

ESTRUTURAÇÃO DE HIPERTEXTOS: PROPOSTA DE CONVERSÃO DE TEXTOS

Elaine Diamantino Oliveira¹
Benildes Coura Moreira dos Santos Maculan²
Ivo Pierozzi Júnior³

RESUMO: Trata-se de uma pesquisa em andamento, que se apoia em revisão bibliográfica nas áreas da Ciência da Informação, Ciência da Computação e Linguística. Tem por objetivo criar um instrumento de representação da informação no formato de um hipertexto navegacional sobre o domínio da agropecuária brasileira. Propõe-se sistematizar uma metodologia para conversão de texto linear em hipertexto que possa auxiliar o processo de construção de hipertextos semanticamente enriquecidos. Aborda conceitos relacionados à produção hipertextual, verificando a existência de metodologias de conversão de hipertexto, seja de forma automática ou intelectual. Parte-se do pressuposto de que os estudos que abarcam metodologias sobre esse processo de conversão não deixam explícitos todos os procedimentos que são necessários para organização conceitual de hipertextos de forma consistente.

Palavras-chave: Hipertextos. Conversão de textos em hipertextos. Sistemas de autoria.

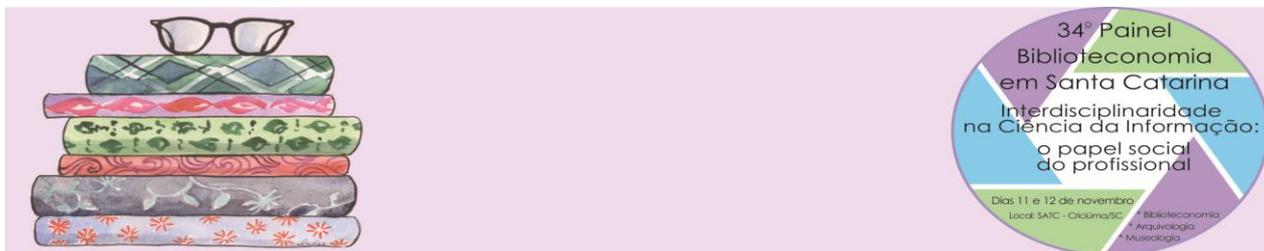
1 INTRODUÇÃO

A utilização da internet como espaço de leitura e produção de texto é consequência de um cenário de intensa transformação nas formas não apenas de organizar, mas também de acessar e disponibilizar conteúdos, no qual a cultura do digital é uma realidade inquestionável. A internet, em sua face mais popular, a *Web*, está cada vez mais presente no cotidiano das pessoas. Diante desse contexto, torna-se improvável que o volume de

¹ Mestranda em Ciência da Informação pela Universidade Federal de Minas Gerais (2016) e especialista em Arquitetura e Organização da Informação, também pela UFMG. Atualmente é bibliotecária da Escola de Ciência da Informação, UFMG. E-mail: elained@eci.ufmg.br

² Doutora e Mestre em Ciência da Informação pelo Programa de Pós-Graduação em Ciência da Informação, na Escola de Ciência da Informação (ECI) da Universidade Federal de Minas Gerais (UFMG), e forma em Bibliotecária pela mesma Universidade. Atualmente, Professora Adjunto no Departamento de Organização e Tratamento da Informação, na ECI/UFMG. Pesquisadora no Grupo de Pesquisa MHTX - Modelagem Conceitual para Organização Hipertextual de Documentos e no Grupo de Pesquisa RECRI - Representação do Conhecimento e Recuperação da Informação e no Programa de Pós-Graduação em Gestão & Organização do Conhecimento (PPGGOC/ECI/UFMG). E-mail: benildes@gmail.com

³ Graduação em Ciências Biológicas (1981), mestrado (1985) e doutorado (1989) em Ecologia, pela Universidade Estadual de Campinas (UNICAMP). Atualmente é Pesquisador A da Embrapa Informática Agropecuária. Ocupou cargos gerenciais em P&D e C&T por mais de uma década, na Embrapa Monitoramento por Satélite. Desenvolveu pesquisas na área de Acridologia Operacional e atualmente trabalha com os seguintes temas: Terminologia Agropecuária, Organização e Representação do Conhecimento e da Informação, Arquitetura da Informação, Sistemas de Informação e Internet. E-mail: ivo.pierozzi@embrapa.br



informações produzidas no ambiente *Web* seja manipulado apenas de forma intelectual. Porém, em contextos de especialidade, aplicações somente automáticas de tratamento de dados pode não ser a melhor solução. O tratamento de dados e sua disponibilização em formato eletrônico superam as condições de uso impostas pelo formato impresso, trazendo mais facilidade e rapidez no acesso a partes de um mesmo documento ou de diferentes documentos.

Intensamente ligada ao desenvolvimento da *Web*, a tecnologia de hipertexto causou forte impacto no modo de interação e manipulação da informação. A dinamicidade na produção, divulgação e atualização das informações, além da facilidade de navegação, são algumas entre as muitas vantagens do hipertexto. Assim, pode-se inserir a temática do hipertexto, enquanto um instrumento de representação do conhecimento, no inventário das novas formas de organizar a informação contida em documentos textuais.

No novo contexto de produção, organização e recuperação de objetos digitais, as metas de trabalho não se restringem à criação de representações simbólicas dos documentos constantes de um acervo, mas compreendem a criação de novas formas de escrita para os hipertextos (ALVARENGA, 2003, p. 36).

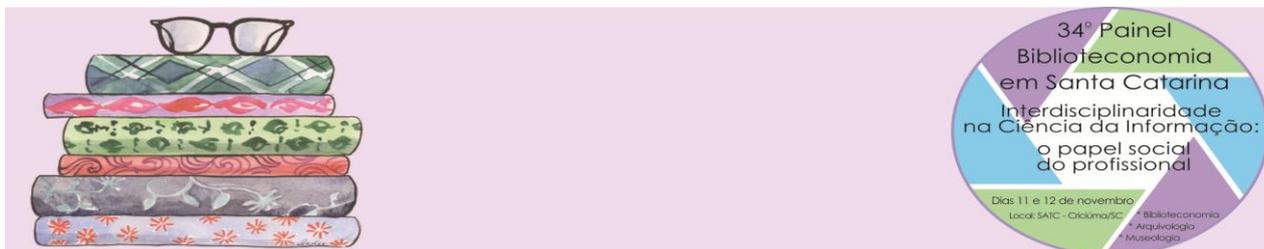
Entretanto, a ideia básica do hipertexto não é nova, tendo em vista que é uma forma de inter-relacionar frações de informação, inclusive em meio impresso, o que propicia uma efetiva interação do leitor com o conteúdo textual. É o caso, por exemplo, das citações, que fazem referência a outros textos, das notas de rodapé, onde constam trechos de informação que esclarecem algum ponto que não cabe discutir no corpo do texto, dos índices ao final dos livros ou das referências bibliográficas. Pode-se considerar que todos esses elementos são *links* que levam o leitor a optar por uma leitura contínua e linear, desconsiderando as informações complementares, ou a escolher interromper a leitura, tornando-a não linear, e consultar os novos elementos de informação.

Nessa perspectiva, o hipertexto vem se configurando, desde os primórdios de sua criação, e provocou um questionamento sobre os tradicionais métodos de escrita em papel, vislumbrando-se o seu potencial para a escrita na *Web* (MONTEIRO, 2000). No meio digital, o hipertexto permite passar rapidamente de um *link* a outro, assim como acoplar uma grande quantidade de informações, dando ao leitor uma condição bastante ativa. Assim, o hipertexto facilita a interação do leitor com o texto e suas hiperligações, tornando o texto um sistema semântico mais amplo que o seu próprio conteúdo isolado.

Apesar das vantagens proporcionadas pelos hipertextos, a sua produção ainda se configura como um desafio. Muitas vezes, os hipertextos são criados sem qualquer planejamento, ou seja, não há preocupação com a estrutura semântica do documento ou com as necessidades informacionais de seus possíveis usuários.

A partir da análise da literatura, percebe-se que as metodologias recuperadas, que direcionam o processo de estruturação de hipertextos, muitas vezes já começam numa etapa avançada de modelagem, e não contemplam a fase inicial de organização conceitual. Apesar do uso recorrente de hipertextos na *Web*, percebe-se que ele ainda carece de uma formalização metodológica e o desenvolvimento de ferramentas mais efetivas para a sua elaboração.

Ao longo dos anos, e com o desenvolvimento de ferramentas computacionais, foram propostos diversos modelos de *design* e sistemas para construção de hipertextos. No entanto,



as pesquisas ainda apontam para a limitação dessas ferramentas, que estão aquém do potencial do hipertexto.

A reflexão sobre essas considerações suscitou as seguintes questões: como um hipertexto deve ser construído? Quais recursos/elementos podem, ou devem, ser acrescentados para facilitar a navegação do usuário final? Quais são os procedimentos necessários para transformar um texto linear em uma estrutura hipertextual semântica? Como organizar (modelar) uma área do conhecimento para que explicita as relações semânticas existentes entre os diferentes *links*? É possível realizar a conversão automática de texto linear em hipertexto, sem prejuízo semântico do conteúdo?

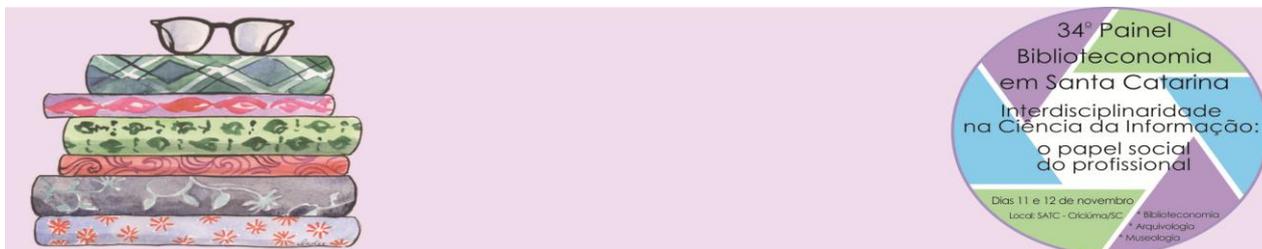
Considera-se que a relevância deste estudo, no âmbito da Ciência da Informação (CI), pauta-se na premissa de que os fundamentos teórico-metodológicos advindos do campo da organização do conhecimento são competências capazes de dar suporte à criação de estruturas hipertextuais de forma lógica e coerente. Com isso, evita-se um dos maiores problemas do hipertexto: a desorientação do usuário na navegação. Por outro lado, na contemporaneidade, nota-se a incipiência de trabalhos, com o aporte da CI, que investiguem o hipertexto, tanto no momento da sua produção quanto do seu uso. Acredita-se, também, que estudos com esse viés auxiliam na consolidação da CI como uma ciência interdisciplinar, uma vez que analisa o objeto com contribuições teóricas da Ciência da Computação e da Linguística.

Este trabalho tem por propósito divulgar uma pesquisa de mestrado proposta ao Programa de Pós-graduação em Ciência da Informação, da Escola de Ciência da Informação (ECI) da Universidade Federal de Minas Gerais (PPGCI/ECI-UFMG). Tem como objetivo criar um instrumento de representação da informação, a partir da conversão de um livro (leitura linear) em um hipertexto navegacional, de modo a facilitar, ao usuário final, o entendimento do conteúdo semântico contido no documento. No decurso inicial desta pesquisa, serão abordados os elementos históricos e os conceitos relacionados à produção de hipertextos, verificando, também, a existência de metodologias de conversão do texto linear em hipertexto, seja de forma automática ou intelectual. O estudo parte do pressuposto de que as metodologias existentes não apresentam os procedimentos completos para a estruturação do hipertexto, desde a modelagem conceitual do objeto em questão, e que estão pulverizadas nas áreas de Ciência da Computação e Linguística, com pouca ocorrência de estudos na Ciência da Informação.

2 METODOLOGIA

O estudo, de natureza exploratória e experimental, se apoia em revisão bibliográfica nas áreas da Ciência da Informação, Ciência da Computação e Linguística. Como domínio de aplicação, foi escolhida a agropecuária brasileira, uma vez que esta pesquisa está inserida em um convênio formalizado entre a Empresa Brasileira de Pesquisa Agropecuária (Embrapa), sediada em Campinas, e a UFMG. O recorte temático é a Intensificação Agropecuária, processo descrito no livro de Boserup (1965). Nesse livro, a autora descreve sua teoria cujo teor determina que o aumento da produção agrícola, em uma mesma quantidade de área, ou a manutenção da mesma produção, em uma área menor, aparece quando ocorre a Intensificação Agropecuária. O conteúdo desse livro será o objeto a ser convertido em um hipertexto.

O levantamento bibliográfico inicial está sendo realizado visando recuperar conteúdos sobre hipertextos (histórico-conceituais) e as metodologias para a sua construção. A busca foi



efetuada em fontes de informação científica: (1) gerais: Portal de Periódicos da Capes, Biblioteca Digital de Teses e Dissertações do Instituto Brasileiro de Informação em Ciência e Tecnologia (BDTD-IBICT), Catálogo do Sistema de Bibliotecas da UFMG e da Rede Pergamum, Google Acadêmico, *Annual Reviews*; (2) especializadas: *Library and Information Science Abstracts* (LISA), *Library and Information Science Technology Abstracts* (LISTA), *Information Science Technology Abstracts* (ISTA), *ACM Digital Library*, *Educational Resources Information Center* (ERIC), *CiteSeer*, Base Peri e Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (BRAPCI). Foram formuladas estratégias de buscas com as seguintes expressões (também traduzidas para os idiomas: inglês e espanhol): hipertextos, hiperdocumentos, sistemas de hipertextos, conversão de hipertextos, metodologias para hipertextos, conversão automática de hipertextos, estruturação de hipertextos, construção de hipertextos, modelagem de hipertextos, design de hipertextos, métodos para hipertextos; técnicas para hipertextos, navegação hipertextual.

A pesquisa bibliográfica e teórica irá permear todas as etapas desta pesquisa, dando respaldo ao percurso que será delineado. De modo geral, ao final desta pesquisa, pretendem-se os seguintes procedimentos para a conversão do texto: (a) leitura do documento (livro de Boserup, 1965); (b) seleção dos conceitos relevantes para a representação do conteúdo do livro; (c) organização dos conceitos; (d) determinação das relações semânticas entre conceitos; (e) transformação do livro em formato digital: hipertexto (informação hipertextual).

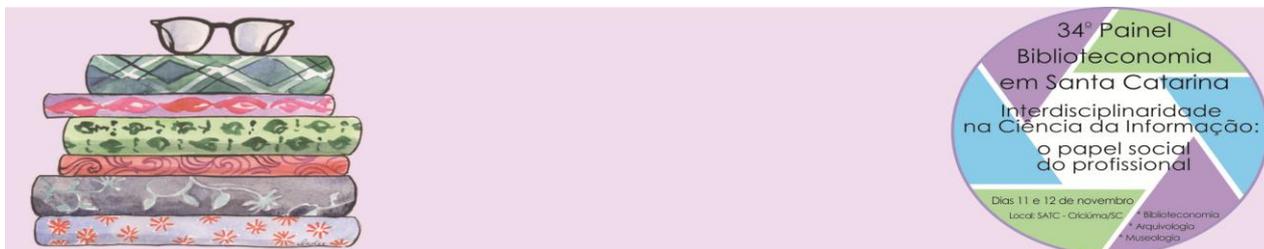
3 RESULTADOS PARCIAIS

Os estudos realizados até o momento já permitiram alguns resultados, sobretudo a visão histórica do hipertexto e algumas metodologias que já foram propostas para a sua construção. Esses dados estão descritos a seguir.

3.1 HIPERTEXTOS – PERSPECTIVA HISTÓRICA

A passagem da cultura oral para a escrita foi proporcionada pelo surgimento da imprensa, que acelerou a reprodução de documentos, viabilizando o acesso ao conhecimento. A passagem da cultura escrita para a cultura eletrônica também é um marco substancial na produção de conhecimento, facilitada pelos avanços das tecnologias de informação e comunicação (TIC's). Tanto numa fase quanto na outra, a passagem não ocorreu de forma simples.

Harnad (1991) destaca o fato de a linguagem falada ser mais compreensível ao raciocínio humano, enquanto a escrita é lenta e não consegue acompanhar esse raciocínio. A linguagem na forma escrita já vem condicionada a uma lógica de encadeamento de ideias, de forma rígida, ao definir uma escrita em parágrafos. Assim, ao propor uma forma linear para externalizar o pensamento, a escrita condicionou emissores e receptores a um novo modelo de comunicação, mais formal. Com as inovações advindas das tecnologias informáticas, foi possível a criação de hipertextos, que teve como consequência amenizar essa rigidez, através da associação não-linear de textos (SILVA; SANTOS, 2006). Esse processo de criação de hipertextos é denominado “autoria” (BROWN, 1990; LIMA, 2004; SHNEIDERMAN, 1989). Assim, pode-se afirmar que o processo de construção do hipertexto é o sistema de autoria e o hipertexto, em si, é o produto desse processo.



Dentre as várias indagações sobre o surgimento do hipertexto, talvez a primeira menção seja a publicação do artigo de Vannevar Bush “*As we may think*” em 1945. Bush descreve uma máquina, o Memex, como um dispositivo mecanizado em que uma pessoa poderia guardar todos os seus documentos (livros, fotos, jornais, revistas) e poderia consultá-los de forma rápida e flexível, como se fosse uma extensão de sua memória (DIAS, 1999). O objetivo da máquina era ampliar os poderes da memória e associação humana. Assim como a mente humana forma memórias através de associações, seria possível ao usuário guardar, recuperar e fazer *links* entre documentos.

Ribeiro (2008) faz uma discussão sobre o recorrente título dado a Vannevar Bush de idealizador do hipertexto, analisando seu famoso artigo “*As we may think*”. Segundo a autora, a importância dada ao artigo de Bush parece basear-se em pelo menos dois pontos: “a ideia de algo que hoje chamamos de hipertexto (e que, ressalte-se, até ali era uma ideia) e a afirmação (também relativa ao hipertexto) de que a mente humana não ‘pensa’ linearmente, mas por associação” (RIBEIRO, 2008, p. 46). Nesse sentido, a proposta de Bush não era a criação de um protótipo tecnológico, mas de um modelo mental a partir da tecnologia.

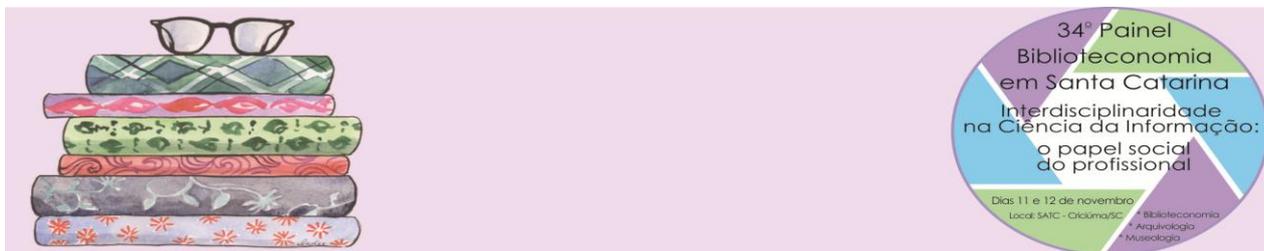
A proposta do Memex não chegou a ser concretizada, mas serviu de inspiração para outras iniciativas. Em 1965, Theodore Nelson, ao idealizar a leitura/escrita não-linear em sistemas informatizados, cunhou o termo “hipertexto”. Seu projeto, Xanadu, tinha por objetivo implementar uma rede de publicações eletrônica, instantânea e universal – um verdadeiro sistema hipertexto (DIAS, 1999). Seria possível não apenas guardar e processar os conteúdos, mas também ligá-los por vínculos *online*.

Nonato (2009) descreve a década de 1980 no contexto da evolução tecnológica, que permitiu a construção de sistemas de autoria (hipertextos) utilizando a potencialidade gráfica – passou-se a utilizar imagens, sons e textos – vista em publicações impressas. Já no início da década de 1990, os hipertextos passaram a ser implementados comercialmente. Segundo o autor, com a concepção – final da década de 1980 – e a criação da *Web* – início da década de 1990 –, os sistemas de hipertextos alcançaram popularidade e visibilidade mundial. Hoje, a *Web* tem alcance mundial e o navegar no hipertexto é uma realidade para milhões de pessoas em todo o mundo (NONATO, 2009).

Desde a implantação da *Web*, os hipertextos vêm sendo estudados no âmbito da Ciência da Computação, e, neste contexto, Marques (1995, p. 85), define os sistemas de hipertextos como

[...] documentos eletrônicos que permitem uma leitura não sequencial, mais de acordo com o raciocínio humano. Propõe uma maneira de interconectar porções de informação e de acessá-las, seguindo o curso natural do raciocínio do usuário, que executa a pesquisa, elabora um documento ou realiza uma série de tarefas que envolvem o uso do computador como suporte à recuperação de informação textual.

Apesar do surgimento do hipertexto estar associado aos usos da informática e não por meio de reflexões linguísticas, o hipertexto é, antes de qualquer coisa, um tipo de intertextualidade que está intrinsecamente ligado à evolução dos modos de leitura e organização da escrita, área que diz respeito à linguística textual. Koch (2000, p. 48) define a intertextualidade como “a relação de um texto com outros textos previamente existentes, isto é, efetivamente produzidos”. No hipertexto, a intertextualidade apresenta-se como uma



característica que lhe é inerente, no sentido de levar a um novo texto, dada a presença dos *hyperlinks*.

No campo da Ciência da Informação (CI), um dos primeiros pesquisadores que tratou do hipertexto foi Pierre Lévy, que avaliou o hipertexto do ponto de vista técnico e funcional: “Tecnicamente, um hipertexto é um conjunto de nós ligados por conexões. [...] funcionalmente, um hipertexto é um tipo de programa para a organização de conhecimentos ou dados, a aquisição de informações e a comunicação” (LÉVY, 1993, p. 33).

Naquele momento, o hipertexto era uma nova forma de organizar a informação e de compartilhá-la com outras pessoas. Também na área da CI, Jayme Leiro Vilan Filho (1994), precursor ao explorar o tema no Brasil, define hipertextos como

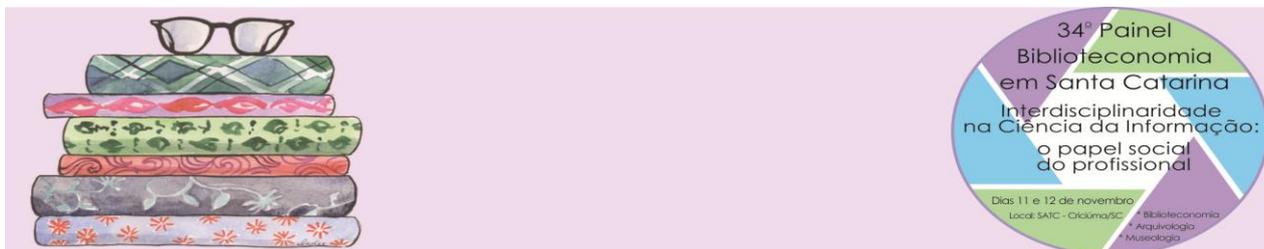
[...] conjuntos de programas de computador (suporte lógico), suas tabelas e dados de controle necessários para a operação de um sistema construído para operar com hiperdocumentos, segundo a filosofia de hipertexto. Em geral, esses sistemas são compostos de: (a) um subsistema de autoria, e (b) um subsistema de navegação (VILAN FILHO, 1994, p. 297).

Numa abordagem do hipertexto enquanto um sistema de navegação semântico, Vilan Filho (1994) afirma que a estruturação do hipertexto é similar a uma rede semântica, já que os nós representam conceitos e ideias, e as ligações representam os relacionamentos entre conceitos. Nesse sentido, esses relacionamentos são determinados a partir da estrutura do assunto abordado no hipertexto. Assim, o autor enfatiza que a rede de relacionamentos do hipertexto deve refletir a estrutura de uma rede semântica. Nessa mesma abordagem, com uma das primeiras aplicações no campo da CI, Lima (2004) considera que os hipertextos são projetados para efetuar a navegação dentro de um espaço de informações. Assim, a determinação da estrutura da rede de conexões é um projeto crucial para a construção do sentido, tanto pelo autor que desenvolve o hipertexto, quanto pelo o usuário que irá navegar.

Diferentemente da forma que a informação é estruturada num livro tradicional, o leitor do hipertexto pode atravessar conexões entre as partes do documento numa fração de tempo mais rápida e fluida. Em seus estudos, Campos (2001) aborda essa estrutura do hipertexto sob dois aspectos: um físico-tecnológico e outro semântico. No primeiro aspecto, o físico/tecnológico, a autora considera que o hipertexto é formado por dois componentes: nós e *links*. Os nós são as unidades de informação do hipertexto, são as partes que formam o ‘todo’ hipertextual. Essas partes podem conter diferentes tipos de informação: sons, imagens ou informação textual. Os *links* - que podem ser chamados de elos, ligações, vínculos, âncoras ou botões – têm a função inconfundível de conectar um nó a outro. Os *links* podem ser representados pela forma textual – utilizando palavras ou frases –, ou pela forma iconográfica, através do uso de imagens.

Em relação à abordagem semântica, Campos (2001) discorre que cada nó deve ser tratado como um conceito e os *links* devem ser tratados como relacionamentos entre os conceitos. Assim, ao pensar a estrutura semântica do hipertexto, é parte fundamental apresentar as informações de forma organizada, fazendo com que o usuário não se perca na navegação, mas sim encontre sentido no conteúdo disposto no hipertexto.

Com outra nomenclatura, aplicada na área da Ciência da Computação, Marques (1992) define os componentes do hipertexto como nodos (nós) e elos (*links*). Para a autora, os nodos representam unidades discretas de informação contendo um conceito. A estruturação deste



componente deve ser cuidadosa, uma vez que as ideias contidas nos nodos devem-se modularizar de forma lógica. Já os elos são os elementos que estabelecem as ligações no hipertexto, mas não necessariamente hierárquicas ou lineares, entre os nós. A autora cita algumas funções dos *links*, considerados por ela como os principais elementos na estruturação do hipertexto: (1) conectar uma referência de um documento ao próprio documento; (2) anexar ao documento uma observação ou comentário sobre seu conteúdo; (3) interligar textos; (4) associar tabelas, figuras, gráficos ou sons; (5) conectar documentos independentes e fornecer informações sobre a organização da informação contida no hipertexto.

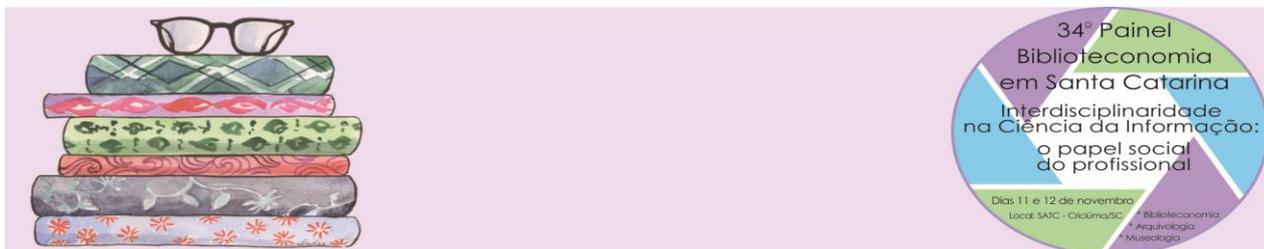
Desconsiderar quaisquer dessas funções pode comprometer a qualidade do produto final. O uso do hipertexto tem aspectos positivos e negativos, conforme descrito na próxima seção.

3.1.1 Vantagens e problemas do hipertexto

A tecnologia de hipertexto apresenta vantagens significativas se comparado à manutenção de informações no formato impresso. A estruturação da informação em forma de hipertexto possibilita, segundo Martin (1992): (1) facilidade muito maior para investigação de referências, o que auxilia no processo de pesquisa e aprendizado; (2) facilidade de manipulação de grandes bases de conhecimento, sem os inconvenientes do meio impresso (peso, volume, necessidade de deslocamento); (3) capacidade de organização de formas dinâmicas e flexíveis, através de maneiras diversas de estruturação (hierárquica, não hierárquica, com diversos níveis ou visões); (4) possibilidade do leitor realizar anotações e referências, sem a necessidade de utilizar margens ou folhas separadas; (5) facilidades adicionais de pesquisa e recuperação de informação, por meio de palavras-chave, assuntos, referência a sessões anteriores de pesquisa, histórico de visita a hiperdocumentos, entre outros; (6) possibilidade de busca de informações através de sistemas gráficos (*browsers* ou folheadores) que permitem um acesso mais direto e intuitivo; (7) facilidades para criação/manutenção de hiperdocumentos, apresentando um meio mais simplificado para alterações/modificações que o meio impresso; (8) utilização dos recursos dos equipamentos para criar audiovisualizações, combinar texto com imagens estáticas (gráficos) e também com imagens animadas, além de sons e programas executáveis.

Entretanto, apesar de apresentar todas essas vantagens, os hipertextos são criticados por algumas de suas limitações. Os maiores problemas encontrados e discutidos na literatura sobre hipertextos ligam-se, diretamente, ao momento da navegação do usuário. Conklin (1987) foi o primeiro autor que elaborou uma lista de problemas cognitivos relacionados ao hipertexto, classificando esses problemas em duas principais questões: a “desorientação” e o “transbordamento cognitivo”. O primeiro problema – desorientação – surge da ausência de um senso de localização e direção em um documento não linear. Já o problema do transbordamento cognitivo apresenta-se pela dificuldade do usuário em adaptar-se à sobrecarga mental derivada do grande número de operações (criação, rotulação e memorização de novas ligações), ou de simples consultas em pedaços de informação diferentes e contextualmente heterogêneas, num curto espaço de tempo.

Dessa forma, a não linearidade do hipertexto exige cuidados tanto no momento de uso do produto (na leitura do hipertexto) quanto na criação do sistema (autoria), quando se faz necessário o uso de uma metodologia adequada ao propósito final desejado.



3.2 MODELOS, FERRAMENTAS E METODOLOGIAS PARA CONSTRUÇÃO DE HIPERTEXTOS

Shneiderman (1989, p. 3, tradução nossa) acredita que o processo de autoria precisa obedecer às “regras de ouro do hipertexto: existe um grande corpo de informação organizado em numerosos fragmentos; os fragmentos relacionam-se entre si; o usuário precisa somente de uma fração de cada vez”. Por sua vez, Lima (2004) considera que a autoria de hipertextos, em especial durante o seu processo de organização conceitual, exige critérios claros e consistentes, pois a falta destes leva a um direcionamento ineficaz do leitor.

Uma das maiores dificuldades no processo de criação de hipertextos é a fragmentação das unidades de conhecimento sem causar prejuízo ao seu conteúdo semântico. Na tentativa de amenizar essa problemática, algumas metodologias e modelos de design foram criados para a estruturação de sistemas de hipertextos, seja para o momento da autoria (criação), seja para a recuperação (folheio e busca) (LIMA, 2004).

As próximas seções descrevem os modelos, ferramentas e as metodologias que já foram recuperadas na pesquisa bibliográfica, que são parte do resultado desta pesquisa.

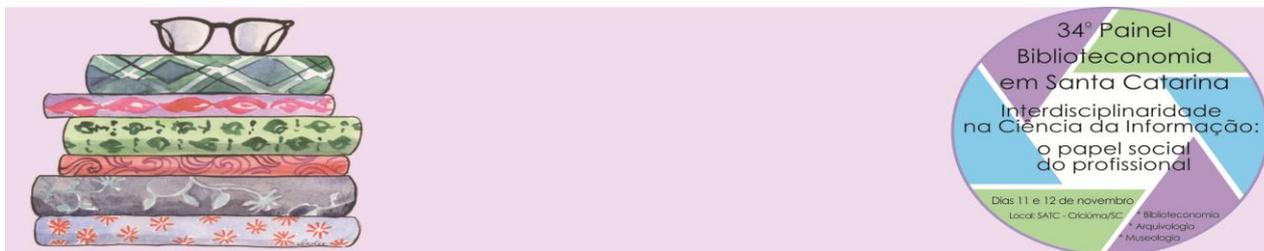
3.2.1 Modelos de design para estruturação de hipertextos

Lima (2004) faz uma revisão dos modelos de *design* de hipertexto (Quadro 1) pré e pós-*Web* mais conhecidos desde o final da década de 1980.

Quadro 1 – Modelos de *design* de hipertextos

Modelos de design	Autores	Data
Asynchronous Design	Perlman	1989
Dexter Reference Model	Halasz e Schwartz	1990
Nested Context Model – NCM	Casanova et al.	1991
Hypermedia Design Model – HDM	Garzotto, Paolini e Schwabe	1993
Relationship Management Methodology – RMM	Isakowitz, Stohr e Balasubramanian	1995
Object Oriented Hypermedia Design Model – OOHDM	Schwabe e Rossi	1995
Enhanced Object-Relationship Model – EORM	Lange	1996
Web Site Design Method – WSDM	De Troyer e Leune	1997
Scenario-Based Object-Oriented Hypermedia Design Methodology – SOHDM	Lee, Lee e Yoo	1999
Web Modeling Language – WebML	Ceri, Fraternali e Bongio	2000
ONTOPORTAL	Universidade de Southampton (Inglaterra)	2000
Model-Driven Ontology-Based Web Site Management – OntoWebber	Yuhui, Decker e Wiederhold	2001
CREating Metadata - CREAM	Handschuh, Staab e Maedche	2002
Conceptual Open Hypermedia Service Environment - COHSE	Universidade de Southampton (Inglaterra)	2002

Fonte: elaborado pelas autoras, adaptado de Lima (2004).



Em seus estudos, Lima (2004) conclui que os modelos até então criados e descritos por ela não satisfaziam às necessidades de produção de hipertextos devido a problemas de fragmentação e de recuperação semântica nos contextos do próprio hiperdocumento.

Apesar da existência, atualmente, de diferentes ferramentas para a produção de hipertextos, Khalifa e Shen (2010) consideram que estes estão longe de alcançar o potencial dos hipertextos de transmitir conhecimento de modo mais eficiente que aquele possível com o texto linear, tradicional.

3.2.2 Ferramentas para autoria de hipertextos

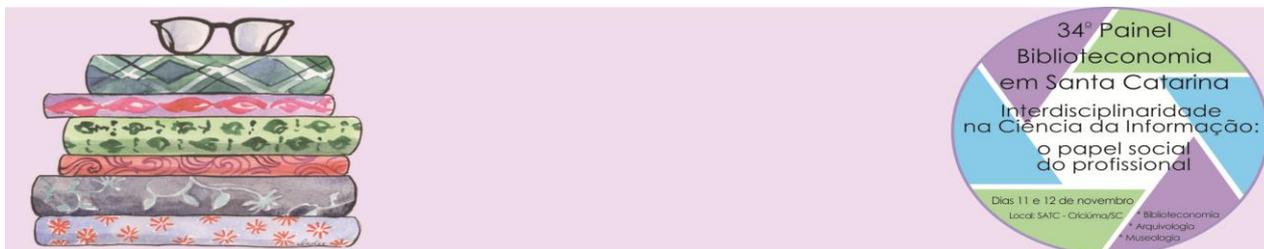
Sistemas de hipertextos são programas de computador que possibilitam a produção (autoria) e a leitura de hipertextos. Nesta pesquisa, serão tratados os do primeiro tipo, que podem ser classificados como sistemas de autoria manual ou sistemas de autoria automática. Nos sistemas de autoria manual, o próprio autor do hipertexto possui a tarefa de segmentar os fragmentos textuais e criar as ligações intra-documento e também ligações inter-documento. Já nos sistemas para autoria automática utilizam-se mecanismos computacionais para segmentar o texto em fragmentos e produzir ligações relevantes (inter-documento e intra-documento) de maneira automática (PAULA, 2013).

Paula (2013), em revisão recente, apresenta sistemas capazes de apoiar a autoria manual de microtextos, isto é, sistemas capazes de converter texto em hipertexto: (1) *Viki*, um dos primeiros sistemas que se preocupou com a representação espacial do hipertexto; (2) *HART*, desenvolvido na segunda geração de hipertextos com a tarefa de facilitar o processo de conversão; (3) *Storyspace*, que apesar de ser da segunda geração, ainda mantém-se atualizado e é um dos mais utilizados na produção de hipertextos; (4) *HypeDyn*, criado na terceira geração com foco no processo de autoria de histórias complexas em hipertextos.

Mitchell e Mcgee (2009) fazem uma crítica aos sistemas que dão suporte à autoria de hipertexto, uma vez que a maior parte desses limita-se a representar mapas de nós e ligações que não são publicados na *Web*. Dessa forma, não são úteis para que os autores possam criar estruturas narrativas mais complexas, como por exemplo, textos que contemplem múltiplos pontos de vista sobre um mesmo tema.

No âmbito da estruturação de hipertextos, Vilan Filho (1994) já mostrava-se preocupado com o processo de conversão de textos em estruturas hipertextuais. Para o autor, essa área precisa de muitas pesquisas para o desenvolvimento de técnicas e instrumentos apropriados a essa tarefa. Ainda segundo o autor, todo o conhecimento que vem sendo produzido em formatos convencionais poderia ser melhor utilizado se transformado em hipertexto. Pesquisas mais recentes ainda apontam nesta direção, de que não há ainda uma compreensão clara sobre os processos para a conversão de um texto linear em um hipertexto que possa ser lido segundo múltiplos pontos de vista (PAULA, 2013). Destacam-se alguns tipos de conversão de textos:

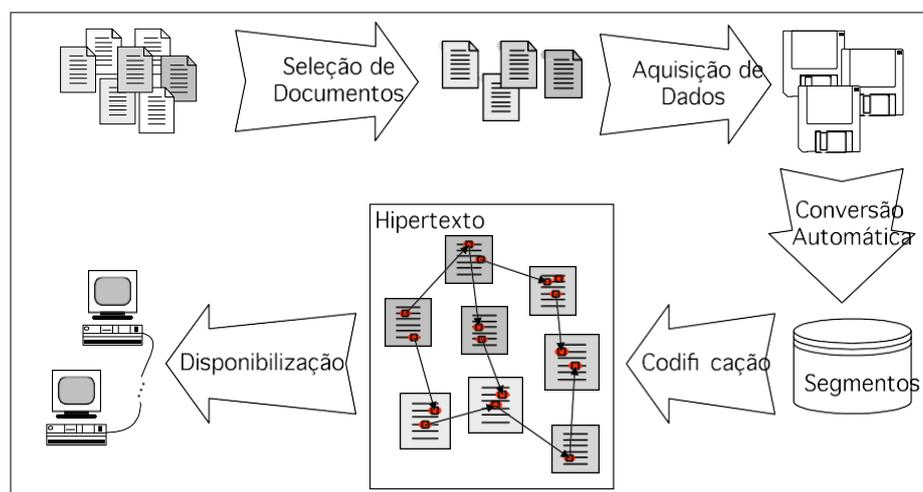
(1) **Conversão manual (intelectual):** Em um processo de criação manual de hipertextos, “os projetos são abordados individualmente, e a cada hiperdocumento a ser produzido todo o processo é repetido” (DORANTE, 1997, p. 3). A conversão manual poderia ser feita por meio dos diversos sistemas de autoria disponíveis, como já mencionados neste



trabalho. Nesse caso, o processo se torna mais sujeito a falhas humanas, muitas vezes associadas à decisão do que deve ser um nó ou um *link*.

(2) **Conversão automática:** O processo de conversão automática de texto para hipertexto é a criação de estruturas hipertextuais a partir de textos em sua forma linear. Compreende todas as etapas de desenvolvimento, desde a escolha do texto que se queira converter até a disponibilização do hipertexto para navegação do usuário. Dorante (1997), em sua dissertação de mestrado, investiga mecanismos para aproveitamento automatizado de material impresso como fonte de informação para a criação automática de hipertextos. Segundo a autora, a geração de hipertextos pode ser apresentada em cinco macro-etapas, ilustradas na Figura 1.

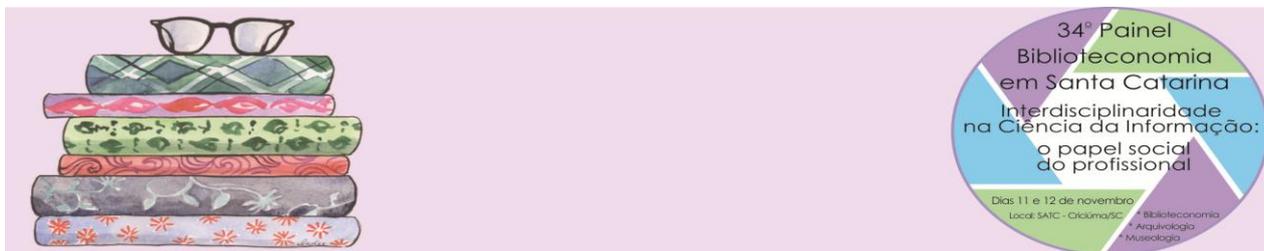
Figura 1 – Macro-etapas da conversão automática



Fonte: Dorante, 1997, p. 4.

Observa-se que o processo de conversão automática engloba: (1) seleção de documentos: definição do conjunto de documentos que serão convertidos em hiperdocumentos; (2) aquisição de dados: transformação dos documentos selecionados em dados digitais; (3) aplicação da conversão automática: segmentação da informação em módulos e criação de hiper-elos ou *links*; (4) codificação em linguagem específica: transformação da informação em uma linguagem com suporte a hipertextos; (5) disponibilização: integração a um servidor conectado (DORANTE, 1997, p. 4). A autora identifica que esse processo é de difícil tratamento para sistemas computadorizados em virtude de dois problemas: o reconhecimento da formatação e estruturação dos documentos e o tratamento das relações entre esses documentos.

Nesse primeiro mapeamento, pôde ser observado que não existem muitos sistemas de conversão disponíveis na literatura, sendo que os encontrados são iniciativas isoladas, tornando-se difíceis de implementar, pois são limitados ao escopo de aplicação para o qual foram desenhados.



3.2.3 Metodologias para estruturação de hipertexto

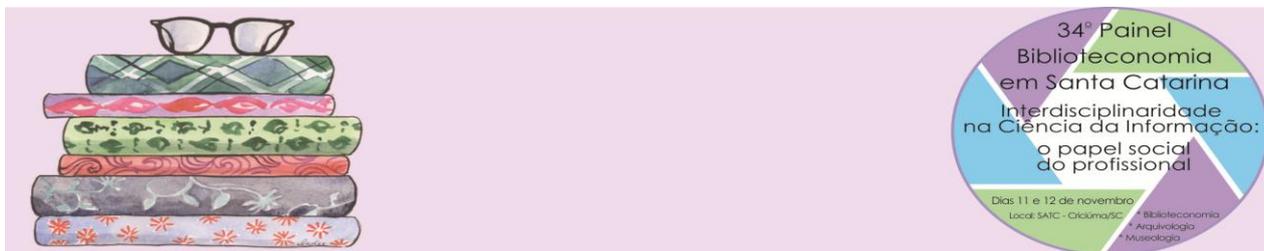
As metodologias para criação de hipertextos são definidas por Avison e Fitzgerald (1995, p. 10) como:

[...] uma coleção de procedimentos, técnicas, ferramentas e documentos os quais ajudarão os criadores de sistemas em seu esforço de implementar um novo sistema de informação. As metodologias consistem em fases, que se subdividem em subfases, as quais irão guiar os criadores de sistemas na sua escolha de técnicas que podem ser mais apropriadas para cada estágio do projeto e também ajudá-los a planejar, administrar, controlar e avaliar projetos de sistemas de informação.

Algumas abordagens da literatura internacional sobre metodologias para estruturação de hipertextos podem ser encontradas em Marchionini (1994, p. 87). O autor sugere a utilização de oito algoritmos para criar um hipertexto baseado nos moldes de elaboração de um índice de um documento: (1) identificação das principais facetas do tópico; (2) criação de uma lista exhaustiva de termos e frases; (3) mapeamento de termos e frases para facetas, revisando-as se necessário; (4) determinação dos termos/conceitos preferidos (*label nodes*); (5) escrita ou introdução dos textos (nodos de informação), como o estabelecimento de referências cruzadas (*links*) para outros nodos durante a criação; (6) revisão do conjunto de textos (nodos) e *links* para outros nodos de acordo com os critérios gramaticais, de estilo, de facilidade de leitura, entre outros; (7) importação dos arquivos revisados para o sistema de hipertexto, implementando os *links*; (8) teste e edição final do hiperdocumento. Ainda segundo o autor, todos os nós do documento devem ser indexados, pois são os vínculos que conduzem o leitor ao conteúdo.

Outra abordagem, apresentada por Yoo e Bieber (2000), baseia-se na análise de relacionamentos entre as unidades de informação do hipertexto. Os autores apresentam os procedimentos em cinco etapas, com objetivo de determinar a estrutura navegacional apropriada ao domínio do conhecimento de que trata o hipertexto: (1) análise de grupos de interesse; (2) elementos de análise; (3) análise de relacionamentos; (4) análise de navegação; (5) análise da viabilidade de implementação.

Na literatura nacional, destacam-se as contribuições de Campos (2001, p. 145), ao apresentar uma metodologia para criação de hipertextos, considerando-se os aspectos conceituais envolvidos na determinação de relacionamentos entre *links*. A autora estabelece os procedimentos em três níveis de entendimento, a partir de sete requisitos (Quadro 2).



Quadro 2 – Metodologia para criação de hipertextos

1º nível:	Nível de entendimento da forma de abordagem do assunto do hiperdocumento.	Requisito 1 – determinação do domínio de conhecimento do hiperdocumento; Requisito 2 – determinação do método de raciocínio utilizado para representar as unidades de conhecimento do hiperdocumento; Requisito 3 – determinação do tipo de leitor do hiperdocumento; Requisito 4 – determinação da tipologia documental do hiperdocumento.
2º nível:	Nível de organização das unidades de conhecimento na construção da narrativa do hiperdocumento.	Requisito 5 – definição da natureza do conteúdo das unidades de conhecimento de um hiperdocumento; Requisito 6 – estabelecimento das relações entre as unidades de conhecimento do hiperdocumento;
3º nível:	Nível do estabelecimento de um veículo de comunicação e expressão sobre a temática do hiperdocumento.	Requisito 7 – elaboração de uma representação gráfica para os nós conceituais e seus relacionamentos em um hiperdocumento.

Fonte: elaborado pelas autoras, adaptado de Campos (2001).

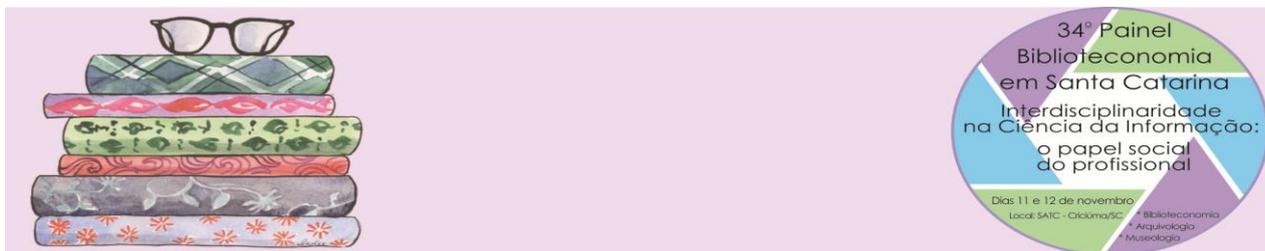
Usando os procedimentos de Campos (2001) e os aportes da Teoria da Classificação Facetada, Lima (2004) também propõe uma metodologia para a organização hipertextual de documentos, com a finalidade de estruturar a informação fragmentada em teses e dissertações, visando facilitar a recuperação em bibliotecas digitais.

A princípio, os resultados encontrados até esta fase dos estudos apontam que esses dois trabalhos irão ser balizares para o desenvolvimento desta pesquisa, que pretende a estruturação de um livro impresso em um hipertexto semanticamente enriquecido, com as relações expressas nos *links*.

4 CONSIDERAÇÕES FINAIS

A comunicação e transmissão de informações é um objetivo que vem sendo buscado desde antes da escrita. Com as tecnologias avançadas, a sociedade vai desenvolvendo formas alternativas de agregar valor às informações disponibilizadas, seja para aprimorar a relevância na sua recuperação, seja para dar maior semântica aos seus conteúdos.

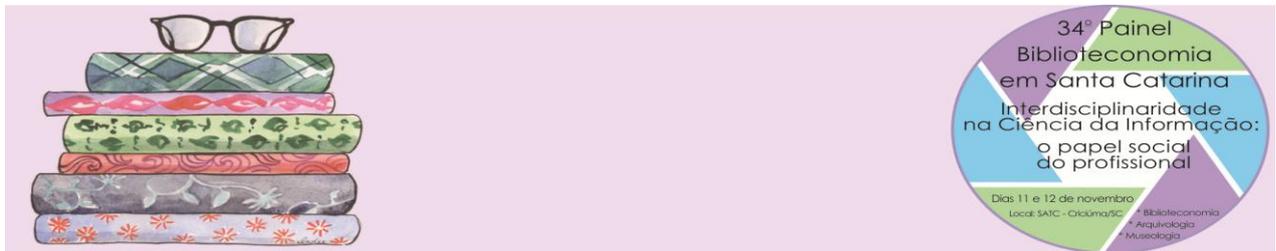
No caso específico desta pesquisa, espera-se que a conversão do livro de Boserup (1965) em um hipertexto, cujo conteúdo versa sobre a teoria por trás da Intensificação Agropecuária, possa facilitar a sua compreensão. Considera-se que esse hipertexto irá se converter num modelo de representação do conhecimento que possa ser utilizado por usuários interessados na compreensão da complexidade agrônômica, ambiental e socioeconômica, embutida nesse processo agropecuário. Com isso, auxiliará na identificação e busca por



soluções de problemas reais, facilitadas pela apreensão da dinâmica sistêmica envolvida, que entende os subsistemas levando em consideração a sua relação com o todo. Essa é uma aproximação à visão holística, que procura perceber uma imagem única de todos os elementos que compõem os sistemas complexos, assim como busca conhecer cada um de seus elementos, aproximando-se, também, da visão reducionista do método analítico das partes. É preciso considerar essas duas visões para entender os elementos do sistema e as inter-relações que ocorrem no todo. Dessa forma, o hipertexto converte-se em ferramenta cognitiva, estabelecendo conexões entre os conhecimentos teóricos e a sua aplicação pragmática em soluções para o setor agropecuário.

REFERÊNCIAS

- ALVARENGA, Lídia. Representação do conhecimento na perspectiva da Ciência da Informação em tempo e espaço digitais. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 8, n. 15, p. 18-40, 2003. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2003v8n15p18/5233>>. Acesso em: 11 maio 2016.
- AVISON, D.; FITZGERALD, G. **Information systems development**. London: McGraw-Hill. 1995.
- BOSERUP, E. **The conditions of agricultural growth: the economics of agrarian change under population pressure**. Chicago: Aldine, 1965. 128 p.
- BROWN, P. J. Assessing the quality of hypertext documents. In: RIZK, A; STREITZ, N.; ANDRE, J. (Ed.). **Hypertext: concepts, systems, and application**. Cambridge, England: Cambridge University, 1990. p. 865-884.
- CAMPOS, Maria Luíza de Almeida. **A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para realização da autoria**. 2001. 190 p. Tese (Doutorado em Ciência da Informação) – CNPq/ IBICT/ UFRJ/ECA: Rio de Janeiro, 2001. Disponível em: <<http://ridi.ibict.br/bitstream/123456789/679/1/mariacampos2001.pdf>>. Acesso em: 23 abr. 2016.
- CONKLIN, Jeff. Hypertext: an introduction and survey. **Computer**, v. 20, n. 9, p. 17-41, set. 1987.
- DIAS, Cláudia. Hipertexto: evolução histórica e efeitos sociais. **Ciência da Informação**, Brasília, v. 28, n. 3, p. 269-277, set./dez. 1999. Disponível em: <<http://www.scielo.br/pdf/ci/v28n3/v28n3a4.pdf>>. Acesso em: 28 abr. 2016.
- DORANTE, Alessandra. **Investigação de processo de conversão automática de textos estruturados para hiperdocumentos**. 1997. 142 p. Dissertação (Mestrado em Física Computacional) - Instituto de Física de São Carlos, Universidade de São Paulo, São Paulo,



1997. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/76/76132/tde-15092010-164303/pt-br.php>>. Acesso em: 28 abr. 2016

HARNAD, S. Post-Gutenberg Galaxy: the fourth revolution in the means of production of knowledge. **Public-Access Computer Systems Review**, v. 2, n. 1, p. 39-53, 1991.

KHALIFA, Mohamed; SHEN, Kathy Ning. Applying semantic networks to hypertext design: effects on knowledge structure acquisition and problem solving. **Journal of the American Society for Information Science**, v. 61, n. 8, p.1673-1685, 2010.

KOCH, Ingedore G. Villaça. **O texto e a construção dos sentidos**. 4 ed. São Paulo: Contexto, 2000.

LEVY, Pierre. **As tecnologias da inteligência**: o futuro do pensamento na era da informática. Rio de Janeiro: Ed. 34, 1993. (Coleção TRANS).

LIMA, Gercina Ângela B. de O. **Mapa hipertextual (MHTX)**: um modelo para organização hipertextual de documentos. 2004. 204 f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2004.

MARCHIONINI, Gary. Designing hypertext: start with an index. In: FIDEL, Raya et al. **Challenges on indexing electronic text and images**. Medford, NJ: ASIS, 1994. cap. 4, p. 77-89.

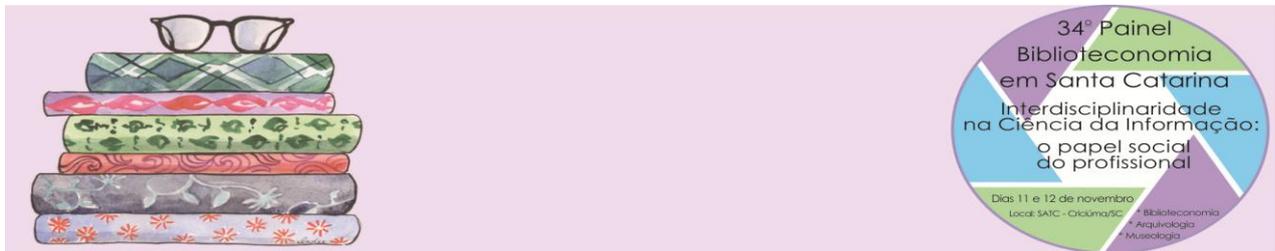
MARQUES, Eugênia Vale. Introdução aos sistemas de hipertexto. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, v. 24, n. 1, p. 85-111, jan./jun. 1995.

MARTIN, J. *Hiper documentos e como criá-los*. Rio de Janeiro: Editora Campus, 1992.

MITCHELL, Alex; MCGEE, Kevin. Designing hypertext tools to facilitate authoring multiple points-of-view stories. In: ACM CONFERENCE ON HYPERTEXT AND HYPERMEDIA, 20, 2009, Torino. **Proceedings...** New York: ACM, 2009. p. 309–316. Disponível em: <<http://dl.acm.org/citation.cfm?id=1557966>>. Acesso em: 27 maio 2016.

MONTEIRO, Silvana Drumond. A forma eletrônica do hipertexto. **Ciência da Informação**, Brasília, v. 29, n. 1, 2000. Disponível em: <http://www.scielo.br/scielo.php?pid=S0100-19652000000100003&script=sci_arttext>. Acesso em: 12 maio 2016.

NONATO, Rafael dos Santos. **Teoria do conceito e hipertextos: uma proposta para determinação de relacionamentos em links conceituais**. 2009. 122 f. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2009. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/ECID-7V2GYA/dissertacao_nonato_2009.pdf?sequence=1>. Acesso em: 16 maio 2016.



PAULA, Antônio Robson de. **A autoria de hipertexto apoiada em tecnologia: um processo para a retextualização digital e uma ferramenta para criação iterativa.** 2013. 94 f. Dissertação (Mestrado em Engenharia de Elétrica) – Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, 2013. Disponível em: <<http://unicamp.sibi.usp.br/handle/SBURI/11339>>. Acesso em: 28 abr. 2016.

RIBEIRO, Ana Elisa. Hipertexto e Vannevar Bush: um exame de paternidade. **Informação & Sociedade**, v. 18, n. 3, p. 45-58, set./dez. 2008. Disponível em: <<http://www.ies.ufpb.br/ojs/index.php/ies/article/viewFile/1811/2271>>. Acesso em: 17 jun. 2016.

SHNEIDERMAN, Ben. Reflections on authoring, editing, and managing hypertext. In: BARRET, Ed (Ed.). **The society of text.** Cambridge, MIT Press, 1989. p. 115-131.

SILVA, Fabio Mascarenhas e; SANTOS, R. N. M. Um estudo das contribuições do hipertexto para o fluxo de informação em meio eletrônico. **Biblios** (Lima), v. 24, p. 12-26, 2006. Disponível em: <<http://repositorio.ufpe.br/handle/123456789/10098>>. Acesso em: 23 abr. 2016.

VILAN FILHO, Jayme Leiro. Hipertexto: visão geral de uma nova tecnologia de informação. **Ciência da Informação**, Brasília, v. 23, n. 3, p. 295-308, set./dez. 1994. Disponível em: <<http://revista.ibict.br/ciinf/article/view/527/527>>. Acesso em: 23 abr. 2016.

YOO, Joonhee; BIEBER, Michael. Finding linking opportunities through relationship-based analysis. In: ACM CONFERENCE ON HYPERTEXT AND HYPERMEDIA, 11, 2000, San Antonio. **Proceedings...** New York: ACM, 2000. p. 181-190. Disponível em: <<http://dl.acm.org/citation.cfm?id=336359>>. Acesso em: 20 maio 2016.

HYPertext STRUCTURE: PROPOSED CONVERSION OF TEXTS

ABSTRACT: This paper is about an ongoing research which is based on literature review in the areas of Information Science, Computer Science and Linguistics. It aims to create an information representation instrument in the form of a navigational hypertext on the field of brazilian agriculture. It proposes to systematize a methodology for the conversion of linear text into hypertext that can assist the process of building hypertexts semantically enriched. It will also approach concepts related to hypertext production, checking for hypertext conversion methodologies, either automatic or intellectual ones. This is on the assumption that the studies that include methodologies on this conversion process do not explain all the procedures that are necessary for conceptual organization of hypertexts in a consistent manner.

Keywords: Hypertext. Conversion of texts into hypertexts. Authoring systems.