

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017

GT- 2 – Organização e Representação do Conhecimento

HIPERTEXTO COM ABORDAGEM SEMÂNTICA PARA A REPRESENTAÇÃO DO CONHECIMENTO

Elaine Diamantino Oliveira - Universidade Federal de Minas Gerais (UFMG)

Benildes Coura Moreira dos Santos Maculan - Universidade Federal de Minas Gerais
(UFMG)

Ivo Pierozzi Júnior - Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA)

HYPertext WITH SEMANTIC APPROACH TO REPRESENTATION OF KNOWLEDGE

Modalidade da Apresentação: Pôster

Resumo: Pesquisa de mestrado em andamento, que propõe o uso do hipertexto como instrumento de representação do conhecimento acerca da área temática Intensificação Agropecuária, cujo teor conceitual está apresentado no livro da autora Ester Boserup (1965). Objetiva-se propor um modelo para conversão de texto linear em hipertexto que possa auxiliar o processo de construção de hipertextos semanticamente enriquecidos, criando um ambiente propício à compreensão do conteúdo apresentado no texto. 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. Recorre-se aos fundamentos do campo da Representação da Informação e do Conhecimento, com ênfase na Teoria da Classificação Facetada e na Teoria do Conceito, destacando as suas contribuições para a modelagem conceitual de hipertextos. Utiliza-se a proposta de refinamento das relações semânticas entre conceitos, evidenciada por Maculan (2015), na aplicação do modelo hipertextual. Consideram-se, também, os aportes da Teoria da Aprendizagem Significativa no processo de aquisição de novos conhecimentos. Espera-se que, a partir da conversão do livro (texto linear) em um hipertexto, que as relações expressas nos *links* sejam controladas semanticamente para facilitação da aprendizagem do conteúdo do documento.

Palavras-Chave: Conversão de Hipertextos. Modelagem Conceitual. Relações Semânticas.

Abstract: A master's degree in progress, which proposes the use of hypertext as an instrument for representing knowledge about the thematic area Agricultural Intensification, whose conceptual content is presented by the author Ester Boserup's book (1965). The aim of this work is to propose a model for converting linear text into hypertext that can help the construction of semantically enriched hypertexts, creating an environment conducive to understanding the content presented in the text. It is assumed that studies that cover methodologies about this conversion process do not make explicit all the procedures that are necessary for the systematic organization of hypertexts consistently. It is

based on the foundations of the field of Representation of Information and Knowledge, with emphasis on the Theory of Concept and the Theory of Faceted Classification, highlighting their contributions to the conceptual modeling of hypertexts. We use the proposal of refinement of semantic relations between concepts, evidenced by Maculan (2015), in the application of the hypertextual model. We also consider the contributions of Significant Learning Theory in the process of acquiring new knowledge. It is hoped that, from the conversion Boserup's linear text of into a hypertext, that the relations expressed in the links be semantically controlled to facilitate the learning of the document content.

Keywords: Conversion of Hypertexts. Conceptual Modeling. Semantic Relationships.

1 INTRODUÇÃO

A tecnologia do 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 das vantagens do hipertexto. Considerando-se isso, o hipertexto se manifesta como uma nova forma de leitura e escrita hiperconectadas, regida pelo princípio da não linearidade, e se materializa enquanto instrumento de representação do conhecimento, a partir de uma estrutura composta por nós (conceitos) e *links* (relações).

Através da sua estrutura não linear associativa, “os documentos hipertextuais apresentam inúmeras inovações em relação aos documentos tradicionais impressos, sobretudo nos processos de pesquisa e aprendizado, uma vez que a leitura hipertextual baseia-se, predominantemente, no processo de navegação” (SILVA, 2008, p. 13). O ato de navegar conduz o leitor a interagir, de forma dinâmica, com as porções de informação do hipertexto e estabelecer novos relacionamentos entre essas porções, além dos previstos pelo autor do hipertexto.

Apesar das vantagens proporcionadas pelos hipertextos, a sua produção ainda se configura como um desafio. Percebe-se que os mesmos ainda carecem de uma formalização metodológica e do desenvolvimento de ferramentas mais efetivas para a sua elaboração. Ao longo dos anos, e com o avanço tecnológico, foram propostos diversos modelos de *design* e sistemas para a construção de hipertextos. No entanto, as pesquisas ainda apontam para a limitação dessas ferramentas e modelos, que estão aquém do potencial do hipertexto (KHALIFA; SHEN, 2010; PAULA, 2013). Como consequência, os hipertextos são estruturados do ponto de vista da intencionalidade de seu criador, não refletindo adequadamente a estrutura semântica do documento, tampouco as necessidades informacionais de seus possíveis usuários.

A partir de uma análise da literatura que versa sobre a construção de hipertextos, entre os anos de 2004 a 2016 (OLIVEIRA; MACULAN; PIEROZZI JÚNIOR, 2016), este estudo parte do pressuposto de que as metodologias empregadas nos processos de estruturação de hipertextos, em geral, não apresentam os procedimentos completos, desde a modelagem conceitual do objeto em questão, até a sua disponibilização para uso, e que estão pulverizadas nas áreas de Ciência da Computação, Linguística e Educação, com pouca ocorrência de estudos na Ciência da Informação. Foi a partir dessa lacuna que este estudo foi concebido e delineado.

Nessa perspectiva, este trabalho tem como objetivo propor um modelo para a conversão de texto linear em um hipertexto semanticamente enriquecido. O domínio de aplicação desta pesquisa é a Agropecuária, mais especificamente, a área temática da Intensificação Agropecuária. A modelagem do hipertexto baseia-se na teoria de Ester Boserup (1965), cujo teor conceitual está apresentado em seu livro *“The Conditions of Agricultural Growth: The Economics of Agrarian Change under Population Pressure”*, publicado em 1965. Pretende-se, a partir da conversão desse livro (texto linear) em um hipertexto, facilitar, ao usuário final, o entendimento do conteúdo semântico contido no documento.

2 FUNDAMENTOS TEÓRICOS E METODOLÓGICOS

Os princípios da Teoria da Classificação Facetada (TCF) e da Teoria Conceito, ambas desenvolvidas no contexto da Biblioteconomia e da Ciência da Informação, irão nortear a construção do modelo, pois possuem potencial para modelar um domínio do conhecimento de forma consistente.

A TCF, através de princípios formulados por Ranganathan (1892-1972) e dos estudos do *Classification Research Group* (CRG), propõe categorias fundamentais para mapear o conhecimento de um domínio: as cinco categorias de Ranganathan, o PMEST (Personalidade, Matéria, Energia, Espaço e Tempo) e as dez categorias do CRG: Tipos de Produto Final, Materiais, Propriedades, Processos, Operações, Agentes, Espaço, Tempo e Forma de Apresentação (LIMA, 2004). A TCF possui uma abordagem analítico-sintético que envolve dois processos distintos: (1) fase analítica: fazer a análise de assunto em facetas, fragmentando-o pelas partes que o constituem e decompondo os assuntos complexos em seus elementos básicos; (2) fase sintética: analisar cada uma das partes fragmentadas e decompostas para que, posteriormente, possam ser combinadas para representar o domínio (ou o documento) mapeado (TRISTÃO; FACHIN; ALARCON, 2004).

Tendo em vista os princípios que regem a TCF, Lima (2002) considera que a mesma fornece subsídios para a modelagem conceitual em hipertextos uma vez que “reconhece muitos aspectos em um único assunto, e tenta sintetizá-los de maneira que os descreva de forma mais adequada” (LIMA, 2002, p. 191). Nesse sentido, a TCF contribui para o estabelecimento de uma rede coerente de conceitos e relações, pois fornece critérios mais confiáveis a serem aplicados na organização e representação do conteúdo semântico de sistemas hipertextuais (CAMPOS, 2000; LIMA, 2004).

Já a Teoria do Conceito pode ser aplicada como instrumento de auxílio para clarificar e sistematizar conceitos de um domínio, tendo em vista que a teoria permite a construção de sistemas conceituais, ao definir adequadamente seus conceitos por meio da identificação de suas características (DAHLBERG, 1978). Ela será aplicada como etapa de modelagem da terminologia do domínio dentro do modelo que será proposto.

As relações entre conceitos, propostas pela teoria de Dahlberg, podem ser agrupadas em: 1) lógicas: quando existem características comuns nas definições dos conceitos, implica que há relacionamento direto entre os conceitos; 2) hierárquicas: relacionamento entre gênero-espécie; 3) partitivas: relacionamento entre o todo e suas partes; 4) de oposição: relacionamento de oposição contraditória; 5) funcionais: relacionamento entre os componentes de uma afirmação (DALHBERG, 1978). Acredita-se que os relacionamentos que cada conceito mantém com os outros conceitos podem revelar grande parte do conhecimento registrado em um determinado domínio, orientando para que o processo de modelagem conceitual seja bem fundamentado. Sendo assim, evidencia-se a pertinência do uso da Teoria do Conceito para a modelagem de hipertextos, uma vez que a aplicação das relações apresentadas pode ser útil na identificação de conceitos e na elaboração de relacionamentos semânticos no hipertexto.

Considerando o hipertexto como um recurso instrucional propício à aquisição de conhecimentos, utilizam-se as contribuições da Teoria da Aprendizagem Significativa, desenvolvida por David Ausubel em 1963. A ideia básica da teoria envolve a interação cognitiva entre conhecimentos prévios e novos conhecimentos, sendo que esses conhecimentos não são, necessariamente, conceitos, podendo ser ideais ou modelos, representações que servem de apoio para que novos conhecimentos possam ser reconstruídos significativamente pelo aprendiz (MOREIRA, 2011).

A partir da interação entre esses conhecimentos – prévios e novos – ocorrem modificações em ambos, pois os primeiros ganham maior estabilidade cognitiva e os segundos adquirem significado, podendo, cada vez mais, facilitar novas aprendizagens. (AUSUBEL; NOVAK; HANESIAN, 1980). De forma semelhante, a aprendizagem significativa pode ocorrer pela interação em grupos, comunidades ou organizações que buscam a construção de conhecimentos novos sobre teorias, conceitos e metodologias, que devem ser apreendidas para que possam ser implementadas e aperfeiçoadas na prática (COUTINHO; KALLÁS, 2005). Nesse contexto, parte-se do pressuposto de que os princípios da Teoria da Aprendizagem Significativa podem auxiliar a compreensão dos processos da área a ser modelada, de maneira que potencialize a aprendizagem do conteúdo proposto pelo hipertexto.

Para a elaboração do modelo de hipertexto proposto, os fundamentos teórico-metodológicos foram definidos a partir das metodologias recuperadas na literatura internacional e nacional, nas quais destacam-se as seguintes: a) Riner (1991) sugere uma estrutura composta por macro etapas, de todo o processo de conversão, desde a escolha do item a ser transformado até a disponibilização do hipertexto para uso; b) Campos (2001) apresenta uma metodologia para a criação de hipertextos considerando os aspectos conceituais envolvidos na determinação de relacionamentos entre *links*; c) Lima (2004), usando os procedimentos de Campos (2001) e os aportes da TCF, propõe uma metodologia para a organização hipertextual de documentos, visando facilitar a recuperação em bibliotecas digitais.

A partir da análise das metodologias mencionadas, percebeu-se que as mesmas não preveem uma etapa para o refinamento semântico da estrutura conceitual. No entanto, Maculan (2015) defende a ideia de que o refinamento das relações semânticas auxilia a organização do conhecimento no domínio a ser modelado, pois especifica o tipo de relação existente entre os conceitos e termos, com vistas a indicar toda a semântica conceitual do domínio. Para a autora, criar estruturas conceituais semanticamente mais refinadas oferece ao usuário maior valor agregado às informações disponibilizadas, facilitando a sua compreensão e a ampliação do entendimento acerca do conhecimento da área modelada.

3 METODOLOGIA E PROCEDIMENTOS

O estudo, de natureza exploratória e experimental, se apoia em revisão bibliográfica nas áreas da Ciência da Informação, Ciência da Computação, Linguística e Educação. Como domínio de aplicação, foi escolhida a agropecuária brasileira, uma vez que esta pesquisa está

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

inserida em um convênio formalizado entre a Embrapa Informática Agropecuária, sediada em Campinas, e a Universidade Federal de Minas Gerais (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.

Além das três propostas mencionadas de Riner (1991), Campos (2001) e Lima (2004) – insumos metodológicos – também serão utilizados insumos de natureza semântica (refinamento de relações entre conceitos) e tecnológica (*softwares* de visualização). Os insumos semânticos se originam dos resultados da tese de Maculan (2015), na qual a autora analisa e modela a área da Intensificação Agropecuária com base na teoria de Boserup (1965), contando com uma amostra composta por trinta conceitos e representada por cerca de seiscentos relacionamentos refinados. Esses insumos serão utilizados para explicitar a relação entre os conceitos do domínio, compondo uma estrutura conceitual mais significativa. Já a aplicação de insumos tecnológicos prevê o uso de *softwares* para duas finalidades: (a) permitir a visualização de partes da estrutura modelada no formato de um mapa conceitual; (b) permitir a conversão de texto linear em hipertexto.

No Quadro 1 estão apresentadas a descrição das etapas metodológicas e as ações, dentro de cada etapa, necessárias para atingir o objetivo proposto pela pesquisa.

Quadro 1: Descrição das etapas metodológicas.

Etapas	Descrição das Etapas	Ações
ETAPA I	Procedimentos necessários à conversão de um texto linear em hipertexto	<ul style="list-style-type: none">➤ Analisar os requisitos propostos por Riner, Campos e Lima;➤ Sistematizar os requisitos em um modelo para conversão de texto linear em hipertexto semanticamente enriquecido.
ETAPA II	Aplicação do modelo proposto para conversão no domínio da Intensificação Agropecuária	<ul style="list-style-type: none">➤ Aplicar os procedimentos do modelo no domínio da Intensificação Agropecuária;➤ Inserir os insumos semânticos resultantes da tese de Maculan (2015) no estabelecimento das relações entre os conceitos.
ETAPA III	Edição da estrutura semântica final em <i>software</i> de visualização hipertextual	<ul style="list-style-type: none">➤ Elaborar uma representação gráfica (mapa conceitual) para os nós conceituais e seus relacionamentos semânticos;➤ Converter o conteúdo conceitual do livro de Boserup no formato de um hipertexto semântico;➤ Disponibilizar a estrutura hipertextual final para os usuários.

Fonte: Elaborada pela autora, 2017.

4 RESULTADOS PARCIAIS E CONSIDERAÇÕES FINAIS

Os estudos realizados até o momento já permitiram alguns resultados, como a sistematização da proposta do modelo para conversão de texto em hipertexto (Quadro 2).

Quadro 2: Modelo para conversão de texto em hipertexto.

PROPOSTA DE REQUISITOS	ORIGEM DOS REQUISITOS
1 Seleção do documento	Riner, Campos, Lima
2 Forma de aquisição de dados	Riner
3 Modelagem conceitual a) Determinação do domínio do conhecimento b) Método de raciocínio utilizado para a representação das unidades do conhecimento c) Tipo de leitor d) Tipologia documental do hiperdocumento e) Natureza do conteúdo das unidades de conhecimento f) Estabelecimento das relações entre as unidades de conhecimento ✓ Identificação do documento básico de trabalho ✓ Leitura do documento ✓ Análise facetada do assunto do documento g) Explicitação das relações semânticas entre conceitos	Campos, Lima, Maculan
4 Elaboração de uma representação gráfica para os nós conceituais e seus relacionamentos	Riner, Campos, Lima
5 Disponibilização: integração a um servidor conectado	Riner

Fonte: elaborado pela autora, adaptado de Riner (1991), Campos (2001), Lima (2004) e Maculan (2015).

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 o entendimento da teoria por usuários interessados na compreensão da complexidade agronômica, ambiental e socioeconômica, embutida nesse processo agropecuário. Espera-se, também, que o modelo de conversão possa ser utilizado para direcionar a construção de hipertextos semanticamente enriquecidos em qualquer domínio que se pretenda representar, contribuindo para dispor uma estrutura semântica que propicie o aprendizado.

REFERÊNCIAS

AUSUBEL, D. P.; NOVAK, J.; HANESIAN, H. **Psicologia educacional**. Rio de Janeiro: Interamericana, 1980.

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

BOSERUP, E. **The conditions of agricultural growth**: the economics of agrarian change under population pressure. Chicago: Aldine, 1965. 128 p.

CAMPOS, M. L. A. **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.

COUTINHO, A. R.; KALLÁS, D. **Gestão da estratégia**: experiências e lições de empresas brasileiras. Oxford, United Kingdom: Gulf Professional Publishing, 2005. 273p.

DAHLBERG, I. Uma teoria para o interconcept: teoria analítica do conceito voltada para o referente. [Publicado originalmente na revista] **International Classification**, v. 5, n. 3, p. 142-151, 1978.

KHALIFA, M.; SHEN, K. N. 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.

LIMA, G. Â. B. O. Análise facetada na modelagem conceitual de sistema de hipertexto: uma revisão de literatura. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 7, n. 2, p. 189-196, jul./dez. 2002. Disponível em:
<<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/407>>. Acesso em: 17 jun. 2016.

_____. **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.

MACULAN, B. C. M. S. **Estudo e aplicação de metodologia para reengenharia de tesauro**: remodelagem do THESAGRO. 2015. 339 f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2015.

MOREIRA, M. A. **Aprendizagem significativa**: a teoria e textos complementares. São Paulo: Editora Livraria da Física, 2011. 179 p.

OLIVEIRA, E. D.; MACULAN, B. C. M dos S.; PIEROZZI JÚNIOR, I. Estruturação de hipertextos: proposta de conversão de textos. Revista ACB: Biblioteconomia em Santa Catarina, Florianópolis, v. 21, n. 3, p. 564-578, ago./nov. 2016. Disponível em:
<<https://revista.acb.org.br/racb/article/view/1195>>. Acesso em: 13 jun. 2017.

PAULA, A. R. 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://bdtd.ibict.br/vufind/Record/CAMP_dcd659a3a2bf288f23a372e5542f7f12>. Acesso em: 28 abr. 2016.

RANGANATHAN, S. R. **Prolegomena to library classification**. 3. ed. London: Asia Publishing House, 1967.

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

RINER, R. Automated conversion. In: BERK, E.; DEVLIN, J. (Ed.). **Hypertext / hypermedia handbook**. Hightstown, NJ: McGraw-Hill Inc., 1991. p. 95-111.

SILVA, G. B. S. **A utilização de mapas de tópicos na compatibilização de conteúdos hipertextuais semanticamente estruturados**. 2008. 143 f. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

TRISTÃO, A. M. D.; FACHIN, G. R. B.; ALARCON, O. E. Sistemas de classificação facetada e tesouros: instrumentos para organização do conhecimento. **Ciência da Informação**, Brasília, v. 33, n. 2, p. 161-171, maio/ago. 2004.