

Expert profiling: um modelo para extração automática de perfil de especialista para apoio à *Expertise Location* e à Gestão do Conhecimento

Rudger N. N. Taxweiler, M.Eng. (rudger@stela.org.br)

Denilson Sell, Dr. (denilson@stela.org.br)

José L. Todesco, Dr. (tite@stela.org.br)

Fernando B. Montenegro, M.Eng. (fborges@stela.org.br)

Paulo M. Selig, Dr. (pauloselig@gmail.com)

Resumo: As ferramentas de *Expertise Location* podem ser utilizadas amplamente na Gestão do Conhecimento para apoiar a identificação e o compartilhamento do conhecimento. Porém, manter os dados dos colaboradores de uma organização atualizados nessas ferramentas pode ser desafiador. Muitas vezes, os colaboradores precisam preencher as mesmas informações em diversos sistemas. Como uma abordagem alternativa para simplificar esse processo de atualização dos dados, este trabalho propõe uma ferramenta para a extração automática de perfis de especialistas a partir de seus documentos. Para construção dessa ferramenta, realizou-se uma revisão integrativa da literatura, a qual resultou na identificação das abordagens atuais para a extração de perfil de especialista. A partir dessas abordagens, elaborou-se um modelo conceitual e uma ferramenta baseada em Processamento de Linguagem Natural para a tarefa de extração de informações de perfil de especialistas que possam fornecer insumos para a identificação de seus conhecimentos e áreas de interesse. A implementação do protótipo resultou também em uma ferramenta de código aberto. Com o intuito de verificar a viabilidade do modelo e da ferramenta proposta, a partir de documentos de voluntários, foram executados testes comparando os perfis gerados pela ferramenta com os perfis presentes na rede social LinkedIn. Os resultados demonstraram que a ferramenta pode representar uma alternativa viável para a geração de perfis de especialistas de forma automática para apoiar as ferramentas de *Expertise Location* em uma organização. Conseqüentemente, a adoção desse modelo pode reduzir a necessidade de atualizações constantes dos perfis de especialistas de forma manual.

Palavras-chave: *Expertise Location*; Gestão do Conhecimento; Engenharia do Conhecimento; Processamento de Linguagem Natural.

1 Introdução

Segundo Schreiber et al. (1999), a cada dia, o conhecimento torna-se mais importante na economia e pesquisas apontaram que os executivos consideram o conhecimento como um dos fatores mais importantes para o sucesso de suas organizações. Em indústrias de manufatura, por exemplo, o conhecimento é relacionado à cerca de 65% a 75% do valor total dos produtos (SCHREIBER et al. 1999).

Dada a importância do conhecimento e a sua visão como um recurso-chave, a Gestão do Conhecimento (GC) torna-se uma atividade crucial nas organizações modernas (WIIG, 1997). O seu objetivo está no planejamento, na organização, na motivação e no gerenciamento de pessoas, de processos e de sistemas em organizações para assegurar que os seus ativos de conhecimento sejam empregados de forma eficaz (STEIL, 2007). Esses ativos de conhecimento incluem o conhecimento na forma de documentos impressos, conhecimentos armazenados em repositórios eletrônicos, o conhecimento dos funcionários sobre a melhor maneira de fazer o seu trabalho, o conhecimento que é realizado por equipes e o conhecimento que está incorporado em produtos, em processos e em relacionamentos da organização (SUN, 2005).

Segundo a APO - Asian Productivity Organization (2010), os *Expertise Location Systems* (ELS) – ou Sistemas para Localização de Especialistas - tem sido uma das principais ferramentas para a Gestão do Conhecimento, pois possuem o potencial de auxiliar as etapas de (a) identificação de conhecimento (b) criação de conhecimento, (c) armazenamento de conhecimento, (d) compartilhamento de conhecimento e (e) aplicação de conhecimento em um programa de GC. Recentemente, a APQC (2015), em seu relatório anual sobre tendências e prioridades em programas de GC, destacou a *Expertise Location* como uma das três principais iniciativas a serem implementadas em 2015 na área de Gestão de Conhecimento, demonstrando a importância da EL no contexto da GC na atualidade.

De acordo com os estudos de Balog et al. (2012), a construção de ELS ocorreu a partir da década de 1960 e uma série de sistemas foram desenvolvidos para suportar a busca por especialistas em determinados assuntos. Os ELSs abordam problemas típicos que grandes organizações enfrentaram: muitas vezes não é possível encontrar o conhecimento adequado que existe em algum lugar na organização. Esse problema pode induzir a organização a “reinventar a roda” para resolver um mesmo problema já resolvido anteriormente, só porque não se sabe quem fez o quê ou quem possui determinado conhecimento (BALOG et al., 2012).

Nesse contexto, a EL envolve ferramentas que permitem compartilhar o conhecimento existente, conectando pessoas que precisam de determinados conhecimentos às pessoas que possuem esses conhecimentos (APO, 2010). Às vezes, o sistema também auxilia o processo de formação de novas equipes de projetos, agregando os conhecimentos necessários. Essa ferramenta pode ser materializada em uma simples página amarela eletrônica ou sistemas mais sofisticados para buscar automaticamente expertise, com vistas a apoiar a localização e a conexão de pessoas que necessitam de um conhecimento a pessoas que detêm esse conhecimento (BALOG et al., 2012; APO, 2010; MAYBURY, 2006).

Embora seja uma ferramenta importante no papel da GC, muitos sistemas de EL ainda dependem do preenchimento das informações dos colaboradores de forma manual. Vários estudos sugerem que sistemas de EL devem incluir a integração de tecnologias de Inteligência Artificial, Text Mining, Web Mining, Processamento de Linguagem Natural, entre outras, para melhorar o processo de atualização dos perfis dos colaboradores com base em repositórios eletrônicos de documentos e informações provenientes da internet (BALOG et al., 2012; BALOG; RIJKE, 2007; BECERRA-FERNANDEZ, 2006; MARWICK, 2001).

Assim, os sistemas de EL possibilitariam uma menor dependência do preenchimento de informações de cada colaborador e permitiriam que os dados dos seus perfis fossem mais facilmente atualizados. Esses estudos também apontam que o problema com uso de perfis explícitos, isto é, preenchidos pelos próprios colaboradores, é a necessidade de atualização constante das informações dos colaboradores o que demanda mais tempo para essas atividades de atualização. Desse modo, é preferível reunir informações automaticamente, se possível a partir de fontes existentes na organização, para extrair informações relevantes sobre os especialistas da organização (BECERRA-FERNANDEZ, 2006).

De modo a corroborar com essa visão, os estudos realizados pela IDC - International Data Corporation (2010) descrevem que as fontes de informações não estruturadas (e.g. documentos em linguagem natural) representarão aproximadamente 90% de todo o conteúdo gerado nesta década (2010 a 2020). Por meio de técnicas da Engenharia do Conhecimento, essas informações, quando trabalhadas com a metodologia e ferramentas adequadas, podem explicitar ativos de conhecimento tais como redes de relacionamento de especialistas, competências e interesses dos especialistas da organização, informações sobre projetos de uma organização, entre outros (GONCALVES, 2006). Esses ativos do conhecimento podem, desse modo, auxiliar no processo de criação e atualização dos perfis de especialistas de uma organização.

Dessa forma, este trabalho se aprofunda no que diz respeito a extração de perfis de especialistas de forma automática a partir de documentos não estruturados para melhorar as ferramentas de *Expertise Location* utilizadas em iniciativas de Gestão de Conhecimento. Assim, teve por objetivo propor um modelo para a extração do perfil de especialista de forma automática a partir de documentos não estruturados para apoiar o desenvolvimento de soluções de *Expertise Location*. Esse objetivo foi decomposto em três objetivos específicos listados a seguir.

- Identificar as abordagens atuais que possibilitem a extração de perfis de especialistas a partir de documentos não estruturados no contexto de *Expertise Location*.
- Definir um modelo conceitual, de acordo com as abordagens existentes, que suporte a extração automática de perfis de especialistas e identifique os elementos necessários para a implementação do modelo.
- Desenvolver um protótipo funcional para verificar a viabilidade do modelo proposto utilizando documentos não estruturados para extração do perfil de especialista.

É importante destacar que o foco do modelo proposto foi a extração e criação de perfis de especialistas a partir de documentos. Não fez parte desta pesquisa a especificação de etapas para a construção de um sistema de *Expertise Location* em sua plenitude. Por exemplo, este modelo não previu como se faz a busca de especialistas de uma organização, mas se além quanto à criação do perfil que será apresentado como resultado de buscas nessas ferramentas.

2 Referencial teórico

Nesta seção, apresenta-se a conceitualização dos elementos utilizados neste trabalho com o intuito de situar e relacionar os temas quanto ao trabalho proposto. Esses elementos são conhecimento, gestão do conhecimento, engenharia do conhecimento e *expertise location*.

2.1 Conhecimento

Nessas últimas décadas, tem se observado um interesse cada vez maior em utilizar o conhecimento como um recurso organizacional, tornando-se um importante ativo nas organizações (ALAVI; LEIDNER, 2001; SCHREIBER et al., 1999; VENZIN; KROGH; ROOS, 1998; TEECE; PISANO; SHUEN, 1997; SPENDER; GRANT, 1996). Apesar de sua crescente importância, o conhecimento é um conceito bastante amplo e abstrato que tem sido debatido de forma epistemológica desde era clássica grega (ALAVI; LEIDNER, 2001; SCHREIBER et al., 1999). Sócrates em Teeteto, um diálogo sobre a natureza do conhecimento escrito em aproximadamente 369 A.C., definiu o conhecimento como “uma crença verdadeira justificada”, mas essa definição vem se tornando inadequada (ANAND; SINGH, 2011).

Anand e Singh (2011) observam que o termo conhecimento está associado a três diferentes significados: o estado de saber; capacidade de ação; e o lugar onde esse conhecimento está codificado. O estado do saber tem relação com a familiaridade de um indivíduo com fatos, métodos, técnicas e princípios. Já a capacidade de ação se refere ao entendimento e compreensão de fatos, métodos, técnicas e princípios na execução de uma ação (i.e., *know how*). E, por último, o lugar onde esse conhecimento está codificado aborda o como o conhecimento é articulado e capturado em forma de livros, manuais, código de computador, fórmulas, entre outros.

Anand e Singh (2011) destacam ainda que o conhecimento é gerado na mente humana e pode ser obtido a partir de fontes externas. Uma vez gerado, o conhecimento pode ser armazenado em um repositório para ser acessado e compartilhado por outros indivíduos ou grupos dentro de uma organização. Dessa forma, existem vários meios de armazenamento no

qual o conhecimento pode residir, como na mente humana, na organização, nos documentos e em computadores.

Quanto ao grau de acessibilidade, também conhecido pela dimensão acessibilidade do conhecimento (ANAND; SINGH, 2011), tem sido discutido por diversos autores ao longo dos anos (e.g. NONAKA; TAKEUCHI, 1995; DAVENPORT; PRUSAK, 1998; COOK; BROWN 1999). Os autores Nonaka e Takeuchi (1995), a partir dos estudos de Polanyi (1967), dividiram a acessibilidade em duas categorias: tácito e explícito.

2.2 Gestão do conhecimento

A crescente literatura de Gestão do Conhecimento oferece diversas definições, conceitos e formas de aplicá-la na prática (ANAND; SINGH, 2011; CHOI; LEE, 2002; ALAVI; LEIDNER, 2001; SMITH, 2001). No que tange as definições, é possível observar uma classificação de Gestão de Conhecimento nas dimensões de objetivos, demandas, processos, tecnologia de informação, estratégia, práticas, natureza holística e capital intelectual (ANAND; SINGH, 2011). Essas diferentes classificações representam uma visão geral do desafio que são as iniciativas de Gestão de Conhecimento nas organizações, uma vez que a GC incorpora uma série de disciplinas que são necessárias para a sua implantação em uma organização.

Contudo, a GC pode ser definida de forma geral como a coordenação sistemática de pessoas, tecnologias, processos e estrutura organizacional, com o objetivo de agregar valor por meio da reutilização do conhecimento em uma organização (DALKIR, 2005). Essa coordenação é obtida na criação, compartilhamento e aplicação do conhecimento, bem como na aplicação de lições aprendidas e de melhores práticas para uma aprendizagem organizacional contínua. Logo, essa reutilização do conhecimento cria novas capacidades organizacionais que permitem melhorar seu desempenho e atender melhor os seus clientes.

A coordenação sistemática de pessoas, tecnologias processos e estrutura organizacional refere-se ao processo da GC. Vários modelos são observados na literatura: Wiig (1993), Mayer e Zack (1996), Mc Elory (1999), Bukowitz e Williams (2003), Wong e Aspinwall (2004) e Lee et.al. (2005). De forma geral, esses modelos consideram quatro processos: criação do conhecimento; armazenamento e recuperação do conhecimento; transferência do conhecimento; e aplicação do conhecimento (ANAND; SINGH, 2011; ALAVI; LEIDNER, 2001).

Por fim, para apoiar os processos de conhecimento, (i.e. criação, armazenamento e recuperação, transferência e aplicação do conhecimento), há uma série de técnicas e ferramentas que contribuem na implantação da Gestão do Conhecimento em uma organização. A APO (2010) apresenta uma lista de métodos e ferramentas classificados tecnológicos (i.e., utilizam TI) e não tecnológicos (i.e., não utilizam TI) que suportam esses processos, conforme demonstrado no Quadro 1 a seguir.

Quadro 1 – Métodos e ferramentas de apoio aos processos da GC.

Métodos e ferramentas não tecnológicos	Métodos e ferramentas tecnológicas
<ul style="list-style-type: none"> • Brainstorming • Captura de ideias e aprendizagem • Assistência em pares • Revisão de aprendizagem • Revisão pós-ação • Storytelling • Estrutura Colaborativa Física • Café do conhecimento • Comunidades de prática • Taxonomia 	<ul style="list-style-type: none"> • Expertise Locator • Sistema de Gestão de Documentos • Bases de Conhecimento • Blogs • Redes Sociais • VOIP (Voice and Voice-over-Internet) • Buscas Avançadas • Clusters para Criação de Conhecimento • Estrutura Colaborativa Virtual

Fonte: Adaptado de APO (2010, p. 9).

2.3 Engenharia do Conhecimento

A Engenharia do Conhecimento teve o seu surgimento a partir da Inteligência Artificial como uma subárea dedicada à concepção, desenvolvimento e implementação de Sistemas Baseados em Conhecimentos (SBC). Nesse período, definido como Engenharia do Conhecimento Clássica, os especialistas extraíam o conhecimento e transferiam para uma base de conhecimento computacional (STUDER; BENJAMINS; FANSEL, 1998). Mais recentemente, a nova Engenharia do Conhecimento teve sua ênfase no paradigma de modelagem de conhecimento (SCHREIBER et al, 1999).

Partindo de um mesmo objeto – o conhecimento como ativo intangível - a Engenharia do Conhecimento e a Gestão do Conhecimento diferenciam-se pela forma com que se posicionam nas organizações. Para Kendal e Creen (2006), o termo “gestão” da Gestão do Conhecimento relaciona-se ao exercício executivo e à direção administrativa e de supervisão, enquanto que o termo "engenharia" está relacionada ao ato de construir, inventar ou planejar. Enquanto a GC estabelece uma visão estratégica para posicionamento do conhecimento organizacional, a EC concebe e desenvolve mecanismos inteligentes (modelos formais e tecnologias) para que a tecnologia da informação esteja alinhada a essa visão. Em termos de produtos, a GC propõe Sistemas de Gestão do Conhecimento (SGC) e a EC os Sistemas Baseados em Conhecimento (SBC).

A Engenharia do Conhecimento atual, segundo Schreiber (1999), oferece vários benefícios: permite identificar as oportunidades e os gargalos na forma como as organizações criam, distribuem e aplicam os seus conhecimentos, de modo a fornecer ferramentas para gestão do conhecimento organizacional; fornece métodos para obter uma compreensão completa das estruturas e processos na criação, distribuição e aplicação do conhecimento, possibilitando uma melhor integração da tecnologia da informação e a utilização do conhecimento; e, por último, possibilita a construção de sistemas de conhecimento mais fáceis de usar, bem estruturados arquiteturalmente e mais simples de se manter.

2.4 Expertise Location

Frequentemente, pessoas consultam outras pessoas para ajudá-las na busca do caminho correto para resolver um determinado problema, seja no trabalho, em ambientes acadêmicos, ou até mesmo em seus lares. Inclusive, pode existir um grande volume de informações disponíveis acerca do problema, mas contar com uma pessoa que possui conhecimento sobre o tema (i.e., um especialista), torna a sua resolução mais rápida (BALOG et al., 2012).

Nesse contexto, as ferramentas de *Expertise Location* (i.e. *Expertise Location Systems*, ou *Expertise Locator*) contribuem na identificação e localização de especialistas que possuem

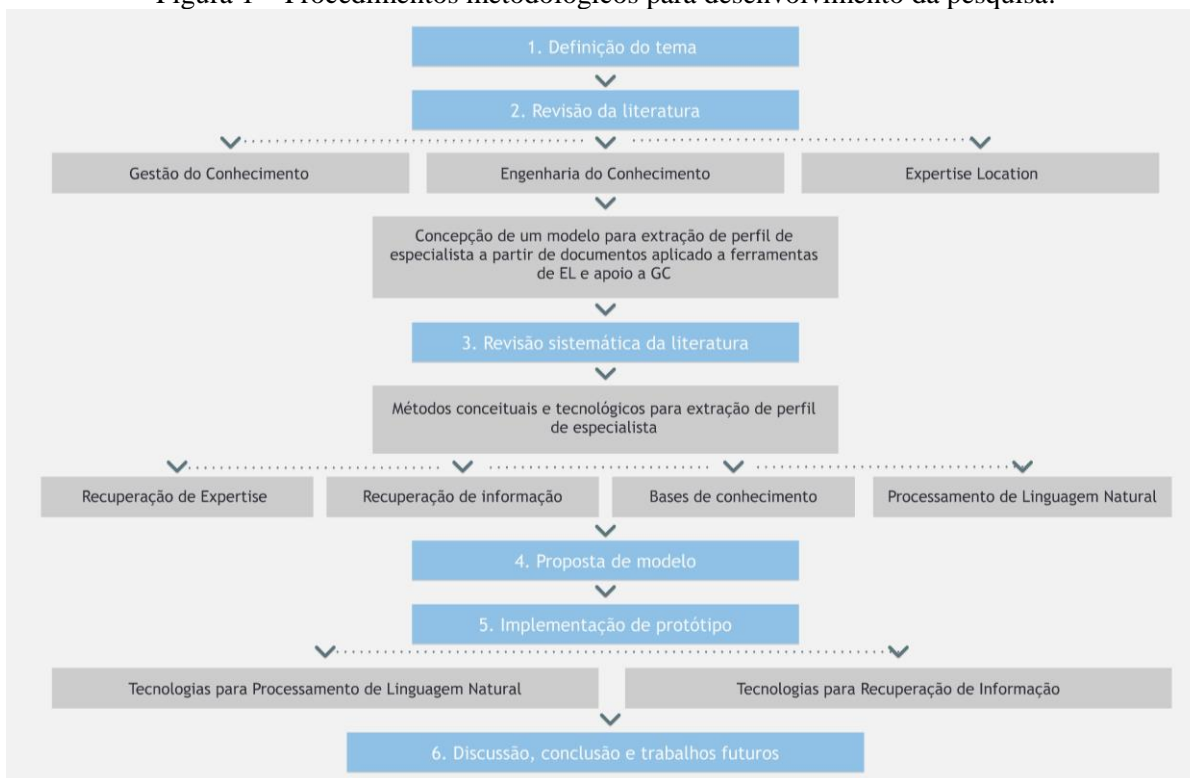
conhecimento e experiência necessários às outras pessoas e têm ganhado destaque na Gestão do Conhecimento (APQC, 2015; BALOG et al., 2012; APO, 2010; MARWICK, 2001; MCDONALD; ACKERMAN, 1998). Principalmente, em grandes empresas, onde saber quem são as pessoas que possuem experiências, habilidades e conhecimento é difícil pela quantidade de pessoas (APO, 2010).

De acordo com Balog et al. (2012), as evoluções nos sistemas de *Expertise Location* nas últimas décadas destacaram o surgimento de duas categorias: *Expert Finding Systems* (i.e., Sistemas de Busca de Especialistas) e *Expert Profiling Systems* (i.e., Sistemas de Perfis de Especialistas). Enquanto o objetivo dos Sistemas de Busca de Especialista é justamente encontrar um especialista em um determinado tópico, o objetivo dos Sistemas de Perfis de Especialistas é apresentar quais são os tópicos que uma pessoa é especialista.

3 Metodologia

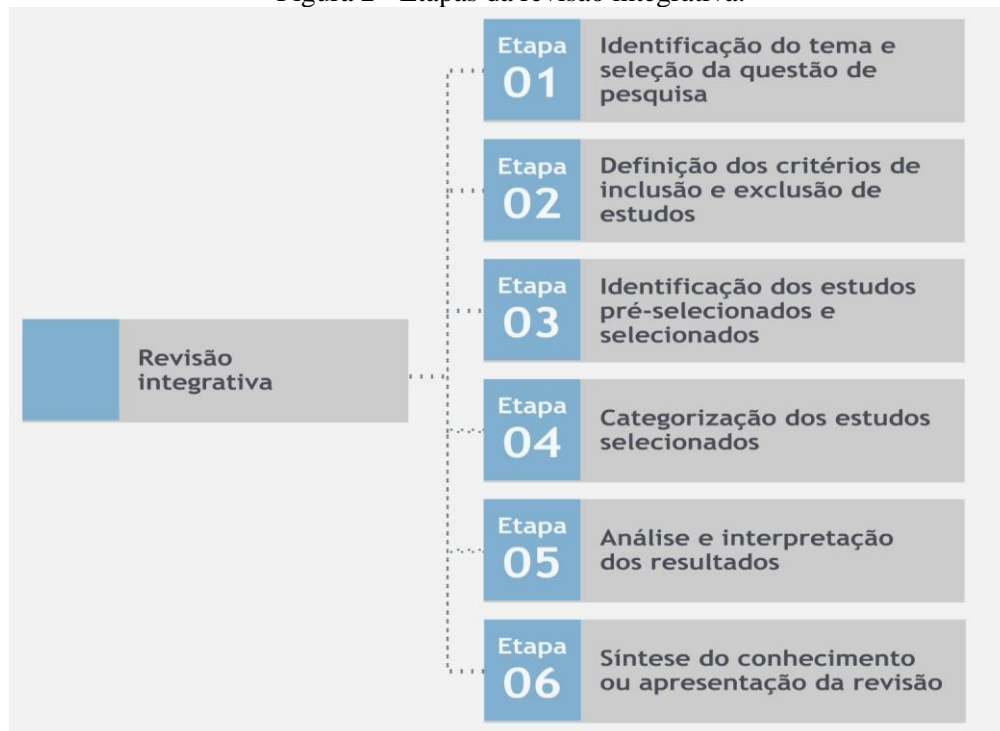
Para a realização dessa pesquisa, utilizou-se três etapas: o levantamento bibliográfico para identificação das abordagens atuais que permitem a extração de perfis de especialistas; a concepção do modelo conceitual e a implementação de um protótipo de ferramenta para verificar a viabilidade do modelo proposto. Dessa forma, a natureza desta pesquisa se enquadrou em pesquisa aplicada e exploratória por prover uma aplicação prática e dirigida à solução de problemas específicos (SILVA; MENEZES, 2005). Quanto à modalidade, esta pesquisa é caracterizada como pesquisa tecnológica, com o objetivo de projetar artefatos tecnológicos à luz do conhecimento científico. Para embasamento e exploração dos trabalhos relacionados ao levantamento bibliográfico desta pesquisa, adotou-se a revisão integrativa da literatura (i.e., revisão sistemática). A Figura 1 apresenta os procedimentos metodológicos para se alcançar os objetivos desta pesquisa e a Figura 2 os procedimentos da revisão sistemática com base no modelo de Botelho, Cunha e Macedo (2011).

Figura 1 – Procedimentos metodológicos para desenvolvimento da pesquisa.



Fonte: Elaboração própria.

Figura 2 - Etapas da revisão integrativa.



Fonte: Adaptado Botelho, Cunha e Macedo (2011, p. 129).

É importante destacar que a revisão integrativa da literatura é o tipo mais amplo dos métodos de avaliação de pesquisa, que permitem a inclusão simultânea de investigação experimental e não experimental, a fim de melhor compreender um fenômeno de interesse do pesquisador (WHITEMORE, 2005). O termo “integrativa”, segundo Botelho, Cunha e Macedo (2011), tem origem na integração de opiniões, conceitos ou ideias provenientes das pesquisas utilizadas no método. Assim, revisões integrativas também podem combinar dados teóricos e empíricos, além de possibilitarem a aplicação em diferentes propósitos como definir conceitos, rever teorias, avaliar evidências e analisar questões metodológicas de um tema específico (BROOME, 1993 apud WHITEMORE; KNAFL, 2005). A seção a seguir detalha os passos executados na revisão integrativa da literatura e seus resultados quanto aos estudos selecionados.

3.1.1 Execução do método da revisão integrativa

A primeira etapa para a realização de uma revisão integrativa foi a identificação do tema e a seleção da questão de pesquisa. Essa é a etapa mais importante da revisão integrativa e é a etapa que guia todas as etapas subsequentes. Além da pergunta de pesquisa bem definida, também há a necessidade da especificação das palavras-chaves utilizadas na busca em base de dados. Os passos seguintes definem os critérios de exclusão e inclusão, de identificação dos estudos pré-selecionados e selecionados, de categorização dos estudos e de análise e interpretação dos resultados. O último passo da revisão é a síntese do conhecimento ou a apresentação da revisão.

Como resultado da primeira etapa, foram definidos os termos de busca listados a seguir. Utilizou-se estratégia de combinação das palavras-chaves na busca de documentos da base Scopus:

- busca pelo grupo de palavras-chaves e seus termos correlatos de “Perfil de especialista” e “*Expertise Location*” e “Gestão do Conhecimento” e “Engenharia do Conhecimento”;
- busca pelo grupo de palavras-chaves e seus termos correlatos de “Perfil de especialista” e “*Expertise Location*”;
- busca pelo grupo de palavras-chaves e seus termos correlatos de “Perfil de especialista” e “Gestão do Conhecimento”;
- busca pelo grupo de palavras-chaves e seus termos correlatos de “Perfil de especialista” e “Engenharia do Conhecimento”;
- combinação da lista de resultados e remoção de redundâncias.

A partir dos termos listados anteriormente, chegou-se ao resultado listado no Quadro 2 a seguir. Após uma leitura crítica dos trabalhos resultantes, foram selecionados os artigos listados no Quadro 3 e que fomentaram a identificação das abordagens atuais para extração do perfil de especialista a partir de documentos.

Quadro 2 – Total de estudos encontrados na base Scopus a partir das palavras-chaves de pesquisa.

Grupos de termos de busca	Estudos encontrados na Scopus
“Perfil de especialista” e “ <i>Expertise Location</i> ” e “Gestão do Conhecimento” e “Engenharia do Conhecimento”	2
“Perfil de especialista” e “ <i>Expertise Location</i> ”	20
“Perfil de especialista” e “Gestão do Conhecimento”	52
“Perfil de especialista” e “Engenharia do Conhecimento”	49
Total de documentos sem redundância	70

Fonte: elaboração própria.

Quadro 3 – Estudos selecionados na revisão integrativa.

#	Autores	Título do trabalho	Ano de publicação
1	Balog K., De Rijke M.	Determining expert profiles (with an application to expert finding)	2007
2	Becerra-Fernandez I.	Searching for experts on the Web: A review of contemporary expertise locator systems	2006
3	Li M., Liu L., Li C.-B.	An approach to expert recommendation based on fuzzy linguistic method and fuzzy text classification in knowledge management systems	2011
4	Macdonald C., Ounis L.	Expertise drift and query expansion in expert search	2007
5	Yang K.-W., Huh S.-Y.	Automatic expert identification using a text categorization technique in knowledge management systems	2008
6	Liu D.-R., Chen Y.-H., Kao W.-C., Wang H.-W.	Integrating expert profile, reputation and link analysis for expert finding in question-answering websites	2013
7	Yao L., Tang J., Li J.	A unified approach to researcher profiling	2007

#	Autores	Título do trabalho	Ano de publicação
8	Yang K.-H., Chen C.-Y., Lee H.-M., Ho J.-M.	EFS: Expert finding system based on wikipedia link pattern analysis	2008
9	Afzal M.T., Maurer H.	Expertise recommender system for scientific community	2011
10	Boeva V., Krusheva M., Tsiporkova E.	Measuring expertise similarity in expert networks	2012
11	Sohail A., Afzal M.T., Ahmad J.	Discovery of medical experts	2013

Fonte: elaboração própria.

4 Resultados

Com base no objetivo desta pesquisa e na revisão integrativa da literatura, obtiveram-se insumos para a identificação de elementos que permitissem a extração de perfil de especialista a partir de documentos. Esse resultado satisfaz o primeiro objetivo específico que é (a) identificar as abordagens atuais que possibilitem a extração de perfis de especialistas a partir de documentos não estruturadas no contexto de *Expertise Location*. O seu conteúdo é detalhado na seção 4.1. A partir desse resultado, se propôs um modelo de extração de perfil de especialista que cumpre o objetivo específico de (b) definir um modelo conceitual, de acordo com as abordagens existentes, que suporte a extração automática de perfis de especialistas e identifique os elementos necessários para a implementação do modelo. Esse resultado é detalhado na seção 4.2. Por fim, chegou-se no terceiro objetivo específico que é (c) desenvolver um protótipo funcional para verificar a viabilidade do modelo proposto utilizando documentos não estruturadas para extração do perfil de especialista, detalhado na seção 4.3.

4.1 Identificação das abordagens para extração de perfil de especialista

Essa seção, resultado da revisão integrativa da literatura sobre os temas de perfil de especialista, Gestão do Conhecimento, Engenharia do Conhecimento e *Expertise Location*, apresenta um panorama sobre dois aspectos importantes a esta pesquisa: a contribuição das ferramentas de *Expertise Location* à Gestão do Conhecimento; e os métodos conceituais e tecnológicos para a extração de perfis de especialistas a partir de documentos.

Sobre o primeiro aspecto, a contribuição das ferramentas de *Expertise Location* à Gestão do Conhecimento, observa-se um conjunto de trabalhos que correlacionam esses temas e onde se destacam os trabalhos realizados por Becerra-Fernandez (2006) da Universidade Internacional da Florida dos EUA; Balog e Rijke (2007) da Universidade de Stavanger, na Noruega, e da Universidade de Amsterdam, na Holanda, respectivamente; Yang e Huh (2008) da Universidade de Keimyung na Coréia do Sul; Li, Liu e Li (2011) da Universidade do Petróleo, na China. Nessa revisão destaca-se que os estudos selecionados na busca de trabalhos com os temas de EL, GC e EC, no contexto de perfis de especialistas, abordam as ferramentas de *Expertise Location* como Sistemas de Gestão do Conhecimento.

Com relação ao segundo aspecto, métodos conceituais e tecnológicos para a extração de perfis de especialistas a partir de documentos, observou-se um diversificado conjunto de soluções para extrair informações e conhecimentos de documentos com o objetivo de contribuir com a *Expertise Location*, seja pela extração e geração de perfil de especialista, ou seja, pela contribuição na busca e recomendação de especialistas. É importante ressaltar que, apesar de descrever como métodos conceituais e tecnológicos, essa seção abordou também

processos, técnicas, algoritmos e tecnologias que pudessem contribuir com o tema e fornecer subsídios para a proposta de modelo de extração de perfis de especialistas deste trabalho.

Observa-se que trabalhos analisados, que apresentam métodos conceituais e tecnológicos, também descrevem o uso de técnicas provenientes da área de Recuperação da Informação e da Inteligência Artificial. Destacam-se a utilização de Processamento de Linguagem Natural e suas técnicas de tokenização, normalização, remoção de *stop-words*, *lemmatization*, *stemming*, *Part-Of-Speech Tagging* e *Named Entity Recognition*; *Web-mining*; *Support Vector Machines (SVM)*; *2-Tuple Linguistic Model*; C-value/NC-value; e TF-IDF.

Entretanto, observou-se que esses trabalhos, apesar de fornecerem visões sobre diferentes propostas conceituais e tecnológicas para a exploração do tema de *Expertise Location*, não apresentam soluções para uma questão importante quando se trata de *expertise*: a dimensão tempo. Saber se um especialista explorou ou ainda explora um determinado tema pode ser um fator crucial em um momento de contato, por exemplo. Crescentes demonstrações de exploração de um tema, quando observado pela dimensão tempo, também podem indicar uma maior familiaridade de um especialista nesse tema.

4.2 Modelo de extração de perfil de especialista proposto

A proposta de modelo de extração de perfil de especialista foi concebida com base nas abordagens atuais observadas na literatura. Essas abordagens, conceituais e tecnológicas, que são compostas por processos, técnicas, algoritmos e tecnologias, forneceram subsídios para a definição do modelo proposto. Além disso, as referidas abordagens, principalmente pautadas nos trabalhos realizados por Becerra-Fernandez (2006), Balog e Rijke (2007), Yang e Huh (2008) e Li, Liu e Li (2011) e complementadas pelos trabalhos de Gonçalves (2006), Ceci (2010) e Bovo (2011), permitiram a criação do modelo proposto.

De forma resumida, este modelo proposto espera como entrada os diversos documentos produzidos por um especialista e, com base no conteúdo identificado nesses documentos, cria uma série de instrumentos visuais que exploram essas informações. Esse modelo é composto por cinco etapas, (a) extração do conteúdo, (b) reconhecimento de entidades, (c) extração e armazenamento de entidades, (d) associação de entidades e, por fim, (e) extração do perfil de especialista.

4.3 Protótipo do modelo proposto

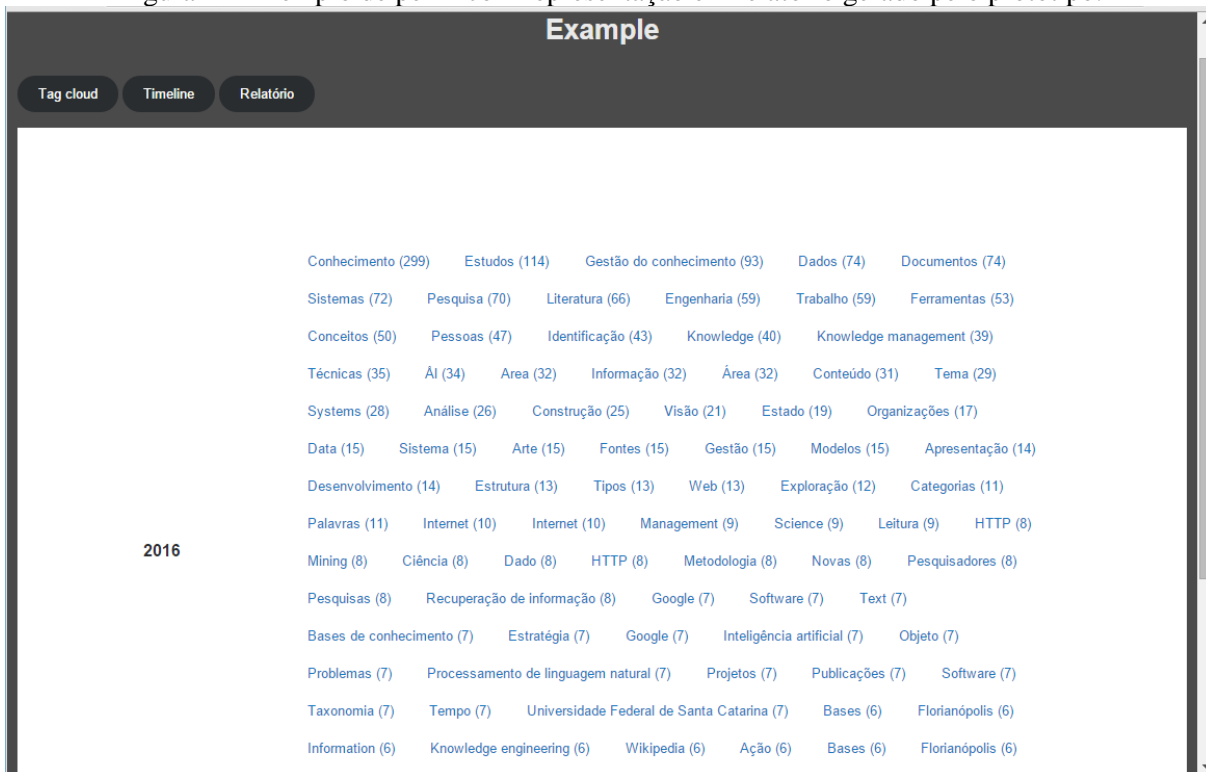
O modelo proposto nesta pesquisa foi concebido a partir das abordagens identificadas na literatura que possibilitam a extração e a criação automática do perfil. Este modelo especifica etapas para o processamento de documentos de especialistas com vistas à extração de conceitos e à apresentação dessas informações em instrumentos visuais como *tag cloud* e *timeline*. A Figura 3 e Figura 4 ilustram esses resultados.

Figura 3 – Exemplo de perfil com representação em tag cloud gerado pelo protótipo.



Fonte: elaboração própria.

Figura 4 – Exemplo de perfil com representação em relatório gerado pelo protótipo.



Fonte: elaboração própria.

Essa implementação corresponde ao último objetivo específico desta pesquisa. Assim, foi realizada a geração de três perfis especialistas que se colocaram à disposição para executar a ferramenta a partir de seus documentos. Os perfis extraídos desses especialistas permitem

observar a viabilidade do modelo proposto. Para fins de demonstração, foram selecionados os conteúdos de *tag cloud* dos resultados da extração do perfil. Os três especialistas voluntários concluíram que os perfis gerados estavam alinhados com as atividades e os conhecimentos pessoais de cada um deles. Esses perfis, além de possibilitarem a demonstração de viabilidade de implementação do modelo conceitual proposto, possibilitam, ainda, a sua comparação com perfis existentes em sites como o da rede social LinkedIn. Tal comparação permitiu que se verifique, por exemplo, o grau de aproximação que um perfil gerado a partir dos documentos de um especialista tem em relação a um perfil preenchido e mantido de forma manual. Como resultado, obtiveram-se os índices 0,6569 para o perfil do especialista #1, 0,7796 para o perfil do especialista #2 e, por fim, 0,6858 para o especialista #3. Esses resultados geraram uma média de **0,7074** de similaridade entre os perfis gerados com os perfis mantidos no LinkedIn.

5 Conclusões

Esta pesquisa teve por objetivo propor um modelo para a extração do perfil de especialista de forma automática, a partir de documentos em linguagem natural, para apoiar a Gestão do Conhecimento na aplicação de ferramentas de *Expertise Location* em organizações. Um dos objetivos específicos deste trabalho foi justamente identificar as atuais abordagens que possibilitem a extração de perfis de especialistas a partir de documentos não estruturadas no contexto de *Expertise Location*. Tal objetivo foi alcançado por meio da realização de uma revisão sistemática da literatura, que resultou em 70 trabalhos para análise. Desses 70 trabalhos, foram selecionados 11, que contribuíram para a identificação dessas abordagens atuais. De posse da visão das abordagens atuais que possibilitaram atingir o objetivo desta pesquisa, foi proposto um modelo conceitual que combinou as várias propostas dos trabalhos observados na literatura. Inclusive, combinaram-se abordagens de Bovo (2011) e Ceci (2010). Dessa maneira, foi atingido também o segundo objetivo específico deste trabalho, que era definir um modelo conceitual. Esse modelo teve como proposta a realização de cinco etapas para o processamento de documentos de um especialista até que fossem extraídas as referências aos possíveis conhecimentos (i.e. tácito ou explícito) e interesses que esse especialista possui. Finalmente, com a proposição do modelo conceitual para a extração de perfil de especialista a partir de documentos, implementou-se um protótipo funcional que materializava o modelo proposto neste trabalho. Tal protótipo, gerou ainda uma ferramenta denominada Expert Profiling, que foi disponibilizada com código fonte aberto.

Com essa implementação do protótipo, atingiu-se, então, o último objetivo específico desta pesquisa que foi desenvolver um protótipo funcional para verificar a viabilidade do modelo proposto. Além disso, a pesquisa apresentou exemplos para a demonstração de viabilidade do modelo proposto, extraíndo perfis de forma automática de três especialistas voluntários, cujos perfis foram comparados aos perfis públicos dos especialistas mantidos na rede social LinkedIn. Dando sequência a esse procedimento, ao se realizar uma verificação de similaridade com o algoritmo de cosseno dos perfis extraídos pelo protótipo em relação ao perfil mantido na rede social, chegou-se ao índice de **0,7074** de similaridade. Esse resultado demonstrou que o modelo conceitual e o protótipo funcional podem representar uma robusta opção para a geração automática de perfis de especialista a partir de seus documentos.

6 Referências bibliográficas

ALAVI, M., LEIDNER, D. E. Knowledge management and knowledge management systems: conceptual foundations and research issues. **MIS Quarterly**, Minnesota, v. 25, n. 1, p. 107-136, 2001.

ANAND, A; SINGH, M. D. Understanding Knowledge Management: A Literature Review", **International Journal of Engineering Science and Technology**, v. 3 N. 2, p. 926- 939. Feb. 2011.

APO - Asian Productivity Organization. **Knowledge Management Tools and Techniques Manual**. Japan: Hirakawacho, Chiyoda-ku, 2010.

APQC - American Productivity & Quality Center. **2015 Knowledge Management Priorities**. Disponível em < https://www.apqc.org/knowledge-base/download/340305/K05834_2015_KM_Priorities_ExecSum.pdf >. Acesso em 02 abr. 2015.

BALOG, K, et al. Expertise Retrieval. **Foundations and Trends in Information Retrieval**, [S.L.], v. 6, Nos. 2–3, p. 127-256. 2012.

BECERRA-FERNANDEZ, I. Searching for experts on the web: a review of contemporary expertise locator systems. **ACM Transactions on Internet Technology**, New York, v. 6, n. 4, p. 333-355, 2006.

BOTELHO, L. L. R.; CUNHA, C. C. A.; MACEDO, M. O método da revisão integrativa nos estudos organizacionais. **Gestão e Sociedade**, Belo Horizonte, v. 5, n. 11, p. 121-136, 2011.

BOVO, A. B. **Um modelo de descoberta de conhecimento inerente à evolução temporal dos relacionamentos entre elementos textuais**. 2011. Tese (Doutorado em Engenharia de Produção) - Universidade Federal de Santa Catarina, Florianópolis, 2011.

CECI, F. **Um Modelo semi-automático para a construção e manutenção de ontologias a partir de bases de documentos não estruturados**. 2010. 131 f. Dissertação (Mestrado em Engenharia de Gestão do Conhecimento) – Universidade Federal de Santa Catarina. Florianópolis, 2010.

CHOI, B.; LEE, H. Knowledge management strategy and its link to knowledge creation process. **Expert Systems with Applicatios**, v. 23, p. 173–187 2002.

COOK, D. J.; MULROW, C. D. RAYNES, R. B. Systematic reviews: synthesis of best evidence for clinical decisions. **Annals of Internal Medicine**, Philadelphia, v. 126, n. 5, p. 376-380, Mar. 1997.

DALKIR, K. **Knowledge management in theory and practice**. Burlington: Elsevier Butterworth-Heinemann, 2005.

DAVENPORT, T.H.; PRUSAK, L. **Working Knowledge: How Organizations Manage What They Know**. Boston: Harvard Business School Press, 1998.

GONÇALVES, A. L. **Um modelo de descoberta de conhecimento baseado na correlação de elementos textuais e expansão vetorial aplicado à engenharia e gestão do**

conhecimento. 2006. 196 f. Tese (Doutorado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 2006.

IDC - International Data Corporation. **Extracting Value from Chaos**. Boston: Emc Corporation, 2010. 12 p.

KENDAL, S.; CREEN, M. **An Introduction to Knowledge Engineering**. Springer, 2006.

LEE, K. C. et al. Knowledge Management: measuring Knowledge Management performance. **Information and Management**, 42(3), 469-82. 2005.

MARWICK, A. Knowledge management technology. **IBM Systems Journal**, v. 40, n. 4, p.814-830, 2001.

MAYBURY, M. **Expert Finding Systems**. MITRE - Center for Integrated Intelligence Systems. Bedford, Massachusetts. 2006.

MAYER, M.; ZACK, M. The design and implementation of information Products. **Sloan Management Review**, v. 37, n. 3, p. 45 – 59, 1996.

NONAKA, I; TAKEUCHI, H. **The knowledge-creating company: How Japanese companies create the dynamics of innovation**. Oxford: Oxford University Press, 1995.

POLANYI, M. **The tacit dimension**. New York: Doubleday, 1967.

SCHREIBER, G. et al. **Knowledge engineering and management: the CommonKADS methodology**. Massachusetts: MIT Press, 1999.

SILVA, E. L. da; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. 4. ed. Florianópolis: UFSC, 2005.

SMITH, E. A. The role of tacit and explicit knowledge in the workplace. **Journal of Knowledge Management**, v. 5, n. 4, p. 311 – 321, 2001.

SPENDER, J. C. Making knowledge the basis of a dynamic theory of the firm. **Strategic Management Journal**, v. 17, p. 45-62, 1996.

STEIL, A. V. **Estado da arte das definições de gestão do conhecimento e seus subsistemas**. Technical Report. Florianópolis: Instituto Stela, 2007.

STUDER, R.; BENJAMINS, R; FANSEL, D. Knowledge Engineering: Principles and methods. **Data & Knowledge Engineering**, v. 25, n. 1, p. 161-197, 1998.

TEECE, D. J.; PISANO, G.; SHUEN, A. Dynamic capabilities and strategic management. **Strategic Management Journal**, v. 18, n. 7, p. 509-533,1997.

VENZIN, M.; KROGH, G; ROOS, J. Future Research into Knowledge Management. In: KROGH, G; ROOS, J; KLEINE, D. **Knowing in Firms**. Londres: Sage Publication, 1998.

WHITEMORE, Robin; KNAFL, Kathleen. The integrative review: updated methodology. **Journal of advanced nursing**, v. 52, n. 5, p. 546-553, 2005.

WIIG, K. M. Integrating Intellectual Capital and Knowledge Management. **Long Range Planning**, v. 30, n. 4, P. 399-405, 1997.

WIIG, K. M. **Knowledge Management Foundations: Thinking about Thinking – How People and Organizations Create, Represent and use Knowledge**. Arlington: Schema Press, 1993.