

AS INSTITUIÇÕES ARQUIVÍSTICAS NA INTERNET: uma análise webométrica a partir dos arquivos nacionais da ibero-américa na World Wide Web

Thiago de Oliveira Vieira

RESUMO: O uso da internet permite às instituições arquivísticas alcançarem um público muito mais amplo, rompendo com as fronteiras físicas das instituições. Para que isso seja possível, é necessário que as instituições possuam *sites* que forneçam acesso às suas informações institucionais e disponibilizem conteúdos, recursos audiovisuais, bases de dados, instrumentos de pesquisa e representantes digitais acerca de seus acervos custodiados. Além de um *site*, é preciso um conjunto de estratégias para que o *site* possua presença e visibilidade na *World Wide Web* (*WWW*) e com isso, alcance de fato um número maior de usuários. Esta pesquisa procura averiguar a presença e os relacionamentos dos *sites* das instituições arquivísticas nacionais, dos países Ibero-americanos, na *WWW*, por meio da webometria. Para isso, realizou-se um aporte teórico que fundamente a métrica utilizada na análise do fluxo informacional na *web*, por meio dos *websites*. O percurso metodológico consistiu na identificação dos arquivos nacionais, dos países que constituem a América Ibérica, que possuem *sites* na *WWW*. A partir da identificação dos *sites*, realizou-se uma análise webométrica, buscando verificar os seguintes indicadores: tamanho do *site*, visibilidade (*inlinks*), luminosidade (*outlinks*) e o *Web Impact Factors*. Semelhante à relação de citações e referências na medição de impacto na comunicação científica, os *inlinks* e *outlinks*, respectivamente, propiciam o estabelecimento de relações com outros *sites* e páginas *web*, potencializando a sua presença na *WWW* e permitindo um maior número de pontos de acesso a um determinado *site*. O tamanho do *site*, calculado por meio da soma de todas as páginas que compõem um *site* também é um indicador importante para identificar a presença de um *site* na *WWW*, mas altamente dependente das relações de *inlinks* e *outlinks*. O *Web Impact Factors*, criado por Ingwersen (1998), estabelece uma relação entre a visibilidade e o tamanho de um *site*. Os resultados apontam uma presença considerável das instituições arquivísticas analisadas nos espaços virtuais emergentes da *WWW*, mas ainda carecem de aumentar seus relacionamentos com o universo de *sites* que constituem a *WWW*, buscando maior alcance e relevância neste espaço informacional. A partir dos resultados, percebe-se que os *sites* das instituições arquivísticas ibero-americanas são, em média, mais citados por outros *websites* do que eles

Thiago de Oliveira Vieira

thiagoov@globo.com

<http://lattes.cnpq.br/6225474368983715>

Arquivista do Arquivo Nacional (Brasil) e Membro da Câmara Técnica de Documentos Audiovisuais, Iconográficos e Sonoros do Conselho Nacional de Arquivos (CONARQ). Mestre (2014) pelo Programa de Pós-Graduação em Gestão de Documentos e Arquivos (PPGARQ) e graduado em Arquivologia (2006) pela Universidade Federal do Estado do Rio de Janeiro (UNIRIO).

Submetido em: 21/04/2018

Publicado em: 27/01/2019

próprios referenciam outros *sites* externos. Apesar de altamente dinâmicos, os indicadores webométricos podem auxiliar as instituições arquivísticas a potencializarem seus *websites*, aspirando atingir um número cada vez maior de usuários na *WWW*.

PALAVRAS-CHAVE: Webometria. *World Wide Web*. Instituição Arquivística. Arquivo Nacional.

1 INTRODUÇÃO

O surgimento do Arquivo Nacional da França, em plena Revolução Francesa, em 1789, marca um novo modo de "reorganização" dos arquivos produzidos pelo Estado. O modelo Francês, que estabelece uma instituição arquivística como órgão responsável pela guarda, preservação e acesso aos documentos produzidos pela administração pública, foi amplamente reproduzido por outros países (FONSECA, 2005).

Além do marco inaugural da relação arquivo-Estado com a criação do Arquivo Nacional Francês e o reconhecimento do dever do Estado com a preservação do patrimônio documental, a Revolução Francesa inaugura, ainda que não em primeiro momento, o direito de acesso público aos arquivos por todos os cidadãos (DUCHEIN, 1983; FONSECA, 2005; SILVA et al., 2009).

Desde então, ao longo do tempo, as instituições arquivísticas passaram por diferentes (re)configurações, fruto das mudanças administrativas do Estado, da ampliação dos direitos de acesso e uso dos arquivos para o exercício da cidadania e transparência do Estado, do desenvolvimento das técnicas de organização e acesso aos documentos e por fim, do desenvolvimento das Tecnologias da Informação e Comunicação (TIC). Nesse sentido, as instituições arquivísticas são o reflexo das transformações da sociedade e devem estar em constante processo de remodelagem.

No que tange ao surgimento e evolução das TIC, é possível perceber que os impactos nas instituições arquivísticas são bastante significativos, principalmente na relação arquivo (instituição) e sociedade. Para Jardim e Fonseca (2004, p. 2), um serviço de informação que centra-se no usuário, pressupõe considerar a importância da informação para quem a pesquisa e utiliza. Esta percepção induz a criação de novos espaços informacionais, longe dos tradicionais meios físicos de acesso à informação.

A necessidade de novas formas de interações entre arquivo e sociedade leva às instituições arquivísticas a reverem seus

modelos. Este novo modelo vigente pressupõe a utilização de diversos espaços virtuais, de ferramentas digitais para proporcionar o acesso aos documentos sob sua custódia (garantindo a transparência, o exercício da cidadania, o incentivo à produção de conhecimento e a garantia de direitos), além da constituição de um canal bidirecional entre a instituição e seus usuários.

Para Jardim e Fonseca (2004, p. 1), "a emergência de práticas inéditas de produção, transferência e uso da informação abrem possibilidades ilimitadas para seu uso e oferta, envolvendo um conjunto cada vez mais amplo e "anônimo" de indivíduos". Nesse sentido, as tecnologias de informação e comunicação impõem novas demandas dos usuários aos arquivos, provocando uma realocação ou desaparecimento de fronteiras físicas, reforçando com isso a ideia de um novo modelo de instituição arquivística que emerge dos espaços físicos aos espaços virtuais (JARDIM; FONSECA, 2004; MARIZ, 2005).

Estes espaços virtuais podem ser constituídos por meio de *sites* (também conhecidos como *websites* ou conjunto de páginas *web*) que são um servidor ou endereço que se encontra na *World Wide Web (WWW)* (MICROSOFT, 2000, p. 360) e disponibilizam um conjunto de documentos e informações de uma instituição, pessoa ou localidade aos usuários da internet, denominado de *content provider* (fornecedor de conteúdo) (MICROSOFT, 2000, p. 79).

É importante distinguir as diferenças entre a internet e a *WWW*. A internet se constitui como "agrupamento mundial de redes e portas de ligação que utilizam o conjunto de protocolos TCP/IP [*Transmission Control Protocol/Internet Protocol*] para comunicar umas com as outras" (MICROSOFT, 2000, p. 179–180), enquanto a *WWW* é um "conjunto de todos os documentos de hipertexto interligados que residem em servidores HTTP [*Hypertext Transfer Protocol*] em todo o mundo. Os documentos da *WWW* (designados por páginas ou páginas Web) são escritos em HTML [*HyperText Markup Language*] e identificados por URL [*Uniform Resource Locator*]" (MICROSOFT, 2000, p. 364). Portanto, a *WWW* é apenas um formato de comunicação utilizado na internet.

A *WWW*, enquanto espaço de informação, proporciona inúmeras possibilidades aos arquivos, constituindo um desafio às instituições arquivísticas a repensarem suas ações teórico-práticas (MARIZ, 2005). Um *site* de uma instituição arquivística deve ser visto como um canal de prestação de serviços dinâmico e atualizável. Deve ser encarado como um espaço virtual entre a instituição e seus usuários, como parte de uma política de informação institucional (CONSELHO NACIONAL DE ARQUIVOS (BRASIL), 2000).

A partir da década de 1990, é possível perceber que as instituições arquivísticas passaram a utilizar o ambiente *web*, como espaço facilitador, para transferência da informação arquivística sob custódia das instituições (OLIVEIRA; MATOS, 2013).

A problemática que se coloca é que esses espaços, em particular os *sites*, para além do uso de tecnologias digitais de maneira a proporcionar o acesso rápido, fácil e interativo às informações, necessitam estar relacionados com outros *sites* para promover uma comunicação dinâmica e o compartilhamento de informações, visando aumentar sua presença e visibilidade na rede.

O objetivo do artigo é verificar a presença das instituições arquivísticas na *WWW*, por meio dos *websites* e analisar os *sites* que mais se destacam a partir dos indicadores webométricos, tendo como universo os arquivos nacionais dos países da América Ibérica (ORGANIZAÇÃO DOS ESTADOS IBERO-AMERICANOS, [s.d.]). A escolha destes países como campo de coleta de dados deve-se a sua representatividade geográfica, abrangendo países de três continentes (África, América e Europa), de línguas espanhola e portuguesa. Justifica-se a escolha desta análise webométrica, pois estes indicadores propiciam uma avaliação da presença e do relacionamento dos *sites* das instituições arquivísticas com a *web*, por meio da análise de *links* (internos e externos). Constituem uma importante ferramenta para aferir e potencializar o alcance dos serviços prestados pelos arquivos nas suas páginas *web*, por meio de estratégias que permitam aumentar seus relacionamentos com a *WWW* e com isso, abranger um número maior de pessoas.

Cabe ressaltar que a webometria, e seus indicadores, é utilizada como uma métrica para a *WWW* (AGUILLO, 2012; AGUILLO et al., 2010; ALMIND; INGWERSEN, 1997; BJÖRNEBORN, 2004; BJÖRNEBORN; INGWERSEN, 2004; INGWERSEN, 1998; MARQUES, 2013; SMITH, 1999; THELWALL, 2007, 2009; THELWALL; VAUGHAN; BJÖRNEBORN, 2005; VANTI, 2002, 2005, 2007, 2010), sendo utilizada inclusive como métrica de avaliação de Universidades, como é o caso do *Ranking Web of Universities*, que tem como um de seus objetivos aferir a visibilidade das universidades, por meio de indicadores como número de páginas *web* e *links* externos recebidos.

Procura-se, a partir deste estudo, para além do mapeamento dos *websites* das instituições arquivísticas nacionais dos países Ibero-americanos, com base na webometria, demonstrar a potencialidade desta análise na concepção e construção dos *websites* de instituições arquivísticas nacionais, visando potencializar sua presença e relacionamentos com a *WWW*.

2 A WEBOMETRIA COMO MÉTRICA PARA AVALIAÇÃO DO FLUXO INFORMACIONAL E DO CONHECIMENTO CIENTÍFICO NA WEB

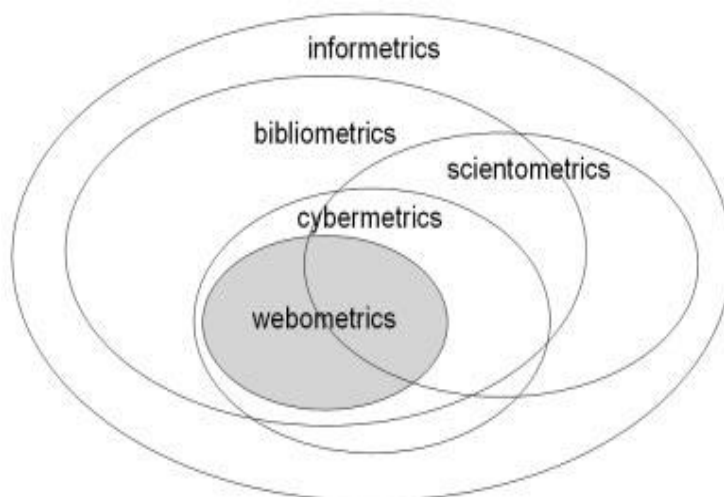
Há diversos métodos quantitativos que permitem avaliar o fluxo informacional e do conhecimento científico: a bibliometria, a cienciometria, a informetria, a webometria e a cibermetria. Este trabalho centra-se na webometria, nesse sentido, as outras métricas podem ser exploradas em: AGUILLO et al., 2006; BORGMAN; FURNER, 2002; HOOD; WILSON, 2001; MACIAIS-CHAPULA, 1998; PRITCHARD, 1969; THELWALL, 2007.

A partir de enfoques diferenciados, estes métodos permitem medir dados específicos de um *corpus* de conhecimento.

A expansão das tecnologias digitais permitiu a expansão do uso e disseminação da ciência e da academia a partir da rede mundial de computadores. Nesse sentido, foram estudadas métricas para quantificação científica neste espaço informacional emergente. É neste contexto que surge a webometria que enfoca nos métodos informativos, enquanto estudos quantitativos da arquitetura e utilização de recursos informacionais e tecnologias voltados para a WWW (ALMIND; INGWERSEN, 1997; BJÖRNEBORN, 2004).

Björneborn (2004) estabelece as relações entre as métricas quantitativas da área de biblioteconomia e ciência da informação, inserindo o espaço e as relações da webometria em relação às outras métricas.

Figura 1: Relações entre as métricas quantitativas da área de biblioteconomia e ciência da informação



Fonte: Björneborn (2004).

Thelwall (2009) destaca que a webometria ocupa-se da medição de aspectos *web* (*sites*, páginas da *web*, *hyperlinks*, resultados de mecanismos de pesquisa na *web*).

A importância da web em si como um meio de comunicação e de hospedar um leque cada vez maior de documentos, desde artigos de revistas a folhetos de destino de férias, não precisa de introdução. Dada essa enorme e facilmente acessível fonte de informação, existem possibilidades ilimitadas de medir ou contar em larga escala (por exemplo, o número de *sites*, o número de páginas da Web, o número de blogs) ou em menor escala (por exemplo, o número de *sites* na Irlanda, o número de páginas da web no *site* da CNN, o número de blogs mencionando Barack Obama antes da campanha presidencial de 2008) (THELWALL, 2009, p.1).

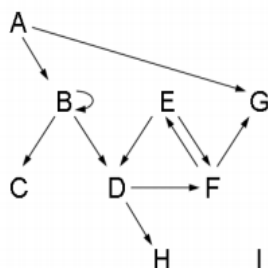
A linguagem HTML, amplamente utilizada na construção de páginas *web* (MICROSOFT, 2000), é para além de uma linguagem de formatação, também muito utilizada para pesquisa e recuperação da informação na *web*, além de permitir realizar análises informativas (ALMIND; INGWERSEN, 1997). Para os autores, “hoje o HTML é amplamente utilizado em um contexto de recuperação de informação (IR), mas muito pouco em um contexto de análise” (ALMIND; INGWERSEN, 1997).

A webometria tem como um de seus métodos a análise quantitativa de *hiperlinks* (THELWALL, 2007). Ainda segundo o autor, o uso dos estudos quantitativos de *links* foi desencadeado por Ingwersen (1998), com o Fator de Impacto Web (FIW), criado em analogia ao *Journal Impact Factor* (JIF), situando a relação entre um *link* a uma citação (THELWALL, 2007).

Um *link*, ou *hiperlink*, é uma “ligação entre um elemento de um determinado documento de hipertexto, como uma palavra, expressão, símbolo ou imagem, e outro elemento do mesmo documento, doutro documento de hipertexto, de um ficheiro ou de um script” (MICROSOFT, 2000, p.166). Um *link* também é utilizado para conectar um *site* ou páginas na *web* (SAWAYA, 1999). Nesse sentido, os *links* estabelecem conexões entre diversos *sites* na WWW.

Procurando estabelecer uma conceituação terminológica de *link*, para uso em pesquisas webométricas, Björneborn (2004) define as ligações básicas de *links*:

Figura 2: Relações básicas entre links



- B has an *inlink* from A; B is *inlinked*; A is *inlinking*; A is an *in-neighbor* of B
- B has an *outlink* to C; B is *outlinking*; C is *outlinked*; C is an *out-neighbor* of B
- B has a *selflink*; B is *selflinking*
- A has no inlinks; A is *non-linked*
- C has no outlinks; C is *non-linking*
- I has neither in- nor outlinks; I is *isolated*
- E and F have *reciprocal links*; E and F are *reciprocally linked*
- D, E and F have in- or outlinks connecting each other; they are *triadically interlinked*
- A has a *transversal* outlink to G: functioning as a shortcut
- H is *reachable* from A by a directed *link path*
- C and D are *co-linked* by B; C and D have *co-inlinks*
- B and E are *co-linking* to D; B and E have *co-outlinks*
- Co-inlinks and co-outlinks are both cases of *co-links*

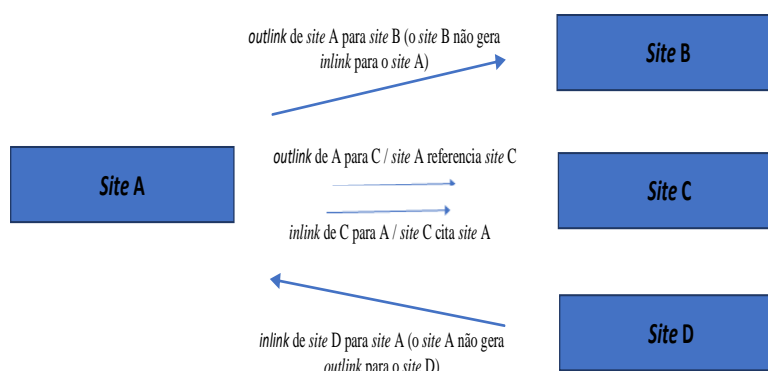
Fonte: Björneborn (2004).

A análise de *links* utiliza um algoritmo para observar *links* de páginas *web*. Os *links* que se relacionam entre as páginas *web* constituem uma fonte potencial de conhecimento para vários fins, incluindo uma classificação de qualidade e autoridade entre páginas similares ou uma comunidade na internet (KHOSROWPOUR, 2013).

Os *inlinks* são ligações que uma página *web* ou *site* recebe de outra página *web* ou *site* e os *outlinks* são ligações que uma página *web* ou *site* faz para outra página *web* ou *site*. Importa perceber que os *inlinks* e *outlinks* são *links* que estabelecem ligações de texto, imagem ou objeto para outra página *web*, *site* ou objeto (MARQUES, 2013).

Segundo Thelwall, Vaughan e Björneborn (2005), se comparado com a bibliometria, os *outlinks* fazem uma analogia ao termo referência e os *inlinks* fazem uma analogia ao termo citação.

Figura 3: Relações entre os *sites* na web por meio de links (*inlinks* e *outlinks*)



Fonte: Elaborado pelo autor.

A partir da figura 3, é possível observar que um determinado *website* pode estabelecer relações bidirecionais com outro *website* (o *site* A aponta para o *site* C, assim como o *site* C aponta para o *site* A) ou apenas uma relação de apontamento simples, onde um *site* aponta - *outlink* - para um *site* externo ou é apontado - *inlink* - por outro *site* (o *site* A aponta o *site* B por meio de *outlink* e o *site* A é apontado, por meio de *inlink*, pelo *site* D).

No Quadro 1 explicita-se, segundo Vanti (2002), o objeto de estudo, as variáveis, os métodos e os objetivos da webometria.

Quadro 1: Objeto, variáveis, métodos e objetivos da webometria

WEBOMETRIA	
Objeto de estudo	Sítio na WWW (URL, título, tipo, domínio, tamanho e links), motor de busca.
Variáveis	Número de páginas por sítio, número de links por sítio, número de links que remetem a um mesmo sítio, número de sítios recuperados.
Métodos	Fator de impacto da <i>web</i> , densidade dos links, "situações", estratégia de busca.
Objetivos	Avaliar o sucesso de determinados sítios, detectar a presença de países, instituições e pesquisadores na rede e melhorar a eficiência dos motores de busca na recuperação das informações.

Fonte: Vanti (2002, p. 160).

As limitações da aplicação das métricas para avaliação dos fluxos informacionais na *web* residem, como apontado por Almind e Ingwersen (1997), na falta de padronização de forma e conteúdo nas páginas *web* e a dinâmica de atualização no ambiente *web*, que opera em "tempo real". Apesar disso, os autores, afirmam que a WWW é bem projetada para pesquisas informétricas de *links*. Sendo os *links* uma ligação entre

"objetos" na *web*, estes são facilmente acessíveis (ALMIND; INGWERSEN, 1997).

3 METODOLOGIA

Esta pesquisa caracteriza-se como descritiva (GIL, 2008) de cunho quantitativo, utilizando os indicadores da webometria para coleta de dados, com base em AGUILLO, 2012; AGUILLO et al., 2010; INGWERSEN, 1998; SHINTAKU; ROBREDO; BAPTISTA, 2011; SMITH, 1999; VANTI, 2002, 2005, 2007, 2010. Os indicadores webométricos selecionados são apresentados no Quadro 2.

Foram aplicados os indicadores da webometria nos *sites* das instituições arquivísticas, com o objetivo de aferir o tamanho do *site* (número de páginas de um *site*), seus *inlinks* (visibilidade - número total de *links* externos únicos recebidos por um *site*), seus *outlinks* (luminosidade - número de *links* externos que um *site* apresenta) e seu fator de impacto *web* (proporção de *links* que um *site* recebe em relação ao seu tamanho).

Quadro 2: Indicadores webométricos

INDICADORES	CONCEITO/MEDIÇÃO
Tamanho do <i>site</i>	O tamanho <i>site</i> corresponde ao número total de páginas que ele apresenta no seu domínio, calculado por meio de um motor de busca (AGUILLO, 2012; AGUILLO et al., 2010; SHINTAKU; ROBREDO; BAPTISTA, 2011; VANTI, 2002, 2005, 2007, 2010) Indicadores do tamanho do <i>site</i> foram extraídos por meio de pesquisa avançada do Google, com a seguinte expressão "domínio <i>site</i> :domínio", onde o domínio representa o endereço <i>web</i> do <i>site</i> a ser aferido.
Visibilidade	O número de <i>links</i> externos recebidos (<i>inlinks</i>) pelo sítio analisado, desconsiderados os <i>links</i> internos e auto <i>links</i> (AGUILLO, 2012; SHINTAKU; ROBREDO; BAPTISTA, 2011; VANTI, 2002, 2005, 2007, 2010). A visibilidade foi apurada por meio do <i>site</i> https://ahrefs.com , considerando um <i>inlink</i> para cada domínio de referência, a partir do filtro "live" (um <i>inlink</i> é considerado "vivo" até voltar a rastrear a página de ligação e ver o contrário - o que geralmente nos leva de algumas horas a poucas semanas). Este filtro permite obter uma medição de <i>inlinks</i> mais atual, dado que as ligações entre <i>links</i> na <i>web</i> são dinâmicas.
Luminosidade	O número de <i>links</i> externos (<i>outlinks</i>) que apresenta um sítio (não pode ser confundido com o número de páginas ligadas, visto que uma página pode conter múltiplos <i>outlinks</i>), apontando para outras URLs diferentes (SHINTAKU; ROBREDO; BAPTISTA, 2011; VANTI, 2002, 2005, 2007, 2010). A luminosidade foi medida por meio do <i>site</i> https://moz.com/ , considerando o total de <i>outlinks</i> (<i>links</i> externos), no nível da página, excluindo os <i>links</i> do mesmo domínio raiz.
<i>Web Impact Factors -WIF</i> (INGWERSEN, 1998)	O <i>WIF</i> de um espaço na <i>web</i> é o número de páginas que se vinculam a um espaço na <i>web</i> , dividido pelo número de páginas no espaço da <i>web</i> (SMITH, 1999). Para o autor, existem três formas distintas de cálculo de <i>WIF</i> ("external <i>WIF</i> ", "self-link <i>WIF</i> " e um que combina os dois anteriores). Nesta pesquisa será utilizado o "external <i>WIF</i> ", pois segundo (SMITH, 1999), eles são mais significativos. O fator de impacto <i>web</i> indica a proporção de links que uma homepage recebe em relação ao seu tamanho ou número total de páginas e é o resultado da divisão do número de páginas que "linkam" determinados <i>sites</i> (indicador visibilidade) pelo número de páginas do <i>site</i> "linkado" (indicador tamanho do <i>site</i>) (VANTI, 2010).

Fonte: Elaborado pelo autor, a partir de Aguillo (2012); Aguillo et al. (2010); Ingwersen (1998); Shintaku; Robredo; Baptista (2011); Smith (1999); Vanti (2002, 2005, 2007, 2010).

O *corpus* desta pesquisa são as instituições arquivísticas nacionais dos países Ibero-americanos. O quadro 3 apresenta a relação de países que constituem a América Ibérica, além das suas respectivas instituições arquivísticas nacionais e endereços *web*.

Quadro 3: Instituições arquivísticas nacionais dos países Ibero-Americanos

Países	Instituição Arquivística	Website
Argentina	Archivo General de la Nación	www.agnargentina.gob.ar/
Bolívia	Archivo y Biblioteca Nacionales de Bolivia	www.archivoybibliotecanacionales.org.bo/
Brasil	Arquivo Nacional	www.arquivonacional.gov.br
Chile	Archivo Nacional	www.archivonacional.cl/
Colômbia	Archivo General de la Nación	www.archivogeneral.gov.co/
Costa Rica	Archivo Nacional de Costa Rica	www.archivonacional.go.cr/
Cuba	Archivo Nacional de la República de Cuba	www.arnac.cu/
El Salvador	El Archivo General de la Nación	www.cultura.gob.sv/agn/
Equador	Archivo Nacional del Ecuador	ane.mas.ec/
Espanha	Há uma rede de arquivos e centros estatais com características nacionais.	
Guatemala	Archivo General de Centro América	<i>website</i> não localizado ou inexistente
Guiné Equatorial	Instituição não localizada ou inexistente	<i>website</i> não localizado ou inexistente
Honduras	Archivo Nacional de Honduras	www.cdihh.ihah.hn/
México	Archivo General de la Nación	www.gob.mx/agn
Nicarágua	El Archivo General de la Nación	www.inc.gob.ni/
Panamá	Archivo Nacional de Panamá	www.archivonacional.gob.pa/
Paraguai	Archivo Nacional de Asuncion	www.archivonacionaldeasuncion.org/
Peru	Archivo General de la Nación	agn.gob.pe/
Portugal	Arquivo Nacional da Torre do Tombo	antt.dglab.gov.pt/
República Dominicana	Archivo General de la Nación	www.agn.gov.do/
Uruguai	Archivo General de la Nación	www.agn.gub.uy/
Venezuela	Archivo General de la Nación	www.agn.gob.ve/

Fonte: Elaborado pelo autor.

Guatemala e Guiné Equatorial foram excluídos do universo da pesquisa, pois não possuem (ou não foram localizados) *website*. Os arquivos nacionais de Nicarágua e de El Salvador não possuem um *site* próprio, sendo apenas uma página dentro de um conjunto de páginas (*site*), do Instituto Nicaraguense de Cultura e da Secretaría de Cultura de la Presidencia de la República de El Salvador, respectivamente, razão pela qual serão excluídos da amostra da pesquisa.

A Espanha possui um conjunto de dez arquivos e dois centros estatais, geridos pela Subdireção Geral de Arquivos Estatais, subordinados a Secretaria de estado de Cultura. Esta pesquisa pretende analisar os indicadores webométricos de uma instituição nacional por país, nesse sentido, os indicadores de apenas um arquivo ou centro não irá refletir a realidade nacional, razão pela qual será excluída do universo apresentado.

O Archivo Nacional de Honduras é ligado ao Centro Documental de Investigaciones Históricas de Honduras CDIHH, no âmbito do Instituto Hondurenõ de Antropología e Historia (IHAH). O CDIHH é constituído pelo Archivo Nacional de Honduras (ANH), el Archivo Etnohistórico y la Biblioteca Especializada "Jesús Núñez Chinchilla" del IHAH.

Conforme o exposto acima, dos 22 países que faziam parte do universo inicial da pesquisa, somente serão aplicados os indicadores webométricos em *websites* de instituições arquivísticas de 17 países que compõem a Ibero-América.

Uma primeira coleta de dados foi realizada nos dias 24 e 25 de outubro de 2017 objetivando verificar a viabilidade da pesquisa por meio destes indicadores webométricos. Uma nova coleta de dados foi realizada entre os dias 23 e 30 de dezembro de 2017.

Na próxima seção estão expostos os resultados da coleta de dados, conforme critérios estabelecidos acima.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

No Quadro 4 são apresentados os resultados da coleta de dados, realizada entre os dias 23 e 30 de dezembro de 2017.

Quadro 4: Indicadores webométricos das instituições arquivísticas nacionais dos países Ibero-Americanos

Países	Tamanho do <i>site</i>	Visibilidade	Luminosidade	Fator de Impacto <i>Web</i>
Argentina	141	1.510	12	10,71
Bolívia	2.440	390.935	30.224	160,22
Brasil	1.880	207.605	16.534	110,43
Chile	3.910	140.588	15.729	35,96
Colômbia	5.350	338.126	28.782	63,20
Costa Rica	7.060	3.566	140	0,51
Cuba	569	11.135	3.253	19,57
Equador	77	2	171	0,03
Honduras	4.830	32	0	0,01
México	1.230	70.842	401	57,60
Panamá	2.390	1.210	420	0,51
Paraguai	24	98	4	4,08
Peru	608	4.475	444	7,36
Portugal	757	113.712	101.000	150,21
República Dominicana	2.760	69.352	524	25,13
Uruguai	205	3.128	699	15,26
Venezuela	148	16.600	567	112,16

Fonte: Elaborado pelo autor.

A partir dos indicadores webométricos, sintetizou-se os dados no quadro abaixo para análise e discussão dos resultados.

Quadro 5: Síntese dos indicadores webométricos

	Maior Indicador	Menor Indicador	Média dos Indicadores dos 17 países	Número de instituições ACIMA da média dos indicadores	Número de instituições ABAIXO da média dos indicadores
Tamanho do site (em páginas)	7060	24	2022	7	10
	Archivo Nacional de Costa Rica	Archivo Nacional de Asuncion, (Paraguai)			
Visibilidade (em inlinks)	390.935	2	80.759	5	12
	Archivo y Biblioteca Nacionales de Bolivia	Archivo Nacional del Ecuador			
Luminosidade (em outlinks)	101.000	0	11.700	5	12
	Arquivo Nacional da Torre do Tombo (Portugal)	Archivo Nacional de Honduras			
Fator de Impacto Web	160,22	0,01	45,46	6	11
	Archivo y Biblioteca Nacionales de Bolivia	Archivo Nacional de Honduras			

Fonte: Elaborado pelo autor.

O tamanho do *site* é um indicativo do espaço que o *site* ocupa na *WWW*, ou seja, quanto maior o *website*, maior tende a ser a sua relevância entre as instituições com as mesmas funções, pois maior é o número de ferramentas/informações disponíveis para acesso online. Em se tratando de instituições arquivísticas, quanto maior o *site*, maior é o número de interações e acervos disponíveis para os usuários.

Além disso, o tamanho do *site* pode influenciar nos indicadores de visibilidade e luminosidade, ou seja, quanto maior o número de páginas, maior é o estabelecimento de relações entre *sites* na *web*, por meio dos *links*.

Como exemplo, utilizando os dados acima, o Archivo Nacional de Costa Rica possui 7.036 páginas a mais do que o Archivo Nacional de Asuncion, ou seja, são 7.036 páginas a mais para estabelecerem *inlinks* e *outlinks* com outros *sites* na *web*.

A visibilidade tem como objetivo medir o número total de *links* externos recebidos (*inlinks*) por um *site*. O número de *inlinks* externos recebidos por um domínio é uma medida que representa visibilidade e impacto do material publicado na *WWW*.

A luminosidade mostra como as páginas das instituições arquivísticas se relacionam com outras páginas na *web*, ou seja, como elas se interligam por meios de *outlinks* e como elas evidenciam conteúdos externos.

Chama atenção o caso do Arquivo Nacional da Torre do Tombo, a instituição possui números similares de visibilidade e luminosidade, mesmo com um número abaixo da média no indicador tamanho do *site*. Um alto índice de luminosidade indica que o *site* da instituição aponta para muitos *links* externos na *web*, detendo um alto relacionamento com a *WWW*.

Cabe frisar que o Archivo Nacional de Honduras é a terceira maior instituição em tamanho do *site*, com 4.830 páginas. O tamanho do *site*, neste caso, não contribuiu com um indicador elevado de *outlinks*.

Os dados coletados revelam que o indicador visibilidade é muito maior do que o indicador luminosidade. Estes números indicam que é muito mais fácil ser "citado" pela *web* (relação de muitos para um, ou seja, muitos *sites* da *WWW* para um *site* na *WWW*) do que "referenciar" a *web* (relação de um para muitos, ou seja, um *site* da *WWW* para muitos *sites* da *WWW*).

O *WIF*, que é o resultado da divisão dos indicadores de visibilidade pelo indicador de tamanho do *site*, demonstra o impacto de uma página por meio da relação entre o tamanho da página e o número de *links* recebidos por outras páginas na *web*.

Neste caso, o indicador *WIF* afere a força de um *site* na *WWW* por meio de um cálculo de proporção 1:1, procurando medir o quão o *site* é *linkado* na *WWW*, levando em consideração o seu tamanho.

Ressalta-se que o Archivo y Biblioteca Nacionales de Bolivia, com maior indicador de *WIF*, possui o maior número de *inlinks* (visibilidade) e a sexta posição do indicador tamanho do *site*, com 2.440 páginas (próximo da média do indicador, que é de 2.022 páginas). Apesar do Archivo y Biblioteca Nacionales de Bolivia possuir um número de páginas próximo da média das instituições analisadas nesta pesquisa, a instituição possui uma alta quantidade de *inlinks* apontados para o seu *website*. Além disso, o Archivo Nacional de Honduras, com menor *WIF*, é a terceira maior instituição em tamanho do *site*, mas com apenas

32 *links* apontados para o seu *website*, sendo a segunda instituição com o menor indicador de visibilidade.

5 CONCLUSÃO

Analisar uma instituição arquivística por meio da análise de *links* dá uma dimensão dos relacionamentos e impactos de suas páginas e seus conteúdos na *WWW*. Quanto mais *links* apontados para um *website*, por outros *sites*, maior é o seu alcance na *WWW* e maior é a sua relevância em um mecanismo de pesquisa na *web*.

As instituições arquivísticas publicam em seus *websites* conteúdos geralmente relativos aos acervos sob sua custódia ou suas ações geralmente relacionadas às suas funções/atividades. Assim, este indicador representa o quanto o conteúdo publicado pelas instituições arquivísticas são *linkados* por outras páginas *web*, não sendo necessariamente instituições congêneres.

Quanto maior o *site* (número de páginas), maior tende a ser a proporção de *inlinks* e *outlinks*, em outras palavras, maior tende a ser a visibilidade e luminosidade do *site*. A razão disso reside no fato de que quanto mais páginas um *site* tem, maior é a probabilidade de ele apontar para outro *site* ou ser apontado por outro *site*. Nos indicadores coletados, de uma maneira geral, as instituições com maiores números de páginas obtiveram maiores indicadores de visibilidade e luminosidade.

A partir do que foi analisado, corrobora-se com as afirmações de Vanti e Thelwall et al. que apontam uma relação entre a visibilidade na *web* (*inlinks*) e "ser citado" e a luminosidade (*outlinks*) e "citar" na bibliometria (THELWALL; VAUGHAN; BJÖRNEBORN, 2005; VANTI, 2010).

Todos os indicadores utilizados na coleta de dados possuíram como objetivo mensurar a relação dos *websites* com a *web*. Nesse sentido, observou-se que as métricas webométricas constituem uma métrica importante para aferir informações na *web*, apesar de a *web* ser um ambiente complexo e instável. Estas métricas podem colaborar nas ações de elaboração e manutenção dos *websites* por parte das instituições arquivísticas.

Percebeu-se entre as duas coletas de dados (um intervalo de 60 dias) uma alteração significativa em alguns indicadores de determinadas instituições arquivísticas, que revelam a complexidade e instabilidade deste espaço informacional emergente. Apesar dessas limitações, a webometria, como campo da Ciência da Informação (VANTI, 2010), pode auxiliar as instituições arquivísticas a potencializar o acesso às informações em seus *websites*, apoiando as decisões

governamentais, o cidadão na defesa de seus direitos e o incentivo à produção de conhecimento científico e cultural.

ARCHIVAL INSTITUTIONS ON THE INTERNET: a webometric analysis from the national archives of ibero-america on the World Wide Web

ABSTRACT: The use of the internet allows to the archival institutions reach a much wider public, breaking up with its physical borders. For that to be possible, it is necessary that the institutions possess websites that supply access to their institutional information and make available contents, audiovisual resources, data base, research instruments and digital representatives concerning their guarded fonds. Besides a website, it is necessary a group of strategies for the website to possess presence and visibility in World Wide Web (WWW) and with that, reach a larger number of users. This research aims to verify the presence and the relationships between the websites of the national archival institutions of the Ibero-American countries in the WWW, through the webometry. For that, a theoretical contribution was made that bases the metric used in the analysis of the informational flow in the web, through the websites. The methodological route consisted in the identification of the national archives, of the countries that constitute Iberian America, that have websites in the WWW. As of the identification of the sites, a web-based analysis was carried out, seeking to verify the following indicators: website size, visibility (inlinks), luminosity (outlinks) and Web Impact Factors. Similar to the relation of citations and references in the measurement of impact in scientific communication, inlinks and outlinks, respectively, allow establishing relationships with other websites and web pages, enhancing their presence on the WWW and allowing a greater number of access points to a website. The size of the website, calculated through the sum of all the pages that make up a website is also an important indicator to identify the presence of a website on the WWW, but highly dependent on the relationships of inlinks and outlinks. Web Impact Factors, created by Ingwersen (1998), establishes a relationship between the visibility and size of a website. The results point to a considerable presence of archival institutions analyzed in the emerging virtual spaces of the WWW, but still need to increase their relationships with the universe of WWW sites, seeking greater reach and relevance in this informational space. From the results, it can be seen that the websites of Ibero-American archival institutions are, on average, more cited by other websites than they refer to other external websites. Although highly dynamic, web-based indicators can help archival institutions

leverage their websites, aiming to reach more and more users on the WWW.

KEYWORDS: Webometry. World Wide Web. Archival Institution. National Archive.

REFERÊNCIAS

AGUILLO, Isidro F. et al. Indicators for a webometric ranking of open access repositories. **Scientometrics**, v. 82, p. 477–486, fev. 2010.

_____. Rankings de universidades: el ranking web. **Higher Learning Research Communications**, v. 2, n. 1, mar. 2012.

_____ et al. Scientific research activity and communication measured with cybermetrics indicators. **Journal of the American Society for Information Science and Technology**, v. 57, n. 10, p. 1296–1302, ago. 2006.

ALMIND, T. C.; INGWERSEN, Peter. Informetric analyses on the world wide web: methodological approaches to “webmetrics”. **Journal of Documentation**, v. 53, n. 4, p. 404–426, 1997.

BJÖRNEBORN, Lennart. **Small-world link structures across an academic web space: A Library and Information Science Approach**. 2004. 399 f. Royal School of Library and Information Science, Copenhagen, 2004.

_____; INGWERSEN, Peter. Toward a Basic Framework for Webometrics. **Journal of the American Society for Information Science and Technology**, v. 55, n. 14, p. 1216–1227, 2004.

BORGMAN, Christine L.; FURNER, Jonathan. Scholarly communication and bibliometrics. **Annual Review of Information Science and Technology**, v. 36, p. 3–72, 2002.

CONSELHO NACIONAL DE ARQUIVOS (BRASIL). **Diretrizes gerais para a construção de *websites* de instituições arquivísticas**. Rio de Janeiro: CONARQ, 2000. Disponível em: http://www.conarq.arquivonacional.gov.br/images/publicacoes_textos/Diretrizes_Construcao_websites.pdf. Acesso em: 23 dez. 2017.

DUCHEIN, Michel. **Los obstáculos que se oponen al acceso, a la utilización y a la transferencia de la**

información em los archivos: un estudio del Ramp. Paris: UNESCO, 1983.

FONSECA, Maria Odila. **Arquivologia e ciência da informação.** Rio de Janeiro: FGV, 2005.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social.** 6. ed. São Paulo: Atlas, 2008.

HOOD, William W.; WILSON, Concepcion Shimizu. The literature of bibliometrics, scientometrics, and informetrics. **Scientometrics**, v. 52, n. 2, p. 291–314, out. 2001.

INGWERSEN, Peter. The calculation of Web impact factors. **Journal of Documentation**, v. 54, n. 2, mar. 1998.

JARDIM, José Maria; FONSECA, Maria Odila. Estudos de usuários em arquivos: em busca de um estado da arte. **DataGramZero**, v. 5, n. 5, p. 1–13, 2004.

KHOSROWPOUR, Mehdi. **Dictionary of information science and technology.** 2. ed. [S.l.]: Idea Group Inc., 2013. v. 2.

MACIAIS-CHAPULA, Cesar A. O papel da informetria e da cienciometria e sua perspectiva nacional e internacional. **Ciência da Informação**, v. 27, n. 2, p. 134–140, ago. 1998.

MARIZ, Anna Carla Almeida. **Arquivos públicos brasileiros:** a transferência da informação na internet. 2005. 187 f. Tese – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

MARQUES, Liliana Marisa Carvalho. **A visibilidade das universidades portuguesas na Web:** estudo webométrico. 2013. 223 f. Dissertação – Universidade Lusófona de Humanidades e Tecnologias, Lisboa, 2013.

MICROSOFT. **Dicionário prático de informática.** Lisboa: McGraw- Hill de Portugal, Ltda, 2000.

OLIVEIRA, Louise Anunciação Fonseca de; MATOS, Maria Teresa Navarro de Britto. *Websites* de instituições arquivísticas nacionais de tradição Ibérica e ferramentas web 2.0: uma reflexão sobre a cultura participativa. **Informação e Informação**, v. 18, n. 2, p. 168–192, ago. 2013.

ORGANIZAÇÃO DOS ESTADOS IBERO-AMERICANOS. **Regulamento da OEI.** Disponível em: <https://oei.org.br/sobre-a-oei/regulamento>. Acesso em: 5 nov. 2018.

PRITCHARD, Alan. Statistical Bibliography or Bibliometrics? **Journal of Documentation**, v. 25, n. 4, p. 348-, 1969.

SAWAYA, Márcia Regina. **Dicionário de informática e internet**. São Paulo: Nobel, 1999.

SHINTAKU, Milton; ROBREDO, Jaime; BAPTISTA, Dulce Maria. Webometria dos repositórios institucionais acadêmicos. **Ciência da Informação**, v. 40, n. 2, p. 312–326, ago. 2011.

SILVA, Armando Malheiro et al. **Arquivística: teoria e prática de uma ciência da informação**. Porto: Afrontamento, 2009.

SMITH, A. G. A tale of two web spaces: comparing *sites* using web impact factors. **Journal of Documentation**, v. 55, n. 5, p. 577–592, dez. 1999.

THELWALL, Michael. Bibliometrics to webometrics. **Journal of Information Science**, v. 34, n. 4, p. 1–18, 2007.

_____. **Introduction to webometrics: quantitative web research for the social sciences**. Chapel Hill: Morgan & Claypool Publishers, 2009.

_____; VAUGHAN, Liwen; BJÖRNEBORN, Lennart. Webometrics. **Annual Review of Information Sciences and Technology**, v. 39, p. 81–135, 2005.

VANTI, Nadia. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ciência da Informação**, v. 31, n. 2, p. 152–162, ago. 2002.

_____. **Links hipertextuais na comunicação científica: análise webométrica dos sítios acadêmicos latino-americanos em Ciências Sociais**. 2007. 292 f. Tese - Universidade Federal do Rio Grande do Sul, 2007.

_____. Mapeamento das instituições federais de ensino superior da região Nordeste do Brasil na Web. **Informação e Informação**, v. 15, n. 1, p. 55–67, jul. 2010.

_____. Os links e os estudos webométricos. **Ciência da Informação**, v. 34, n. 1, 2005.