

Webmetria:

origens e usos contemporâneos

Fábio Castro Gouveia

Ronaldo Ferreira de Araújo

Como citar: GOUVEIA, F. C.; ARAÚJO, R. F. Webmetria: *origens e usos contemporâneos*. In: GRÁCIO, M. C. C.; MARTÍNEZ-ÁVILA, D.; OLIVEIRA, E. F. T.; ROSAS, F. S. (org.). **Tópicos da bibliometria para bibliotecas universitárias**. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2020. p. 208-228. DOI: <https://doi.org/10.36311/2020.978-65-86546-91-0.p208-228>



All the contents of this work, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0).

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença Creative Commons Atribuição- NãoComercial-SemDerivações 4.0 (CC BY-NC-ND 4.0).

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia Creative Commons Reconocimiento-No comercial-Sin derivados 4.0 (CC BY-NC-ND 4.0).

Capítulo 9

WEBOMETRIA: ORIGENS E USOS CONTEMPORÂNEOS

Fábio Castro Gouveia
Ronaldo Ferreira de Araújo

1 A WEB É UM ORGANISMO VIVO

Desde seu surgimento, tem sido comum encontrarmos na literatura sobre a internet o uso de metáforas para auxiliar na compreensão tanto de seus aspectos conceituais quanto de suas aplicações práticas. Entre as metáforas mais comuns encontramos comparações e analogias da internet, por exemplo, como uma estrada, uma comunidade ou mesmo uma biblioteca.

Para Wu e Chen (2013) no caso da estrada, a metáfora destacaria o processo, referindo-se ao tráfego e as instalações da internet. Pode-se pensar, por exemplo, no Protocolo Internet (IP) que permite que o tráfego seja roteado de uma rede para outra, ou mesmo no fato do gerente do sistema poder ser o único “*gateway*” que todos os outros PCs na rede podem usar. A segunda, da ideia de comunidade é outra metáfora dominante sobre a Internet em uma sociedade da informação, cuja analogia enquadra a compreensão da internet como um meio de comunicação no qual novos tipos de relações sociais e vida cívica podem agora ser possíveis no chamado “ciberespaço” (WU; CHEN, 2013).

A terceira, que nos interessa mais de perto, coloca no centro da discussão o seu conteúdo, e considera a internet como uma biblioteca,

na qual podemos percorrer navegando on-line e que possui informação abundante em diferentes formatos (WU; CHEN, 2013). Nessa analogia podemos considerar a internet como uma biblioteca em ebulição, na qual o conhecimento registrado é “coletivo, dissociado de regras convencionais, sem semântica privada, com estímulo à criação de arquivos de documentos digitais independentes e de iniciativas que rompem com as tradições” (FUJINO; JACOMINI, 2007, p.79).

Segundo Wu e Chen (2013) quando queremos pesquisar um site na Internet, podemos encontrá-lo em mecanismos de buscas bastando clicar e acessar com um navegador, e ao navegar em um site ou livro eletrônico, podemos pular, digitalizar ou ler em detalhes, como também virar ou retroceder as páginas.

Uma vez que as analogias nos aproximam de questões que envolvem aspectos estruturais, funcionais e de recursos e serviços, torna-se importante começarmos a reforçar a diferença entre a internet e a web que algumas vezes são erroneamente consideradas como sinônimas. A Internet é a rede mundial de computadores, ou seja, a estrutura que oferece uma série de serviços (dentre os quais, transferência de dados e arquivos entre servidores e os mais variados dispositivos).

A web (*world wide web*), por sua vez, se restringe ao conteúdo que é oferecido e trafegado fazendo uso de um protocolo parte do TCP/IP (Protocolo de Controle de Transmissão / Protocolo Internet ou *Transmission Control Protocol / Internet Protocol*) chamado HTTP (Protocolo de Transferência de Hipertexto ou *Hypertext Transfer Protocol*) bem como sua versão segura HTTPS (Protocolo de Transferência de Hipertexto Seguro ou *Hypertext Transfer Protocol Secure*).

Apesar de muito do nosso contato com as diversas fontes de dados providas por intermédio da internet se dar utilizando um assim chamado “Navegador Web”, em alguns casos (como, por exemplo do e-mail) ele apenas faz a intermediação pela web de um processo que vai ser consolidado por outro protocolo da internet (no caso do e-mail o Protocolo de Transferência de Correio Simples ou *Simple Mail Transfer*

Protocol– SMTP). Esta distinção se fará relevante na delimitação dos campos de metrias da informação que tem na internet e na web seus campos de estudo.

Quando pensamos numa biblioteca e no seu conteúdo, com diversos volumes devidamente catalogados e organizados, podemos nos lembrar das cinco leis da biblioteconomia de Ranganathan (1931). Nestas leis, nas quais podemos identificar elementos que contemporaneamente poderiam dialogar com tranquilidade com às preconizadas por Donald A. Norman quanto a importância da UX (*User Experience*– Experiência do usuário), temos como conclusão de que bibliotecas são um organismo em crescimento. Noruzi (2004) faz a transposição das leis para o ambiente web enunciando a quinta lei como “A web é um organismo em crescimento”¹ (NORUZI, 2004, s.p.). É do nosso entendimento que mais do que em crescimento, a web é um organismo vivo e dinâmico, onde conteúdos são adicionados e outros removidos, com uma taxa de crescimento exponencial não só em termos de arquivos, quanto em volume de dados gerados ao longo do tempo. Fazer com que este universo em expansão possa ser catalogado é hoje tarefa de softwares específicos, onde apenas parte do conteúdo pode ser visível aos que nele buscam informações.

Os estudos estatísticos de uso de itens de uma unidade de informação, e o acesso ao conteúdo da internet via web, guardam similaridades pelo fato de os últimos nascerem alicerçados numa abordagem como a utilizada nas ações de gestão de acervos de uma biblioteca. E se a bibliometria passa do olhar de gestão por fora das obras e mergulha para dentro delas, entendendo o seu conteúdo como importante para estabelecer relações métricas a serem estudadas, não surpreende que se tenha vislumbrado para este novo espaço com informações e acesso registrado a possibilidade de aplicações similares. Na criação do algoritmo do *Page-Rank*, raiz do sistema de classificação do Google, por Brin e Page (1998), ideia similar a descrita por Love e Sloman (1995) e que poderia ser entendida como já proposta

¹ No original “*The Web is a growing organism*”

para análise de artigos em Pinski e Narin (1976), temos a declaração de que a citação à literatura acadêmica já havia sido aplicada usando dos recursos web de contagem de links, tendo Page et al. (1999) inclusive citado um trabalho de Eugene Garfield².

Assim, seguiu-se pelo entendimento de que técnicas de estudos de citação poderiam ser transpostas para estudos onde o link³ é a unidade de vinculação entre documentos. Nasce então a Webometria, mas não sem uma pequena trajetória antes de sua consolidação terminológica.

2 O ÚLTIMO DA NINHADA?

O processo de abertura para uso comercial da internet se inicia em 1995 e é neste momento que, com cerca de 12 anos desde a sua configuração como rede científica, Marcia Bossy apresenta sua visão para as potencialidades de estudos métricos da informação na internet. No artigo intitulado “*The last of the litter: ‘netometrics’*”⁴, Bossy (1995) apresenta um novo campo propositivamente nomeado como “netometrics”, o então “último da ninhada” à época dentre os subcampos dos estudos métricos da informação.

No texto Bossy (1995) revela seu interesse de análise na repercussão da internet nas comunidades científicas e, conseqüentemente, na cientometria e como essas interações mereciam ser sistematizadas enquanto campo de observação. Num olhar latouriano (de Bruno Latour) para o ambiente virtual, a autora considera que na internet seria possível ver a “ciência em ação”, dado que os processos de comunicação entre os cientistas estariam mais e mais registrados neste espaço. Em sua concepção a autora considera que

² Citado como “GARFIELD, E. New international professional society signals the maturing of scientometrics and informetrics. *The Scientist*, v. 9, n. 16, Aug. 1995. http://www.the-scientist.library.upenn.edu/yr1995/august/issi_950821.ht%ml” mas acessível em [<https://www.the-scientist.com/news/new-international-professional-society-signals-the-maturing-of-scientometrics-and-informetrics-58398>]

³ Para um aprofundamento recomendamos ver Vanti (2005)

⁴ “O último da ninhada: ‘netometrics’” (tradução nossa)

Devemos integrar a realidade da rede eletrônica em nosso modelo de fluxo de informações entre os cientistas, detectar os elos que consolidam os colaboradores, conceber meios de entender as práticas dessas comunidades científicas globais, estabelecer indicadores relevantes para a atividade sócio-cognitiva que ocorre na internet e finalmente, incorporar nossos resultados em ferramentas de gerenciamento de fluxo de informações voltadas para as novas comunidades científicas (BOSSY, 1995 - tradução nossa)⁵.

Infelizmente, o texto foi publicado online em inglês em uma revista francesa, possivelmente resultando em menor visibilidade, sendo rara a menção por autores não brasileiros deste artigo visionário. Lembremo-nos que estávamos na década de 90 do século passado, onde somente algumas iniciativas pioneiras estavam usando e explorando o formato online, donde aqui podemos destacar o pioneirismo de Aldo de Albuquerque Barreto com a DataGramZero, como proeminente exemplo brasileiro.

As propostas para nomear o campo se seguiram, e de forma não exaustiva⁶ citamos a proposta de “webometry” por Abraham (1996) até “webometrics” postumamente por Almind, que faleceu em um acidente de carro, e Ingwersen (ALMIND; INGWERSEN, 1997). Já a proposta de Chakrabarti et al. (2002) de nomear o campo como “web bibliometry” não teve eco na comunidade.

Ao destacar o crescimento explosivo da *World Wide Web* (WWW ou web) nos últimos anos, de uma nova ideia de Tim Berners-Lee ao sistema nervoso de uma nova sociedade planetária, Abraham (1996) destaca alguns paradigmas históricos e dentre eles situa na

⁵ No original: “We must integrate the electronic network reality in our model for information flow among scientists, detect the links that consolidates laboratories, devise means of understanding the practices of these global scientific communities, establish relevant indicators for the socio-cognitive activity taking place on the Net and finally, embody our results into information flow management tools aimed at the new scientific communities”.

⁶ Para uma visão histórica das terminologias ver Vanti (2002)

visão pós-moderna do mundo, no qual, as esferas terrestre, humana e celeste estão todas em um processo de coevolução concomitante, como na embriogênese de uma nova sociedade planetária. Segundo o autor, nesse hábito de pensar, a web pode ser considerada como a neurogênese do cérebro global, intrínseca e essencial para a coevolução geral de tudo, mas que, para que essa evolução tenha um resultado favorável:

[...] devemos participar da consciência emergente do cérebro global e, portanto, devemos visualizar, observar e interagir com a explosão da web. É por causa dessa crença que desenvolvemos as ferramentas da *webometry* que são descritas neste artigo: as ferramentas de observação da Web. (ABRAHAM, 1996 - tradução e grifo nosso⁷).

O campo, no entanto, já havia sido semeado com o primeiro estudo de *co-links*, o equivalente para a webometria para um estudo de co-citação, efetuado por Larson (1996). Seguindo a lógica do fator de impacto para as revistas, Rodrigues I Garin (1997) faz uma proposição de fator de impacto web, mas novamente o reconhecimento da primazia fica comprometido pela publicação ter sido efetuada em revista de circulação mais restrita e em espanhol. A atribuição para o Fator de Impacto Web (FIW ou *Web Impact Factor* - WIF) costuma ser feita para Ingwersen (1998). Dentre as alterações propostas durante o tempo, convém destacar a de uso de Logaritmo Natural para atenuação do denominador – que consiste no número de páginas do site – proposta por Vanti, Costa e da Silva (2013), o qual foi aplicado a um conjunto de páginas web de universidades brasileiras e replicado por Vanti e Sanz Casado (2015) a outro conjunto de páginas web, desta vez de universidades espanholas para ratificar sua eficácia.

Para Björneborn (2004, p. 12 et seq. – tradução e grifo nosso), a webometria seria “o estudo dos aspectos quantitativos da construção

⁷ No original: “we must participate in the emerging consciousness of the global brain, and thus, we must visualize, observe, and interact with the explosion of the WWW. It is because of this belief that we have developed the tools of webometry which are described in this paper: the tools of Web Watch”

e uso dos recursos de informação, estruturas e tecnologias da **web** a partir de abordagens informétricas e bibliométricas”⁸, enquanto a cibermetria seria “o estudo dos aspectos quantitativos da construção e uso dos recursos de informação, estruturas e tecnologias da **internet** como um todo a partir de abordagens informétricas e bibliométricas”⁹. Assim, a cibermetria contém a webometria dentro de si, pela lógica de que a internet engloba a web (THELWALL; VAUGHAN; BJÖRNEBORN, 2005).

Pesquisas que sejam sobre o uso da internet em formas não majoritariamente dependentes do protocolo HTTP – como acessos diretos a arquivos, trocas de e-mail, ambientes virtuais ou jogos eletrônicos ou mesmo acesso a conteúdos digitais online – seriam, portanto, cibermétricos. Já estudos que sejam fundados nas métricas de acesso a servidores web configuram o que em geral é qualificado pelos termos “web metrics” ou “webmetrics” (em português, webmetria sem o “o” derivado da analogia com bibliometria) (GOUVEIA, 2012)

3 ESTUDOS WEBOMETRICOS: ASCENSÃO E “CRISE”

A webometria como campo se estabeleceu com foco principal em estudos de Fator de Impacto Web, estudos de *co-links* (similar a co-citação) e de *interlinks* (similar aos estudos de citação mútua). Fortemente baseados na coleta de dados por intermédio de mecanismos de busca, podiam inferir a partir de resultados da web visível (aquela que é mapeada pelos mecanismos de busca) as proximidades entre sites do ponto de vista externo (*co-links*), da troca de links entre eles (*interlinks*) e seu “impacto” contado pelo número de links recebidos dividido pelo número de páginas (conteúdos) catalogadas.

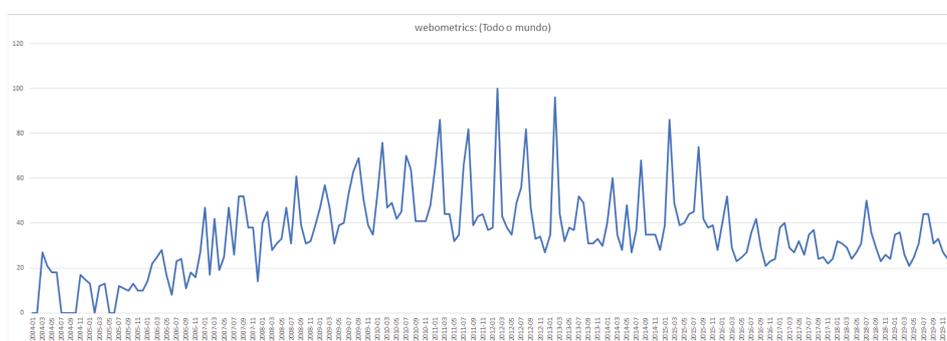
Os indicadores webométricos foram fonte também para um

⁸ No original “*The study of the quantitative aspects of the construction and use of information resources, structures and technologies on the Web, drawing on bibliometric and informetric approaches*”

⁹ No original “*The study of the quantitative aspects of the construction and use of information resources, structures and technologies on the whole Internet, drawing on bibliometric and informetric approaches*”

ranking, o *Webometrics Ranking*¹⁰ desenvolvido pelo *Cybermetrics Lab*, grupo de pesquisa do *Consejo Superior de Investigaciones Científicas (CSIC)* da Espanha. Nele, vários parâmetros são utilizados para sua concepção¹¹. Na Figura 1 podemos ver que o interesse de busca pelo termo “webometrics” (de 2004 até 2019) tem se mantido ao longo do tempo, com picos associados à divulgação do *Webometrics Ranking*.

Figura 1 - Diagrama de Venn da intercessão entre as *hashtags* atribuídas para a primeira imagem, por curso



Fonte: Google Trends¹²¹³

Entretanto, interesses comerciais começaram a surgir em relação aos dados que davam alicerce à webometria. Um a um, os mecanismos de busca começaram a fechar o acesso aos seus dados, sendo hoje possível obtê-los apenas por intermédio de assinaturas de provedores ou fazendo o levantamento num universo de sites pré-definido. Assim se deu a “crise” da webometria que obrigou a busca por novos caminhos de pesquisa (GOUVEIA, 2012).

¹⁰ Disponível em: <http://www.webometrics.info/>

¹¹ Os parâmetros mudaram ao longo do tempo, não sendo os resultados do ranking comparáveis de forma longitudinal, mas em geral eles sempre incluíram o número de páginas, número de links para as páginas. Além disso já utilizou o número de documentos disponibilizados (pdfs, xls, ppts, etc...) e alguma forma de se levantar citações à produção da instituição (atualmente no Google Scholar), porém hoje usa o número de citações para os pesquisadores nos seus perfis do Google Scholar e para os artigos no Scimago.

¹² Disponível em: <http://trends.google.com>

¹³ Dados coletados em 20 de junho de 2020.

Propostas de uso de termos no lugar de links como unidade de análise foram efetuados, e alternativas com levantamentos mais exaustivos usando *crawlers*¹⁴ personalizados também passaram a ser o foco, mas na prática o campo que era emergente seguiu com algum interesse, mas perdeu sua tendência de crescimento.

4 A REINVENÇÃO DA WEBOMETRIA

Seguindo seu caminho como campo dos estudos métricos da informação, a webometria tenderia naturalmente a abarcar estudos que levassem em consideração aspectos de links e acessos às páginas web e citações. Vaughan e Hysen (2002) observaram relações entre os links para páginas de revistas e as citações para seus artigos. Thelwall Vaughan e Björneborn (2005) já falavam da relação entre downloads de *preprints*, e da rápida citação desses artigos.

Já a conectividade em sites de redes sociais veio em seguida. Ao se fazer uma postagem, está se criando um elemento em uma página que tem vínculos (*links*) com o autor, local de publicação e cujas reações (curtidas, comentários e compartilhamentos) geram conexões entre outros elementos que são representados por *links* no site de rede social. Taraborelli (2008) discutiu o cenário da avaliação científica por pares na web 2.0, enquanto Neylon e Wu (2009) pensou nas métricas ao nível de artigo diante dessas novas possibilidades.

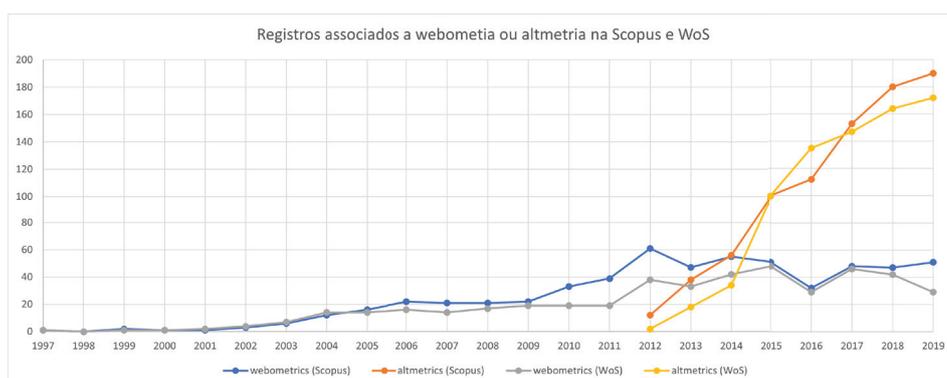
Entretanto, conteúdos em sites de redes sociais são percebidos como uma nova segmentação para os estudos métricos da informação. Indiretamente retornando ao visionado por Bossy (1995), o Manifesto Altmétrico (PRIEM et al., 2010) expressa a possibilidade de acelerar o processo de percepção da ciência e seus impactos ao olhar para o ambiente dos sites de redes sociais como fonte de dados.

Assim seria observada a “ciência em ação”, pensada por Bossy

¹⁴ Um *crawler*, ou rastreador da rede, pode ser definido como um software que navega uma rede a partir dos links entre as páginas de forma automática de forma a mapear o seu conteúdo.

(1995), onde todas as interações contariam, como pedido por Piwowar (2013). Por outro lado, se formos olhar de forma estrita, estas ações observadas estão na forma de *links*, o que resulta em se considerar que a altmetria trata de estudos cientométricos que fazem uso de dados webométricos e cibernétricos (da internet, ou, para além da web) (GOUVEIA, 2013). Na Figura 2 podemos ver como estes campos (a webometria e a altmetria) se configuraram em termos de publicações ao longo do tempo nas bases *Web of Science* e *Scopus* (de 1997 até 2019).

Figura 2 - Registros associados à webometria ou altmetria na Scopus e Web of Science de 1997 a 2019



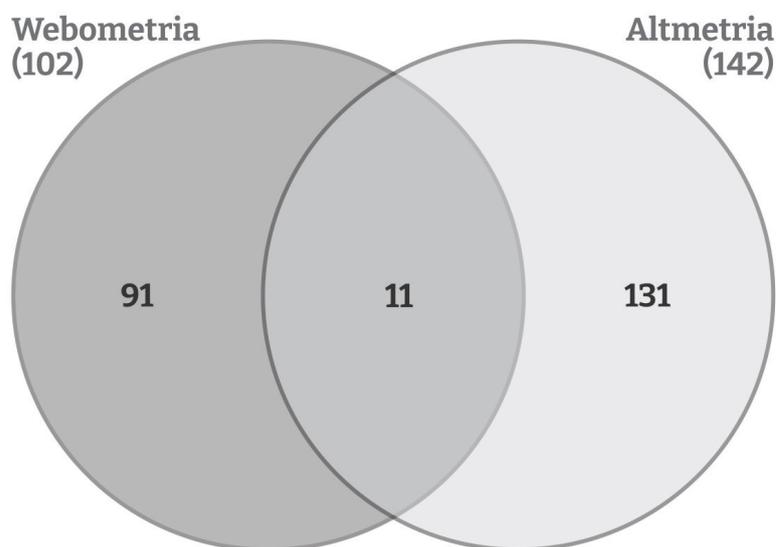
Fonte: Dados da pesquisa¹⁵

Ambos estão concentrados de acordo com a classificação da *Web of Science* em “*Information Science Library Science*” (67% para webometria e 56% para altmetria) tendo com áreas próximas representando de 13% a 29% dos estudos “*Computer Science Information Systems*” e “*Computer Science Interdisciplinary Applications*”. Entretanto, percebe-se a perda de tendência dos estudos em webometria a partir do ano de 2012 com o surgimento das primeiras publicações indexadas nas bases sobre altmetria. Estudos que não se enquadram nas análises de dados de sites de redes sociais ainda persistem, mas num nicho específico, sem crescimento aparente.

¹⁵ Dados coletados em 20 de junho de 2020.

Já uma busca por assunto no Currículo Lattes retorna 102 CVs com menções aos termos associados à webometria¹⁶ contra 142 para menções a termos associados à “altmetria”¹⁷. Se representarmos estes currículos encontrados em um diagrama de Venn, apenas 11 são encontrados com menções aos dois campos (Figura 3).

Figura 3 - Diagrama de Venn entre CVs Lattes com menções a termos associados à “webometria” e “altmetria”



Fonte: Dados da Plataforma Lattes¹⁸¹⁹ - Desenvolvido com Interactivenn²⁰

Aparentemente, no panorama de pesquisa do Brasil, pouca é a sobreposição dos campos se considerarmos que as menções num currículo Lattes podem ser até mesmo uma participação em banca.

¹⁶ “webometria”, “webométrica”, “webométrico”, “webmetric” ou “webometrics”

¹⁷ “altmetria”, “altmétrico”, “altmétrica”, “altmetric” ou “altmetrics”

¹⁸ Acessível em: <http://lattes.cnpq.br>

¹⁹ Dados coletados em 20 de junho de 2020.

²⁰ Disponível em: <http://www.interactivenn.net> - Heberle et al. (2015).

5 USOS CONTEMPORÂNEOS PARA MÉTRICAS WEB E DE REDES DE CONTEÚDOS

É comum nas análises quantitativas dos subcampos dos estudos métricos da informação a utilização de instrumentos para coleta, extração, tratamento e visualização de dados. Os estudos webométricos também contam com recursos para estas finalidades e entre os mais comuns estão os motores de busca, as aplicações de rastreamento e os programas de representação e visualização de redes. Algumas soluções combinam um ou mais destes recursos e nos permitem obter desde dados de acesso como de correlação entre conteúdos web e sites.

Ferramentas de *Web Analytics* que fazem análise de acesso e tráfego permitem a obtenção de registros de navegação, ou seja, o levantamento de dados sobre o comportamento de acesso e uso de sites, possibilitando aos seus gestores ter melhor compreensão sobre quem são seus visitantes, de onde eles vêm, quais páginas acessam, se encontram o que procuram, e quanto tempo permanecem nas páginas visitadas.

Entre as opções para análise de acesso (perfil de usuário e padrão de uso) o *Google Analytics* (<https://analytics.google.com/>) talvez seja a ferramenta mais conhecida. Trata-se de um serviço gratuito, oferecido pela *Google LLC*, que funciona a partir do monitoramento e mensuração de dados de acesso e uso de um determinado site por seus usuários com disponibilização de relatórios estatísticos de desempenho.

Uma outra forma de ver os processos de tráfego da rede, mas a partir de um olhar externo, e acessando o site da ferramenta *Alexa* (<http://www.alexa.com/>) uma empresa da *Amazon*²¹. Os dados abertos são limitados na versão gratuita, mas é possível saber em que lugar do ranking de acesso aos conteúdos o *Alexa* estima que o site esteja tanto mundialmente quanto no país de maior nível encontrado (<https://>

²¹ Disponível em: <http://www.amazon.com>

www.alexa.com/siteinfo).

Para os bibliotecários, também pode ser estratégico acessar versões anteriores de sites ou mesmo páginas que foram citadas e que não estão mais disponíveis online provisoriamente (caso comum durante períodos eleitorais) ou definitivamente. O *Internet Archive* (<http://archive.org>), uma organização sem fins lucrativos criada em 1996 que atua preservando uma biblioteca digital da internet e de outros artefatos digitais, dispõe de um buscador chamado *Wayback Machine* (<https://archive.org/web/>) onde é possível obter versões passadas de sites arquivados.

Tendências de buscas também podem indicar caminhos de pesquisa. O interesse relativo por um termo ou conceito em diferentes localidades ao longo do tempo é o que oferece o *Google Trends* (<http://trends.google.com>) cujo exemplo pode ser visto na própria Figura 1 apresentada previamente.

Ainda dentro das análises que fazem uso da rede de links temos o projeto *carrot2* (disponível no site: <https://search.carrot2.org/>) que faz uso de dados do site <http://etools.ch>, um meta-mecanismo de busca da Suíça. De maneira muito simples e direta, permite que seja feita uma busca por um termo e que se faça uma navegação a partir de uma projeção bidimensional em árvore. Há três algoritmos de agrupamento disponíveis além da possibilidade de se fazer um levantamento usando a base *PubMed* (<https://pubmed.ncbi.nlm.nih.gov/>). Na Figura 4 temos uma representação de busca pelo termo “webometrics” onde foram aplicadas as configurações padrão.

Por último, mas não menos importante, podemos citar uma ferramenta robusta, que combina algumas das características presentes nas listadas anteriormente e que tem sido utilizada em pesquisas acadêmicas denotando a forte aproximação dos estudos webométricos e suas variações no contexto científico. Trata-se do *Webometric Analyst* (<http://lexiurl.wlv.ac.uk/>), um programa gratuito, desenvolvida pelo *Statistical Cybermetrics Research Group* utilizado para coleta e visualização de dados de citações, de menções em fontes e recursos da web social e webometria, incluindo análise de links. O programa coleta dados da web por meio de Interface de Programação de Aplicações (*Application Programming Interface - APIs*) ou *download* direto e inclui uma ampla variedade de opções de processamento.

As aplicações listadas não esgotam as ferramentas disponíveis para o uso prático na pesquisa ou atuação profissional dos interessados na webometria, mas dão uma boa ideia de seus benefícios e do seu lugar nos campos dos estudos métricos da informação. Como ferramentas de *Web Analytics* o principal problema da análise pode girar em torno de se obter estatísticas confiáveis, uma vez que os resultados obtidos em diferentes sistemas analíticos podem variar significativamente. Além disso, como desvantagens, podem apresentar imprecisões relacionadas aos variados algoritmos de coleta de dados; privacidade/anonimato dos usuários da web; falhas técnicas; e a complexidade das diferenças entre os diferentes tipos de tráfego na web (UDARTSEVA, 2018).

Assim, interessados nessas aplicações webométricas no contexto das bibliotecas²³ precisam prestar atenção a esses problemas e tentar considerá-los em suas próprias avaliações, a fim de obter o resultado final mais objetivo e evitar erros na interpretação dos indicadores quantitativos (UDARTSEVA, 2018).

Com o crescente interesse nos espaços online para

²³ Para um olhar mais aprofundado de aplicações webométricas no contexto das bibliotecas, recomendamos consultar a revisão de literatura: UDARTSEVA, O. M. An Overview of Webometrics in Libraries: History and Modern Development Tendencies. *Sci. Tech. Inf. Proc.*, v. 45, p.174-181, 2018. Disponível em: <https://doi.org/10.3103/S0147688218030115>. Acesso em: 20 jun. 2020.

comunicação científica, e o constante desenvolvimento de ferramentas disponíveis para o monitoramento de indicadores de acesso e uso, os bibliotecários estão em uma posição-chave para assumir a liderança no reforço do conhecimento das atuais tendências de mensuração do impacto da informação científica (ARAÚJO, 2015). Assim como em aplicações altmétricas, os bibliotecários devem conhecer e explorar as ferramentas para coleta e geração de indicadores e encontrar formas de integrar tais métricas às práticas de treinamento da biblioteca (NASCIMENTO, 2016).

E uma vez que os bibliotecários e profissionais da informação estão envolvidos, em grande parte, no ciclo de pesquisa, comunicação científica, sistemas de publicação de revistas, bases de dados, bibliotecas digitais e repositórios, parece um caminho natural que eles estejam envolvidos na aplicação de métricas nestes produtos e recursos de informação em suas unidades. Podemos destacar alguns estudos desenvolvidos com aplicações webométricas nestes contextos.

Com o objetivo de medir e analisar o uso e a impressão dos usuários, da biblioteca protótipo “Biblioteca nas Nuvens” Andrade e Araújo (2014) desenvolveram um estudo webométrico, ancorado numa abordagem metodológica mista com coleta dos dados via *Google Analytics* e estatísticas disponibilizadas pela fanpage do *Facebook*. Segundo os autores as métricas aplicadas para medir a utilização do protótipo permitiram aferir o acesso e identificar a satisfação dos usuários, com resultados considerados positivos tendo em vista que 87% dos usuários retornaram à biblioteca. Para eles “a Webometria deve ser utilizada pelas bibliotecas digitais, buscando medir e analisar a impressão dos usuários acerca dos serviços e acervo oferecidos pela biblioteca” (ANDRADE; ARAÚJO, 2014, p.3288).

Silveira, Afonso e Matias (2019) analisaram o desempenho dos websites institucionais das Bibliotecas Nacionais no MERCOSUL por meio de métricas da webometria e constataram que utilização dos indicadores webométricos indicaram que os websites destas instituições podem servir como fonte de informação, para investigar

sua atratividade, e o seu impacto na web, os quais estão associados às informações e ao conteúdo disponibilizados na rede.

A webometria foi aplicada em repositórios institucionais acadêmicos por Shintaku, Robredo e Baptista (2011). Para os autores métricas aplicadas em repositórios podem apresentar informações importantes sobre sua instituição e os indicadores webométricos “fornecem um panorama geral da produção intelectual de uma instituição, verificando aspectos que permitem situar a instituição sob vários aspectos, com destaque à visibilidade e impacto” (SHINTAKU; ROBREDO; BAPTISTA, 2011, p.324).

Com o propósito de oferecer aos editores das revistas de ciências da comunicação parceiras do Portal Revcom subsídios para análises quanto a acessibilidade, visibilidade e uso pela comunidade de seus usuários, Ferreira e Cunha (2008) implementaram uma ferramenta estatística para análise de *logs* de acesso com base no software Google Analytics. As aplicações webométricas foram feitas tanto das páginas introdutórias e iniciais do Portal como individualmente das onze revistas nele disponíveis na época. O estudo conseguiu verificar variáveis como visitas e visitantes, padrões de uso e perfil do usuário e seus resultados indicaram que as revistas do Portal Revcom têm tido uma procura considerável, além de comprovar que a participação em um Portal de periódicos, que congrega várias outras revistas da mesma área, potencializa a visibilidade e acessibilidade destas revistas.

Com o objetivo de revisar os conceitos e aplicações da webometria no âmbito dos periódicos científicos e assim contribuir com a discussão atual sobre seus indicadores, o estudo de Shintaku (2017), com coleta de dados via buscadores o Google e Google Acadêmico em consultas por URLs da revista (página inicial; edição; página de sumário) e do artigo (página de identificação do artigo), revelou que os *inlinks* de uma revista eletrônica online são como citações em sua grande parte, alinhando a bibliometria à webometria, mas que podem também ser indicação de artigos em redes sociais, o que aproximaria o campo da altmetria.

Assim, as possibilidades de aplicações de estudos webométricos seguem abertas. Pode-se levar em conta dados de acesso e disponibilização de conteúdo, volume de links atribuídos à sites, co-ocorrência de *links* (*co-links*) para páginas de estudo, trocas de links entre páginas (*interlinks*), ou outras formas mais recentes onde se utilizam termos no lugar de links como unidade de estudo. Mas para além disso, o entendimento da web como organismo vivo coloca aos bibliotecários a responsabilidade de lidar com uma “biblioteca” com alguma instabilidade no seu acervo, onde pode ser necessário obter um registro perdido que foi arquivado no *Internet Archive*, compreender as dinâmicas de interesses de busca por temas ao longo do tempo, perceber as interrelações dos termos conforme são gerados por usuários da internet e ser capaz de lidar com retratos gerados pelo “macroscópio” dos estudos métricos da informação aplicados à web.

REFERÊNCIAS

ABRAHAM, R. H. **Webometry**: measuring the complexity of the World Wide Web. 1996. Disponível em: <http://www.ralph-abraham.org/articles/MS%2385.Web1/>. Acesso em: 20 jun. 2020.

ALMIND, T. C.; INGWERSEN, P. Informetric analyses on the world wide web: methodological approaches to 'Webometrics'. **Journal of Documentation**, London, v. 53, n. 4, p. 404-426, 1997.

ANDRADE, R. L. V. ; ARAUJO, W. J. . Análise Webométrica aplicada ao site 'Biblioteca nas Nuvens'. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO: além das nuvens, expandindo as fronteiras da Ciência da Informação, 14., 2014, Belo Horizonte. **Anais** [...]. Belo Horizonte: ECI, UFMG, 2014. p. 3288-3306.

ARAUJO, R. F. Estudos métricos da informação na web e o papel dos profissionais da informação. **Bibliotecas Universitarias: Pesquisas, Experiencias e Perspectivas**, v. 2, p. 42-64, 2015.

BJÖRNEBORN, L. **Small-World link structures across an academic web space**: a Library and Information Science approach. 2004. 469 f. Tese (Doutorado) - Royal School of Library and Information Science. Disponível em: https://static-curis.ku.dk/portal/files/47039808/lennart_bjorneborn_phd.pdf. Acesso em: 20 jun. 2020.

BOSSY, M. J. The Last of the Litter: "Netometrics". **Solaris**, v. 2, p. 1-5, 1995. Disponível em: <http://gabriel.gallezot.free.fr/Solaris/d02/2bossy.html>. Acesso em: 20 jun. 2020.

BRIN, S.; PAGE, L. The anatomy of a large-scale hypertextual Web search engine. **Computer Networks and ISDN Systems**, v. 30, n. 1-7, p. 107-117, abr. 1998.

CHAKRABARTI, S.; JOSHI, M. M.; PUNERA, K.; PENNOCK, D. M. The structure of broad topics on the Web, Proceedings of the WWW2002 Conference. 2002. Disponível em: <https://web.archive.org/web/20120708004804/http://www2002.org/CDROM/refereed/338/>. Acesso em: 20 jun. 2020.

FERREIRA, S. M. S. P.; CUNHA, A. S. Portal Revcom & Google Analytics: acessando a caixa-preta da informação. **Em Questão**, n. 1, v. 14, p. 41-61, 2008.

FUJINO, A.; JACOMINI, D. D. Produtos e Serviços de informação na sociedade do conhecimento: da identificação ao uso. In: GIANNASI-KAIMEN, M. J.; CARELLIANA, A. E. (org.). **Recursos informacionais para compartilhamento da informação**: redesenhando acesso, disponibilidade e uso. Rio de Janeiro: E-Papers, 2007. p. 73-98

GOUVEIA, F. C. Novos caminhos e alternativas para a Webometria. **Em Questão**, v. 18, n. 3, p. 249-261, 2012.

GOUVEIA, F. C. Altméria: métricas de produção científica para além das citações. **Liinc em Revista**, v. 9, n. 1, p. 214-227, 2013.

HEBERLE, H.; MEIRELLES, G. V.; DA SILVA, F. R.; TELLES, G. P.; MINGHIM, R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. **BMC Bioinformatics**, v. 16, n. 169, p. 1-7, 2015. DOI 10.1186/s12859-015-0611-3.

INGWERSEN, P. The Calculation of Web Impact Factors. **Journal of Documentation**, London, v. 54, n. 2, p. 236-243, 1998.

LARSON, R. Bibliometrics of the world wide web: An exploratory analysis of the intellectual structure of the cyberspace. In: ANNUAL MEETING OF THE AMERICAN SOCIETY OF INFORMATION SCIENCE, 1996. **Proceedings** [...]. [S.l.: s.n.], 1996.

LOVE, B. C.; SLOMAN, S. A. Mutability and the determinants of conceptual transformability. In: ANNUAL CONFERENCE OF THE

COGNITIVE SCIENCE SOCIETY, 17., 1995. **Proceedings** [...]. [S.l.: s.n.], 1995. p. 654–659.

NASCIMENTO, A. G. **Altmetria para bibliotecários: guia prático de métricas alternativas para avaliação da produção científica**. Rio de Janeiro: Revolução eBook, 2016. Disponível em: <http://eepurl.com/dbCRbX>. Acesso: 20 jun. 2020.

NEYLON, C.; WU, S. Article-level metrics and the evolution of scientific impact. **PLoS biology**, v. 7, n. 11, p. e1000242, 2009.

NORUZI, A. Application of Ranganathan's laws to the web. **Webology**, v. 1, n. 2, a. 8, s.p., 2004. Disponível em: <http://www.webology.org/2004/v1n2/a8.html>. Acesso em: 20 jun. 2020.

PAGE, L. BRIN, S. MOTWANI, R. WINOGRAD, t. **The PageRank citation ranking: bringing order to the web**. [S.l.]: Stanford InfoLab, 1999. Disponível em: <http://ilpubs.stanford.edu:8090/422/>. Acesso em: 20 jun. 2020.

PIWOWAR, H. Value all research products. **Nature**, v. 493, p. 159, 2013.

PRIEM, J.; TARABORELLI, D.; GROTH, P.; NEYLON, C. **Altmetrics: a manifesto**. v. 1.0, 26 Oct.2010. Disponível em: <http://altmetrics.org/manifesto>. Acesso em: 20 jun. 2020.

RANGANATHAN, S. R. **The five laws of library science**. Madras: The Madras Library Association, 1931. Disponível em: [https://hdl.handle.net/2027/uc1.\\$b99721](https://hdl.handle.net/2027/uc1.$b99721). Acesso em: 20 jun. 2020.

RODRÍGUEZ I GARÍN, J. M. Valoración del impacto de la información en Internet: Altavista, el "Citation Index" de la red. **Revista Española de Documentación Científica**, v. 20, n. 2, p. 175-181, 1997.

SHINTAKU, M.; ROBREDO, J.; BAPTISTA, D. M. Webometria dos repositórios institucionais acadêmicos. **Ciência da Informação**, v. 40, n. 2, sep. 2012. ISSN 1518-8353. Disponível em: <http://revista.ibict.br/ciinf/article/view/1319>. Acesso em: 13 jun. 2020.

SHINTAKU, M. Webometria e os periódicos científicos eletrônicos. **Informação & Sociedade: Estudos**, v. 27, n. 3, 2017. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/34796>. Acesso em: 20 jun. 2020.

SILVEIRA, E.; AFONSO, R. D.; MATIAS, M. R. Bibliotecas nacionais do

Mercosul: um estudo webométrico em seus websites institucionais. *In: ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA*, 5., 2016, São Paulo. **Anais [...]**. São Paulo: USP, 2016.

TARABORELLI, D. Soft peer review: social software and distributed scientific evaluation. *In: INTERNATIONAL CONFERENCE ON THE DESIGN OF COOPERATIVE SYSTEMS*, 8., 2008, Carry-le-Rouet. **Proceedings [...]**. França: Institut d'Etudes Politiques d'Aix-en-Provence, 2008. p. 99–110.

THELWALL, M.; VAUGHAN, L.; BJÖRNEBORN, L. Webometrics. **Annual Review of Information Science and Technology**, Medford, v. 39, n. 1, p. 81-135, 2005.

UDARTSEVA, O.M. An Overview of Webometrics in Libraries: History and Modern Development Tendencies. **Sci. Tech. Inf. Proc.**, v. 45, p.174-181, 2018.

VANTI, N. A. P. 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, 2002.

VANTI, N. Os links e os estudos webométricos. **Ciência da Informação**, v. 34, n. 1, p. 78-88, 2005.

VANTI, N.; COSTA, J. A. F.; DA SILVA, I. C. O. Nova fórmula revisada para o cálculo do fator de impacto web (FIW). **Liinc em Revista**, v. 9, n. 1, p. 228-236, 2013.

VANTI, N.; SANZ CASADO, E. O uso do Fator de Impacto Web alternativo para avaliar as universidades públicas espanholas. *In: ARAÚJO, R. (org.). Estudos métricos da informação na web: atores, ações e dispositivos*. 1. ed. Maceió: EDUFAL, 2015. p. 109-128.

VAUGHAN, L.; HYSEN, K. Relationship between links to journal web sites and impact factors. **Aslib Proceedings**, v. 54, n. 6, p. 356–361, 2002.

WU, J.; CHEN, R. Metaphors ubiquitous in computer and internet terminologies. **Journal of Arts and Humanities**, v. 2, n. 10, p. 64-78, 2013.