

Os links e os estudos webométricos

Nadia Vanti

Doutoranda do Programa de Pós-graduação em Comunicação e Informação (PPGCOM) da UFRGS. Bibliotecária do Instituto Latino-Americano de Estudos Avançados (Ilea) da UFRGS.
E-mail: nvanti@hotmail.com

Resumo

Neste artigo são abordados os estudos quantitativos aplicados à Internet e à Web. É dada especial atenção ao conceito de webometria, à sua abrangência, aplicações e possíveis denominações. São assinaladas as suas diferenças com a cibermetria e estabelecidas as relações que estes dois métodos mantêm com os tradicionais campos da bibliometria, cientometria e informetria. É enfatizado também um recurso que vem despertando grande interesse entre os investigadores, não só por facilitar a navegabilidade entre sítios dentro da Web, mas também por constituir um dos indicadores mais relevantes dentro dos estudos webométricos: os weblinks. Na parte final, é feita breve revisão das pesquisas que têm sido realizadas sobre a estrutura, as funções e as motivações para a criação de links no espaço Web.

Palavras-chave

Webometria; Cibermetria; Métodos quantitativos; Links; Web.

The links and the webometric studies

Abstract

This article deals with quantitative studies applied to the Internet and the Web. It addresses the concept of webometrics, its scope, uses and possible designations. It points out the differences between this concept and cybermetrics and it establishes the relationship that these two methods maintain with the traditional fields of bibliometrics, scientometrics and informetrics. It also focuses on a resource that has been calling much attention among researchers, not only because it facilitates surfing on the Web sites, but also because it represents one of the most relevant indicators in webometric studies: weblinks. Toward the end, this article reviews the research that has been done about structure, functions and motivations for creating links on the Web space.

Keywords

Webometrics; Cybermetrics; Quantitative methods; Weblinks.

INTRODUÇÃO

Com o despontar das recentes tecnologias, surgem novas áreas de interesse dentro da biblioteconomia e ciência da informação que necessitam ser examinadas e explicadas. A webometria é uma destas áreas de estudo que vem adquirindo importância crescente para as análises quantitativas na Internet e, mais especificamente, na Web. Atualmente a World Wide Web constitui o maior repositório e a mais rica fonte de informação já conhecida pela humanidade. Assim sendo, não surpreende que pesquisadores que habitualmente se dedicavam a estudar os sistemas de informação tradicionais voltem-se agora para a investigação de como este novel ambiente pode ser utilizado, organizado e avaliado. Este é um dos motivos pelo qual se pretende analisar, neste trabalho, a temática referente às possíveis formas de medir a informação na Web e examinar como se dá a interconexão entre os sítios neste meio. Busca-se, portanto, trazer à luz o conceito de webometria e suas diversas definições, desenvolvidas por renomados autores que se dedicam a estudar este tema e a delinear sua abrangência, aplicações e possíveis denominações. Procura-se também assinalar suas diferenças com a não menos recente área da cibermetria e identificar as relações existentes entre estes dois métodos e os já conhecidos subcampos da bibliometria, cientometria e informetria, já abordados em artigo anterior (2002).

Toda a primeira parte do trabalho tem como objetivo servir de arcabouço para, na segunda parte, examinar um recurso que vem despertando grande interesse entre os investigadores no assunto, não só por facilitar a navegabilidade entre sítios dentro da Web, mas também por constituir um dos indicadores mais relevantes dentro dos estudos webométricos: os chamados *weblinks*.

WEBOMETRIA

Embora o subcampo da webometria seja considerado um dos mais recentes dentro dos estudos quantitativos na área de biblioteconomia e ciência da informação, já existem diversos trabalhos internacionais que abordam este tema. Muitos autores têm direcionado seu foco de estudos para este novo ambiente, por encontrarem na Web imensa rede diversificada de recursos de informação, facilmente acessível e ainda pouco explorada. Neste sentido, Cronin e McKim (1996) argumentam que, como a Web está se tornando um meio de comunicação cada vez mais importante para a ciência e a academia, é lógico que os estudos quantitativos se estendam também a este meio. Também Thelwall,

Vaughan e Björneborn (2003) consideram que, sendo a Web uma rede global de documentos, desenvolvida inicialmente para uso acadêmico e depois estendida para os usuários em geral, é óbvio que ela constitui um campo de pesquisa fértil para a bibliometria, a cientometria e a informetria. Além dos citados, muitos outros investigadores vêm desenvolvendo pesquisas sobre o conteúdo e a estrutura das *homepages* na Web.

Almind e Ingwersen (1997) foram dois dos primeiros autores que se dedicaram a estudar esta nova área, devendo-se a eles o termo *webometrics* que tem se consagrado dentro da ciência da informação para definir os estudos informétricos aplicados à World Wide Web. Abraham e Foresta (1996) também foram autores que se referiram a uma técnica original, cujo propósito consistia em elaborar mapas cognitivos e modelos matemáticos que ajudassem a entender melhor o funcionamento da WWW e suas dimensões social, cultural, política e organizacional. Mais recentemente, Björneborn (2002) definiu a webometria como “o estudo dos aspectos quantitativos da construção e uso dos recursos de informação, estruturas e tecnologias na Web, a partir das abordagens bibliométricas e informétricas” (tradução nossa).

Entre os objetivos e aspectos principais deste subcampo, Ingwersen e Christensen (1997) destacam a aplicação dos métodos bibliométricos tradicionais na medição dos fenômenos relacionados ao mundo da Web. Tais métodos possibilitam investigar os modelos de comunicação, a identificação de áreas de pesquisa, os estudos históricos sobre o desenvolvimento de uma disciplina ou domínio e a avaliação da pesquisa por países, instituições ou indivíduos. Abraham (1997), por sua parte, assinala que a webometria busca elucidar a relação entre os principais elementos da Web, seus nós (domínios, sítios e páginas), suas conexões (nexos ou *links* que se estabelecem entre estes nós) e a matriz de ligações resultantes que envolvem toda a extensão da rede.

OUTRAS DENOMINAÇÕES UTILIZADAS PARA FENÔMENOS SEMELHANTES

Além do termo *webometrics*, da autoria de Almind e Ingwersen, também se encontra na literatura a expressão *cybermetrics* ou cibermetria que, segundo Björneborn (2002), abrange “os estudos quantitativos de toda a Internet, incluindo *chats*, *mailing lists*, *new groups*, *MUDs* e a própria WWW” (tradução nossa). *Cybermetrics* também é o nome da revista acessível em formato eletrônico* lançada oficialmente durante a VI Conferência Internacional de

Cientometria e Informetria, em Jerusalém, no ano de 1997, pelo pesquisador Isidro Aguillo.

Estas duas novas áreas de estudos também têm sido apresentadas por outros autores com distintas denominações. De acordo com Cronin (2001), Elisabeth Davenport empregou o vocábulo *influmetrics* para se referir ao estudo quantitativo das linhas imperceptíveis e difusas da influência acadêmica no meio eletrônico, expressas em notas de agradecimento ou reconhecimento incluídas nas publicações, trabalhos e pesquisas científicas. Em outro artigo, Cronin e Weaver (1995), utilizando a mesma terminologia, exploram a relação entre autores, agradecimentos e citações no contexto da avaliação acadêmica, relação denominada ‘triângulo do reconhecimento’.

Shiri (1998), por sua parte, menciona estudo realizado em 1997 pela Royal School of Librarianship que pretendia analisar quantitativamente as páginas web dinamarquesas, sob a denominação *internetmetrics*. Já Bossy (1995) utilizou o termo *netometrics* para aludir à aplicação de técnicas cientométricas à Web. Em uma pesquisa de Almind e Ingwersen (1996) e em outra de Quoniam e Rostaing (1997), o termo adotado é *internetometrics* para se referir a estudos semelhantes. Outras expressões empregadas foram *webometry* por Abraham (1997) e *web bibliometry* por Chakrabarti *et alii* (2002). Entretanto, concorda-se com Thelwall (2003, p. 3, tradução nossa), quando afirma que “a razão de ser do termo *webometrics* seria o de denotar uma herança na bibliometria e na informetria e destacar uma perspectiva da ciência da informação para os estudos da Web”. Coincidindo com o último autor, optou-se neste trabalho pelo termo webometria, por ser aquele que melhor expressa a natureza dos estudos quantitativos aplicados à Web e ser, ademais, aquele que adquiriu maior difusão na crescente literatura internacional produzida sobre o tema.

Cabe ressaltar, todavia, que os mesmos objetos investigados pela webometria têm sido estudados também por outras áreas como a ciência da computação. Por este motivo, um leque de enfoques tem surgido a partir da metade dos anos 90, com nomes tais como “ecologia da Web”, “inteligência Web”, “exploração Web” e “análise gráfica da Web”.

WEBOMETRIA, CIBERMETRIA E SUAS RELAÇÕES COM OS TRADICIONAIS SUBCAMPOS QUE SE DEDICAM AOS ESTUDOS QUANTITATIVOS

Existem, portanto, dois novos campos de conhecimento relacionados às análises quantitativas dentro da biblioteconomia e da ciência da informação que surgiram com o desenvolvimento das tecnologias de informação

* Esta revista tem como proposta disponibilizar a todos os pesquisadores do mundo as análises e medições da comunicação no âmbito científico e, em especial, as medições do fluxo deste tipo de informação na WWW. URL: www.cindoc/csic/cybermetrics/cybermetrics

e comunicação. Cada um deles, porém, apresenta características distintas no que diz respeito à sua abrangência, recursos e objetos de estudo. De acordo com Björneborn (2002), a cibermetria teria um escopo mais amplo do que a webometria, pois compreende a aplicação das tradicionais técnicas informétricas a qualquer tipo de informação disponível na Internet. A webometria, segundo o mesmo autor, seria mais restrita, pois utiliza técnicas quantitativas para medir, especificamente, a informação disponível na Web, sendo assim uma parte do universo maior da cibermetria.

Conforme as definições levantadas, cabe determinar a posição ocupada por estes dois subcampos emergentes dentro da área de biblioteconomia e ciência da informação e mostrar sua inter-relação com os demais métodos quantitativos tradicionais: a bibliometria, a cientometria e a informetria².

Apesar de existirem posições divergentes, opiniões discordantes e conceitos diferentes a respeito destes métodos, considera-se que o mais adequado seria entender a informetria como um campo mais abrangente, que inclui tanto a bibliometria quanto a cientometria. Thelwall Vaughan e Björneborn (2003), Björneborn e Ingwersen (2003) e Björneborn (2004) reafirmam esta posição, ao seguirem as definições de informetria desenvolvidas por Brooks (1990), Tague-Sutckiffe (1992) e Russell (1994). Eles adotam especialmente o conceito de Tague-Sutckiffe (1992, p. 1, tradução nossa), para quem este subcampo refere-se ao “estudo dos aspectos quantitativos da informação em qualquer forma, não somente as registradas ou bibliográficas e de qualquer grupo social, não apenas dos cientistas”. Com o surgimento dos subcampos da webometria e da cibermetria, os primeiros autores citados consideram que a informetria inclui também estes dois novos subcampos, explicando que a informetria é mais ampla e que, pela definição de Tague-Sutckiffe, comportaria estes dois métodos quantitativos emergentes.

Em se tratando dos últimos métodos citados, Björneborn (2002) e Thelwall Vaughan e Björneborn (2003) apresentam a idéia de que a cibermetria compreende a webometria, dado que a primeira estuda os aspectos quantitativos de toda a Internet, incluindo a Web. Ao relacionar a cibermetria e a webometria com a bibliometria,

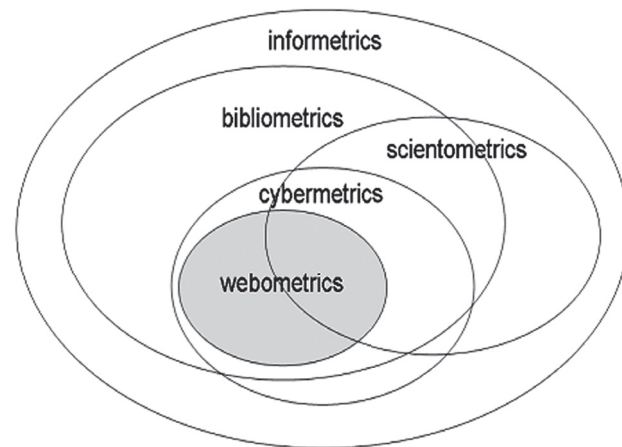
Thelwall Vaughan e Björneborn (2003, p. 3, tradução nossa) afirmam que “a cibermetria excede as fronteiras da bibliometria, pelo fato de algumas atividades no ciberespaço não serem normalmente registradas” e que a webometria estaria totalmente abraçada pela bibliometria, dado que “os documentos web – sejam eles textos ou multimídia – não deixam de ser informação registrada e armazenada, mesmo que em servidores web”.

Com relação à cientometria, os autores citados estão de acordo com o fato de que este método cobre parcialmente tanto a cibermetria quanto a webometria, entendendo que muitas atividades acadêmicas e científicas são atualmente apoiadas na Internet ou, mais especificamente, na Web.

A figura 1 reflete a visão de Björneborn (2002), retomada por Björneborn e Ingwersen (2003) e Thelwall Vaughan e Björneborn (2003) sobre a vinculação dos diferentes subcampos, com a qual se concorda preliminarmente.

FIGURA 1

Inter-relação entre os subcampos das métricas dentro da biblioteconomia e ciência da informação segundo Björneborn (2002), Björneborn e Ingwersen (2003) e Thelwall Vaughan e Björneborn (2003)



Entende-se, porém, que não fica suficientemente claro que a webometria esteja totalmente absorvida pela bibliometria, como mostra o diagrama (figura 1). Caberia questionar se é passível afirmar que a informação que se encontra em servidores web é informação registrada. O que ocorre com as páginas ou sítios que foram retirados da rede ou com as versões anteriores de páginas ou sítios que sofrem constantes processos de atualização? Há alguma forma de recuperá-los? Tais versões permanecem gravadas em um servidor remoto ou estão definitivamente extintas? O caráter dinâmico da Web nos obriga a repensar o

² Para uma revisão sobre o histórico da pesquisa brasileira nestes subcampos tradicionais, consultar o artigo de Miranda e Barreto (2000) “Pesquisa em Ciência da Informação no Brasil: síntese e perspectivas”, no qual eles destacam o relevante papel desempenhado pelo IBICT, especialmente no desenvolvimento e aplicação das técnicas bibliométricas no país, assim como a contribuição de Frederic W. Lancaster, Tefko Saracevic e Ingetrout Dahlberg e a importância da revista *Ciência da informação* na divulgação destes trabalhos.

conceito de informação registrada, não sendo possível trasladar automaticamente a terminologia tradicionalmente utilizada em formato impresso ao âmbito da Web. Como assevera Bustelo (1997, p. 47), “[...] es necesario reformular las bases teóricas de la gestión documental, de forma que adapten los conceptos a la realidad actual”.

Na biblioteconomia e na ciência da informação, existem pelo menos três definições para o termo registro. Uma delas, mais tradicional, diz respeito ao ciclo documental (conjunto de etapas pelas quais passa o documento dentro de uma unidade de informação para ser disponibilizado ao usuário), no qual registrar é uma das etapas constitutivas deste processo, em que cada documento que chega à biblioteca é numerado seqüencialmente e tem seus dados bibliográficos básicos descritos, de modo que possa ser inventariado ou tombado e identificado.

Em outro sentido, registrar uma obra significa indicar que a obra pertence ou é patrimônio de alguém. Existem locais predeterminados para registrar uma obra, de acordo com a sua natureza. Este tipo de registro regula o direito autoral e dá plenos poderes ao autor de usá-la como e onde quiser, enquanto limita o seu uso por outras pessoas que não tenham autorização para fazê-lo.

Ainda com outro significado, o ato de registrar está relacionado à gravação e fixação de informações em qualquer suporte, seja ele impresso, multimídia, mecanizado ou digital.

Em qualquer destes casos, é necessário revisar os habituais significados do termo registro, estender sua abrangência, adaptá-los a outros meios, ou criar novos conceitos que contemplem os recursos das novas tecnologias de informação e comunicação, cada vez mais presentes nos círculos acadêmicos e na sociedade de modo geral.

A Web é, ao mesmo tempo, fonte, suporte e sistema de informação descentralizado. Ela é constituída, basicamente, de sítios e links que corresponderiam, respectivamente, aos documentos e citações ou referências em um suporte impresso ou ainda aos itens ou registros e remissivas nos catálogos de bibliotecas tradicionais. Com o advento do mundo digital, por sua vez, novas possibilidades tecnológicas diretamente ligadas ao processo de produção, armazenagem, tratamento e recuperação de documentos e informação alteraram de forma substantiva não somente o modo como são realizadas tais tarefas, mas também os produtos finais deste processamento. Uma característica fundamental de tal mudança é a “desterritorialização do documento”, como destaca Alvarenga (2001), acrescentando que, “a partir da implementação e desenvolvimento da plataforma WWW, o documento passa a ter sua materialidade desvinculada da forma física, assumindo a forma digital, que possibilita

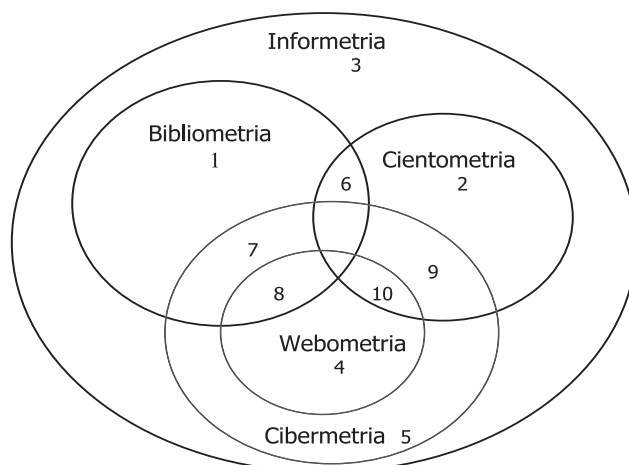
uma organização integrada de textos, imagens e sons”.

A mesma autora, por sua vez, afirma que o principal motivo que leva a Internet a ser muitas vezes desqualificada é, exatamente, por ela ser composta de “fontes transitórias” e de “proveniência pouco segura”. Sabe-se que, neste meio, muitos sítios não estão corretamente indexados, saem do ar ou são constantemente atualizados. Portanto, quando se conceitua registro como a gravação ou fixação do documento em um suporte, a fim de que ele possa ser acessado e consultado em qualquer momento sem alterações, os fatores anteriormente citados vêm reforçar a dúvida quanto ao fato de a informação em meio digital ser considerada efetivamente como ‘informação registrada’.

Dentro desta perspectiva, propõe-se um diagrama alternativo ao apresentado por Björneborn (2002), Björneborn e Ingwersen (2003) e Thelwall Vaughan e Björneborn (2003) que dê conta do assinalado acima, o que, conseqüentemente, repercutirá na inter-relação dos diferentes subcampos e na sua representação gráfica*.

Para se chegar à visão alternativa da inter-relação dos distintos subcampos exibidos na figura 1, tentou-se definir que recursos ou objetos de estudo e suportes estariam envolvidos em cada um deles e, especialmente, em cada espaço de intersecção entre tais subcampos, representados aqui pelos números que aparecem no gráfico (figura 2) e explicados a seguir.

FIGURA 2
Inter-relação entre os subcampos das métricas dentro da biblioteconomia e da ciência da informação, segundo Vanti



* Com isso retoma-se e atualiza-se o diagrama publicado pela autora (2002), no seu artigo “Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento”. Pode-se consultar este artigo, também, para obter mais informações a respeito da definição e histórico da bibliometria, cientometria e informetria.

1 Bibliometria: registros impressos, citações, agradecimentos, autores, usuários; livros, revistas, artigos de revistas.

2 Cientometria: áreas do conhecimento, cientistas, profissionais de um mesmo campo de atuação, colégios invisíveis, atividades científicas; dissertações, teses, documentos tecnológicos (patentes, normas técnicas etc.).

3 Informetria: todo o tipo de informação; fluxo, busca, recuperação, acesso à informação, sistemas de recuperação, comunicações informais entre quaisquer grupos sociais e de qualquer forma, inclusive oral; qualquer tipo de suporte.

4 Webometria: toda a Web (domínios, sítios, páginas web, URLs, motores de busca, *links*, agrupamentos de sítios - *clusters*, pequenos mundos).

5 Cibermetria: Internet, ciberespaço (*chats*, *mailing lists*, grupos de discussão, *muds* e a WWW).

6 Bibliometria versus cientometria: registros impressos, citações, agradecimentos dentro de uma área do conhecimento.

7 Bibliometria versus cibermetria: mensagens de *chats*, de *mailing lists* ou de grupos de discussão que permanecem disponíveis em um servidor web.

8 Bibliometria versus webometria: *E-books*, artigos eletrônicos de revistas disponíveis na Web.

9 Cientometria versus cibermetria: *Chats*, *mailing lists*, grupos de discussão, *muds* de uma região, área do conhecimento específica, ou entre cientistas pela Internet ou ciberespaço.

10 Cientometria versus webometria: Domínios, sítios, páginas web, URLs, agrupamentos (*clusters*) de sítios, pequenos mundos de uma região ou área do conhecimento específica.

Ao interpretar este diagrama alternativo, percebe-se que, diferentemente do que se observa no primeiro, a webometria aqui não está totalmente compreendida pela bibliometria. Considera-se que pelo menos uma parte dos recursos web não permanecem disponíveis de maneira indefinida, ou seja, não ficam registrados, assim como acontece na cibermetria – disciplina que estuda tanto recursos registrados, quanto recursos de informação não registrados dentro do âmbito da Internet. Sabe-se que a maioria dos *chats*, mensagens eletrônicas, mensagens

em listas de discussão etc. não estão disponíveis nem assumem caráter permanente, o que os torna recursos de informação não registrados.

ESCOPO E APLICAÇÕES DA WEBOMETRIA

A webometria tem direcionado seu foco de cobertura às metodologias e aos resultados de pesquisas bibliométricas, cientométricas e informétricas, com ênfase nos aspectos relacionados à Web.

De acordo com Thelwall Vaughan e Björneborn (2003), a webometria trata dos aspectos quantitativos tanto da construção quanto do uso da Web, compreendendo quatro áreas principais de pesquisa:

a) análise de conteúdo das páginas web;

b) análise da estrutura dos *weblinks*;

c) análise do uso da Web (exploração dos programas que registram os comportamentos de pesquisa e busca na Web);

d) análise de tecnologias na Web (incluindo o desempenho dos motores de busca.

O mais recente fascículo do *Journal of the American Society for Information Science and Technology* (JASIST, v. 55, n. 14, 2004), apresentado pelos pesquisadores Mike Thelwall e Liven Vaughan, foi totalmente dedicado à webometria. Na introdução, os autores fazem referência ao âmbito deste campo de estudos e esclarecerem quais os tópicos de interesse incluídos nesta edição. Destacam ainda que a maior parte dos *papers* escolhidos para fazer parte deste número especial enfoca os *hyperlinks* como uma fonte potencial de novas informações. Percebe-se, desta forma, a importância dos *links* para os estudos webométricos.

WEBLINKS

O *link** tem suas raízes na idéia de Paul Otlet da criação do Livro Universal e na proposta teórica do sistema Memex** de Vannevar Bush, sendo o conceito básico mais importante no hipertexto, já que é ele que permite estabelecer conexões entre diferentes tipos de informação,

* *Link*, *weblink* e *hyperlink* são utilizados neste trabalho como sinônimos.

** A partir da observação de que a mente opera por meio de associação de idéias, Bush imaginou um aparato que reunisse todos os documentos (livros, comunicações, arquivos etc.) de uma pessoa, de forma que fosse possível consultar este acervo com rapidez e agilidade por meio de algum tipo de mecanismo. Esta descrição do Memex guarda grande semelhança com os caminhos que o interagente percorre hoje quando está em um ambiente regido pela hipertextualidade.

produzindo diferentes resultados. Estes resultados podem ser outro tópico, uma referência, uma informação adicional, uma ilustração, um esquema, uma fotografia, uma seqüência de vídeo, um índice, outro programa ou ainda outra página web. Cada um destes destinos, considerados unidades de informação, são denominados também nós ou subnós na rede hipertextual. As relações ou uniões entre os nós descrevem o vínculo entre conteúdos e a ação que um exerce sobre o outro.

O *link* é também fonte de informação que apresenta as relações sociais que se estabelecem em torno do documento, refletindo o contexto social em que ele surge e as relações que ele e o seu autor mantêm com outros documentos, outros autores e seus leitores. Isto traduz a visão de Paul Otlet, quando defendia a idéia de que cada documento tem três dimensões, sendo a terceira justamente o contexto social (relação com o lugar, tempo, língua, outros leitores, escritores e assuntos (WRIGHT, 2003).

Nos estudos webométricos, além dos conceitos citados, o *link* é considerado indicador importante para determinar a relevância e o lugar que ocupa determinado sítio ou página no espaço web. Atualmente, os mais destacados motores de busca vêm utilizando informação dos *weblinks* para mostrar os resultados de uma pesquisa na Web. O número de *links* que levam a um sítio é considerado fator de hierarquização no momento de exibir na tela os sítios recuperados por motores como o Google – pioneiro em aplicar este sistema, em 1998 – e o Northern Light e o Alta Vista – mais tarde, a partir de 2002. O entendimento, neste caso, é que o número de *links* apontando para um sítio configuraria indício da sua importância ou qualidade (VAUGHAN; HYSEN, 2002). Seguindo esta lógica, ser *linkado* na Web significa estar visível e ter valor. Quanto mais vezes isto ocorre, mais visível se torna e mais valor adquire um sítio.

Os *links* entre sítios permitem traçar um mapa do relacionamento entre diferentes instituições e a intensidade de tais relações. Da mesma forma, contabilizar *links* entre instituições em diferentes países é um exercício útil para a identificação de padrões do fluxo da informação e de reconhecimento internacional. Pode-se calcular também nestas análises o tamanho médio de uma página em *bytes*, o número médio de *links* por página e a densidade média por *link*. Para tal fim, há certos elementos que devem ser considerados, como o URL (*Uniform Resource Locators*)*,

* Cada arquivo WWW disponibilizado (*site*) tem seu próprio endereço na Internet, conhecido como URL (*Uniform Resource Locator*) ou endereço *Internet*. O endereço *Internet* é composto por um indicador de protocolo (ex.: "http"), um nome de máquina *Internet* (ex.: "milonga.ilea.ufrgs.br") e um nome de arquivo (ex.: "index.html"). Além disso, através do endereço URL, é possível determinar que tipo de organização produziu o *site* (ex.: ".com, .org, .edu", etc.) (JACKES, 1997).

o título, as palavras-chave, o tipo de *homepage*, o domínio, o tamanho e o número de *links*.

Como explicam Almind e Ingwersen (1997), a densidade média dos *links* consiste na relação que pode ser estabelecida entre o tamanho de uma página e a quantidade de *links* que ela inclui. É uma medida que consegue reunir e normalizar em um só valor duas informações: tamanho da página e quantidade de *links*. Os autores sustentam que, quanto menor o tamanho da *homepage* ou o número de *bytes*, mantendo-se igual o número de *links*, menor será a taxa de densidade destes últimos. Tanto as páginas pessoais quanto as organizacionais apresentam freqüentemente grande quantidade de *links*, mesmo quando têm tamanho pequeno (baixa densidade por *link*). Já aquelas que estão direcionadas para um assunto específico costumam ser pequenas e ter poucos *links*, o que faz com que ostentem densidade por *link* mais alta, tornando este último tipo muito mais descritivo e auto-suficiente.

Os *links* vêm ocupando lugar cada vez mais destacado dentro das análises webométricas. Traçando uma analogia com as tradicionais citações bibliográficas que remetem a outros textos, os *links* têm sido considerados por alguns autores como citações, pois cumpriram função semelhante no meio eletrônico, ao estabelecerem relações entre páginas web. Estes são vistos como um indicador da importância global de um sítio* ou de um espaço web** para a comunidade externa (INGWERSEN, 1998; SMITH, 1999). Há número crescente de estudos que contabilizam *links* e sítios buscando identificar quais dentre os últimos são mais citados.

Existe um neologismo para definir a citação entre *homepages* "situation", que neste trabalho será traduzido como "situação". McKiernan (1996) foi quem utilizou pela primeira vez o termo no sentido de sítios citados, expressão também usada por Aguillo, em 1996, no IV Encontro da Associação Européia para o Estudo da Ciência e da Tecnologia, em Bielefeld (Alemanha). Rousseau (1997) é outro autor que recorreu a esta expressão em trabalho publicado na *Cybermetrics* intitulado "Situations: an exploratory study", no qual estudou os *links* entre *websites* e a distribuição de freqüências de "situações".

Nos estudos webométricos, os *links* funcionam como conectores entre os diferentes nós, entendendo por nó

* Conjunto de páginas web localizadas no mesmo endereço IP (*Internet Protocol*) e que se relacionam entre si por disponibilizarem um mesmo tipo de informação (OCASO; SPINAK, 1999).

** Espaço de armazenamento de dados acessado via *Internet*, usualmente utilizado para hospedar sítios da Web e arquivos de dados (WEBHOSTING, 1991).

qualquer unidade de informação como as páginas web, os diretórios, os sítios e os domínios. É possível classificar os *links* em diferentes tipos, de acordo com a direção que eles assumem e com a função que exercem no espaço Web. Assim sendo, Björneborn (2004) propõe uma tipologia de *links* que compreende as distintas formas de relacionamento entre estas unidades: *inlink*, *outlink*, *selflink*, *co-links*, *link* interno e *link* externo.

Os *inlinks*, também conhecidos como *incoming link*, *inbound link*, *inward link*, *back link* e *sitation*, são aqueles *links* recebidos por um nó dentro da Web, enquanto os *outlinks* são aqueles que apontam para outras páginas. Os *inlinks* e os *outlinks*, portanto, têm relação com o direcionamento que é dado a eles pelo nó ou unidade de informação. Já o *self-link* (*linkar* a si mesmo) ou o *co-link* (*linkar* dois nós ou ser *linkado* por dois nós ao mesmo tempo) têm relação com a função que cada *link* exerce dentro de um agrupamento de nós (*cluster*).

O critério que diferencia os *links* entre internos e externos possui outra conotação, dado que, neste caso, o que se privilegia na análise é o espaço geográfico onde o *link* atua. O *link* interno seria aquele que, estando em uma página, remete à outra página existente dentro do mesmo domínio ou unidade de análise. Já o *link* externo é aquele que remete a um sítio que está fora do domínio ou unidade de análise à qual pertence. Isto significa que um *inlink* ou um *outlink* pode ser tanto interno quanto externo.

Para tornar mais claras estas diferenciações, construiu-se uma matriz que mostra os cruzamentos entre estes tipos de *links*:

TABELA 1
Cruzamento entre as diferentes categorias de *links*

	Outlink*	Inlink**
Link Externo	Link visto da perspectiva do sítio que linka (sítio-fonte) e que aponta para um sítio localizado fora do seu domínio ou unidade de análise (sítio-alvo)	Link visto da perspectiva do sítio que é linkado (sítio-alvo) e que provém de um sítio localizado fora do seu domínio ou unidade de análise (sítio-fonte)
Link Interno	Link visto da perspectiva do sítio que linka (sítio-fonte) e que aponta para um sítio localizado dentro do seu domínio ou unidade de análise (sítio-alvo)	Link visto da perspectiva do sítio que é linkado (sítio-alvo) e que provém de um sítio localizado dentro do seu domínio ou unidade de análise (sítio-fonte)

* Corresponde a citar ou referência em bibliometria.

** Corresponde a ser citado ou citação em bibliometria.

Já têm sido pesquisadas no campo da bibliometria as razões que levam certos autores a citar outros autores. Egghe e Rousseau (1990), por exemplo, têm desenvolvido trabalhos sobre o tema. Também Gilbert (1977) e Brooks (1985) escreveram sobre o assunto, chegando à conclusão de que a persuasão é, na maioria das vezes, o que explicaria a necessidade de invocar a autoridade daqueles que produziram investigações e textos prévios sobre a matéria. Entretanto, as razões pelas quais as pessoas *linkam* outras páginas na sua ainda não têm sido suficientemente esclarecidas, e as diferenças entre citações nos artigos científicos e “situações” na Web começaram a ser investigadas só mais recentemente. Boudourides, Sigrist e Alevizos (1999), por exemplo, são alguns dos poucos autores que têm trabalhado sobre o tema e estabeleceram uma diferenciação entre citar e “sitar”, ao afirmarem que “enquanto citar é persuadir, *linkar* é lutar por atenção. O recurso que é ao mesmo tempo escasso e desejável, no espaço virtual é exatamente a atenção” (tradução nossa).

INVESTIGAÇÕES SOBRE A ESTRUTURA, FUNÇÕES, NATUREZA E MOTIVAÇÕES PARA A CRIAÇÃO DE LINKS

No do campo da webometria, existem vários pesquisadores que se dedicam ao estudo específico dos *links*. Com a proposta de reunir e conhecer que aspectos relacionados a *links* tais autores abordam em suas pesquisas, realizou-se uma revisão bibliográfica atualizada, porém não exaustiva, destes trabalhos, encontrando-se assim várias investigações muito interessantes. Borgman e Furner (2002), por exemplo, mostram em seu trabalho os trajetos dos *links* web como estruturas equivalentes às citações das revistas. Além disto, sugerem a utilização da expressão análise de *links* em lugar dos termos *links* e *linking*, “para dar conta de todas as técnicas quantitativas em que as conexões são classificadas e contadas com vistas à descrição, explanação, prognóstico e avaliação do fenômeno relação-documento” (BORGMAN e FURNER, 2002, *on-line*, tradução nossa). Propõem, ademais, uma distinção entre análise de *links* relacionais e análise de *links* avaliativos. Na primeira, a contabilização de *links* seria usada como indicador do nível de conectividade, da força do relacionamento ou da direção do fluxo entre documentos, pessoas, revistas, grupos, organizações, domínios ou nações. Já na segunda, na análise de *links* avaliativa, a contagem de *links* seria utilizada como indicador ou medida do nível de qualidade, importância, influência ou desempenho individual de documentos, pessoas, revistas, grupos, domínios ou nações. Vaughan e Hysen (2002), por sua parte, desenvolveram um estudo no qual encontraram uma correlação entre o número de *links* externos e o fator de impacto de revistas em biblioteconomia e ciência da informação.

Thelwall é um dos autores que mais têm se dedicado ao estudo de *links*. Em 2001, ele publicou artigo no *Journal of information science* explorando o diagrama de rede como ferramenta importante para visualizar a força da interconexão entre áreas da Web, mostrando que é possível a contabilização de *links* de quatro diferentes formas com base na sua intensidade, cada uma salientando um aspecto diferente dos dados (THELWALL, 2001). Somente no ano de 2002, foram encontrados cinco artigos dele sobre o assunto. Na revista *Scientometrics*, em co-autoria com Smith, Thelwall descreve a investigação que realizaram sobre os padrões de *links* internacionais entre sítios Web universitários da Ásia e Pacífico, verificando que a Austrália e o Japão eram o centro da região na Web, sendo a Austrália particularmente alvo comum de recebimento de *links*, enquanto o Japão apresentou desempenho mais balanceado de *links* recebidos de *links* que apontavam para outros domínios (THELWALL; SMITH, 2002). No volume 28 do *Journal of information science*, o mesmo autor (THELWALL, 2002a) apresentou outro artigo questionando se o número de *links* que uma página de Web recebe pode ser considerado um indicador de qualidade acadêmica, constatando que os recursos que apresentam maior número de *links* são aqueles que facilitam o acesso a amplo leque de informação – conhecidos como *hubs* –, mais do que aqueles que fornecem conteúdo específico – chamados de autoridade.

No mesmo ano, Thelwall publicou dois *papers* sobre *interlinks* no *Journal of documentation*. O primeiro deles apresenta o desenvolvimento de uma metodologia para analisar os padrões de *interlinkagem* entre sítios de universidades do Reino Unido e o uso destes padrões para indicar que o grau de *interlinkagem* decresce de acordo com a distância geográfica das universidades (THELWALL, 2002b). O segundo *paper* aborda com grande nível de detalhamento a contabilização de *links* entre pares de universidades, mostrando evidências de um relacionamento linear importante entre tamanho e qualidade da pesquisa tanto das instituições-fonte quanto das instituições-alvo. O autor propôs que este modelo fosse aplicado a outros sistemas universitários nacionais, com a expectativa de vê-lo confirmado por estudos realizados em outros países, abrindo assim as portas para o aprofundamento da exploração de dados sobre os *links* web acadêmicos (THELWALL, 2002c). Ainda neste ano, Vaughan e Thelwall (2002) disponibilizaram na Web o *preprint* de um artigo, relatando estudo que realizaram com o intuito de investigar os fatores que influenciam na criação de *links* em sítios web de revistas e descobrir se o tempo de existência e o conteúdo de um sítio Web induzem a criação de *links* para um sítio web de revista, medidos tanto pela relação da contabilização de *links* quanto pelo Fator de Impacto de Revistas. O resultado

encontrado foi que tanto o tempo de existência quanto o conteúdo dos sítios web são fatores significativos, pelo menos para as disciplinas estudadas (biblioteconomia, ciência da informação e direito).

Em 2003, Thelwall escreveu outro artigo referindo-se a *links*, no qual se propõe a dar início a um processo de diferenciação entre as motivações de criação de *links* em sítios web acadêmicos e de citações em revistas, partindo do pressuposto de que eles são fenômenos muito diferentes. Neste estudo, amostra aleatória de 100 *links* entre sítios e páginas de universidades do Reino Unido como ponto de partida para uma exploração qualitativa, definindo-se quatro tipos de motivações: “reconhecimento”, “social”, “navegacional geral” e aqueles sem motivação comunicacional específica. Para o autor, cada um destes tipos desempenha função única na Web (THELWALL, 2003).

Kim (2004), por sua vez, apresentou um *paper* que descreve a estrutura dos *hiperlinks* acadêmicos incluídos em sítios web universitários da Coreia e examina a associação entre a estrutura das redes de *hiperlinks* e os padrões de colaboração entre tais universidades, baseados na frequência de artigos em co-autoria. O autor acredita que estudar estas duas redes pode ser uma contribuição para o entendimento da mudança dinâmica no processo atual de comunicação criado e sustentado pelos meios de comunicação tradicionais e emergentes. A sua análise confirma que os padrões de redes de *hiperlinks* estão correlacionados com as redes de autoria e mostra que, estendendo tais redes, a identidade da investigação pode ser um indicador de produtividade dos pesquisadores e dos institutos de pesquisa.

Já Smith (2004), no mesmo ano, publicou, na *Information research* um estudo que investiga em que medida os *links* web são análogos às citações na literatura tradicional impressa. Ele desenvolveu uma classificação para os *weblinks*, a partir da natureza das páginas-fonte e das páginas-alvo e das razões destes *links*. Descobriu que, no total, 20% dos *links* estudados podem ser considerados como *links* de pesquisa, assim como ocorre com as citações.

Rousseau e Thelwall (2004), também neste ano, mostraram em seu artigo que os chamados *escher staircases*¹¹, isto é, os ciclos de quatro nós em um gráfico com ligações recíprocas formam um elemento estrutural básico na World Wide Web. Os autores descobriram que a existência

¹¹ Referência à “Ascending and Descending”, litografia criada em 1960 pelo artista gráfico holandês Maurits Cornelis Escher (1898-1972). Sua arte é apreciada no mundo inteiro, e a fama que adquiriu se deve às suas chamadas estruturas impossíveis. Seus trabalhos podem ser vistos em muitos sítios na Internet <www.mcescher.com>.

destes ciclos é parcialmente determinada pelo tamanho do sítio, sendo menos provável de encontrá-los em um sítio pequeno do que em um sítio grande. Apontam, ainda, que este fenômeno também é parcialmente determinado pela estrutura do sítio, ou seja, um sítio com muitos *links* por página tenderá naturalmente para o crescimento dos *escher staircases*.

Por fim, cabe citar a tese de Björneborn (2004), na qual o autor buscou, tanto conceitual quanto empiricamente, identificar e caracterizar o fenômeno do “pequeno mundo”*, mediante a análise da estruturação dos *links* no espaço web acadêmico. Para isto, escolheu um conjunto de *links* que cobria 109 universidades do Reino Unido, descobrindo que os subsítios web acadêmicos deste país mostravam propriedades características de um “pequeno mundo”, com alto coeficiente de agrupamentos e baixa extensão de *links* entre os subsítios. Também percebeu que os criadores de *links* pessoais, tais como pesquisadores e estudantes, podem ser conectadores importantes de sítios e de assuntos na investigação ou pesquisa em tais espaços acadêmicos web. Ao realizar exame mais acurado das páginas e *links* de dez redes de trilhas, constatou a existência de um gênero de conectividade muito rico na Web. No final, o autor ofereceu uma prospecção com relação ao que pode vir a acontecer no futuro, com base nos achados da sua investigação empírica. E ainda discutiu as possíveis implicações da organização do conhecimento dos pequenos mundos não planejados pela biblioteconomia e ciência da informação, argumentando que é necessário redefinir os objetivos e a rede conceitual da pesquisa nesta área para que ela possa acompanhar tanto o comportamento da informação convergente – aquele que aponta para determinadas metas –, quanto o comportamento de informação divergente – *serendipitous*** – conduzida por usuários tanto de bibliotecas tradicionais e bases de dados bibliográficas, quanto a partir de sistemas como a Web.

Como se pôde observar, existe uma diversidade de estudos sobre os *links*, abordando vários aspectos: estrutura, funções, natureza, motivações para a sua criação etc. A quantificação, nestes estudos, opera como método de pesquisa e como

instrumento para avaliar os relacionamentos entre os sítios web, com a finalidade de desvendar, em muitos casos, como este fenômeno vem se desenvolvendo dentro da mais recente e mais complexa fonte de informação de que se tem conhecimento: a Web.

CONCLUSÃO

Diante do exposto, pode-se concluir que a webometria representa área fundamental dentro dos estudos quantitativos e dentro da biblioteconomia e ciência da informação. Como foi discutido nas primeiras páginas deste trabalho, a webometria tem suas raízes na bibliometria, informetria e cientometria, mantendo com elas importantes espaços de convergência, porém avançando para novos ambientes no mundo da informação. Pelos seus objetivos, características e escopo, esta nova métrica constitui método crucial para quem queira mergulhar na análise da informação contida na Internet. A revisão de literatura mostrou as ricas possibilidades e as potencialidades que esta área emergente oferece para futuros desenvolvimentos no campo das análises quantitativas no âmbito da Web.

No que se refere especificamente aos *links*, foi possível observar a utilidade que eles revestem não só para os internautas como uma ferramenta que facilita a movimentação dentro da Web, mas também a informação que eles proporcionam ao pesquisador que busca estabelecer as relações e inter-relações que estruturam determinada área do conhecimento. A classificação dos diferentes tipos de *links*, conforme o espaço virtual em que eles se encontram e o direcionamento que assumem do ponto de vista do sítio que *linka* (sítio-fonte) ou do sítio que é *linkado* (sítio-alvo), fornece amplo leque de dados para pesquisas deste gênero. Os *links* entre sítios permitem também traçar mapas que mostrem os relacionamentos existentes entre diferentes pesquisadores, comunidades acadêmicas, instituições, áreas do conhecimento, países e regiões do planeta, assim como a intensidade de tais relações. Ao mesmo tempo, eles possibilitam medir a importância relativa de cada sítio dentro do mundo virtual e o grau de visibilidade que estes alcançam.

Cabe destacar, pelos motivos já citados, a importância de se continuar pesquisando sobre os *links*. Pelo fato de serem recurso muito recente, ainda há espaço para novas descobertas que possam apontar para tendências e padrões do fluxo da informação e do relacionamento entre documentos eletrônicos disponibilizados na Web.

Também não restam dúvidas de que a área de biblioteconomia e ciência da informação tem de ser revisada no que diz respeito a certos conceitos tradicionalmente utilizados, como, por exemplo, a definição de “informação

* O fenômeno do pequeno mundo refere-se à teoria desenvolvida por Watts e Strogatz (1998). Nela se afirma que qualquer pessoa pode ser encontrada através de uma pequena rede social de conhecidos. O conceito tem sua origem no experimento de Stanley Milgran realizado em 1960 que demonstra que cada pessoa estaria separada da outra no mundo por uma cadeia de, no máximo, seis pessoas, conhecida como “seis graus de separação”. Desta forma, a própria Web pode ser considerada um pequeno mundo, já que os sítios estão conectados uns aos outros por meios de *links* e, através do acesso a um sítio, pode-se chegar a conhecer muitos outros que estão na rede.

** Refere-se à pesquisa na Web, onde o internauta vai em busca de uma determinada informação e, casualmente, acaba encontrando outra.

registrada”. A informação “registrada” do suporte impresso perde espaço para a informação “disponível” nos servidores web. Cabe ampliar e atualizar os conceitos teóricos desta área do conhecimento que ainda estão imbuídos por um paradigma fortemente centrado no papel. Com o advento da Internet e da Web, surgiram novas demandas e novas formas de organizar, processar, disponibilizar e acessar a informação. Porém, em termos teóricos, não houve evolução tão rápida quanto na prática profissional. É indispensável, portanto, que tais mudanças entrem na pauta das discussões teóricas, reformulando e criando novos conceitos dentro da disciplina, sob pena de esta se tornar um campo obsoleto diante das rápidas mudanças que ocorrem na sociedade da informação.

Neste sentido, é importante incorporar os diversos setores sociais que produzem e consomem informação, para o que é necessário flexibilizar os critérios classificatórios da disciplina, abrindo suas fronteiras para os novos fenômenos informacionais. Vale insistir na urgência de atualizar o arcabouço conceitual da biblioteconomia e ciência da informação, para assim poder acompanhar a dinâmica dos avanços tecnológicos e as concomitantes mudanças que vêm se produzindo dentro da sociedade e nas relações que seus diferentes setores estabelecem com o mundo da informação.

Artigo recebido em 15/12/2004 e aceito
para publicação em 15/08/2005.

REFERÊNCIAS

AGUILLO, Isidro. A preliminary approach to citation phenomena in the world wide web. In: THE EUROPEAN ASSOCIATION FOR THE STUDY OF SCIENCE AND TECHNOLOGY (EASST) CONFERENCE, 4., 1996, Bielefeld. *Anais...* Alemanha: Bielefeld, 1996. Signatures Of Knowledge Societies.

ALVARENGA, Lídia. A teoria do conceito revisitada em conexão com ontologias e metadados no contexto das bibliotecas tradicionais e digitais. *DataGramaZero*, v. 2, n. 6, dez. 2001. Disponível em: <http://www.dgz.org.br/dez01/F_I_art.htm>. Acesso em: 5 set. 2004.

ABRAHAM, Ralph. *Webometry: measuring the complexity of the world wide web*. Basead on a talk in Vienna at FIS96, 6.15.96. Appeared in *World Futures*, 1997. Disponível em: <<http://deleafproductions.com/utopicalibrary/text/abraham-web.html>>. Acesso em: 10 maio 2004.

_____; FORESTA, D. *Webometry: chronotopography of the world wide web*. 1996. Disponível em: <<http://www.ralph-abraham.org/articles/MS%2389.Web3/webometry.html>>. Acesso em: 24 jun. 2004.

ALMIND, Tomas C.; INGWERSEN, Peter. Informetric analysis on the world wide web: a methodological approach to “internetometrics”. *CIS Report 2*, 1996.

_____; INGWERSEN, Peter. Informetric analyses on the world wide web: methodological approaches to “webometrics”. *Journal of Documentation*, v. 53, n. 4, p. 404-426, 1997.

BJÖRNEBORN, Lennart. *Small-world link structures on the web*. Copenhagen, DK: School of Library and Information Science, 2002. Disponível em: <www.db.dk/lb/2002smallworld.pps>. Acesso em: 28 jul. 2003.

_____. *Small-world strutures across an academie web space: a library and information science approach*. 2004. Dissertação (PhD em Ciência da Informação)- Department of Informations Studies, Royal School of Library and Information Science, Copenhagen, Denmark, 2004. 399 p.

_____; INGWERSEN; Peter. Towards a basic framework of webometrics. *Journal of American Society for Information Science and Technology*, 2003. Preprint.

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

BOSSY, M. G. *The last of the litter: “netometrics”*. 1995. Disponível em: <<http://biblio-fr.info.unicaen.fr/bnum/jelec/Solaris/d02/2bossy.html>>. Acesso em: 10 set. 2004.

BOUDOURIDES, Moses A.; SIGRIST, Beatrice; ALEVIZOS, Philippos D. *Webometrics and the Self-Organization of the European Information Society*. 1999. Disponível em: <<http://hyperion.math.upatras.gr/Webometrics/>>. Acesso em: 24 jun. 2004.

BROOKES, B.C. Biblio-, sciento-, infor-metrics??? what are we talking about? In: EGGHE, L.; ROUSSEAU, R. (Ed.). *Informetrics 89/90*. Amsterdam: Elsevier, 1990. p. 31-43.

BROOKS, T. A. Private acts and public objects: an investigation of citer motivations. *Journal of the American Society for Information Science*, v. 36, n. 4, p. 223-229, July 1985.

CAMPELLO, Bernadete Santos; MAGALHÃES, Maria Helena de Andrade. *Introdução ao controle bibliográfico*. Brasília: Briquet de Lemos, 1997. 108 p.

CHAKRABARTI, S.; JOSHI, M. M.; PUNERA, K; PENNOCK, D. M. The structure of broad topics on the web. In: WWW2002 CONFERENCE, 11., 2002, Honolulu, Hawaii. *Electronic proceedings...* Disponível em: <<http://www2002.org/CDROOM/refereed/338/>>. Acesso em: 5 set. 2004.

CRONIN, Blaise. Bibliometrics and beyond?: some thoughts on web based citation analysis. *Journal of Information Science*, v. 27, n. 1, p. 1-7, 2001.

_____; MCKIM, Geoffrey. Science and scholarship on the world wide web: a north american perspective. *Journal of Documentation*, v. 52, n. 2, p. 163-171, 1996.

_____; WEAVER S. The praxis of acknowledgement: from bibliometrics to influmetrics. *Revista Española de Documentacion Cientifica*, v. 18, n. 2, p. 172-177, 1995.

EGGHE, Leo; ROUSSEAU, Ronald. *Introduction to informetrics: quantitative methods in library, documentation and information science*. Amsterdam: Elsevier, 1990.

- GILBERT, G.N. Referencing as persuasion. *Social Studies of Science*, n. 7, p. 113-122, 1977.
- KIM, Hyo. Comparing academic hyperlink structure with co-authorship patterns in Korea. In: AOIR-ARIST 2004 WORKSHOP ON WEB SCIENCE RESEARCH METHODS, 2004, Brighton. *Proceedings...* Brighton: [s.n.], 2004.
- INGWERSEN, Peter. The calculation of web impact factors. *Journal of Documentation*, v. 54, n. 2, p. 236-243, 1998.
- _____ ; CHRISTENSEN, F.H. Data set isolation for bibliometric online analyses or research publications: fundamental methodological issues. *Journal of the American Society for Information Science*, v. 48, n. 3, p. 205-217, 1997.
- _____ ; INGWERSEN, P.; CHRISTENSEN, F. H. Online citation analysis: a methodological approach. *Scientometrics*, v. 37, n. 1, p. 39-62, 1996.
- MARTIN VEGA, Arturo. *Fuentes de información general*. Gijón: TREA, 1995. 320 p.
- MCKIERNAN, G. CitedSites(sm): citation indexing of the web resources. 1996. Disponível em: <<http://www.public.iastate.edu/~CYBERSTACKS/Cited.html>>. Acesso em: 5 set. 2004.
- MIRANDA, Antônio; BARRETO, Aldo de Albuquerque. Pesquisa em ciência da informação no Brasil: síntese e perspectivas. *DataGramaZero*, v. 1, n. 6, dez. 2000. Disponível em: <http://www.dgz.org.br/dez00/Art_04.htm>. Acesso em: 27 ago. 2005.
- QUONIAM, L.; ROSTAING, H. From scientometrics, informetrics to internetometrics, cybermetrics or is it possible to neglect Internet nowadays?. In: CYBERMETRICS'97, 1997, Jerusalém, Israel. *Electronic proceedings...* Disponível em: <<http://www.cindoc.csic.es/cybermetrics/cybermetrics.html>>. Acesso em: 2 out. 2001.
- ROUSSEAU, Ronald. Sitations: an exploratory study. *Cybermetrics*, v. 1, n. 1, 1997. Disponível em: <http://www.cindoc.csic.es/Cybermetrics_Volume_1_Issue_1_Paper1_Sitations_An_exploratory_study_by_Rousseau.htm>. Acesso em: 25 set. 2001.
- _____ ; THELWALL, Mike. Escher staircase on the word wide web. *First Monday*, v. 9, n. 6, June 2004. Disponível em: <http://www.firstmonday.org/issues/issue9_6/rousseau/index.html>. Acesso em: 22 jul. 2004.
- RUSSELL, J. M. Back to the future for informetrics?. *Scientometrics*, v. 30, n. 1, p. 407-410, 1994.
- SMITH, Alastair. A tale of two web spaces: comparing sites using web impact factors. *Journal of Documentation*, v. 55, n. 5, p. 577-592, Dec. 1999.
- _____ . Web links as analogues of citations. *Information Research*, v. 9, n. 4, July 2004.
- SHIRI, A. A. Cybermetrics: a new horizon in information research. In: FID CONFERENCE AND CONGRESS, 49., 1998, New Delhi, India. *Anais...* New Delhi: [s.n.], 1998.
- TAGUE-SUTCKIFFE, J. An introduction to informetrics. *Information Processing & Management*, v. 28, n. 1, p. 1-3, 1992.
- THELWALL, Mike. Exploring the link structure of the web with network diagrams. *Journal of Information Science*, v. 27, n. 6, p. 393-402, 2001.
- _____ . The top 100 linked-to pages on UK university web sites: high inlink counts are not usually associated with quality scholarly content. *Journal of Information Science*, v. 28, n. 6, 2002a.
- _____ . Evidence for the existence of geographic trends in university web site interlinking. *Journal of Documentation*, v. 58, n. 5, 2002b.
- _____ . A research and institucional size-based model for national university Web site interlinking. *Journal of Documentation*, v. 58, n. 6, 2002c.
- _____ . What is this link doing here?: beginning a fine-grained process of identifying reasons for academic hyperlink creation. *Information Research*, v. 8, n. 3, Apr. 2003.
- _____ ; SMITH, Alastair G. 2002. Interlinking between Asia-Pacific University Web sites. *Scientometrics*, v. 55, n. 3, p. 362-376, 2002.
- _____ ; VAUGHAN, Liwen; BJÖRNEBORN, Lennart. Webometrics. In: *Annual Review of Information Science And Technology*, v. 39, 2003. Preprint.
- VANNEVARBUSH. Aswemaythink. *The Atlantic Monthly*, July 1945. Disponível em: <<http://www.csi.uottawa.ca/~dduchier/misc/vbush/awmt.html>>. Acesso em: 12 ago. 2004.
- 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*, Brasília, v. 31, n. 2, p. 152-162, maio/ago. 2002.
- VAUGHAN, Liwen; HYSEN, Kathy. Relationship between links to journal Web sites and impact factors. *Aslib Proceedings*, v. 54, n. 6, 2002.
- _____ ; THELWALL, Mike. Scholarly use of the web: what are the key inducers of links to journal web sites?. *Journal of the American Society for Information Science and Technology*, v. 57, n. 1, p. 29-38, 2002.
- WATTS, D. J.; STROGATZ, S. H. Collective dynamics of small-world networks. *Nature*, v. 393, p. 440-442, June 1998.
- WRIGHT, Alex. *Forgotten forefather*: Paul Otlet. 2003. Disponível em: <http://www.bboxesandarrows.com/archives/forgotten_forefather_paul_otlet.php>. Acesso em: 3 jul. 2005.