

UNIDADE 3

MÉTODOS QUANTITATIVOS ESPECÍFICOS DA BIBLIOTECONOMIA: COMO QUANTIFICAR COMPORTAMENTOS CULTURAIS?

3.1 OBJETIVO GERAL

Apresentar técnicas quantitativas específicas utilizadas pelos profissionais de Biblioteconomia, bem como as ferramentas apropriadas a elas.

3.2 OBJETIVOS ESPECÍFICOS

Esperamos que, ao final desta Unidade, você seja capaz de:

- a) descrever o objeto de estudo, as variáveis, os métodos e os objetivos da bibliometria, infometria, cientometria e webometria;
 - b) definir indicadores bibliométricos, reconhecendo algumas de suas limitações;
 - c) descrever as Leis de *Bradford*, *Lotka* e *Zipf*;
 - d) efetuar a análise bibliométrica das bases de dados *Web of Science*, *Scopus* e *SciELO*;
 - e) operar *softwares* bibliométricos.
-

3.3 BIBLIOMETRIA, INFOMETRIA, CIBERMETRIA, WEBOMETRIA, CIENTOMETRIA, PATENTOMETRIA

Na Unidade anterior, estudamos ferramentas quantitativas gerais, que são aplicadas ao levantamento e análise de dados de pesquisas e à própria Biblioteconomia.

Nesta Unidade, você entrará em contato com algumas ferramentas quantitativas que foram desenvolvidas especificamente para o bibliotecário e/ ou cientista da informação.

Inicialmente, precisamos delimitar as várias subáreas quantitativas relacionadas com a Biblioteconomia.

Vamos lá?

A partir dos investimentos massivos realizados em Ciência e Tecnologia, pelos diversos países envolvidos na Segunda Guerra Mundial, a gestão e a produção do conhecimento mudaram de escala, atingindo níveis sem precedentes na história. A Ciência e a Tecnologia passam a ser produzidas por grandes equipes de pesquisadores, exigindo enormes montantes de recursos e laboratórios, que, por sua vez, também exigem grandes recursos para serem administrados. A produção de artigos científicos, patentes e livros, o surgimento de novas subáreas do conhecimento e, mais recentemente, da internet, bem como a criação de periódicos científicos eletrônicos, eventos e congressos virtuais, elevaram a complexidade da avaliação da produção científica e tecnológica. Tudo isso aconteceu dentro de uma mesma área do conhecimento e em um mesmo conjunto de países.

3.3.1 As várias “metrias”

Não há um consenso entre os especialistas das fronteiras entre as várias “metrias”. Seguindo a definição proposta por *Bjorneborne* e *Ingwersen* (2004), citados por *Lucas, Zorita* e *Casado* (2013, p. 257. Grifos dos autores), a

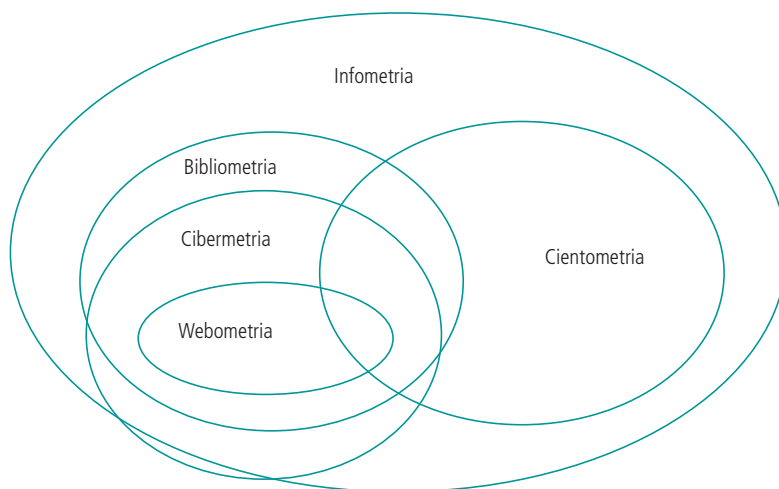
[...] *Cientometria* diz respeito aos estudos da atividade científica; a *Bibliometria*, aos estudos da produção da informação; os *Estudos de usuários* (área de estudos também chamada de *Bibliotecometria*) reportam aos estudos de consumo de informação e das necessidades de informação; a *Patentometria* diz respeito aos estudos da inovação; a *Cibermetria* aos estudos de internet; e a *Webometria*, os estudos de sites.

Semestre

7

Na citação anterior, não é apresentado o termo infometria, que estudaremos nas próximas páginas. As Figuras 29 e 30, a seguir, extraídas dos respectivos artigos citados, apresentam duas propostas para o entendimento das relações e fronteiras entre as várias “metrias”:

Figura 29 – Relações entre as disciplinas métricas

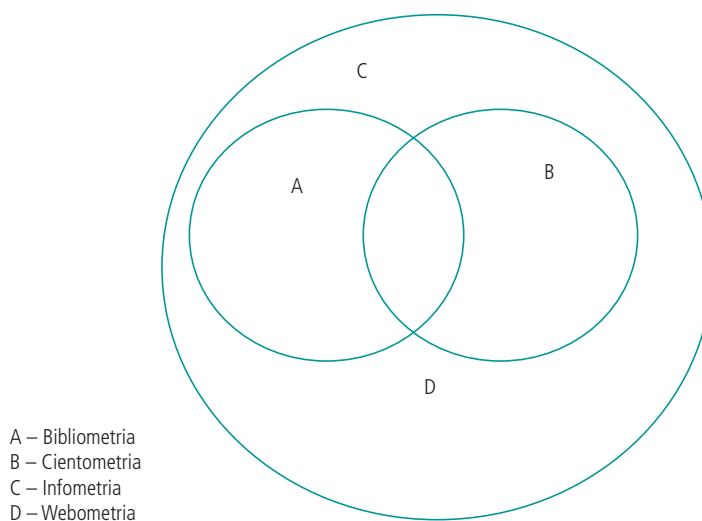


Fonte: Lucas; Zorita; Casado (2013, p. 258).

Observe, na Figura 30, que há uma intersecção entre o foco de estudo da bibliometria e da cientometria, mas essas duas disciplinas não se confundem. Por essa figura, a infometria abrange todas as outras “metrias”. Não há consenso entre os autores sobre as fronteiras entre as várias “metrias”. Para alguns deles (por exemplo, *Bjorneborn* e *Ingwesen, op. cit.*), a webometria é tratada como uma área dentro da cibermetria. Já para *Bar-Ilan* (2008), *cibermetria* e *webometria* são considerados sinônimos.

De *Vanti* (2002) extraímos o diagrama ilustrativo (Figura 30):

Figura 30 – Intersecção entre as “metrias”



A – Bibliometria
B – Cientometria
C – Infometria
D – Webometria

Fonte: *Vanti* (2002, p. 10).

Na Figura 30, não está incluída a cibermetria. Com esse diagrama, você pode observar melhor as relações de intersecção entre as “metrias” apresentadas.

De acordo com *Le Coadic* (2005) (apud SANTIN, 2011, p. 109), as variáveis mais utilizadas nas análises métricas são: título, palavras-chave e nome dos autores. Com as palavras-chave, o pesquisador pode efetuar inferências a respeito dos temas em destaque em uma área do conhecimento. Ainda segundo o autor,

a ocorrência de dois termos num mesmo documento ou base de dados representa a associação entre os assuntos, permitindo destacar as tendências de investigação e os principais centros de interesse dos cientistas (LE COADIC, 2005 apud SANTIN, 2011, p. 109).

Na sequência, você terá um detalhamento de cada uma das “metrias”: bibliometria, cientometria, infometria e webometria.

3.3.2 Bibliometria

Do inglês, *bibliometrics*. De acordo com *Chaviano* (2004), o primeiro estudo bibliométrico foi realizado em 1917. *Cole e Eales*, em 1917, efetuaram uma análise estatística de publicações sobre Anatomia Comparada, entre 1550 e 1860. Em 1923, *Edward Hulme* apresentou uma análise estatística da História da Ciência, tendo introduzido o termo *statistical bibliography* (bibliografia estatística) (ARAÚJO, 2006; VANTI, 2002). Segundo *Guedes e Borschiver* (2005), o termo *statistical bibliography* ficou esquecido por 22 anos, voltando a ser utilizado em 1944, por *Gosnell* (1944). Ressurge, em 1962, empregado por *L. M. Raisig*. Em 1927, *Gross e Gross*, com o objetivo de fornecer subsídios para determinar quais periódicos deveriam ser adquiridos, fizeram uso da técnica de contagem das referências citadas no periódico *Journal of the American Chemical Society* (ARAÚJO; ALVARENGA, 2011). Em 1934, *Paul Otlet* foi o primeiro autor a apresentar o termo “bibliometria”, em sua obra *Traité de Documentation* (ARAÚJO, 2006).

A disseminação do termo bibliometria ocorre, contudo, a partir de 1969, com a publicação de um artigo de *Allan Pritchard* sobre a questão: “bibliografia estatística ou bibliometria?” (ARAÚJO, 2006). Nesse artigo, a bibliometria é caracterizada como área de estudo que usa métodos matemáticos e estatísticos para investigar e quantificar os processos de comunicação escrita. Assim, a bibliometria não se restringe à avaliação da produção científica, nem foca seus esforços unicamente na *web* (distinguindo-se, assim, da webometria).

Segundo *Figueiredo*, que é citado por *Araújo* (2006, p. 13),

a bibliometria desde sua origem é marcada por uma dupla preocupação: a análise da produção científica e a busca de benefícios práticos imediatos para bibliotecas (desenvolvimento de coleções, gestão de serviços bibliotecários) (FIGUEIREDO, 1977 apud ARAÚJO, 2006, p. 13).



Silva, Hayashi e Hayashi (2011, p. 114) esclarecem que, “a partir da década de 1990, a Bibliometria se tornou uma ferramenta padrão da política científica”. Tendo se tornado uma valiosa forma de medir a comunicação científica, o que podemos observar é que, desde então, reforçou-se o papel do profissional de Ciência da Informação e Biblioteconomia no planejamento e na gestão de políticas científicas.

A análise de citação e a de cocitação são duas técnicas da análise bibliométrica. Na subseção “Softwares bibliométricos”, você conhecerá uma lista de programas utilizados nas diversas “metrias”.

A bibliometria faz análises sobre aspectos quantitativos de autores, grupos de pesquisa, área de conhecimento e outros indicadores da produção científica.

As três principais leis da bibliometria são:

- c) Lei de *Lotka* (que foca na produtividade de autores);
- d) Lei de *Zipf* (Lei do mínimo esforço, que analisa a frequência de palavras);
- e) Lei de *Bradford* (ou Lei de dispersão, que estuda a produtividade de periódicos).

Em seção posterior, estudaremos em detalhes cada uma dessas leis bibliométricas.



Multimídia

No endereço <<http://www.feg.unesp.br/~fmarins/seminarios/Bibliometria/>>, você encontrará mais informações sobre bibliometria, com *links* para outros artigos e textos sobre o tema. Vale a pena a visita!

3.3.3 Cientometria

A Cientometria (em inglês, *scientometrics*) quantifica a produção científica (artigos, revistas e livros). Ela avalia, quantitativamente, a produção do conhecimento científico de pesquisadores, núcleos de pesquisa, universidades, instituições e até mesmo países. O planejamento e a avaliação das políticas científicas, as decisões sobre investimentos financeiros em determinadas áreas do conhecimento nas universidades e as verbas para pesquisadores individuais são, muitas vezes, baseados em dados cientométricos. Observe que, enquanto a bibliometria estuda questões mais gerais da informação, a cientometria foca na produção científica, mas claramente há uma intersecção entre elas.

De acordo com *Vanti* (2002), o termo cientometria – ou cientometria – foi criado na antiga União Soviética (URSS) e Europa Oriental. O autor explica que o termo alcançou notoriedade, no final da década de 1970, com a publicação da revista *Scientometrics*, editada originalmente na Hungria e, atualmente, na Holanda.



Multimídia

Acesse o *link* para a revista *Scientometrics*: <<http://www.springer.com/computer/database+management+%26+information+retrieval/journal/11192>>. Acesso em: 15 jul. 2021.

Entre as variáveis estudadas na cientometria, *Santin* (2011) destaca a comunicação entre os cientistas e os fatores de diferenciação das subdisciplinas. *Macias-Chapula* (1998) e *Spinak* (1998) apresentam, como objetos de estudo da cientometria,

- a) as várias áreas da ciência que evoluem em termos quantitativos;
- b) o estudo de como as várias subáreas científicas evoluem quantitativamente;
- c) as relações entre ciência e tecnologia;
- d) as relações entre o desenvolvimento científico e o crescimento econômico.

Na subseção sobre “Indicadores bibliométricos e cientométricos”, você conhecerá os principais indicadores utilizados nessas “metrias”.

Com a fundação, por *Eugene Garfield*, do *Institute for Scientific Information (ISI)*, em 1958, foi criado um dos principais centros para a quantificação e geração de insumos na área da produção científica.

3.3.4 Infometria

A infometria (em inglês, *informetrics*) ou informetria, para alguns autores, como área de estudo, surge após a constituição da bibliometria e da cientometria. É considerada a “metria” mais abrangente. Segundo *Vanti* (2002), *Otto Nacke* foi o primeiro a apresentar o termo infometria, em 1979. Ela não se restringe a estudar a produção científica e, assim, extrapola os limites da cientometria, possuindo, com ela, uma intersecção. Por outro lado, a infometria é o estudo dos aspectos quantitativos da informação em qualquer formato, e não apenas em registros catalográficos ou bibliografias, tampouco é restrita aos cientistas (VANTI, 2002).

Convém destacar que a infometria

[...] utiliza-se de unidades definidas, tais como palavras, documentos, textos, fontes ou bases de dados, como focos de análise, podendo priorizar variáveis como recuperação, a relevância, a revocação ou outras características da informação que possam ser consideradas relevantes. (BUFREM; PRATES, 2005, p. 11).

Wormell (1998), em um importante artigo, destaca as novas perspectivas que a análise infométrica apresenta, seja para investigação das bases de dados, seja como instrumento de análise. Nele, a autora relata três experiências do *Centro de Estudos Infométricos (CIS)*, criado em

1997, e que faz parte da *Real Escola de Biblioteconomia e Informação*, em inglês, *Royal School of Library and Information Studies (RSLIS)*, situada em Copenhague e em Aalborg, na Dinamarca. O *CIS* aplica os estudos bibliométricos, além da avaliação da produção em ciência e tecnologia, nos estudos do impacto sobre as relações sociais e econômicas. Dessa forma, a infometria estende sua aplicação para além de ferramenta unicamente quantitativa.

Tague-Sutckiffe, citado por *Vanti* (2002), destaca, como aspecto adicional de distinção entre infometria, bibliometria e cientometria, a relação delas com os objetos e sujeitos de estudo. A infometria não tem como escopo apenas a informação registrada. Estuda, também, fenômenos de comunicação informal (comunicação oral, inclusive) e de grupos sociais diversos (não apenas aqueles pertencentes às elites intelectuais).



Atenção

Segundo *Araújo* (2006), os principais grupos de pesquisa em bibliometria são:

- a) *BIRG* (Sidney, Austrália);
- b) *CEST* (Berna, Suíça);
- c) *CINDOC* (Madrid, Espanha);
- d) *CIS* (Copenhague, Dinamarca);
- e) *CRRM* (Marselha, França);
- f) *CWTS* (Leiden, Holanda);
- g) *FhG-ISI* (Karlsruhe, Alemanha);
- h) *Inforsk* (Umea, Suécia);
- i) *OST* (Montreal, Canadá);
- j) *OST* (Paris, França);
- k) *SPRU* (Sussex, Inglaterra);
- l) *ISI Research Service Group* (Filadélfia, EUA).



Atenção

Para uma visão mais completa da evolução histórica das “metrias”, sugerimos que você consulte a linha do tempo dos principais eventos na história dos estudos métricos da informação, em *Lucas, Zorita e Casado* (2013).

3.3.5 Webometria

Esta “metria” distingue-se por utilizar as técnicas da bibliometria na mensuração de processos de comunicação em *sites* na *web*. A partir dessas técnicas, é possível avaliar a evolução de uma ou mais áreas do conhecimento ao longo do tempo, efetuar a análise da produção de indivíduos, universidades ou instituições, regiões geográficas ou mesmo países. A webometria estuda a *web* para avaliar as quantidades de padrões de acessos, *hiperlinks* e *sites*.

A webometria (*webmetrics*, em inglês), conforme definição apresentada por *Almind e Ingwersen* (1997, p. 404), é a aplicação de métodos infométricos à WWW. *Vanti* (2002, p. 157) ensina:

Entre as medições que podem ser realizadas no campo da webometria, encontra-se, por exemplo, aquela que diz respeito à frequência de distribuição das páginas no ciberespaço. Esta medição aponta para o estudo ou análise comparativa da presença dos diversos países na rede, das proporções de páginas pessoais, comerciais e institucionais.

A webometria produz “mapas de fluxos de informação” baseados na internet e não em outros suportes eletrônicos ou em papel, como aqueles utilizados pela cientometria. As análises webométricas também podem ser feitas por domínio (por país ou por instituições).



3.3.6 Atividade

Esta atividade atende aos objetivos *a* e *b*: “descrever o objeto de estudo, as variáveis, os métodos e os objetivos da bibliometria, infometria, cientometria e webometria” e “definir indicadores bibliométricos, reconhecendo algumas de suas limitações”.

Para desenvolvê-la, consulte *Vanti* (2002), e verá que a webometria é tratada em seção específica. A partir dessa consulta, faça o que se pede:

- a) apresente três exemplos de mensurações que podem ser utilizadas na webometria;
- b) demonstre quais inconsistências lógicas são apresentadas pelos motores de busca, de tal forma que tornam mais complexa a tarefa de mensuração dos dados obtidos por essa via.

Resposta comentada

- a) Apresentamos, a seguir, algumas das mensurações que podem ser realizadas no campo da webometria, elencadas em *Vanti* (2002):
 - frequência de distribuição de páginas no ciberespaço;
 - classificações sobre o tipo de páginas (setor público x setor privado, por exemplo);
 - efetuar a medição dos mesmos elementos em momentos diferentes no tempo;



- pode-se mensurar o tamanho médio de uma página (expressa em *bytes*) ou em relação ao número médio de *links* por página;
- um indicador relevante na área da webometria é o fator de impacto na *web* (*web impact factor – WIF*).

b) Um problema bastante relevante é a reprodutibilidade das mensurações. Na página 159 do referido artigo, de acordo com *Judit Bar-Ilan* e *Thelwall*, destacamos:

Os motores costumam perder informação: URLs relevantes recuperados em um determinado momento por um determinado motor de busca não são encontrados por esse mesmo motor em um momento posterior (apesar de continuarem existindo e sendo relevantes) (BAR-ILAN; THELWALL apud VANTI, 2002, p. 159).

Outro ponto destacado no artigo diz respeito ao fato de que URL recuperados em um levantamento posterior nem sempre trazem informação idêntica à da primeira pesquisa, inclusive tendo baixa sobreposição de resultados entre os vários buscadores. Além da questão das inconsistências no que tange à contabilização, existe o importante problema do processamento das sintaxes de busca. Os motores de busca podem trazer resultados espúrios, além do fato de que a hierarquização dos documentos é função de análises subjetivas.

3.3.7 Cibermetria

Do inglês, *cibermetrics*. Como já mencionamos anteriormente, enquanto alguns autores utilizam cibermetria e webometria como sinônimos, outros, como *Bjorneborn* e *Ingwesen*, citados por *Lucas, Zorita* e *Casado* (2013, p. 260), definem cibermetria como o “estudo dos aspectos quantitativos da construção e dos usos dos recursos de informação, estruturas e tecnologias em internet, de acordo com as aproximações bibliométricas e informétricas”.

Em *Bufrem* e *Prates* (2005), você encontrará um quadro que apresenta a diversidade de conceituações para os termos bibliometria, biblioteconometria, cientometria, infometria e webometria, com as respectivas referências.



3.3.7 Atividade

Esta atividade atende ao objetivo *a* da Unidade, a saber: “descrever o objeto de estudo, as variáveis, os métodos e os objetivos da bibliometria, infometria, cientometria e webometria”.

Vanti (2002) apresenta possibilidades de aplicação das técnicas bibliométricas, cientométricas e infométricas. Apresente cinco aplicações, dentre as citadas no artigo.

Resposta comentada

Certamente, é relevante conhecer as definições teóricas das várias “metrias”, seus alcances e limites. Atingimos melhor compreensão desses métodos conhecendo suas possibilidades de aplicação. A seguir, citamos as aplicações apresentadas em *Vanti* (2002, p. 155):

- identificar as tendências e o crescimento do conhecimento em uma área;
- identificar as revistas do núcleo de uma disciplina;
- prever tendências de publicação;
- medir o grau e os padrões de colaboração entre autores;
- analisar os processos de citação e cocitação.

O referido artigo ainda apresenta outras oito possibilidades de aplicação.



3.3.8 Atividade

Esta atividade atende ao objetivo a: “descrever o objeto de estudo, as variáveis, os métodos e os objetivos da bibliometria, infometria, cientometria e webometria”.

Estudos de usuários e de seu comportamento frente aos serviços ofertados devem ser realizados, seja na fase de planejamento de um novo serviço, seja após sua implantação. A bibliometria proporciona ferramentas para análises quantitativas de grupos de pesquisa, autores e outros indicadores de produção científica.

Com isso em mente, faça o que se pede:

- a) apresente a caracterização da bibliometria segundo *Allan Pritchard*;
- b) *Figueiredo* (1977) destaca que a bibliometria, desde sua origem, apresenta dupla preocupação. Apresente quais as duas preocupações explicitadas pela autora;
- c) A respeito da segunda preocupação explicitada por *Figueiredo* (1977) no item *b*, dê exemplos específicos de aplicações “práticas” da bibliometria para bibliotecas.

Resposta comentada

- a) De acordo com *Pritchard*, a bibliometria é caracterizada como uma área de estudo que usa métodos matemáticos e estatísticos para investigar e quantificar os processos de comunicação escrita.
- b) Segundo *Figueiredo*, citada por *Araújo* (2006, p. 13),

a bibliometria desde sua origem é marcada por uma dupla preocupação: a análise da produção científica e a busca de benefícios práticos imediatos para



bibliotecas (desenvolvimento de coleções, gestão de serviços bibliotecários (FIGUEIREDO, 1977 apud ARAÚJO, 2006, p. 13).

- c) Como aplicações práticas voltadas a bibliotecas, citamos: efetuar estimativas de crescimento dos acervos e efetuar avaliações quantitativas de suas características.

3.4 AS TRÊS LEIS DA BIBLIOMETRIA

Sem dúvida, as três principais leis bibliométricas são:

- a) Lei de *Bradford* (relativa à produtividade de periódicos);
- b) Lei de *Lotka* (voltada ao estudo da produtividade de autores);
- c) Lei de *Zipf* (associada aos estudos sobre frequência de palavras).

Cada uma dessas leis está associada a determinados “padrões” ou regularidades encontrados nos seus respectivos focos de estudo. Alguns desses padrões apresentam pequenas variações nos coeficientes ou expoentes numéricos, em função da área do conhecimento dos dados.

3.4.1 Lei de *Bradford* (Lei da dispersão)

A importância prática da Lei de *Bradford* consiste no seu uso para identificação de quais são os periódicos (ou outros tipos de publicação) a serem adquiridos (ou descartados) pelas bibliotecas. Os periódicos responsáveis pela produção de maior quantidade de artigos de determinada área do conhecimento são denominados “núcleo de periódicos”.

Os estudos de *Bradford* resultaram na publicação, em 1934, da sua Lei de dispersão. O conjunto de periódicos da área de Geofísica utilizados por ele foi constituído de um núcleo menor, com periódicos mais próximos ao assunto, e um outro conjunto maior, menos próximo do assunto. *Bradford* observou que, à medida que a quantidade de periódicos aumentava, a produtividade dos pesquisadores caía. Em sua análise, foram utilizados 326 periódicos. Desses, nove continham 429 artigos, 59 continham 499 artigos e 258 continham 404 artigos. A Lei de *Bradford* foi proposta nos seguintes termos, tal como citado por *Coutinho* (1988, p. 219):

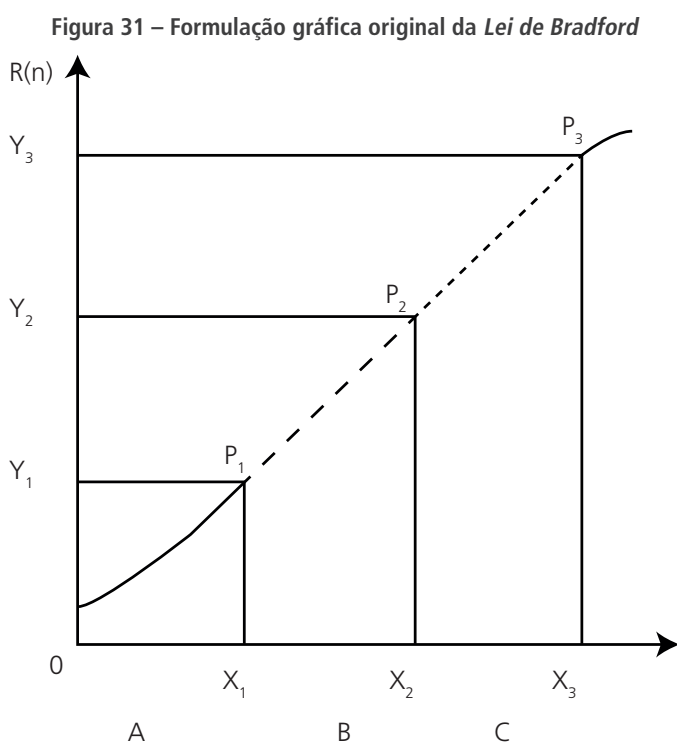
se os periódicos forem ordenados em ordem de produtividade decrescente de artigos sobre determinado assunto, poderão ser distribuídos num núcleo de periódicos mais particularmente dedicados a esse assunto e em diversos grupos ou zonas contendo o mesmo número de artigos que o núcleo, sempre que o número de periódicos e das zonas sucessivas for igual a $1:n:n^2$.

No seu trabalho original, *Bradford* encontrou três zonas de produtividade. Ele utilizou um gráfico semilogarítmico para apresentar graficamente sua lei. Esse gráfico apresenta uma leve curvatura quando nos aproximamos da origem, tornando-se, praticamente, uma reta, a partir do ponto denotado por P1.

Pinheiro (1983) alerta para uma das “armadilhas” da Lei de *Bradford*: o núcleo de periódicos determinado por ela não corresponde ao dos mais particularmente devotados ao assunto, mas aos periódicos mais produtivos em determinado período. *Coutinho* (1988) também alerta para fatores como a periodicidade e o número de artigos publicados por fascículo do periódico, que podem resultar em mudanças na posição de sua produtividade.

Brooks (1969), citado por *Lousada et al.* (2012), apresenta condições para a aplicação adequada da Lei de *Bradford*:

- a) o tema da bibliografia deve ser definido com clareza;
- b) todos os periódicos relevantes devem estar incluídos;
- c) o período do levantamento deve ser longo o bastante para permitir maior representatividade.



Log(n), n é o número de periódicos.

Fonte: *Pinheiro* (1983).

R(n) é o número acumulado de artigos sobre um assunto e n é o número acumulado de periódicos contendo aqueles artigos.

Na Tabela 8, a seguir, extraída de *Aymard* (1980), temos os dados utilizados por *Bradford* em seu levantamento original e, na sequência, o gráfico com os valores $\log_{10}(n)$, no eixo das abscissas, e $\sum_{i=1}^m (R_i \cdot A_i)_{\text{observado}}$, no eixo das ordenadas.

Tabela 8 – Dados utilizados por *Bradford* em seu levantamento

R_i	A_i	$n = \sum_{i=1}^{i=m} R_i$	$R_i \cdot A_i$	$\text{Log}_{10}(n)$	$\sum_{i=1}^{i=m} (R_i \cdot A_i)$ observado
1	22	1	22	0	22
1	18	2	18	0,30103	40
1	15	3	15	0,477121	55
2	13	5	26	0,69897	81
2	10	7	20	0,845098	101
1	9	8	9	0,90309	110
3	8	11	24	1,041393	134
3	7	14	21	1,146128	155
1	6	15	6	1,176091	161
7	5	22	35	1,342423	196
2	4	24	8	1,380211	204
13	3	37	39	1,568202	243
25	2	62	50	1,792392	293
102	1	164	102	2,214844	395

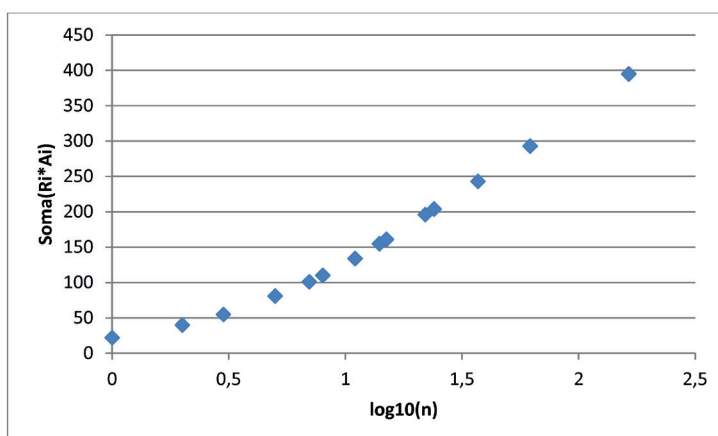
Fonte: Aymard (1980).

R_i é a quantidade de periódicos que produziram uma quantidade correspondente de artigos.

A_i é a quantidade correspondente de artigos durante o período da pesquisa.

Na Figura 31, temos, no eixo x, o logaritmo na base 10 do número de periódicos ($\log_{10}(n)$). No eixo y, temos a soma do produto associada. Já na Figura 32, adiante, pode ser observada a tendência linear à medida que o número de periódicos aumenta.

Figura 32 – Lei de *Bradford*



Fonte: Aymard (1980).

Para mais detalhes matemáticos sobre a aplicação da Lei de *Bradford*, consultar *Aymard* (1980) e *Nunez* (2014, p. 61).



3.4.2 Atividade

Esta atividade atende ao objetivo c: “descrever as Leis de *Bradford* e *Lotka*”.

Atividade: a Lei de *Bradford* é uma das leis bibliométricas. A respeito dela, faça o que se pede:

- a) descreva a Lei de *Bradford*;
- b) *Pinheiro* (1983) destaca possíveis “armadilhas” da Lei de *Bradford*. Por sua vez, *Brooks* (1969) destaca condições para a adequada aplicação da Lei de *Bradford*. Quais são essas condições?

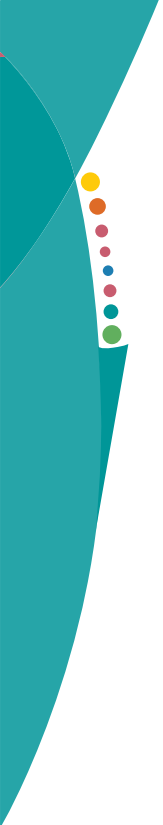
Resposta comentada

- a) A Lei de *Bradford* foi formulada em 1934. Em seu artigo original, *Bradford* utilizou artigos da área de Geofísica. Ele apresentou três zonas de produtividade, cada uma com uma porcentagem do número total de artigos relacionados com o assunto em investigação. *Bradford* identificou que 50% ou mais dos artigos constam de um número bastante pequeno de periódicos.
- b) De acordo com *Brooks* (1969), as condições a seguir devem estar satisfeitas para que a Lei de *Bradford* seja adequadamente aplicada:
 - a definição do tema da bibliografia deve ser efetuada com clareza;
 - não devem ser excluídos da aplicação os periódicos relevantes da área em estudo;
 - o levantamento deve ser desenvolvido por período suficientemente longo para assegurar representatividade.

3.4.3 Leis de *Zipf* (Lei do mínimo esforço)

A Lei de *Zipf* é associada aos estudos sobre frequência de palavras-chave, sendo utilizada para efetuar estimativas de frequências de palavras em um texto técnico. De acordo com a Lei de *Zipf*, um grupo relativamente pequeno de palavras apresenta frequência muito maior que um grande número de palavras.





Foi *George Zipf*, linguista da *Universidade de Harvard*, que, em 1935, pela primeira vez destacou que as palavras apresentam determinada frequência nos textos. *Zipf* observou que há uma relação entre a frequência com que uma palavra aparece em um texto suficientemente longo e sua posição (ordem de série, ou *rank*, em inglês) na lista de palavras ordenadas de acordo com a frequência com que as palavras aparecem no texto.

A primeira Lei de *Zipf* afirma que, se ordenarmos a frequência com que as palavras aparecem nos textos e multiplicarmos o número de ordem pela frequência, obteremos um valor constante:

$$r.f = c,$$

onde *r* é a ordem de série da palavra, *f* é a frequência e *c*, uma constante.

Como um exemplo simples, extraímos de *Maia* (1973, p. 99-100) o texto a seguir:

A palavra *informação* é usada, na linguagem cotidiana, de diferentes maneiras. Falamos de informação "útil", "valiosa", "real", "fidedigna", "precisa", "verdadeira", "relevante"; mas nenhuma dessas expressões ocorre na teoria matemática da comunicação, que descreve a informação tão somente como uma probabilidade estatística de sinais provindos de uma determinada fonte. Apesar de todos sabermos o que é informação, não há maneira de defini-la, exceto através de suas características e propriedades, estudadas pela Ciência da Informação. Dentre as propriedades, a representação da informação é um dos campos mais investigados, em consequência de uma premissa antiga, na qual se identifica o bom funcionamento de um sistema de informação com a representação adequada desta informação – e isto acontece frequentemente. Assim, os diferentes tipos de representação se refletem nas bibliotecas, através do sistema de classificação adotado, dos cabeçalhos de assunto e da catalogação; nos sistemas de recuperação da informação, através das linguagens de indexação; nos sistemas que utilizam o computador, através das linguagens de programação e sua codificação; nos sistemas de transmissão de sinais, através de sua codificação.

Contando as palavras diferentes e ordenando-as, construímos a Tabela 9, a seguir:

Tabela 9 – Exemplo para a Lei de Zipf

	F	R	R * F = C
de	19	1	19
maneiras	11	2	22
informação	8	3	24
através	5	4	20
da	5	5	25
e	4	6	24
a	3	7	21
é	3	7	21
na	3	7	21
nos	3	7	21
o	3	7	21
que	3	7	21
sistemas	3	7	21
uma	3	7	21
codificação	2	8	16
das	2	8	16
diferentes	2	8	16
dos	2	8	16
linguagens	2	8	16

Fonte: Produção do próprio autor (2017).

A frequência de ocorrência de uma palavra, como função da sua posição n em uma lista ordenada, é dada pela expressão $f = \frac{b}{n^a}$, onde b e a são constantes.

Zipf apresentou uma segunda lei (aplicada às palavras de baixa frequência). Em 1967, *Booth* apresentou uma revisão e modificação para esta segunda Lei de *Zipf* (GUEDES; BORSCHIVER, 2005). Por isso, encontramos, na literatura especializada, referências à segunda Lei de *Zipf* como Lei de *Zipf-Booth*.

A segunda Lei de *Zipf* propõe que, para as palavras de baixa frequência em um texto, várias delas apresentarão a mesma frequência. A apresentação desta segunda lei em símbolos matemáticos por *Booth* resultou na expressão:

$$\frac{l_1}{l_n} = \frac{n(n+1)}{2}, \text{ em que:}$$

l_1 representa a quantidade de palavras com frequência igual a 1;

l_n representa a quantidade de palavras com frequência igual a n ;

2 é a constante válida para a língua inglesa.



Ora, se a constante da segunda Lei de Zipf para a língua inglesa é 2, qual seria essa constante na língua portuguesa? Maia (1973) apresenta pesquisa realizada com três textos em língua portuguesa para buscar essa resposta. Ela determinou o valor 1,5 para a segunda Lei de Zipf em língua portuguesa.

Guedes e Borschiver (2005) explicam que Goffman observou que a primeira Lei de Zipf era válida apenas para a região de palavras de alta frequência de ocorrência. Por outro lado, a segunda Lei de Zipf descreve o comportamento de palavras de baixa frequência de ocorrência. Parece razoável imaginar que existe uma região crítica na qual ocorre a mudança no comportamento das palavras de alta frequência para a região de baixa frequência. O número de palavras com frequência "n" que demarca a região de transição é denominado ponto de transição T de Goffman. Para o autor (GOFFMAN, 1966), nessa região de transição estão as palavras com elevado conteúdo semântico (são os termos de indexação e as palavras-chave). A quantidade de palavras com alta frequência apresentaria $l_n = 1$. Temos a equação:

$$\frac{l_1}{1} = \frac{n(n+1)}{2}$$

Resolvendo a equação de 2.º grau: $n^2 + n - 2l_1 = 0$, teremos a raiz positiva

$$n = \frac{-1 + \sqrt{1 + 8l_1}}{2}$$

Denomina-se esse valor "n" de ponto de transição T de Goffman. Esse valor identifica o ponto de transição das palavras de baixa frequência para as de alta frequência (GUEDES; BORSCHIVER, 2005). A importância das Leis de Zipf está na construção de processos automáticos de indexação. Deve-se destacar, além da questão da aplicação, a importância específica das Leis de Zipf: elas se relacionam a estudos sobre a representação da informação.



Atenção

Existem outros conceitos e linhas de investigação na Bibliometria, além dos que foram mencionados aqui, mas está fora do escopo deste texto apresentá-los. Uma sugestão para você se aprofundar nesse tema é a consulta ao artigo de Guedes e Borschiver (2005). Nele, são abordados: as frentes de pesquisa, colégios invisíveis, acoplamento bibliográfico e cocitação, teoria epidêmica de Goffman, obsolescência e vida média da literatura.

Outra sugestão é o artigo de Araújo (2006) que discute a obsolescência da literatura e sua vida média. O propósito no desenvolvimento das referidas leis foi descrever a queda da validade ou utilidade de informações com o decorrer do tempo. Assim, com os estudos realizados a partir delas, disciplinas como Matemática, Geologia e Botânica apresentam

[...] forte componente de literatura clássica, enquanto outras são quase exclusivamente compostas por literatura efêmera (Física, Engenharia), sendo algumas de caráter intermediário (Fisiologia, Química) (ARAÚJO, 2006, p. 21).

3.4.4 Lei de Lotka

Um estudo sobre a produtividade de autores foi o ponto de partida para a Lei de Lotka, formulada em 1926. Lotka estava interessado em estudar a produtividade científica dos pesquisadores e estudou a produção científica de químicos e físicos (com publicações no *Chemical Abstracts* entre 1909 e 1916 e no *Auerbach's Geschichtstafeln der Physik*, a partir do início da publicação deste periódico, em 1900). Segundo a relação apresentada por Lotka, pequena quantidade de autores é responsável por uma porcentagem bastante grande da produção científica. A relevância do resultado de Lotka está no fato de ele observar que existe uma distribuição desigual na produção científica.

A Lei de Lotka é uma lei de potência: se x representa o número de trabalhos publicados, c é uma constante que varia de acordo com as áreas do conhecimento científico e y é o número de autores que publicaram a quantidade x de artigos então,

$$y = \frac{C}{x^2}.$$

Em termos matemáticos, dizemos que a Lei de Lotka segue a lei do inverso do quadrado.

Talvez você não se lembre do que vem a ser uma lei que segue o padrão "inverso do quadrado". É quando uma variável ("y", no caso anterior) apresenta associação com uma constante dividida por uma variável ("x", no caso anterior) elevada ao quadrado. Só isso. Ok?

A descoberta de Lotka significa que um número pequeno de autores produz grande parte dos artigos científicos e uma grande quantidade de autores publica relativamente poucos artigos. A importância da Lei de Lotka, segundo Guedes e Borschiver (2005), está na avaliação da produtividade de pesquisadores, na identificação dos centros de pesquisa mais desenvolvidos, em dada área de assunto, e no reconhecimento da "solidez" de uma área científica.

Segundo a Lei de Lotka, para determinado período com um número "n" de artigos, a quantidade de pesquisadores que são autores de dois deles é igual a $\frac{1}{4}$ da quantidade de pesquisadores que publicaram um artigo. A quantidade de pesquisadores autores de três artigos é igual a $\frac{1}{9}$ dos pesquisadores que escreveram um artigo.

Desde a publicação do trabalho de Lotka, muitos pesquisadores publicaram refinamentos na lei que leva seu nome, enquanto outros apontaram questionamentos a ela. Um deles é a Lei do elitismo, de Price.



3.4.4.1 Lei do elitismo, de Price

Segundo Araújo (2006, p. 14), “o número de membros da elite corresponde à raiz quadrada do número total de autores e a metade do total da produção é considerada o critério para se saber se a elite é produtiva ou não”. Em símbolos matemáticos,

$$N_{elite} = \sqrt{N_{autores}}.$$

Observe que uma contagem que leva em consideração os coautores dos trabalhos pode acabar resultando em um valor para o expoente distinto da contagem que exclui os coautores. Uma crítica frequente à Lei de Lotka é a ausência, em muitos estudos, de testes estatísticos para provar a hipótese assumida e poder validar o ajuste, ou não, a uma distribuição do tipo Lotka. A esse respeito, consultar o trabalho de Sobrino, Caldes e Guerrero (2008).

3.5 ÍNDICES E INDICADORES

Silva (2001) separa os indicadores bibliométricos em duas categorias: (1) indicadores quantitativos da atividade científica, mensurados pelo número de publicações e (2) indicadores de impacto (consideram o número de citações realizadas sobre cada artigo ou produção científica). Os indicadores de produção científica indicam quais instituições, departamentos, países e pesquisadores possuem maior produção. Contudo, de acordo com o autor, esses indicadores não conseguem apresentar uma estimativa da importância (ou impacto) do trabalho científico produzido. Dois indicadores foram desenvolvidos para aferir esse aspecto: o número de citações de um artigo e o fator de impacto do periódico no qual ele foi publicado. Deve-se, adicionalmente, lembrar que a qualidade de um trabalho científico não é a mesma coisa que o seu impacto.

3.5.1 Fator de impacto

O fator de impacto (FI) calcula a relação entre o número de citações, nos dois anos anteriores à publicação, e o número de artigos publicados nesses dois anos. Em símbolos (LOPES et al., 2012):

$$FatorImpacto_{ano} = \frac{citações_{ano-2} + citações_{ano-1}}{artigos_{ano-2} + artigos_{ano-1}}.$$

Lopes et al. (2012) destacam outro ponto de análise com o FI: diferentes áreas científicas possuem taxas distintas para a chamada vida média de citação (mediana da antiguidade das referências), bem como diferentes taxas de citação.

O FI pode ser utilizado para duas funções:

- a) analisar o impacto de um periódico, para fazer avaliação de coleção e aquisição de um novo título;
- b) avaliar o impacto dos pesquisadores a partir do número de citações aos seus trabalhos.

Atualmente, já não é mais importante apenas publicar; os autores devem ser citados. Nessa questão, o FI é uma ferramenta de muito valor.

Como alerta *Garfield* (2004, p. 4),

o termo fator de impacto evoluiu gradativamente, especialmente na Europa, para descrever tanto o impacto do periódico quanto do autor. Essa ambiguidade frequentemente causa problemas. Uma coisa é usar fatores de impacto para comparar periódicos, e outra bem diferente é usá-los para comparar autores.

Observe que *Garfield* (2004) alerta para o fato de que é bem diferente utilizar fatores de impacto para comparar periódicos e autores. Os FIs para periódicos são calculados a partir de uma “população” razoavelmente grande de artigos e citações. Por outro lado, quando observamos um autor individual, ele produzirá, em geral, números menores.

Garfield também alerta para uma crença equivocada: o tamanho da comunidade que um periódico atende alteraria, de forma significativa, o seu FI. De fato, o elemento determinante no impacto não é nem o número de autores, nem o de artigos em uma área do conhecimento, mas a densidade de citações e a idade da literatura citada. Ou seja, o pesquisador bibliométrico, cientométrico ou webométrico deve estar consciente dos alcances, das limitações e especificidades (que variam de acordo com as áreas de conhecimento) de indicadores como o FI.

3.5.2 H-índice (*h-index*)

Em 2005, *J. Hirsch* propôs a quantificação do impacto e da produtividade dos pesquisadores considerando os artigos mais citados. Suponha que um pesquisador possua h-índice igual a 10. Então, de todos os artigos publicados por esse pesquisador, 10 deles possuem, no mínimo, 10 citações.

O h-índice também pode ser aplicado a países, grupos de pesquisa e periódicos. *Lopes et al.* (2012) alertam para o fato de que, utilizando o h-índice para avaliar investigadores em início de carreira, com baixa produção e/ou com uma estratégia de publicação mais seletiva, podemos incorrer em alguns erros. Os autores explicam que, entre as críticas ao *h-index*, estão a longevidade da carreira do investigador, sua estratégia de publicação e as diferenças existentes entre as áreas científicas (*LOPES et al.*, 2012).

3.5.3 Fator de impacto na web

Segundo *Ingwersen* apud *Vanti* (2002), o FI na web (ou *Web Impact Factor – WIF*) é dado pela expressão adiante:

$$\text{Fator Impacto Web} = \frac{n.^{\circ} \text{ da página que linka determinado site}}{n.^{\circ} \text{ de páginas do site linkado}}$$



O propósito desse indicador é avaliar a atratividade e a importância que os vários *sites* apresentam na *web*. *Vanti* (2002) esclarece que os indicadores da *web* podem diminuir (e até mesmo deixar de existir), comportamento diferente do das citações científicas tradicionais.

Destacamos este trecho de *Rodrigues* (2002, p. 51), citando *Velho* (1989), além de *Meis* e *Leta* (1996):

Há uma série de questionamentos quanto aos critérios adotados pelo *ISI*, o principal deles é que sua abrangência não é total, ou seja, um significativo contingente de publicações é desconsiderado, sobretudo aquelas de países periféricos como os latino-americanos [...] (VELHO, 1989; MEIS; LETA, 1996 apud RODRIGUES 2002, p. 51).

3.5.4 Outros indicadores

Existem outros indicadores bibliométricos (que, por questões de espaço, não abordaremos neste texto), como apontam *Lopes et al.* (2012): o *Eigenfactor™ Metrics (EF)*, o *Article Influence™ Score (AI)* e o *SCImagoJournalRank (SJS)*.

O *Google Acadêmico*, ou *Google Scholar*, criou uma interessante ferramenta para medir o *h-index*, ou índice-h, dos autores. A vantagem é que todos os trabalhos do autor que receberam citações são contabilizados, diferentemente das bases de dados seletivas, que limitam as citações apenas às revistas indexadas. Considerando a seletividade de bases de dados como a *Web of Science* ou *Scopus*, que você vai estudar na disciplina de “Fontes de Informação”, o *Google Acadêmico* promete ser olhado com atenção pelos avaliadores da produção científica. A *Fundação de Amparo à Pesquisa (FAPESP)* solicita aos pesquisadores que submetam projetos ao *h-index*, de acordo com as medidas do *Google Scholar*.

Spinak (1998) apresenta dezenas de indicadores de publicação, indicadores de citação e técnicas de mapeamento. Já o trabalho de *Santos e Kobashi* (2009) apresenta técnicas tradicionais e inovadoras de visualização de informação: diagramas de dispersão, gráficos de setores, pictogramas, *tree-map*, dendogramas, grafos ou redes de relacionamentos e rede de relacionamentos dos descritores de registros de uma base de dados. Sugerimos fortemente que você, estudante, consulte esses artigos, para ter uma visão mais detalhada da “floresta” de indicadores bibliométricos existentes.

3.5.5 Limitações dos indicadores bibliométricos

Nem tudo são flores. Há inúmeras críticas aos indicadores bibliométricos:

- a) o número de publicações não é uma medida de qualidade da produção científica. Sugere-se a utilização de indicadores bibliométricos em conjunto com a avaliação por pares;
- b) a produtividade das áreas do conhecimento não é idêntica. Não é razoável supor que Humanidades, Engenharias, Artes Plásticas e Matemática possuam, todas, os mesmos valores numéricos para os diversos indicadores, ou as mesmas tradições de publicação.

Lopes et al. (2012) destacam que, geralmente, as áreas clínicas se caracterizam por menor produtividade, em relação às áreas básicas ou experimentais.

Indicadores de impacto possuem limitações tais como a existência de artigos (raros, mas existem) com muitas citações, que, contudo, são negativas. Por vezes, os autores apresentam referências em seus trabalhos que não foram diretamente utilizadas neles (LOPES et al., 2012). Também existe a barreira da língua: artigos publicados em inglês ou francês tendem a ter mais citações que aqueles publicados em coreano ou russo;

- c) é reconhecido que artigos de revisão (*review*) e artigos metodológicos apresentam número elevado de citações. Isso também é aceito para periódicos de revisão;
- d) as autocitações (quando um autor cita a si mesmo) podem distorcer os resultados;
- e) Lopes et al. (2012) apontam que o FI é calculado apenas com os dados registrados na *Web of Science*. Como a base de dados é extremamente seletiva, o número de periódicos indexados é reduzido, fazendo com que um grupo pequeno de artigos possa ser analisado pela ferramenta do FI. O *Scielo Citation Index* foi uma alternativa do *Scielo* para minimizar os efeitos de distorção do FI da *Web of Science*. Vamos falar sobre a *Web of Science* e *SciELO* logo adiante.

3.6 AS BASES DE DADOS UTILIZADAS EM PESQUISAS BIBLIOMÉTRICAS

Todos os indicadores bibliométricos apresentados seriam inúteis se não fosse possível conseguir os dados para estudo. Existem bases de dados que são utilizadas nos estudos bibliométricos. As mais usadas atualmente são a *Web of Science*, *Scopus* e, mais recentemente, o *Google Scholar Metrics*, de acordo com Lopes et al. (2012).

3.6.1 *Web of Science*

Em 1955, *Eugene Garfield* apresentou um índice de citação da ciência baseando-se no *Shepard's Citations*, que é um índice de citações de recursos legais utilizado nos Estados Unidos desde 1873. As causas jurídicas se relacionam umas com as outras por meio de uma codificação específica. O sistema então proposto por *Garfield* permite a recuperação de um artigo por meio de outros que o citam. A partir disso, *Garfield* criou o *Institute for Scientific Information (ISI)* e as bases de dados *Arts &*

Humanities Citation Index, Social Science Citation Index, Science Citation Index e Conference Proceedings Citation Index. Todas essas bases estão, na atualidade, na base de dados *Web of Science (Editora Thomson Scientific)*. A *Web of Science* indexa milhares de periódicos, nas diversas áreas da ciência, com dados que datam desde o início do século XX. O *Journal Citation Reports (JCR)* é considerado, por Lopes et al. (2012, p. 5), a “ferramenta líder de mercado dos *rankings* de periódicos”.



Multimídia

Sugerimos fortemente que você consulte o *SCImago*, cujo *link* é: <<http://www.scimagojr.com/index.php>>. Por favor, explore a grande diversidade de produção de *rankings* e mapas que o *site* oferece.

No gráfico *Shape of Science* do *SCImago*, você pode acessar indicadores bibliométricos dos periódicos da base de dados do *SJR SCImago Journal & Country Rank*. Nossa sugestão é a de que você consulte este *site* e o explore para observar a potencialidade de mapas e *rankings* que lá podem ser desenvolvidos.

Por exemplo, você poderá produzir um “mapa de bolhas” (*bubblemap*). As variáveis nos eixos x e y podem ser escolhidas entre:

- a) *documents*;
- b) *citable documents*;
- c) *cites*;
- d) *self cites*;
- e) *cites per doc*;
- f) *h-index*;
- g) *% cited documents*.

3.6.2 SciELO

O *Scientific Electronic Library Online (SciELO)* (<<http://www.scielo.org/php/index.php?lang=pt>>) oferece acesso a indicadores bibliométricos (de publicação, de coleção e de citação). No *link* <http://statbiblio.scielo.org/stat_biblio/index.php?lang=en>, você pode consultar dados de citação de periódicos (fator de impacto para dois anos, fator de impacto para três anos, meia-vida, citações recebidas e citações concedidas). Vale lembrar que o *SciELO* foi a primeira iniciativa de publicação de revistas em acesso aberto no mundo e o primeiro a lançar dados de citações dessas publicações.

3.7 SOFTWARES BIBLIOMÉTRICOS

Existe grande variedade de *softwares* para facilitar os trabalhos de pesquisa na bibliometria e cientometria. No Portal da Escrita Científica da *Universidade de São Paulo (USP)* da cidade de São Carlos, você encontrará uma relação de *softwares* bibliométricos e de colaboração científica. Veja a tela copiada do *site* na Figura 33:

Figura 33 – Tela com *softwares* bibliométricos da USP São Carlos



Fonte: *Escrita Científica* (c2012).⁵

Alguns desses *softwares* são de uso gratuito e outros, não. Outros *softwares* utilizados em pesquisas bibliométricas são: *Ucinet* e *Sitkis*. O *link* de acesso ao *Ucinet*, que é um *software* voltado à análise de redes sociais, é: <<https://sites.google.com/site/ucinetsoftware/home>>. Você pode efetuar o *download* (gratuito) do *Ucinet* e usá-lo por um período de 90 dias. Como outros *softwares*, existem tarifas especiais para estudantes. O *Sitkis* é uma ferramenta gratuita para análises bibliométricas e pode ser acessado no *link*: <<https://sites.google.com/site/sitkisbibliometricanalysis/>>. Você não terá dificuldades em encontrar tutoriais e ajuda na *web* para cada um desses *softwares*.



3.7.1 Atividade

Esta atividade atende ao objetivo a: “descrever o objeto de estudo, as variáveis, os métodos e objetivos da bibliometria, infometria, cientometria e webometria”.

Para realizar esta atividade, considere a Figura 30, Intersecção entre as “metrias”, e o Quadro 8, a seguir, ambos extraídos de *Vanti* (2002).

⁵ Disponível em: <<http://www.escritacientifica.sc.usp.br/gerenciadores-bibliometricos/>>.

Quadro 8 – Bibliometria, infometria, cientometria e webometria

Tipologia/ subcampo	Bibliometria	Cientometria	Infometria	Webometria
Objeto de estudo	Livros, documentos, revistas, artigos, autores, usuários	Disciplinas, assuntos, áreas e campos científicos e tecnológicos, patentes, dissertações e teses	Palavras, documentos, bases de dados, comunicações informais (inclusive em âmbitos não científicos), <i>homepages</i> na WWW	Sítios na WWW (URL, título, tipo, domínio, tamanho e <i>links</i>), motores de busca
Variáveis	Número de empréstimos (circulação) e de citações, frequência de extensão de frases	Fatores que diferenciam as subdisciplinas, como os cientistas se comunicam	Difere da cientometria no propósito das variáveis, por exemplo, medir a recuperação, a relevância, a revocação	Número de páginas por sítio, de <i>links</i> por sítio, de <i>links</i> que remetem a um mesmo sítio e de sítios recuperados
Métodos	<i>Ranking</i> , frequência, distribuição	Análise de conjuntos e de correspondência, co-ocorrência de termos, expressões, palavras-chave	Modelo vetor-espço, modelos booleanos de recuperação, modelos probabilísticos, linguagem de processamento, abordagens baseadas no conhecimento, tesouros	Fator de impacto da <i>web</i> (FIW), densidade dos <i>links</i> , "situações", estratégias de busca
Objetivos	Alocar recursos: pessoas, tempo, dinheiro etc.	Identificar domínios de interesse. Onde os assuntos estão concentrados. Compreender como e quanto os cientistas se comunicam	Melhorar a eficiência da recuperação da informação, identificar estruturas e relações dentro de diversos sistemas de informação	Avaliar o sucesso de determinados sítios, detectar a presença de países, instituições e pesquisadores na rede e melhorar a eficiência dos motores de busca na recuperação das informações

Fonte: Vanti (2002).

- Apresente a distinção dos objetos de estudo da infometria da cientometria;
- identifique, de acordo com a Figura 30, qual das "metrias" pode ser considerada como a mais ampla e justifique sua resposta;
- de acordo com o Quadro 8, quais os principais métodos da bibliometria?;
- de acordo com a Figura 30, existem intersecções entre os campos da bibliometria e da webometria. Apresente exemplos nos quais temos essa intersecção;
- de acordo com a Figura 30, existem intersecções entre os campos da cientometria e da webometria. Apresente exemplos nos quais temos essa intersecção.

Resposta comentada

- a) Enquanto a cientometria tem como objeto de estudo as disciplinas, assuntos, áreas e campos científicos e tecnológicos, a infometria trata de palavras, documentos e bases de dados (mesmo que não científicos).
- b) A Figura 30 está de acordo com análises efetuadas por *Russel* apud *Vanti* (2002, p. 10) e *Tague-Sutckiffe* apud *Vanti* (2002, p. 10), segundo as quais a infometria compreende a bibliometria, a cientometria e a webometria. A infometria é o termo considerado “guarda-chuva”, abrangendo os conceitos de bibliometria, cientometria e webometria.
- c) De acordo com o Quadro 8, os principais métodos da bibliometria são *ranking*, frequência e distribuição.
- d) No caso da intersecção entre bibliometria e webometria, em particular no que diz respeito aos objetivos, um pesquisador pode estar interessado na alocação de recursos, pessoas, tempo e dinheiro, que se relaciona com o sucesso de determinados sítios.
- e) No caso da intersecção entre cientometria e webometria, em particular no que diz respeito às variáveis, um pesquisador pode estar interessado em estudar fatores que diferenciam as subdisciplinas e em como os cientistas se comunicam avaliando o número de *links* que remetem a um mesmo sítio.

3.8 CONCLUSÃO

Concluimos esta Unidade destacando que os métodos quantitativos são de grande importância, seja ela teórica (ao propor leis e padrões de dispersão), seja ela prática (ao possibilitar estimativas e previsões sobre acervos e coleções).

Contudo, nem todos os valores numéricos dos parâmetros das leis bibliométricas são idênticos entre as várias áreas ou línguas (por exemplo, lembremo-nos da diferença entre a constante da segunda Lei de *Zipf* para a língua inglesa e para outras línguas).

De forma análoga à discussão sobre os alcances e limitações que fizemos na Unidade anterior, com relação aos métodos quantitativos gerais, temos também a discussão sobre a limitação dos indicadores bibliométricos. Em particular, não podemos comparar a produtividade das diferentes áreas de conhecimento. Não podemos esperar que Artes Plásticas, Física Nuclear e Direito possuam, todas, os mesmos valores numéricos e de produtividade para os diferentes indicadores quantitativos. É possível encontrar artigos que apresentam citações que não foram, efetivamente, utilizadas nos trabalhos. Outro aspecto: é natural esperar que periódicos voltados a artigos de revisão (*review*) e artigos metodológicos apresentem elevado número de citações. Você deve se lembrar, ainda, de ressalvas quanto a eventuais inconsistências lógicas apresentadas pelos motores de buscas.



RESUMO

Nesta Unidade, você estudou alguns dos principais instrumentos quantitativos utilizados pelos bibliotecários e cientistas da informação para análise da produção científica e avaliação de coleções, bem como para efetuar mensurações no registro e difusão da informação e do conhecimento. Você também estudou as Leis de *Bradford*, *Zipf* e *Lotka*. Conheceu o ponto T de *Goffman* e a Lei do elitismo, de *Price*.

Você estudou as diversas “metrias” da Biblioteconomia: a bibliometria, a cientometria, a infometria e a mais recente delas, a webometria, abordando suas semelhanças e diferenças com relação a objetos de estudo, objetivos, variáveis e métodos em cada uma delas.

Entrou em contato com os principais índices e indicadores utilizados na Biblioteconomia: fator de impacto, h-índice e fator de impacto na *web*.

Também viu o quão fundamentais à Biblioteconomia e à Ciência da Informação são a *Web of Science*, o *Scopus*, *SCImago*, *Scielo* e o *Google Acadêmico*.



Sugestão de Leitura

TAGUE-SUCKFFE, J. An introduction to informetrics. **Information Processing & Management**, [S.l.], v. 28, n. 1, p. 1-3, 1992.

REFERÊNCIAS

ALMIND, T. C.; INGWEERSEN, P. Informetric analyses on the World Wide Web: methodological approaches to “webometrics”. **Journal of Documentation**, [S.l.], v. 53, n. 4, p. 404-426, 1997.

ARAÚJO, C. R. Bibliometria: evolução histórica e questões atuais. **Em Questão**, Florianópolis, v. 12, n. 1, p. 11-32, 2006.

ARAÚJO, R. F.; ALVARENGA, L. A. Bibliometria na pesquisa científica da pós-graduação brasileira de 1987 a 2007. **Encontros Bibli**, Florianópolis, v. 16, n. 31, p. 51-70, 2011.

AYMARD, M. A Lei de dispersão bibliográfica de Bradford. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 13, n. 3-4, p. 147-156, 1980.

BAR-LLAN, J. Informetrics at the beginning of the 21st century: a review. **Journal of Informetrics**, [S.l.], v. 2, p. 1-52, 2008.

BUFREM, L.; PRATES, Y. O saber científico registrado e as práticas de mensuração da informação. **Ciência da Informação**, Brasília, v. 34, n. 2, p. 9-25, 2005. Disponível em: <<http://www.scielo.br/pdf/ci/v34n2/28551.pdf>>. Acesso em: 21 fev. 2015.

CHAVIANO, O. G. Algunas consideraciones teórico-conceptuales sobre las disciplinas métricas. **Acimed**, Ciudad de La Habana, v. 12, n. 5, 2004. Disponível em: <http://bvs.sld.cu/revistas/aci/vol12_5_04/aci07504.htm>. Acesso em: 21 fev. 2015.

COUTINHO, E. As armadilhas da Lei de Bradford. **Revista de Biblioteconomia**, Brasília, v. 16, n. 2, p. 217-225, 1988. Disponível em: <<http://www.brapci.ufpr.br/documento.php?dd0=0000002566&dd1=2410e>>. Acesso em: 21 fev. 2015.

FIGUEIREDO, Nice. **Tópicos modernos em Bibliometria**. Brasília: Associação dos Bibliotecários do Distrito Federal, 1977.

GOFFMAN, W. Mathematical approach to the spread of scientific ideas: the history of mast cell research. **Nature**, [S.l.], v. 212, p. 449-452, 1966.

GOSNELL, C. Obsolescence of books in college library. **College & Research Libraries**, [S.l.], v. 5, p. 115-125, 1944.

GUEDES, V. L.; BORSCHIVER, S. Bibliometria: uma ferramenta estatística para a gestão da informação e do conhecimento, em sistemas de informação, de comunicação e de avaliação científica e tecnológica. In: ENCONTRO NACIONAL DE CIÊNCIA DA INFORMAÇÃO, 6., 2005, Salvador. **Anais...** Salvador: [s.n.], 2005.

LOPES, S. et al. A Bibliometria e a avaliação da produção indicadores e ferramentas. In: CONGRESSO NACIONAL DE BIBLIOTECÁRIOS, ARQUIVISTAS E DOCUMENTALISTAS, 11., 2012, Lisboa. **Actas...** Lisboa: [s.n.], 2012.

LOUSADA, M. et al. Produção científica sobre gestão do conhecimento e gestão da informação no âmbito da Ciência da Informação: uma aplicação da Lei de Bradford. **Anales de Documentación**, Murcia, v. 15, n. 2, 2012. Disponível em: <<http://eprints.rclis.org/18034/>>. Acesso em: 21 fev. 2015.

LUCAS, E. O.; ZORITA, J. C. G.; CASADO, E. S. Evolução histórica de investigação em informetria: ponto de vista espanhol. **Liinc em Revista**, Rio de Janeiro, v. 9, n. 1, p. 255-270, 2013. Disponível em: <<http://www.ibict.br/liinc>>. Acesso em: 21 dez. 2014.

MAIA, E. L. S. Comportamento bibliométrico da língua portuguesa, como veículo de representação da informação. **Ci. Inf.**, Rio de Janeiro, v. 2, n. 2, p. 99-138, 1973.

MEIS, L. J. **O perfil da ciência no Brasil**. Rio de Janeiro: Ed. UFRJ, 1996. p. 13-37.



NUNEZ, Z. A. G. **A produção científica brasileira em Medicina Tropical indexada nas bases de dados Web of Science e Scopus entre os anos de 2005 e 2012**. 2014. Dissertação (Mestrado em Comunicação e Informação) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2014. Disponível em: <<https://www.lume.ufrgs.br/bitstream/handle/10183/95387/000917848.pdf?sequence=1>>. Acesso em: 21 fev. 2015.

PINHEIRO, L. V. R. Lei de Bradford: uma reformulação conceitual. **Ciência da Informação**, Brasília, v. 12, n. 2, p. 59-80, 1983. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/1498/1116>>. Acesso em: 21 fev. 2015.

RODRIGUES, N. A. **Pesquisa agropecuária oficial do Rio Grande do Sul**: a temática da produção técnico-científica no período 1990/1998. 2002. Dissertação (Mestrado em Comunicação e Informação) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2002.

SANTIN, D. M. Avanços e perspectivas da infometria e dos indicadores multidimensionais na análise de fluxos da informação e estruturas do conhecimento. **Encontros Bibli**, Florianópolis, v. 16, n. 32, p. 107-122, 2011.

SANTOS, R. N. M.; KOBASHI, N. Y. Bibliometria, Cientometria, Infometria: conceitos e aplicações. **Pesquisa Brasileira em Ciência da Informação**, João Pessoa, v. 2, n. 1, p. 155-172, 2009. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/view/21/43>>. Acesso em: 21 fev. 2015.

SILVA, J. A.; BIANCHI, M. L. P. Cientométrica: a métrica da Ciência. **Paideia**, Ribeirão Preto, v. 11, n. 20, p. 5-10, 2001. Disponível em: <<http://www.scielo.br/pdf/paideia/v11n21/02.pdf>>. Acesso em: 21 fev. 2015.

SILVA, M. R.; HAYASHI, C. R.; HAYASHI, M. C. P. I. Análise bibliométrica e cientométrica: desafios para especialistas que atuam em campo. **InCID: Revista de Ciência, Informação e Documentação**, Ribeirão Preto, v. 2, n. 1, p. 110-129, 2011.

SOBRINO, M. I. M.; CALDES, A. I. P.; GUERREIRO, A. P. Lei de Lotka aplicada à produção científica da área de Ciência da Informação. **Brazilian Journal of Information Science**, Marília, v. 2, n. 1, p. 16-32, 2008. Disponível em: <<http://www.bjis.unesp.br>>. Acesso em: 21 fev. 2015.

SPINAK, E. Indicadores cientométricos. **Ciência da Informação**, Brasília, v. 27, n. 2, p. 141-148, 1998. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/349/1690>>. Acesso em: 21 fev. 2015.

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. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/171/150>>. Acesso em: 25 fev. 2015.

VELHO, L. Avaliação acadêmica: a hora e a vez do “baixo clero”. **Ciência e Cultura**, Campinas, v. 41, n. 10, p. 57-68, out. 1989.