceu elo com iniciativas que se configuram como alternativas para o sistema de comunicação científica instituído.

Uma lista de categorias emergiram do estudo do fenômeno e embasaram a análise dos *outlinks* dos quatro periódicos. (ver Gráfico 1).

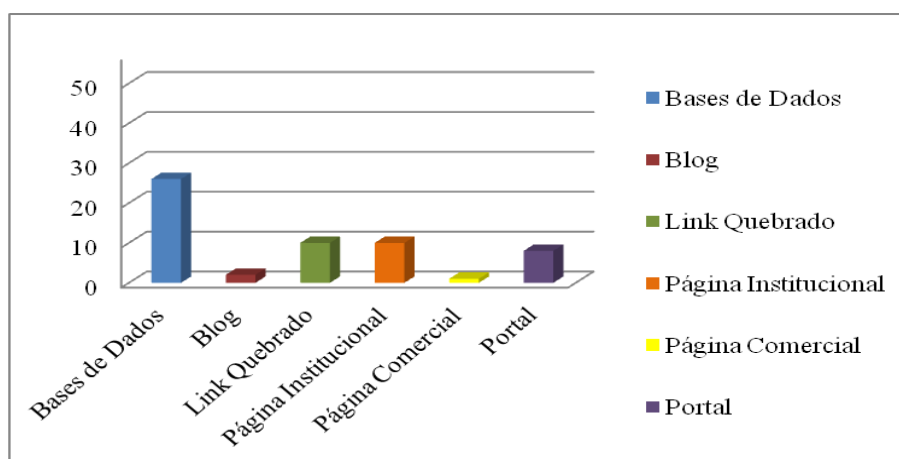


Gráfico 1 – Categorias dos *Outlinks*

Ao verificar as 9.967 URLs encontrou-se 71 diferentes *outlinks* distribuídos entre os quatro periódicos, sendo que 12 deles foram categorizados como *broken links*. Ao realizar a conferência desse universo, visando dirimir a duplicidade dos dados, obteve-se um número de 57 *outlinks* distintos e 10 *broken links*. Dentre eles, destaca-se a presença predominante de Bases de Dados (45,61%), o que reproduz uma estreita relação entre o canal e os indexadores. Em contrapartida, a baixa incidência de *outlinks* para *Blogs* (3,51%) pode ser explicada em razão desse meio ainda não estar inserido formalmente entres os canais referendados de comunicação científica.

No que concerne ao âmbito desses *outlinks*, 75,44% deles remetem para páginas que possibilitam acesso público às informações, esse dado reforça o caráter público da Ciência. (ZIMAN, 1979; MEADOWS, 1999). Para complementar a análise de âmbito, procurou-se verificar o Domínio/Extensão dos *outlinks*. Optou-se pela verificação da última unidade da URL. (ver Tabela 2).

Tabela 2 – Domínio dos Outlinks

Domínio/Extensão	Freq. Absoluta	Freq. Relativa (%)
.br	16	28,07%
.com	3	5,26%
.edu	2	3,51%
.gov	6	10,53%
.org	15	26,32%
Link Quebrado	10	17,54%
Internacionais	5	8,77%
TOTAL	57	100,00%

Fonte: dados da pesquisa

A predominância de URLs com extensão .br (28,07%) indica o estabelecimento de relações restritas ao âmbito nacional, reproduzindo, em certa medida, a necessidade de conferir visibilidade e referendar autoridade aos canais e instituições característicos e consagrados de cada área que compõe as CSA1. Outro dado que corrobora os resultados acima destacados está imbricado à incidência dos 15 *outlinks* de domínio .org.

Verificou-se também o Caráter da Informação veiculada nos *outlinks* objetivando contemplar o propósito firmado pelas relações hipertextuais efetivadas nesses periódicos. (ver Tabela 3).

Tabela 3 – Caráter da Informação

Caráter da Informação	Freq. Absoluta	Freq. Relativa (%)
Informação Científica	29	50,88%
Informação Utilitária	16	28,07%
Produtos e Serviços	2	3,51%
Link Quebrado	10	17,54%
TOTAL	57	100,00%

Fonte: dados da pesquisa

De forma não surpreendente, os *outlinks* encontrados versam, em sua maioria, a respeito de conteúdo acadêmico e/ou científico. Isso corrobora a identificação temática entre o canal e um universo de informações dispersas na *web*. Nesse sentido o *link*, desempenha o papel de dispositivo que atribui identidade ao canal, testemunhando em favor de sua qualidade e de sua afirmação como veículo de comunicação da Ciência.

5 CONSIDERAÇÕES FINAIS

Os dados foram quantificados e categorizados, o que permitiu a verificação de variações significativas no número de *outlinks* nos quatro periódicos analisados. Em relação ao periódico Dados, o único da amostra a figurar no estrato A1, a baixa ocorrência de *outlinks* não pode ser relacionada com a qualidade atestada pelo Qualis/CAPES 2008, visto que os *outlinks* pouco dizem sobre a visibilidade do canal. Essa afirmação encontra respaldo no fato de que o periódico Perspectivas em Ciência da Informação apresentou o maior número de *outlinks* mesmo estando no estrato A2.

Dentre as possibilidades de futuros estudos, elenca-se a substituição do centro de interesse da análise dos *outlinks* para os *inlinks*, objetivando identificar o impacto e a visibilidade dos periódicos ante a comunidade científica.

Diante das análises foi possível evidenciar que os periódicos investigados transitam entre perspectivas tradicionais, em relação as fontes de informação científica aos quais estão ligados, e inovadoras quanto as informações de caráter utilitário.

Em relação aos *outlinks* dos periódicos contemplados os mesmos permitiram compor um quadro analítico no que diz respeito a perspectivas que necessitam de investigações baseadas não apenas em categorias sucitadas, mas na combinação de outras metodologias que imbriquem diferentes dados e informações centradas nos atores que compõem o ciclo de comunicação científica.

Finalmente, sobre a metodologia empregada, acredita-se que as planilhas, geradas a partir do Excel, auxiliaram na padronização dos dados. Quanto à escolha do *software*



SocSciBot4, ela demonstrou-se adequada para os fins da pesquisa. Este programa possibilitou o crawling nos quatro periódicos verificados de modo confiável. As escolhas demonstraram estar de acordo com as expectativas e atenderam às necessidades da investigação.

REFERÊNCIAS

ALMIND, T., INGWERSEN, P. Informetric analyses on the World Wide Web: methodological approaches to "webmetrics". **Journal of Documentation**, London, v. 53, n. 4, p. 404-426, Sept. 1997.

BARDIN, L. **Análise de conteúdo**. São Paulo: Martins Fontes, 1977.

BJÖRNEBORN, L.; INGWERSEN, P. Perspectives of webometric. **Scientometrics**, Budapest, v. 50, n. 1, p. 65-82, 2001.

MEADOWS, A. J. **A Comunicação Científica**. Brasília, DF: Briquet de Lemos/Livros, 1999.

MENDONÇA, Marília Alvarenga Rocha *et. al.* Avaliação de Periódicos Eletrônicos Brasileiros em Ciência e Tecnologia: uma proposta de metodologia baseada na análise de "links" para o "site" do periódico. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 13., 2004, Natal. **Anais**, 2004. [1 CDROM].

THELWALL, M., VAUGHAN, L; BJÖRNEBORN, L. Webometrics. **Annual Review of Information Science and Technology**, Medford, v. 39, 2003.

THELWALL, M. **Link analysis: an information science approach**. Amsterdam: Elsevier Academic, 2004.

ZIMAN, John Michael. **Conhecimento Público**. Belo Horizonte: Itatiaia, 1979. (Coleção O Homem e a Ciência, 8).