

WEB SEMÂNTICA E ONTOLOGIAS APLICADA A TECNOLOGIA DA INFORMAÇÃO NA CONSTRUÇÃO CIVIL

**Maria do Carmo D. Freitas (1), Denise F. Tsunoda (2) Ricardo Mendes Junior(3)
Avanilde Kemczinski (4), Viviane H. Kuntz (5).**

(1) Departamento de Ciência e Gestão da Informação – Universidade Federal do Paraná, Brasil
e-mail: mcf@ufpr.br

(2) Departamento de Ciência e Gestão da Informação – Universidade Federal do Paraná, Brasil
e-mail: dtsunoda@ufpr.br

(3) Departamento de Construção Civil– Universidade Federal do Paraná, Brasil
e-mail: mendesjr@ufpr.br

(4) Departamento de Ciência da Computação – Universidade do Estado de Santa Catarina, Brasil
e-mail: avanilde@joinville.udesc.br

(5) Departamento de Ciência e Gestão da Informação – Universidade Federal do Paraná, Brasil
e-mail: vivianekuntz@ufpr.br

RESUMO

O grande volume de informações disponíveis na Internet dificulta, aos profissionais das diferentes áreas, a recuperação da informação de interesse específico. No setor da construção civil não é diferente. É crescente o número de grupos de pesquisas na área de construção civil que se formam nas Instituições de Ensino Superior espalhadas pelo país, produzindo materiais, a serem divulgados e precisando de novas fontes informacionais para consulta. O objetivo da pesquisa é salientar os conceitos de Web Semântica a qual visa facilitar a obtenção, classificação e organização de informações sobre o tema, estruturando na web os documentos de forma que sejam legíveis tanto para humanos quanto para máquinas. Divide-se a pesquisa em três fases: conceito de Web semântica, ontologias, para indexação e posteriormente a recuperação da informação, tendo como estudo a base tecnológica da área de construção civil, o Centro de Referência e Informação em Habitação. As sugestões de melhoria são no sentido de facilitar a recuperação da informação pelo usuário.

Palavras-chave: web semântica, ontologias, indexação, recuperação da informação, construção civil

ABSTRACT

The large volume of information available in the Internet makes it difficult for professionals from different fields to find information of specific interest. In the construction sector it's not different. The number of research groups in this area inside Colleges and Universities across the country is growing. Those groups produce material that need to be mad public besides new information sources. The goal pursued by the research is emphasize the concepts of Semantic Web, which aims to facilitate the acquisition, classification and organization of information on a theme, structured in such a way that documents will be legible by both humans and machines. The research is divided in three phases: concept of the Semantic Web, onthologies and indexing for information retrival, which ones will work over the technological base of construction (Center for Reference and Information on Housing). The suggestions for improvement are to facilitate the retrieval of information by the user.

Keywords: Web Semantic, anthologies, indexing, retrieval of the information, civil construction

1 INTRODUÇÃO

Os avanços nos meios de comunicações e a velocidade de transmissão das informações por meio dos satélites e redes de fibra ótica estão provocando rápidas mudanças nas organizações. A facilidade de distribuição da informação e o crescimento da quantidade de dados armazenados acarretam um crescente problema de abundância de dados para as áreas da ciência, negócios e governo, dentre outras.

Percebe-se que a velocidade de coleta de informações é maior que a velocidade de processamento ou análise destas. O advento da Internet transformou e triplicou o volume de informações que circulam em muitos ambientes. Surge então a necessidade da aplicação de técnicas e ferramentas que agilizem o processo de extração e recuperação de informações relevantes diante dos grandes volumes de dados armazenados.

Outra tecnologia inserida neste estudo é a web semântica. Berners-Lee (2001) pai da Internet, afirma que a mesma é capaz de estruturar programas de buscas para melhorar o resultado das pesquisas trazendo compreensão, isto é, significado exato das palavras e relações conceituais entre elas. Acrescente-se que por estar baseada no uso de ontologias, fornecem os vocabulários necessários para permitir a definição das relações entre os conceitos dos sistemas.

Este estudo torna-se oportuno, já que os métodos tradicionais de análise de dados, baseados, principalmente, nos procedimentos humanos de tratamento direto com os dados, simplesmente não funcionam para grandes volumes de dados. É crescente o volume de arquivos de pesquisas nacionais sobre a construção civil disponíveis na internet, no entanto torna-se ineficiente e impraticável a sua recuperação para análise.

É comum as buscas de dados coletados em bibliotecas digitais resultar em arquivo não utilizado por pesquisadores ou suas instituições. Além disso, importantes decisões são frequentemente tomadas baseando-se, não na informação armazenada, e sim na intuição do tomador da decisão, simplesmente porque ele não dispõe da ferramenta apropriada para efetuar a extração da informação de maneira segura.

Este artigo apresenta a ferramenta web semântica que passa a ser primordial na recuperação dos dados que hora estavam ocultos e transformá-los em informação estratégica, seja em um ambiente empresarial, governamental ou científico.

2 IMPORTÂNCIA DE TECNOLOGIAS DE INFORMAÇÃO

O século XXI tem trazido ao debate, as transformações do conhecimento e a necessidade de estudos interdisciplinares mais profundos, baseados em especificidades e correlações.

A tecnologia da informação em tempo real possibilita trabalhar colaborativamente, uma vez que permite de qualquer parte do mundo ocorra a troca e o acesso instantâneo de idéias, reduzindo as distâncias.

Segundo Rabelo e Amorin (2008) é integrando recursos das novas tecnologias de informação “no armazenamento da documentação técnica, na otimização das atividades de gerenciamento e no controle da documentação (emissão, revisão, aprovação e distribuição)” que se tem a melhoria do fluxo de trabalho, padronização de arquivos, aumento da segurança desses arquivos, e garantia de ganhos nas atividades de coordenação, compatibilizando projetos, melhorando a rastreabilidade das informações e aumentando a confiabilidade e velocidade da produção.

O trabalho colaborativo entre equipes de engenheiros e projetistas geograficamente distribuídos tem o potencial de reduzir tempo para o desenvolvimento do produto, reduzindo custo e aumentando a qualidade do produto final. Dillenbourg *et al.* (1996). enfatiza que a colaboração envolve o acoplamento mútuo dos participantes, em um esforço coordenado na solução do problema. A formação de equipes de trabalho colaborativo propicia processos de trabalhos eficientes, diminui o número de viagens e aumento de oportunidades de interação pessoal em diferentes campos de trabalho (JOHANSON & TÖRLIND, 2004).

No entanto, perde-se ainda muito tempo. A linguagem de indexação não é padronizada e reconhecida pelos pares. Daí a motivação das novas ferramentas que farão uma maior aproximação entre os homens, as máquinas e os arquivos gerados. A idéia é criar ambientes onde agentes de *software* e usuários possam trabalhar de forma colaborativa fornecendo estruturas e dando significado semântico ao conteúdo das páginas *web*.

2.1 Web Semântica

A inovação nas atividades colaborativas conta com a arquitetura para colaboração móvel inteligente baseada em tecnologias Web Semântica, *Web Services* e Agentes (AZIZ et al., 2004). A Web Semântica é uma extensão da *web* atual, onde a informação possui um significado bem definido, permitindo a cooperação entre computadores e pessoas. Portanto, segundo Hendler *et al.* (2002) *apud* Aziz *et al.*, (2004) a Web Semântica é baseada na idéia de ter dados na *web* já definidos e relacionados de modo que possam ser usados mais eficientemente para: descoberta, automação, integração e reuso através de várias aplicações.

O *Web Service* é uma aplicação modular que possui um conteúdo e uma descrição própria, podendo ser publicado, alocado e chamado por meio da *web*. Uma vez que um *Web Service* é desenvolvido, outras aplicações e outros *Web Services* podem descobrir e chamar os serviços, desconsiderando o sistema operacional e a linguagem de programação. (Kreger, 2001 *apud* Aziz *et al.*, 2004).

Conforme Wooldridge *et al.* (1995) *apud* Aziz *et al.* (2004), agente é um programa com conteúdo próprio capaz de controlar suas próprias tomadas de decisões e agir de acordo com sua percepção do seu ambiente, em busca de um ou mais objetivos. Segundo Aziz *et al.* (2004), quando vários agentes são chamados para trabalharem juntos resulta em um sistema chamado “multi-agentes”.

No ambiente móvel de construção, agentes são essenciais para auxiliar as pessoas em relação aos dados e serviços específicos. Os agentes garantem uma relação semântica aos dados de recursos, aumentando o nível de identificação do conhecimento e prevenindo o usuário de ser esmagado por informações irrelevantes através de filtros baseados em regras e adaptação da informação ao contexto e preferência individuais (AZIZ *et al.*, 2004).

Portanto, tecnologias como Web Semântica, *Web Services* e Agentes possibilitam alavancar a capacidade da computação móvel na construção. Essas tecnologias têm o potencial de causar uma mudança de paradigma no gerenciamento da construção, permitindo o acesso do usuário móvel a uma grande variedade de dados e serviços quando necessário. Entretanto, a concretização dos serviços baseado na Web Semântica para computação móvel na construção devem satisfazer restrições introduzidas pela complexidade da tecnologia, custo, necessidades do usuário e interoperabilidade (AZIZ *et al.*, 2004).

2.2 Ontologias

A ontologia é uma especificação categórica de uma conceituação. O termo ontologia é advindo da filosofia e é definido como parte sistemática da Existência. Para sistemas de Inteligência Artificial o que “existe” pode ser representado.

Quando o conhecimento de domínio é representado em um formalismo declarado, o grupo de objeto selecionado que pode ser representado é chamado de universo do discurso. Este grupo de objetos, e o relacionamento descritivo entre eles, são reflexos de um vocabulário representacional com o qual um programa de conhecimento de base representa o conhecimento.

No contexto de Inteligência Artificial, pode-se descrever a ontologia de um programa pela definição de um grupo de termos representacionais. Deste modo as definições de ontologia associam os nomes de entidades no universo do discurso (por exemplo, classes, relações, funções ou outros objetos) com textos legíveis descrevendo o que os nomes significam, e axiomas formais que limitam a interpretação

e o uso bem formado desses termos. Formalmente, uma ontologia é uma afirmação da lógica teórica (GRUBER 1993).¹

Moreira (2003) esquematiza um sistema de conceito de ontologia conforme a figura 1.

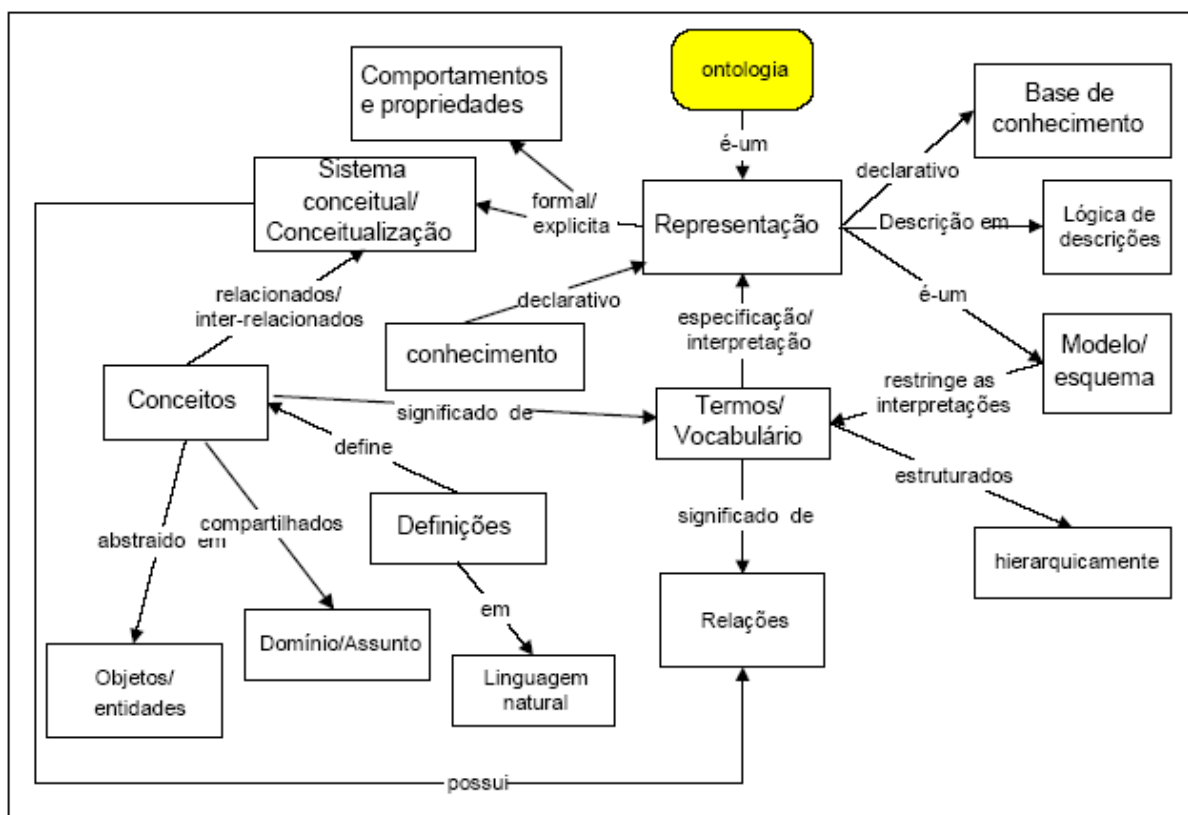


Figura 1 – Sistema de Conceitos para ontologia

Carlan (2006) explica qual a motivação para se definir uma ontologia. A proposta é obter um vocabulário comum para pesquisadores que necessitam compartilhar informações em uma área de domínio. Isso inclui interpretações de definições de conceitos básicos do domínio e das relações entre eles, feitas por máquinas e/ou humanos.

A Ontologia, na web semântica estabelece uma ligação terminológica entre membros de uma comunidade podendo ser estes membros, agentes humanos ou máquinas. No jargão dos pesquisadores em inteligência artificial, uma ontologia é um documento ou arquivo que define formalmente a relação entre termos. (OLIVEIRA, 2008).

2.3 Sistema de Recuperação da Informação

Os sistemas de recuperação de informações proporcionam o armazenamento, recuperação e gerenciamento de informações. “Informação, neste contexto, pode ser composta de textos (incluindo o formato numérico e datas), imagens, áudio, vídeo e outros objetos multimídia” (AMORIN, 2006).

¹ “An ontology is an explicit specification of a conceptualization. The term is borrowed from philosophy, where an Ontology is a systematic account of Existence. For AI systems, what “exists” is that which can be represented. When the knowledge of a domain is represented in a declarative formalism, the set of objects that can be represented is called the universe of discourse. This set of objects, and the describable relationships among them, are reflected in the representational vocabulary with which a knowledge-based program represents knowledge. Thus, in the context of AI, we can describe the ontology of a program by defining a set of representational terms. In such an ontology, definitions associate the names of entities in the universe of discourse (e.g., classes, relations, functions, or other objects) with human-readable text describing what the names mean, and formal axioms that constrain the interpretation and well-formed use of these terms. Formally, an ontology is the statement of a logical theory”. (GRUBER, 1993)

Com o aumento de informações a serem recuperadas, mesmo com a evolução das tecnologias é bastante comum, a apresentação de resultados aquém dos desejados (AMORIN, 2006).

Para Cardoso (2008), “Um sistema de recuperação de informações textuais é um sistema desenvolvido para indexar e recuperar documentos do tipo textual, ou seja, documentos cujas informações estão descritas através da linguagem natural”.

Um Sistema de Recuperação de Informação (SRI) pode ser estruturado conforme a Figura 2 elaborada por Gey (1992) *apud* Cardoso (2008).

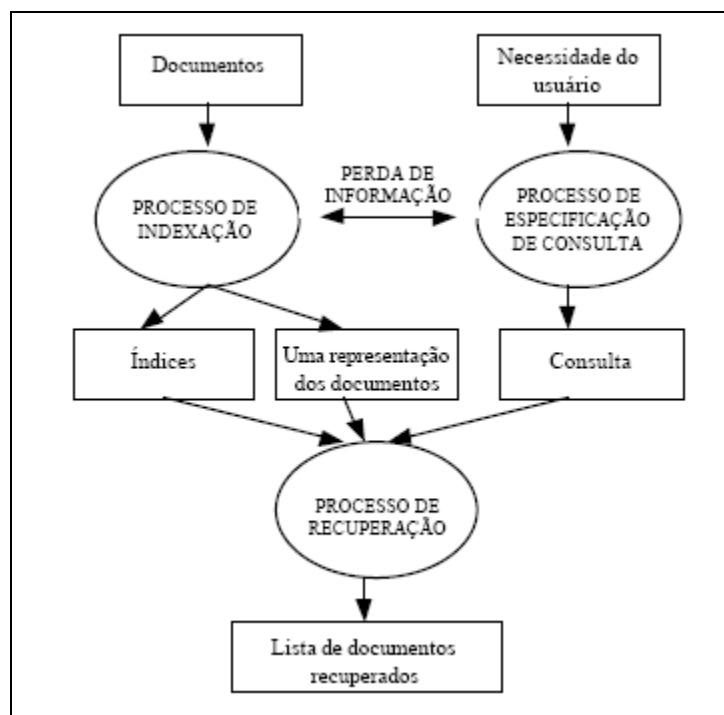


Figura 2 – Componentes de um sistema de recuperação de informação

Conforme visto na figura 2, em dois momentos tem-se a perda de informação, no processo de indexação e no processo de especificação de consulta.

2.3.1 Processo de Indexação

Esse processo é realizado pelo indexador, que tem um papel fundamental, pois é o mediador entre o texto escrito pelo autor e o interesse advindo do usuário.

O cotidiano do indexador é “concentrado no ato da leitura, de forma a viabilizar o acesso à informação contida nos documentos aos usuários dos sistemas de informação anteriormente mencionados”. Entretanto a leitura dos documentos na íntegra pode torna-se inviável com relação a tempo, portanto “é instruído a ater-se a partes do documento, tais como o título e o resumo”. (NEVES; DIAS; PINHEIRO, 2008)

Segundo Ward (1996) *apud* Neve (2008) o indexador precisa saber julgar o que indexar e com que profundidade; ler (inclusive dados não-verbais); analisar e avaliar textos; gerar textos específicos do sistema de informação, catalogar e classificar.

Segundo Bruzinga, Maculan e Lima (2008), “Indexar é o ato de selecionar ou definir termos (palavras ou expressões) que irão descrever o conteúdo de um determinado documento, sempre levando em consideração uma clientela específica”.

A indexação compreende ainda a tarefa de identificação do conteúdo, ou seja do assunto que se trata; e a geração de respectivos textos desse assunto, utilizados como instrumentos da linguagem de indexação, os tesauros, por exemplo (NEVES *et al* , 2008)

Segundo Navarro (1988) *apud* Bruzina (2008), “o ato de indexar define-se em traduzir o conteúdo de um documento em palavras que torne possível sua recuperação, entretanto, observam-se significativas dificuldades na interseção de lingüística e indexação nesse processo”.

2.3.2 Processo de especificação

Esse processo é realizado pelo usuário. Para efetuar uma busca com sucesso é preciso uma estratégia de busca.

Para Lopes (2002) “a estratégia de busca pode ser definida como uma técnica ou conjunto de regras para tornar possível o encontro entre uma pergunta formulada e a informação armazenada em uma base de dados”.

Entretanto verifica-se o impacto dos erros ortográficos nas bases de dados bibliográficas que afetam a produtividade das buscas. Erros ortográficos que ocorrem na linguagem natural, isto é, nos campos de títulos e resumos. (BOURNE, 1977 *apud* LOPES, 2002).

2.4 O valor da informação

Para definição de informação, Mcgee e Prusak (1994) extrapolam o limite dos simples dados coletados, classificando-a como: dados coletados, organizados, ordenados, aos quais são atribuídos significados e contexto.

Davenport (2002) propõe seis características que determinam o valor da informação e Stair (1998) reconhece onze características da boa informação, destacadas do quadro abaixo.

Características	Observação	Davenport	Stair
Exatidão/precisão	Deve ser percebida como valiosa e utilizada com confiança, com ausência de erros na coleta de dados	x	x
Oportunidade/atualidade	Ser enviada quando necessária	x	
Acessibilidade/economia	Obtenção rápida do que é de interesse Ser de produção econômica e oferecer boa relação entre a informação e o custo de sua produção	x	x
Envolvimento/completude	Apresentada como útil	x	x
Aplicabilidade	Utilizada diretamente para equacionar problemas ou apoiar a decisão	x	
Escassez	A raridade de uma informação pode ter grande influência em seu valor.	x	
Flexibilidade	Pode ser utilizada para diferentes finalidades		x
Confiabilidade	Ser obtida de fontes confiáveis e por meio de métodos de coleta igualmente confiáveis;		x
Simplicidade	Ser facilmente compreensível e sem detalhes desnecessários;		x
Verificabilidade	Poder ser verificada a fim de aferir sua correção		x

Quadro 1 – Valor da informação na visão de Davenport e Stair

3 OBJETIVO

Ressaltar a importância de aplicar conceitos de web semântica e ontologias em bases de dados, utilizando como objeto de estudo o InfoHab, sugerindo possíveis melhorias no que diz respeito a recuperação da informação feita pelo usuário.

4 METODOLOGIA

A investigação caracteriza-se quanto aos seus objetivos como uma pesquisa exploratória, envolvendo um levantamento de literatura e a análise de experiências práticas com o problema pesquisado.

Quanto a sua forma de abordagem trata-se de uma pesquisa qualitativa utilizando-se de métodos das ciências humanas que pesquisam, explicitam e analisam fenômenos (visíveis ou ocultos) (HOLANDA, 2006). Caracteriza-se ainda, como uma pesquisa de estudo de caso já que envolve um estudo profundo da base do InfoHAB.

4.1 Centro de Referência e Informação em Habitação – InfoHAB

A Associação Nacional do Ambiente Construído – ANTAC e a Financiadora Nacional de Projetos - FINEP em projeto coordenado pela Universidade Federal de Santa Catarina – UFSC desenvolveram e tornaram públicas uma das maiores bases de artigos científicos da construção civil (www.infohab.org.br). O setor ganhou em inovação e trouxe inúmeros benefícios para quem pesquisa os temas relacionados ao setor.

Funcionando desde 2000, com a iniciativa de entidades associativas, agências de governo e universidades, a Biblioteca Digital InfoHab, é o Centro de Referência e Informação em Habitação, área de Construção Civil.

Além de oferecer aos pesquisadores, profissionais e empresas um acervo digital gratuito composto pela produção técnica e acadêmica brasileira na área de construção civil, permite que o usuário pesquise por publicações dos domínios que envolvem o tema “Ambiente Construído”. Além disso, permite a divulgação e/ou participação em eventos da área (FERREIRA, 2003). A seguir o quadro de estatísticas das publicações divulgadas na base do InfoHAB.

Documento	Quantidade
Artigos de congresso	16791
Artigo de periódico	4115
Outros	6212
Total	27118

Quadro 2 – Estatístico das Publicações de 2007.
Fonte: InfoHAB (2008)

4.2 Breve estudo da base

A aplicação, que está em andamento, de uma referência para indexação da Base de Dados do INFOHAB é feita pelo sistema de classificação ONTOARQ (MALIK, 2006 *apud* RABELO; AMORIN, 2008). Segundo Amorin (2006) esse sistema foi desenvolvido utilizando os conceitos teóricos do projeto Desenvolvimento de Terminologia e Codificação de Materiais e Serviços para a Construção – CDCON –, que é baseado na listagem de termos e relacionamentos para o desenvolvimento da ontologia para Arquitetura, Engenharia e Construção.

É possível pesquisar e inserir novos termos no “sistema ONTOARQ e estes ficarão em “quarentena” para verificação e validação antes de sua total inserção no sistema. Porém para inserir um novo termo, serão solicitadas informações referentes às facetas, significado, relacionamentos (origem e destino), etc”. (RABELO, AMORIN, 2008)

A recuperação no acervo do InfoHAB é feita por meio da busca com a escolha dos três campos; e entre eles, optando entre os operadores booleanos: “e” e “ou”; podendo escolher os campos por: palavra-chave, autor, resumo, sumário e título. Também é possível selecionar o tipo de material, conforme visto na figura a seguir.

Figura 3 – Busca do acervo

4.3 Constatações na base

A escolha de uma amostra da base foi aleatória, mas fundamentada por ser o assunto tratado neste artigo. A busca foi realizada a partir da palavra chave tecnologia da informação. Tratou-se de uma pesquisa simples, com apenas um campo, ou seja, sem a utilização dos operadores booleanos e sem optar pelo tipo de documento.

Obteve-se 118 ocorrências, destas segue as constatações dentro da amostra definida:

Constatações	Observações	Problema foco
Verifica-se a repetição de artigos e teses	1 dissertação – 3 artigos duplicados	Duplicidade de informação
Palavras chaves indexadas diferem entre os duplicados (citado acima) e das utilizadas pelo autor no texto.	Palavras chave que obtém busca em alguns artigos/teses, e outros não mesmo esta palavra fazendo parte do texto do autor.	Exatidão da informação
Nomes com escritas diversas	Abreviação dos nomes após sobrenome e vírgula; ora JR, ora JUNIOR. Nome completo, em outras publicações, o nome está incompleto.	Padronização da informação

Quadro 2 – Pesquisa por amostragem

Outras constatações foram verificadas fora da amostra, por exemplo:

Constatações	Observações	Problema foco
Diferença ao buscar por: 1 - tecnologia da informação, 2 - tecnologia de informação, 3 - tecnologias da informação.	Obteve como resultados respectivamente os seguintes número de documentos disponíveis 118, 12 e 3.	Acessibilidade da informação
Possibilidade de buscar uma referencia pela palavra chave e seu respectivo autor	A caixa de seleção permite apenas um filtro, ou seja, ou a busca é feita por palavra chave, ou por autor.	Flexibilidade de informação

Quadro 3 – Pesquisa aleatória por tema e/ou autor

5 RESULTADOS

O InfoHAB por se tratar de uma base com alta relevância na área de construção, merece estudos e sugestões de melhorias vindas tanto de profissionais capacitados, quanto de usuários do ambiente, buscando sempre a inovação e a evolução. Para as constatações encontradas no InfoHAB, parte-se para sugestões de possíveis melhorias a partir do problema foco observado na base.

5.1 Duplicidade da Informação

A duplicidade dos artigos e/ou teses é evitada dependendo da situação enfrentada. Na indexação manual, é preciso fazer uma busca verificando a existência do material na base. Já para indexação automática, ao inserir o título e os dados pertinentes à publicação, esta já existindo, o sistema deve ser capaz de identificar e notificar a possível duplicidade.

5.2 Exatidão da Informação

A exatidão da informação é com relação às palavras chave indexada serem deferentes das designadas pelo autor, que ocasiona o insucesso na busca caso a referencia venha a ser procurada pela palavra-chave que o usuário lembrava ter visto no material do autor. Isso se deve possivelmente por se tratar de um tesouro que só tem a inserção de novas palavras, quando esta tem sua importancia. Para tanto é interessante a divulgação do tesouro utilizado pelo InfoHAB, tendo assim, uma referência a buscar quando o autor estiver elaborando seu artigo e/ou dissertação/tese, etc.

5.3 Padronização da Informação - Um paralelo com o Lattes

A Plataforma Lattes do CNPq é exemplo de serviço e pesquisa, que gera banco de dados sobre cientistas brasileiros disponível à sociedade. Objetiva a integração de bases de informação e conhecimento sobre a ciência e tecnologia. Para isso criou ontologia comum a esse tipo de informação, em que foi desenvolvida a linguagem LMPL – Linguagem de Marcação da Plataforma Lattes (PACHECO & KERN, 2001 apud OLIVEIRA, 2008).

Assim a relação foi feita com o Lattes, no sentido de um benchmarking, pois para inserir uma produção, sendo ela técnica, bibliográfica ou cultural, ao informar sua autoria, basta escrever o nome completo dos autores que o sistema já normaliza, em forma de citação, segundo a ABNT (Associação Brasileira de Normas Técnicas), e esse autor já fica armazenado para futuras inserções.

Propoe-se o uso de sistema de padronização dos nomes de autores, caso seja constatado a existencia de nomes com escritas diversas, a linguagem utilizada deverá oferecer opções de resolução da situação.

5.4 Acessibilidade de informação

A informação procurada deve ser recuperada independentemente de se utilizar a(s) palavra(s) de interesse no plural, singular ou com algum artigo que a modifique morfológica, mas não semanticamente. É necessária uma reestruturação do banco para a recuperação total das palavras, dando acessibilidade ao usuário.

5.5 Flexibilidade de informação

Ao fixar a caixa de seleção que permite filtro para palavras chave, autor, resumo, título, etc., não possibilita ao usuário a recuperação de uma palavra chave que o interesse de um determinado autor. Tendo a possibilidade de três campos para busca, é interessante que cada campo tenha um filtro independente do outro.

Outra sugestão para tornar a busca mais flexível é a inserção na caixa de seleção da busca pela instituição, quando se trata de teses e dissertações e na busca por eventos e por periódicos, no caso de artigos e resumos.

6 CONSIDERAÇÕES FINAIS

Com a explosão de informações na internet, são muito válidas iniciativas em que o objetivo é reunir toda informação da área, desenvolvendo uma fonte de informação confiável para os pesquisadores em um processo de interação contínua, na qual todos os membros do grupo trabalham para construir juntos suas idéias e desenvolverem determinado trabalho de forma coletiva.

A proposta desta pesquisa é colaborar com essa atitude, ao analisar, sob o ponto de vista do usuário, e sugerir melhorias, visando à acessibilidade, exatidão e completude, enfim, todas as características, vistas na literatura, que proporcionam a recuperação de informações com valor.

REFERÊNCIAS

- AMORIM, S. R. L.; CHERIAF, M.. Sistema de indexação e recuperação de informação em construção baseado em ontologia. In: **III Encontro de Tecnologia de Informação e comunicação na Construção Civil**. Porto Alegre: jun. 2006.
- AZIZ, Z; ANUMBA, C; RUIKAR, D; CARRILO, P; BOUCHLAGHEM, D. Semantic web based services for intelligent mobile construction collaboration. In: **Electronic Journal of Information Technology in Construction (Itcon)**, v. 9, p. 367-379, 2004.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O.. **The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities**. **Scientific American**. 2001. Disponível em: <<http://www.cs.nyu.edu/rgrimm/teaching/reading/semantic-web.pdf>> Acessado em: nov. 2006.
- BRUZINGA, G.S.; MACULAN, B. C. M. dos S.; LIMA, G. Â. B. de O.. **Indexação automática e semântica: estudo da análise do conteúdo de teses e dissertações**. In: Encontro Nacional de Pesquisa em Ciência da Informação. Bahia: out. 2007.
- CARDOSO, O. N. P.. **Recuperação de Informação**. Disponível em: <<http://www.dcc.ufla.br/infocomp/artigos/v2.1/olinda.pdf>>. Acesso em: jan. 2008
- CARLAN, E. **Ontologias e web semântica**. Monografia. Departamento de Ciência da Informação e Documentação - Universidade de Brasília. 2006.
- DAVENPORT, T. H. **Ecologia da Informação**: porque só a tecnologia não basta para o sucesso na era da informação. São Paulo: Futura, 2002.
- DILLENBOURG, P., BAKER, M., BLAYE, A., & O'MALLEY, C. (1996). **The evolution of research on collaborative learning**. In E. Spada & P. Reiman (Eds.), *Learning in Humans & Machines: Towards an Interdisciplinary Learning Science* (pp.189-211). Oxford, UK:Elsevier.
- FERREIRA, S. M.. PITHAN, Denise Nunes. **Estudo de usuários e de usabilidade na biblioteca infohab** : relato de uma experiência. Disponível em: < <http://eprints.rclis.org/archive/00011621/>>. Acesso em: jan. 2008.
- GRUBER, T. R. Toward principles for the Design of Ontologies Used for Knowledge Sharing. In: **Formal Ontology in Conceptual Analysis and Knowledge Representation**. March 1993. Padova. Italy. Available as Technical Report KSL 93-04, Stanford University.
- JOHANSON, M; TÖRLIND, P. Mobility support for distributed collaborative teamwork. In: **Electronic Journal of Information Technology in Construction (Itcon)**, v. 9, p. 355-366, 2004.
- MC GEE, J.; PRUSAK, L.. **Gerenciamento estratégico da informação**. (tradução de Astrid Beatriz de Figueiredo). 12 ed. Rio de Janeiro: Campus, 1994.
- MOREIRA, A.. **Tesauros e ontologias**: estudo de definições presentes na literatura das áreas das ciências da computação e da informação, utilizando-se o método analítico-sintético. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação, 2003.
- NEVES, D. A.; DIAS, E.; PINHEIRO, Â. M.. **Uso de estratégias metacognitivas na leitura do indexador**. Ciência da Informação, Brasília, DF, 35.3, 07 07 2007. Disponível em: <<http://www.ibict.br/cienciadainformacao/viewarticle.php?id=775>>. Acesso em: jan. 2008.
- OLIVEIRA, R. M. V. B.i. **Web semântica**: novo desafio para os profissionais da Informação. Disponível em: <www.sibi.ufrj.br/snbu/snbu2002/oralpdf/124.a.pdf>. Acesso em: fev. 2008.
- RABELO, P. F. R.; AMORIM, S. R. L. de. **ONTOARQ** – ontologia para arquitetura, engenharia e construção. Visualização e gerenciamento na web. Disponível em: <<http://noriegec.cpgec.ufrgs.br/tic2007/artigos/A1131.pdf>>. Acesso em: fev. 2008.
- STAIR, R. **Princípios de Sistema de Informação**: uma abordagem gerencial. Rio de Janeiro: LTC, 1998.