# ANÁLISE DA REPRESENTAÇÃO DA INFORMAÇÃO EM MODELOS ENTIDADE RELACIONAMENTO COM BASE EM METADADOS

*Célia Medeiros Dantas*[1]
*Flavio Ribeiro Cordula*[2]
*Wagner Junqueira Araújo*[3]

**RESUMO**

No contexto atual, a informação, atrelada às tecnologias, têm-se tornado imprescindíveis para as atividades cotidianas dos seres humanos, pois a sociedade contemporânea necessita de elementos que possibilitem a organização, disseminação e preservação da informação, independentemente do suporte. O trabalho desenvolvido para este artigo teve como objetivo refletir sobre a representação da informação para o processo de recuperação informacional no ambiente virtual. Tendo como base os campos representativos do modelo de metadados Dublin Core e a elaboração de modelos de Entidade-Relacionamento para banco de dados. A relevância dos metadados para a representação da informação é enfatizada quando esta é aplicada para o acesso as bases de dados. Assevera que a informação adquiriu novas características a partir de sua interseção com a tecnologia e a Internet, as quais são agentes responsáveis pelas mudanças no fluxo informacional. Para a análise e reflexão foi elaborado um exemplo de como se dá a interação entre representação da informação aplicando-o modelo de metadados Dublin Core em uma ficha catalográfica e como seriam as possíveis alternativas de modelagem para uma base de dados. Como resultado, demonstra que uma representação errônea dos campos no modelo de banco de dados, pode provocar transtornos como: demora, ou dificuldade, ou até mesmo a irrecuperabilidade das informações contidas em documentos registrados nas bases de dados. Ressalta a necessidade do profissional da informação participar do processo de desenvolvimento dos modelos de representação para bases de dados.

**Palavras-chave**: Representação da Informação. Recuperação da Informação. Bases de Dados.

## 1 INTRODUÇÃO

A informação é ao mesmo tempo mercadoria e insumo para a produção de conhecimento. Ela é de vital importância para o desenvolvimento e evolução da sociedade. Foi apenas com a explosão informacional, ocorrida durante a Segunda Guerra Mundial que intelectuais, dentre eles, Bush (1945), passaram a observar, analisar,

[1]Mestre em Ciência da Informação pela UFPB- Universidade Federal da Paraíba. Graduada em Arquivologia pela UEPB-Universidade Estadual da Paraíba.
[2] Mestre em Ciência da Informação pela Universidade Federal da Paraíba (2015) e Graduado em Ciências da Computação pelo Centro Universitário de João Pessoa (2008).
[3] Doutor em Ciência da Informação Pela Universidade de Brasília (2009); Mestre em Ciência da Informação - UNB (2001) e Bacharel em Ciência da Computação Pela Universidade do Oeste Paulista (1993). Professor do Programa de Pós-Graduação-em Ciência da Informação (PPGCI)/UFPB. Professor do Programa de Pós Graduação em Gestão em-Organizações Aprendentes (MPGOA)/UFPB. Professor Adjunto III - Departamento de Ciência da Informação da UFPB.


(cc) BY-NC-ND

estudar a informação a partir de sua interseção com a tecnologia. Embora esta escolte o homem desde muito antes da invenção do ábaco, objetivando na melhor forma de produção, disponibilização, recuperação, armazenamento e preservação da informação seja ela digital ou em suporte tradicional (papel, por exemplo).

A Internet, principalmente a partir da década de 90, com o declínio dos custos do computador pessoal e dos serviços de acesso às redes, foi a grande responsável pelas mudanças ocorridas em parte dos canais de comunicação e no fluxo de informação. Entretanto, essa inserção tecnológica não veio sem algum tipo de resistência por parte de alguns profissionais, o pessoal mais antigo nas organizações sempre apresenta resistência as mudanças. Corroborando com esse raciocínio, Veloso (2011) explicita que certas vezes essa resistência ocorre devido a incompreensão da relevância das tecnologias de informação e comunicação (TICs) no ambiente de trabalho, impossibilitando o desenvolvimento de suas atividades.

Logo, percebe-se um acréscimo a produções acadêmicas e relatos técnicos informando sobre as contribuições da aplicação das TICs como forma de aperfeiçoar o trabalho e, consequentemente, oferecendo ferramentas necessárias para o desenvolvimento do trabalho corriqueiro. A Ciência da Informação, é uma das áreas de estudo que contribuem para a disponibilização e o tratamento das informações, tendo seu aporte teórico em constante evolução.

Nesse viés, este trabalho tem como objetivo discutir elementos sobre a representação e recuperação da informação (RRI) no ambiente tecnológico e defendendo a premissa de que um dos principais desafios da recuperação da informação está relacionado com a forma de como os documentos devem ser representados.

## 2 TECNOLOGIA DA INFORMAÇÃO E COMUNICAÇÃO

No campo científico, percebe-se que há uma diversidade nas perspectivas e análises, dos impactos das tecnologias no âmbito social, principalmente, no que tange as transformações decorrentes do uso comum e cotidiano das tecnologias da informação e comunicação.


(cc) BY-NC-ND

O paradigma tecnológico do século XX apresentou novas possibilidades e modificou os processos de economia e relações socioculturais. A tecnologia, para Castells e Lévy, não pode ser considerada um ator aparte da sociedade. Para Lévy (2000, p. 26) há a necessidade de "situar as irreversibilidades às quais um de seus usos nos levaria, de formular os projetos que explorariam as virtualidades que ela transporta e de decidir o que fazer com ela", consequentemente, não há um imperativo em avaliar os seus impactos.

No campo da Ciência da Informação, se percebe que ao decorrer dos anos, esta, passou a manter um forte elo com as tecnologias por meio de trabalhos sobre bibliotecas digitais, documentos eletrônicos, repositórios digitais, trabalhos referentes a infometria, bases de dados etc. É fato que, esse elo se fortificou também em outras áreas, como da psicologia, comunicação, administração, dentre tantas outras. Considerando CI como uma Ciência interdisciplinar, apesar desta, ser apresentada por diferentes pontos de vista como intradisciplinar, interdisciplinar, multidisciplinar e transdisciplinar. Segundo a definição de Borko (1968), a Ciência da Informação é complexa e multidimensional e, por esta razão, deve ser intencionalmente abrangente. Sobre o exposto afirma:

> Ciência da Informação é a disciplina que investiga as propriedades e comportamento da informação [...]. Ela refere-se à origem, disseminação, coleta, organização, recuperação, interpretação, transmissão, transformação e uso da informação. Isso inclui a representação da informação em ambos os sistemas naturais e artificiais, o uso dos códigos para uma transmissão de mensagens eficiente e o estudo de mecanismos de processamento de informação e técnicas, tais como computadores e seus sistemas. (BORKO, 1968, p. 3, tradução nossa).

Segundo a concepção de Freire e Freire (2010, p. 13), um dos objetivos da CI seria "contribuir para a informação se tornar, cada vez mais, um elemento de inclusão social, trazendo desenvolvimento para pessoas e nações". Em outras palavras, o que os autores enfocam é que tal ciência não deve ser tratada como uma abordagem apenas teórica ou tecnicista e sim preocupada com fatores como a responsabilidade social, que o cientista da informação deve perceber e pregar em suas atividades cotidianas, usando a informação de modo a beneficiar um ou mais indivíduos.

Segundo Le Coadic (2004, p. 5-7) o futuro da informação está designado por dois pontos,


(cc) BY-NC-ND

1. A explosão quantitativa da informação. Fundamentando-se o progresso técnico e social no poder criativo da linguagem e do raciocínio lógico que daí resulta, pode-se compreender a importância da comunicação verbal da informação. Com o advento da escrita, a comunicação passou de oral a escrita.
2. A implosão do tempo de comunicação da informação. Não há mais distância que seja obstáculo à velocidade, nenhuma fronteira detém a informação. A velocidade dos computadores se mede em milionésimos de segundos.

Ainda sobre este tema, Mattos (2010) constata que se outrora a escassez informacional era um empecilho, hoje a questão chave é justamente esta produção desenfreada. Mais do que isso, na verdade o "tempo do produtor geralmente é curto. O do consumidor mais ainda. Por isso não convém gastar muito tempo com o desenvolvimento histórico e exigir do consumidor muito esforço a mais, por mais interessantes que o contexto possa ser." (MATTOS, 2010, p. 40).

Todavia, é importante avaliar a riqueza que essa torrente de informações pode trazer para os usuários. As tecnologias e a Internet disponibilizam uma fonte inesgotável de informações e possibilitam a abertura da maior biblioteca do mundo dentro da casa de cada pessoa, ao alcance de um simples clique do mouse. Destaca-se, ainda, a nova noção de tempo na era da Internet que passa a ter os minutos comparados com horas, horas com dias, e assim por diante. Que desencadeou uma necessidade em divulgar informações e uma urgência da sociedade em recebê-la e consumi-la.

Talvez a principal finalidade das novas tecnologias seja o processamento da informação que passa a mediar à apreensão da necessidade informacional e do conhecimento. Quanto a esse fator, Chu (2007, p. 17, tradução nossa) revela que "o problema essencial na representação e recuperação da informação permanece em como obter a informação correta para o usuário correto no momento certo, apesar da existência de outras variáveis (p. ex., características do usuário ou cobertura da base de dados) no ambiente de RRI".

Destrinchando sobre a representação da informação, Novellino (1998) destaca que inicialmente se refere a um processo de transferência informacional apresentando as partes importantes do documento visando sua posterior recuperação. A dificuldade está em estabelecer quais pontos são relevantes para uma maior parcela de usuários e não separar os processos de representação e recuperação da informação, como ocorria

na perspectiva tradicionalista, como se um não causasse interferência no outro e vice-versa.

## 3 RECUPERAÇÃO DA INFORMAÇÃO

É fato que a informação, independente de seu formato, precisa ser representada para poder ser recuperada. Blair (1990) afirma que a principal preocupação da recuperação da informação é como os documentos devem ser representados para que possam ser recuperados. Sobre o exposto, Heting Chu compartilha da mesma ideia e assevera que a representação da informação

> [...] inclui a extração de alguns elementos (p.ex, palavras-chave ou frases) de um documento ou a atribuição de termos (p.ex, descritores ou cabeçalhos de assunto) a um documento de modo que a sua essência pode ser caracterizada e apresentada. Tipicamente, a representação da informação pode ser feita via qualquer combinação dos seguintes meios: abstração, indexação, categorização, sumarização e extração (CHU, 2010, p. 15, tradução nossa).

A recuperação de documentos e informações em bases de dados tem sido assunto discutido por profissionais da Ciência da Informação. A Internet e a *Web* tornam esse tema ainda mais relevante na medida em que facilitou um fluxo maior de informações. Segundo Chu (2010), a *Networked Era* (era da rede de conexões, tradução nossa) teve seu início nos anos de 1990 e se faz presente até os dias de hoje. A Internet é o grande símbolo dessa era e proporciona uma nova plataforma para a representação e recuperação da informação.

A recuperação da informação (RI) é um procedimento pela qual se seleciona as informações contidas nos documentos[4] a partir de uma grande coleção ou acervo – para documentos digitais, pode ser entendida como um repositório ou uma base de dados digital – ou, ainda, a operação que fornece uma resposta, até certo ponto elaborada, a uma demanda, e tal resposta é transformada em produto. Segundo Ferneda (2003, p. 14), "há ainda autores que conceituam a recuperação de informação de forma muito

[4] De acordo com o Dicionário de Terminologia Arquivística (2005, p. 73) considera-se como documento a "Unidade de registro de informações, qualquer que seja o suporte ou formato". Assim, o documento pode ser informações registradas em CD, *pen drive*, papel, etc. em formatos variados como livro, caderno, folha, etc.


(cc) BY-NC-ND

mais ampla, ao subordinar à mesma o tratamento da informação (catalogação, indexação, classificação)".

Mooers (1951), criador do termo *Information Retrieval* (Recuperação de Informação), indica que, "trata dos aspectos intelectuais da descrição da informação e sua especificação para busca, e também de qualquer sistema, técnicas ou máquinas que são empregadas para realizar esta operação".

A recuperação de informação digital segue o conceito do *two steps information retrieval* ou recuperação de informação em duas etapas (tradução nossa). Para González de Gómez (2004),"o primeiro *step* consistiria em um dispositivo de representação que descreve, sumariza e codifica as fontes primárias de informação para facilitar sua busca seletiva e sua localização efetiva". Hjorland (2002) afirma que no primeiro *step* são construídas representações que oferecem diferentes pontos de acesso às fontes primárias de informação como tema, autor, título, palavras-chave, etc. O segundo *step*, por sua vez, seria a localização e disponibilização das fontes primárias escolhidas, nelas mesmas, dando acesso a seu texto completo (GONZÁLEZ DE GÓMEZ, 2004).

Lancaster (2004) afirma que os sistemas de recuperação de informação (SRIs) são originários das imensas bases de dados de instituições americanas como: *National Library of Medicine* (*NLM*), *Department of Defense* (*DoD*) e *National Aeronautics and Space Administration* (*NASA*). Hoje, um SRI é capaz de enviar em questão de segundos, a qualquer canto do mundo, grandes volumes de dados e de armazenar bilhões de itens de informação ou *terabytes* de arquivos em computadores pessoais comuns.

Os sistemas de recuperação de informação viabilizam acesso a documentos, tais como leis, notícias, fotografias, jornais, notícias, revistas, periódicos científicos e informações a partir da execução de tarefas de representação da informação. Tal processo se materializa pela representação e, também, pela recuperação das informações representadas e dos documentos armazenados com a finalidade de atender às necessidades dos usuários (SOUZA, 2006).

Sobre os SRIs, Lopes (2002, p. 60) assevera que "esses sistemas possibilitam o planejamento de estratégias de busca com maior nível de complexidade envolvendo vários conceitos na mesma estratégia". Os SRIs permitem a utilização de variados tipos de busca sendo que no caso específico dos sistemas de recuperação de informação


(cc) BY-NC-ND

científica, essas buscas podem ser realizadas utilizando-se dos metadados "Título", "Resumo", "Palavras-chave", "Autor" etc. Ou seja, os sistemas de recuperação da informação usam termos da linguagem natural (LOPES, 2002).

Além dos modelos ou técnicas clássicas de recuperação de informação, a saber, booleano, vetorial e probabilístico (CARDOSO, 2000; FERNEDA, 2003), várias outras técnicas, algumas mais atuais, outras mais avançadas, são utilizadas na modelagem de sistemas de recuperação de informação.

Chu (2010) divide as principais técnicas que envolvem a recuperação da informação em dois grupos, sendo o primeiro grupo, de técnicas das abordagens básicas, formado pela pesquisa booleana, *case sensitive*, de truncamento, por proximidade e pesquisa de campo, e o segundo grupo, dos métodos avançados, formado pelas técnicas de pesquisa *fuzzy*, pesquisa ponderada, expansão de consulta (*query*) e pesquisas em múltiplas bases de dados.

Nesse contexto, ressalta-se que a pesquisa *case sensitive* distingue as letras maiúsculas das minúsculas, e, assim, uma expressão de busca "Base de Dados" torna-se diferente de "base de dados". O truncamento é outra técnica interessante e bastante usada na *web*. Ele consiste em retornar resultados a partir da variação de uma palavra-chave em combinação com o uso de um símbolo, geralmente o asterisco (*) ou a porcentagem (%). Por exemplo, a expressão de busca "* de dados" pode indicar ao mesmo tempo, banco de dados ou base de dados ou ainda qualquer outra expressão chave que finalize com "de dados".

A pesquisa *fuzzy*, entretanto, é uma técnica mais avançada que consiste em um processo que localiza as páginas *web* que possam ser relevantes para uma determinada expressão de busca, mesmo, quando o argumento não corresponde exatamente à informação desejada.

### a) Bases de dados e o Armazenamento de Informação

As bases de dados (BD) provocaram – e ainda provocam – um grande impacto no uso de computadores. Eles representam um papel crítico em todas as áreas de conhecimento em que o computador e suas tecnologias específicas são utilizados: nas


(cc) BY-NC-ND

áreas afins de Negócio e *Business*, nas Engenharias, na Matemática e na Física, na Medicina, na Educação, na Ciência da Informação, entre outras áreas.

Zaninelli e Catarino (2004) afirmam que atualmente o meio eletrônico é uma forma rápida de armazenar, transmitir, compartilhar e recuperar documentos e informações. As autoras ainda expõem que:

> Um recurso muito utilizado e eficaz para a disponibilização e troca de informações entre organizações públicas e privadas, é o banco de dados e, consequentemente, as suas bases de dados que ampliam significativamente a qualidade das buscas informacionais, visto que essas bases apresentam diversificados pontos de acesso à informação. (ZANINELLI; CATARINO, 2004, p.2).

Elmasri e Navathe (2005) definem dados como sendo fatos que podem ser gravados e possuem um significado implícito e um banco de dados como uma coleção de dados relacionados e que possuem propriedades implícitas, são elas:

- a representação de aspectos do mundo real, chamado de minimundo ou universo do discurso. As mudanças no minimundo são refletidas em um banco de dados;
- a coleção lógica e coerente de dados com algum significado inerente. Uma organização de dados ao acaso não pode ser considerada um banco de dados;
- um projeto construído e povoado por dados, atendendo a uma proposta específica;
- a posse de um grupo de usuários definido e algumas aplicações preconcebidas de acordo com o interesse desse grupo de usuários.

Um banco de dados apresenta enorme capacidade de armazenamento e tende a tornar-se mais complexo dependendo das características do mundo real nele modeladas. Por exemplo, uma lista de clientes pode conter apenas uma centena de registros de nomes, cada um com uma estrutura simples. Já um catálogo de biblioteca pode conter, facilmente, alguns milhões de registros – nome e sobrenome dos autores, títulos e subtítulos, resumos e palavras-chave, entre outros – sendo cada categoria de registros relacionada entre si e organizada de maneira específica.

Por sua vez, um Sistema Gerenciador de Banco de Dados (SGBD) é um programa ou *software* que permite a criação e manutenção de uma coleção de bancos de dados. Para Elmasri e Navathe um SGBD é:


(cc) BY-NC-ND

> Um *sistema de software de propósito geral* que facilita os processos de definição, construção, manipulação e compartilhamento de bancos de dados entre vários usuários e aplicações. A **definição** de um banco de dados implica especificar os tipos de dados, as estruturas e as restrições para os dados a serem armazenados em um banco de dados. [...] Outras funções importantes do SGBD são a proteção e manutenção do banco de dados por longos período (ELMASRI; NAVATHE, 2005, p.4, grifos dos autores).

Os termos "Banco de Dados" e "Base de Dados" não devem ser usados de forma intercambiável. Cianconi (1987, p.55) distingue esses conceitos ao afirmar que uma base de dados é "um conjunto de dados inter-relacionados, organizados de forma a permitir a recuperação de informação" e banco de dados é "um conjunto de bases de dados".

A definição de base de dados é vista por Heemann (1997, p. 2) como "um arquivo ou um conjunto de arquivos computacionais no qual são armazenados dados, permitindo a recuperação e atualização de informações". Já para Rowley (1994, p.66) uma base de dados é "uma coleção de registros similares entre si e que contém determinadas relações entre esses registros". Esses registros a que Rowley se refere são dados, informações que a base possui sobre um determinado documento ou item.

O objetivo de uma base de dados é prover informação atualizada, precisa e confiável para de acordo com a demanda apresentar o que o usuário necessita. Um requisito fundamental de base de dados, além de armazenamento eficiente de dados, é o fornecimento de mecanismos eficientes de recuperação de informação.

Para este fim, o profissional da informação deve estar atento as medidas a serem tomadas para a representação da informação. Diante disso, a próxima seção trata sobre a representação por meio da utilização dos atributos estipulados em padrões de metadados, usando como exemplo o padrão Dublin Core.

## 4 REPRESENTAÇÃO DA INFORMAÇÃO E METADADOS

Na contemporaneidade a informação é um elemento de poder e provocador de mudanças nas estruturas sociais, políticas, econômicas e culturais. Contudo, para se alcançar esse feito "a informação deve ser ordenada, estruturada ou contida de alguma forma, senão permanecerá amorfa e inutilizável" (MCGARRY, 1999, p.11). Logo, os processos de criação, armazenamento e disponibilização das informações


(cc) BY-NC-ND

destrinchavam e exigiam um olhar apurado por parte dos profissionais da informação. Todavia, como ressalta Kobashi e Fernandes (2009, p. 3,) "representar é um termo ambíguo. Pode ser entendido como processo ou como produto".

Por esse motivo, na Ciência da Informação as preocupações têm-se voltado para os componentes representativos da informação, tanto em ambientes físicos quanto em ambientes digitais, para garantir a salvaguarda e a disponibilização informacional sempre que solicitado. Nesse contexto, Novellino (1998, p. 137) assevera que:

> Uma das atribuições da ciência da informação é construir teorias e elaborar métodos para a transferência da informação, fundamentando o estabelecimento de canais formais para a comunicação da informação. A ação de transferência da informação engloba representação, recuperação e comunicação. Tradicionalmente, porém, ênfase maior é dada à representação.

A partir dessa perspectiva têm-se percebido a evolução e as contribuições de demais áreas consolidadas como, por exemplo, a Biblioteconomia para dar o aporte sobre a representação da informação e garantir que ela seja acessada de modo eficiente. Todavia, como toda evolução, as atividades corriqueiras que outrora eram realizadas apenas em suporte físico, necessitaram da disposição em ambiente digital de aprimoramentos pragmáticos como teóricos.

Com esse pensamento, Novellino (1996) constatou que diante da necessidade de um tratamento temático da informação, devido à especificidade dos acervos e à conveniência em tornar os processos representativos e de recuperação eficientes, deixou-se de lado algumas práticas como, por exemplo, o controle do vocabulário. Dessa forma, a representação deve estar atrelada a comportamentos organizativos, bem como ao monitoramento de procedimentos indexadores, por exemplo.

Não obstante, qual é o intuito do método de representação, em si? Para Ferneda (2003, p. 16), o processo de representação "busca descrever ou identificar cada documento do corpus através de seu conteúdo. Tal representação geralmente é realizada através do processo de indexação. Durante a indexação são extraídos conceitos do documento através da análise". Ratificando essa linha de pensamento, podemos concluir que a representação proporciona a extração dos elementos essenciais contidos


(cc) BY-NC-ND

em um documento, por exemplo. E é por essa razão que Chu (2007, p. 25, tradução nossa) declara que:

> [...] como o esquema de representação pode ser aplicado de forma consistente e precisa, ainda apresenta desafios para os profissionais da informação, mesmo se as escolhas dos esquemas não forem a preocupação e alguns métodos (p. ex., abstração) não necessitarem de um esquema para ser implementados. Representação da Informação, em poucas palavras, significa desafio, complexidade e sutileza.

Sobre esse pensamento de Chu, atenta-se aqui a dificuldade dos profissionais da informação no que cerne a descrição informacional, principalmente no ambiente digital, haja vista a subjetividade dos usuários e a complexidade em estabelecer parâmetros para os estudos de usuários com vista na determinação adequada quanto a busca de informação para a posteriori ser usada.

Enfatizamos que a representação da informação digital, tendo como foco atender às exigências informacionais dos usuários, não visa apenas entender o processo de busca-acesso- uso da informação, mas, consecutivamente a isso, estabelecer mecanismos de acesso que facilitem tal relação de busca com poucos cliques. Esses estudos, portanto, aperfeiçoam os procedimentos de acesso à informação e, consequentemente, o próprio sistema que o armazena.

Dentre os principais métodos convencionais que envolvem a representação da informação, Chu (2007) atenta para três. São eles: a indexação, a categorização e a sumarização. As atividades de indexação, de acordo com Leiva (2012), foram iniciadas pelos antigos escribas da Mesopotâmia. A indexação visa extrair as informações essenciais do documento para uma boa representação em que sua qualidade se dará por meio da exaustividade, da especificidade, da correção e da consistência. Para este fim, Manini (2002) informa que a indexação apresenta duas principais fases em que a primeira caracteriza-se pela análise conceitual, ou seja, é analisado o assunto evidenciado no documento, e a segunda corresponde à tradução em si.

No que cerne à categorização, conforme Jacob e Shaw (1998) se configuram por ser um processo cognitivo dos sujeitos com os quais as experiências são classificadas, através do estabelecimento de entidades ou categorias, possibilitando um ordenamento


(cc) BY-NC-ND

físico e social. Nesse sentido, são utilizadas categorias como forma de classificar determinadas informações e toma como base as categorias pré-estabelecidas na *Dewey Decimal Classification* (*DDC*) e *Library of congress classification* (*LCC*).

Já a sumarização, a partir da visão Chu (2007, p. 29-30, tradução nossa), "[...] tenta representar um documento com um ou dois parágrafos - um sumário ou uma cópia resumida do original. O grau de condensação e método utilizado define as diferentes abordagens de compactação". Ou seja, são sínteses de um documento e tentam extrair a essência do conteúdo apresentado no suporte. Quanto aos tipos de sumarização, Chu (2007) elenca três, sendo eles: resumo se caracteriza por ser uma síntese do documento original; sumário, que visa trazer os principais tópicos do documento original podendo ser encontrado no início ou no final do livro, por exemplo; e a extração é compreendida como porções selecionadas dos documentos.

**b) Metadados como Componente Representativo da Informação**

Dentre as técnicas para organizar/classificar a informação para recuperação e intercâmbio, encontram-se os metadados. Apesar de possuir diferentes definições, metadados são, muitas vezes, conceituados como dados sobre dados. Determinados dados são gerados para descrever outros dados. Ou seja, são dados que têm o intuito de representar de forma qualitativa outros dados para uma recuperação eficiente por parte de um usuário, seja ele em potencial ou não. Essa representação contribui para o acesso à informação desejada. Entretanto, vale ressaltar que essa não é sua principal função. Ela possibilita desde a identificação informacional, intercâmbio, até seu armazenamento. Nesse viés, Lourenço (2005, p. 58) define o termo metadados como sendo:

> Um "identificador" que descreve, contextualiza, administra e recupera um objeto digital, além de relacioná-lo a outros objetos digitais semelhantes ou relacionados a ele dentro de uma biblioteca digital ou no ambiente da web como um todo. É representado pelas tags das linguagens de marcação, pelos hiperlinks que ligam os objetos digitais entre si e até mesmo pelas URLs que identificam os sites da web.

Assim, os metadados são elementos representativos da informação no ambiente digital que visam proporcionar o acesso a uma grande variedade de informações e


(cc) BY-NC-ND

multiplicidade de formatos. No âmbito arquivístico é possível comparar os metadados, respeitando, obviamente suas particularidades, com os instrumentos de pesquisa arquivísticos nos quais tende-se a descrever o arquivo todo, parte da documentação ou o documento em si para facilitar à recuperação informacional. Os tipos básicos de instrumentos de pesquisa podem ser guias, inventários, catálogos e repertórios (PAES, 2004).

Sem dúvidas, para que haja uma competência tanto na elaboração quanto na aplicação dos metadados e dos instrumentos de pesquisa no âmbito dos arquivos, é preciso seguir parâmetros. No campo da Arquivologia esses critérios podem ser encontrados na utilização de *thesaurus*, dos próprios vocabulários controlados e até na forma como os documentos ou o acervo de um modo geral são descritos. No caso dos metadados, eles seguem os denominados "padrões", dos quais pode-se listar:

- *Dublin Core* (*DC- Dublin Core Metadata Element Set*);
- *Machine Readable Card* (*MARC*);
- *Government Information Locator Sevice* (*GILS*);
- *Spatial Archive and Interchange Format* (*SAIF*);
- *Federal data Geographic Committee* (*FGDC*);
- *Consortium for the interchange of Museum Information* (*CIMI*);
- *Meta Content Format* (*MCF*);
- *Text Enconding Iniciative* (*TEI*);
- *Eletronic Archive description* (*EAD*);
- *Resource Description Framework* (*RDF*);
- Padrão de Metadados do Governo Eletrônico (e-PMG).

Dentre esses, os principais padrões utilizados em bibliotecas e demais instituições informacionais são o *Dublin Core* e o *Marc* 21. O Dublin Core é constituído por 15 elementos básicos e 3 componentes adicionais. Os elementos básicos são: Título (*title*), Criador (*creator*), Assunto (*subject*), Descrição (*description*), Publicador (*publischer*), Colaborador (*contributor*), Data (*date*), Tipo (*type*), Formato (*format*), Identificador (*identifier*), Fonte (*source*), Idioma (*language*), Relação (*relation*), Cobertura (*coverage*) (THE DUBLIN CORE METADATA INITIATIVE, 2014). É composto por quatro níveis de interoperabilidade. São eles:


(cc) BY-NC-ND

- **Nível 1 (definições de termos compartilhados)** - vocabulários compartilhados definidos por meio da linguagem natural;
- **Nível 2 (interoperabilidade semântica formal)** - vocabulários compartilhados com base na semântica formal;
- **Nível 3 (Descrição Definir interoperabilidade sintática) -** vocabulários formais compartilhados no registro de troca;
- **Nível 4 (Descrição Definir interoperabilidade *Profile*)** - vocabulários e restrições formais compartilhados nos registros.

O *MARC*21 é um dos padrões mais antigos, criado em 1960 pela *Library of Congress* (LOURENÇO, 2005). É composto por três elementos: estrutura, indicação do conteúdo e conteúdo propriamente dito. A estrutura do registro é uma implementação das normas internacionais ANSI Z39.2 e ISO 2709. No que diz respeito às designações de conteúdo, esses são códigos e convenções estabelecidos para identificar e caracterizar os dados dentro do registro, permitindo, posteriormente, sua manipulação. Os conteúdos dos dados são normalmente estabelecidos por padrões externos ao formato, tais como o *International Standard Bibliographic Description* (*ISBD*), *Library of Congress Subject Headings* (*LCSH*) e *Anglo-American Cataloguing Rules* (*AACR2*).

Já o *Eletronic Archive description* (*EAD*) utiliza campos descritivos propostos pela ISAD G (Norma Geral Internacional de Descrição Arquivística) para descrever os documentos encontrados em meio eletrônico. A *EAD* surgiu por volta de 1993, na Universidade da Califórnia, Berkeley, sendo o foco do *EAD* a constituição de elementos descritivos semelhantes ao padrão *MARC*. Um dos exemplos encontrados no qual se utiliza o *EAD* é a Biblioteca Digital de *Kentucky*.

Outro componente utilizado para a representação da informação é a Norma Brasileira de Descrição Arquivística (NOBRADE) sendo essa, a adaptação de duas normas internacionais: a *ISAD* (*G*) e a *ISAAR* (*CPF*). Nesse sentido, a NOBRADE apresenta pressupostos básicos aos fundos e a descrição multinível: descrição do geral para o específico; informação relevante para o nível de descrição; relação entre descrições; não repetição da informação. A NOBRADE (CONSELHO NACIONAL DE ARQUIVOS, 2006) compreende 28 elementos de descrição, constituindo-se em oito áreas, sendo elas: Área


(cc) BY-NC-ND

de identificação; Área de contextualização; Área de conteúdo e estrutura; Área de condições de acesso e uso; Área de fontes relacionadas; Área de notas; Área de controle da descrição; Área de pontos de acesso e descrição de assuntos.

Diante disso, percebe-se que, por meio dos elementos nos quais se queiram representar, será identificado o padrão mais adequado. O *Dublin Core*, por exemplo, compreende-se em um padrão sintético sendo útil para acervo com uniformidade de conteúdos, formatos e tipologias. No caso, a NOBRADE e a *EAD* são compostas por campos provindos das normas internacionais arquivísticas e, por esta razão, abarcam elementos nem sempre tão relevantes ao usuário, como é o caso do estado de conservação documental. Por fim, o Marc 21, sendo este formato de metadados recomendável quando o acervo a ser representado engloba diversos formatos, gêneros e tipologias documentais haja vista sua completude representativa.

## 5 REPRESENTAÇÃO DE INFORMAÇÃO E SUAS QUESTÕES PRÁTICAS

Este contexto procurou deixar evidente o auxílio proveniente da representação para a eficiência na recuperação informacional. Com efeito, as contribuições proporcionadas pelos metadados na representação do mundo real e os banco de dados como recursos essenciais na disponibilização e ampliação da qualidade das buscas informacionais. Portanto, um projeto de banco de dados deve descrever a realidade de forma abstrata para que suas informações possam ser representadas fisicamente nas bases de dados.

O modelo de entidade relacionamento (MER) tem por base a descrição de algo do mundo real realizada através de um conjunto de objetos chamados "entidades" e pelo conjunto de "relacionamentos" entre esses objetos. Entidade é uma coisa (física ou conceitual) do mundo real. Ela pode ser forte, se não depender de outra entidade para existir ou fraca, se sua existência no modelo estiver condicionada à presença de outra entidade. (ELMASRI; NAVATHE, 2005, p. 39).

Além disso, as entidades possuem "atributos", que são propriedades particulares que as descrevem. Os valores assumidos por cada atributo descrevem as entidades e formam o conjunto de dados armazenados na base de dados (ELMASRI; NAVATHE,


(cc) BY-NC-ND

2005, p. 39). Esses valores são as representações da informação fornecendo pontos de acesso ao texto, ao documento, a informação completa.

Pode-se representar, por exemplo, a realidade de um cliente em um *site* de comércio *online* ou um acervo digital de artigos científicos.

Para criar um cenário prático, foi elaborado para este trabalho de reflexão, um exemplo de como a representação da informação é necessária à construção de uma base de dados e por consequência a recuperação de informação.

Desta forma o Quadro 1 demonstra a representação de um acervo digital de artigos científicos por meio de uma ficha catalográfica elaborada com base no modelo Dublin Core.

**Quadro1**: Ex. de artigo representado em ficha catalográfica - Dublin Core.

| Elementos Descritivos | Descrição |
|---|---|
| **Título** | Um Estudo de Caso na Base de Dados da Empresa ABCX |
| **Autor** | Jackson Brasileiro Silva; Maria Rita Braz |
| **Assunto** | Representação da Informação; Recuperação de Informação; Ciência da Informação; Banco de Dados Bibliográfico; Metadados. |
| **Descrição** | Realiza uma comparação dos métodos de representação e recuperação da informação nas bases de dados bibliográficas da empresa ABCX. |
| **Editora** | Periódicos Imaginários |
| **Colaborador** | Maria e João |
| **Data** | 19 de Janeiro de 2011 |
| **Tipo** | Artigo Científico |
| **Formato** | PDF |
| **Identificador** | www.periodicosimaginarios.com.br/ufpb/reperecubiblio1918.pdf |
| **Fonte** | (100_5641) |
| **Idioma** | Português |
| **Itens Relacionados** | www.periodicosemrepresentacao.com.br; www.periodicosemrec.com.br; www.periodicosreprec.com. |
| **Cobertura** | UFPB; João Pessoa; Paraíba; Brasil (2011) |
| **Direitos** | Acesso livre para divulgação. Favor dar os devidos créditos. A submissão de artigo implica compromisso de não submetê-lo a outro periódico e autoriza, caso aprovado, a sua publicação. A responsabilidade por eventuais plágios nos artigos publicados é de responsabilidade do(s) autor(es). |

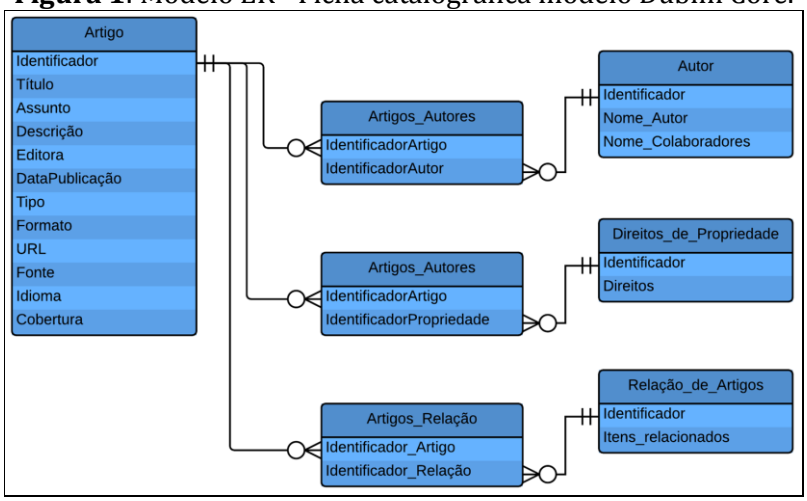
**Fonte**: Elaborado pelos autores (2015).

A partir dessa representação, é possível à construção de um modelo de entidade relacionamento servindo de alicerce para a criação de uma base de dados. Os elementos descritivos serão, por conseguinte, os atributos e as suas descrições os valores assumidos pelos atributos. É por meio desses valores que se chega ao texto completo.


(cc) BY-NC-ND

Observe na Figura 1 o modelo conceitual representativo baseada na ficha catalográfica do Quadro 1.

**Figura 1**: Modelo ER - Ficha catalográfica modelo Dublin Core.




**Fonte:** Elaborado pelos autores (2015).

Pode-se afirmar que ao representar de forma errônea os dados em algum dos campos explicitados, provoca a irrecuperabilidade da informação desejada. É por este fator que os campos representativos são delimitados de forma concisa para que no ato de representar, o indexador– a pessoa responsável por desenhar o modelo ER dos campos do documento – não fique com dúvidas.

Um ponto crítico a ser observado é que um mesmo atributo não pode receber valores diferentes. Veja na Figura 1; na tabela "Autor", o atributo "Nome_Autor" pode receber o nome do autor (Jackson), o sobrenome (Silva) ou o nome completo (Jackson Brasileiro Silva). Dessa forma, se apenas o nome desse autor tivesse sido cadastrado no sistema, nenhum documento dele seria passível de ser recuperado se a palavra "Silva" fosse usada como expressão de busca.

Outro ponto retratado na Figura 1 refere-se à duplicidade de conteúdos em um mesmo atributo. Na tabela "Artigo", o atributo "Cobertura" designa a descrição tanto da localização espacial quanto da duração temporal. Para que uma informação possa ser representada no banco de dados, o atributo tem que ser atômico, ou seja, ele só pode possuir um valor ou designar uma característica.

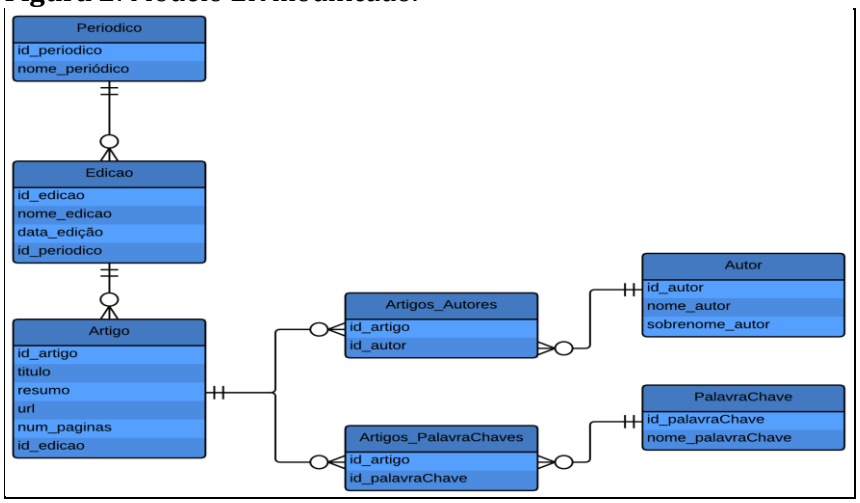
Caso um sistema de recuperação de artigos científicos seja projetado tendo como suporte o modelo de base de dados exemplificado na Figura 1, o usuário poderia ter


(cc) BY-NC-ND

dificuldade, por exemplo, de recuperar artigos baseado sem atributos como "palavra-chave" ou "resumo". Isso se dá pelo fato dos campos "Assunto" e "Descrição" não representarem exclusivamente os atributos "palavra-chave" e "resumo". De acordo com o documento oficial do The Dublin Core Metadata Initiative (2015), o elemento "Assunto" deve ser expresso por palavras-chave, frases, ou códigos de classificação que descrevam o conteúdo do recurso. Já o elemento "Descrição" pode ser representado por um resumo, um índice, uma referência ou uma descrição textual. Outro ponto que merece destaque é o fato de alguns desses atributos serem irrelevantes no que tange a recuperação de artigos científicos digitais. Os atributos "Formato", "Idioma" e "Direitos", por exemplo, não são usuais para termos de busca.

Com base nestas reflexões, é possível melhorar esse exemplo de modelo de base de dados, sendo necessárias algumas modificações: a adição de alguns atributos, tais como resumo, *abstract* e palavras-chave; o desmembramento dos atributos "Nome_Autor", sendo este substituído por "Nome", "Sobrenome" e "Nome_Completo"; e a remoção dos atributos "Direitos_Autorais" e "Editora". A partir dessas alterações, é possível aperfeiçoar o modelo de entidade relacionamento com vistas a facilitar a recuperação da informação armazenada, conforme apresentado na figura 2.

**Figura 2**: Modelo ER modificado.




**Fonte:** Elaborado pelos autores (2015).

A partir desse MER modificado (Figura 2) os artigos são vinculados às edições e que essas edições pertencem a periódicos. Assim, pelo menos um periódico necessita ser


(cc) BY-NC-ND

cadastrado e, a partir de então, edições podem ser a ele vinculadas. Um periódico pode ter várias edições, mas uma edição pode pertencer a apenas um único periódico. Dessa forma, apenas depois de realizado os cadastros dos periódicos e suas respectivas edições, as referências dos artigos podem ser armazenadas.

Com o artigo cadastrado, autores e palavras-chave podem ser vinculados. Um artigo pode ter sido escrito por diversos autores e um autor por ter escrito mais de um artigo. Essa mesma lógica é usada para o atributo palavra-chave. Assim, a descrição de algo do mundo real, nesse caso, artigos científicos, está de acordo com a percepção do usuário, evitando possíveis erros de busca provenientes de uma tarefa de inclusão de dados no sistema mal executada.

A partir de uma base de dados desenvolvida seguindo esse modelo de entidade relacionamento (Figura 2), os usuários poderão recuperar artigos científicos em formato digital por "Título", "Palavra-chave", "Resumo" e "Autor" – nome ou sobrenome.

Essas modificações, entretanto, não devem ser tomadas como uma atividade finda, pois é por meio desses atributos e seus valores que se chega ao texto completo de um artigo científico, deve-se manter sempre o foco na necessidade informacional dos usuários. Esse exemplo prático indica que quão melhor for a representação da informação armazenada nas bases de dados, melhor e mais relevante poderá ser o processo de recuperação da informação. Esta alteração demostra a necessidade de um profissional da informação no processo de desenvolvimento e elaboração do modelo de dados, pois este pode indicar os campos e a forma como o usuário irá interagir com o sistema durante o processo de recuperação da informação.

## 6 CONSIDERAÇÕES FINAIS

A informação é um produto de extremo valor e que caminha lado a lado com a evolução e o desenvolvimento das sociedades. A partir do referencial desenvolvido, demonstrou-se que a informação adquiriu novas características a partir de sua interseção com a tecnologia e a Internet. A representação da informação na *Web* é um desafio, visto que há um aumento constante no volume de informações produzidas,


(cc) BY-NC-ND

atualizadas e compartilhadas quase que instantaneamente, dificultando ainda mais a sua recuperação.

Assim, fica evidente a importância de refletir sobre os desafios encontrados pelos profissionais da informação, assumindo seus papeis de mediadores da informação, no que cerne a representação e recuperação informacional no ambiente tecnológico, pois este está em constante evolução.

Percebe-se, que a ação de interferência com o propósito de disseminação, realizada pelos profissionais da CI, deve ter ao seu lado a representação e recuperação da informação como processos fundamentais. Pois, uma informação deve ser representada para poder ser recuperada e, então, satisfazer, plena ou parcialmente, a necessidade informacional dos usuários. Nesse contexto, julga-se imprescindível, compreender o funcionamento de todo o processo representativo para uma posterior recuperação de informação. Infere-se que as atividades representativas devem ser construídas por meio da análise de perfil de usuário de uma dada fonte de informação.

No exemplo exposto neste texto, percebeu-se que quando as bases de dados seguindo os padrões de metadados foram direcionadas a um perfil de recuperação de determinado grupo de usuários, o modelo elaborado apresentou um maior grau de campos representativos, o que potencializa o êxito na recuperação da informação. Demonstrou que uma representação errônea dos campos no modelo de banco de dados, pode provocar a irrecuperabilidade das informações contidas em documentos registrados nas bases de dados.

O que sugere a necessidade do profissional da informação participar do processo de desenvolvimento dos modelos de representação para bases de dados e não delegar essa tarefa aos profissionais de TI.

## ANALYSIS OF REPRESENTATION OF INFORMATION IN MODELS ENTITY RELATIONSHIP BASED ON METADATA


### ABSTRACT

In the current context, information, linked to technology, have become essential to the daily activities of human beings, for contemporary society needs elements to empower the organization, dissemination and preservation of information, regardless of the support. This study intended to consider on the representation of information for information retrieval

process in the virtual environment. Based on the fields representation of the Dublin Core metadata model and the development of Entity-Relationship models for database. The relevance of metadata for the representation of information is highlighted when it is applied to the databases reach. Asserts that the information acquired new features from its intersection with technology and the Internet, which are agents that change the informational flow. For analysis and reflection it was elaborated an example of how is the interaction representation of information applying the Dublin Core metadata model in a cataloguing data as how would possible alternatives for a database modeling. As a result, demonstrates that a misrepresentation of the fields in the database model, could result in irrecoverable of the information contained in documents filed in databases and emphasizes the need for the information professional participate in the development process of representation models for bases data.

**Keywords**: Information Representation. Information Retrieval. Database.

**REFERÊNCIAS**

ANDRADE, Maria Eugênia Albino; OLIVEIRA, Marlene de Oliveira. A Ciência da Informação no Brasil. In: OLIVEIRA, Marlene de. **Ciência da Informação e Biblioteconomia:** novos conteúdos e espaços de atuação. Belo Horizonte: Editora UFMG, 2005. p. 45-60.

BLAIR, David C. **Language and representation in information retrieval**. Elsevier Science Publishers; Distributors for the US and Canada, Elsevier Science Pub. Co., 1990.

BORKO, Harold. Information science: what is it?. **American documentation**, v. 19, n. 1, p. 3-5, 1968.

BUSH, Vannevar. As we may think. In: **Evolution of an information society.** Londres: 1945.

CAPURRO, Rafael ; HJØRLAND, Birger. The Concept of Information. **Annual Review of Information Science and Technology**. Ed. B. Cronin, v. 37, Ch. 8, p. 343-411, 2003.

CARDOSO, Olinda Nogueira Paes. Recuperação de Informação. **INFOCOMP Journal of Computer Science**, v. 2, n. 1, p. 33-38, 2004.

CASTELLS, Manuel. A era da informação: economia, sociedade e cultura. O poder da identidade. **São Paulo: Paz e Terra**, v. 2, p. 24, 1999.

CIANCONI, Regina. Banco de Dados de acesso público. **Ciência da informação.** Brasília, v. 16, n. J, p. 53-59, jan./jun. 1987.

CHU, Heting. **Information Representation and Retrieval in the Digital Age**. New Jersey: ASIS& T, 2007.


(cc) BY-NC-ND

**DUBLIN CORE METADATA INITIATIVE**. Disponível em: <http://dublincore.org>. Acesso em 02 Dez. 2014.

CONSELHO NACIONAL DE ARQUIVOS (Brasil). **NOBRADE**: Norma Brasileira de Descrição Arquivística. Rio de Janeiro: Conselho Nacional de Arquivos, 2006.

ELMASRI Ramez; NAVATHE Shamkant B. **Sistemas de Banco de Dados**. 4. ed. São Paulo: Pearson, 2005.

FERNEDA, E. **Recuperação da informação**: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação. 2003. 137 f. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes da Universidade de São Paulo, São Paulo, 2003. Disponível em: < http://www.teses.usp.br/teses/disponiveis/27/27143/tde-15032004-130230/pt-br.php>. Acesso em: 19 fev. 2015.

FREIRE, Gustavo Henrique de Araújo; FREIRE, Isa Maria. **Introdução à Ciência da Informação**. João Pessoa: Editora Universitária da UFPB, 2009.
GONZÁLEZ DE GÓMEZ, Maria Nélida. Novas fronteiras tecnológicas das ações de informação: questões e abordagens. **Ci. Inf**, v. 33, n. 1, p. 55-67, 2004.

HEEMANN, V M. Avaliação **ergonômica de interfaces de bases de dados por meio de Checklist especializado**. 1997. Dissertação (Mestrado em Engenharia da Produção) – Universidade Federal de Santa Catarina, Florianópolis, 1997.
HJØRLAND, Birger. Domain analysis in information science: eleven approaches-traditional as well as innovative. **Journal of documentation**, v. 58, n. 4, p. 422-462, 2002.

JACOB, Elin K.; SHAW, Debora. Sociocognitive Perspectives on Representation. **Annual Review of Information Science and Technology (ARIST)**, v. 33, p. 131-85, 1998.

KOBASHI, Nair Yumiko; FERNANDES, Joliza Chagas. Pragmática lingüística e organização da informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 2009, João Pessoa. **Anais**... João Pessoa: PPGCI/UFPB, 2009.

LANCASTER, F. W. **Indexação e Resumos**: teoria e prática. Brasília: Briquet de Lemos, 2004.

LE COADIC, Yves–François. **A Ciência da Informação**. Tradução Maria Yêda F.S. de Filgueiras Gomes. França. Briquet de Lemos. 2004.

LÉVY, Pierre. **Cibercultura**. Tradução Carlos Irineu da Costa. São Paulo: Editora 34, 1999.

LOPES, Ilza Leite. Estratégia de busca na recuperação da informação: revisão da literatura. **Ciência da Informação**, v. 31, n. 2, p. 60-71, 2002. Disponível em:


(cc) BY-NC-ND

<http://revista.ibict.br/index.php/ciinf/article/view/158>. Acesso em: 05 nov. 2014.

LOURENÇO, Cíntia de Azevedo. **Análise do Padrão Brasileiro de Metadados de Teses e Dissertações segundo o Modelo Entidade-Relacionamento**. 2005. 165 f. Tese. (Doutorado em Ciência da Informação) – Universidade Federal de Minas Gerais. Minas Gerais. Belo Horizonte. 2005.

MANINI, Míriam Paula. **Análise documentária de fotografias: um referencial de leitura de imagens fotográficas para fins documentários**. 2002. Tese de Doutorado. Universidade de São Paulo. Disponível em: <http://www.teses.usp.br/teses/disponiveis/27/27143/tde-23032007-111516/>. Acesso em: 28 Mar. 2015.

MARC STANDARDS. Disponível em: <http://www.loc.gov/marc/>. Acesso em: 02 Dez. 2013.

MATTOS, Alessandro Nicoli de. **Informação é prata, compreensão é ouro** - um guia para todos sobre como produzir e consumir informação na Era da Compreensão. [S.l.: s. n.], 2010.

MOOERS, C. Zatocodingapplied to mechanical organization of knowledge. **American Documentation**, v. 2, n. 1, p.20-32, 1951.

NOVELLINO, María Salet Ferreira. Instrumentos e metodologias de representação da informação. **Informação & Informação**, Londrina, v. 1, n. 2 p. 37-45, jul.-dez. 1996.

_____, Maria Salet Ferreira. A linguagem como meio de representação ou de comunicação da informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.3, n.2, p.137-146, jul./dez. 1998. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/602/371>. Acesso em 29 nov. 2014.

PAES, Marilena Leite. **Arquivo**: teoria e prática. Rio de Janeiro: Editora FGV, 2004.

ROWLEY, Jennifer. **Informática para bibliotecas**. Tradução de Antônio Agenor Briquet de Lemos. Brasília: Briquet de lemos/Livros, 1994.

SARACEVIC, Tefko. Ciência da Informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.

SILVA, Jonathas Luiz Carvalho; FREIRE, Gustavo Henrique de Araújo. Um Olhar sobre a Origem da Ciência da Informação: Indícios Embrionários para sua Caracterização Identitária. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, v. 17, n. 33, p. 1-29, jan./abr., 2012.

SOUZA, Renato Rocha. Sistemas de Recuperação de Informações e Mecanismos de Busca na Web: panorama atual e tendências. **Perspectivas da Ciência da Informação**. Belo


(cc) BY-NC-ND

Horizonte, v. 11, n. 2, p. 161-173, maio/ago. 2006. Disponível em: <http://www.scielo.br/pdf/pci/v11n2/v11n2a02.pdf>. Acesso em: 06 nov. 2013.

VELOSO, Renato. **Tecnologia da informação e da comunicação:** desafios e perspectivas. São Paulo: Saraiva, 2011.

ZANINELLI, T. B.; CATARINO, M. E. Recuperação de informação no banco de dados da Embrapa soja: estudo da inf. **inf. inf.**, londrina, v. 9, n. 1/ 2, jan./ dez. 2004.


(cc) BY-NC-ND