

**UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE INFORMAÇÃO E COMUNICAÇÃO
CURSO DE BIBLIOTECONOMIA**

TATIANA LIMA EVANGELISTA

**ARQUITETURA DA INFORMAÇÃO, RECUPERAÇÃO DA INFORMAÇÃO E
SEARCH ENGINE OPTIMIZATION (SEO): CONEXÕES CONCEITUAIS.**

**GOIÂNIA
2016**

**UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE INFORMAÇÃO E COMUNICAÇÃO
CURSO DE BIBLIOTECONOMIA**

TATIANA LIMA EVANGELISTA

**ARQUITETURA DA INFORMAÇÃO, RECUPERAÇÃO DA INFORMAÇÃO E
SEARCH ENGINE OPTIMIZATION (SEO): CONEXÕES CONCEITUAIS.**

Monografia apresentada à Coordenação do Curso de Biblioteconomia da Universidade Federal de Goiás, como requisito parcial para a obtenção do título de bacharel em Biblioteconomia.

Orientador: Prof. Ms. Arnaldo Alves Ferreira Júnior.

**GOIÂNIA
2016**

FICHA CATALOGRÁFICA

E92a Evangelista, Tatiana Lima

Arquitetura da Informação, Recuperação da Informação e *Search Engine Optimization* (Seo): Conexões Conceituais / Tatiana Lima Evangelista -- Goiânia, 2016.

56 f. il.

Trabalho de conclusão de curso (Graduação) (Curso de Biblioteconomia) Universidade Federal de Goiás, Faculdade de Informação e Comunicação, 2016.

Orientador: Ms. Arnaldo Alves Ferreira Júnior

1. Arquitetura da Informação. 2. Recuperação da Informação. I. Título. II. Ferreira Júnior, Arnaldo Alves (orientador).

..

CDU – 004.2: 347.99 (813.8)

**UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE INFORMAÇÃO E COMUNICAÇÃO
CURSO DE BIBLIOTECONOMIA**

TATIANA LIMA EVANGELISTA

FOLHA DE APROVAÇÃO

**ARQUITETURA DA INFORMAÇÃO, RECUPERAÇÃO DA INFORMAÇÃO E
SEARCH ENGINE OPTIMIZATION (SEO): CONEXÕES CONCEITUAIS.**

Banca Examinadora:

Orientador

Prof. Ms. Arnaldo Alves Ferreira Júnior
Faculdade de Informação e Comunicação - Universidade Federal de Goiás
Curso de Gestão da Informação

Membro Convidado

Profa. Esp. Werônica Yamacyra Cordeiro
Faculdade de Informação e Comunicação - Universidade Federal de Goiás
Curso de Biblioteconomia

**GOIÂNIA
2016**

Dedico aos meus queridos Pais e a minha
Avó Anna, por ser meu melhor exemplo!

AGRADECIMENTO

A Deus, por mais essa oportunidade!

Aos meus irmãos, pelo companheirismo e amor!

Aos meus amigos, por sempre estarem presentes. Em especial, à Denilza, Cinthia, Cleia, Ana Paula, Itana, Eduardo, Daniel e Raphael, que “sofreram” comigo as angústias da construção dessa pesquisa.

Aos professores de Biblioteconomia, pois possibilitaram a minha chegada até aqui. Em especial, à Professora Luciana, cuja paixão e dedicação pelo curso inspirou-me, à Professora Eliany, uma vez que seu conhecimento possibilitou-me a inteligibilidade, e sua boa vontade para com o bem público, sobretudo por ajudar-me tanto em momentos difíceis. Suas orientações despertaram, em mim, grande admiração, bem como mais força e garra para seguir “escrevendo”.

Ao meu Orientar Arnaldo, pelas enriquecedoras orientações, incontáveis conselhos, pela constante paciência, apoio e disponibilidade para comigo.

A toda e qualquer outra figura social que, direta ou indiretamente, ainda que não mencionado, tenha feito parte do meu processo de construção discursiva, de conhecimento, de compreensão...

Muitíssimo obrigada a todos!!!

“Quem nunca errou nunca experimentou nada novo”

Albert Einstein

RESUMO

A presente pesquisa procurou conhecer como funciona a modelagem de rede entre Arquitetura da Informação, *Search Engine Optimization* e Recuperação da Informação para que o objetivo fosse alcançado foi realizado um referencial teórico para ter melhor embasamento sobre as três áreas estudadas. Através do grafo percebeu-se que o nó mais conectado na rede é do termo Informação, no qual os três campos estão fortemente interligados, onde cada área trabalha com a informação em sentidos diferentes, mas todas voltadas para a facilidade de uso da web pelo usuário.

Palavras-chave: Arquitetura da Informação. Search Engine Optimization. Recuperação da Informação.

ABSTRACT

This research sought to know how the network modeling of information architecture, Search Engine Optimization and Information Recovery for the goal to be reached was a theoretical framework to have better basis on the three areas studied. Through the graph it was realized that the most connected node in the network is the term information in which the three fields are closely interconnected, where each area works with the information in different ways, but all geared for ease of user web use.

Keywords: Information Architecture. Search Engine Optimization. Information Recovery.

LISTA DE FIGURAS

Figura 1 -	Representação do resultado de uma expressão booleana conjuntiva (AND).....	33
Figura 2 -	Resultado de uma busca booleana disjuntiva (OR).....	34
Figura 3 -	Representação de uma busca booleana NOT.....	34
Figura 4 -	Rede conceitual das três áreas de pesquisa.....	37
Figura 5 -	Modelo estatístico de intervalo de grau na rede.....	40
Figura 6 -	Grafo.....	41

LISTA DE QUADROS

Quadro 1 - Métricas aplicáveis aos estudos de redes utilizadas neste trabalho.....	14
Quadro 2 - Tipos de navegação.....	19
Quadro 3 - Tipos de rótulos.....	21
Quadro 4 - Indicadores e/ou Valores de Referência das métricas alcançados.....	42

SUMÁRIO

1	INTRODUÇÃO.....	13
2	ARQUITETURA DA INFORMAÇÃO (AI).....	16
2.1	SISTEMA DE ORGANIZAÇÃO.....	17
2.2	Sistema de Navegação.....	19
2.3	Sistema de Rotulagem.....	20
2.4	Sistema de Busca.....	22
3	Motores de Busca.....	23
3.1	Search Engine Optimization (SEO).....	24
3.1.1	Título.....	26
3.1.2	Tags.....	26
3.1.3	Palavras Chave.....	27
3.1.4	Arquivo Robots.Txt.....	28
3.1.5	Linguagem de Marcação de Texto (html).....	28
3.1.6	Conteúdo da Página.....	28
3.1.7	Nomenclatura de Páginas e Diretórios.....	29
3.1.8	A Classificação dos Mecanismos de Busca para os Sites.....	29
3.1.9	Popularidade e Correlação de Links.....	30
4	RECUPERAÇÃO DA INFORMAÇÃO (RI).....	31
4.1	Modelo Booleano.....	33
4.2	Modelo Vetorial.....	35
4.3	Modelo Probabilístico.....	35
5	RESULTADOS ALCANÇADOS.....	37
6	CONSIDERAÇÕES FINAIS.....	43
	REFERÊNCIAS.....	44
	ANEXO - REFERÊNCIAS DOS TEXTOS UTILIZADOS NA REDE	48

1 INTRODUÇÃO

Em 1980, foi desenvolvido o World Wide Web (WWW), popularmente conhecido como web, pelo pesquisador Tim Berners-Lee. A web é conhecida como um conjunto de hipertextos que têm alcance mundial. Esses hipertextos funcionam como troca de informação. Neste sentido, para que haja disseminação, é necessário acesso à internet. A internet é uma rede interligada de computadores a partir de pesquisas de professores universitários, estudantes e empresas de tecnologia. Mas, sua divulgação começou após a Segunda Guerra Mundial, para uso exclusivo dos Estados Unidos, com objetivo de armazenamento de informações. Com o avanço da tecnologia, a internet melhorou e, hodiernamente, é utilizada para compartilhar conteúdo de forma rápida, facilitando o acesso à informação (ALMEIDA, 2007; GARCIA, 2011).

Sendo assim, a partir dessa perspectiva, a web pode ser vista como uma biblioteca, que precisa estar organizada para que o usuário encontre o que veio buscar na rede. Dessa forma, surgiu a Arquitetura da Informação, *Search Engine Optimization* e a Recuperação da Informação, áreas do conhecimento distintas, mas que trazem a preocupação de facilitar o acesso do usuário à informação desejada.

A Arquitetura da Informação trabalha com a estruturação de ambientes informacionais para melhorar a usabilidade e recuperação da informação. O *Search Engine Optimization* (SEO) são técnicas de ranqueamento¹ de páginas online utilizadas pelos mecanismos de busca para indexar os sites e colocá-los em suas bases de dados. Por sua vez, a Recuperação da Informação estuda a relevância dos materiais disponibilizados na rede de acordo com termos buscados pelos clientes.

Redes complexas são utilizadas como ferramentas para descrição dos mais variados tipos de sistemas. Os tipos de redes reais que foram descritos mais recentemente com o auxílio de redes complexas são as redes sociais e as redes biológicas. Até pouco tempo essas redes eram difíceis de modelar pela ausência de um modelo adequado e pela insuficiência de capacidade computacional no processamento de dados necessários (LOPES, 2011, p. 7).

¹ Determinação de uma posição hierárquica, pela contagem de pontos etc. (Dicionário Michaelis). <http://michaelis.uol.com.br/busca?r=0&f=0&t=0&palavra=ranqueamento>.

As ligações em rede seguem regras preestabelecidas, entretanto, o crescimento da rede acontece naturalmente e, justamente por isso, é imprevisível. Portanto, compreendendo-se as regras de ligações entre os elementos da rede é possível analisar o comportamento global da rede a partir de suas características físicas, como tamanho, diâmetro, grau de conexão médio, densidade, etc. (LOPES, 2011, p. 8).

Os principais modelos de rede são: Redes Aleatórias: o mais antigo possui ligações baseadas em probabilidades de conexão; Redes de Mundo, usadas na modelagem de problemas em redes sociais; Redes Livre de Escala, utilizadas para estudo da dinâmica de crescimento de redes (LOPES, 2011, p. 9).

Quadro 1- Métricas aplicáveis aos estudos de redes utilizadas neste trabalho.

Dimensões	Medidas	Definições
Rede	Distribuição de Grau (Conectividade)	Média do número de conexões de todos os nós da rede.
	Componentes Conectados	Determina o número de componentes conectados na rede.
Arestas	Comprimento Médio dos Caminhos	<p>Centralidade de intermediação: mede a frequência com que um nó aparece nos caminhos mais curtos entre nós da rede.</p> <p>Centralidade de proximidade: Distância média de um determinado nó inicial para todos os demais nós da rede.</p> <p>Excentricidade: Distância de um determinado nó inicial até o nó mais distante dele na rede.</p>

Fonte: Elaborado pela autora.

Diante do exposto, este trabalho será dividido em quatro seções, sendo a primeira Arquitetura da informação com ênfase na web, seguirá o modelo de Morville e Rosenfeld com os seguintes sistemas: sistema de organização, sistema de busca, sistema de rotulagem e sistema de navegação. A segunda seção será de Search Engine Optimization definirá o que é esse tipo de otimização e abordará algumas técnicas utilizadas para ranqueamento de website

A terceira seção falará de Recuperação da Informação também tem o foco na web e a recuperabilidade da informação, visto que existem diversos mecanismos de

recuperação da informação. Também serão discutidos os modelos de se recuperar informações mais utilizadas que são: Modelo booleano, Modelo vetorial e Modelo probabilístico. A última seção mostrará o resultado desse trabalho com o uso de Rede Conceitual.

Considerado o contexto exposto acima, a questão problema que norteará o trabalho foca especificamente, em como a Arquitetura da Informação, o Search Engine Optimization e Recuperação da Informação trabalham na pretensão de otimizar a busca de informação na Web. Sendo assim, essa pesquisa tem como problema procurar saber qual o tipo de relação existente entre essas áreas do conhecimento?

Os objetivos a serem cumpridos para a realização da presente pesquisa estão assim organizados, objetivo geral: "Modelar a conectividade da Arquitetura da Informação, Search Engine Optimization e Recuperação da Informação com o uso de redes conceituais". Objetivos Específicos:

- Levantar textos científicos sobre Arquitetura da Informação, Search Engine Optimization e Recuperação da Informação através de busca bibliográfica orientada;
- Modelar a distribuição temática das áreas de conhecimento a partir da formalização de redes conceituais;
- Identificar, com o uso de grafos, a conectividade existente entre as três áreas de conhecimento citadas na pesquisa;

Os Procedimentos Metodológicos necessários à operacionalização da pesquisa estão divididos e encadeados nas seguintes etapas:

Procedimento 1 – Realização da fundamentação teórica sobre Arquitetura da Informação, *Search Engine Optimization* e Recuperação da informação.

Procedimento 2 – Fazer o tratamento dos resumos de artigos, dissertações e teses no txt para a criação da rede conceitual no gephi.

Procedimento 3 – Analisar os resultados obtidos pelas métricas e comentários.

Os resultados esperados por essa pesquisa é procurar entender se existe algum vínculo entre as três áreas estudada, se tiver, buscar entender como funciona essa ligação e como elas estão conectadas.

2 ARQUITETURA DA INFORMAÇÃO (AI)

Arquitetura da informação (AI) foi um termo criado por Richard Saul Wurman que é arquiteto por formação para designar a tarefa de tornar o complexo em algo de fácil compreensão, facilitando assim a recuperação da informação, também foi o homem que criou a própria profissão ao dizer que era arquiteto da informação, no seu livro “Ansiedade de Informação” (WURMAN, 1996).

Wurman traz demonstrações de como tornar a informação mais compreensível a ponto de ser útil para as pessoas, o autor cita vários exemplos de informações que tiveram mais clareza depois que o mesmo arrumou, por exemplo, organizou uma determinada lista telefônica por cidade em vez de ser por ordem alfabética, dessa forma Wurman afirma que fica mais fácil das pessoas entenderem quais estabelecimentos que tem em uma determinada região e quais não existem (WURMAN, 1996).

Dessa forma, várias áreas do conhecimento adaptaram a arquitetura da informação para si, Morville e Rosenfeld (2006) transformaram a arquitetura da informação para a web e publicaram o livro *Information Architecture for the world wide*, onde definem a arquitetura da informação como: a arte e ciência de organizar e rotular websites, intranets, comunidades on-line e software, para suportar a usabilidade. Nesse contexto, a preocupação é torna a web um lugar onde a compreensão seja mais fácil para os usuários e que, além disso, este cliente não precise gastar tempo para aprender acessar um site, mas que o indivíduo possa compreender de uma forma mais intuitiva (MORVILLE; ROSENFELD, 2006).

Inakufu (2013, p.55) reafirma o que Morville e Rosenfeld disseram ao expressar que “Encontrabilidade da informação se diz sobre a recuperação de informações relevantes com mais rapidez, de modo que a navegação seja mais intuitiva”.

Com o uso da AI é possível reduzir o tempo de encontrar informação, o tempo de não encontrar informação, custos com construção e manutenção de websites, despesas com treinamentos de funcionários e ainda prover um aumento da valorização da marca (REIS, 2004 *apud* SILVA *et al.*, 2011, p. 12).

Percebe-se que através desses autores a arquitetura da informação na web está centrada em diminuir a dificuldade de localizar a informação desejada, fazer com que o website ou software se torne algo mais intuitivo no momento da busca

por informação para que o usuário tenha mais facilidade na procura, então nota-se que a AI está centrada no usuário, em facilitar o acesso à informação, ou seja, tornar o que é considerado um labiríntico na web em algo mais simples de ser absorvido.

Para que os websites se tornem mais acessível aos seus usuários Morville e Rosenfeld (2006) constataram a necessidade de estruturar e organizar a informação para ficar mais fácil a recuperação da mesma, então notaram que para organizar a informação na web era preciso separar um website por partes, cada uma dessas foi chamada de sistema, sendo: sistema de organização, sistema de navegação, sistema de rotulagem e sistema de busca.

2.1 Sistema de Organização

Para Morville e Rosenfeld (2006) o sistema de organização, possui o compromisso de organizar o conteúdo dentro de um website para torná-lo mais compreensível, afim de que o usuário consiga atingir o objetivo de encontrar a informação desejada dentro do site. Silva *et al.* (2011) resumem tudo isso em: o resultado final do sistema de organização é recuperar a informação para o seu cliente de forma ágil.

Brasileiro e Freire (2012) afirmam que a organização é feita utilizando dois contextos de organização: esquemas de organização e estrutura de organização. Esquemas de organização são definidos como os traços gerais do elemento, o qual vai influenciar no agrupamento lógico do conteúdo do site podendo ser por esquemas alfabético, cronológico, geográfico, assunto, tarefa-orientada, público-alvo e metáfora.

Já a estrutura da organização está relacionada com a forma que os usuários irão navegar pelo site, sendo essa estruturação feita por: modelos hierárquicos de organização, modelos de hipertextos ou modelo de banco de dados relacionais.

Mas, para organizar essa informação da forma que Morville e Rosenfeld (2006) propõem Ribeiro e Vidotti (2009) afirmam que vários aspectos entram em jogo, por exemplo, a classificação que de acordo com Anjos (2008) é a forma pela qual nosso cérebro difere um objeto do outro, dessa forma o ser humano faz isso de maneira inconsciente e consciente já que separamos tudo por suas características como: as cores, animais, sapatos, ideias e até mesmo classificamos o que é para se comer e o não é alimento e etc. Só que esse processo de classificação acontece de

forma subjetiva com cada classificador, pois cada sujeito tem uma forma única de ver o mundo e termina classificando da maneira que melhor combina para o próprio, o que torna complicada a classificação, mas esse procedimento precisa ser feito pensando no usuário que vai utilizar o sistema. Outra preocupação no momento de classificar é a ambiguidade dos termos que para Reis (2007) acontece por causa da classificação ou da categorização realizada, e também por causa do contexto de cada conteúdo que influencia em muitas esferas, por exemplo, a famosa manga que em uma circunstância pode ser manga um tipo de fruta e em outra pode ser a parte de uma camiseta. Então estruturar e organizar a informação não são tarefas fáceis, requer que o classificador pense no momento em como o usuário procuraria por uma determinada informação.

Alguns autores como Ribeiro e Vidotti (2009), Reis (2007) e Silva *et al.*(2011) afirmam que a heterogeneidade dos sites e dos seus conteúdos dificulta a organização, esses autores consideram heterogeneidade como a diversidade dos tipos e formatos de conteúdos dentro de um website, o que termina complicando a padronização e criação de uma política única.

2.2 Sistema de Navegação

O sistema de navegação serve como guia para nortear o usuário como percorrer no ambiente digital, coordena a forma que será feito o movimento dentro do site, para Silva (2014) o sistema de navegação foi elaborado para impedir a confusão e insatisfação no momento de navegar, fazendo com que o usuário não corra o perigo de ser perder no meio do caminho. Para isso Reis (2007) diz que o sistema de navegação deve responder as seguintes perguntas para os seus usuários: Onde estive? Onde estou? Aonde posso ir? , uma vez que se o usuário consegue responder essas perguntas ele tem total autonomia para navegar nesse site e encontrar o que deseja.

Percebe-se que através das respostas dos usuários as perguntas acima têm como saber como vai o funcionamento desse site. Silva *et al.* (2011) informam que existem dois tipos de sistema de navegação:

O sistema de navegação de websites pode ser dividido em duas categorias: sistema de navegação embutido, incluindo-se componentes como: logotipo, menu de navegação global, menu de navegação local, componentes de navegação contextual, bread crumb e cross content; e sistema de navegação remoto onde se inclui componentes suplementares como: mapas do site, índices e guias (SILVA *et al.*, 2011, p.16).

Aquino e Oliveira (2012) também citam os componentes de navegação apresentado acima por Silva *et al.* (2011) e salientam que esses elementos trazem compreensão ao usuário, pois oferece flexibilidade e contexto para se entender onde está e para onde pode seguir, esses sistemas são subdivididos das seguintes formas (Quadro 1).

Quadro 2 – Tipos de navegação

Sistema de navegação embutidos	Sistema de navegação complementar	Sistema de navegação avançado
Sistemas de navegação global	Mapa do site	Personalização e customização
Sistemas de navegação local	Índices	Visualização
Sistemas de navegação contextual	Guias	Navegação social
	Wirzards	

Fonte: Elaborado pela autora.

Essa divisão serve para uma melhor visão sobre a abrangência do sistema de navegação. Por não existir uma forma universal de navegação na web, cada website tem a sua missão de forma particular que é encontrar a maneira ideal, na qual o cliente vai encontrar nesse ambiente o que necessita.

E, por fim, o sistema de navegação pode ajudar ou atrapalhar a usabilidade no site, pois esse procedimento orienta o caminho que um usuário segue nesse site, caso o site não esteja bem direcionado causará angustia nas pessoas que navegam por ele, podendo a chegar não acessar mais aquele endereço e terminar falando mal desse site para outros usuários em potencial (REIS, 2007; RIBEIRO; VIDOTTI, 2009).

2.3 Sistema de Rotulagem

A palavra rótulo no dicionário Priberam² online significa “Impresso que identifica o conteúdo, as características ou a composição de um produto ou outras informações complementares”, rotulagem para o mesmo dicionário é: “Ato ou efeito de rotular”, rótulos para Reis (2007) é um símbolo linguístico que serve para comunicar um conceito sem ocupar bastante espaço na página e não causar dificuldade no usuário para compreendê-lo, similarmente Silva (2014) diz que rotulagem é uma forma de representação, no qual tem o objetivo, transmitir o

² <http://www.priberam.pt/DLPO/rotulo>

conteúdo de forma competente e ocupar o menor espaço possível dentro nas páginas.

O sistema de rotulagem descreve e apresenta itens no website, normalmente eles são usados nos títulos das páginas, na indexação das páginas para o sistema de busca, nos meta-dados das páginas e nos sistemas de navegação. Segundo Reis (2007) os rótulos mais encontrados nas páginas são os ícones (pequenas imagens que representa um conceito). A rotulagem apresenta as seguintes dificuldades: Conseguir falar na linguagem do usuário, ambiguidade e o feedback.

Conseguir falar na linguagem do usuário é um dos maiores problemas na hora de fazer rotulagem, uma vez que para compreender a linguagem do usuário é necessário certo conhecimento de como funciona o dia-a-dia desse público-alvo (INAFUKO, 2013; REIS, 2007).

Agora a ambiguidade no sistema de rotulagem ocorre da seguinte maneira, o cliente pode esbarrar com duas ou mais categorias na qual um mesmo item pode estar englobado, por exemplo: em uma categoria de livros, o mesmo livro pode está nas subcategorias de romance, novela e romance policial. Do mesmo jeito acontece com as palavras que são sinônimas, essas podem causar certo tipo de ambiguidade, pois tem palavras com diferenças mínimas e que a maioria dos usuários pode desconhecer (INAFUKO, 2013; REIS, 2007).

É preciso ter feedback para saber se o usuários estão entendendo o conteúdo transmitido e com a falta desse pode ficar complicado para o arquiteto saber se o sistema está corretamente compreendido ou não (INAFUKO, 2013; REIS, 2007).

Quadro 3 – Tipos de rótulos

Rótulos como cabeçalhos	É a representação do conteúdo. Na maioria das vezes os cabeçalhos são ordenados por hierarquia, na qual determina as categorias e subcategorias.
Rótulos como links contextuais	São links com informações, na qual remetem a outras páginas.
Rótulos como termos de indexação	Descreve os tipos de conteúdos no website como: palavras-chave, tags, metadados descritivos, taxonomias,

	vocabulários controlados.
Rótulos dentro do sistema de navegação	Esses rótulos demandam uma consistência maior, pois eles precisam construir uma familiaridade entre site e o usuário, exemplos: Home, apresentação, mapa do site, contatos e fale conosco.
Rótulos iconográficos	Representa o conteúdo de uma informação, o único problema do ícone é que ele possui uma linguagem mais limitada do que o texto.

Fonte: Elaborado pela autora.

Nota-se que rótulos ou ícones ambíguos causam obscuridade de interpretação, desse modo é necessário tomar bastante cuidado, se não, todo o trabalho efetuado na elaboração do sistema de organização e navegação pode ser jogado fora simplesmente porque o arquiteto utilizou rótulos no qual o usuário não consegue compreender (REIS, 2007).

Segundo Rosenfeld e Morville (2006) a rotulagem é considerada a tarefa mais complicada da arquitetura da informação, sendo assim, rótulos feitos de forma que não alcança a linguagem do usuário, termina prejudicando todo o site.

2.4 Sistema de Busca

Sistema de busca contribui para o usuário encontrar facilmente as informações desejadas nas páginas, pois o mesmo fornece a localização da informação, é através dele que o usuário pode escolher se vai fazer uma busca simples ou mais rebuscada com buscadores booleanos e afinamento de pesquisa (MARINHO, 2012). Para Silva *et al.*(2011) não há um modelo universal e padronizado de interface a ser seguido, sem contar que nesse sistema o que o arquiteto da informação deve ter em mente, é o objetivo principal de criar uma interface que possua mecanismos favoráveis à interpretação das informações nela contida, para que assim o usuário possa achar relevante no momento de buscar informação.

3 MOTORES DE BUSCA

O motor de busca é um banco de dados que ajuda o usuário a encontrar informações de acordo com os termos requisitados em sua caixa de busca, nessa conformidade Silveira (2002 *apud* ALMEIDA, 2007, p.37) define “o mecanismo de busca é um banco de dados que ajuda as pessoas a encontrar informações na Internet de acordo com palavras ou termos digitados pelos usuários”, Ferneda (2003, p. 96) expressa de forma mais concisa:

Grande parte dos mecanismos de busca encontrados na web é de uso geral. Esses mecanismos, chamados de *search engines*, *sites de busca* ou *portais*, permitem ao usuário submeter sua expressão de busca e recuperar uma lista (geralmente ordenada) de endereços de páginas (URLs) que presumivelmente são relevantes para a sua necessidade de informação.

Dessa forma, subentende que os mecanismos de busca são utilizados por todos os públicos de usuários, pois é possível encontrar todos os tipos de informação, visto que não se restringe apenas um tipo de informação como as bases de dados.

Kent (2008 *apud* QUITANA, 2012) determina que os mecanismos de busca, é um site manejado por usuários que procura documentos pela internet e que também é considerado como todo sistema que rastreia, armazenam e indexam documentos da web. Dessa forma, nota-se que os mecanismos de busca facilitaram o acesso à informação eletrônica para o usuário, pois segundo Costa (2011) antes dos motores de busca não era possível que usuários comuns encontrassem informação na internet.

Sendo assim, os motores de busca vieram para facilitar e otimizar o acesso a páginas da internet, fazendo com que cliente com pouco grau de instrução tenha acesso a conteúdos, os mecanismos mais populares hoje são: Google (possui o maior número de páginas indexadas), Yahoo, Bing, Altavista, AOL, Lycos (COSTA, 2011).

Segundo Quintana (2012) a indexação pelos mecanismos de busca é feito através de recolhimento de informação como: palavras-chave ou frase que são: indicadores, URL, código, links internos e externos. Sendo assim é de suma importância que os sites invistam em uma arquitetura e conteúdo das páginas.

3.1 *Search Engine Optimization (SEO)*

Conforme o dicionário Priberam (*On-line*) o termo “otimização” tem origem da palavra “ótimo” que por sua vez tem sentido de aprovação ou que possui a mais alta qualidade. Ainda segundo o Priberam (*On-line*) para compreender o significado de otimização é necessário definir o que é otimizar, sendo que no Dicionário de Significados (*On-line*) esse termo é caracterizado por propiciar meios para desenvolver um produto da melhor forma possível. Priberam (*On-line*) diz que otimizar é “dar uma máquina, a uma empresa, a uma ação, etc., o rendimento ótimo”.

Desta forma, pode-se considerar otimização uma ação de otimizar, já que no dicionário o termo é definido como o “Processo pelo qual se determina o valor ótimo de grandeza, pode ser também ato ou efeito de otimizar, acepção (MICHAELIS, *On-line*). Nota-se que otimização é um método pelo qual se trabalha para que um produto chegue ao ponto de excelente ou fique no seu ponto ideal, no dicionário de informática é um “processo utilizado com o objetivo de melhorar o desempenho dos diversos componentes de hardware e/ou software. “Alguns dos componentes são: compiladores, placas de circuitos, conexões de redes, discos, etc.” (MOTA, 2010 p. 121), ao fazer comparações percebe-se que na área de informática pode-se dizer que é aperfeiçoar um site para facilitar a encontrabilidade tanto por parte dos *crawling*, quanto para o usuário.

Existe uma discordância sobre quem deu origem ao termo *Search Engine Optimization*, mas a primeira vez que foi documentado foi em 1997 pela empresa Multimedia Marketing Group (MMG), no qual utilizou a expressão para designar um conjunto técnicas que melhorasse o desempenho de um site fazendo com que esse ficasse com melhor posição em relação aos outros, desde então, a expressão ficou definida para todos os tipos de métodos que visa melhorar o ranqueamento de sites nos motores de busca e também aumentar o número de usuário de um website (ALVAREZ, 2011; RICOTTA, 2008; RIBEIRO, 2013).

O *Search Engine Optimization (SEO)* é traduzido para o português como otimização de sítios para os mecanismos de busca, Almeida (2007, p.59) define SEO da seguinte maneira “É o uso de um conjunto de técnicas de projeto para que um sítio se posicione bem na busca orgânica realizada pelo usuário através de palavras-chave, sem que este sítio precise pagar para isso”, já Davis (2006 *apud*

Ribeiro, 2013, p.9) afirma que SEO “é como a arte, habilidade e ciência de direcionar tráfego da web para websites”.

Dessa forma toma-se por base que o objetivo final da otimização de websites é melhorar o posicionamento do site com relação aos outros no momento da busca feita através dos mecanismos de busca, também visa um maior tráfego de usuários.

Sendo assim, Almeida (2007) afirma que as técnicas de *Search Engine Optimization* foram criadas porque os usuários costumam olhar nos mecanismos de busca os dez primeiros resultados, então o restante das opções na maioria das vezes não é visualizado, desta maneira torna-se interessante ter um site bem estruturado para facilitar a indexação feita pelo *spider*.

E alguns critérios utilizados por esses mecanismos de busca para classificar as páginas são: pela facilidade de leitura, pela facilidade de navegação, pela facilidade de encontrabilidade, pela consistência visual (identidade visual) da página, pela velocidade de carregamento, facilidade de reconhecimento do público-alvo e retorno de tráfego. Vale ressaltar que os parâmetros de avaliação dos motores de busca mudam constantemente (ALMEIDA, 2007).

Já ao voltar o olhar para os usuários, pode dividi-los em duas categorias: iniciantes e avançados. Os iniciantes tendem a utilizar termos mais genéricos ao realizarem suas buscas e os avançados tendem a fazer uma pesquisa utilizando termos mais específicos e utilizam recursos de busca avançada com diferentes combinações lógicas (SILVEIRA, 2002 *apud* ALMEIDA, 2007).

Quanto mais as respostas estiverem localizadas nos lugares que os usuários procuram por elas, mais fácil o design será percebido pelos usuários e maior será o sucesso do projeto (ALMEIDA, 2007).

Dessa forma, percebe-se que a busca feita pelos usuários nos mecanismos de busca é representada como um desejo de recuperar algum tipo de informação o que é bastante importante para quem trabalha com SEO saber por quais palavras os usuários estão buscando, quais representam melhor determinados sites e até mesmo tomar cuidado com palavras em desuso.

Dessa maneira será destacado nessa pesquisa algumas técnicas de SEO que são divididas em duas áreas, sendo a primeira área consiste na otimização que faz dentro do website e a segunda incide em *links* que outros sites apontam para o website em questão (ALMEIDA, 2007; COSTA, 2011; SÁ, 2013).

Técnicas internas são ajustes feitos no site visando facilitar o uso da página pelo usuário, sendo algumas delas: Título, meta tags, meta keyword, arquivo robots, criação do código HTML, conteúdo, nomenclatura da página, imagens, qualidade dos links externos, idade de um domínio.

3.1.1 Título

O título (*tag title*) é a primeira informação exibida pelo site, pois os mecanismos de busca utilizam o título e uma breve definição ou descrição da página, deve possuir entre 60 e 70 caracteres, necessita haver palavras-chave e o nome da página precisa estar em todos os títulos. Além disso, é recomendado que utilize tag com título distinto para cada página, pois ajudará o google na diferença entre as páginas (COSTA, 2011; GOOGLE, ONLINE).

Também vai acompanhado do título e da descrição da página o link para disponibilizar para o usuário no momento da busca. Dessa forma, uma das formas dos motores de pesquisa classificar a página é pelo título, uma vez que o cabeçalho decidirá se o usuário vai acessar ou não a página, no momento da recuperação da informação sobre determinado termo e pode fazer com que esse cliente volte mais vezes ou não depende da exatidão da palavra-chave adicionada ao título (ALMEIDA, 2007; COSTA, 2011).

3.1.2 Tags

Tags meta são informações que não aparecem para o usuário, pois conforme Sarti (2011) são linhas de código HTML ou etiquetas que descrevem o conteúdo de um site, sendo assim essa *tags* serve para os motores de busca, pois apresenta: descrição, palavra-chave e robôs.

A descrição necessita ser feita página por página e precisa ser concisa, recomenda-se que gaste mais tempo fazendo essa meta descrição, pois se o mecanismo de busca encontrar alguma incoerência no seu site pode acontecer de ser desclassificado ou até mesmo considerado spam (ALMEIDA, 2007). Além disso, nessa opção que é inserido as palavras-chave auxiliará o usuário no momento da busca, percebe-se que essa *tag* é de suma importância tanto para a classificação nos motores de busca, quanto para o usuário. Aconselha-se que tenha de 200 a 250 caracteres, também nota que essa opção de descrição ajuda os mecanismos a

qualificar o site, além ajudar o usuário no momento de clicar na página (ALMEIDA, 2007; SARTI, 2011).

3.1.3 Palavras Chave

As palavras-chave devem ser feita página por página do site, assim como a meta *tag* descrição, visto que cada página tem um conteúdo específico, também sugere que tenha de 50 a 80 palavras (ALMEIDA, 2007; COSTA, 2011).

Do mesmo modo recomenda utilizar as palavras mais buscadas pelos usuários e termos em singular e plural, ordene-as de acordo a relevância para a página. Outro aspecto importante é não diferenciar as letras em maiúsculas e minúsculas, já que não existe essa distinção na busca e pode chegar a ser considerado *spam* pelos mecanismos de busca por causa da repetição de palavras (ALMEIDA, 2007).

Já os espaço e vírgula são utilizados para separar as palavras-chave alguns desenvolvedores de sites utilizam espaço apenas para usar menos caracteres outros preferem vírgula por apresentar esteticamente mais organizada (ALMEIDA, 2007).

Através do google nas ferramentas do google Analytics e Webmasters é possível saber quais as palavras-chave mais utilizadas pelos clientes, sendo assim, ao apurar quais os termos que levam a determinada página é plausível que consiga ter um site com palavras-chave mais específicas e atualizada para um determinado público-alvo (SÁ, 2013).

Também é necessário analisar quais termos e expressões representam corretamente os produtos oferecidos pelo site, vale averiguar qual é o grau de relevância dessa palavra para o artefato ofertado (SÁ, 2013).

3.1.4 Arquivo Robots.Txt

O arquivo robots.txt conduz os robôs dos mecanismos de busca ao caminho que o desenvolvedor do site acha correto para indexar as páginas. Esse arquivo também é utilizado para que os robôs não considerem certos diretórios dentro do domínio do site, como: intranets, áreas restritas, administrativas e outras. Esse arquivo pode ser criado em um editor de texto e inserido na raiz do site (ALMEIDA, 2007; SARTI, 2011).

3.1.5 Linguagem de Marcação de Texto (HTML)

A linguagem HTML (Linguagem de marcação de texto) é a interface entre o conteúdo demonstrado para o usuário e os robôs dos mecanismos de busca, por esse motivo precisam ser bem elaborados (ALMEIDA, 2007). Essa linguagem é estabelecida pela IETF-RFC1866** e composta por um tipo específico do SGML utilizado na internet. O HTML é trata-se de documentos simples, possui marcação fixas projetadas, exercido para relatório técnico, são mais utilizados em: escritórios (ALMEIDA, 2002).F

Dessa forma, um site que tem o cuidado de preparar seu conteúdo seguindo o padrão da W3C segundo Oliveira (2011, p.143) “É um consórcio internacional liderado pelo criador do world wide web”, essa estrutura é voltada para melhorar a experiência do usuário na web. O HTML é criado de acordo com o padrão de semântica da internet o que facilita a experiência do cliente (ALMEIDA, 2007).

3.1.6 Conteúdo da Página

O conteúdo da página é o artifício mais importante a ser trabalhado na otimização de websites, já que é através dele que surgem as diferenças na concorrência. Então necessário trabalhar página por página do site, pois cada item de um site tem uma função e uma conexão de sentido, é importante que o título descreva com exatidão o conteúdo da página, também é necessário ter palavras-chave relacionadas a esse conteúdo (GOOGLE, ONLINE).

As palavras que aparecem no topo da página são caracterizadas como mais importante do que as aparecem mais abaixo pelos motores de busca. Tanto é que esses mecanismos caracterizam as 25 primeiras palavras do primeiro parágrafo como a mais relevante, por isso torna-se importante incluir as principais palavras-chave da página no parágrafo principal, mas de forma coordenada para que faça sentido para o usuário (ALMEIDA, 2007).

Segundo Almeida (2007, p.70-71) os locais considerados mais relevantes para os mecanismos de busca são:

Texto mais acima da página, palavras-chave nos títulos e subtítulos do conteúdo dos textos, texto próximo às outras palavras-chave, palavras e expressões enfatizadas por negrito ou itálico, texto com palavras-chave nos links, utilização de tags H1, H2, H3, H4, H5 e H6 para definição de textos importantes, ordenando de H1 para H6.

Dessa forma, é necessário que o desenvolvedor no momento utilize as palavras-chave corretas para representar o conteúdo até porque essas palavras precisam fazer sentido para o usuário e manter o foco do propósito do site, segundo o Google (Online) recomenda-se que o título represente de forma clara e concisa o conteúdo da página.

Existe uma regra para não exceder em palavras-chave que é chamado de densidade da palavra-chave, onde divide o número de vezes que a palavra-chave aparece no texto da página pelo número total de palavras que aparecem na página, esse cálculo não pode passar de 30%, pois o excesso da expressão pode ser considerado *spam* pelos mecanismos (ALMEIDA, 2007; HISTÓRIA SOBRE OS SITES DE BUSCA, ONLINE).

3.1.7 Nomenclatura de Páginas e Diretórios

A encontrabilidade de informações e a fácil navegação dentro do site são mantidas como um dos melhores resultados pelos mecanismos de busca. Sendo assim, os nomes da página precisam estar relacionada a palavras-chave do conteúdo, se o site tiver uma estrutura grande recomenda-se a divisão do domínio em subdomínios e que utilize a palavra-chave referente a cada categoria, também aconselha-se que utilize os diretórios para mostrar a característica do conteúdo encontrado na URL e evite nomes que não de diretório que não tenha ligação com a página (ALMEIDA, 2007; GOOGLE, ONLINE).

3.1.8 A Classificação dos Mecanismos de Busca para os Sites

A classificação dos sites é feita pela facilidade de leitura, facilidade de navegação, facilidade de encontrabilidade, consistência visual, velocidade de carregamento, direcionamento de público-alvo. Também pela relevância do conteúdo, a densidade textual, proximidade (proximidade do termo com outras páginas), qualidade dos links internos e externos, particularidade do conteúdo e a atualização constante do site (ALMEIDA, 2007; HISTÓRIA SOBRE OS SITES DE BUSCA, ONLINE).

3.1.9 Popularidade e Correlação de Links

A popularidade de um site é calculada pela citação que outro website faz dele através de *links*, também pela incomplexidade que os motores de busca encontram em navegar pelo seu conteúdo e assim classifica-lo. Existem dois tipos de popularidade a de links e a de cliques, conforme mencionado por Miranda (2011), Thurow (2003 *apud* ALMEIDA, 2007):

Popularidade de links é caracterizada pela quantidade de websites com conteúdo de grande relevância e bem classificados que direcionam links para o seu site. Dessa forma, nesse critério de relevância é a quantidade de links que conduzem os visitantes até o site, para os mecanismos de busca quanto mais links direcionado um website teve maior será importância no momento de ranquear (MIRANDA, 2011; ALMEIDA, 2007).

Já a popularidade de clique é quando o usuário clica no seu site em um resultado de busca, a regularidade que o mesmo cliente visita o site, a citação do site nas redes sociais e nas páginas, no entanto se o usuário voltar rapidamente para a página de resultado de busca os mecanismos entenderá que aquele conteúdo não é representativo para aquele perfil de busca e terminará colocando seu site em posições inferiores, e em consequência faz sua popularidade cair (MIRANDA, 2011; SÁ, 2013).

4 RECUPERAÇÃO DA INFORMAÇÃO (RI)

A recuperação da informação é um termo bastante utilizado nas bibliotecas, pois caracteriza a quantidade de material ou documento existente sobre determinado assunto e a quantidade do mesmo que é possível recuperar no momento da busca. Figueredo (2006, p. 17) define essa expressão da seguinte forma: “recuperação da informação consiste em identificar, no conjunto de documentos (corpus), quais informações atendem à necessidade de informação do usuário”, já Barros e Vieira (2010, p.141) descrevem assim: “A recuperação da informação é um processo que visa o atendimento das necessidades informacionais do usuário”.

Dessa forma, percebe-se que a recuperação da informação está ligada diretamente com a organização da informação, pois se o conteúdo não estiver estruturado a encontrabilidade fica ameaçada, neste caso, um recurso inevitável no momento do tratamento da informação conforme Hjørland (2007 *apud* BRANDT, 2009) é a indexação, pois é o modo de representação de um documento através de termos e tem como foco a descrição do conteúdo a ser armazenado, sem contar que é por meio desses termos que ocorre a recuperação de arquivos.

Nota-se que nesse processo de indexação o profissional precisa ser o mais subjetivo possível e conhecer bem a área do conhecimento trabalhada para melhor representar os documentos através dos termos, pois desta maneira o indexador saberá qual é a linguagem controlada de um determinado grupo, sem contar que terá controle terminológico melhor e assim poderá extrair o significado da informação com mais eficiência (INGWERSEN, 1992 *apud* BARROS; VIEIRA, 2010; HJORLAND, 2007 *apud* BRANDT, 2009).

Nas unidades de informação é possível saber a quantidade de documentos existentes sobre determinado assunto e o quanto dele foi recuperado, o qual é chamado de revocação e também faz-se a precisão que é o total de material recuperado que será relevante para o usuário. Na recuperação da informação voltada para web é complicado trabalhar com a revocação, Yang (2005 *apud* GIORDANO; BIOLCHINI, 2012) expõe:

[...] que em sistemas web, as medidas convencionais mostram-se inadequadas, pois sem saber o tamanho do acervo, torna-se impossível

precisar quanto dele foi recuperado ou mesmo quanto é relevante (YANG, 2005 *apud* GIORDANO; BIOLCHINI, 2012, p. 128).

Essa problemática da revocação na web se dá porque qualquer indivíduo que tenha um pouco de domínio da ferramenta pode inserir informação na internet, e na maioria das vezes disponibilizar materiais, porém não faz uma representação de termos plausível o que causa uma dificuldade no momento da recuperação, por mais que os motores de busca hoje façam a indexação automática, os *softwares* robôs chamados *spiders* ou *crawlers* buscam a todo o momento informações atualizadas presentes em sites e a cada conteúdo encontrado o *spider* faz essa indexação através da relevância das palavras no texto e *tags* dos códigos do página, o que pode ser arriscado se o texto não foi representado com os termos corretos (FEITOSA, 2006 *apud* BRANDT, 2009).

Por outro lado existe o usuário final que precisa recuperar a informação, Barros e Viera (2010) descreve sobre a importância de o usuário conhecer as ferramentas de busca e qual tipo de informação deseja resgatar, pois desta forma facilita a encontrabilidade da informação, mas na web temos todo o tipo de usuário desde aquele que não sabe nem digitar corretamente a aquele que tem um entendimento avançado sobre a ferramenta e isso se tornou um desafio o qual vem sendo trabalhado pelos mecanismos de busca, tanto é que o google uma das principais ferramentas de motores de busca utilizado na atualidade, tem a opção de busca simples e avançadas (BRANDT, 2009; PINTO *et al.*, 2004).

A busca simples do google permite o usuário fazer uma pesquisa sem exigir conhecimento da ferramenta, mas é possível afunilar essa consulta e fazer uma procura mais direcionada, pois tem a opção de filtrar o que deseja, por exemplo, buscar somente por: imagens, textos, livros, vídeos, notícias e etc.

Já na busca avançada é possível utilizar os buscadores booleanos, tem a página específica de pesquisa rebuscada do google, aonde é mais fácil de fazer uma pesquisa avançada, já para os programadores tem a alternativa do código-fonte público, onde digita o código e já aparece a informação procurada, tudo isso para atender a necessidade do usuário de recupera informação relevante (SMAAL, 2010; PINTO, 2004).

Então, para uma boa recuperação da informação, faz se necessário que o usuário saiba o que está procurando na web para facilitar a recuperação de informações relevantes, pois a procura pode causar frustração pela demora em

encontrar o que almeja na web fazendo muitos desistirem da busca (BRANDT, 2009; BARROS; VIERA, 2010).

Existem vários modelos de recuperação da informação e esses tratam basicamente da representação arquivo e suas expressões de busca, os mais utilizados são: modelo booleano, modelo vetorial e modelo probabilístico, os quais serão descritos abaixo (FERREIRA, 2015).

4.1. Modelo Booleano

O sistema booleano é fundamentado na teoria dos conjuntos e na álgebra de boole, nessa expressão o conteúdo dos documentos é interpretado por termos de indexação. O modelo booleano é um dos mais utilizados nos sistemas de recuperação da informação, pois é considerado simples, formal e conciso. As buscas são feitas através das expressões booleana ao utilizar os termos conectivos lógicos AND, OR E NOT (SILVA; SANTOS, 2013).

Dessa forma, no momento de recuperar informações o usuário poderá buscar por duas expressões e inserir o AND, sistema subtemde que o cliente deseja informações sobre os dois assuntos, Ferneda (2003, p.22) ressalta de forma concisa ao definir o AND na busca booleana como: “Uma expressão conjuntiva de enunciado t_1 AND t_2 recuperará documentos indexados por ambos os termos (t_1 e t_2)” compreende que o AND funciona nas buscas como sinal de soma (Figura 1).

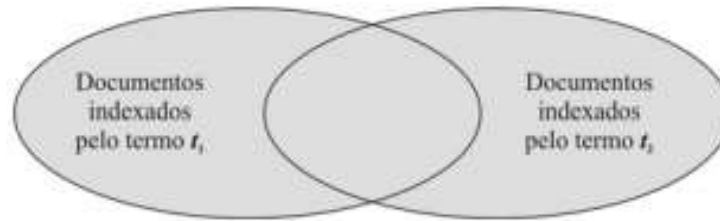
Figura 1 – Representação do resultado de uma expressão booleana conjuntiva (AND)



Fonte: Ferneda (2003)

Já o OR é considerado por Ferneda (2003) e Ferreira (2015) como uma expressão que deve ser utilizada quando o usuário desejar recuperar informações indexadas pelo termo 1 ou pelo termo 2, veja a figura 2.

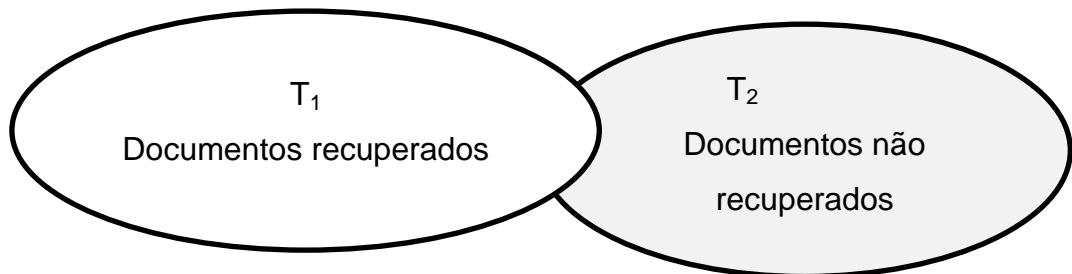
Figura 2 – Resultado de uma busca booleana disjuntiva (OR)



Fonte: Ferneda (2003)

Agora a sentença NOT recupera apenas um termo, se digitar t_1 NOT t_2 , o sistema recuperar apenas documentos do t_1 , figura 3 (FERNEDA, 2003).

Figura 3 – Representação de uma busca booleana NOT



Fonte: Elaborado pela autora.

Desse modo, os termos booleanos podem ser utilizados para fazer uma busca mais avançada, pois esses operadores lógicos influenciam no resultado da recuperação de informação. Autores como Ferneda (2003), Ferreira (2015) e Silva, Santos e Ferneda (2013) dizem que as desvantagens do modelo booleano é que o resultado da busca não vem ordenado pela relevância, pois trabalha com relevância binária, a qual não ordena os documentos pelo seu grau de prioridade, por isso que não é apropriado para os mecanismos de busca da web, onde os documentos são recuperados de forma ordenada é de extrema importância.

4.2 Modelo Vetorial

O modelo vetorial foi inventado por Gerald Salton com o objetivo de utilizar esse modelo em um sistema de recuperação de informação também criado por ele com o nome de SMART (*System for the Manipulation and Retrieval of Text*), segundo Ferneda (2003) esse sistema teve um papel significativo no desenvolvimento da área da recuperação da informação e é um complemento do modelo vetorial apresentado por Salton nos anos 60.

No SMART cada arquivo é representado por um vetor numérico, sendo que cada um desses vetores descreve a importância de um termo na descrição de um documento. Já o modelo vetorial é definido por Kuramoto (2002, p.2) por “[...] comparação parcial entre a representação dos documentos e a da consulta do usuário”.

Sendo assim, pode entender que esse modelo de recuperação da informação busca pesa a similaridade do documento em relação à busca feita pelo usuário, pois esse modelo consiste em balancear cada vetor de termos a um documento e pode se afirmar através de Ferneda (2003, p. 27) “O modelo vetorial propõe um ambiente no qual é possível obter documentos que respondem parcialmente a uma expressão de busca”.

Outro aspecto interessante é que ao contrario do modelo booleano no vetorial é utilizado formas de avaliar a relevância dos termos e esses são representados numericamente demonstrados através de números entre zero e um (FERREIRA, 2015; SILVA; SANTOS, FERNEDA, 2013).

4.3 Modelo Probabilístico

O modelo probabilístico é pautado na matemática das probabilidades, pois procura representar o método de recuperação da informação através do ponto de vista da probabilidade, Silva, Santos e Ferneda (2013, p. 31) afirmam essa definição ao declarar que “O modelo Probabilístico se pauta na teoria matemática das probabilidades”.

Esse modelo trabalha com a seguinte expressão: N (Números de documentos) é dividido em quatro subconjuntos diferentes, no qual **Rel** são os documentos relevantes, o **Rec** documentos recuperados, **RR** documentos relevantes que foram recuperados, e por último documento que não foram recuperados e que não são relevantes (FERNEDA, 2003).

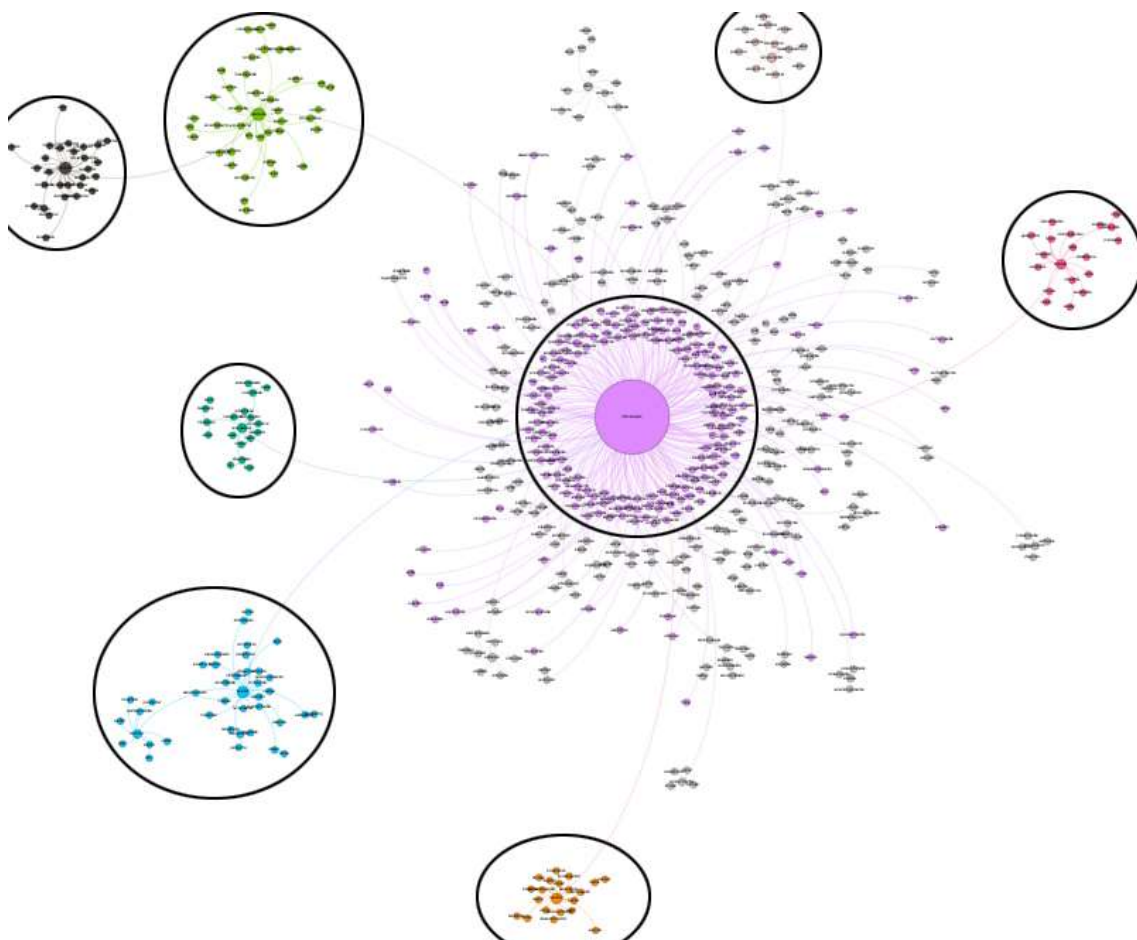
Conforme Ferneda (2003) a expressão ideal ou o conjunto é aquele no qual o resultado da busca seja aquele que contenha o documento relevante para o cliente, tanto é que o autor ainda ressalta que: Se o usuário soubesse como funciona todo o conjunto de documentos seria mais fácil à recuperação da informação de forma apropriada, uma vez que o modelo busca saber a probabilidade de um documento ser relevante, no caso o termo buscado pelo usuário apareçam nele (FERREIRA, 2015).

Ferneda (2003) cita algumas vantagens do modelo probabilístico que são: reconhecer que atribuição de relevância deve ser feita pelo usuário, por isso que o modelo é baseado nas probabilidades matemática, sem contar que pode ser facilmente implementado ao utilizar a estrutura proposta pelo modelo vetorial e assim podem-se utilizar as vantagens desses dois modelos de RI. A desvantagem é baseada na hipótese quais grupos de documentos são relevantes e quais não são relevantes (FERREIRA, 2015).

5 RESULTADOS ALCANÇADOS

Para a criação da rede foram recuperados 35 documentos sobre Arquitetura da Informação, 13 documentos sobre *Search Engine Optimization* e 43 documentos de Recuperação da Informação, um total de 91 artigos, dissertação e tese. A construção da rede foi pelo sistema Gephi, os resultados foram considerando as redes conceituais. Com a aplicação do modelo de modularidade de redes, que diz respeito à identificação de grupos dentro de uma rede, obtemos um total de 82 comunidades, no entanto apenas 8 dessas se mostram realmente expressivas nessa rede, os termos de maior impacto estão relacionados ao tema do estudo, sendo que o termo de maior peso dentro da rede é “Informação” como mostra a imagem abaixo.

Figura 4 – Rede conceitual das três áreas de pesquisa



Fonte: Elaborado pela autora.

Legenda:

- Informação - Rosa meio roxo
- Pesquisa - Verde
- Arquitetura - verde meio azulado
- Resultado - Preto
- Usuário - Laranja
- Estudo - Rosa pink
- Busca - azul
- Recuperação - marrom claro

Como é possível observar o cluster ou comunidade ligada ao termo Informação se encontra no centro da rede e a sua volta os outros termos, todos na ordem: Informação, Pesquisa, Arquitetura, Resultado, Usuário, Estudo, Busca e Recuperação. Todos esses termos serão contextualizados com as áreas de trabalho estudado.

O termo Informação é o mais forte, pois as três áreas do conhecimento estudado estão diretamente ligadas a esse assunto. Ao considerar a definição dada por Silva (2014, p.4) a Arquitetura da Informação da para perceber de forma nítida o porquê desse laço tão forte com a Informação envolvendo essas áreas, “A arte e a ciência de dar forma a serviços e produtos de informação, melhorar a usabilidade e recuperação da informação”.

Nessa definição constatamos que a Arquitetura da Informação tem a preocupação de organizar e estruturar a informação, depois vem o SEO às técnicas estão voltadas para subir o posicionamento de websites com relação aos mecanismos de busca que uma das exigências é a clareza e organização da informação no site para que o mesmo fique em uma boa posição e por último a recuperação da informação que o usuário só consegue recuperá-la porque foi organizada e estruturada para facilitar o acesso.

O termo Pesquisa na AI esta mais voltado para o Sistema de busca e navegação, pois é através desses dois sistemas que o usuário faz a pesquisa dentro do site, segundo Reis (2007) o sistema de busca determina as perguntas que o usuário fará para o sistema, já o sistema de navegação conduz o cliente dentro do site. Quanto a Recuperação da Informação a pesquisa esta envolvida na questão de procurar colocar as palavras chave mais utilizadas nos sistemas de recuperação da

informação, quanto ao SEO é a junção desses dois serviços citados acima para que o site fique em uma melhor posição no momento da pesquisa do usuário.

Arquitetura envolve toda organização e estruturação para receber o cliente, sem a mesma complicaria a Recuperação da Informação, pois a AI envolve a organização da informação e termo que serão utilizados para recuperar o site ou conteúdo contido no mesmo. O resultado é trabalho executado pelas três áreas do conhecimento que permite o resultado chegar até o usuário. Quanto ao usuário a Arquitetura da Informação, *Search Engine Optimization* e Recuperação da Informação são áreas que trabalham com o tratamento das informações nos quesitos estruturar, organizar, ranquear e recuperar essas informações para que o usuário tenha acesso.

O termo Estudo esta voltada para a questão tanto a AI quanto a RI estão sempre sendo estudadas para melhorar e evoluir de acordo com a internet. Busca pode ser associada ao sistema de busca dentro da Arquitetura da Informação, o qual quando o usuário não quer acessar o site pelo menu de navegação, digita diretamente a expressão pela busca, na Recuperação da Informação pode ser retratado no momento que o usuário faz a busca para recuperar informação. A Recuperação só acontece depois que foi feito a Arquitetura da Informação e SEO para facilitar o acesso à informação.

Para contextualizar melhor os dados apresentados, faz-se uso de métrica para calcular o intervalo de grau na rede, obtemos uma visualização dos nós da rede com maior influência dessa forma foi utilizado um intervalo para se demonstrar apenas os 10 nós mais importantes da rede, além daqueles aqui apresentados mais dois nós se sobressaíram, sendo eles, dados e técnica.

Figura 5 – Métrica do intervalo de grau na rede



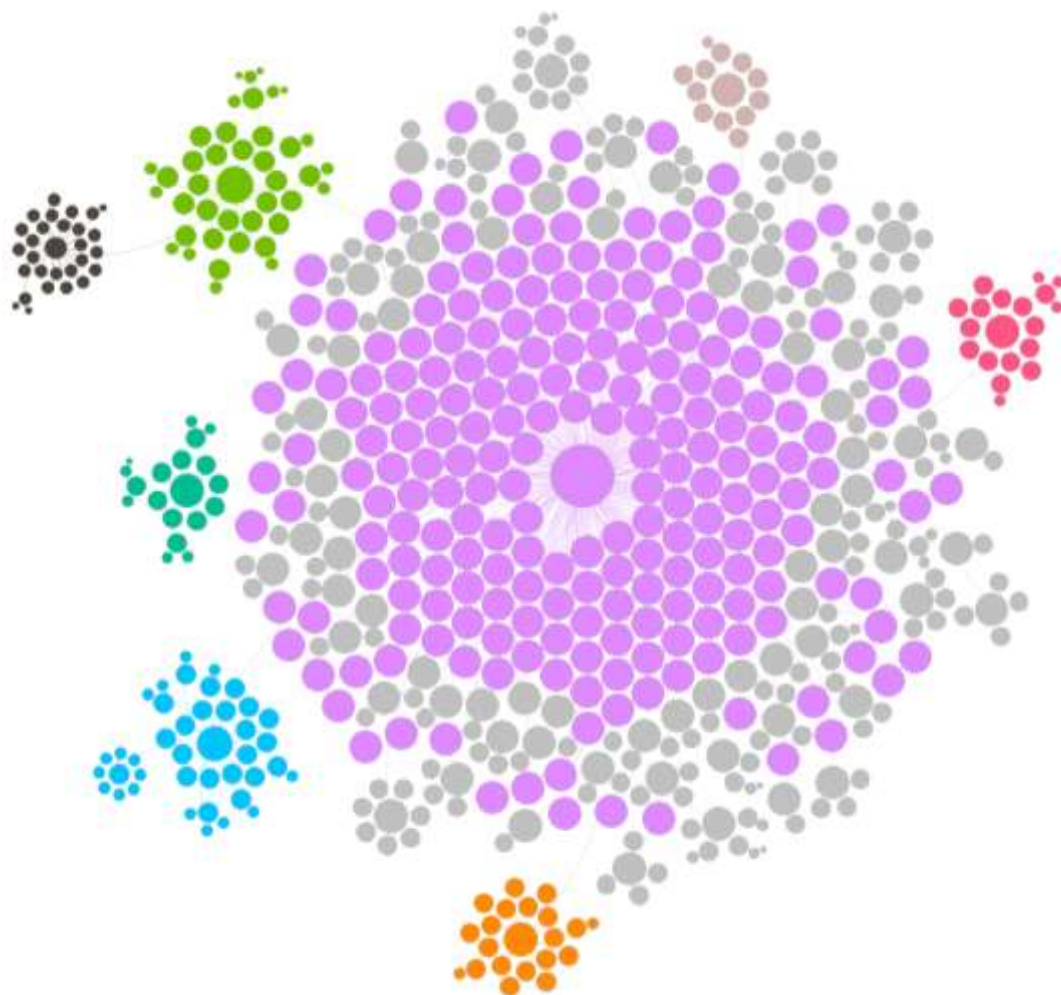
Fonte: Elaborado pela autora

Os termos dados e técnicas podem ser explorados em pesquisas futuras. Ao se aplicar o modelo de cálculo do diâmetro da rede obtemos outros resultados para a rede elaborada, sendo eles centralidade de proximidade e centralidade de intermediação, na rede produzida pelo estudo não se obteve uma visualização de centralidade de intermediação que pudesse indicar fatores abrangentes, isso se dá, talvez pelo fato de se tratar de uma rede de termos e por conta do tamanho da rede.

Já em relação à centralidade de proximidade a rede muda completamente de forma, isso ocorre, pois os termos em si são próximos de acordo com os temas os quais abrangem, dessa forma temos uma rede muito mais uniforme e que mostra que a partir de um termo existe a possibilidade de se chegar a todos os outros, seria algo como uma proximidade de similaridade por conta dos temas que retratam esses termos, isso provavelmente essa uniformidade acontece porque a rede foi criada

através de uma análise bibliográfica de temas que são comuns uns com outros, como podemos ver na imagem abaixo.

Figura 6 – Grafo



Fonte: Elaborado pela autora

Resultados de métricas obtidos pelo Gephi no quadro abaixo.

Quadro 4 – Indicadores e/ou Valores de Referência das métricas alcançados.

Dimensões	Medidas	Indicação/Valores de Referência
Rede	Distribuição de Grau (Conectividade)	1,997
	Componentes Conectados	1
Arestas	Comprimento Médio dos Caminhos	3,312

Fonte: Elaborado pela autora.

Os componentes conectados possuem uma ligação fraca com a rede, no contexto geral isso é ruim. Mas ao mesmo tempo pode ser bom, pois não existe componente menos conectado e sim um componente que possui vínculo com outras redes, fazendo com que esse nó seja a ponte entre as áreas estudada.

6 CONSIDERAÇÕES FINAIS

A Execução dessa pesquisa possibilitou a realização de uma fundamentação teórica sobre Arquitetura da Informação, *Search Engine Optimization* e Recuperação da Informação, o objetivo de Modelar a conectividade das três áreas o uso de redes conceituais foi alcançado.

Através da rede conceitual percebeu-se que todos os termos estão atrelados ao tema de pesquisa e todos os termos convergem em Informação, pois Informação é o nó principal dentro da rede. No desenvolvimento do trabalho notou-se que as três áreas trabalham com a Informação em vertentes diferentes.

O termo SEO não apareceu na rede, através da literatura recuperada sobre *Search Engine Optimization* e o resultado apresentado na rede notou-se que esse termo não precisa necessariamente aparecer, logo que o SEO é um conjunto de técnicas para melhorar o posicionamento do site em uma busca orgânica, sendo assim, tanto a Arquitetura da Informação e Recuperação da Informação cooperam para essas técnicas de SEO, fazendo com que não haja a necessidade do termo Otimização de motores de busca aparecer na rede.

Através desse trabalho notou se a preocupação de estudar essas três áreas em conjunto, uma vez que trabalham com a disponibilização da informação para o usuário em extensões diferentes. Dessa forma, com a junção das três áreas e um estudo mais aprofundado, ambas as áreas podem se beneficiar, principalmente a Otimização de sites e o usuário, pois o SEO ganha um site mais estruturado e capaz de atender as necessidades dos clientes, fazendo com que essa plataforma melhore seu posicionamento em relação a outros sites, uma vez que, um cliente indica outro pela qualidade do serviço ofertado e esse fator também faz a credibilidade do website subir não só em relação aos mecanismos de busca, mas por causa da divulgação para outros usuários.

Nessa pesquisa percebeu-se que para fazer um trabalho mais aprofundado é necessário coletar mais dados e fazer uma análise minuciosa da rede, para assim entender melhor como funciona a conectividade existente entre os termos. Para trabalhos futuros sugere-se que de ênfase não só para a união dos três temas de pesquisa, mas que procure criar uma rede para cada área, pois assim entenderá de forma mais aprofundada os nós mais conectados de cada tema.

REFERÊNCIAS

- ALMEIDA, P. S. B. Otimização de sítios para mecanismos de busca na internet. In:_____. **Otimização de websites para mecanismos de busca na internet: uma contribuição do ergodesign.** 2007. 187f. Mestrado (Dissertação) – Pontifícia Católica do Rio de Janeiro: Rio de Janeiro, 2007. p.58-80.
- ALMEIDA, M. B. Uma introdução ao XML, sua utilização na internet e alguns conceitos complementares. **Ci. Inf.**, Brasília, v. 31, n. 2, p. 5-13, maio/ago. 2002
- ACIOLI, S. Redes sociais e teoria social: revendo os fundamentos do conceito. **Inf.Inf.** Londrina, v.12, n. esp. 2007.
- ALVAREZ, B. Z. **Cartilha de recomendações de SEO para jornalistas online.** Porto Alegre: PUC-RS, 2011. Disponível em:<
http://seonojornalismo.com.br/libraries/down/Cartilha_SEO_no_Jornalismo.pdf>
Acesso em: 01 jun. 2016.
- ANJOS, L. **Sistemas de classificação do conhecimento na filosofia e na biblioteconomia:** uma visão histórico-conceitual crítica com enfoque nos conceitos de classe, de categoria e de faceta. 2008. 277f. Tese (Doutorado) Escola de Comunicações e Artes, Universidade de São Paulo, 2008.
- AQUINO, M. A; OLIVEIRA, H. P.C. Contribuições da arquitetura da informação para website “A Cor da Cultura”. **Inf. & Soc. Est.**, João Pessoa, v.22, n.1, p.129-143, jan./abr.2012.
- BARROS, C. M.; VIEIRA, A. F. G. E a recuperação da informação de objetos multimídia. **Inf & Soc.: Est.**, João Pessoa, v. 20, n. 3, p.135-144, set./dez. 2010.
- BRANDT, M. B. **Etiquetagem e folksonomia:** uma análise sob a óptica dos processos de organização e recuperação da informação na web. 2009. 142 f. Dissertação (Mestrado) – Universidade de Brasília, Departamento de Ciência da Informação e Documentação, Brasília, 2009.
- BRÄSCHER, M. A ambiguidade na recuperação informação. **DataGramZero**, Rio de Janeiro, v.3, n.1, p. 1-22, fev. 2002.
- BRASILEIRO, F. S; FREIRE, G. H. A. O marketing e a arquitetura da informação para web no contexto do processo de mediação da informação. **Biblioline**, João Pessoa, v.8, n. esp., p.161-174, 2012.
- COSTA, F. M. S. P. **Search Engine and Optimization.** 2011. 65f. Dissertação (Mestrado) – Instituto Politécnico de Viseu, Escola Superior de Tecnologia e Gestão de Viseu, 2011.
- FACHIN, G. R. B. Recuperação inteligente da informação e ontologias: um levantamento na área da ciência da informação. **Biblos**, Rio Grande, v. 23, n. 1, p. 259-283, 2009.

FERNEDA, E. **Recuperação de informação**: análise sobre a contribuição da ciência da computação para a ciência da informação. 2003. 137f. Tese (Doutorado) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 2003.

_____. **Introdução aos modelos computacionais de recuperação de informação**. Rio de Janeiro: Ciência Moderna, 2012.

FERREIRA, Juliano Benedito. **Recuperação de informação de música e dados ID3**: possíveis aplicações. 2015. 62 f. Dissertação (mestrado) - Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências, 2015. Disponível em: <<http://hdl.handle.net/11449/132089>> Acesso em: 08 jul. 2016.

FIGUEREIDO, D. E. A. **Recuperação da informação**: uma análise sobre os sistemas de busca da web. 2006. 61 f. Monografia (Graduação) – Universidade de Brasília, Departamento de Ciência da Informação e Documentação, Brasília, 2006.

GARCIA, L. M. L. S. **Investigação e implementação de ferramentas computacionais para otimização de websites com ênfase na descrição de conteúdo**. 2011. 126 f. Dissertação (mestrado). Universidade Estadual Paulista “Julio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas. Bauru, 2011.

GIORDANO, R. B.; BIOLCHINI, J. C. A. Busca e recuperação da informação científica na web: comportamento informacional de profissionais da informação. **InCID: R. Ci. Inf. e Doc.**, Ribeirão Preto, v. 3, n.1, p. 125-145, jan./jun. 2012.

GOOGLE. **Otimização de sites para mecanismos de pesquisa (SEO)**: guia do google para iniciantes. Disponível em:<
http://static.googleusercontent.com/external_content/untrusted_dlcp/www.google.ru/ru/intl/pt-BR/webmasters/docs/guia-otimizacao-para-mecanismos-de-pesquisa-pt-br.pdf> Acesso em: 27 jun. 2016.

GUIMARÃES, Dilva; CABRAL, Paulo. **Significados.com.br**. 2015. Disponível em:<<http://www.significados.com.br/otimizacao/>> Acesso em: 01 jan. 2015.

HISTÓRIA SOBRE SITES DE BUSCA. **Otimização de sites**. Disponível em:<
<https://sites.google.com/site/historiasobreossitesdebusca/otimizacao-de-sites>>
Acesso em 29 jun. 2016.

LIMA, F. **Mas o que significa um domínio adequado**. 2009. Disponível em <http://www.agenciamestre.com/seo/importancia-dominio-rankeamento/>> Acesso em: 29 jun. 2016.

MARINHO, R. B. **Arquitetura da informação para web**: projetando a experiência do usuário no portal de periódicos CAPES. 2012. 142 f. Dissertação (mestrado) – Universidade Federal da Bahia, Instituto de Ciência da Informação, Salvador, 2012.

MARTINHAGO, A. Z. **Desenvolvimento de um sistema de recuperação de informação para a área hospitalar**. 2003. 55 p. Monografia (Graduação) - Universidade Federal de Lavras, Lavras, 2015.

MICHAELIS. Ed. Melhoramentos Ltda, 2009. Disponível em:<<http://michaelis.uol.com.br/moderno/portugues/index.php?typePag=sobre&languageText=portugues-portugues>> Acesso em: 01 fev. 2015.

MIRANDA, R. **Conceito de popularidade do link**. 2011. Disponível em;<<http://www.otimizacao-sites-busca.com/links/popularidade-link.htm>> Acesso em 29 jun. 2016.

MORVILLE, P; ROSENFELD, L. **Information Architecture for the world wide web**. 3 ed. Sebastopol: O'Reilly, 2006.

MOTA, José Carlos. **Dicionário de Computação e Informática**. Rio de Janeiro: Ciência Moderna, 2010.

OLIVEIRA, A. M. Search engine optimization SEO: a contribuição do bibliotecário na otimização de websites para os mecanismos de busca. **Perspectivas em Gestão e Conhecimento**, João Pessoa, v.1, n. esp. p.137- 159, 2011.

PINTO, M. D. S. *et al.* Diversidade da recuperação de informações via internet na pesquisa e ensino. **Rev. ACB**, Florianópolis, v. 9, n.1, p. 143-154, 2004.
PRIBERAM. 2015. Disponível em:<<http://www.priberam.pt/dlpo/otimiza%C3%A7%C3%A3o>> Acesso em: 31 jan. 2015.

QUINTANA, A. D. C. **Avaliação das técnicas de otimização para motores de busca**. 2012. 105 f. Dissertação (Mestrado) – Universidade do Minho, Escola de Engenharia, 2012.

REIS, G. A. **Centrando a arquitetura de informação no usuário**. 2007. 250f. Dissertação (Mestrado), Universidade de São Paulo, Escola de Comunicação e Artes: São Paulo, 2007.

RIBEIRO, F. B. **Otimização de motores de busca como método de melhoria do índice de qualidade na AdClick**. 2013. 41f. Dissertação (Mestrado) – Universidade do Porto, Faculdade de Engenharia: Porto, 2013.

RIBEIRO, O. B; VIDOTTI, S A B G. Otimização do acesso à informação científica: discussão sobre a aplicação de elementos da arquitetura da informação em repositórios digitais. **Biblos**, Rio Grande, v.23, n.2, p.105-106, 2009.

RICOTTA, F. **A origem do termo “Search Engine Optimization”**. 2008. Disponível em:< <http://www.fabioricotta.com/marketing-digital/a-origem-do-termo-search-engine-optimization.html>> Acesso em: 01 jun. 2016.

SÁ, R. M. P. **Gestão de Search Engine Optimization**. 2013. 121 f. Dissertação (Mestrado), Universidade de Trás-os-Montes e Alto Douro: Vila Real, 2013.

SARTI, E. **O básico do SEO Search Engine Optimization**. 2011. Disponível em:<<http://www.infowester.com/seo.php>> Acesso em: 27 jun. 2016.

SAWYA, Márcia Regina. **Dicionário de Informática e Internet**. São Paulo: Nobel, 1999. Disponível em:<
<http://comp.ist.utl.pt/aaa/Prog/Dicion%20rio%20De%20Inform%20tica%20&%20Inter%20net%20Ingl%20EAs-Portugu%20EAs.pdf>> Acesso em: 30 jan. 2015.

SILVA, E. L. **OntoAI**: uma ontologia para modelar o domínio da Arquitetura da Informação para web. 2010. 115f. Dissertação (Mestrado), Universidade Federal de Pernambuco, Centro de Informática: Pernambuco, 2010.

SILVA, M. F. Arquitetura da Informação. 2014. Slide. Disponível em:<
<http://marcelferrante.com.br/ai-2015/>> Acesso em: 16 set. 2015.

SILVA, M. A. T *et al.* O que é arquitetura da informação. **Biblionline**, João Pessoa, v.7, n.1, p.11-21, 2011.

SILVA, R. E; SANTOS, P.L.V. A; FERNEDA, E. M. Modelos de recuperação de informação e web semântica: a questão da relevância. **Inf. Inf.**, Londrina, v. 18, n.3, p.27-44, set./dez. 2013.

SMAAL, B. **Guia do google**: encontre até a agulha no palheiro. 2010. Disponível em:<
<http://www.tecmundo.com.br/google/3291-guia-do-google-encontre-ate-a-agulha-no-palheiro-.htm>> Acesso em: 3 maio 2016.

WU GARCIA, Léo M L S. **Investigação e implementação de ferramentas computacionais para otimização de websites com ênfase na descrição de conteúdo**. 2011. 126 f. Dissertação (mestrado). Universidade Estadual Paulista “Julio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas. Bauru, 2011.

GOMES, G. M. R.; CENDON, B.V. Análise da integração da recuperação da informação, information search behaviour e interação humano-computador para avaliação de sistemas de recuperação da informação. **TransInformação**, Campinas,v. 27(3):277- 284, set./dez., 2015.

ANEXO
REFERÊNCIAS DOS TEXTOS UTILIZADOS NA REDE

AGUIAR, W. F. Otimização para mecanismos de pesquisa (SEO): um estudo sobre a otimização de websites para sites de pesquisa de conteúdo na Internet. **Revista da Católica**, Uberlândia, v.5, n. 9, 2013.

ALBUQUERQUE, A. R. R. Sobre os fundamentos da Arquitetura da Informação. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 1, n. esp., p. 60-72, out. 2011.

ALBUQUERQUE, F. A. A. C. **Modelo de framework de Arquitetura da Informação baseado em roteiros**. 2014. 145 f. Dissertação (Mestrado) – Universidade de Brasília, Faculdade de Ciência da Informação, Brasília, 2014.

ANTUNES, M. A.; PIMENTA, F.J. P. A relação pauta versus palavra-chave: Possíveis implicações dos mecanismos de busca na agenda da mídia online. In: Intercom. 17., 2012. **Anais eletrônicos...** Universidade Federal de Juiz de Fora, Ouro Preto. 2012. Disponível em:<
<http://www.intercom.org.br/papers/regionais/sudeste2012/resumos/R33-0130-1.pdf>>
Acesso em: 19 maio de 2016.

AQUINO, M. A; OLIVEIRA, H. P.C. Contribuições da arquitetura da informação para website “A Cor da Cultura”. **Inf. & Soc. Est.**, João Pessoa, v.22, n.1, p.129-143, jan./abr.2012.

ARAÚJO JÚNIOR, R. H.; TARAPANOFF, K. Precisão no processo de busca e recuperação da informação: uso da mineração de textos. **Ci. Inf.**, Brasília, v. 35, n. 3, p. 236-247, set./dez. 2006.

AZEVEDO, M. M. V.; MARCONDES, C. H. Aportes da visualização da informação para sistemas de recuperação da informação. IN: ENANCIB. 16., 2015. João Pessoa. **Anais Eletrônicos...** Disponível em:<
<http://www.ufpb.br/evento/lti/ocs/index.php/enancib2015/enancib2015/paper/view/2685>> Acesso em: 19 maio 2016.

BARBANTI, C. H. **Representação e recuperação da informação em Centros de Memória**. 2015. 77f. Universidade de São Paulo, Escola de Comunicações e Artes, São Paulo, 2015.

BARBOSA, J. S.; SILVA, L. C. O.; SILVA, H. O. P. A recuperação de informação em trabalhos apresentados em encontros nacionais e regionais de estudantes de biblioteconomia, documentação, ciência e gestão da informação: uma proposta de utilização do programa open conference systems. **Biblos**, Rio Grande, v.23, n.2, p. 9-22, 2009.

BARROS, C. M.; VIERA, A. F. G. Mpeg-7 e a recuperação da Informação de objetos multimídia. **Inf. & Soc.:Est.**, João Pessoa, v.20, n.3, p. 135-144, set./dez. 2010.

BODÊ, E. C.; SOUSA, R. T.B. Preservação Digital, Recuperação da Informação e Linguagem. **RICI: R.Ibero-amer. Ci. Inf.**, Brasília, v. 8, n. 2, p. 122-141, jul./ dez. 2015.

BRÄSCHER, M. A Ambigüidade na Recuperação da Informação. **DataGramZero**, Rio de Janeiro, v. 3, n.1, fev. 2002.

BRASILEIRO, F. S; FREIRE, G. H. A. O marketing e a arquitetura da informação para web no contexto do processo de mediação da informação. **Biblioline**, João Pessoa, v.8, n. esp., p.161-174, 2012.

CAMARGO, L. S. A. **Metodologia de desenvolvimento de ambientes informacionais digitais a partir dos princípios da arquitetura da informação**. 2010. 287 f. Tese (Doutorado) – Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Marília, 2009.

CAMARGO, L. S. A.; VIDOTTI, S. A.B. G. Arquitetura da informação para biblioteca digital personalizável. **Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, n. esp., 2006.

CASELLI, L. M. V. **Utilização de Conhecimento Semântico comparada em domínio específico, baseado em expressões multipalavras**. 177f. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2013.

CORREA, R.. F.; VIEIRA, J. M. L. Representações visuais para recuperação de informação na BDTDUFPE. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.18, n.4, p.18-34, out./dez. 2013.

COSTA, F. M. S. P. **Search Engine and Optimization**. 2011. 65f. Dissertação (Mestrado) – Instituto Politécnico de Viseu, Escola Superior de Tecnologia e Gestão de Viseu, 2011.

COSTA, Ismael de Moura. **Um método para arquitetura da informação: fenomenologia como base para o desenvolvimento de arquiteturas da informação aplicadas**. 2009. 149 f. Dissertação (Mestrado em Ciências da Informação e Documentação)-Universidade de Brasília, Brasília, 2010.

DANTAS, C. M.; SILVA, H. O. Arquitetura da informação, acessibilidade e usabilidade: princípios básicos para análise do website da fundação centro integrado de apoio à pessoa portadora de deficiência (FUNAD). In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 14., 2013. Florianópolis. **Anais eletrônicos...**Florianópolis. Disponível em:< <http://enancib2013.ufsc.br/index.php/enancib2013/XIVenancib>> Acesso em: 19 maio 2016.

DANTAS, C. S.; SILVA, T. V. G.; SOUZA, A. C. B. Processo de recuperação da informação: barreiras encontradas pelos usuários. **Biblionline**, João Pessoa, v. 9, n. 1, p. 16-26, 2013.

DIAS, G. A.; VIDOTTI, S. A. B. G. Arquitetura da informação no ambiente digital: avaliando as relações com o direito da propriedade intelectual. **Inf. & Soc.:Est.**, João Pessoa, v.22, n.3, p. 115-132, set./dez. 2012.

DIAS, G. A.; VIDOTTI, S. A. B. G. O direito da propriedade intelectual: relações com os entregáveis da arquitetura da informação. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 1, n. Esp. p. 73-85, out. 2011.

DIAS, J. L. S. P. L. **Práticas de otimização integrada de sítios Web para leitores e motores de pesquisa**: um estudo empírico sobre sítios Web noticiosos. 2012. 113f. Dissertação (Mestrado) – Universidade do Porto, Faculdade de Economia, Porto, 2012.

DUQUE, C. G.; LYRA, M.R. O posicionamento da arquitetura da informação na governança de ti. **BJIS**, Marília (SP), v.4, n.2, p.41-46, jun./dez. 2010. Disponível em:< <http://www2.marilia.Unesp.br/revistas/index.php/bjis/index> > Acesso em: 22 maio 2016.

ESPANTOSO. J. J.P. A gestão dos espaços informacionais em ambientes de arquitetura da informação organizacional. **Inf. & Soc.:Est.**, João Pessoa, v.22, n.3, p. 33-39, set./dez. 2012.

FACHIN, G. R. B. Recuperação inteligente da informação e ontologias: um levantamento na área da Ciência da Informação. **Biblos**, Rio Grande, v.23, n.1, p. 259-283, 2009.

FERREIRA, S. M.S.P.; REIS, G. A prática de Arquitetura de Informação de websites no Brasil. **TransInformação**, Campinas, 20(3): 285-307, set./dez., 2008.

FUJITA, M. S. L.; GILVA-LEIVA, I. Avaliação da indexação por meio da recuperação da informação. **Ci. Inf.**, Brasília, v. 41 n. 1, p.50-66, jan./abr., 2014.

GARCIA, L. M. L. S. **Investigação e implementação de ferramentas computacionais para otimização de websites com ênfase na descrição de conteúdo**. 2011. 126 f. Dissertação (mestrado). Universidade Estadual Paulista “Julio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas. Bauru, 2011.

GASPAROTTO, M. Projetistas de interface para internet: a navegação nas mãos do usuário e a importância dos conceitos de Usabilidade e Arquitetura da Informação. **Comunicação & Inovação**, São Caetano do Sul, v. 8, n. 15, P.68-70, jul./dez. 2007.

GOMES, A. P. S. Recuperação da informação no acervo curt lange- ufmg: ensaio para levantamento de dados. **Biblos**, Rio Grande, v. 29, n.1, 2015.

GOMES, C. A.; ARAUJO, N. C. DESCRIÇÃO ARQUIVÍSTICA: a construção de um sistema de banco de dados para recuperação da informação. **Archeion Online**, João Pessoa, v.3, n.1, p. 45-64, jan./jun. 2015.

GOMES, G. M.; CENDON, B. V. Análise da integração da recuperação da informação, information search behaviour e interação humano-computador para avaliação de sistemas de recuperação da informação. **TransInformação**, Campinas, 27(3):277- 284, set./dez., 2015.

JANNUZZI, A. H. L.; AMORIM, R. C.R.; SOUZA, C. G. Implicações da categorização e indexação na recuperação da informação tecnológica contida em documentos de patentes. **Ci. Inf.**, Brasília, v. 36, n. 2, p. 27-34, maio./ago. 2007.

LACERDA, F.; LIMA-MARQUES, M. Da necessidade de princípios de Arquitetura da Informação para a Internet das Coisas. **Perspectivas em Ciência da Informação**, v.20, n.2, p.158-171, abr./jun. 2015.

LARA FILHO, D. O fio de Ariadne e a arquitetura da informação na WWW. **DaraGramZero**, Rio de Janeiro, v.4, n.6, p. 1-15, dez. 2003.

LEHMKUHL, K.M. **Os nativos digitais e a recuperação da informação científica on-line**. 2012. 165f. Dissertação (Mestrado) – Universidade Federal de Santa Catarina, Centro de Ciências da Educação. Santa Catarina, 2012.

LOPES, I. L. Estratégia de busca na recuperação da informação: revisão da literatura. **Ci. Inf.**, Brasília, v. 31, n. 2, p. 60-71, maio/ago. 2002.

MACHADO, I.M.R. **um gazetteer ontológico para recuperação de informação geográfica**. 2011. 85f. Dissertação (Mestrado) – Universidade Federal de Minas Gerais, Departamento de Ciência da Computação. Belo Horizonte, 2011.

MAIA, M. E *et al.* Análise sobre sistemas de busca na perspectiva da arquitetura da informação em ambiente de cordéis. **Biblos**, Rio Grande, v. 29, n.2, p.76-107, 2015.

MAPUTERE, M. C. J. **Influência dos fatores cognitivos e afetivos no processo de recuperação da informação: um estudo de caso em instituições de ensino superior na cidade de chimoio, manica, Moçambique**. 2015. 165f. Dissertação (Mestrado) - Universidade Federal de Santa Catarina, Centro de Ciências da Educação. Florianópolis, 2015.

MARTINS, E. C. D.; CARVALHO, T. Recuperação da informação em psicologia: LILACS e Index Psi Revistas Técnico-Científicas. **Perspectivas em Ciência da Informação**, v.19, n.2, p.118-130, abr./jun. 2014.

MELO, D. O. **Recuperação da informação com realimentação de relevância apoiada em visualização**. 2014. 95f. Dissertação (Mestrado) – Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação, São Carlos, 2014 .

MIRANDA, Z. D. Análise do ambiente SIS médicos e a cultura a partir da arquitetura da informação: enfocando o sistema de organização. **TPBCI**, Belo Horizonte, v.5, n.1, 2012.

MIROIR, J.C. A“otimização dos mecanismos de busca” (seo)como ferramenta de coleta automatizada de documentos para elaboração de corpora. **Blucher Social Sciences Proceedings**, São Paulo, v.2, n.3, mar. 2016.

NASCIMENTO NETO, G. H. **Arquitetura da Informação no comércio eletrônico de livros no Brasil**: dimensões que norteiam a e-satisfação do usuário. 2010. 113 f. Dissertação (Mestrado) – Universidade Federal da Paraíba, Centro de Ciências Sociais Aplicadas, 2010.

NONATO, R. S *et al.* Arquitetura da informação em bibliotecas digitais: uma abordagem da ciência da informação e da biblioteconomia. **Inf. Inf.**, Londrina, v. 13, n. 2, p. 125-141, jul./dez. 2008.

OKADA, S. I.; SOUZA, E. M. S. Estratégias de marketing digital na era da busca. **REMark**, São Paulo, v. 10, n. 1, p 46-72, jan./abr. 2011.

OLIVEIRA, A. M. Search engine optimization SEO: a contribuição do bibliotecário na otimização de websites para os mecanismos de busca. **Perspectivas em Gestão e Conhecimento**, João Pessoa, v.1, n. esp. p.137- 159, 2011.

OLIVEIRA, G. A.; ARAÚJO, W. J. Usar ou não usar : qual a relevância das meta tags na recuperação da informação pelos mecanismos de busca. **Biblionline**, João Pessoa, v. 8, n. 1, p. 60-77, 2012.

OLIVERIA, H. P. C. **Arquitetura da informação pervasiva**: contribuições conceituais. 2014. 202 f. Tese (Doutorado) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Filosofia e Ciências. Marília, 2014.

ORRICO, E. G. D. A Representação Metafórica como Filtro de Recuperação da Informação. **DataGramZero**. Rio de Janeiro, v.2 n.5, out. 2001.

PAES, D. M. B.; TABOSA, H. R. Biblioteca digital de teses e dissertações: reflexões sobre representação da informação com vistas á recuperação da informação. **Revista ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 20, n. 2, p. 225-239, maio/ago., 2015.

PASSOS, P. C. S. J. **Perspectivas para as revistas científicas no contexto da colaboração em rede**: um enfoque da arquitetura da informação. 2016. 256 f. Tese (Doutorado) – Universidade Federal do Rio Grande do Sul, Faculdade de Biblioteconomia e Comunicação, Porto Alegre, 2016.

PATO, R. G. **Imagens**: polissemia *versus* indexação e recuperação da informação. 2015. 340f. Tese (Doutorado) – Universidade de Brasília, Brasília, 2015.

PINTO, M. D. S. Diversidade da recuperação de informações via internet na pesquisa e ensino. **Rev. ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 8/9, p. 143, 2004.

PONTES JÚNIOR, J.; CARVALHO, R. A.; AZEVEDO, A. W. Da recuperação da informação à recuperação do conhecimento: reflexões e propostas. **Perspectivas em Ciência da Informação**, v.18, n.4, p.2-17, out./dez. 2013.

POPADIUK, S *et al.* Arquitetura da Informação e Mensuração do Desempenho: Um Estudo na Indústria de Artefatos e Utensílios de Plásticos no Estado de São Paulo. **Gestão e Produção**. São Carlos, v.13, n.1, p.151-165, jan./abr. 2008.

QUINTANA, A. D. C. **Avaliação das técnicas de otimização para motores de busca**. 2012. 105 f. Dissertação (Mestrado) – Universidade do Minho, Escola de Engenharia, 2012.

RAMOS, C. L *et al.* Search engine optimization: um estudo de caso sobre os benefícios de sua aplicação em websites. **Tekhne e Logos**, Botucatu, SP, v.7, n.1, abril, 2016.

RAMOS, C.; MUNHOZ, D.P. A subjetividade da relevância na recuperação da informação: análise a partir de imagens representativas. **Biblos**, Rio Grande, v. 25, n.1, p.69-79, jan./jun. 2011.

REIS, G. A. **Centrando a arquitetura de informação no usuário**. 2007. 250f. Dissertação (Mestrado), Universidade de São Paulo, Escola de Comunicação e Artes: São Paulo, 2007.

RIBEIRO, F. B. **Otimização de motores de busca como método de melhoria do índice de qualidade na AdClick**. 2013. 41f. Dissertação (Mestrado) – Universidade do Porto, Faculdade de Engenharia: Porto, 2013.

RIBEIRO, F. F; MONTEIRO, S. D. Arquitetura da informação em sites de pré-reitorias de graduação: um enfoque nas instituições estaduais de ensino Superior do Paraná. **Inf. Inf.**, Londrina, v. 17, n. 3, p. 125 – 164, set./dez. 2012.

RIBEIRO, O. B; VIDOTTI, S A B G. Otimização do acesso à informação científica: discussão sobre a aplicação de elementos da arquitetura da informação em repositórios digitais. **Biblos**, Rio Grande, v.23, n.2, p.105-106, 2009.

RIGO, W. **Semântica e visualização para anotação e recuperação de informação**. 2011. 158f. Dissertação (Mestrado) - Universidade Federal de Santa Catarina, Florianópolis, 2011.

RODRIGUES, C. B.; CRIPPA, G. A recuperação da informação e o conceito de informação: o que é relevante em mediação cultural. **Perspectivas em Ciência da Informação**, v.16, n.1, p.45-64, jan./mar. 2011.

SÁ, R. M. P. **Gestão de Search Engine Optimization**. 2013. 121 f. Dissertação (Mestrado), Universidade de Trás-os-Montes e Alto Douro: Vila Real, 2013.

SÁ, R.C. **Recuperação da informação fotográfica: o arquivo**

SAMPAIO, D. A.; SOUZA, A. M. C.; SILVA, T.P. Interfaces entre controle bibliográfico e recuperação da informação. **Biblionline**, João Pessoa, v. 8, n. 1, p. 3-11, 2012.

SANTOS, L. C. M.; VIERA, A.F.G. Avaliação da recuperação da informação em acervos digitais de jornais. **Em Questão**, Porto Alegre, v. 21, n. 2, p. 49-73, mai/ago. 2015.

SILVA, D. B. **Recuperação de informação de comunicados à imprensa**. 2011. 147 f. Dissertação (Mestrado) – Universidade de São Paulo, Faculdade de Economia, Administração e Contabilidade. São Paulo, 2011.

SILVA, Edson Marchetti da. **Recuperação da informação através de busca comparada em domínio específico, baseado em expressões multipalavras**. 177f. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2013.

SILVA, M. A. T *et al.* O que é arquitetura da informação. **Biblionline**, João Pessoa, v.7, n.1, p.11-21, 2011.

SILVA, M. A. T.; PINHO NETO, J. S.; DIAS, G. A. Arquitetura da Informação para quem e para quem?: uma reflexão a partir da prática em ambientes informacionais digitais. **Encontros Bibli.** v. 18, n. 37, p. 283-302, mai./ago., 2013.

SILVA, N. B. X.; SOUSA, M. R. F. A dimensão tecnológica da gestão do conhecimento e a contribuição da arquitetura da informação: Uma análise da plataforma pódio. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 5, n. 2, p. 186-200, jul./dez. 2015.

SILVA, P. M. A arquitetura da informação centrada no usuário: estudo do website da biblioteca virtual em saúde (BVS). **Enc. Bibli.** Florianópolis, n. 26, 2008.

SOUSA, E. A. A. **Arquitetura da informação e usabilidade em ambientes informacionais digitais**: avaliação da intranet da Creduni João Pessoa. 2014. 112f. Dissertação (Mestrado) – Universidade Federal da Paraíba, Programa de Pós-Graduação em Ciência da Informação, João Pessoa, 2014.

SOUSA, M. R. F. O acesso a informações e a contribuição da Arquitetura da informação, usabilidade e Acessibilidade. **Inf. & Soc.:Est.**, João Pessoa, v.22, p. 65-76, n. esp. 2012.

SOUSA, R. P. M *et al.* O exercício da arquitetura da informação na Web: reflexões relativas à Lei dos Direitos Autorais. **Encontros Bibli.** v. 18, n. 36, p.107-128, jan./abr., 2013.

SOUSA, R. T. B.; ARAÚJO JÚNIOR, R. H. A classificação e a taxonomia como instrumentos efetivos para a recuperação da informação arquivística. **Ci. Inf.**, Brasília, DF, v. 42 n. 1, p.131-144 jan./abr., 2013.

SOUZA, O.; TABOSA, H. R. A eficácia dos modelos de recuperação de

Informações: um estudo particularizado na Comunicação científica na web. In: ENACIB. 16., 2015. João Pessoa. **Anais eletrônico...**
Disponível: <<http://www.ufpb.br/evento/lti/ocs/index.php/enancib2015/enancib2015/paper/view/3018/1188>> Acesso em: 19 maio, 2016.

STEINMETZ, E. F. P. S. **Processo de Organização da Informação para a Superior do Paraná.** *Inf. Inf.*, Londrina, v. 17, n. 3, p. 125 – 164, set./dez. 2012.

TEIXEIRA, R. S. Serviço de recuperação da informação na biblioteca de um laboratório farmacêutico: um estudo prático. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 2, n.2, p 80-89, jan./jun. 2005.

TEIXEIRA, T. M.C.; VALETIM, M. L. P. Estratégias para disseminação do Uma análise da plataforma pódio. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 5, n. 2, p. 186-200, jul./dez. 2015.

VASQUEZ, L.C.C; SILVA, P.C. **Solução para desenvolvimento da ferramenta k-seo de análise de palavras-chave.** Disponível em: <https://www.researchgate.net/profile/Paulo_Caetano_da_Silva/publication/268258842_Solucao_para_Development_da_Ferramenta_K-SEO_de_Analise_de_Palavras-Chave/links/54bd06ca0cf218da93905783.pdf > Acesso em: 19 maio de 2016.

VÉRAS NETO, C.; SILVA, M. B.; DIAS, G. A. Avaliação da implantação do ged nos arquivos da UNIMED/NNE como forma de recuperação da informação. **Biblionline**, João Pessoa, v. 6, n. 1, p. 125-141, 2010.

VIEIRA, J. M. L. **A contribuição da organização e da visualização da informação para os sistemas de recuperação da informação.** 2014. 227f. Dissertação (Mestrado) - Universidade Federal de Pernambuco, Centro de Artes e Comunicação. 2014.

VIEIRA, J. M. L.; PINHO, F. A. A contribuição da organização e da visualização da informação para os sistemas de recuperação de informação. *Inf. Inf.*, Londrina, v. 20, n. 1, p. 110 - 136, jan./abr. 2015.