

A BIBLIOMETRIA E A GESTÃO DA INFORMAÇÃO E DO CONHECIMENTO CIENTÍFICO E TECNOLÓGICO: uma revisão da literatura

RESUMO - Este estudo aborda o campo da Bibliometria salientando, sobretudo, a aplicação das leis e princípios bibliométricos na gestão da informação e do conhecimento científico e tecnológico. O objetivo principal é apresentar uma breve revisão da literatura sobre Bibliometria, distinguindo a utilização dos indicadores bibliométricos, como instrumentos matemáticos e estatísticos básicos de organização, gestão e difusão da informação e do conhecimento, como também elementos-chave para o desenvolvimento de sistemas mais precisos, econômicos e cientificamente constituídos. Inicialmente, é traçado um pequeno histórico da Bibliometria. Dando sequência, diferentes metrias como a Cientometria, a Informetria e a Webmetria são apresentadas, apontando a distinção conceitual entre esses domínios do conhecimento e seus campos de aplicação. A seguir, são citados autores que se destacaram no estudo e na enunciação das leis e princípios da Bibliometria, com especial ênfase às abordagens teóricas e descritivas, às leis e princípios bibliométricos, como ferramentas empíricas utilizadas na quantificação dos processos de comunicação científica. Finalmente, conclui-se evidenciando a importância da Bibliometria para a gestão dos sistemas de recuperação da informação, no âmbito dos sistemas de comunicação e de avaliação científica e tecnológica.

Palavras-chave: Bibliometria. Gestão da Informação. Gestão do Conhecimento. Sistemas de Recuperação da Informação. Informação Científica e Tecnológica.

BIBLIOMETRICS AND SCIENTIFIC AND TECHNOLOGICAL INFORMATION AND KNOWLEDGE MANAGEMENT: a review of the literature

ABSTRACT – This study discusses the field of Bibliometrics highlighting the application of bibliometrics laws and its assumptions in the scientific and technological information and knowledge management. The main objective is to present a brief review of the literature on Bibliometrics, distinguishing the use of bibliometric indicators as basics mathematical and statistical tools of knowledge organization, information and knowledge management and dissemination as well as key elements for the development of more accurate, economical and scientifically established systems. Initially, a brief history of the subject area is traced. Different denominations as scientometrics, informetrics and webometrics are presented, pointing out the conceptual distinction between those fields of knowledge and its applications. Following are cited authors who have studied and enunciated the bibliometric laws and principles, with special emphasis on its theoretical and practical approaches to their objective empirical indicators as a tool to quantify the processes of scientific communication. Finally, it appears evident the importance of Bibliometrics in the information retrieval systems management within communication systems as well as in scientific and technological evaluation systems.

Key-words: Bibliometrics. Information Management. Knowledge Management. Information Retrieval Systems. Scientific and Technological Information.

Vania Lisboa da Silveira Guedes

Professora Adjunta da UFRJ, ligada ao Curso de Biblioteconomia e Gestão de Unidades de Informação e ao Curso de Pós-graduação em Políticas de Informação e Organização do Conhecimento. Pós-doutoranda do Programa de Pós-graduação em Linguística da UFRJ. Doutora em Linguística. Mestre em Ciência da Informação. Atua na área de Organização do Conhecimento, Bibliometria e Indexação na Ciência da Informação, assim como na área de Sociolinguística, na Linguística, com ênfase em Análise de Discurso, Análise da Escrita Científica do Gênero Acadêmico, Análise Léxico-morfológica.

[.vanielisboa@facc.ufrj.br](mailto:vanielisboa@facc.ufrj.br)

1. INTRODUÇÃO

Este artigo apresenta uma revisão da literatura sobre marcos teóricos e aplicações práticas importantes para a composição do campo da Bibliometria, no âmbito da Ciência da Informação. Sob essa perspectiva, o estudo menciona questões ligadas à nomeação da área de assunto, descreve as leis e princípios que integram a Bibliometria e se vale ainda de alguns exemplos da utilização dos indicadores bibliométricos, como ferramenta de organização, gestão e difusão da informação e do conhecimento científico e tecnológico. O estudo fundamenta-se, sobretudo, nas pesquisas sobre Bibliometria editadas na dissertação de mestrado e tese de doutorado de Guedes (1992, 2010) e considera uma versão preliminar apresentada no VI Encontro Nacional de Ciência da Informação, por Guedes e Borschiver (2005), salientando, ainda, algumas potencialidades e tendências contemporâneas da pesquisa no âmbito das metrias. Nesse sentido, defende a importância do desenvolvimento de estudos teóricos e descritivos bibliométricos, para o avanço do campo conceitual da Bibliometria e obtenção de indicadores matemáticos e quantitativos importantes para o estabelecimento de critérios científicos na gestão e organização da informação e do conhecimento. Em síntese, o objetivo principal é desenvolver uma revisão da literatura sobre a criação e composição da Bibliometria, destacando a importância fundamental da produção de indicadores bibliométricos, para a tomada de decisão no âmbito da organização, disseminação, avaliação e gestão da informação e do conhecimento.

A estrutura do artigo apresenta os seguintes itens. O primeiro introduz de forma panorâmica o estudo. O segundo traça um breve histórico da nomeação e conceituação do campo da Bibliometria e inclui a apresentação da interseção da área com a Cientometria¹, Informetria e Webmetria. As leis de Bradford, de Lotka, de Zipf e o Ponto de Transição (T) de Goffman, desde as enunciações e os pressupostos teóricos até as aplicações, acham-se resumidos no terceiro item. O item quatro aborda as leis e princípios bibliométricos que

¹ Devido à verificação de variação na escrita dos termos utilizados para a denominação das Metrias da comunicação, optou-se, neste estudo, pela terminologia adotada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).

utilizam as citações como objeto de análise. Menciona, portanto, a Frente de Pesquisa e os Colégios Invisíveis, o Fator de Imediatismo ou de Impacto, o Acoplamento Bibliográfico e a Cocitação, a Obsolescência da Literatura e a Vida-média, a Lei do Elitismo, a Teoria Epidêmica de Goffman, finalizando com a Lei dos 80/20. Dando continuidade, é delineado um quadro síntese que relaciona as principais leis e princípios bibliométricos, seus focos de estudo e suas principais aplicações na gestão da informação e do conhecimento. A partir daí, são destacadas algumas considerações que evidenciam a importância do desenvolvimento de estudos teóricos, descritivos e práticos bibliométricos, para a organização e gestão da informação e do conhecimento, nos sistemas de recuperação da informação, assim como nos sistemas de avaliação e de comunicação científica e tecnológica. Finalmente, são apresentadas as referências dos estudos consultados e citados no desenvolvimento do artigo.

2. BREVE HISTÓRICO DA NOMEAÇÃO DA BIBLIOMETRIA E A INTER-RELAÇÃO COM OUTRAS METRIAS

A área de Bibliometria na Ciência da Informação é constituída por um conjunto de leis e princípios empíricos, que contribuem para o estabelecimento da fundamentação teórica da Ciência da Informação. Miranda Pao (1989) refere-se à Bibliometria como um termo introduzido por Allan Pritchard, em seu artigo *Statistical Bibliography or Bibliometrics*, editado em 1969, para denotar a área de assunto, que utiliza métodos matemáticos e estatísticos, com o objetivo de analisar os processos de comunicação escrita. Pritchard, segundo Pao (1989), percebeu que a literatura é o “ingrediente chave” no processo de comunicação científica e que os atributos de uma unidade de literatura, representada pelo artigo de periódico, patente, livro, entre outros, podem ser analisados estatisticamente. Pao (1989) acrescenta que publicações, autores, palavras chave, usuários, citações e títulos de periódicos são alguns dos parâmetros pesquisados em estudos bibliométricos da literatura, como instrumentos de quantificação e prognosticação dos processos de comunicação escrita.

O termo *Statistical Bibliography*, hoje Bibliometria, foi usado pela primeira vez em 1922 por E. Wyndham Hulme, antecedendo à data a qual se atribui a formação da área de Ciência da Informação, com a conotação de esclarecimento dos processos científicos e tecnológicos por meio da contagem de documentos. Após Hulme, pressupõe-se que o termo *Statistical Bibliography* foi ignorado por 22 (vinte e dois) anos, até ser usado por Gosnell, em 1944, em um artigo sobre obsolescência da literatura, sem o reconhecimento de seu uso prévio. Consta da literatura sobre Bibliometria (PAO, 1989) que existiu um intervalo de cerca de 20 (vinte) anos, até 1962, quando o termo *Statistical Bibliography* foi mencionado pela terceira vez, por L. Miles Raisig, em um estudo sobre análise de citações, intitulado *Bibliography in Health Sciences*. Raisig (1962) apresenta uma revisão da literatura sobre métodos quantitativos de pesquisas na área das ciências da saúde e conceitua o termo *Statistical Bibliography* como reunião e interpretação de estatísticas relacionadas a livros e periódicos. Todavia, existia um consenso entre pesquisadores de que o termo *Statistical Bibliography* não era de todo satisfatório, o que se verifica pelo seu escasso emprego na literatura.

Raisig (1962) salienta que a análise estatística da literatura sobre anatomia comparativa, no período de 1550 a 1860, relatada em 1917 por Cole e Eales, é um exemplo pioneiro de estudo bibliográfico utilizado para demonstrar movimentos históricos. Em uma discussão textual ilustrada por uma série de gráficos, os autores demonstram claramente a distribuição da literatura sobre anatomia comparativa entre países, no período, e fornecem as fontes utilizadas para a análise estatística. O estudo de Cole e Eales demonstra como uma pesquisa histórica da área de anatomia comparativa, representada estatisticamente e correlacionada com outros dados, pode contribuir para a atribuição de novo significado e surgimento de novas interpretações sobre o nascimento e a morte.

Guedes (1992, p. 38) destaca que,

distintamente de sua concepção atual, em 1934, o termo “Bibliometrie” foi utilizado por Otlet, em sua obra intitulada “*Traité de Documentation*”. Entretanto, Paul Otlet referia-se à mensuração física do livro, demonstrando uma preocupação com as características extrínsecas do livro, em contraste com a abordagem de Pritchard (1969), que revela preocupação com as características intrínsecas, isto é

com a informação registrada.

Pinheiro (1983, p. 60) afirma que a Bibliometria antecede à Ciência da Informação como preocupação com as análises documentais voltadas, no entanto, para a contagem de livros. Martyn e Lancaster (1981), ao analisarem as definições de Pritchard, inferem que as técnicas bibliométricas básicas envolvem a contagem de material impresso, qualificado como escrito por Pritchard (1969). Okubo (1997) afirma que as abordagens bibliométricas são baseadas na noção de que a produção de conhecimento é a essência da pesquisa científica e, nesse contexto, a literatura científica é a manifestação constituinte do conhecimento. Okubo esclarece que os dados e indicadores bibliométricos podem servir como ferramentas ao menos para a descrição e expressão de questões que surgem no mundo da Ciência.

Tague-Sutcliffe (1992, p. 1) salienta que a Bibliometria “é o estudo dos aspectos quantitativos dos processos de produção, disseminação e uso da informação registrada”. O autor acrescenta que a “Bibliometria desenvolve medidas e modelos matemáticos visando à sua utilização em estudos de prospecção e tomadas de decisão”. Já a Cientometria, segundo Tague-Sutcliffe, “é o estudo dos aspectos quantitativos da ciência como uma disciplina ou atividade econômica” aplicada no desenvolvimento de políticas científicas. Envolve, portanto, “estudos quantitativos das atividades científicas, sobrepondo-se à Bibliometria”. Finalmente, ele ressalta que a Informetria refere-se ao “estudo dos aspectos quantitativos da informação, em qualquer formato”, voltada para a comunicação, oral ou escrita, bem como para as necessidades e usos de informação de qualquer indivíduo, não apenas de uma elite intelectual.

Braun e Schubert (2003) mencionam que a Cientometria investiga os aspectos quantitativos ligados à produção, disseminação e uso de informações científicas, com a finalidade de contribuir para uma melhor compreensão do processo de pesquisa científica, enquanto atividade social. Vanti (2002, p. 156) acrescenta a Webmetria como a área emergente na Ciência da Informação que consiste “na aplicação de métodos infométricos à World Wide Web”. Bar-Ilan (2008) apresenta uma revisão da literatura sobre o campo da Informetria no início do século XXI, focalizando estudos publicados entre 2000 e 2006. A

autora ressalta que o início do século é marcado pela criação de novos temas de pesquisa, como a comparação entre base de dados de citações, notadamente a partir da introdução das bases Scopus e Google Scholar que possibilitam o livre acesso e a visualização de mapeamento de citações, entre outros índices bibliométricos. Nesse cenário, observa-se também a intensificação de tópicos previamente existentes, como por exemplo, o Fator de Impacto e o ranqueamento de títulos de periódicos, artigos altamente citados e instituições de ensino, de grande interesse público. Santos e Kobashi (2009) ressaltam que a Informetria é um modelo mais recente, no campo das metrias, que se vale da apropriação dos métodos bibliométricos e cientométricos para a obtenção de conhecimento acerca dos aspectos de natureza cognitiva da atividade científica. Ao discutirem os estudos métricos da informação, os autores discorrem também sobre a constituição e institucionalização da Bibliometria, da Cientometria e da Informetria, salientando os seus principais conceitos, semelhanças e diferenças, bem como suas potencialidades, fronteiras e as tendências atuais das abordagens teóricas e práticas, nesses domínios do conhecimento.

Segundo o International Society of Scientometrics and Informetrics (2011), a Informetria, a Cientometria, a Bibliometria e a Webmetria são subáreas estreitamente relacionadas, ou interligadas, que pertencem à área de Ciência da Informação. Essas subáreas são utilizadas nas análises quantitativas e na mensuração de todas as formas de informação registradas, fundamentando a investigação da distribuição, circulação e amplo uso da informação, no contexto ou entre indivíduos, disciplinas, organizações e países. Nesse sentido, Saracevic e Garfield (2010) ressaltam que os avanços contemporâneos das tecnologias de informação e comunicação permitiram a criação inovadora de grandes bases de dados e a incorporação de publicações e dados sobre citações, a partir dos quais, entre outras, uma variedade de métricas são derivadas. Produtividade e impacto acadêmico estão sendo investigados quantitativamente pela análise de dados obtidos em grandes bases de dados. Os resultados são frequentemente utilizados para uma variedade de propósitos avaliativos. Assim, é feita uma distinção entre Bibliometria/Cientometria

relacional, medindo entre outros a produtividade, bem como Bibliometria/Cienciometria avaliativa, medindo, por exemplo, o fator de impacto da literatura.

Vanti (2002), ao confrontar a Bibliometria, a Cientometria, a Informetria e a Webmetria, ressalta que as quatro subdisciplinas possibilitam a mensuração dos fluxos de informação, da comunicação científica e da difusão do conhecimento. Segue o diagrama, desenvolvido por Vanti, representando graficamente a superposição da Bibliometria, Cientometria, Informetria e Webmetria.

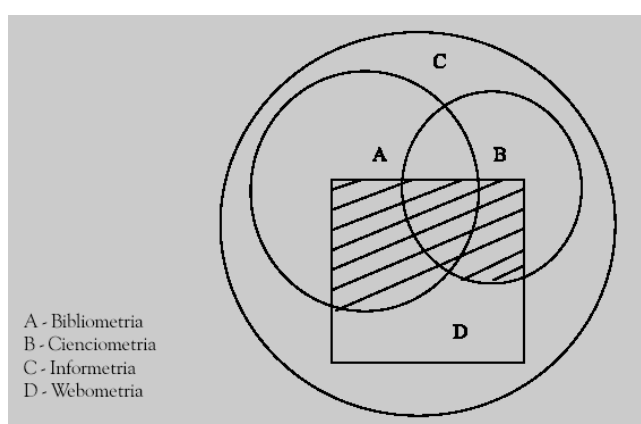


Figura 1 Diagrama representando graficamente a inter-relação entre as quatro subáreas de Metrias da comunicação.

Fonte: Vanti (2002, p. 161)

O presente estudo volta-se para a Bibliometria, no âmbito da Ciência da Informação, como área do conhecimento que possibilita a produção de indicadores, a partir da análise dos aspectos matemáticos e estatísticos da comunicação científica em domínios específicos do conhecimento. Em suma, o objetivo do estudo é apresentar uma revisão da literatura sobre Bibliometria, destacando a criação e denominação da área, as leis e princípios bibliométricos, como também a utilidade de seus indicadores como ferramenta estatística de organização, disseminação e gestão da informação e do conhecimento. É apresentada ainda uma breve reflexão sobre os conceitos das subáreas que distinguem os campos de aplicação dos indicadores bibliométricos nas metrias da comunicação, ou seja, da Bibliometria, da Ciencometria, da Informetria e da Webmetria.

3. LEIS E PRINCÍPIOS BIBLIOMÉTRICOS

As principais leis bibliométricas são: Lei de Bradford, (produtividade de periódicos), Lei de Lotka (produtividade científica de autores) e Leis de Zipf (frequência de palavras).

O padrão de distribuição das leis e princípios bibliométricos segue a máxima “poucos com muito e muito com poucos”. Essa máxima é conhecida como o Efeito Mateus na Ciência, que diz: aos que mais têm será dado em abundância e, aos que menos têm, até o que têm lhes será tirado (MERTON, 1968, p. 58). Trata-se de uma abordagem ao Efeito Mateus na Ciência baseada na análise de processos psicossociais, que afetam o sistema de avaliação e distribuição de recompensas científicas pelas contribuições. Sob essa perspectiva, Crane (1965), citado por Merton (1968), em um estudo de caso sobre a produtividade científica conclui que cientistas altamente produtivos, de universidades conceituadas, obtêm frequentemente mais reconhecimento que cientistas igualmente produtivos de outras universidades.

Verifica-se na Lei de Bradford, que permite estimar o grau de relevância de periódicos em uma área do conhecimento, que os periódicos que produzem maior número de artigos sobre o assunto formam um núcleo de periódicos, supostamente de maior qualidade ou relevância para aquela área. A Lei de Lotka considera que alguns pesquisadores, supostamente de maior prestígio em uma área do conhecimento, produzem muito e muitos pesquisadores, supostamente de menor prestígio, produzem pouco. Nas Leis de Zipf, que permitem estimar as frequências de ocorrência das palavras em textos científicos e tecnológicos e delimitar a região de concentração de termos de indexação, que um pequeno grupo de palavras ocorre muitas vezes e um grande número de palavras é de pequena frequência. A seguir são enunciadas e discutidas as leis acima mencionadas.

3.1 LEI DE BRADFORD

Samuel Clement Bradford (1878-1948), químico e bibliotecário da British Library, no

âmbito de iniciativas voltadas para o desenvolvimento de uma literatura periódica científica, compilou e analisou a bibliografia sobre geologia e óleos lubrificantes, publicada nas décadas de 20 e 30 do século XX. Observou então que alguns periódicos produziam muitos artigos, outros produziam alguns artigos e muitos, poucos artigos. Bradford verificou que o núcleo de periódicos mais produtivos na área em análise continha maior número de artigos sobre o assunto do que os periódicos que tratavam de assuntos distintos; possibilitando, conseqüentemente, a ordenação de periódicos segundo zonas de produtividade decrescente de artigos.

Assim, a Lei de Bradford, relacionada à dispersão da literatura periódica científica, enuncia que:

Se periódicos científicos forem ordenados em ordem decrescente de produtividade de artigos sobre determinado assunto, poderão ser divididos em um núcleo de periódicos mais particularmente dedicados ao assunto e em vários grupos ou zonas, contendo o mesmo número de artigos que o núcleo. O número de periódicos (n), no núcleo e zonas subseqüentes, variará na proporção $1:n:n^2 \dots$ (BROOKES, 1969, p. 953)².

A Lei de Bradford sugere que, na medida em que os primeiros artigos sobre um novo assunto são produzidos, eles são submetidos a uma pequena seleção, por periódicos apropriados, e são aceitos. Esses periódicos, inicialmente selecionados, atraem mais e mais artigos, no decorrer do desenvolvimento da área de assunto. Porém, ao mesmo tempo, outros periódicos começam a publicar seus primeiros artigos sobre o assunto. Se o assunto continua a se desenvolver, emerge eventualmente um núcleo de periódicos mais produtivos, no que se refere à edição de artigos sobre o assunto. Brookes (1969, p. 954) refere-se a esse fenômeno como o “mecanismo do sucesso gerando o sucesso”.

Goffman e Morris (1970) acrescentam que a definição do núcleo é arbitrária na medida em que depende do número de artigos selecionados para a distribuição de títulos de periódicos em zonas. Pinheiro (1983) apresenta uma proposta de mudança conceitual. A autora entende que os periódicos que compõem o núcleo são os mais produtivos. Pinheiro (1983, p. 60) acrescenta que a lei de Bradford vem sendo objeto de pesquisa em

² Onde n é igual a um número x de periódicos (CHEN; CHONG; TONG, 1994: 536).

campos distintos do conhecimento.

A Lei de Bradford é de grande importância para o estabelecimento de políticas de desenvolvimento de coleções, sobretudo para a aquisição e descarte de títulos de periódicos. Sua aplicação possibilita estimar a magnitude de determinada área do conhecimento e o custo de toda e qualquer fração específica da bibliografia, no todo. Segue o quadro 1 ilustrando a aplicação idealizada da Lei de Bradford, exemplificado por Saracevic (2004).

Nº de periódicos	Nº de artigos por periódico	Nº total de artigos
3 {	1	60
	2	35
9 {	1	30
	2	25
	2	9
	4	8
27 {	10	6
	7	5
	5	4
	5	3

Quadro 1 Lei de Bradford: um exemplo idealizado
 Fonte: Saracevic (2004).

3.2 LEI DE LOTKA

A Lei de Lotka, relacionada à produtividade de autores e fundamentada na premissa básica de que alguns pesquisadores publicam muito e muitos publicam pouco (VOOS, 1974), enuncia que a relação entre o número de autores e o número de artigos publicados por esses, em qualquer área científica, segue a Lei do Inverso do Quadrado: $1/n^2$. Isto é, em um dado período de tempo, analisando um número n de artigos, o número de cientistas que escrevem dois artigos seria igual a

$1/4$ do número de cientistas que escreveram um. O número de cientistas que escreveram três artigos seria igual a $1/9$ do número de cientistas que escreveram um, e

assim sucessivamente. O mesmo autor revela que Price (1963), em *Little Science, Big Science*³, observou que, para as ciências em geral, o número de autores decresce mais rapidamente que o inverso do quadrado, mais aproximadamente à Lei do Inverso do Cubo $1/n^3$.

Na gestão da informação e do conhecimento, assim como no planejamento científico e tecnológico, sua aplicabilidade se verifica na avaliação da produtividade de pesquisadores, na identificação dos centros de pesquisa mais desenvolvidos e no reconhecimento da “solidez” de uma área científica. Ou seja, quanto mais solidificada estiver uma ciência, maior a probabilidade de seus autores produzirem múltiplos artigos, em dado período de tempo. Segue o quadro 2 ilustrando a aplicação idealizada da Lei de Lotka.

Nº de Artigos	Nº de Autores	Proporcionalidade de Autores
1 artigo	280 autores	
2 artigos	70 autores	1/4 ($1/n^2$) (n=2)
3 artigos	31,11 autores	1/9 ($1/n^2$) (n=3)
4 artigos	17,5 autores	1/16 ($1/n^2$) (n=4)
5 artigos	11,2 autores	1/25 ($1/n^2$) (n=5)
n= número de artigos		

Quadro 2 Representação da aplicação idealizada da Lei de Lotka.
Fonte: Fonte Própria

3.3 LEIS DE ZIPF

As Leis de Zipf, relacionadas à frequência de ocorrência de palavras em um determinado texto, enriquecida pelo conceito do Ponto de Transição (T) de Goffman⁴,

³ *Little Science, Big Science*, refere-se à denominação do Seminário de Estudos em Ciência da Informação, onde pela primeira vez, em 1963, Price apresentou dados estatísticos sobre o fenômeno do crescimento exponencial da literatura.

⁴ Será detalhado a partir da página 12, item 2.4. (PAO, 1989).

relacionam-se diretamente com a indexação automática da informação.

Zipf observou que, num texto suficientemente longo, existia uma relação entre a frequência que uma palavra ocorria e sua posição⁵ na lista de palavras ordenadas segundo sua frequência de ocorrência. Essa lista era confeccionada, levando-se em conta a frequência decrescente de ocorrências. À posição nesta lista dá-se o nome de ordem de série (rank). Assim, a palavra de maior frequência de ocorrência tem ordem de série 1, a de segunda maior frequência, ordem de série 2 e, assim, sucessivamente. Zipf verificou que o produto da ordem de série (r) de uma palavra pela sua frequência de ocorrência (f) era aproximadamente constante (C). Enunciou, assim, a Primeira Lei de Zipf, representada matematicamente por

$$(1) r \times f = C$$

Fairthorne (1969) ressaltou que Zipf considera essa relação como uma consequência do Princípio Geral do Menor Esforço. Pao (1978) reconhece que essa lei é elegante em sua simplicidade. Entretanto, ela se aplica somente a palavras de alta frequência de ocorrência, em um texto. Para palavras de baixa frequência, Zipf propôs uma segunda lei, revisada e modificada por Booth (1967).

A Segunda Lei de Zipf enuncia que, em um determinado texto, várias palavras de baixa frequência de ocorrência (alta ordem de série) têm a mesma frequência. Booth (1967), ao modificá-la a representa matematicamente da seguinte forma:

$$(2) I_n = \frac{2I_1}{n(n+1)}$$

Onde I_1 é o número de palavras que têm frequência 1, I_n é o número de palavras que têm frequência n , sendo 2 a constante válida para a língua inglesa. Assim, para se

⁵ Localização em uma lista, confeccionada segundo a ordem decrescente de frequência de ocorrência das palavras, que compõem um determinado texto.

calcular o total de palavras que ocorrem três vezes ($n = 3$), tendo por hipótese que o total de palavras com frequência 1 no ranking (I_1) é 250, o cálculo é:

$$(3) I_3 = \frac{2 \times 250}{3(3+1)} = \frac{500}{12}$$

Logo, segundo Zipf, o total de palavras que ocorrem 3 (três) vezes é igual a 41,666..., ou seja, a aproximadamente 42. A Segunda Lei de Zipf é também denominada Lei de Zipf-Booth.

Guedes e Valois (1988) observam que a expressão $n(n+1)/2$ corresponde à soma dos n primeiros números naturais. Por exemplo: a soma de $1 + 2 + 3 + 4 + 5 = 15$ pode ser calculada, levando-se em conta que $n = 5$, por

$$(4) \frac{5(5+1)}{2} = 15$$

Segundo Pao (1978), Goffman observou que a Primeira Lei de Zipf era válida apenas para a região de palavras de alta frequência de ocorrência. Essas palavras têm a propriedade de ocupar rank único na lista de distribuição de palavras, isto é, das palavras de alta frequência de ocorrência, dificilmente, existem duas palavras com a mesma frequência. Por outro lado, a Segunda Lei de Zipf, modificada por Booth, descreve o comportamento das palavras de baixa frequência de ocorrência. Nessa região, observa-se que existem muitas palavras com a mesma frequência. Esses dois comportamentos, inteiramente distintos, definem as duas extremidades da lista de distribuição de palavras de um dado texto. Assim, é razoável esperar uma região crítica, na qual há a transição do comportamento das palavras de alta frequência para as de baixa frequência. Goffman, segundo Pao (1978), admitiu como hipótese que nessa região de transição estariam as palavras de maior carga semântica de um texto em análise.

Os estudos estatísticos, baseados na frequência de ocorrência de palavras de

artigos científicos e tecnológicos, têm revelado grande potencial para automatização do processo de indexação temática. Luhn (1957) ressalta que esses estudos partem da premissa básica de que, ao produzir um artigo científico e tecnológico, um autor tende a decidir por um determinado nível de especificidade de linguagem e comunicação de palavras, para transmitir sua mensagem. Existe também, segundo Luhn (1957), a probabilidade de quanto mais frequentemente uma noção e combinação de noções ocorrerem, maior importância o autor lhes atribuirá para expressar a essência de sua idéia no todo. Ele observa que a probabilidade de um autor empregar palavras diferentes, para expressar a mesma idéia, é pequena. Infere, ainda, que as palavras de frequência média são as de maior carga semântica e, conseqüentemente, as mais adequadas para a indexação temática de um texto. Assim, Hans Peter Luhn (1957) propôs que a indexação deveria ser baseada na própria literatura. Tópicos relevantes, termos e a relação entre eles, sobre um assunto, deveriam ser derivados da análise da escrita científica de documentos sobre aquele assunto. Vários estudos têm se baseado nos pressupostos lançados por Luhn e utilizado a frequência de ocorrência como indicador do grau de relevância das palavras em relação a um texto em análise.

3.4 PONTO DE TRANSIÇÃO (T) DE GOFFMAN

De acordo com Goffman, para chegar a essa região de transição, onde estariam as palavras de alto conteúdo semântico, a expressão da Segunda Lei de Zipf teria que fornecer o comportamento típico das palavras de alta frequência, isto é, o número de palavras que têm frequência n tenderia a 1 (unidade). Substituindo-se, na expressão da Segunda Lei de Zipf-Booth, I_n por 1, obtém-se:

$$(5) \frac{I_1}{1} = \frac{n(n+1)}{2}$$

Ou ainda, rearranjando

$$(6) n^2 + n - 2I_1 = 0$$

cujas raízes são

$$(7) n = \frac{-1 \pm \sqrt{1 + 8I_1}}{2}$$

Da expressão acima, interessa somente determinar a raiz positiva, assim

$$(8) n = \frac{-1 + \sqrt{1 + 8I_1}}{2}$$

Ao valor de n assim determinado corresponde o Ponto de Transição (T) de Goffman. O Ponto T de Goffman determina graficamente a localização onde ocorre a transição das palavras de baixa frequência para as de alta frequência. Existe uma determinada região, ao redor desse ponto, com probabilidade de concentrar as palavras de alto conteúdo semântico, que seriam portanto utilizadas na indexação de um texto em análise. Goffman apresenta, com o Ponto de T, a primeira oportunidade de se decompor um texto sintaticamente, objetivando a sua indexação.

Pao (1978), ao testar a hipótese de que a Região de Transição de Goffman produziria as palavras de maior conteúdo semântico de um texto, delimitou a região projetando para baixo o mesmo número de palavras que ocorreu acima da frequência associada ao Ponto T. Nessa região, foi calculado o total de palavras e, desse total, excluídas as palavras sem conteúdo semântico, tais como: preposições, artigos, conjunções, pronomes e verbos auxiliares. Essas palavras foram excluídas dos dados, na medida em que o objetivo do estudo era a investigação da carga semântica dos termos localizados na Região de Transição de Goffman. Ao analisar os resultados atingidos, ela conclui que eles foram satisfatórios, para os textos analisados, e parecem indicar que a

seleção automática de termos de indexação, a partir de uma lista de frequência de ocorrência de palavras, é uma perspectiva promissora para a indexação automática (GUEDES, 2010).

No caso específico de textos “suficientemente longos”, a aplicação das leis de Zipf, Zipf/Booth, Ponto T de Goffman e similares produz uma listagem de palavras acompanhadas de respectivas frequências de aparecimento no texto. Segundo Braga (1996, p. 56):

Os estudos feitos até agora parecem indicar que realmente é possível falar-se de desconstrução do contexto das palavras para construção de um conteúdo semântico de interesse imediato para os processos de representação da informação e de indexação.

Rouault (1987), em sua discussão sobre métodos estatísticos de indexação temática automática, apesar de questionar a validade das leis de Zipf, destaca que podemos delimitar aproximadamente três zonas em qualquer índice e verifica que as fronteiras entre elas não são claras. Acrescenta, entretanto, que essas observações são válidas para todo o corpus analisado. Segundo Rouault (1987), a primeira zona é constituída de formas com um número elevado de ocorrências. Ela contém, acima de tudo, palavras cuja presença deve-se a razões de sintaxe (estrutura linguística do idioma, no qual é escrito o corpus). No idioma francês, encontramos, essencialmente, determinantes, tais como: “le” (o), “la” (a); preposições: “des” (dos, das), representando cerca de 10% do corpus; conjunções, verbos auxiliares etc. Essa zona contém poucos representantes das categorias de substantivos, adjetivos, verbos. Já a segunda zona, na opinião de Rouault, caracteriza-se, sobretudo, pelo fato de conter uma quantidade maior que a primeira zona de representantes de categorias morfológicas “informativas”, tais como: substantivo, adjetivo, verbo, entretanto, as formas gramaticais que caracterizam a primeira zona ocorrem, ainda, sobretudo no topo dessa segunda zona. A terceira zona contém as formas que jamais serão utilizadas como descritores, como palavras chave. A parte mais importante da terceira zona é composta por formas que ocorrem uma única vez no corpus e que representam frequentemente 50% das formas distintas de um corpus. O autor volta a enfatizar que o limite entre a primeira, a segunda e a terceira

zonas é muito tênue. Segundo ele, a fronteira deve ser resultado de uma decisão arbitrária de fixação de um limite entre as efetivas formas de cada zona e depende do corpus utilizado. Finalmente, ao abordar a identificação de três zonas de ocorrência de palavras, ele ressalta que a segunda e o final da primeira caracterizam-se pela presença de palavras com potencial para representar tematicamente os textos analisados (GUEDES, 2010). Seguem a representação da Região de Transição de Goffman, e resultados da aplicação da Segunda Lei de Zipf, apresentados em Guedes (2010) a partir da análise de artigo de periódico científico, sobre vinicultura, de autoria de Mamede e Pastore (2004).

Rank	Word	Frequency	C		
1	De	197	197		PONTO T DE GOFFMAN: 27 (27,07)
2	A	109	218		
3	E	72	216		TOTAL DE PALAVRAS: 2369
4	Fermentação	59	236	I ₁	380
5	Da	51	255	I ₂	91
6	O	41	246	I ₃	43
7	Em	37	259	I ₄	29
8	Foi	33	264	I ₅	15
9	Do	32	288		
10	As	29	290		
11	Kloeckera	25	275		
12	Apiculata	23	276		
13	Saccharomyces	23	299		
14	Com	22	308		
15	Leveduras	22	330		
16	Mostos	22	352		
17	Cerevisiae	21	357		
18	Concentração	21	378		
19	Foram	21	399		
20	Produção	21	420		

Quadro 3 Representação da Região de Transição de Goffman no artigo de Mamede e Pastore (2004).
Fonte: Guedes (2010)

Hrebicek (2002) alerta que, do ponto de vista estatístico, nenhuma objeção razoável é feita às Leis de Zipf. Do ponto de vista linguístico, entretanto, verifica-se uma ausência dificilmente admissível de fenômenos relevantes ligados às estruturas lexicais, atualmente ocorrendo em textos. Qualquer unidade lexical de uma linguagem natural tem uma abstração intuitiva de seu significado deduzido a partir de seu uso em textos. As propriedades semânticas das palavras podem ser mais adequadamente observadas

quando são testadas como unidades lexicais em textos individuais.

Bybee e Hopper (2001, p. 1), citados por Guedes (2010), ao discutirem a relação da frequência de uso e da frequência de tipo com a emergência de estruturas linguística, mencionam que George K. Zipf foi o pioneiro nos anos 30. A obra *“The Psycho-biology of Language”* editada em 1935 se baseia no *Princípio do Menor Esforço*. Zipf antecipou muitos temas de investigação explorados recentemente sobre a relação entre a frequência e a estruturas linguísticas. Nos anos 80, outros linguistas procuraram respostas nos usos reais no âmbito do discurso. A propósito os tesouros são introduzidos nas linguagens documentárias valendo-se da noção de domínio e da frequência de uso dos itens lexicais.

Guedes (1992) fundamenta-se na Bibliometria, especificamente nas Leis de Zipf e Ponto de Transição de Goffman, ao propor contribuição para o desenvolvimento de um critério para indexação automática derivativa de textos científicos e tecnológicos. Guedes (2010) salienta que na Bibliometria, especialmente no que se refere à aplicação das Leis de Zipf e Ponto de T de Goffman à indexação da informação, os pressupostos teóricos apontam para uma ligação entre a frequência de ocorrência de uma determinada palavra e o seu potencial de representação temática em determinado *corpus* em análise. Na interface Linguística e Indexação, são analisados *corpora* textuais, utilizando-se conhecimentos linguísticos e estatísticos para a construção de referenciais, com o objetivo de contribuir para o desenvolvimento de sistemas de indexação automática. Como exemplos de trabalhos interdisciplinares dessa natureza, podemos citar Pao (1978), Basílio (1979), Basílio, Braga e Peirotti (1979), bem como Mauceri (1995) e Polanco, Grivel e Ro Auté (1995). Os estudos, fundamentados na frequência de ocorrência de palavras, têm evoluído para tentativas de desenvolvimento de algoritmos. O objetivo principal é o de contribuir para a automatização, em parte ou no todo, da indexação temática da informação. A autora acrescenta que essa linha de raciocínio representa um passo importante, na busca de um critério para o processamento automático da indexação. As Leis de Zipf vêm sendo aplicadas, inclusive, para identificar estilos distintos de autores, na redação de artigos científicos e como ferramenta estatística, em diferentes áreas do conhecimento, tais como: Linguística, urbanismo, física, medicina, Economia. Nesses

domínios do conhecimento, a estatística utiliza-se da frequência de ocorrência, com aporte de teorias probabilísticas, para a investigação e obtenção de conhecimentos acerca de determinados fenômenos, possibilitando o diagnóstico e a prospecção de eventos de importância crucial para as tomadas de decisão.

A figura 2 representa espacialmente as três principais leis bibliométricas e seus respectivos focos de estudo, considerando-as inseridas em um sistema de informação científica e tecnológica e este, em um sistema de comunicação científica e tecnológica.

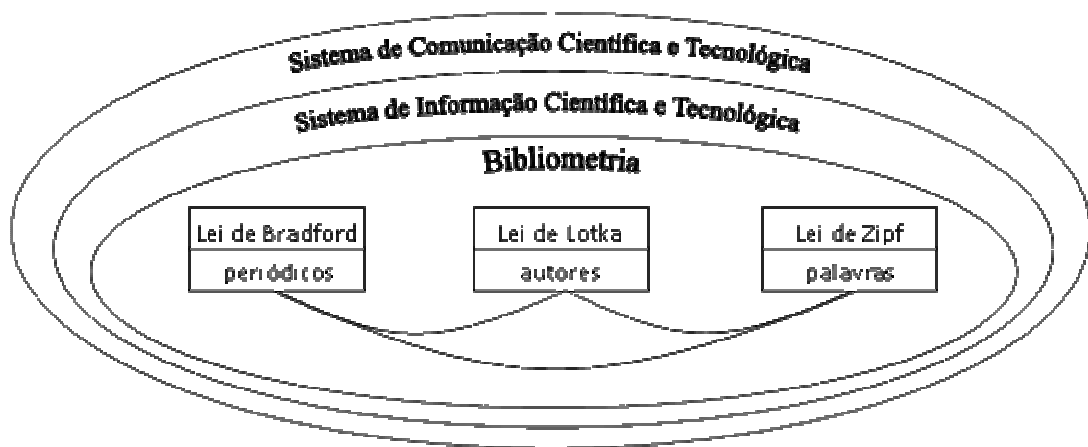


Figura 2 Principais leis da Bibliometria, seus focos de estudo e suas relações com os sistemas de comunicação e de informação científica e tecnológica.
Fonte: Guedes e Borschiver (2005).

Vale ressaltar que o conceito de sobrecidade⁶, na Ciência da Informação, se faz absolutamente indispensável à aplicação das Leis de Bradford, Zipf e Ponto T de Goffman. A Lei de Bradford tem como objetivo estimar o grau em que um periódico é sobre dado assunto, da mesma forma que as de Zipf e Ponto T de Goffman têm como objetivo determinar sobre que assunto é um determinado texto científico e tecnológico.

⁶ Sobrecidade é um conceito da área de Ciência da Informação, que significa tematicidade, “sobre o que trata” um documento (ZIMAN, 1977).

4. OUTROS ESTUDOS E CONCEITOS APLICADOS À BIBLIOMETRIA

Neste item, são mencionadas algumas abordagens teóricas e descritivas ligadas à análise de citações na Bibliometria. A análise de citações⁷ fundamenta-se na hipótese de que citação é um indicador válido da influência de um trabalho citado sobre outro em análise, evidenciando conexões intelectuais. Nesse sentido, a análise de citações visa à mensuração da produção científica de um determinado pesquisador, de grupos de pesquisa, de instituições de pesquisa, países etc., como também à obtenção de indicadores acerca do desenvolvimento científico como um todo ou de um domínio específico do conhecimento.

A análise de citações foi utilizada inicialmente por Gross e Gross (1927, p. 386), em um estudo estatístico das referências mencionadas no volume editado em 20 de outubro de 1926 do periódico científico *The Journal of the American Chemistry Society*. Garfield (1972a, p. 527), em seu artigo sobre a aplicação da análise de citações como ferramenta de avaliação de periódicos científicos, salienta que desde 1927 outros estudos foram publicados, periodicamente, citando como exemplos Allen (1929), Gross e Woodford (1931), Hooker (1935), Henkle (1938), Fussler (1949), Bradford (1953) e Brown (1956). Nos anos 60, a análise de citações é intensificada dando origem à produção do Science Citation Index (SCI) do Institute of Scientific Information (ISI), por Garfield, o que possibilita o mapeamento de redes de citações nos periódicos científicos e tecnológicos mais importantes do mundo. Swales (1986) menciona que a análise de citações foi indubitavelmente intensificada pelo surgimento do Science Citation Index do Institute of Scientific Information, em 1961, e também pela acessibilidade na época a fitas de computador como suporte para o índice. Quanto à motivação para a análise de citações, Swales acrescenta que certamente tem existido uma expectativa de que a análise de citações seja de grande utilidade para os historiadores de áreas específicas do conhecimento em suas tentativas de estabelecimento da origem e distribuição de ideias particulares e descobertas, bem como no mapeamento de redes de influência,

⁷ Citação é a “menção no texto de uma informação extraída de outra fonte” (ABNT, 2002).

colaboração e dependência entre autores. Uma segunda e completamente diferente motivação vem sendo a validação e o refinamento de leis estatísticas bibliométricas. Uma terceira abordagem refere-se à investigação das diferenças e similaridades entre modelos de citações, ao longo do tempo, e entre áreas distintas do conhecimento. Bensman (2007), em seu estudo sobre Garfield e o Fator de Impacto, reitera que Garfield foi o fundador do Institute of Scientific Information, atualmente ThomsonReuters Scientific, responsável pelo lançamento do Science Citation Index, Social Citation Index e Arts & Humanities Citation Index.

Hyland (1999) investiga, do ponto de vista quantitativo e qualitativo, como as práticas de citação acadêmica contribuem para a construção de uma área do conhecimento. Com base na análise de citações em um corpus em meio eletrônico, composto por 80 artigos de periódicos e entrevistas com autores renomados, o autor verifica a variabilidade contextual de citações em oito áreas. Suas conclusões sugerem que as convenções para a produção textual, em cada comunidade discursiva, apresentam formas distintas típicas de negociação e validação do conhecimento. Nesse sentido, ele argumenta que essas diferenças nas práticas de citação estão relacionadas ao fato de que os acadêmicos participam ativamente na construção do conhecimento e que suas decisões discursivas são influenciadas profundamente pelos acordos epistemológicos e sociais como também pelas convenções de suas áreas de assunto. Hyland (1999) acrescenta que a citação é um elemento mediador da relação entre os argumentos desenvolvidos pelo autor e os argumentos compartilhados por sua comunidade discursiva. Para Bazerman (1983) os estudos de citações sugerem que a maioria dos cientistas leva em consideração um conjunto limitado de artigos que corresponde à estrutura de sua especialidade e à sua rede de contatos pessoais e profissionais. Price (1965, p. 512) em seu artigo intitulado *Networks of Scientific Papers* verifica que apenas um número limitado de artigos publicados recentemente está interligado formando uma frente de pesquisa.

4.1 FRENTE DE PESQUISA E COLÉGIOS INVISÍVEIS

A análise de citações possibilita, portanto, a identificação da Frente de Pesquisa de uma área científica específica, representada por um conjunto de autores que se citam na literatura recente. Na Frente de Pesquisa, observa-se um estreito padrão de relações múltiplas entre os autores. Nesse pequeno grupo de artigos entrelaçados, a análise de citações possibilita, também, o reconhecimento do trabalho de algumas centenas de colaboradores que formam os Colégios Invisíveis. Price (1963) observou que um grupo pequeno de autores e de publicações exerce maior influência, em uma área específica de assunto. Esses autores são líderes, na área, e seus trabalhos são, conseqüentemente, mais citados. Outras citações são distribuídas, regularmente e uniformemente, por toda a literatura passada, com frequência decrescente.

4.2 FATOR DE IMEDIATISMO OU DE IMPACTO

A análise de citações também pode ser utilizada para calcular o Fator de Imediatismo de um artigo publicado, investigando-se a concentração de citações a esse artigo, em documentos publicados nos últimos quinze anos (PRICE, 1965). A hipótese é de que, em uma área científica específica, artigos de periódicos citados mais frequentemente são mais relevantes do que artigos menos citados. Jones (2003), ao externar sua preocupação com o uso e abuso do Fator de Impacto, na avaliação da importância e do prestígio de periódicos científicos e de cientistas, defende que o Fator de Impacto de um título de periódico específico é calculado pela divisão do número de citações no ano corrente aos artigos do periódico, publicados nos últimos dois anos, pelo número total de itens, artigos e *reviews*, publicados no mesmo período de tempo, ou seja, no período de dois anos. Pao (1989) infere que o Fator de Impacto de um título de periódico é baseado no número de citações feitas aquele periódico por todos os títulos de periódicos indexados, por exemplo, na base de dados do *Science Citation Index* do *Institute of Scientific Information*. Garfield (1972b) explica que o Fator de Impacto pode ser calculado de várias formas. O primeiro método é computar a razão entre citações a determinados anos de publicação de um título de periódico e o número de artigos publicados naqueles

anos no título de periódico em análise.

4.3 ACOPLAMENTO BIBLIOGRÁFICO E CO-CITAÇÃO

O Acoplamento Bibliográfico é um procedimento de análise retrospectiva de citações que consiste na união de artigos citantes, citando o(s) mesmo(s) documento(s). Segundo Kesller (1963), dois artigos que compartilham a mesma referência contêm uma unidade de acoplamento.

Por outro lado, a Co-citação é um procedimento de análise prospectiva de citações que consiste na união de artigos, citados pelo(s) mesmo(s) documento(s). Small (1973) examina a co-citação como uma medida de associação entre pares de documentos citados frequentemente. Na opinião de Marshakova (1981), o Acoplamento Bibliográfico mede o grau de ligação entre dois ou mais artigos, segundo o número de documentos idênticos citados por esses artigos, e a Co-citação mede o grau de ligação de dois ou mais artigos, pelo número de documentos onde esses artigos são citados, simultaneamente.

Segundo Saracevic na análise de co-citações os artigos que citam um mesmo artigo são ligados pelo interesse dos autores pelo artigo citado. Seguem as figuras 3 e 4, adaptadas de Saracevic (2004), representando respectivamente o acoplamento bibliográfico e a análise de co-citações.

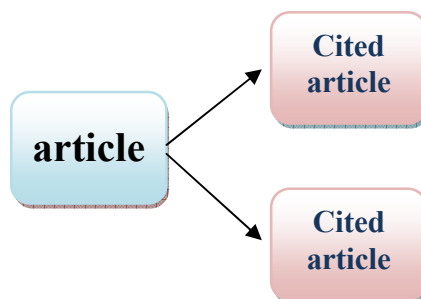


Figura 3 Representação da Análise de Co-citações.
Fonte: Adaptação de Saracevic (2004).

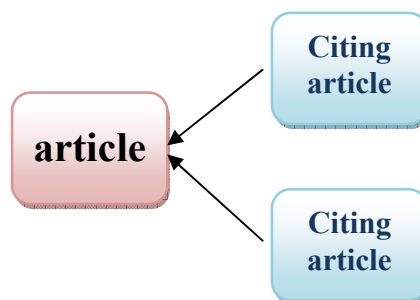


Figura 4 Representação do Acoplamento Bibliográfico.
Fonte: Adaptação de Saracevic (2004).

4.4 OBSOLESCÊNCIA DA LITERATURA E VIDA-MÉDIA

A Obsolescência da Literatura é estimada pela análise do declínio do uso da literatura, em um período de tempo, e a Vida-Média é calculada a partir da razão de obsolescência e da razão de crescimento, de um determinado corpus da literatura (LINE, 1970). No que se refere ao uso da literatura, a Vida-Média tem sido determinada pela análise do número de citações feitas a um determinado item. Burton e Kebler (1960) consideram a analogia entre a vida-média de substâncias radioativas e a razão de obsolescência da literatura científica. Os autores mencionam que a vida-média da literatura científica refere-se ao tempo requerido para a obsolescência da literatura simplesmente por tornar-se não usada, em contraposição ao que ocorre com as substâncias radioativas, que se tornam não usáveis devido ao processo de desintegração das substâncias. Eles investigam a vida-média e a obsolescência de nove áreas científicas, coletando dados a partir da contagem de citações em fontes selecionadas da literatura periódica nesses campos do conhecimento. Os autores apontam ainda a possibilidade da literatura periódica de uma área ser composta de dois ou mais tipos de literatura, cada qual com sua vida-media própria. A literatura periódica clássica, por exemplo, tem vida-média mais longa que a literatura denominada efêmera, tais como a composta por artigos encontrados nas publicações do gênero weekly news. A literatura das ciências estáveis tem vida-média mais longa do que a literatura que passa por maiores mudanças no conteúdo e nas técnicas. Miranda Pao (1989), ao analisar as citações na área de química,

acrescenta que metade das referências citadas na literatura relaciona-se a artigos com menos de oito anos de publicação, enquanto que, na matemática, a Vida-Média de uma unidade da literatura é de aproximadamente vinte anos de publicação.

4.5 LEI DO ELITISMO

A Lei do Elitismo enuncia que toda população de tamanho N tem uma elite efetiva tamanho \sqrt{N} (PRICE, 1965). Miranda Pao (1989) menciona que Price (1963), Crawford (1971), Crane (1972), Cole e Cole (1972) e Griffith e Mullins (1972) encontraram, em seus estudos, evidências de elites e elitismo na ciência. ~~(PAO, 1989)~~. Price (1965) ao investigar as citações em uma área de assunto considera que somente uma parte da literatura existente é estreitamente entrelaçada e identifica, nesse pequeno grupo de artigos, a frente de pesquisa de uma área do conhecimento. Yablonsky (1980) apresenta um estudo sobre as principais regularidades relacionadas à distribuição da produtividade científica, do ponto de vista de duas abordagens importantes: frequência e classificação em um ranking. O autor estuda a relação entre essas abordagens e apresenta a dedução de um número de fórmulas matemáticas significantes para a compreensão da informação e dos dados estudados.

4.6 TEORIA EPIDÊMICA DE GOFFMAN

A Teoria Epidêmica de Goffman fundamenta-se na analogia entre a transmissão de uma doença infecciosa, no decorrer de um processo epidêmico, e a transmissão e desenvolvimento de idéias, informações registradas, em uma comunidade científica (GOFFMAN; NEWILL,1964). Segundo esse modelo, as ideias científicas são materiais infecciosos, no curso de uma epidemia intelectual, transmitidas, por exemplo, por comunicações diretas, entre um conferencista e o público, ou por meio de conversações. Essas ideias podem também ser expostas por um autor, em um artigo de periódico, para um determinado público. A análise matemática de Goffman foi capaz de prognosticar as condições de controle da epidemia, a razão de crescimento e de declínio, de uma dada área do conhecimento, e permitiu definir as condições sob as quais a epidemia declinaria e

se tornaria estável (PAO, 1989). Segundo Goffman (1966), sua teoria possibilita estimar os níveis de importância de linhas de pesquisa, em uma área específica, assim como prognosticar os seus comportamentos.

Em síntese, as citações possibilitam a busca e o acesso ao artigo citado, indicando o uso desse artigo pelo artigo citante. Pela análise de citações é possível estimar, por exemplo, o número de artigos, títulos de periódicos, autores, departamentos, universidades, institutos de pesquisa, mencionados na literatura. Smith (apud PAO, 1989) destaca a citação como meio de quantificação, aceitando como premissa que a citação normalmente indica a relação de assunto entre o artigo citado e o artigo citante, ainda que essa relação não seja especificada.

Basicamente, a análise de citações tem três grandes campos de aplicação: bibliotecas (gestão de coleções de periódicos, busca e recuperação da informação etc), ciência (mapeamento, prognosticação e avaliação do desempenho dos autores, áreas de conhecimento e instituições, entre outros) e administração (utilização de indicadores bibliométricos para orçamento e investimento em bibliotecas, financiamento de pesquisa, auxílio e bolsas etc).

4.7 LEI DOS 80/20

Finalmente, embora não baseada na análise de citações, vale destacar a Lei dos 80/20, que consiste em um fenômeno, inicialmente observado no comércio e na indústria, segundo o qual, em sistemas de recuperação da informação, 80% da demanda de informação são supridos com 20% do conjunto de fontes de informação. (TRUESWELL, 1969). Em sistemas de informação, esta lei pode ser utilizada nas tomadas de decisão relacionadas à composição e redução de acervos.

Segue o quadro 4, adaptado de Guedes e Borschiver (2005), que relaciona as principais leis e princípios bibliométricos, seus focos de estudo e suas principais aplicações.

Ciência da Informação		
Bibliometria		
Leis e Princípios	Focos de Estudo	Principais Aplicações
Lei de Bradford	Títulos de Periódicos	Estimativa dos graus relativos de relevância de títulos de periódicos em área (s) específica (s) do conhecimento.
Lei de Lotka	Autores	Estimativa dos graus relativos de relevância de autores em área (s) específica(s) do conhecimento.
Leis de Zipf	Palavras	Análise conceitual da escrita científica e indexação automática, ou semiautomática, de artigos científicos.
Ponto de Transição de Goffman	Palavras	Análise conceitual da escrita científica e indexação automática, ou semiautomática, de artigos científicos e tecnológicos.
Colégios Invisíveis	Citações	Identificação da elite de pesquisadores em área(s) específica (s) do conhecimento.
Fator de Imediatismo ou Fator de Impacto	Citações	Estimativa dos graus relativos de relevância de artigos, cientistas e títulos de periódicos científicos em área (s) específica (s) do conhecimento.
Acoplamento Bibliográfico	Citações	Estimativa do(s) grau(s) relativo(s) de ligação de dois ou mais artigos (análise retrospectiva).
Co-citação	Citações	Estimativa do(s) grau(s) relativo(s) de ligação de dois ou mais artigos (análise prospectiva).
Obsolescência da Literatura	Citações	Estimativa do grau de declínio da literatura em área (s) específica (s) do conhecimento.
Vida média	Citações	Estimativa da vida média de uma unidade da literatura em área (s) específica (s) do conhecimento.
Teoria Epidêmica de Goffman	Citações	Estimativa do grau de crescimento e de declínio de uma área de assunto, bem como da importância de linhas de pesquisa em área (s) específica (s) do conhecimento.
Lei do Elitismo	Citações	Identificação e descrição da elite formada por autores que participaram intensivamente da produção científica em área (s) específica (s) do conhecimento.
Frente de Pesquisa	Citações	Identificação de um padrão de relações múltiplas entre autores que se citam na literatura e reconhecimento dos Colégios Invisíveis em área (s) específica (s) do conhecimento.
Lei dos 80/20	Demanda de informação	Processos de tomada de decisão ligados à composição, ampliação e redução de acervos em sistemas de recuperação da informação.

Quadro 4 Leis e princípios bibliométricos, focos de estudo e principais aplicações.

Fonte: Adaptado de Guedes e Borschiver (2005).

Enfim, as leis e princípios bibliométricos são enunciados a partir da observação de fenômenos que se repetem com frequência, em um domínio específico do conhecimento.

A generalidade no comportamento, observada em suas aplicações, é que os caracteriza

como leis e princípios, considerando suas possíveis limitações. O conceito fundamental à aplicação de tais leis e princípios é a probabilidade.

5. CONSIDERAÇÕES FINAIS

A Bibliometria é uma ciência constituída por leis e princípios empíricos estatísticos que contribuem para o estabelecimento da fundamentação teórica da área de Ciência da Informação. A aplicação das leis e princípios bibliométricos possibilita a produção de diferentes indicadores de grande relevância para o tratamento e a gestão da informação e do conhecimento, especialmente em sistemas de recuperação da informação, de comunicação e de avaliação científica.

Pode-se observar, ao longo deste estudo, a diversidade de leis, princípios e conceitos que constituem a Bibliometria. As Leis de Bradford, Lotka e Zipf, focando respectivamente a produtividade de periódicos, a produtividade de autores e a frequência de ocorrência de palavras são consideradas as de maior interesse, do ponto de vista da dinâmica dos processos operados no âmbito dos sistemas de recuperação da informação. Por outro lado, a análise de citações, em gêneros textuais acadêmicos, é um dos métodos mais utilizados principalmente para as tomadas de decisão na política científica e tecnológica. Sua aplicação possibilita o diagnóstico, mapeamento e prognóstico dos fenômenos que norteiam, sobretudo, a comunicação científica e o desenvolvimento do conhecimento de uma instituição, área do conhecimento ou país. Portanto, os indicadores bibliométricos são também de grande importância para a avaliação, planejamento e gestão da ciência e da tecnologia como um todo.

Em suma, as abordagens teóricas, práticas e descritivas da Bibliometria contribuem sobretudo para a adequada mensuração da informação e produção de indicadores, que funcionam como ferramenta de relevância crucial para a cadeia de tomadas de decisão em sistemas de recuperação da informação, de avaliação e de comunicação científica; otimizando, em última análise, o gerenciamento de recursos e de resultados pretendidos, no âmbito da gestão da informação e do conhecimento científico, entre indivíduos, disciplinas, organizações e países.

REFERÊNCIAS

ARAÚJO, C. A. Bibliometria: evolução histórica e questões atuais. *Em Questão*, Porto Alegre, v.

12, n. 1, p. 11-32, jan./jun. 2006.

Disponível em:

<http://revistas.univerciencia.org/index.php/revistaemquestao/article/viewFile/3707/3495>

Acesso em: 07/03/2011.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 10520: informação e documentação. Citações em documentos: apresentação*. Rio de Janeiro, 2002. 7 p.

BAR-ILAN, J. Informetrics at the beginning of the 21st century. A review. *Journal of Informetrics*,

v. 2, n. 1, p. 1-52, 2008.

Disponível em: <http://www.sciencedirect.com/science/article/pii/S1751157707000740>

Acesso em: 07/03/2011.

BASÍLIO, M. M. de P. Interface Linguística e Ciência da Informação: potencialidade na análise de estruturas lexicais. In: REUNIÃO BRASILEIRA DE CIÊNCIA DA INFORMAÇÃO, 2., 1979, Rio

de Janeiro. *Trabalhos apresentados...* Rio de Janeiro: IBICT, 1979.

BASÍLIO, M. M. de P.; BRAGA, L. M.; PIEROTTI, M. de L. C. Estrutura de textos científicos em língua portuguesa: estudo bibliométrico-linguístico. In: REUNIÃO BRASILEIRA DE CIÊNCIA DA INFORMAÇÃO, 2., 1979, Rio de Janeiro. *Anais...* Rio de Janeiro: IBICT, 1979.

BAZERMAN, C. Scientific writing as a social act: a review of the literature of the sociology of science. In: ANDERSON, P. V.; BROCKMAN, R. J. *New Essays in Technical Writing and Communication*. Farmingdale: Baywood, 1983.

Disponível em:

<http://education.ucsb.edu/bazerman/chapters/documents/Bazerman1983ChapterSociologyofScience.pdf>

Acesso em: 27/03/2012.

BENSMAN, Stephen J. Garfield and the impact factor. In: CRONIN, B. *Annual Review of Information Science and Technology*, v. 41. Medford, NJ: Information Today, Inc., 2007. p. 93–155.

Disponível em: <http://onlinelibrary.wiley.com/doi/10.1002/aris.2007.1440410110/pdf>

Acesso em: 27/03/2012.

BOOTH, A. D. A "law" of occurrences for words of low frequency. *Information and Control*, [s.l.], v. 10, n. 4, p. 386-393, April 1967.

Disponível em: [http://dx.doi.org/10.1016/S0019-9958\(67\)90201-X](http://dx.doi.org/10.1016/S0019-9958(67)90201-X)

Acesso em: 27/03/2012.

BRADFORD, S. C. Sources of information on specific subjects. *Engineering*, [s.l.], v.137, p. 85-86, 1934.

BRAGA, G. M. A representação da informação na desconstrução do contexto, *Informare*, Rio de Janeiro, v. 2, n. 2, p. 53-57, jul./dez. 1996.

Disponível em: <http://ibict.phlnet.com.br/anexos/bragav2n2.pdf>

Acesso em: 27/03/2012.

BRAUN, T.; SCHUBERT, A. A quantitative view on the coming of age of interdisciplinarity in the sciences 1980-1999, *Scientometrics*, Budapest, v. 58, n. 1, p. 183-189, sept. 2003.

Disponível em: <http://www.springerlink.com/content/q80v88u45642662l/fulltext.pdf>

Acesso em: 24/11/2009.

BROOKES, B. C. Bradford's law and the bibliography of science. *Nature*, [s.l.], v.224, p. 953-956, Dec. 1969.

Disponível em: <http://www.nature.com/nature/journal/v224/n5223/pdf/224953a0.pdf>

Acesso em: 27/03/2012.

BURTON, R. E.; KLEBER, R. W. The "half -life" of some scientific and technical literatures. *American Documentation*, [s.l.], v. 11, n.1, p. 18-22, Jan. 1960.

Disponível em: <http://onlinelibrary.wiley.com/doi/10.1002/asi.5090110105/pdf>

Acesso em: 27/03/2012.

BYBEE, J. L.; HOPPER, P. *Frequency and the emergence of language structure*. Amsterdam: John Benjamins, 2001.

Disponível em: <http://books.google.com.br/books?id=6-bQUClxn1IC&printsec=frontcover&hl=pt-BR#v=onepage&q&f=false>

Acesso em: 27/03/2012.

CHEN, Y.; CHONG, P. P.; TONG, M. Y. The Simon-Yule approach to bibliometric modeling. *Information Processing and Management*, v. 30, n.4, p. 535- 556, 1994.

Disponível em: <http://www.sciencedirect.com/science/article/pii/0306457394900388>

Acesso em: 27/03/2012.

CHENG, K. H. Automatic identification for topics of electronic documents. *Bulletin of Library Association of China*, v. 59, p. 43-58, Dec. 1997.

CHENG, Y. S. Booth's Law of word frequency. *Journal of American Society for Information*

Science, v. 41, n. 5, p. 387-388. July 1990.

Disponível em:

[http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)10974571\(199007\)41:5%3C387::AID-ASI10%3E3.0.CO;2-I/pdf](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)10974571(199007)41:5%3C387::AID-ASI10%3E3.0.CO;2-I/pdf)

Acesso em: 27/03/2012.

FAIRTHONE, R. A. Empirical hyperbolic distribution (Bradford-Zipf-Mandelbrot) for bibliometric description and prediction. *Journal of Documentation*, [s.l.], v. 25, n. 4, p. 319-343, Dec. 1969.

Disponível em: <http://garfield.library.upenn.edu/classics1987/A1987F496000001.pdf>

Acesso em: 27/03/2012.

GARFIELD, E. Citation analysis as a tool in journal evaluation, *Science*, v. 178, p. 471-479.

1972a. Disponível em: <http://www.garfield.library.upenn.edu/essays/V1p527y1962-73.pdf>

Acesso em: 10/02/2010.

GARFIELD, E. "Citations-to" divided by "items-published" gives journal impact factor; ISI lists the top fifty high-impact journals of science, *Essays of an Information Scientist*, v. 1, p. 270-273. Feb. 1972b.

Disponível em: <http://www.garfield.library.upenn.edu/essays/V1p270y1962-73.pdf>.

Acesso em: 10/02/2010.

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

Disponível em: <http://www.nature.com/nature/journal/v212/n5061/pdf/212449a0.pdf>

Acesso em: 27/03/2012.

GOFFMAN, W.; MORRIS, T. G. Bradford's Law and Library Acquisitions ~~Bradford's law applied to the maintenance of library collections~~. *Nature*, [s.l.], v. 226, n. 5249, p. 922-923, June 1970.

Disponível em: <http://www.nature.com/nature/journal/v226/n5249/pdf/226922a0.pdf>

Acesso em: 27/03/2012.

GOFFMAN, W.; NEWILL, V. A. Generalization of epidemic theory: an application to the transmission of ideas. *Nature*, [s.l.], v. 204, n. 4955, p. 225-228, Oct. 1964.

Disponível em: <http://www.nature.com/nature/journal/v204/n4955/pdf/204225a0.pdf>

Acesso em: 27/03/2012.

GROSS, P. L. K.; GROSS, E. M. College libraries and chemical education. *Science, New Series*, v. 66, Issue 1713, p. 385-389, Oct. 1927.

Disponível em:

http://www.garfield.library.upenn.edu/papers/grossandgross_science1927.pdf

Acesso em: 27/03/2012.

GUEDES, V. L. da S. *Um estudo para indexação automática de textos de mecânica dos solos, Engenharia Civil*. 20 mar. 1992. Dissertação (Mestrado em Ciência da Informação) – CNPq/IBICT-UFRJ/ECO, Rio de Janeiro, 1992.

GUEDES, V. L. da S. *Nominalizações deverbais em artigos científicos: uma contribuição para a análise e a indexação temática da informação*, 2010. ... f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

GUEDES, V. L. da S.; 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 eletrônico...* Salvador: ICI/UFBA, 2005. Disponível em: http://www.cinform.ufba.br/vi_anais/trabalhos.htm
Acesso em: 26/10/2011

GUEDES, V. L. da S.; VALOIS, E. C. *Adequação das Leis de Zipf (1a e 2a) e Ponto T de Goffman à indexação de documentos científicos: uma aplicação em Mecânica dos Solos (engenharia civil)*. Rio de Janeiro, 1988. Trabalho não publicado apresentado á Disciplina de Bibliometria da ECO/UFRJ. P.3.

HREBICEK, L. Zipf's law and text, *Glottometrics*, Czech Republic, v. 3, n. 22, p. ~~23~~27-38, 2002.
Disponível em: <http://www.arteuna.com/talleres/lab/ediciones/libreria/Glottometrics-zipf.pdf>
Acesso em: 27/03/2012.

HYLAND, K. Academic attribution: citation and the construction of disciplinary knowledge. *Applied Linguistics*, v. 20, Issue 3, p. 341-367, 1999.
Disponível em: <http://apliij.oxfordjournals.org/content/20/3/341.full.pdf>
Acesso em: 27/03/2012.

INTERNATIONAL SOCIETY OF SCIENTOMETRICS AND INFORMETRICS (ISSI), *Scope of 13th ISSI Conference 2011*.
Disponível em: <http://www.issi2011.uzulu.ac.za/index.php/announcement>
Acesso: 05/03/2011.

JONES, A. W. Impact factors of forensic science and toxicology journals: what do the numbers really mean? *Forensic Science International*, v. 133, n. 1-2, p. 1-8, April 2003.
Disponível em: <http://www.sciencedirect.com/science/article/pii/S0379073803000422>
Acesso: 05/04/2005.

KESSLER, M. An experimental study of bibliographic coupling between technical papers, *IEEE Transactions on Information Theory*, v. 9, n. 1, p. 49-51, Jan. 1963.

Disponível em: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1057800>
Acesso: 27/03/2012.

LINE, M. B. The "half-life" of periodical literature: apparent and real absolescence. *Journal of Documentation*, [s.l.], v.26, n.1, p. 48-54, Mar. 1970.

LUHN, H. P. A statistical approach to mechanized encoding and searching of literary information. *IBM Journal of Research and Development*, New York, v. 1, n.4, p. 309-317, Oct. 1957

Disponível em: http://www.phil-fak.uni-duesseldorf.de/fileadmin/Redaktion/Institute/Informationswissenschaft/downloadcenter/infocenter/Informationretrieval/Luhn_1957_statistical_approach.pdf
Acesso: 27/03/2012.

LOTKA, A. J. The frequency of distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, v. 16, n.12, p. 317-323, 1926.

MAMEDE, M. E. de O.; PASTORE, G. M. Avaliação da produção dos compostos majoritários da fermentação de mosto da uva por leveduras isoladas da região da "Serra Gaúcha" (RS). *Ciência e Tecnologia de Alimentos*, Campinas, v. 24, n. 3, p. 453-458, jul./set. 2004.

Disponível em: <http://www.scielo.br/pdf/cta/v24n3/21942.pdf>
Acesso em: 27/05/2009

MARSHAKOVA, I. V. Citation networks in information science. *Scientometrics*, [s.l.], v. 31, n. 1, p. 13- 26, 1981.

MARTYN, J.; LANCASTER, F. W. *Investigative methods in library and information science*. Arlington: Information resources Press, 1981.

MAUCERI, C. L'anal se de corpus textuels à l'aide de techniques linguistiques et statistiques au service de la construction deréférentiels et du développement des s stèmes d'indexation automatique. In : JOURNÉES D'ETUDES ORGANISÉES PAR LA SOCIÉTÉ FRANÇAISE DE BIBLIOMÉRIE APPLIQUÉE AVEC LE CONCOURS. *Resumos...* Ilê Rousse: Société Française de Bibliométrie Appliquée, 1995.

Disponível em: <http://crrm.u-3mrs.fr/sfba/ile-rousse/1995/cd-rom.pdf>
Acesso em: 27/06/2007.

MERTON, R. K. The Mathew effect in science. *Science*, [s. l.], v. 159, n. 3810, p. 56-63, Jan. 1968.

Disponível em: <http://www.garfield.library.upenn.edu/merton/matthew1.pdf>
Acesso em: 27/03/2012.

OKUBO, Y. Bibliometric indicators and analysis of research systems: methods and examples. Paris: OECD Publishing, 1997. *Science, Technology and Industry Working Papers*, 1997/01.

Disponível em:

[http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=OCDE/GD\(97\)41&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=OCDE/GD(97)41&docLanguage=En)

Acesso em: 27/03/2012.

PAO, M. L. Automatic text analysis based on transition phenomena of word occurrences. *Journal of the American Society for Information Science*, New York, v. 29, n.3, p. 121-124, May 1978.

Disponível em: <http://onlinelibrary.wiley.com/doi/10.1002/asi.4630290303/pdf>

Acesso em: 27/03/2012.

PAO, M. L. *Concepts of information retrieval*. Englewood, Colorado: Libraries Unlimited, Inc., 1989. 285 p.

PINHEIRO, L. V. R. Lei de Bradford: uma reformulação conceitual, *Ciência da Informação*, Brasília, v. 12, n. 2, p. 59-80, jul./dez. 1983.

Disponível em: <http://revista.ibict.br/ciinf/index.php/ciinf/article/download/1498/1116>

Acesso em: 27/03/2012.

POLANCO, X.; GRIVEL, L. ; RO AUTÉ, J. Infométrie et linguistique informatique: une approche linguistico-infométrique au service de la veille scientifique et technologique. In : JOURNÉES D'ÉTUDES ORGANISÉES PAR LA SOCIÉTÉ FRANÇAISE DE BIBLIOMÉRIE APPLIQUÉE AVEC LE CONCOURS. *Resumos...* Ilê Rouse: Société Française de Bibliométrie Appliquée, 1995.

Disponível em: <http://crrm.u-3mrs.fr/sfba/ile-rousse/1995/cd-rom.pdf>

Acesso em: 27/08/2007.

PRICE, D. J. de S. *Little science, big science*. New York: Columbia University Press, 1963.

Disponível em: <http://www.garfield.library.upenn.edu/classics1983/A1983QX23200001.pdf>

Acesso em: 27/03/2012.

PRICE, D. J. de S. Networks of scientific papers. *Science*, [s.l.], v. 149, n. 3683, p. 510-515, July 1965.

Disponível em: <http://www.garfield.library.upenn.edu/papers/pricenetworks1965.pdf>

Acesso em: 27/03/2012.

PRITCHARD, A. Statistical bibliography or bibliometrics? *Journal of Documentation*, [s. l.], v. 25, n.4, p. 348-349, Dec. 1969.

RAISIG, L. M. Statistical bibliography in the health sciences. *Bull Medical Libr Association*,

v. 50, n. 3, p. 450–461, July 1962. Disponível em:
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC197860/pdf/mlab00192-0151.pdf>
Acesso em: 06/03/2011.

ROUAULT, J. *Linguistique automatique: applications documentaires*. Berne: Peter Lang, 1987.

SANTOS, M. J. V. da C. *A política brasileira de informação em Geociências e Tecnologia Mineral: uma contribuição e sua implementação via análise bibliométrica da literatura*. 13 nov. 1992. Dissertação (Mestrado em Ciência da Informação), UFRJ/ECO, Rio de Janeiro. Orientadora: Gilda Maria Braga.

SANTOS, R. N. M. dos; KOBASHI, N. Y. Bibliometria, Cientometria, Infometria: conceitos e aplicações. *Tendências da Pesquisa Brasileira em Ciência da Informação*, v.2, n.1, jan.-dez, 2009. Disponível em: <http://inseer.ibict.br/ancib/index.php/tpbci/article/view/21/43>
Acesso em: 09/10/2009.

SARACEVIC, T. *Bibliometrics*. New Jersey: Rutgers University, 2004. 43 slides. Disponível em:
<http://comminfo.rutgers.edu/~tefko/Courses/610/Lectures/Bibliometrics.ppt>
Acesso em: 27/10/2011.

SARACEVIC, T.; GARFIELD, E. On measuring the publication productivity and citation impact of a scholar: A case study. In: LARSEN, B.; SCHNEIDER, J. W.; ÅSTRÖM, F. *The Janus Faced Scholar: A Festschrift in Honour of Peter Ingwersen*. Special volume of the e-zine of the International Society for Scientometrics and Informetrics, v. 06-S, June 2010. Disponível em: <http://garfield.library.upenn.edu/papers/ingwersenfestschrift2010.pdf>
Acesso em: 27/03/2012.

SMALL, H. Co-citation in the scientific literature: a new measure of the relationship between two documents, *Journal of the American Society for Information Science*, v. 24, n. 4, p. 265–269, Jul./Aug. 1973. Disponível em: <http://www.garfield.library.upenn.edu/essays/v2p028y1974-76.pdf>
Acesso em: 27/03/2012.

SWALES, J. M. Citation analysis and discourse analysis, *Applied Linguistics*, v. 7, n. 1, p. 39–56, 1986. Disponível em: http://applied.oxfordjournals.org/cgi/pdf_extract/7/1/39
Acesso em: 01/06/2009.

TAGUE-SUTCLIFFE, J. An introduction to informetrics. *Information Processing & Management*, v. 28, issue 1, p.1-3, Jan./Feb. 1992. Disponível em: <http://www.sciencedirect.com/science/article/pii>

Acesso em: 27/10/2011.

TRUESWELL, R. W. Some behavioral patterns of library users: the 80/20 rules. *Wilson Library Bulletin*, [s.l.], v. 43, n. 5, p. 458-461, 1969.

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, mai./ago. 2002.

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

Acesso em: 29/07/2009.

VOOS, H. Lotka and information science. *Journal of the American Society of Information Science*, New York, v. 25, p. 270-272, July/Aug. 1974.

Disponível em: <http://onlinelibrary.wiley.com/doi/10.1002/asi.4630250410/pdf>

Acesso em: 27/03/2012.

YABLONSKY, A. On fundamental regularities of the distribution of scientific productivity. *Scientometrics*, [s.l.], v. 2, n. 1, p. 3-34, 1980.

ZIMAN, J. M. Information, communication and knowledge, *Nature*, v. 224, n. 5217, p. 318-324, Oct. 1969.

Disponível em: <http://www.nature.com/nature/journal/v224/n5217/pdf/224318a0.pdf>

Acesso em: 27/03/2012.

ZIPF, G. K. *Human behavior and the principle of least effort*. Cambridge, Ma: Addison Wesley, 1949.

Agradecimentos

A Cleber Villaça da Silva pela colaboração no processamento de buscas bibliográficas, formatação e normalização dos textos, figuras e referencias apresentados no artigo.