



ID. Knowledge representation

MAPEAMENTO DE CONHECIMENTO: IDENTIFICAÇÃO DE PALAVRAS ATRAVÉS DE COCORRÊNCIA

KNOWLEDGE MAPPING:
IDENTIFICATION OF WORDS THROUGH COOCURRENCE

*Danielly Oliveira Inomata*¹
*Mauricio Cordeiro Manhães*²
*Bruna Devens Fraga*³
*Gregorio Jean Varvakis Rados*⁴

RESUMO

Este artigo propõe um procedimento para a pesquisa bibliográfica, tendo como principal resultado o mapeamento da coocorrência de palavras com base em palavras-chave específicas. Este procedimento foi desenvolvido para identificar as principais áreas de pesquisa em campos onde as palavras têm significados complexos e, muitas vezes, estão em fase de interpretação pré-paradigmáticas. Como é o caso da palavra "serviço", que serve como ilustração para este artigo. O procedimento mapeia a coocorrência de palavras, e identifica as presenças fracas ou fortes de palavras específicas em relação à produção de literatura acadêmica. O procedimento proposto, combinado com o apoio de especialistas, identifica palavras solares, planetárias e satélites em relação a um campo de pesquisa acadêmica específico. Como resultado, ele cria um mapa de conhecimento da produção acadêmica nas áreas de pesquisa abordadas pelo procedimento proposto. Desse modo, ele corrobora alguns entendimentos encontrados na literatura no que diz respeito aos termos identificados na coocorrência, como no caso do termo "serviço": ciência, gestão, engenharia e *design*. No entanto, os resultados acima mencionados apontam para a necessidade de mais pesquisas sobre o termo "inovação em serviço", devido ao fato dele ter um índice menor de ocorrência, quando comparado a outros termos. Recomenda-se fazer, ao aplicar este procedimento em pesquisas futuras, a análise com base no tempo da produção acadêmica.

PALAVRAS-CHAVE: Conhecimento. Literatura científica. Serviços. Mapeamento. Coocorrências.

ABSTRACT

This paper proposes a procedure for literature research, having as its main result a mapping of the co-occurrence of words based on specific keywords. This procedure was developed to identify the main areas of research in research fields where the keywords have complex meanings and, often, are on pre-paradigmatic interpretation stages. As is the case of the word 'service' that serves as illustration for this article. The procedure maps the co-occurrence of words, which can identify both weak and strong presence of specific words on academic literature production. The proposed procedure, combined with the support of experts, enables to identify solar, planetary and satellites words in relation to a specific academic research field. As a result, it creates a conceptual map of academic production in the research areas focused by the proposed procedure. Thereby, it corroborated some understandings found in the literature in respect to the terms identified in the co-occurrence, as in the case of service: science, management, engineering and design. However, the aforementioned results pointed to the need of further research about the term 'service innovation', due to the fact that it had a smaller index of occurrence

¹ Graduada em Biblioteconomia, Mestre e doutoranda em Ciência da Informação. E-mail: inomata.danielly@gmail.com

² Mestre e doutorando em Engenharia e Gestão do Conhecimento (UFSC). E-mail: mcmanhaes@gmail.com

³ Mestranda em Engenharia e Gestão do Conhecimento (UFSC). E-mail: brunadefraga@gmail.com

⁴ Professor do Departamento de Ciência da Informação e do Departamento de Engenharia e Gestão do Conhecimento, da UFSC. Doutor em Manufacturing Engineering - Loughborough University of Technology. E-mail: g.varvakis@ufsc.br

Recebido em: 01/12/2014 - **Aceito em:** 28/04/2015

when compared to other terms. It is recommended to do, when applying this procedure in future research, a time-based analysis of the academic.

Keywords: Knowledge. Scientific Literature. Service. Mapping. Cooccurrence.

1. INTRODUÇÃO

O crescimento da produção científica e a ampliação da cobertura de revistas indexadas justifica o aumento de registros de publicações em bases de dados bibliográficos. Isso, segundo Mugnaini *et al.* (2004) é uma tendência inevitável e esperada. E é por meio das publicações – produto das comunicações científicas – que um campo do conhecimento adquire maturidade e se torna visível, tal maturidade pode ser estudada a partir da análise da produção científica, ao mesmo tempo em que possibilita o mapeamento de uma disciplina ao conhecer as temáticas pesquisadas (MACHADO, 2013). Por isso, torna-se relevante o desenvolvimento de métodos ou procedimentos que possibilitem mapear e entender as temáticas relacionadas aos objetos de interesse do pesquisador, bem como acompanhar ou monitorar a evolução científica do mesmo.

A produção de textos científicos é a parte tangível da ciência e se configura como resultado da atividade científica, o qual é divulgado através dos canais de comunicação. A ciência se materializa pela produção de documentos, que se valem dos canais formais e informais para divulgação, de acordo com as escolhas dos autores (BRAMBILLA; STUMPF, 2012).

Desta forma, o presente artigo tem por objetivo apresentar um procedimento para mapear coocorrências de palavras nos resumos de artigos em determinado campo da produção científica. Nesta proposição, considerou-se a análise de conteúdo, baseado em Bardin (2011), quanto a sua função heurística: “a análise de conteúdo enriquece a tentativa exploratória, aumenta a propensão para a descoberta. É a análise de conteúdo ‘para ver o que dá’” (BARDIN, 2011, p.35).

O trabalho está composto por esta introdução, em seguida são detalhados os procedimentos sugeridos, ilustrando a sistemática proposta como forma de mapeamento de conhecimento da produção científica de artigos. Também são apresentados e discutidos os resultados da pesquisa, realizada na base de dados *Scopus*, para cada resultado encontrado são tecidos comentários analíticos. E por fim apresentam-se as considerações, limitações e propostas para futuras pesquisas.

2. PROCEDIMENTOS METODOLÓGICOS

O procedimento proposto tem a intenção de mapear a produção científica sobre uma determinada temática, por meio da análise de conteúdo. Neste sentido, se faz necessária a análise dos “significados” (análise temática) conduzida por especialistas, devido a sua familiaridade com o objeto a ser estudado.

O objeto de estudo eleito para esta pesquisa foi representado pela palavra *service*. Ressalta-se que o procedimento descrito a seguir foi realizado para tratar a necessidade de pesquisadores do Núcleo de Gestão e Sustentabilidade (EGC/UFSC) em obter um horizonte mais definido de conceitos e pesquisas acadêmicas sobre o tema *serviço*, área de especialidade dos referidos pesquisadores. Os pesquisadores em questão são também coautores do presente artigo.

A análise de conteúdo segundo Bardin (2011) serviu de metodologia de apoio ao procedimento aqui proposto. Ilustrativamente, a sistemática para o mapeamento de conhecimento na literatura científica é apresentada na FIGURA 1.

A **Pré-análise** é a macrofase cujas decisões são tomadas pelos especialistas na área, principalmente no que tange a associação das palavras-chave. Com base nos procedimentos iniciais preconizados pelo Método Delphi (DALKEY, N.; HELMER, O, 1962), os pesquisadores devem identificar e formar um painel de especialistas (YOUSUF, 2007), aos quais deveram fazer a seguinte pergunta: “Quais outros termos podem ser combinados à palavra-chave que representa o tema central, afim de representar as principais perspectivas sobre o tema central?” O resultado final desse processo deve permitir representar, através da combinação de termos, as principais perspectivas acadêmicas sobre o tema central.

Na macrofase seguinte, **Exploração do material**, utiliza-se como técnica a Análise de Conteúdo, principalmente quanto à análise categorial, ou seja, análise temática – a contagem de um ou vários temas ou itens de significação, numa unidade de significação previamente determinada – é transversal, pois recorta o conjunto de informações por meio de uma grade de categorias projetada sobre os conteúdos, nesse momento “não se tem em conta a dinâmica e a organização, mas a frequência dos temas extraídos do conjunto dos discursos, considerados dados segmentáveis e comparáveis” (BARDIN, 2011, p.222).

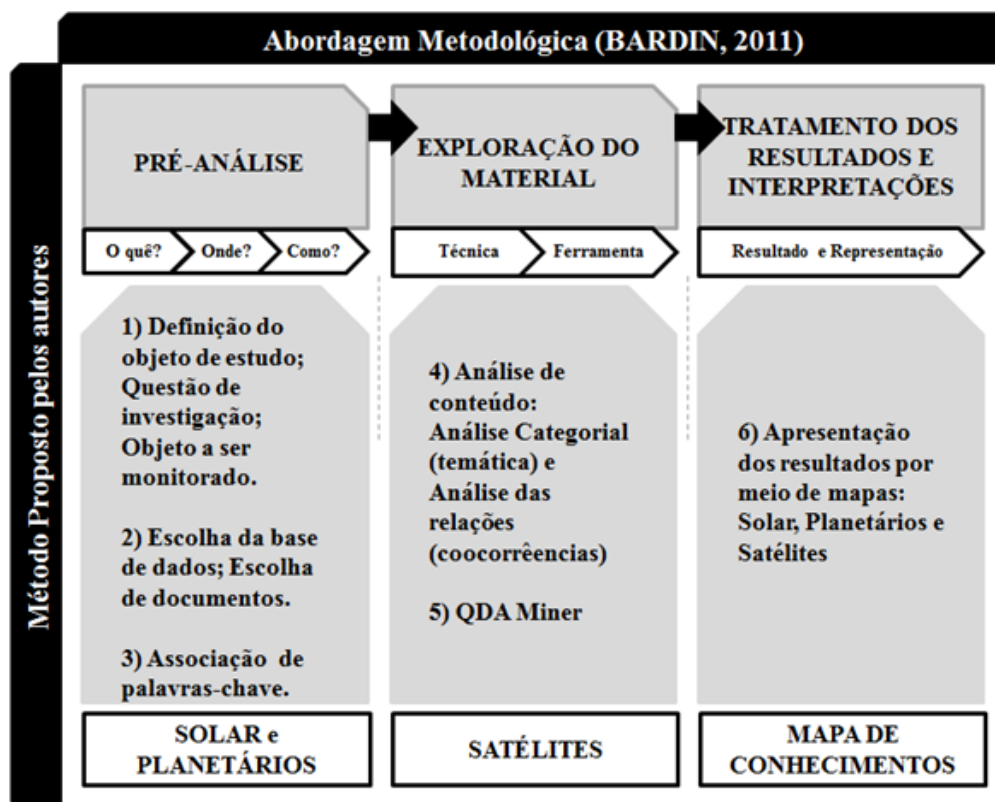


FIGURA 1 – Proposta metodológica de mapeamento de conhecimento na literatura científica

Fonte: Elaborada pelos autores.

E a análise de coocorrências que

procura extrair do texto as relações entre os elementos da mensagem, ou mais exatamente, dedica-se a assinalar as presenças simultâneas de dois ou mais elementos na mesma unidade de contexto, isto é, num fragmento de mensagem previamente definido. (BARDIN, 2011, p.261).

A macrofase **Tratamento dos Resultados e Interpretações** reserva-se para a apresentação dos mapas de conhecimento resultantes da pesquisa. A metáfora adotada para a representação do mapa baseia-se no conceito de *infoticle* ou *infortícula* (MOERE *et al.*, 2004). Trata-se da representação metafórica de informações como se fossem partículas (informação + partícula). Tal forma de representação gráfica e estática facilita a visualização e identificação de diferentes padrões e *clusters* (MOERE *et al.*, 2004, p.42). Neste caso, a representação baseada em infortículas, adota termos metafóricos oriundos da Astronomia. Neste caso, os mapas figuram representações nominadas de: Solar, Planetários e Satélites. O mapeamento do conhecimento serve para representar graficamente os conceitos hierárquicos (palavras-chave) de um determinado domínio do conhecimento, e são desenvolvidos para o domínio de gerenciamento destes ativos (HEGAZY; ALI; ABDEL-MONEM, 2011). A execução desta macrofase torna possível que uma grande quantidade de informação seja representada de forma explícita. A visualização das proximidades espaciais permite de forma mais fácil a compreensão das relações entre os dados e informações (MOERE *et al.*, 2004).

Deste modo, o procedimento desenvolvido pode ser estruturado nas seguintes etapas:

- a) **Definição do termo Solar:** no caso específico, a definição do termo de pesquisa fundacional;
- b) **Definição do conjunto de termos Planetários:** a partir de uma revisão de literatura o termo Solar é relacionado aos seus subconjuntos de pesquisa;
- c) **Levantamento da literatura:** na base *Scopus* (www.scopus.com) foram realizadas pesquisas para a identificação dos artigos contendo cada uma das expressões definidas como identificadoras de áreas de pesquisa;
- d) **Obtenção de dados bibliográficos:** para cada uma das áreas de pesquisa definidas foram obtidos os títulos, autores e resumos de todos os artigos identificados na referida base em todas as áreas e todos os períodos de publicação;
- e) **Análise de conteúdo dos registros:** o conjunto total de títulos e resumos, na forma de arquivos de texto, é submetido à Análise de Conteúdo com a utilização de um programa computacional específico;
- f) **Obtenção dos termos Satélites:** a partir dos índices de coocorrência de termos, é possível estabelecer uma classificação de maior para menor coocorrência tanto dos termos Planetários em relação ao Solar, quanto daquele em relação aos Satélites;
- g) **Elaboração do Mapa de Coocorrências:** com a utilização de um programa de elaboração de mapas de conhecimento, os dados obtidos das coocorrências são representados graficamente, integrando o termo Solar aos termos Planetários e Satélites;
- h) **Entendimento dos Termos:** o mapa de coocorrências facilita a identificação e tratamento dos termos identificados. Isto permite o aprofundamento na investigação de determinados termos. Tal investigação é realizada com a leitura dos termos nos contextos em que eles aparecem. A identificação dos artigos nos quais ocorrem os termos permite classificar esses artigos na base de dados *Scopus* tanto por quantidade de citações, quanto por relevância;
- i) **Descritivo da produção acadêmica:** ao final, com a junção do mapa de coocorrências às definições qualitativas dos termos identificados é possível elaborar um texto descritivo do cenário das pesquisas acadêmicas que gravitam ao redor de determinado termo.

3. APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

A seguir são apresentados os produtos de cada uma das etapas do procedimento proposto.

3.1 Definição do termo Solar

No caso específico desta aplicação ilustrativa do procedimento proposto, a definição do termo de pesquisa foi a palavra inglesa *service*. Com o objetivo de verificar a produção

acadêmica mundial a respeito desse termo, optou-se por utilizar o termo em inglês por ser o idioma com maior produção e abrangência pela comunidade científica ligada ao objeto de estudo: serviço. O termo – *service* – foi definido em razão da necessidade específica da pesquisa de doutoramento que está sendo realizada por um dos autores deste artigo.

3.2 Definição do conjunto inicial de termos Planetários

A partir de uma revisão de literatura sobre *Service*, feita por um dos autores deste artigo, identificaram-se quatro linhas de pesquisa principais: *Service Management*, *Service Engineering*, *Service Design* e *Service Science* (CHESBROUGH; SPOHRER, 2006; LARSON, 2008; MAGLIO *et al.*, 2006, 2009).

Desta forma, o termo Solar foi relacionado aos seguintes subconjuntos de pesquisa: *Science* (representa “*Service Science*”), *Management* (representa “*Service Management*”), *Engineering* (representa “*Service Engineering*”) e *Design* (representa “*Service Design*”).

A FIGURA 2 representa a visualização dos termos Solar e os Planetários principais. Ela apresenta também os Índices Jaccard (IJ) de coocorrência (JACCARD, 1901; SMALL, 1973) que serão tratados no decorrer deste artigo.

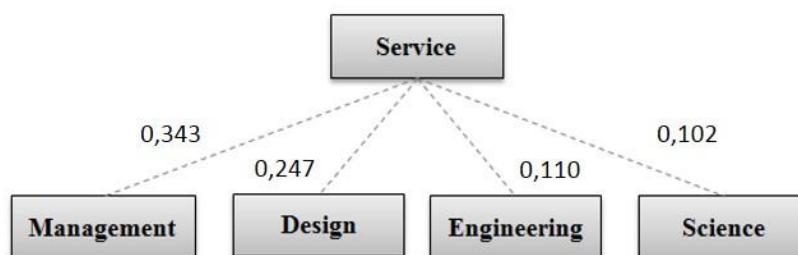


FIGURA 1 – Termo Solar e os Planetários principais

Fonte: Elaborada pelos autores.

Em termos breves, o índice Jaccard de coocorrência indica o percentual de trechos que apresentam os dois termos pesquisados, a partir de um conjunto de textos que possuem um e/ou outro. Ou seja, no caso do par *Service* + *Science*, o índice é de 0,102. Isto quer dizer que, do somatório dos trechos de textos que possuem um e/ou outro, 10,2% apresentam os dois termos em coocorrência.

3.3 Levantamento da literatura

Na base *Scopus*, no dia 14 de Julho de 2014, foram realizadas pesquisas para a identificação dos artigos contendo cada uma das expressões definidas como identificadoras de áreas de pesquisa, resultantes da composição entre o termo Solar e cada um dos termos Planetários.

Na mesma base foram realizadas quatro buscas diferentes com os seguintes termos: *Service Science*, *Service Engineering*, *Service Management* e *Service Design*. Ou seja, na referida base, foram buscados a utilização justaposta dos pares de palavras contendo o termo

Solar e de cada um dos Planetários. Os pares foram informados entre aspas, com requisitos como: *Document Search: Article Title, Abstract, Keywords; Limit to: Date Range (inclusive): All years to Present; Document Type: All; Subject Areas: Life Sciences (> 4,300 titles.), Health Sciences (> 6,800 titles. 100% Medline coverage), Physical Sciences (> 7,200 titles.), Social Sciences & Humanities (> 5,300 titles.)*.

Desta forma, dos grupos Planetários foram obtidos os *abstracts* (resumos) necessários para a pesquisa, ação realizada por meio da exportação dos dados para posterior tratamento. Como a base de dados *Scopus* limita a 2000 a quantidade de resumos que podem ser exportados para cada busca, foi necessário dividir os grupos de Busca em pacotes de 2000 resumos. Tal divisão tomou por base a distribuição dos artigos pelos anos de publicação. Ao final, foram gerados 8 pacotes de resumos, que ficaram assim distribuídos:

- “Service Design” de: *All years* a 2009: 2000 resumos; 2010 a *Present*: 1.785 resumos;
- “Service Engineering” de: *All years* a *Present*: 1.321 resumos;
- “Service Management” de: *All years* a 2000: 2000 resumos; 2001 a 2007: 2000 resumos; 2008 a 2010: 2000 resumos; 2011 a *Present*: 1.556 resumos;
- “Service Science” de *All years* a *Present*: 1.197 resumos.

Esta busca resultou num total de 13.859 resumos, conjunto este que é detalhado a seguir.

3.4 Obtenção de dados bibliográficos

Para cada uma das áreas de pesquisa definidas foram obtidos os títulos, autores e resumos de todos os artigos identificados na referida base, em todas as áreas e todos os períodos de publicação. Ao final, foram identificados 13.859 resumos de artigos, distribuídos da seguinte forma:

TABELA 1 – Grupos de Busca e Resumos obtidos

Grupos de Busca	Resumos	% do Total
<i>Service Science</i>	1.197	8,64%
<i>Service Engineering</i>	1.321	9,53%
<i>Service Management</i>	7.556	54,52%
<i>Service Design</i>	3.785	27,31%
Total	13.859	100%

Fonte: Manhães, Inomata, Fraga e Varvakis (2014).

Ao analisar o conjunto inicial de resumos, fica evidente uma maior concentração deles relacionada ao grupo de busca “Service Management.” A Análise de Conteúdo dos registros também leva em consideração que o conjunto total apresenta a possibilidade de que uma quantidade indefinida de resumos possa estar presente em dois ou mais grupos de busca. A utilização do índice de Jaccard, de certa forma, permite equilibrar a análise visto que ele

leva em consideração o percentual de trechos (no caso, parágrafos) no total dos 13.859 resumos que apresentam o termo Solar em coocorrência com cada um dos termos Planetários.

Ressalta-se que só foi possível estabelecer uma comparação indireta da quantidade de resumos, embora seja interessante estabelecer exatamente a quantidade de resumos que aparecem em mais de um dos grupos de busca.

No que tange à verificação de duplicidade de resumos, ela ocorreu da seguinte forma: todos os títulos dos 13.859 artigos foram inseridos em uma coluna do *MS Excel* e foi executada a função de remoção de duplicadas. Desta forma, identificou-se a existência de 519 artigos com mais de uma presença na lista total. Ou seja, na lista obtida constam 13.340 resumos com título único.

TABELA 2 – Quantidades de Resumos duplicados obtidos

Comparativo	Resumos Duplicados	Resumos Obtidos	Outros Resumos	% Duplicados nos OUTROS
“ <i>Service Design</i> ” X OUTROS	233	3.785	10.074	2.31%
“ <i>Service Management</i> ” X	197	7.556	6.303	3.13%
“ <i>Service Engineering</i> ” X	167	1.321	12.538	1.33%
“ <i>Service Science</i> ” X OUTROS	154	1.197	12.662	1.22%

Fonte: Manhães, Inomata, Fraga e Varvakis (2014).

Com esta verificação, também foi possível estabelecer em quais grupos de busca mais ocorreu a duplicidade de artigos. Na tabela anterior é possível ver que, ao se comparar a lista de títulos obtida pela busca *Service Design* (3.785 resumos) com o total de OUTROS grupos somados (no caso específico, este corresponde a 10.074 resumos de um total de 13.859), 233 artigos que aparecem na primeira lista, também estão presentes no combinado das demais. Em outras palavras, 2,31% dos resumos presentes em OUTROS, aparecem duplicados na busca do grupo “*Service Design*.” Outro exemplo, em termos percentuais é possível constatar que 1,22% dos títulos que aparecem no combinado dos demais grupos de busca (154 dos 12.662 resumos) estão presentes também no resultado da busca por *Service Science* (1.197 resumos). A mesma lógica se aplica aos demais grupos, de acordo com a tabela 2.

3.5. Análise de conteúdo dos registros

O conjunto total de títulos e resumos, na forma de arquivos de texto, foi submetido à Análise de Conteúdo com a utilização de um programa computacional especializado na análise qualitativa de dados. O programa utilizado foi o *QDA Miner* e, em específico, uma das ferramentas oferecidas pelo sistema para a análise de conteúdo denominada *Proximity Plot Cluster*. Esta ferramenta verifica a coocorrência de palavras em determinados trechos de texto, através de um algoritmo de análise.

Os oito casos foram importados pelo *software QDA Miner*, configurando-se as seguintes estatísticas (COLLECTION STATISTICS): Número total de casos: 8; Número total de casos não-vazios: 8; Número total de parágrafos: 123.426; Número total de frases:

137.707; Número total de palavras (token): 3.212.147; Número total de formatos de palavras (type): 65.331; Razão Type/Token: 0,020; Total de palavras excluídas: 1.429.192; Percentual de palavras excluídas: 44,5%; Palavras por frases: 23,3; Palavras por parágrafo: 26,0; Palavras por casos não-vazios: 401.518.

Na tabela a seguir é possível perceber que a quantidade de parágrafos, onde se verificam as coocorrências, é maior do que a quantidade de resumos obtidos para os grupos de busca. Isto indica que, embora a pesquisa na base de dados *Scopus* tenha retornado os resumos em que os termos aparecem imediatamente um ao lado do outro, as coocorrências indicam a proximidade dos termos de forma a representar a presença dos mesmos em um mesmo parágrafo.

Sendo assim, o número maior de coocorrências indica que (i) os termos coocorrem em mais de um parágrafo por resumo e (ii) os termos coocorrem em resumos de outros grupos de pesquisa.

TABELA 3 – Quantidade de Coocorrências nos Resumos obtidos

Descritores de Casos (QDA Miner)	Resumos (SCOPUS)	COOCORRÊNCIAS (Parágrafos - QDA Miner)	% de COOCORRÊNCIAS (Parágrafos - QDA Miner)
<i>Service Science</i>	1.197	3.754	3,04%
<i>Service Engineering</i>	1.321	3.943	3,19%
<i>Service Management</i>	7.556	12.974	10,51%
<i>Service Design</i>	3.785	8.498	6,89%
Total	13.859		

Fonte: Manhães, Inomata, Fraga e Varvakis (2014).

3.6 Obtenção dos termos Satélites

A partir dos índices de coocorrência de termos, é possível estabelecer uma classificação de maior para menor coocorrência dos termos Planetários em relação ao Solar.

A tabela seguinte permite examinar mais detalhadamente os valores numéricos referentes ao cálculo dessas coocorrências. A fim de compreender o distanciamento que ocorre entre os termos com base nas coocorrências, a tabela fornece informações detalhadas tais como o número de vezes que uma dada palavra-chave coocorre com outra (COOCORRUS) e o número de vezes que aparece na ausência da palavra-chave (DO NOT). Também é possível visualizar na tabela o número de vezes que a palavra-chave selecionada aparece na ausência de outra palavra-chave determinada (IS ABSENT).

Em todas as análises a seguir, foi desconsiderada a palavra SERVICE, visto que esta foi o ponto de partida da busca, e desta forma ampliar e aprofundar o grupo de termos satélites encontrados relacionados aos termos planetários.

TABELA 4 – Classificação das Coocorrências para o termo SERVICE

TARGET	KEYWORD	COOCCURS	DO NOT	IS ABSENT	Jaccard
SERVICE	MANAGEMENT	12974	5214	19680	0,343
SERVICE	DESIGN	8498	1739	24156	0,247
SERVICE	SYSTEM	7858	4011	24796	0,214
SERVICE	MODEL	5938	1021	26716	0,176
SERVICE	PROCESS	4796	1013	27858	0,142
SERVICE	INFORMATION	5115	3859	27539	0,140
SERVICE	MARKET	4742	1504	27912	0,139
SERVICE	QUALITY	4035	710	28619	0,121
SERVICE	CUSTOMER	3977	550	28677	0,120
SERVICE	NETWORK	4016	1970	28638	0,116
SERVICE	PROVIDE	3689	328	28965	0,112
SERVICE	ENGINEERING	3943	3186	28711	0,110
SERVICE	DEVELOPMENT	3599	943	29055	0,107
SERVICE	HEALTH	3756	2908	28898	0,106
SERVICE	TECHNOLOGY	3813	3667	28841	0,105
SERVICE	DATA	3488	745	29166	0,104
SERVICE	BUSINESS	3562	2028	29092	0,103
SERVICE	SCIENCE	3754	4206	28900	0,102
SERVICE	SUPPORT	3376	524	29278	0,102

Fonte: Manhães, Inomata, Fraga e Varvakis (2014).

Como exemplo (calculados com base nos critérios de frequência em relação a parágrafos), pode-se ver que a palavra MANAGEMENT coocorre em 12.974 parágrafos com SERVICE, mas MANAGEMENT é encontrado em 5.214 parágrafos sem a palavra SERVICE, enquanto SERVICE é encontrado em 19.680 parágrafos, na ausência de MANAGEMENT.

O coeficiente de Jaccard de 0,343 indica que, de todos os parágrafos que contenham qualquer uma das duas palavras MANAGEMENT ou SERVICE, 34,2% contém as duas palavras em coocorrência. Cabe ressaltar, entretanto, que nem todas as medidas de proximidade podem ser interpretadas de forma simplista ou muito facilmente.

3.7 Elaboração do Mapa de Coocorrências

Com a utilização de um programa de elaboração de mapas de conhecimento, os dados obtidos das coocorrências são representados graficamente, integrando o termo Solar aos termos Planetários e Satélites, conforme pode ser visto na FIGURA 3.

Em função da experiência dos autores em relação à pesquisa sobre serviço, foi sugerida a adição do termo INNOVATION ao mapa, por se tratar de um termo-chave relevante para o campo de pesquisa atual (MAGLIO; SPOHRER, 2007; OSTROM *et al.*, 2010). Na análise de coocorrência, é possível verificar que os referidos termos SERVICE e INNOVATION, coocorrem em 1.363 parágrafos; com INNOVATION aparecendo sozinho em 361 e SERVICE em 31.291 parágrafos. O índice de Jaccard entre os dois termos é de 0,041, isso significa que de todos os parágrafos onde ocorreram as palavras SERVICE e/ou INNOVATION, apenas 4,1% deles continham ambos os termos concomitantemente. Na classificação geral de coocorrência de palavras com o termo SERVICE, o termo INNOVATION aparece como 109º colocado.

3.8 Entendimento dos Termos

O mapa de coocorrências facilita a identificação e tratamento dos termos identificados, isto permite o aprofundamento na investigação de determinados termos. Tal investigação é realizada com a leitura dos termos nos contextos nos quais eles aparecem. Embora não faça parte dos objetivos do presente artigo, a identificação dos artigos nos quais ocorrem determinados termos permite classificar esses artigos na base de dados *Scopus* tanto por quantidade de citações, quanto por relevância.

Além de estabelecer os índices de coocorrência dos termos Planetários em relação ao Solar, é possível fazer o mesmo para a relação deles com os Satélites. Os tabelas seguintes apresentam as referidas coocorrências para os termos Planetários: SCIENCE, ENGINEERING, MANAGEMENT e DESIGN. E, também, são apresentadas análises superficiais das leituras possíveis em relação à classificação das coocorrências dos termos Planetários em relação aos Satélites.

SCIENCE

Na tabela seguinte é possível perceber que as maiores coocorrências do termo SCIENCE, dentro do conjunto de resumos em análise, se dá com as palavras COMPUTER, UNIVERSITY e TECHNOLOGY. Segundo os especialistas, é possível considerar as UNIVERSITY, DEPARTMENT e AFFILIATION como representações da estreita relação que a coocorrência de SERVICE e SCIENCE possui com o meio acadêmico. É possível desconsiderar o termo SERVICE em função das características da própria pesquisa.

Feitas essas considerações, identifica-se que o termo SCIENCE coocorre com os termos COMPUTER, TECHNOLOGY, ENGINEERING, INFORMATION, SYSTEM e INTELLIGENCE.

Para os especialistas, existe uma tendência de que o termo SERVICE SCIENCE ocorra em áreas de pesquisa mais ligadas às tecnologias de informação e comunicação. Isto é explicado pelo fato de que a principal entidade promotora dessa “disciplina” é a *International Business Machines Corporation* (IBM), conforme atesta a literatura (LARSON, 2008).

TABELA 5 – Classificação das Coocorrências para o termo SCIENCE

TARGET	KEYWORD	COOCCURS	DO NOT	IS ABSENT	Jaccard
SCIENCE	COMPUTER	2156	1834	5804	0,220
SCIENCE	UNIVERSITY	2077	8523	5883	0,126
SCIENCE	TECHNOLOGY	1619	5861	6341	0,117
SCIENCE	SERVICE	3754	28900	4206	0,102
SCIENCE	ENGINEERING	1316	5813	6644	0,096
SCIENCE	DEPARTMENT	1041	3471	6919	0,091
SCIENCE	INFORMATION	1414	7560	6546	0,091
SCIENCE	SYSTEM	1553	10316	6407	0,085
SCIENCE	AFFILIATION	1357	8735	6603	0,081
SCIENCE	INTELLIGENCE	601	400	7359	0,072

Fonte: Manhães, Inomata, Fraga e Varvakis (2014).

ENGINEERING

Na tabela seguinte é possível perceber que as maiores coocorrências do termo ENGINEERING, dentro do conjunto de resumos em análise, se dá com as palavras SERVICE, DESIGN e SYSTEM. Segundo os especialistas, é possível considerar as ocorrências de UNIVERSITY e DEPARTMENT como representação de um viés acadêmico em relação à palavra ENGINEERING.

Feitas essas considerações, identifica-se que o termo ENGINEERING coocorre com os termos DESIGN, SYSTEM, SOFTWARE, COMPUTER, SCIENCE, MODEL e TECHNOLOGY.

Ao analisar apenas os resumos referentes ao termo SERVICE ENGINEERING, é possível perceber que a utilização das palavras DESIGN e MODEL enquadra-se, por exemplo, na seguinte definição da referida disciplina (SPATH; FÄHNRIK, 2007, p. 05, tradução nossa):

Engenharia de Serviço pode então ser definida como a sistematização do design e desenvolvimento de serviços utilizando modelos [**models**], métodos e ferramentas apropriadas.

Para os especialistas, existe uma tendência de que o termo SERVICE ENGINEERING também ocorra em áreas de pesquisa mais ligadas às tecnologias de informação e comunicação. Os termos SYSTEM, SOFTWARE, COMPUTER, SCIENCE e TECHNOLOGY permitem suportar essa tendência.

TABELA 6 – Classificação das Coocorrências para o termo ENGINEERING

TARGET	KEYWORD	COOCCURS	DO NOT	IS ABSENT	Jaccard
ENGINEERING	SERVICE	3943	28711	3186	0,110
ENGINEERING	DESIGN	1643	8594	5486	0,104
ENGINEERING	SYSTEM	1789	10080	5340	0,104
ENGINEERING	SOFTWARE	874	1644	6255	0,100
ENGINEERING	COMPUTER	980	3010	6149	0,097
ENGINEERING	UNIVERSITY	1563	9037	5566	0,097
ENGINEERING	SCIENCE	1316	6644	5813	0,096
ENGINEERING	MODEL	1147	5812	5982	0,089
ENGINEERING	TECHNOLOGY	1158	6322	5971	0,086
ENGINEERING	DEPARTMENT	874	3638	6255	0,081

Fonte: Manhães, Inomata, Fraga e Varvakis (2014).

MANAGEMENT

Na tabela seguinte é possível perceber que os maiores índices de coocorrência do termo MANAGEMENT, dentro do conjunto de resumos em análise, são apresentados pelas palavras SERVICE, SYSTEM e NETWORK. Segundo os especialistas, é possível desconsiderar o termo SERVICE em função das características da própria pesquisa.

Feitas essas considerações, é possível inferir que o termo MANAGEMENT coocorre com os termos SYSTEM, NETWORK, INFORMATION, MODEL, PROCESS,

TECHNOLOGY, DESIGN, QUALITY e PROVIDE. Estes termos podem ser enquadrados no entendimento dos especialistas desta pesquisa a respeito do termo SERVICE MANAGEMENT. O termo PROVIDE merece, nesta breve análise, uma atenção especial, visto que ele representa um dos elementos de gestão de serviços. Um exemplo do contexto no qual é utilizado o termo PROVIDE está no seguinte trecho de um resumo (WADE; RICHARDSON, 2000, p.231, tradução nossa):

Um aspecto crucial que enfrentam os Provedores [**Providers**] de Serviços de Telecomunicação é a necessidade de prover [**provide**] sistemas de gestão flexíveis e eficientes do ponto de vista de custos e que possam ser adaptados rapidamente à mudança das demandas dos clientes e das oportunidades de mercado. Uma tendência recente, que tenta tratar dessa necessidade, tem sido o desenvolvimento de soluções de gestão na forma de (*building blocks*) componentes de gestão reutilizáveis. Muita das vezes a integração desses componentes requer a programação de códigos e roteiros para prover [**provide**] uma solução integrada.

Para os especialistas é interessante o destaque recebido pelo termo PROVIDE (com 10% de ocorrências nos respectivos parágrafos). O que poderia configurar que um dos focos do SERVICE MANAGEMENT está em prover sistemas, informações, modelos e processos na forma de soluções integradas de gestão.

TABELA 7 – Classificação das Coocorrências para o termo MANAGEMENT

TARGET	KEYWORD	COOCCURS	DO NOT	IS ABSENT	Jaccard
MANAGEMENT	SERVICE	12974	19680	5214	0,343
MANAGEMENT	SYSTEM	4985	6884	13203	0,199
MANAGEMENT	NETWORK	3033	2953	15155	0,143
MANAGEMENT	INFORMATION	3365	5609	14823	0,141
MANAGEMENT	MODEL	2879	4080	15309	0,129
MANAGEMENT	PROCESS	2671	3138	15517	0,125
MANAGEMENT	TECHNOLOGY	2537	4943	15651	0,110
MANAGEMENT	DESIGN	2769	7468	15419	0,108
MANAGEMENT	QUALITY	2216	2529	15972	0,107
MANAGEMENT	PROVIDE	2024	1993	16164	0,100

Fonte: Manhães, Inomata, Fraga e Varvakis (2014).

DESIGN

As maiores coocorrências do termo DESIGN, dentro do conjunto de resumos em análise, acontecem com as palavras SERVICE, MODEL e SYSTEM.

Diante disto, é possível inferir que o termo DESIGN, coocorre com os termos MODEL, SYSTEM, PROCESS, METHODOLOGY, PROVIDE, GROUP, DEVELOP, PRESENT e SUPPORT. Todos esses termos podem ser considerados como que enquadrados no entendimento que os especialistas possuem a respeito do SERVICE DESIGN.

Um dos termos que os especialistas julgaram merecer destaque é o GROUP. No seguinte texto é possível ver uma ilustração de como ele é usado para se referir a interação com o conceito de grupos tanto para a pesquisa científica quanto para o desenvolvimento de novos produtos (GREENWOOD *et al.*, 1994, p.1999, tradução nossa):

No entanto, a duração do contato não foi aumentada pela gestão de casos, e não houve nenhuma melhora perceptível nos resultados dos grupos [groups] com gestão de casos. As tendências foram a favor do grupo [group] de controle e elas poderiam ser atribuídas a grupos [groups] de diferentes níveis iniciais de severidade dos traumas.

Segundo os especialistas é possível definir a atuação da produção acadêmica do SERVICE DESIGN, a partir da coocorrência de termos dentro do conjunto definido de resumos, como abrangendo o desenvolvimento de modelos, sistemas e processos *por* e com o objetivo de prover suporte *para* grupos sociais.

TABELA 8 – Classificação das Coocorrências para o termo DESIGN

TARGET	KEYWORD	COOCCURS	DO NOT	IS ABSENT	Jaccard
DESIGN	SERVICE	8498	24156	1739	0,247
DESIGN	MODEL	2533	4426	7704	0,173
DESIGN	SYSTEM	3177	8692	7060	0,168
DESIGN	PROCESS	2259	3550	7978	0,164
DESIGN	METHODOLOGY	1740	658	8497	0,160
DESIGN	PROVIDE	1866	2151	8371	0,151
DESIGN	GROUP	1676	1249	8561	0,146
DESIGN	DEVELOP	1693	1989	8544	0,138
DESIGN	PRESENT	1676	2037	8561	0,137
DESIGN	SUPPORT	1688	2212	8549	0,136

Fonte: Manhães, Inomata, Fraga e Varvakis (2014).

3.9 Descritivo da produção acadêmica

Ao final, com a junção do mapa de coocorrências às definições qualitativas dos termos identificados é possível elaborar um texto descritivo do cenário das pesquisas acadêmicas que gravitam ao redor de determinado termo, neste caso, SERVICE.

A tabela seguinte apresenta as classificações dos termos satélites em relação ao solar de acordo com os índices de coocorrência. É possível verificar que o maior índice de coocorrência entre o termo solar e os planetários está na relação entre SERVICE e MANAGEMENT com 34,3%. No outro extremo está a relação entre SERVICE e INNOVATION com apenas 4,1% de coocorrência.

A baixa coocorrência do termo INNOVATION merece destaque pois ela é menos da metade do índice registrado pelo par SERVICE e SCIENCE que é de 10,2%. Em ambos os casos, são duas abordagens relativamente novas na academia.

TABELA 9 – Destaque das Classificações para o termo SERVICE

TARGET	KEYWORD	COOCCURS	DO NOT	IS ABSENT	Jaccard
SERVICE	MANAGEMENT	12974	5214	19680	0,343
SERVICE	DESIGN	8498	1739	24156	0,247
SERVICE	ENGINEERING	3943	3186	28711	0,110
SERVICE	SCIENCE	3754	4206	28900	0,102
SERVICE	INNOVATION	1363	361	31291	0,041

Fonte: Manhães, Inomata, Fraga e Varvakis (2014).

Dentro desse conjunto total dos resumos coletados, também é possível fazer a análise das coocorrências dos termos Planetários com os termos INNOVATION. Com a utilização desse termo como palavra-alvo é possível levantar que SCIENCE é o 9º classificado em termos de coocorrência com 5,1%. DESIGN é o 12º, com 4,9%; ENGINEERING, o 43º, com 3,7% e MANAGEMENT é o 84º, possuindo um índice de coocorrência de 3%.

Segundo os especialistas, o fato de que os termos DESIGN e SCIENCE coocorram em maior grau em relação ao termo INNOVATION é consistente com o entendimento deles sobre a produção de pesquisas acadêmicas na área quando considerado o requisito da busca inicial como abrangendo todos os anos disponíveis na base de dados *Scopus* (Limit to: Date Range (inclusive): All years to Present).

TABELA 9 – Destaque das Classificações para o termo INNOVATION

TARGET	KEYWORD	COOCCURS	DO NOT	IS ABSENT	Jaccard
INNOVATION	SCIENCE	471	7489	1253	0,051
INNOVATION	DESIGN	558	9679	1166	0,049
INNOVATION	ENGINEERING	317	6812	1407	0,037
INNOVATION	MANAGEMENT	572	17616	1152	0,030

Fonte: Manhães, Inomata, Fraga e Varvakis (2014).

Com base nas coocorrências dos termos Planetários em relação aos satélites, é possível estabelecer algumas perspectivas de diferenciação entre as disciplinas representadas pelos grupos de pesquisa. Nos tabelas apresentados é possível notar que, enquanto as coocorrências na coluna de *Keywords* de SCIENCE (TABELA 5) aponta para uma perspectiva de pesquisas universitárias em relação a sistemas de computação e tecnologia da informação, na mesma coluna de DESIGN (TABELA 8) há referências a modelos e sistemas de serviço envolvendo suporte a grupos sociais.

Enquanto que na coluna ENGINEERING (TABELA 6), as ocorrências são de *Keywords* ligados a sistemas e programas de computador, com uma alta coocorrência do DESIGN; na de MANAGEMENT (TABELA 7) aparece a coocorrência de termos ligados a modelos, sistemas, redes e informação.

A FIGURA 3 permite ter uma visão completa das relações entre os termos e seus índices de coocorrências.

4. CONCLUSÕES

A proposição detalhada nesse trabalho é direcionada para mapear coocorrências de termos, assim como se configura uma forma de visualizar a produção científica de artigos, elegendo-se o resumo como unidade de análise. Ressalta-se que embora a proposta tenha focado mais em dados quantitativos, a análise de conteúdo – especificamente a análise categorial e de coocorrências – foi essencial para dar solidez a esta proposição.

Acredita-se que é possível fazer um monitoramento do conhecimento produzido e veiculado em periódicos científicos indexados em bases de dados bibliográficas. No entanto, tal monitoria deve ser baseada em conhecimento especializado. O procedimento proposto exige a interpretação competente dos resultados obtidos, para a qual ele fornece uma perspectiva de abordagem a um estoque elevado de dados e informações.

Da mesma forma que o procedimento confirmou as percepções dos especialistas sobre as diferentes abordagens acadêmicas em relação ao tema *service*, ele também gerou novas e interessantes perspectivas. Como ilustração de confirmações das perspectivas dos especialistas, pode-se ressaltar por um lado: a confirmação de que a abordagem identificada como SERVICE SCIENCE está relacionada ao campo acadêmico, com maiores ocorrências de termos como UNIVERSITY, DEPARTMENT e AFFILIATION. De outro, como ilustração de novas perspectivas, pode-se destacar a baixa coocorrência dos termos SERVICE e INNOVATION. O que, de certa forma, representa um resultado surpreendente para os especialistas envolvidos. Segundo os mesmos, seria esperado que os termos SERVICE e INNOVATION apresentassem uma alta coocorrência devido às diversas iniciativas acadêmicas que tem sido realizadas desde o início do século XXI (MAGLIO; SPOHRER, 2007; OSTROM *et al.*, 2010).

Os resultados obtidos por meio do procedimento adotado nesta pesquisa foram considerados de alta relevância tanto por confirmar muitos dos entendimentos dos especialistas a respeito das referidas disciplinas (*Service Management, Service Engineering, Service Design e Service Science*), quanto por questionar outro, como inovação. A análise mais detalhada dos resultados apresentados aponta para a oportunidade de um tratamento aprofundado em um futuro artigo.

Recomenda-se para pesquisas futuras, a aplicação dessa proposição somada à análise qualitativa dos artigos, bem como um recorte temporal por grupos de busca, a fim de se obter mais profundidade a respeito do universo de estudo. Ou estender o campo de busca para títulos e palavras-chave.

REFERÊNCIAS

BARDIN, L. **A análise de conteúdo**. São Paulo: Edições 70, 2011.

BRAMBILLA, S. D. S.; STUMPF, I. R. C. Produção científica da UFRGS representada na WoS (2000-2009). **Perspectivas em Ciência da Informação**, v. 17, n. 3, p. 34-50, jul./set. 2012.

CHESBROUGH, H.; SPOHRER, J. A research manifesto for services science. **Communications of the ACM**, v. 49, n. 7, 2006.

DALKEY, N.; HELMER, O. **An experimental application of the Delphi Method to the use of experts**. Santa Monica - California: 1962. Disponível em:
<http://www.rand.org/content/dam/rand/pubs/research_memoranda/2009/RM727.1.pdf>.

FERREIRA, A. G. C. Bibliometria na avaliação de periódicos científicos. **Data Gramma Zero: Revista de Ciência da Informação**. Rio de Janeiro, v.11, n.3, p. 1-7, jun. 2010. Disponível em: < http://www.dgz.org.br/jun10/Art_05.htm >. Acesso em: 14 jul. 2014.

GREENWOOD, R. J. *et al.* Effects of case management after severe head injury. **British Medical Journal**, v. 308, n. 6938, p. 1199–1205, 1994.

HEGAZY, T.; ALI, A.; ABDEL-MONEM, M. Prospects of mind maps for better visualization of infrastructure literature. **Journal of Professional Issues in Engineering Education & Practice**, v. 137, n. 4, p. 239-247, 2011.

JACCARD, P. Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines. **Bulletin de la Société Vaudoise des Sciences Naturelles**, v. 37, pp. 241-272, 1901.

LARSON, R. Service science: At the intersection of management, social, and engineering sciences. **IBM Systems Journal**, pp. 1–17, 2008.

MACHADO, R. das N. Produtividade dos autores em ceratocone: estudo longitudinal em artigos indexados no SciELO Brasil (2001-2010). **RDBCI**, v. 11, n. 1, p. 1-20, 2013. Disponível em: <<http://www.sbu.unicamp.br/seer/ojs/index.php/rbci/article/view/558/pdf>>. Acesso em: 14 jul. 2014

MAGLIO, P. P. *et al.* Service systems, service scientists, SSME, and innovation. **Communications of the ACM**, v. 49, n. 7, p. 81–85, 2006.

MAGLIO, P. P. *et al.* The service system is the basic abstraction of service science. **Information Systems and e-Business Management**, v. 7, n. 4, p. 395–406, doi:10.1007/s10257-008-0105-1, 2009.

MAGLIO, P. P.; SPOHRER, J. Fundamentals of service science. *Journal of the Academy of Marketing Science*, 36(1), 18–20. doi:10.1007/s11747-007-0058-9, 2007.

MOERE, A. V.; MIEUSSET, K. H., ; GROSS, M. Visualizing abstract information using motion properties of data-driven infoticles. In: ERBACHER, R. F., P. C. (Ed.) et al. **Proc. SPIE 5295**: visualization and data analysis. San Jose, CA, 2004. doi:10.1117/12.539238, 2004.

MUGNAINI, R. *et al.* Indicadores bibliométricos da produção científica brasileira: uma análise a partir da base Pascal. **Ciência da Informação**, v. 33, n. 2, p. 123-131, 2004.

OSTROM, A. L. *et al.*. Moving Forward and Making a Difference: Research Priorities for the Science of Service. **Journal of Service Research**, v. 13, n. 1, 4–36. doi: 10.1177/1094670509357611, 2010.

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*, 24(4), 265–269. doi:10.1002/asi.4630240406, 1973.

SPATH, D.; FÄHNRIK, K.-P. Service engineering: State of the art and future trends. In: SPATH, D.; FÄHNRIK, K.-P. (Ed.). **Advances in services innovations**. [S.l.]: Springer Berlin Heidelberg, 2007. p. 3–16.

VANTI, N. Indicadores web e sua aplicação à produção científica disponibilizada em revistas eletrônicas. In: FERREIRA, S. M.; TARGINO, M. das G. (Org.). **Acessibilidade e visibilidade de revistas científicas eletrônicas**. São Paulo: Editora São Paulo, 2010.

WADE, V. P.; RICHARDSON, T. Workflow - a unifying technology for operational support systems. In: IEEE SYMPOSIUM RECORD ON NETWORK OPERATIONS AND MANAGEMENT SYMPOSIUM. **Anais...** [S.l.: s.n.], 2000. Disponível em: <<http://www.scopus.com/inward/record.url?eid=2-s2.0-0033715088&partnerID=40&md5=5ab4de4aa651890789e0064779ad7f3c>>. Acesso em: 14 jul. 2014.

YOUSUF, M. I. Using experts' opinions through Delphi technique. **Practical Assessment, Research & Evaluation**, v. 12, n. 4, p. Available online: <http://pareonline.net/getvn.asp?>, 2007.

Como citar este documento:

INOMATA, Danielly Oliveira et al. Mapeamento de conhecimento: identificação de palavras através de coocorrência. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, SP, v. 13, n. 2, p. 279-297, maio/ago. 2015. ISSN 1678-765X. Disponível em: <<http://periodicos.bc.unicamp.br/ojs/index.php/rdbci/article/view/2121>>. Acesso em: 31 maio 2015.