



XVI Encontro Nacional de Pesquisa em Ciência da Informação (XVI ENANCIB)
ISSN 2177-3688

GT 4 – Gestão da Informação e do Conhecimento

Comunicação Oral

**ANÁLISE DA PRODUÇÃO CIENTÍFICA SOBRE INTELIGÊNCIA
DE NEGÓCIOS NA WEB OF SCIENCE (WOS)¹**

***REVIEW OF SCIENTIFIC LITERATURE ON BUSINESS
INTELLIGENCE ON THE WEB OF SCIENCE (WOS)***

Alexandre Lucas, UFSC
alexlucas.al@gmail.com

Angel Freddy Godoy Viera, UFSC
godoy@cin.ufsc.br

Adilson Luiz Pinto, UFSC
adilson@cin.ufsc.br

Resumo: O artigo apresenta uma visão da produção científica sobre Inteligência de Negócios. Identifica os principais autores, as citações, a cronologia da produção, as palavras-chave e os principais periódicos utilizados. A análise é obtida através de estudo bibliométrico dos artigos indexados na base de dados Web of Science (WoS) sobre o termo Inteligência de Negócios. Ele compreende somente artigos indexados até o ano de 2013, onde foram utilizados para a busca os campos: título ou resumo ou palavra-chave. O estudo foi desenvolvido como parte da dissertação no programa de Pós-graduação em Ciência da Informação (PGCin) da Universidade Federal de Santa Catarina (UFSC). O estudo, além de gerar informações para o conhecimento da dinâmica e evolução da pesquisa sobre o tema, promove indicadores para gestão informacional considerando a Bibliometria como uma importante ferramenta de aplicação nas organizações que trabalhem com Pesquisa, Desenvolvimento e Inovação (PD&I) como os Institutos de Ciência, tecnologia e Inovação (ICTIs). A Bibliometria poderia inclusive ser utilizada pelos próprios sistemas de Inteligência de Negócios como parte integrante dos sistemas.

Palavras-chave: Inteligência de Negócios. Bibliometria. Web of Science. Gestão da Informação.

Abstract: The article presents an overview of the scientific production published on Business

¹ O conteúdo textual deste artigo, os nomes e e-mails foram extraídos dos metadados informados e são de total responsabilidade dos autores do trabalho.

Intelligence. It identifies the main authors, citations, and the chronology of the production, the keywords and the main specialized journals on the field. The analysis was obtained through bibliometric indicators of articles indexed in the Web of Science database (WoS) regarding the term Business Intelligence and only includes articles indexed up to the year of 2013. The following search fields were used: title or summary or keyword. The study was developed as part of the Master's thesis in a graduate program in Information Science (PGCin) at the Federal University of Santa Catarina (UFSC). The study besides generating information for the understanding of the dynamics and evolution of research on the topic, provides indicators for management informational whereas the Bibliometrics as an important tool for application in organizations that work with Research, Development and Innovation (R&D&I) as Science, Technology and Innovation Institutes. The Bibliometrics could also be used for the own systems of Business Intelligence as an integral part of the systems.

Keywords: Business Intelligence. Bibliometrics. Web of Science, Information Management.

1 INTRODUÇÃO

A informação possui papel importante na sociedade atual. Ela está estruturada em uma Sociedade da Informação e desde o advento da Internet vive o fenômeno de uma Sociedade em Rede (CASTELLS, 2003). Por outro lado, com a globalização acirrando a competitividade em diversas dimensões, aliado ao fato do uso de computadores e sistemas informatizados de forma massiva e intensiva, a informação ganha um reforço no seu posicionamento como elemento estratégico e importante para diversos integrantes da Sociedade. Notoriamente, as empresas fazem parte deste universo, mas também governos e centros de pesquisa e desenvolvimento tecnológicos. Inteligência de Negócios surge, então, para, a partir de uma montanha de dados disseminados em diversas bases, permitir a transformação da informação em conhecimento e apoiar o processo dos decisores, dentro de seu papel funcional, e ganhar vantagem competitiva para sua organização (TODESCO et al, 2007).

Estudar a produção científica sobre Inteligência de Negócios passa a ser relevante, pois como qualquer nova área é a Pesquisa e o Desenvolvimento (P&D) que permitem sua evolução e sua inovação. Uma das melhores maneiras para analisar a produção científica se dá pelos estudos bibliométricos. A bibliometria é capaz de gerar indicadores objetivos da produção científica permitindo sua visualização e mapeamento científico. A Web of Science (WoS) tem se constituído um dos mais importantes indexadores de periódicos científicos e sua estrutura concebida para estudos bibliométricos apresenta uma das melhores plataformas para uso em diferentes *softwares* com esta missão. O objetivo deste artigo é apresentar uma análise bibliométrica dos artigos indexados na base de dados Web of Science (WoS) sobre o termo Inteligência de Negócios. O trabalho não abrange todos os estudos, mas contribui com algumas informações quantitativas e situacionais da área. Este estudo foi desenvolvido como parte da dissertação de mestrado, já defendida no programa de Pós-Graduação em Ciência da

Informação (PGCIN) da Universidade Federal de Santa Catarina (UFSC).

2 INDICADORES BIBLIOMETRICOS E WEB OF SCIENCE (WoS)

Segundo De Fillipo e Fernández (2002 apud PINTO, 2013, p.20), os indicadores bibliométricos são um conjunto de leis e princípios empíricos que contribuem para estabelecer os fundamentos teóricos das ciências, podendo ser aplicados na identificação da produtividade de autores, identificação dos periódicos mais relevantes, autores mais citados, etc. Neste trabalho foram utilizados três importantes indicadores bibliométricos como base para os estudos: Lei de Lotka, Lei de Zipf e Lei de Bradford.

A Lei de Lotka (LOTKA, 1926) está relacionada à produtividade dos autores e, segundo Guedes e Borschiver (2005, p.5), ela:

[...] 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, em “Little Science, Big Science 2, 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 $1/n^3$.

É importante notar que o primeiro passo para aplicação da Lei de Lotka é a quantificação do número de autores que tiveram somente uma contribuição, pois é a partir do seu produto com o fator de 1 dividido por “ n ” contribuições ao quadrado é que se obtém a quantidade de autores para este mesmo “ n ”. Outra observação é que o expoente “ao quadrado” é na verdade uma constante que depende da área científica e também tem sido objeto de diversos estudos (URBIZAGASTEGUI, 2008; MARTÍN SOBRINO; PESTANA CALDES; PULGARÍN GUERRERO, 2008).

A Lei de Zipf (ZIPF, 1949) está relacionada à ocorrência de determinadas palavras em um determinado conjunto de textos, podendo ser utilizada também na análise de palavras-chave. Segundo Guedes e Borschiver (2005, p.6):

Zipf observou que, num texto suficientemente longo, existia uma relação entre a frequência que uma dada palavra ocorria e sua posição 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 (*ranking*). Assim, a palavra de maior frequência de ocorrência tem ordem de série 1, a de segunda maior frequência de ocorrência, ordem de série 2 e, assim, sucessivamente. Zipf observou,

também, que o produto da ordem de série (r) de uma palavra, pela sua frequência de ocorrência (f) era aproximadamente constante (c). Enunciou assim que $r \cdot f = c$, o que ficou conhecido como Primeira Lei de Zipf.

Para esta pesquisa utilizou-se a Lei de Zipf na avaliação da ocorrência de palavras-chave.

A Lei de Bradford (BRADFORD, 1934) está relacionada aos periódicos mais produtivos e visa sua identificação baseado no princípio da atração de mais artigos sobre o assunto. Segundo Guedes e Borschiver (2005, p.4):

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 subsequentes, variará na proporção 1:n:n².

A Web of Science conhecida também como WoS é uma base de dados da Thomson Reuters. Segundo o Portal CAPES (2014):

Ela é uma Base de dados referencial com resumos em todas as áreas do conhecimento. A base referencial multidisciplinar, Web of Science, que está integrada à base ISI Web of Knowledge. Oferece ferramentas para análise de citações, referências, índice h , permitindo análises bibliométricas. Cobre aproximadamente 12.000 periódicos. A assinatura deste conteúdo, oferece a possibilidade de consulta a cinco coleções: Science Citation Index Expanded (SCI-EXPANDED) - com disponibilidade de acesso desde 1945 até o presente; Social Sciences Citation Index (SSCI) - com disponibilidade de acesso desde 1956 até o presente; Arts & Humanities Citation Index (A&HCI) - com disponibilidade de acesso desde 1975 até o presente; Conference Proceedings Citation Index- Science (CPCI-S) - com disponibilidade de acesso desde 1991 até o presente e Conference Proceedings Citation Index - Social Science & Humanities (CPCI-SSH) - com disponibilidade de acesso desde 1991 até o presente.

Segundo Mugnaini e Strehl (2008, p.95), o *Institute for Scientific Information (ISI)*, atual ISI-Thomson-Scientific, ampliou a cobertura temática de suas bases de dados, e passou a contemplar as Ciências Sociais, com a publicação do Social Science Citation Index (SSCI), e as áreas de Artes e Humanidades, com o Arts and Humanities Citation Index. Em 1997 culminou com a incorporação de sofisticados mecanismos de pesquisa e ampliação da cobertura, na produção de uma interface de consulta única para os três índices, chamada Web of Science (WoS).

Neste sentido, a possibilidade de quantificar as vezes que um artigo indexado na base foi citado por trabalhos publicados posteriormente e, ainda que, para agências de fomento a

manipulação deste tipo de dado é praticamente inviável, os indicadores calculados a partir dos dados das citações indexadas na WoS e agrupados por periódicos no Journal Citation Reports® (JCR) têm sido o recurso mais frequentemente utilizado por administradores da ciência no Brasil e no mundo, para fins de avaliação da produção científica (MUGNAINI; STREHL, 2008, p. 96).

3 PROCEDIMENTOS E OPÇÕES METODOLÓGICAS

O estudo utilizou o sistema de recuperação de informações da WoS e toda as suas facilidades. Nos registros recuperados foram analisados os principais autores, periódicos em que os artigos foram publicados, ano de publicação e citações dos mesmos. Os detalhes das opções e algumas definições são apresentados nos itens subsequentes.

3.1 USO DOS TERMOS

Para a busca utilizou-se o termo em inglês entre aspas *Business Intelligence*. Não se utilizou o termo em português devido à baixa indexação de periódicos em português nesta área. De fato, um teste feito com os termos em português “Inteligência de Negócios” e “Inteligência Empresarial” no campo Tópico do SRI da WoS, resultaram com nenhum registro recuperado. Outro motivo baseia-se no fenômeno da Polissemia, no qual um termo possui diversos conceitos; e da Sinonímia, no qual um conceito possui diversos termos. Este assunto é tratado nos trabalhos de Matheus e Parreiras (2004), no qual para o termo em inglês, um sinônimo pode ser *Competitive Intelligence*, assunto também tratado por Bergeron e Hiller (2002).

3.2 CAMPOS DE BUSCA

Utilizou-se o campo de busca denominado “tópico” que compreende ocorrências nos campos Título, Resumo, Palavras-chave de autor e *Keywords Plus*®. Segundo a página de Ajuda da WoS, *KeyWords Plus*® é um índice de terminologias criadas pela Thomson Reuters nas quais as mesmas são derivadas dos títulos originais da obra. O *KeyWords Plus*® aumenta o número de resultados tradicional de palavras-chave ou títulos.

3.3 TIPO DE DOCUMENTO

Para o objetivo da pesquisa foram selecionados somente os artigos indexados pela WoS. Não foram avaliados os: *proceedings paper*, *book review*, *editorial material* e *review* entre outros de menor ocorrência.

3.4 ANO DE PUBLICAÇÃO

Com relação ao ano de publicação foram considerados todos os estudos publicados entre o período de 1958 a 2013.

3.5 SOFTWARE DE PROCESSAMENTO

Para análise dos dados da pesquisa foram utilizados os recursos de análise da WoS que geram gráficos de porcentagem e de normalização dos dados em conjunto com outros *softwares*. Para a maioria dos processamentos foi utilizado o Excel da Microsoft, versão 2013, para geração e interação dos dados recuperados; para a análise de palavras-chave utilizou-se o Wordle (FEINBERG, 2014), e; para a geração de mapas de relações e de temporalidade fez-se uso do CiteSpace (CHEN, 2006).

3.6 NORMALIZAÇÃO DOS DADOS

Durante a análise dos dados identificou-se o efeito da falta de padronização de nomes de autoridades, palavras-chave, no qual diferenças na grafia podem ocasionar contagens indevidas. Como parte das análises ocorreu utilizando CiteSpace, em alguns casos o sistema normalizou automaticamente. Em outros casos esta normalização foi perceptiva e foram reparadas manualmente.

Também se detectou uma diferença de resultados pela lista de autores abreviados e autores completos fornecidos pela WoS. Assim, decidiu-se o uso dos autores abreviados com uma correlação posterior ao nome completo.

3.7 DEFINIÇÕES PARA USO NAS LEIS BIBLIOMÉTRICAS

Lei de Lotka - Utilizou-se o coeficiente ao quadrado inverso conforme enunciado da Lei.

Lei de Zipf - Para esta pesquisa utilizou-se a Lei de Zipf na avaliação da ocorrência de palavras-chave, na qual as palavras consideradas **estratégicas** foram obtidas pela contagem no Ranking começando pela primeira, onde o valor de contagem foi obtido com a raiz quadrada do número de ocorrências da palavra mais frequente. As consideradas **interessantes** foram contadas a partir da última estratégica e o valor obtido com a raiz quadrada do número de ocorrências da segunda mais frequente. O restante das palavras foi considerado **ruído**. Estas

Zonas de distribuição foram definidas com base em Quoniam et al (2001), exceto as denominadas estratégicas, que foram assim atribuídas por se tratar de uma análise nas palavras-chave, na qual não aparecem as palavras do vocabulário geral (triviais).

Lei de Bradford - Utilizada para análise da produtividade de periódicos, na qual foram definidas três zonas: Concentração, com 33,3% dos artigos, Média produtividade (Prod. Média), com 33,3% dos artigos e, Baixa Produtividade (Prod. Baixa), com os restantes 33,3% dos artigos.

4 RESULTADOS: APRESENTAÇÃO E DISCUSSÃO

Nesta seção apresenta-se e discute-se os resultados da análise bibliométrica dos artigos recuperados conforme explicado no item 3.

4.1 RESULTADOS DA BUSCA NA WOS

O resultado geral da pesquisa na WoS (busca realizada no dia 10/01/2014), para o termo em inglês *Business Intelligence*, no período de 1958 a 2013; recuperou 491 artigos; sendo produzido por 1235 autores, e; publicados em 244 periódicos científicos.

4.2 ANÁLISE BIBLIOMÉTRICA DE INTELIGÊNCIA DE NEGÓCIOS

A investigação consistiu na análise dos autores, análise das citações, análise das palavras-chave e análise dos periódicos.

4.1.1 Análise dos Autores

Foi identificado no universo de artigos recuperados 1235 autores, nos quais quatro destes foram identificados como anônimos. Utilizando a análise da produtividade segundo a lei de Lotka foi elaborada a Tabela 1, que apresenta a relação de contribuições por autor. A análise confirma a máxima “Poucos com muitos e muitos com poucos” (GUEDES; BORSCHIVER, 2005, p.3) conhecida na Ciência como efeito Mateus: “aos que mais têm será dado em abundância e, aos que menos têm, até o que têm lhes será tirado”. (MERTON, 1968 apud GUEDES; BORSCHIVER, 2005, p. 3).

Entretanto, a relação de Lotka (baseada no quadrado inverso de Pareto “20/80”) não se confirmou, na verdade a relação apresentada tem um coeficiente maior que dois e a somatória para esta lei ficou em 11,41% de autores que publicaram 24,72% de toda publicação, muito aquém do quadrado inverso.

Na análise da Tabela é possível observar que enquanto 141 autores apresentam mais de uma contribuição chegando ao máximo de nove contribuições por um único autor, 1094 autores tiveram somente uma contribuição cada, no período estudado, no qual os autores mais produtivos ocupam o *ranking* das 20 primeiras posições, conforme apresenta a Tabela 1.

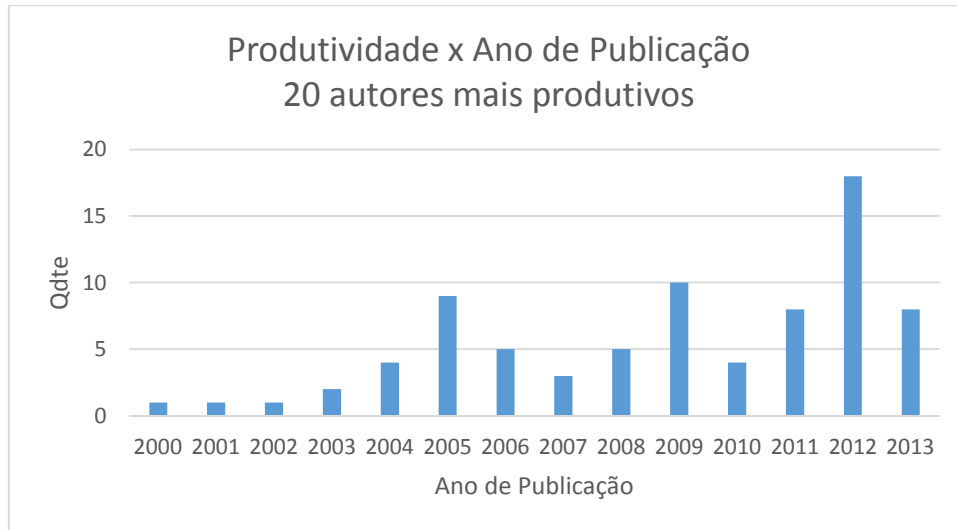
Tabela 1 – Lei de Lotka representada com os autores mais produtivos

Ranking	Autores Abreviados	Autores Completos	Artigos	Porcentagem	Acumulativo
1	CHEN, HC	CHEN, HSINCHUN	9	0,61	0,61%
2	CHUNG, WY	CHUNG, WINGYAN	8	0,54	1,15%
3	RIZZI, S	RIZZI, STEFANO	7	0,48	1,63%
4	GOLFARELLI, M	GOLFARELLI, MATTEO	7	0,48	2,11%
5	SHI, Y	SHI, YONG	6	0,41	2,25%
6	PEDERSEN, TB	PEDERSEN, TORBEN BACH	5	0,34	2,59%
7	DAYAL, U	DAYAL, UMESHWAR	5	0,34	2,93%
8	BEDARD, Y	BEDARD, YVAN	5	0,34	3,27%
9	WIXOM, BH	WIXOM, BARBARA H.	4	0,27	3,54%
10	WANG, H	WANG, HAI	4	0,27	3,81%
11	ZHONG, N	ZHONG, N	4	0,27	4,08%
12	ZHANG, L	ZHANG, GUANGQUAN	4	0,27	4,35%
13	LUFTMAN, J	LUFTMAN, JERRY	4	0,27	4,62%
14	POPOVIC, A	POPOVIC, ALES	4	0,27	4,89%
15	LIU, Y	LIU, YING	4	0,27	5,16%
16	LIU, CN	LIU, CHUNNIAN	4	0,27	5,43%
17	HUSSAIN, FK	HUSSAIN, FAROOKH KHADEER	4	0,27	5,70%
18	LEE, CS	LEE, CHIEN-SING	4	0,27	5,97%
19	DUAN, L	DUAN, LIAN	4	0,27	6,24%
20	CHAU, M	CHAU, MICHAEL	4	0,27	6,51%
21	[ANONYMOUS]		4	0,27	6,78%
22 – 43	22 autores	Com 3 artigos cada	66	4,52	11,30%
44 – 141	98 autores	Com 2 artigos cada	196	13,42	24,72%
142 – 1235	1094 autores	Com 1 artigo cada	1094	75,28	100,00%

Fonte: Elaborado pelos Autores a partir dos dados da pesquisa (2015).

A análise temporal da produção dos 20 principais autores é apresentada no Gráfico 1. É possível identificar que os autores mais produtivos iniciaram sua produção a partir do ano 2000.

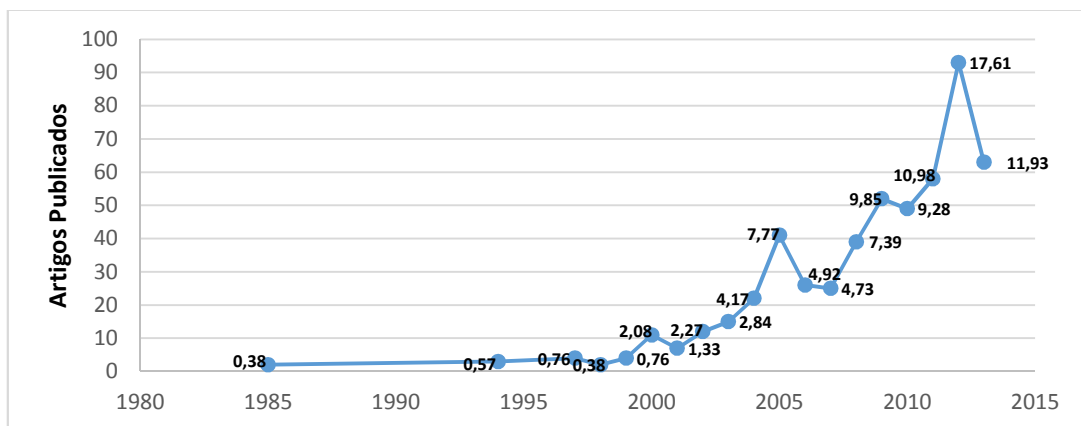
Gráfico 1 – Análise temporal da produção dos Autores mais produtivos



Fonte: Elaborado pelos Autores a partir dos dados da pesquisa (2015)

Além da análise de autoria, também é importante demonstrar a análise temporal de todos os artigos recuperados. Observa-se que existe um crescimento exponencial de produtividade, porém existe uma variação de diminuição científica entre os anos 2006 e 2007 e posteriormente em 2010 e 2013.

Gráfico 2 – Produção por ano de publicação

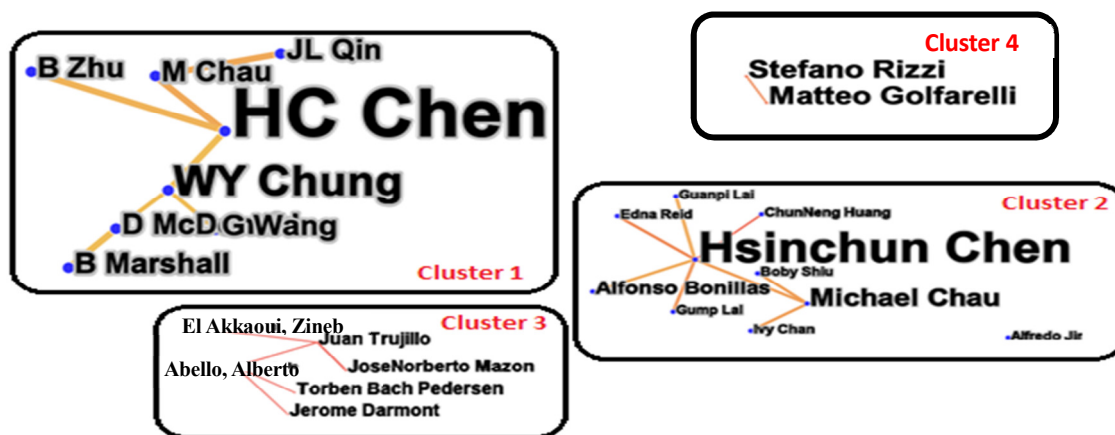


Fonte: Elaborado pelos Autores a partir dos dados da pesquisa (2015).

Relacionada à publicação de autores e ao período de análise, também se observa a disposição dos principais anos de publicação dos autores mais relevantes. Na Figura 1 obtida pelo *software* CiteSpace, observa-se os principais produtores. Em uma análise simplificada de coautoria pode-se notar que o mesmo autor gerou dois agrupamentos em clusters (HC Chen e Hsinchun Chen), onde foram normalizados para a contagem de produção, porém foram mantidos de forma diferente para a geração do grafo (Figura 1).

Pela análise do Excel tem-se que HC Chen produziu cinco artigos em conjunto com WY Chung. Também se tem que Stefano Rizzi e Matteo Golfarelli produziram juntos sete artigos.

Figura 1 – Agrupamento de coautoria de alguns dos principais produtores – CiteSpace



Fonte: Elaborado pelos Autores a partir dos dados da pesquisa (2015).

Neste tipo de análise é fundamental observar-se que existe centralidade nas relações, em especial na forma de poder e estrela, onde o primeiro caso fica evidente nos clusters 1 e 3, havendo liderança para que possa existir a rede (caso de HC Chen “cluster 1” e Juan Trujillo “cluster 3”), pois se os autores principais não existissem dentro das análises, as redes seriam fragmentadas. O outro caso fica por conta do cluster 2, onde existe uma centralidade estrela, com o grau direcionado a Hsinchun Chen (em forma de citação seria HC Chen). Sem este autor este cluster também não existiria. Por fim, temos uma relação de paridade, representada por Stefano Rizzi e Matteo Golfarelli.

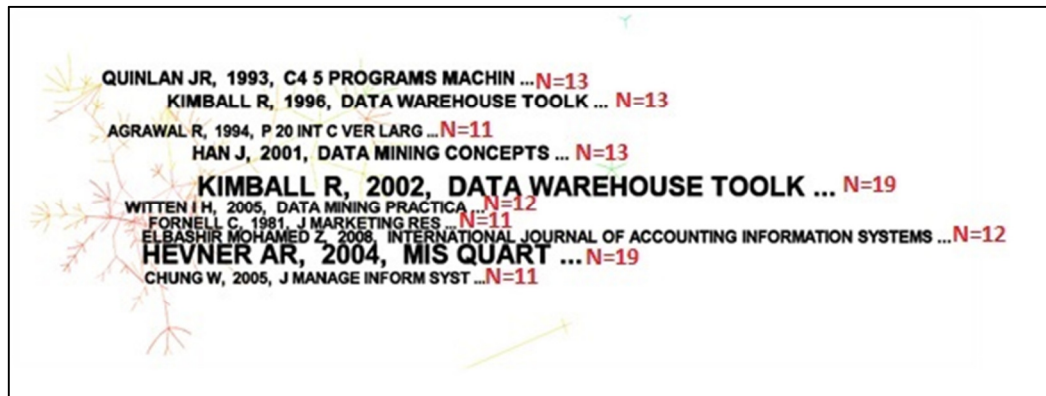
4.1.2 Análise das Citações

A análise de citação pode ser representada por (1) citação documental, que é relacionada à citação como um todo, composta por autor, título, publicação ou imprensa, páginas e série (volume, número e ano); ou (2) por citação autoral, que simplesmente relaciona o nome de referência de autoridade.

Na Figura 2 apresenta-se as principais referências citadas, trata-se do primeiro caso, no qual pode-se observar que Kimball R com suas publicações sobre Data Warehouse de 1996 e 2002 alcança o maior número de citações com 32 ocorrências na somatória geral. Outra referência relevante nesta análise é a obra *Design Science In Information Systems Research* de Alan R. Hevner. As demais referências citadas alcançam um número interessante de citações como o caso de *HAN J, 2001, DATA MINING CONCEPTS, V, P* com 13 citações; o caso de

WITTEN I H, 2005, DATA MINING PRACTICA, V, P e de ELBASHIR MOHAMED Z, 2008, INTERNATIONAL JOURNAL OF ACCOUNTING..., ambos com 12 citações, e; FORNELL C, 1981, J MARKETING RES, V18, P39, CHUNG W, 2005, J MANAGE INFORM SYST, V21, P57 e AGRAWAL R, 1994, P 20 INT C VER LARG, V, P487 todos com 11 citações cada.

Figura 2 – Principais referências citadas - CiteSpace



Fonte: Elaborado pelos Autores a partir dos dados da pesquisa (2015).

A segunda análise trata das citações de autorias. Diferentemente das citações documentais, o foco deste tipo de averiguação parte exclusivamente do número de citações que um determinado autor consegue dentro de um grupo de dados recuperados, no qual os principais neste quesito ficaram por conta de: KIMBALL R, 1996 (n=45), seguido por AGRAWAL R, 1993 (n=34); HAN J, 1998 (n=33); WATSON HJ, 2006 (n=23); SALTON G, 1983 (n= 22); DAVENPORT TH, 2002 – FAYYAD U, 1996 – PORTER ME, 1980 (todos com n=21); GOLFARELLI M, 1998 – TURBAN E, 2008 – CHEN HC, 1996 – HEVNER AR, 2004 (todos com n=20); e INMON WH, 2002 (n=19), conforme representação na próxima figura.

Figura 3 – Autores citados - Citespace



Fonte: Elaborado pelos Autores a partir dos dados da pesquisa (2015).

4.1.3 Análise das Palavras-chave

A Figura 4 demonstra as palavras-chave definidas pelos autores e normalizadas no Excel, apresentadas com um gerador de nuvens em tag (FEINBERG, 2014). Como complemento, a Tabela 2 apresenta a contagem e a classificação segundo a Lei de Zipf das primeiras 20 palavras de um total de 1267 palavras.

Figura 4 – Nuvem de Palavras chaves Autor - Wordle



Fonte: Elaborado pelos Autores a partir dos dados da pesquisa (2015).

Dentro da análise de palavras-chave existem três tipos de classificação que podem ser viáveis, dentro da gestão informacional, para a decisão de quais termos são fundamentais para uma determinada área do conhecimento. O primeiro tipo de palavra-chave é chamado de **estratégica** (termos primários para identificar-se o cerne da árvore terminológica, conforme apresenta a Figura 4). O segundo tipo são as palavras-chave que se considera **interessante** (onde são termos derivativos que em uma primeira análise que não são fundamentais para o entendimento da área de concentração da pesquisa, porém podem ser utilizadas em uma segunda análise, quando se busca correlações de áreas e de terminologias). Este tipo de terminologia pode ser armazenado para uso futuro em outras análises. O terceiro tipo é considerado como ruídos, ou seja, palavras-chave sem importância.

Na Tabela 2 é apresentada a ilustração de como pode-se atribuir poucas palavras-chave para entendimento de como funciona a representação do termo Inteligência de negócios (*Business Intelligence*). O mesmo se relaciona diretamente com mais 12 palavras-chave estratégicas, onde estes termos completam a primeira esfera da teoria de Zipf. Para melhor entendimento, a geração desta esfera se deu pelo valor total de vezes que foi utilizada a palavra *Business Intelligence* pelos autores ($n=179$) elevada a sua raiz quadrada ($n=13$), que determina que as 13 primeiras palavras-chave compõem as informações triviais, denominadas neste estudo como estratégica. As informações denominadas interessantes são fruto da raiz quadrada da segunda palavra mais recorrente ($n=58$), que representa o valor da raiz elevada a $n=7$, porém

como seis palavras continham o mesmo valor (de frequência 6) e uma com valor de cinco, considera-se somente as mais numerosas, conforme identifica-se na Tabela 2.

Tabela 2 – Palavras chaves Autor

Ranking	Palavra-chave do Autor	Fr	Porcentagem	Acumulativo	Zipf
1	Business intelligence	179	9,21%	9,21	Estratégica
2	Data mining	58	2,98%	12,19	Estratégica
3	Data warehouse	46	2,37%	14,56	Estratégica
4	On-Line Analytical Processing (OLAP)	25	1,28%	15,84	Estratégica
5	Knowledge management	24	1,23%	17,07	Estratégica
6	Decision Support System	15	0,77%	17,84	Estratégica
7	E-commerce	12	0,61%	18,45	Estratégica
8	Ontology	9	0,46%	18,91	Estratégica
9	Information management	8	0,41%	19,32	Estratégica
10	Enterprise systems	8	0,41%	19,73	Estratégica
11	Internet	7	0,36%	20,09	Estratégica
12	Intelligence	7	0,36%	20,45	Estratégica
13	Competitive intelligence	7	0,36%	20,81	Estratégica
14	Text mining	6	0,30%	21,11	Interessante
15	Visualization	6	0,30%	21,41	Interessante
16	Information systems	6	0,30%	21,71	Interessante
17	Classification	6	0,30%	22,01	Interessante
18	Data Integration	6	0,30%	22,31	Interessante
19	Customer relationship management	6	0,30%	22,61	Interessante
20 – 25	Palavras com 5 ocorrências	30	1,54%	24,15	Ruído
26 – 45	Palavras com 4 ocorrências	80	4,11%	28,26	Ruído
46 – 78	Palavras com 3 ocorrências	99	5,09%	33,35	Ruído
79 - 182	Palavras com 2 ocorrências	208	10,70%	44,05	Ruído
183 - 1267	Palavras com 1 ocorrência	1085	55,95%	100,00	Ruído

Fonte: Elaborado pelos Autores a partir dos dados da pesquisa (2015).

As palavras-chaves representadas nesta análise são aquelas elaboradas pelos autores e mostram o pensamento destes sobre os assuntos discutidos nos artigos. Os dois primeiros assuntos com maior frequência, depois de BI, são ‘mineração de dados’ e ‘armazém de dados’. Os dois assuntos seguintes, com quase metade da frequência, são ‘OLAP’ (Processamento analítico online) e ‘gestão do conhecimento’. Os outros dois assuntos na sequência, com destaque, são ‘Sistema de Suporte à Decisão’ e ‘E-commerce’.

A mesma toada foi utilizada para a representação das *Keyword Plus* da WoS, que é um complemento de terminologias geradas pelas palavras-chave dos autores, mais a incorporação dos títulos dos artigos, no qual a base de dados gera seus próprios descritores.

Existe a tendência de estes descritores serem representados por termos mais simples como dispostas na Tabela 3.

Tabela 3 – Palavras chaves Autor

Ranking	Palavras-Chaves (Keyword Plus)	WOS	Fr	Porcentagem	Acumulativo	Zipf
1	SYSTEMS		34	2,94%	2,94	Estratégica
2	MANAGEMENT		32	2,77%	5,71	Estratégica
3	MODEL		24	2,07%	7,78	Estratégica
4	FRAMEWORK		23	1,98%	9,76	Estratégica
5	TECHNOLOGY		19	1,64%	11,40	Estratégica
6	PERFORMANCE		18	1,55%	12,95	Estratégica
7	SYSTEM		16	1,38%	14,33	Interessante
8	INFORMATION		16	1,38%	15,71	Interessante
9	MODELS		14	1,21%	16,92	Interessante
10	KNOWLEDGE		14	1,21%	18,13	Interessante
11	WEB		13	1,12%	19,25	Interessante
12	DESIGN		13	1,12%	20,37	Interessante
13 - 574	562 DESCRITORES		920	79,63%	100,00	Ruído

Fonte: Elaborado pelos Autores a partir dos dados da pesquisa (2015).

As palavras-chaves selecionadas pelo sistema já são mais genéricas e mostram 4 assuntos principais: Sistemas, Gestão, Modelo e Estrutura (Framework). O único termo dos 6 destacados anteriormente e de certa forma dissonante dos quatro genéricos escolhido pelo sistema foi ‘E-commerce’.

4.1.4 Análise dos Periódicos

A última análise trata da aplicação da lei de Bradford, que também é conhecida como lei de dispersão e é aplicada a publicações científicas, neste estudo, em periódicos científicos onde os autores publicaram seus estudos sobre *Business Intelligence*.

A Tabela 4 apresenta o *ranking* dos periódicos segundo a Lei de Bradford. São os 17 primeiros de um total de 244, onde este primeiro núcleo representa 32,59%, chamado nesta análise como revistas de concentração.

Tabela 4 – Lei de Bradford aplicado para dispersão de revistas

Ranking	Nome do Periódico	Fr	%	Bradford	
1	EXPERT SYSTEMS WITH APPLICATIONS	28	5,70%	5,70%	Concentração
2	DECISION SUPPORT SYSTEMS	21	4,28%	9,98%	Concentração
3	INFORMATION SYSTEMS MANAGEMENT	14	2,85%	12,83%	Concentração
4	KNOWLEDGE-BASED SYSTEMS	10	2,04%	14,87%	Concentração
5	MIS QUARTERLY	10	2,04%	16,90%	Concentração
6	INFORMATION SYSTEMS FRONTIERS	8	1,63%	18,53%	Concentração
7	INTERNATIONAL JOURNAL OF DATA WAREHOUSING AND MINING	8	1,63%	20,16%	Concentração
8	METALURGIA INTERNATIONAL INTERNATIONAL JOURNAL OF	7	1,43%	21,59%	Concentração
9	INFORMATION MANAGEMENT	7	1,43%	23,01%	Concentração
10	INDUSTRIAL MANAGEMENT & DATA	7	1,43%	24,44%	Concentração

	SYSTEMS				
11	JASIST	6	1,22%	25,66%	Concentração
12	JOURNAL OF COMPUTER INFORMATION SYSTEMS	6	1,22%	26,88%	Concentração
13	DATA & KNOWLEDGE ENGINEERING	6	1,22%	28,11%	Concentração
14	INFORMATION SYSTEMS	6	1,22%	29,33%	Concentração
15	BUSINESS & INFORMATION SYSTEMS ENGINEERING	6	1,22%	30,55%	Concentração
16	JOURNAL OF MANAGEMENT INFORMATION SYSTEMS	5	1,02%	31,57%	Concentração
17	ECONTENT	5	1,02%	32,59%	Concentração
18 – 28	11 Periódicos com 4 ocorrências	44	8,96%	41,55%	Prod. Média
29 – 41	13 Periódicos com 3 ocorrências	39	7,94%	49,49%	Prod. Média
42 – 86	45 Periódicos com 2 ocorrências	90	18,32%	67,82%	Prod. Média
87 – 244	158 Periódicos com 1 ocorrência	158	32,18%	100,00%	Prod. Baixa

Fonte: Elaborado pelos Autores a partir dos dados da pesquisa (2015).

Os demais núcleos foram denominados de produtividade média, com concentração de 69 periódicos científicos; e produtividade baixa, com 158 periódicos sendo representados.

O curioso desta análise é a ampliação do foco das revistas e suas áreas de concentração, onde se pode encontrar *Computer Science, Information Systems; Information Science & Library Science; Management; Computer Science, Artificial Intelligence; Operations Research & Management Science; Computer Science, Theory & Methods; Computer Science, Software Engineering; Metallurgy & Metallurgical Engineering; Computer Science, Interdisciplinary Applications; Engineering, Industrial.*

5 CONSIDERAÇÕES FINAIS

Os resultados do estudo mostraram a relevância do tema Inteligência de Negócios. A produção de artigos é crescente e, além de ter concentração na área da computação, também está presente na Ciência da Informação e na Administração, com enfoque em gestão. Também apresentou os principais autores, principais autores citados, os principais periódicos científicos utilizados para a publicação e as palavras-chave relacionadas.

A *Web of Science* com toda sua abrangência de indexação e sua estrutura e *interface* preparadas para recuperação e estudos bibliométricos e cintométricos, entre outros, não consegue resolver por completo as diferentes grafias de entrada dos dados impactando nos resultados obtidos. No estudo deste artigo a normalização mais intensiva poderia melhorar a exatidão dos dados e informações. Porém julga-se que não houve prejuízo significativo nas análises e considerações.

Como mencionado o termo Inteligência de Negócios sofre com o fenômeno de Polissemia e Sinonímia. Uma pesquisa com o termo *Competitive Intelligence* poderia

complementar este estudo, bem como uma análise conceitual nos artigos de ambos poderia ainda provocar algumas separações sobre os temas e suas produções científicas.

Em outra direção, em toda gestão, incluindo a informacional, necessita-se de indicadores que mostrem quantitativamente os resultados do processo sob a gestão. A bibliometria, no âmbito informacional, é uma das formas de gerar estes indicadores. Além de mostrar aspectos ligados aos resultados, ela também permite, como demonstrado neste estudo, caracterizar uma área ou tema e gerar o elemento fundamental para a tomada de qualquer decisão: conhecimento. No âmbito mais estratégico os indicadores bibliométricos têm permitido a governos, comunidades científicas e programas de P&D, formular políticas e diretrizes incluindo alocação de investimentos e recursos. Na dimensão das organizações, a bibliometria pode ser uma importante ferramenta de aplicação principalmente nas organizações que trabalhem especificamente com Pesquisa, Desenvolvimento e Inovação (PD&I) como os Institutos de Ciência, tecnologia e Inovação (ICTIs). A Bibliometria poderia inclusive ser utilizado pelos próprios sistemas de Inteligência de Negócios como parte integrante dos sistemas nestas organizações.

REFERÊNCIAS

BERGERON, P.; HILLER, C. A. *Competitive intelligence*. **Annual Review Information Science Technology**, v. 36, p. 353-390, 2002. doi: 10.1002/aris.1440360109.

BRADFORD, Samuel Clement. Sources of information on specific subject. **Engineering: an Illustrate Weekly Journal**, Kent, n.137, p.85-86, 1934.

CASTELLS, M. **A sociedade em rede**. 7ed. São Paulo: Paz e Terra, 2003. v.1.

CHEN, C. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. **Journal of the American Society for Information Science and Technology**, v. 57, n. 3, p. 359-377, 2006.

FEINBERG, J. Wordle: *software* de nuvem de palavras. Disponível em: <www.wordle.net>. Acesso em: 10 jan. 2014.

FILIPPO, D. de; FERNÁNDEZ, M. T. Bibliometría: importancia de los indicadores bibliométricos. **El estado de la ciencia: principales indicadores deficiencia y tecnología ibero-americanos / interamericanos**, 2002.

FONSECA, E. N. **Bibliometria**: teoria e prática. São Paulo: Cultrix, 1986.

GUEDES, V. L. 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 PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 6., Salvador, 2005. **Anais...** v. 6, p. 1-18, 2005.

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

MARTÍN SOBRINO, M. I.; PESTANA CALDES, A. I.; PULGARÍN GUERRERO, A. Lei de Lotka aplicada à produção científica da área de ciência da informação. **Brazilian Journal of Information Science**, v. 2, n. 1, p. 16-32, 2008.

MATHEUS, R. F.; PARREIRAS, F. S. Inteligência Empresarial versus *Business Intelligence*: abordagens complementares para o apoio à tomada de decisão no Brasil. In: CONGRESSO ANUAL DA SOCIEDADE BRASILEIRA DE GESTÃO DO CONHECIMENTO, 3., 2004, São Paulo. **Anais...** São Paulo: Kim Brasil, 2004.

MERTON, R. K. The Mathew effect in science. **Science**, [s. l.], v. 159, n. 3810, p. 58, Jan. 1968.

MUGNAINI, R.; STREHL, L. Recuperação e impacto da produção científica na era Google: uma análise comparativa entre o Google acadêmico e a Web of Science. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 13, n. 1, p. 92-105, 2008. doi: 10.5007/1518-2924.

PINTO, A. L.; SOUZA, A. A. de. **Indicadores científicos e tecnológicos de visibilidade nacional e internacional do estado de Mato Grosso**. Cuiabá: EdUFMT, 2013.

PORTAL CAPES. Disponível em:

<http://www.periodicos.capes.gov.br/?option=com_pcollection&mn=70&smn=79&cid=81>. Acesso em: 11 jan. 2014.

TODESCO, J. L.; CARRETERO, L. E.; DURAN, A. **Business Intelligence (Business Intelligence)**. [Slides]. In: CURSO DE BUSINESS INTELLIGENCE. Escuela Complutense Latinoamericana, Florianópolis, 2007.

QUONIAM, L. et al. Inteligência obtida pela aplicação de *data mining* em base de teses francesas sobre o Brasil. **Ciência da Informação**, v. 30, n. 2, p. 20-28, 2001.

URBIZAGASTEGUI, R. A produtividade dos autores sobre a lei de Lotka. **Ciência da Informação**, Brasília, v. 37, n. 2, p. 87-102, 2008.

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**, v. 31, n. 2, p. 152-162, 2002.

ZIPF, George Kingsley. **Human behaviour and the principle of least effort**. Cambridge, Mass.: Addison-Wesley, 1949. 543 p.