

# Recuperação Contextual de Informação na Web para Apoiar Aprendizagem Colaborativa em Redes Sociais

**Autor: Eduardo Fritzen**

**Orientadores: Sean W. M. Siqueira & Leila C. V. de Andrade**

Departamento de Informática Aplicada (DIA/CCET), Universidade Federal do Estado do Rio de Janeiro (UNIRIO): Av. Pasteur, 458, Urca, Rio de Janeiro, Brasil, 22290-240

{eduardo.fritzen, sean, leila}@uniriotec.br

***Abstract.** This work considers the documents retrieval relevance on the Web can be important to encourage collaborative learning dynamics based on discussion. To improve and contextualize the documents retrieval on the Web, this paper proposes: (i) the context modeling from the extraction of messages in social networking groups and (ii) use of the context to improve the relevance in the documents retrieval on the Web. A case study demonstrated that the capture of the enriched context message using the discussion messages can improve the relevance of search results on the Web and contribute to the discussions.*

***Resumo.** Este trabalho considera que a relevância na recuperação de documentos na Web pode ser primordial para estimular a colaboração em dinâmicas de aprendizagem baseadas em discussão. Para melhorar e contextualizar a recuperação de documentos na Web, este artigo propõe: (i) a modelagem do contexto a partir da extração das mensagens em grupos de rede social e (ii) uso do contexto para melhorar a relevância na recuperação de documentos na Web. Um estudo de caso demonstrou que a captura do contexto enriquecido usando mensagens de discussão pode melhorar a relevância dos resultados das buscas na Web e contribuir com as discussões.*

## 1. Introdução

A necessidade de usuários adquirirem documentos a partir da *Web* que os apoiem a participar e entender os assuntos discutidos em suas redes sociais foi o que motivou este trabalho. A recuperação de informação desejada a partir da *Web* não é uma tarefa trivial (KOBAYASHI e TAKEDA, 2000). Os resultados dos buscadores não necessariamente refletem o que se busca em determinado momento. Além disto, o fato de muitos estudantes já usarem sítios de redes sociais (69,3% se enquadra na faixa etária entre 16 e 35 anos (IBOPE, 2011)), fez com que professores começassem a se familiarizar com esta tendência, para usá-la a seu favor, explorando novas possibilidades para melhorar os resultados da aprendizagem. Existem muitos exemplos de sítios de redes sociais usados por professores e alunos como provedores de comunicação (MORA-SOTO, 2009) (MANSUR *et al.*, 2011) (DOTTA, 2011) (WANG *et al.*, 2011).

Sítios de redes sociais geralmente fornecem recursos para o compartilhamento de conteúdo, como a publicação de documentos e *links*, e também a troca de mensagens usando programas de comunicação. Estas funcionalidades permitem o uso destes sítios

como um ambiente oportuno à aprendizagem colaborativa. Em algumas plataformas de redes sociais é possível criar grupos de usuários que compartilham o mesmo interesse (por exemplo, alunos de uma classe ou um curso). Compartilhamento de conteúdo permite fornecer o material de aprendizagem necessário para um curso, enquanto o programa de comunicação permite a troca de ideias.

Entretanto, ao observar o resultado das buscas realizadas nos motores de buscas disponibilizados nos dois sítios de redes sociais mais populares do Brasil, *Facebook*<sup>1</sup> e *Orkut*<sup>2</sup> (COMSCORE, 2011), percebe-se que o *Orkut* restringe os resultados da busca a elementos de sua própria rede, como pessoas e comunidades, enquanto no *Facebook* os resultados da *Web* são idênticos aos obtidos com o sítio de seu motor de busca *Web*, independente se a pesquisa originou-se no sítio do buscador ou a partir de um grupo de discussão no sítio da rede social. Este problema se torna mais crítico quando as redes sociais envolvem atividades de trabalho ou de ensino/aprendizagem, pois as pesquisas na *Web* realizadas a partir destes ambientes visam auxiliar o desenvolvimento dessas atividades. Melhorar a precisão dos resultados das buscas pode estimular as discussões no grupo e promover a colaboração. Por exemplo, em um ambiente educacional, a relevância dos documentos obtidos a partir da *Web* pode apoiar o aprendizado, uma vez que o conteúdo retornado será possivelmente mais adequado à necessidade de informação do aprendiz.

Em geral, a expressão de busca (conjunto de palavras-chave) informada pelo usuário é composta por poucos termos e, portanto, pouco representativa do contexto do domínio do usuário e suscetível a ambiguidades que degradam o processamento e resultado da busca (ou seja, os documentos que compõem o resultado não são tão relevantes para o usuário de acordo com seu contexto). Segundo levantamento feito pela empresa Experian Hitwise, entre 26/08/2011 e 26/11/2011, 66,55% das buscas utilizam no máximo três palavras (EXPERIAN, 2011). Outra informação relevante sobre o perfil das buscas na *Web* é que apenas os cinco (5) primeiros resultados são efetivamente acessados pelos usuários (SPINK e JANSEN, 2004), o que reforça a importância da priorização dos resultados. Além da quantidade, a qualidade dos termos também deve ser considerada, frente à dificuldade que grande parte dos usuários tem em definir quais palavras-chave são boas representantes para os documentos de seu interesse e que deverão compor a sua expressão de busca (FERNEDA, 2003).

O uso do contexto pode melhorar a relevância dos resultados a partir de ajustes na consulta do usuário. Estes ajustes podem ser, por exemplo, o uso da técnica de expansão de consultas. A técnica de expansão de consultas, citada na área de recuperação de informação (CARPINETO e ROMANO, 2012) e adotada neste trabalho, consiste em adicionar termos à consulta original, a fim de diminuir a ambiguação e promover maior acurácia nos resultados. Quanto mais termos, e mais representativos estes termos forem, maior a possibilidade de encontrar documentos relevantes (YATES e NETO, 1999).

---

<sup>1</sup> <http://www.facebook.com/>

<sup>2</sup> <http://www.orkut.com.br/>

A intenção é tornar a recuperação de informação sensível ao contexto das discussões (por intermédio do uso das mensagens para modelagem do contexto), oferecendo, portanto, resultados de busca contextualizados. Contudo, sítios de redes sociais possuem características que dificultam a captura do contexto, como ausência de conteúdo no início das discussões e mensagens escritas de maneira informal, com o uso de abreviações e linguagem coloquial, expressas com poucas palavras. Para transpor esse obstáculo, é proposta a captura do contexto a partir do enriquecimento destas mensagens em dados abertos e o uso deste contexto para melhorar a consulta do usuário, fornecendo-lhe conteúdos mais adequados a partir da *Web*. O simples uso de medidas de frequência para extrair termos relevantes em uma discussão pode não ser suficiente para prover bons resultados, principalmente no início das discussões. Assim, este trabalho deve lidar com problema do arranque a frio (*cold start problem*). Para o processamento das consultas, buscaram-se algoritmos que analisassem a semelhança conceitual entre a expressão de busca e os documentos que compõem o contexto enriquecido. Isso faz com que os termos sugeridos sejam sensíveis ao contexto de domínio e relacionados à expressão de busca do usuário.

O restante deste trabalho está organizado da seguinte forma: a seção 2 apresenta os trabalhos relacionados, a seção 3 apresenta o trabalho preliminar para a captura do contexto a partir de discussões em rede social e seus principais resultados, a seção 4 apresenta a proposta para captura de contexto durante discussões, a seção 5 detalha o estudo de caso, sua metodologia e forma de avaliação, a seção 6 as métricas utilizadas para avaliação, a seção 7 apresenta os resultados a partir das métricas e a seção 8 traz as considerações finais.

## **2. Trabalhos Relacionados**

Este artigo relaciona-se com trabalhos que, independente de técnica, usam algum modelo de conhecimento para tornar a recuperação da informação sensível ao contexto. Informações de domínio do conhecimento, de processo de negócios, fornecidas pelo usuário (explicitamente) ou obtidas a partir de seus comportamentos (implicitamente) podem ser usadas para modelar o contexto. Esta modelagem pode exigir grande esforço humano, como a criação de ontologias por especialistas, preenchimento de preferências e marcação de documentos (modelagem manual) ou técnicas computacionais que supram (modelagem automática) ou minimizem (modelagem semiautomática) o esforço humano, como a análise de cliques, *corpus* textual, dados históricos, sensores ou outras maneiras ubíquas (BHOGAL *et al.*, 2007).

Na literatura é possível encontrar muitas propostas que visam à melhoria da recuperação da informação, com o uso de expansão de consultas sensíveis ao contexto. Serão descritos alguns desses trabalhos, de modo especial, aqueles que empregam o *corpus* textual para a geração do modelo de contexto e, a partir deste modelo, propõe a recuperação da informação mais adequada às necessidades do usuário. A abordagem apresentada nesta dissertação difere das demais na origem dos dados (discussões em plataformas de rede social) e na forma em que o contexto é modelado (enriquecimento das discussões em dados abertos).

Prates e Siqueira (2011a, 2011b) propõem um método automático para apoiar a contextualização das atividades de busca na *Web*. Sua abordagem realiza a extração de

conhecimento dependente de *corpus* (análise e processamento textual) para então utilizar técnica de expansão das consultas e as executa em motores de busca Web. A abordagem para a expansão de consulta assume que os termos mais frequentes em documentos que são representativos de um domínio têm maior probabilidade de ocorrer em sítios e documentos disponíveis na internet e são relevantes e relacionados a este domínio. Além disso, considera que um documento pode tratar de diversos assuntos e com isso, a criação da lista de termos, usada na expansão das consultas, faz uso de técnicas de segmentação de tópicos por assuntos antes da etapa de agrupamento (*clustering*), aplicadas a um conjunto de arquivos de contexto representativos do domínio. O usuário pode visualizar e usar para a expansão os termos de todos os clusters, não só os termos do cluster que mais se relaciona aos termos da consulta. Kang *et al.* (2010) realizaram o agrupamento dos *x* primeiros documentos retornados pela consulta original e extraíram os termos mais relevantes de cada agrupamento. O usuário deve selecionar um dos *x* agrupamentos sugeridos, de acordo com o que julgar relevante. Os termos do agrupamento selecionado são somados aos termos da consulta original e o resultado da consulta expandida é apresentado ao usuário. Ambrósio *et al.* (2009) utilizaram técnicas de mineração de textos em um conjunto de documentos (apresentações) para recomendar documentos armazenados em um repositório e também sugerir a expansão da consulta a buscas na Web.

### **3. Melhoria da relevância das buscas a partir de discussões em redes sociais**

Uma primeira proposta foi desenvolvida para captura de contexto a partir de discussões em redes sociais. O objetivo foi avaliar a melhoria dos resultados das pesquisas com o uso da técnica de expansão de consultas, comparando-se o contexto gerado a partir de documentos (exemplo, anotações de aula e materiais de referência) e de discussões. A arquitetura proposta, bem como o estudo de caso e seus resultados foram apresentados em (PRATES *et al.*, 2012).

No protótipo, quando se considerou apenas as anotações de aula para aquisição do contexto (chamado de “Contexto 1”), a consulta original (sem contexto, usada como *baseline* de avaliação) teve melhores resultados (50%) do que a expansão geral (considerando todos os documentos) (25%) e igual à expansão por assuntos (organizando os documentos de acordo com os assuntos que tratavam) (50%), mas pior do que considerando qualquer expansão (ou seja, comparando-se os resultados das diferentes expansões e da consulta original) (63%). Considerando-se as mensagens de discussão para as expansões de consulta (ou “Contexto 2”), observou-se que as consultas originais proveram melhores resultados em 18% dos casos, enquanto as expansões obtiveram 45% das melhores avaliações. Considerando-se qualquer expansão, os resultados foram melhores em 73% dos casos. Finalmente, a abordagem híbrida (ou “Contexto 3”), composta por anotações de aula e mensagens de discussão, do contexto apresentou melhor resultado em 14% dos casos para a consulta original e expansão com termos específicos dos assuntos. A expansão com os termos gerais do contexto (sem a segmentação por assuntos) apresentou precisão total de 57%. Considerando-se qualquer expansão, o resultado foi melhor em 71% dos casos.

No questionário, ao serem questionados sobre qual contexto trouxe melhores resultados, 47% dos participantes apontaram o Contexto 01 (mensagens das discussões), 33% escolheram o Contexto 02 (material de aula) e os 20% restantes responderam

Contexto 03 (híbrido). Quando perguntados sobre o que eles acharam da experiência de aula colaborativa usando o *Facebook*, todos os alunos responderam que gostaram muito. Em geral, a impressão dos participantes em relação ao protótipo foi muito boa, porém relataram que a avaliação dos resultados foi uma tarefa exaustiva.

Entretanto, a primeira proposta não se adequa ao requisito de execução de buscas que possam sanar dúvidas durante a dinâmica e, assim, possa contribuir com as discussões. Com isso, os principais requisitos evidenciados foram: (i) permitir o uso do protótipo de busca durante o andamento das discussões no grupo e (ii) usar somente as mensagens para a modelagem do contexto (visto que esta apresentou bons resultados neste estudo de caso). Também se observou a necessidade de modificação da arquitetura e técnicas para adequar-se ao dinamismo da geração do contexto (o contexto cresce e molda-se de acordo com as discussões dos alunos).

#### 4. Enriquecimento de termos de discussões em redes sociais

Em uma segunda proposta, o enfoque de solução usa mensagens de softwares de comunicação, como as discussões em grupos das redes sociais para gerar o contexto e usá-lo para extrair e sugerir palavras-chave que mais se aproximem da expressão de busca informada pelo usuário. Estas palavras são enriquecidas a partir de dados abertos e poderão ser combinadas pelo usuário e adicionadas à consulta original. Pode-se dizer então que a consulta é expandida a partir da construção de vocabulários gerados automaticamente por algoritmos de processamento de texto, com a ajuda do usuário.

O protótipo de aplicação desenvolvido, chamado “*CCS Agent*” (*Collaborative Context Search Agent*), possibilita que a busca por documentos Web seja feita, a qualquer momento, com o auxílio de um agente assistente, solicitado a partir da própria interface de discussão. Pensou-se em uma experiência de busca capaz de entregar ao usuário a resposta para a sua necessidade de informação, paralelamente às atividades colaborativas, de maneira a auxiliar a construção do conhecimento desejado sobre o tema no momento que ele ocorre. Essa necessidade de informação por novos conhecimentos deve ser indicada na forma de uma expressão de busca disposta entre sinais duplos de interrogação, conforme exemplificado no fluxo “C” da Figura 1.

O sistema deve ser capaz de processar a solicitação do usuário e devolver uma mensagem com o *link* para o protótipo, que sugere termos a serem incluídos na consulta e usa o Google como provedor de documentos Web. Assim o aluno evita acessar outro ambiente para realizar a consulta, enquanto aguarda a mensagem de resposta no próprio grupo. A intenção é manter o usuário no grupo enquanto o sistema processa sua requisição.

A Figura 1 ilustra a visão geral da proposta de solução, para facilitar o entendimento da arquitetura. Duas macrofuncionalidades podem ser identificadas na figura: processamento do contexto e processamento da consulta do usuário. O processamento do contexto é alimentado por mensagens do grupo (setas A e B), ou seja, a colaboração entre os usuários irá gerar conteúdo que servirá de insumos para a geração do contexto. Já o processamento da consulta inicia-se pela identificação de uma solicitação de consulta (C), processamento da requisição, que extrai os termos sugeridos para expansão e formula uma mensagem de resposta com o *link* para o protótipo associado à requisição gerada (D). O usuário deve clicar nesse *link* (E) para acessar a

interface do protótipo de busca (que é personalizado para sua necessidade de informação), escolher os termos para expandir a consulta e avaliar os resultados da busca, clicando sobre as estrelas (F).



Figura 1. Visão Geral da Proposta

Cada nova requisição gera três conjuntos de termos, extraídos de maneira distinta e vinculados à expressão de busca informada pelo usuário. A busca é realizada pela combinação dos termos selecionados e gera um conjunto de documentos como resultado. O usuário deve marcar todos os termos que julgar necessário para serem utilizados na expansão de consulta. Até seis (6) (possíveis repetições entre grupos são eliminadas da interface) termos dos grupos “A”, “B” e “C” podem ser combinados. Duas consultas são enviadas ao Google Search API. A primeira contém os termos originais da consulta e a segunda, além dos termos da consulta original, os termos escolhidos para expansão da consulta.

## 5. Estudo de Caso

Um estudo de caso único foi conduzido em um ambiente acadêmico (UNIRIO) e teve a participação de alunos na utilização de um protótipo. Participaram do estudo 18 alunos do primeiro período do curso de Bacharelado em Sistemas de Informação, turma BSI 2011.2, matriculados na disciplina “Fundamentos de Sistemas de Informação”. Os participantes possuem idade entre 18 e 26 anos, média e mediana de 20 anos. Todos os participantes possuem computador e acesso a internet. Apenas um participante não possuía perfil na rede social Facebook, que foi prontamente criado para participar da pesquisa.

Os alunos foram instruídos a participar de uma aula baseada em discussão, realizada no Facebook. No protótipo, cada aluno pôde executar o número de buscas e combinação de termos que julgasse necessário e avaliaram a relevância de todos os resultados retornados. Após a realização da dinâmica, os participantes foram convidados a preencher um questionário com questões qualitativas sobre a relevância dos resultados na utilização do protótipo *CCS Agent*. Do total de participantes, 14 responderam o

questionário, que foram analisados qualitativamente e comparados aos resultados quantitativos extraídos do protótipo.

O estudo de caso foi realizado para observar e relacionar a relevância dos resultados retornados pelas duas modalidades de consulta (original e expandida) e a combinação de termos selecionada nos grupos A e B (termos do contexto com enriquecimento) ou C (termos do contexto sem enriquecimento). A relevância de cada documento retornado no protótipo é aferida explicitamente pelos usuários numa escala de cinco (5) estrelas e relaciona-se implicitamente com o(s) grupo(s) do(s) termo(s) selecionado(s). Em outras palavras, espera-se avaliar a relevância dos resultados da consulta original com os resultados da consulta expandida, observando-se o grupo de termos escolhido para a expansão da consulta (ou suas combinações). Foi obrigatória a escolha de pelo menos um termo de qualquer um dos grupos para expansão.

## 6. Métricas

As avaliações coletadas no protótipo são recursos que representam julgamento humano de relevância para cada documento Web retornado. O resultado da consulta é composto por seis (6) documentos retornados com a consulta original e seis (6) documentos com a consulta expandida. Para fins de consistência, os resultados são mesclados em uma única lista e os resultados redundantes são exibidos apenas uma vez e recebem a mesma avaliação. Caso os seis primeiros resultados de cada expansão (expressões de busca original e expandida) sejam diferentes, serão apresentados ao usuário doze resultados para avaliação (valor máximo). Caso os primeiros seis primeiros resultados coincidam (caso extremo) para ambas as consultas (original e expandida), apenas seis resultados serão exibidos aos usuários (valor mínimo). As métricas, utilizadas para comparar os resultados da busca original e expandida, foram: (i) precisão total dos  $x$  primeiros resultados (CHIGNELL et al., 1999), (ii) comprimento da busca (COOPER, 1968) e (iii) correlação de ranking (SU et al., 1998).

A precisão total dos  $x$  primeiros resultados (*first x full precision*) mede a quantidade total de informações relevantes nos  $x$  primeiros documentos. O valor considerado para  $x$  foi seis (6), ou seja, considerou-se a precisão total dos seis (6) primeiros resultados.

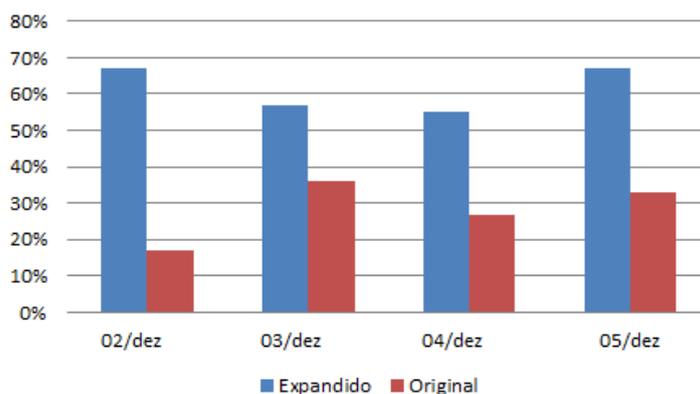
O comprimento da busca (*search length*) mede o número de documentos não relevantes que um usuário deve examinar antes de encontrar uma quantidade  $x$  de documentos relevantes. Utilizou-se o parâmetro  $x=2$  (os dois primeiros documentos consecutivos e relevantes) e considerou-se relevância igual a quatro ou cinco estrelas.

Por fim, correlação de ranking (*rank correlation*) mede a correlação entre a classificação do sistema de busca e o julgamento do usuário (escala de cinco (5) estrelas) para os resultados da busca. Como não existe acesso às notas reais de classificação atribuídas pelo sistema (Google no caso deste trabalho), utilizou-se a posição do documento no conjunto-resposta para presumir sua pontuação. Quanto maior a proximidade do documento ao topo da lista, maior a sua pontuação. Empregou-se o coeficiente de correlação de Pearson para o cálculo da correlação entre a matriz A, que representa as avaliações dos usuários e a matriz B, que representa a ponderação associada a sua posição no conjunto de resultados da busca. A matriz de correlação varia de 4 para o primeiro documento retornado (pontuação máxima atribuída pelo usuário a

um documento) a 0 para o sexto documento retornado (pontuação mínima atribuída pelo usuário a um documento), com passo de 0.8. O resultado da correlação entre as variáveis é apresentado no intervalo  $[-1, +1]$ , onde os resultados próximos de +1 representam uma correlação perfeita positiva, 0 ausência de correlação linear e -1 correlação perfeita negativa entre as variáveis (KENDALL e STUART, 1973).

## 7. Resultados

Foram selecionadas 58 sugestões de termos (nos três grupos) para as 34 avaliações consideradas nos quatro (4) dias de estudo, uma média de 1,7 termos escolhidos por consulta. Em relação à métrica precisão total, a consulta expandida trouxe melhores resultados que a consulta original em todos os dias, conforme observado na Figura 2. As consultas expandidas apresentaram melhores resultados em 67% dos casos nos dias 2 e 5, 57% no dia 3 e 55% no dia 4. As consultas originais foram melhores que as consultas expandidas em 17% (dia 2), 36% (dia 3), 27% (dia 4) e 33% (dia 5). Com isso, a precisão média de pesquisa para a consulta original foi de 29% e de 62% para a consulta expandida.



**Figura 2. Métrica Comprimento da Busca por Dia**

Na segunda métrica, comprimento da busca, verificou-se a superioridade das consultas expandidas em relação às consultas originais. Houve superioridade em 41% dos resultados obtidos com as consultas expandidas e 18% dos resultados obtidos com a consulta original. A última métrica, correlação de ranking, apresentou melhores resultados para a consulta original (62%) em relação à consulta expandida (29%). Apesar das outras duas métricas (precisão total e correlação de ranking) apresentarem melhores resultados para as consultas expandidas com termos do contexto, a métrica precisão total trouxe melhores resultados para as consultas executadas sem expansão, ou seja, a correlação entre as notas dos usuários e a ordem dos documentos retornados pelo motor de busca foi melhor para as consultas originais. Os somatórios para as métricas não totalizam 100%, visto que empates não foram contabilizados.

## 8. Considerações Finais

O objetivo deste artigo foi melhorar a relevância dos resultados das buscas à Web a partir do tratamento de informação contextual obtida em mensagens de grupos de redes sociais online. Verificou-se a viabilidade da proposta de recuperação de informação contextual em um grupo de uma plataforma de rede social mundialmente conhecida

(Facebook). Usou-se um conjunto de bibliotecas *open source* para interagir com Facebook e explorar com sucesso os recursos do grupo.

A abordagem utilizada para a construção coletiva do conhecimento considerou somente discussões e troca de ideias com os colegas. Observou-se grande interação entre os participantes da dinâmica, que aproveitaram para discutir assuntos relacionados ao tema proposto. A correlação de Pearson para a matriz [número de consultas; número de comentários] obteve valor de correlação 0,79, ou seja, existe uma correlação positiva entre o número de consultas e o número de comentários por assunto.

Por fim, este trabalho defendeu a criação do contexto de domínio de maneira automática, ou seja, com o uso de recursos e tecnologias existentes, que independam de esforço humano adicional para a modelagem deste domínio. A modelagem de novas ontologias foi descartada. A engenharia de ontologias tem custo elevado e exige dedicação de especialistas e tempo para a modelagem. Criar modelos de representação e realizar marcações (anotações semânticas) a partir desses modelos é uma tarefa trabalhosa que pode requerer dedicação de especialistas, embora possa apresentar bons resultados se bem realizada.

## Referências

- AMBRÓSIO, A.P., SILVA, L.O., NETO, V.G. (2009). “Automatic Retrieval of Complementary Learning Material for Slide Presentations”. International Conference on Interactive Computer Aided Blended Learning (ICBL 2009).
- BHOGAL, J., MACFARLANE, A., SMITH, P. (2007) “A review of ontology based query expansion”. *Information Processing & Management*, 43(4), 866-886. Elsevier.
- CARPINETO, C., ROMANO, G., (2012) “A Survey of Automatic Query Expansion in Information Retrieval”. *ACM Computing Surveys*, 44(1), 1:1-1:50.
- CHIGNELL, M. H., GWIZDKA, J., BODNER, R. C. (1999 ) “Discriminating meta-search: A framework for evaluation”, *Information Processing and Management: an International Journal*, v. 35, issue 3, pp-337-362.
- COMSCORE. (2011) “It’s a Social World: Top 10 Need-to-Knows About Social Networking and Where It’s Headed”.
- COOPER, W. S. (1968) “Expected search length: A single measure of retrieval effectiveness based on the weak ordering action of retrieval systems”, *Journal of American Society of Information Science*, v. 19, issue 1, pp-30-41
- DOTTA, S., (2011) “Uso de uma Mídia Social como Ambiente Virtual de Aprendizagem”. *Anais do XXII Simpósio Brasileiro de Informática na Educação (SBIE) - XVII WIE*, Aracaju.
- EXPERIAN HITWISE. (2011), “Buscas com uma palavra são maioria”, Disponível em: [http://www.serasaexperian.com.br/release/noticias/2011/noticia\\_00728.htm](http://www.serasaexperian.com.br/release/noticias/2011/noticia_00728.htm).
- FERNEDA, E, (2003) “Recuperação de Informação: Análise sobre a contribuição da Ciência da Computação para a Ciência da Informação”. Tese de doutorado defendida na Escola de Comunicação e Artes da Universidade de São Paulo. USP.

- IBOPE, (2011), “Brasileiros caem na rede social”, *Instituto Brasileiro de Opinião Pública e Estatística*.
- KANG, J. W., KANG, H., KO, M., JEON, H. S., & NAM, J., (2010) “A Term Cluster Query Expansion Model Based on Classification Information in Natural Language Information Retrieval”. In: Proceedings of the International Conference on Artificial Intelligence and Computational Intelligence (pp. 172 – 176)
- KENDALL, M.G., STUART, A., (1973) “The Advanced Theory of Statistics”, Volume 2: Inference and Relationship. Griffin.
- KOBAYASHI, M. E. I., TAKEDA, K., (2000) “Information Retrieval on the Web”, *ACM Computing Surveys*, Vol. 32, No. 2, 144-173.
- MANSUR, A. F. U., CARVALHO, R. A. D., AND BIAZUS, M. C. V., (2011) “Rede de Saberes Coletivos (RESA): Um Ambiente Complexo para Aprendizagem Acadêmica por Meio de Redes Sociais”, *Anais do XXII Simpósio Brasileiro de Informática na Educação (SBIE) - XVII WIE*, Aracaju, 2011. 1284-1293.
- MORA-SOTO, A., (2009) “Collaborative Learning Experiences Using Social Networks. International Conference on Education and New Learning Technologies” (EDULEARN09), 2009.
- PRATES, J. C., SIQUEIRA, S. W. M., (2011a ) “Contextual Query based on Segmentation and Clustering of Selected Documents for Acquiring Web Documents for Supporting Knowledge Management”. Americas Conference on Information Systems (AMCIS), AIS Electronic Library (pp. 1-9).
- PRATES, J. C., SIQUEIRA, S. W. M., (2011b) “Using educational resources to improve the efficiency of Web searches for additional learning material”, In: IEEE International Conference on Advanced Learning Technologies (ICALT) (pp. 563-567).
- PRATES, J. C., FRITZEN, E., SIQUEIRA, S. W. M., ANDRADE, L. C. V., BRAZ, M. H. L. B., (2012) “Contextual Web Searches in Facebook using Learning Materials and Discussion Messages”. In: Computers in Human Behavior (CHB).
- SPINK, A., JANSEN, B.J. (2004) “A study of Web search trends”, *Webology*, 1(2), article 4, Available at: <http://www.webology.ir/2004/v1n2/a4.html>.
- SU, L. T., CHEN, H. L., DONG, X. Y., (1998) “Evaluation of Web-based search engines from an end-user's perspective: A pilot study”. In: Proceedings of the 61st Annual Meeting of the American Society for Information Science, Pittsburgh, PA., pp-348-361.
- WANG, Q., WOO, H. L., QUEK, C. L., YANG, Y., LIU, M., (2011) “Using the Facebook group as a learning management system: An exploratory study”, *British Journal of Educational Technology*, 43(3), no-no. doi:10.1111/j.1467-8535.2011.01195.x.
- YATES, R. B., NETO, B. R., “Modern Information Retrieval”. 1 ed, Addison Wesley, 1999.