



Transinformação

ISSN: 0103-3786

transinfo@puc-campinas.edu.br

Pontifícia Universidade Católica de  
Campinas  
Brasil

Fugisawa Souza, Marcia Izabel; Rosa Alves, Maria das Dores; Ribeiro Queiros, Leonardo;  
Delfino dos Santos, Adriana; Mendonça de Oliveira, Leandro Henrique  
Representação descritiva e temática no Sistema Agência de Informação Embrapa:  
controle de vocabulário  
Transinformação, vol. 22, núm. 1, abril, 2010, pp. 61-75  
Pontifícia Universidade Católica de Campinas  
Campinas, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=384334883005>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

redalyc.org

Sistema de Informação Científica  
Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal  
Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

# Representação descritiva e temática no Sistema Agência de Informação Embrapa: controle de vocabulário

## *Descriptive and thematic representation of the Embrapa Information Agency: vocabulary control*

Marcia Izabel Fugisawa SOUZA<sup>1</sup>

Maria das Dores Rosa ALVES<sup>2</sup>

Leonardo Ribeiro QUEIROS<sup>1</sup>

Adriana Delfino dos SANTOS<sup>1</sup>

Leandro Henrique Mendonça de OLIVEIRA<sup>1</sup>

### RESUMO

Este artigo aborda a concepção e o estabelecimento das atividades de representação descritiva e temática e o controle de vocabulário especializado no website Agência de Informação Embrapa. Adota-se o padrão *Dublin Core* para descrição dos metadados dos recursos de informação, que são representados pelas regras mínimas de catalogação e de indexação. São adotados tesouros especializados em agricultura para atribuição de palavras-chave e categorias de assunto. Faz-se o controle de termos livres, visando à consistência do vocabulário e à manutenção da qualidade. Utilizam-se ferramentas de tratamento de recursos de informação e de controle de vocabulário, integradas ao sistema gestor de conteúdo, que abriga o conjunto de atividades de organização da informação do website Agência de Informação Embrapa. As atividades que envolvem a representação descritiva e temática são apoiadas por procedimentos e ferramentas que sistematizam e padronizam suas rotinas, de modo a garantir melhoria na recuperação da informação.

**Palavras-chave:** Representação descritiva. Representação temática. Catalogação. Indexação. Vocabulário controlado. Agência de Informação Embrapa.

### ABSTRACT

*This paper discusses the conception and establishment of the descriptive and thematic representation activities as well as the specialized vocabulary control in the website Embrapa Information Agency (IEA). It is adopted the standard Dublin Core for metadata description of the information resources, which are represented by the minimum rules for cataloging and indexing. Specializing thesauri are adopted in agriculture for assigning*

<sup>1</sup> Embrapa Informática Agropecuária, Laboratório de Organização e Tratamento da Informação Eletrônica. Av. André Tosello, 209, Cidade Universitária, 13083-886, Campinas, SP, Brasil. Correspondência para/Correspondence to: M.I.F. SOUZA. E-mail: <marcia@cnpia.embrapa.br>.

<sup>2</sup> Consultora em Organização e Tratamento da Informação Digital. Campinas, SP, Brasil.  
Recebido em 11/11/2009 e aceito para publicação em 12/2/2010.

keywords and subject categories. Free terms control is assured, aiming vocabulary consistency and quality maintenance. It uses handling tools for information resources of vocabulary control, integrated to the management system content, which houses the activities set of the IEA website information organization. Activities that involve thematic and descriptive representation should be supported by procedures and tools that systematize and standardize their routines in order to ensure improvement in information retrieval.

**Keywords:** Descriptive representation. Thematic representation. Cataloguing. Indexing. Controlled vocabulary. Embrapa Information Agency.

## INTRODUÇÃO

A Empresa Brasileira de Pesquisa Agropecuária (Embrapa) é detentora de um valioso patrimônio científico e tecnológico, resultante de suas atividades de Pesquisa, Desenvolvimento e Inovação (PD&I), voltadas para a sustentabilidade da agricultura, em benefício da sociedade (Embrapa, 2008). Há 36 anos a Embrapa dedica-se dedicando à geração de conhecimento, tecnologia, produtos e serviços para a agricultura nos trópicos. A incorporação de inovações tecnológicas garantiu ao País grandes saltos de qualidade e produtividade agrícola, e, ainda, o provimento de alimentos e insumos para a população urbana e o setor industrial.

Diante da relevância da agricultura para o País, a Embrapa mantém o compromisso permanente de divulgar e a viabilizar o acesso *on-line* às informações e tecnologias por ela produzidas, de interesse imediato dos produtores rurais e agentes de desenvolvimento, assistência técnica e extensão. Para isso, a Empresa criou a Agência de Informação Embrapa (AIE), um *website* de conteúdo totalmente dedicado à organização, tratamento e divulgação da informação tecnológica na internet (Souza *et al.*, 2006).

Este trabalho aborda a concepção e o estabelecimento das atividades de representação descritiva e temática e o controle de vocabulário especializado no *website* Agência de Informação Embrapa.

## SISTEMA AGÊNCIA DE INFORMAÇÃO EMBRAPA

A Agência de Informação Embrapa é um sistema *web* que possibilita a organização, o tratamento, o armazenamento, a divulgação e o acesso à informação tecnológica e ao conhecimento gerados pela Embrapa e outras instituições públicas de pesquisa. Nesse *website*, as informações são organizadas numa estrutura

ramificada e hierárquica, em forma de árvore de conhecimento.

A Agência de Informação Embrapa reúne um conjunto de inúmeras árvores de conhecimento, desenvolvidas pelas Unidades Descentralizadas da Embrapa, sobre culturas agrícolas, criações animais e temas do agronegócio. Para tal, a Embrapa dispõe de metodologias desenvolvidas especificamente para a construção de árvores de conhecimento, envolvendo a elaboração e tratamento de conteúdo (Cobbe *et al.*, 2007; Alves *et al.*, 2008), e de ferramental tecnológico de gestão de conteúdo (Santos *et al.*, 2007a, b, c).

As árvores de conhecimento têm como princípio a estruturação hierárquica do conteúdo a ser organizado, partindo-se do nível mais genérico para o mais específico. Cada item dessa hierarquia é chamado “nó” e/ou “subnó” (Cobbe *et al.*, 2007). Nos primeiros níveis (nós) de uma árvore são exibidos os conhecimentos mais genéricos e, nos níveis (subnós) mais profundos, os mais específicos. A construção de árvores de conhecimento de cadeias produtivas de culturas agrícolas e criações animais é orientada ao desenvolvimento de conteúdos técnicos, cujo arranjo recomendado é composto de três tópicos (nós) principais: pré-produção, produção e pós-produção (Cobbe *et al.*, 2007).

O processo de organização da informação e gestão do conhecimento no Sistema Agência é composto das etapas: a) seleção do tema/produto a ser abordado pela árvore de conhecimento; b) delineamento da estrutura da árvore de conhecimento; c) elaboração de conteúdo da árvore de conhecimento; d) seleção e catalogação de recursos de informação; e) publicação de conteúdo no *website* da Agência. Para cada árvore de conhecimento é constituída uma Equipe Editorial da Agência (EEA), com a função de coordenar a elaboração dos conteúdos da árvore de conhecimento (Embrapa Informática..., 2005).

A EEA é composta de editor técnico, editores assistentes, profissionais de informação, de

comunicação e de informática (Embrapa Informática..., 2005). O editor técnico orienta e coordena o delineamento da estrutura da árvore de conhecimento a ser construída. A etapa de elaboração de conteúdo consiste na criação, revisão e edição dos textos pelos especialistas do domínio, versando sobre tópicos e questões relevantes aos interesses dos usuários. Essa etapa de elaboração de conteúdo pode ser executada em paralelo à seleção de recursos de informação. Assim, à medida que os editores selecionam os materiais-base para a redação dos textos dos conteúdos de nós, também selecionam recursos de informação para complementar esses textos.

Essas informações complementares incluem os recursos eletrônicos, que, associados aos nós, ampliam a compreensão do seu conteúdo, e podem ilustrá-los com outras formas de expressão, como áudio, imagens e sons, entre outros. De acordo com Alves *et al.* (2008) e Souza e Alves (2009), a etapa catalogação de recurso consiste na representação descritiva e temática dos recursos eletrônicos associados aos nós, em conformidade com o padrão de metadados *Dublin Core*. Ferramentas de indexação e categorização de assuntos apóiam a descrição de recursos eletrônicos (Alves *et al.*, 2008). A etapa de publicação consiste na extração dos conteúdos de nós da árvore de conhecimento e dos metadados dos recursos de informação, na geração automática das páginas HTML e sua consequente disponibilização na internet, formando assim o conteúdo do *website* da Agência de Informação Embrapa.

A infraestrutura computacional do *website* da Agência, desenvolvida em *software* livre (Evangelista, 2005b), dispõe de sistema gestor de conteúdo (Santos *et al.*, 2007a, b, c), de mecanismos de busca em metadados e *full text*, e ainda, oferece visualização da informação em estrutura de árvore hiperbólica. Essa visualização gráfica é proporcionada pela ferramenta HiperNavegador, desenvolvida pela Embrapa (Evangelista, 2005a).

A árvore de conhecimento assim construída e publicada na internet propicia a reunião no *website* da Agência de um extenso conjunto de informações validadas sobre temas diversos, culturas agrícolas e criações animais, envolvendo praticamente todas as etapas das respectivas cadeias produtivas. O Anexo 1 ilustra a página de um nó (nível mais genérico), contendo um texto sucinto, seguido de subnós (níveis mais específicos). Ao final da página, tem-se o item Informações Complementares, por meio do qual o

*website* da Agência oferece ao usuário o acesso instantâneo a recursos de informação, tais como textos, imagens, mapas, vídeos, sons, bases de dados etc. (Embrapa Informática..., 2005).

## REPRESENTAÇÃO DESCRITIVA NO SISTEMA AGÊNCIA

A representação descritiva e temática é um componente importante da metodologia de construção de árvores de conhecimento. Por definição, todos os recursos de informação associados aos nós de árvores de conhecimento são descritos e tratados por meio do padrão de metadados *Dublin Core* (Dublin Core..., 2008). Essa prática está fundamentada na imperativa necessidade de adoção de padrões de descrição internacionais, visando à garantia de integração e troca de dados em sistemas similares e a interoperabilidade entre aplicações. Aliada a isso, a qualidade dos conteúdos descritivos dos metadados certamente contribui para a melhoria e o acesso ágil à informação desejada.

A representação descritiva de recursos de informação eletrônicos no Sistema Agência (Souza; Alves, 2009) requer o uso das regras mínimas de catalogação, as quais são baseadas no AACR2 - Código de Catalogação Anglo-americano (2004) -, além de controle de vocabulário. Os elementos metadados *Dublin Core* são descritos de acordo com as recomendações contidas no Manual de Representação Descritiva de Recursos Eletrônicos do Sistema Agência, elaborado por Alves *et al.* (2008).

A descrição temática adota controle de vocabulário, visando à atribuição de palavras-chave e categorias de assunto, apoiada por vocabulário controlado. De acordo com Smit e Kobashi (2003, p. 20), citados por Aguiar (2008, p.207) o “controle de vocabulário é processo para um objetivo que se deseja atingir”, já o vocabulário controlado é resultado desse processo, é “um instrumento para nomear as atividades/funções, gerando confiança no sistema”.

### Metadados *Dublin Core*

O Sistema Agência adota os 15 elementos *Dublin Core*, versão 1.1, definidos no documento *Request for Comments – RFC 2413* - (Weibel *et al.*, 1998), sendo: Título, Criador, Assunto (palavras-chave e categoria),

Descrição, Publicador, Colaborador, Data, Tipo, Formato, Identificador, Fonte, Idioma, Relação, Cobertura e Direitos. Cada elemento *Dublin Core* é definido por meio do padrão formal de descrição, composto de dez atributos (Quadro 1), da ISO/IEC 11179 (International..., 2004).

Outra característica do *Dublin Core* é a utilização de qualificadores para diferenciar o valor de cada ocorrência de seus elementos. Os qualificadores são formados por esquema e/ou modificador, e ambos servem para informar aos usuários e aplicativos como interpretar o valor de cada ocorrência.

Exemplificando, no Sistema Agência, tem-se:

Para o elemento Assunto:

- Subelemento: Palavras-chave

Esquema: THESAGRO; NAL Agricultural Thesaurus; Termos Livres; Outros.

- Subelemento: Categoria

Esquema: NAL Agricultural Thesaurus; Outros.

Para o elemento Descrição:

Modificador: Resumo, Sumário, Texto livre.

O elemento Assunto é descrito por meio de palavras-chave, utilizando-se as opções de vocabulário controlado, apresentadas pelos Esquemas: THESAGRO, NAL Agricultural Thesaurus, Termos Livres ou Outros. Também se atribui a categoria correspondente ao assunto, extraída do NAL Agricultural Thesaurus. Já para o elemento Descrição, ao invés de esquema, têm-se os modificadores Resumo, Sumário e Texto livre, por meio dos quais é possível relatar o conteúdo do recurso.

## Descrição dos elementos de metadados

Os quinze elementos *Dublin Core*, com definição, qualificadores e valor de ocorrência ilimitada (Esquema e Modificador), são apresentados no Anexo 2.

A representação descritiva e temática de recursos de informação é realizada por meio do gestor de conteúdo do Sistema Agência, que possui duas ferramentas dedicadas exclusivamente a tais atividades. A primeira é a de catalogação, denominada Recursos, e a outra é o Banco de Termos Autorizados (BTA). A ferramenta Recurso foi amplamente detalhada por Souza e Alves (2009), e por essa razão será abordada de forma superficial neste artigo. Portanto neste trabalho apenas a ferramenta BTA será objeto de descrição, com ênfase nos aspectos relativos ao controle de vocabulário (processo documentário) e à adoção de vocabulário controlado (tesauros – instrumento/produto documentário) no Sistema Agência.

## Termos livres na catalogação: vocabulário controlado

A utilização de termos livres para entrada de nomes pessoais e corporativos, bem como para a descrição de afiliação, endereços pessoais e corporativos é feita mediante o registro, normalização e o controle de vocabulário. Dessa maneira, ao descrever os elementos Criador, Publicador e Colaborador, o catalogador deve inserir obrigatoriamente o registro de termos livres no BTA, padronizando as suas entradas e respectivos conteúdos de afiliação e endereços.

**Quadro 1.** Atributos de elementos Dublin Core.

| Nome              | O rótulo atribuído para o elemento de dado.   |
|-------------------|---|
| Identificador     | O único identificador atribuído para o elemento de dado.  |
| Versão            | Versão do elemento de dado.   |
| Autoridade        | A entidade autorizada a fazer o registro do elemento de dado.                                   |
| Idioma            | O idioma no qual o elemento de dado está especificado.  |
| Definição         | Uma declaração que claramente represente o conceito e a natureza essencial do elemento de dado. |
| Obrigatoriedade   | Indica se o elemento de dado necessita estar presente sempre ou esporadicamente.                |
| Tipo de dado      | Indica o tipo de dado que pode ser representado no valor do elemento de dado.                   |
| Ocorrência máxima | Indica qualquer limite para a repetibilidade do elemento de dado.                               |
| Comentários       | Um comentário relativo à aplicação do elemento de dado.   |

## REPRESENTAÇÃO TEMÁTICA

A recuperação de informação no Sistema Agência depende fundamentalmente de uma boa indexação de assuntos contidos nos recursos eletrônicos associados aos nós e subnós das árvores de conhecimento. A propósito, de acordo com Nunes (2004, p.59),

é conveniente jamais esquecer que a indexação de um documento é feita uma única vez, enquanto que a recuperação da informação contida no documento é feita inúmeras vezes, do que se conclui que tempo “gasto”, na indexação é, de fato, economia de tempo para inúmeros usuários no futuro.

A ausência de tratamento terminológico por meio da adoção de ferramentas e procedimentos de controle de vocabulário tem sido a causa frequente de baixa revocação e baixa precisão na recuperação da informação, observou Souza (2004), ao analisar a necessidade de vocabulário controlado no campo do Direito do Trabalho.

A indexação no Sistema Agência corresponde à representação do conteúdo temático de um determinado recurso eletrônico (vídeo, áudio, texto etc.), mediante a determinação do assunto de que trata esse recurso, a seleção dos conceitos relevantes associados ao assunto determinado e à tradução desses conceitos para os termos autorizados de um tesauro.

Souza e Alves (2001), baseados em Lancaster (1993), ao abordarem o perfil profissional do bibliotecário para atuar no Sistema Agência, propõem ao indexador a formulação das seguintes perguntas:

- a) De que trata o recurso que está sendo catalogado? Ou seja, qual o assunto?
- b) Por que está sendo incorporado ao repositório do Sistema Agência?
- c) Quais de seus aspectos são de interesse para os usuários da Agência?

A atividade de análise de conteúdos dos recursos de informação eletrônicos no Sistema Agência é realizada com o emprego de linguagem artificial por meio de tesauros, visando ao controle do vocabulário e à busca de maior consistência na indexação e na recuperação.

A necessidade de adoção de vocabulário controlado é apontada por Cesarino e Pinto (1978, p. 272), em razão de diferenças culturais e de domínio terminológico, relacionados aos aspectos humanos, os quais podem alterar a percepção do indexador e do usuário em relação ao texto. Outro indicador da necessidade de controle de vocabulário refere-se aos aspectos linguísticos, como sinonímia, homônimos e sintaxe. A falta de controle de vocabulário, em outras palavras, o não uso de instrumentos de terminologia, como os tesauros, causa vários problemas que certamente afetam a recuperação da informação (Catarino; Baptista, 2007).

A representação temática no Sistema Agência é parte integrante da metodologia de organização e tratamento da informação, desenvolvida e adotada pela Embrapa (Alves *et al.*, 2008). A atividade de indexação é realizada mediante o controle de vocabulário para atribuição de palavras-chave e de categorias de assunto, apoiada pela adoção de dois tesauros: THESAGRO e NAL Agricultural Thesaurus.

### Termos livres na indexação

A adoção e o controle de termos livres na indexação são justificados por Nunes (2004, p.59), com argumentos que se aplicam perfeitamente ao contexto da indexação do Sistema Agência:

Sobre sua adoção, é preciso lembrar que estes termos são “livres” no sentido de que não integram uma linguagem de indexação institucionalizada, porém isto não significa que possam ser lançados como entradas de assunto sem qualquer preocupação quanto ao seu controle, ao livre arbítrio do bibliotecário, caso a caso. Ao contrário, os “termos livres” devem ser objeto de controle por parte do bibliotecário, através do catálogo decisório, de inclusões na linguagem documentária adotada ou do estabelecimento de uma lista de “termos livres”, sujeita aos procedimentos normais de controle.

No Sistema Agência utiliza-se o BTA para o registro, normalização e gerenciamento de palavras-chave correspondentes ao elemento Assunto. Os termos livres são registrados sob a forma de substantivo ou frase substantivada no singular, sem o emprego de adjetivos e verbos. Apresentam uma definição sucinta que expressa o seu conteúdo, e, ainda, trazem uma

indicação da fonte de consulta do termo, na forma de uma referência bibliográfica, ou o nome do pesquisador que sugeriu o termo.

### Termos livres: cobertura

Termos livres também são atribuídos ao elemento Cobertura (Modificador - Espacial), que compreende os aspectos geográficos e de localização espacial de um determinado recurso de informação. Assim como os termos livres palavras-chave, os termos livres Cobertura (Espacial) também devem ser registrados, normalizados e gerenciados pelo BTA.

### Termos tesouros: THESAGRO e NAL Agricultural Thesaurus

São duas as linguagens de indexação adotadas no Sistema Agência para atribuição de palavras-chave: THESAGRO e NAL Agricultural Thesaurus. Ambos os tesouros abrangem o domínio agrícola, com ênfase nas ciências biológicas (animal e vegetal), físicas e sociais.

O Thesagro – Thesaurus Agrícola Nacional - é um tesouro brasileiro especializado em literatura agrícola, desenvolvido pela Biblioteca Nacional de Agricultura (Brasil, 2006), do Ministério da Agricultura, Pecuária e Abastecimento. Sua estrutura contempla as relações de hierarquia, associação e equivalência. Possui 9351 termos, para os quais não existem categorias de assunto. Está disponível na internet no endereço: <[http://www.agricultura.gov.br/portal/page?\\_pageid=33,959135&\\_dad=portal&\\_schema=PORTAL](http://www.agricultura.gov.br/portal/page?_pageid=33,959135&_dad=portal&_schema=PORTAL)>.

O NAL Agricultural Thesaurus é um tesouro norte-americano para indexação e categorização de assuntos de recursos de informação agrícola, produzido e mantido pela National Agricultural Library (2008). Possui mais de 70 mil termos (incluindo 28 mil referências cruzadas) organizados em 17 categorias de assunto. Os relacionamentos entre termos são de natureza hierárquica, de equivalência e associativa. Apresenta notas de escopo e definições. É bilíngue (espanhol e inglês). Traz glossário contendo coleção de definições de termos agrícolas criados pelo NAL Agricultural Thesaurus. Está disponível na web no endereço: <<http://agclass.nal.usda.gov/>>.

Tanto o THESAGRO como o NAL Agricultural Thesaurus estão integrados na base de dados do serviço web da ferramenta BTA (WS-BTA). O gestor de conteúdo do Sistema Agência utiliza as bases de dados dos dois tesouros, via acessos a esse serviço web, garantindo agilidade, confiabilidade e segurança no gerenciamento de vocabulário controlado inerente às atividades de indexação.

### Categorias de assunto: NAL Agricultural Thesaurus

Todos os recursos de informação eletrônicos associados a conteúdos de nós de árvore de conhecimento no Sistema Agência são catalogados, indexados e classificados por categorias de assunto, utilizando o NAL Agricultural Thesaurus. O amplo conjunto de termos palavras-chave desse tesouro está distribuído em 17 categorias, as quais cobrem o domínio agrícola e áreas correlatas. Tais categorias são as seguintes: *Animal Science and Animal Products; Biological Sciences; Breeding and Genetics Improvement; Economics, Business and Industry; Farms and Farming Systems; Food and Human Nutrition; Forest Science and Forest Products; Geographical Locations; Government, Law and Regulations; Health and Pathology; Insects and Entomology; Natural Resources, Earth and Environment; Physical and Chemical Sciences; Plant Science and Plant Products; Research, Technology and Engineering; Rural and Agricultural Sociology; Taxonomic Classification of Organisms.*

As 17 categorias de assunto NAL Agricultural Thesaurus foram traduzidas do inglês para o português, com o intuito de tornar mais abrangente a categorização dos recursos de informação eletrônicos no Sistema Agência.

### BTA: FERRAMENTA DE CONTROLE DE VOCABULÁRIO

O BTA constitui-se de uma aplicação que permite o registro, a consulta, a alteração e a exclusão de termos autorizados e termos não-autorizados. Possibilita, ainda, o registro de definições, notas, relacionamentos e fontes de informação que fundamentam a referida constituição do termo.

O BTA tem o propósito de eliminar redundâncias e inconsistências na descrição do recurso, reduzir

possibilidades de erro e oferecer maior facilidade operacional ao catalogador. E, para isso, os procedimentos de consulta, seleção e inserção de novos termos estão interligados ao formulário de catalogação.

Assim, ao descrever, na catalogação, os elementos Criador, Publicador, Colaborador, Assunto Palavra-chave, Assunto Categoria e Cobertura (Espacial), deve-se, obrigatoriamente, adotar os termos autorizados referentes a nomes pessoais e corporativos, endereços pessoais e corporativos, afiliação, palavras-chave, categoria e lugares ou localização geográfica controlados pelo BTA.

Portanto, se determinado termo a ser utilizado não se encontra registrado no BTA, deve-se, primeiramente, normalizar a sua forma de entrada, tomando por base as regras mínimas de catalogação, em especial para nomes pessoais e corporativos, os capítulos 20 e 22, respectivamente, do AACR2 (Código..., 2004). Em seguida, faz-se a inserção desse termo no BTA, para que fique disponível ao compartilhamento e possa ser selecionado, via ferramenta Recurso, no momento da catalogação.

### Implementação do banco de termos

O BTA foi concebido como um aplicativo integrado à ferramenta de catalogação do Sistema Gestor de Conteúdo da Agência de Informação Embrapa e composto por quatro formulários: Novo Termo (Anexo 3), Alterar, Excluir e Consultar Termo(s), permitindo ao catalogador inserir, consultar e/ou selecionar termos via formulário de catalogação.

O BTA foi projetado para atender especificamente à ferramenta de catalogação, nas suas necessidades de normalizar e gerenciar o uso de vocabulário controlado. Os pontos de acesso que requerem o uso de vocabulário controlado são:

- Assunto - Palavra-chave
- Assunto - Categoria
- Cobertura (Espacial)
- Nome pessoal
  - Endereço pessoal
  - Afiliação
- Nome corporativo
  - Endereço corporativo

Diante das necessidades e dos requisitos especificados, foram definidos os seguintes campos para o BTA:

- Termo (campo livre para digitação do termo a ser inserido)
- Idioma (a selecionar)
- Tipo de termo (a selecionar)
  - Assunto Palavra-chave
  - Assunto Categoria
  - Cobertura Espacial
  - Nome Pessoal
  - Nome Corporativo
- Esquema (a selecionar)
  - Thesagro
  - NAL Agric. Thes.
  - Termos livres
  - Outros
- Fonte
- Definição
- Notas
- Status de Autorização (Autorizado/Não Autorizado)
- Status do Registro (Em Análise/Validado)
- Data de Criação (dd/mm/aaaa)
- Data da última modificação (dd/mm/aaaa)
- Relacionamento entre termos (a selecionar)
  - Ver também
  - Use
  - Usado para
  - Tradução
  - Nome científico
  - Afiliação

Após o registro de um determinado termo, o nome do catalogador é inserido automaticamente, bem como é gerado um número identificador do registro efetuado.

O acesso ao BTA é controlado, de modo que apenas o profissional de informação/catalogador tem permissão para realizar consultas e inserir termos novos. No entanto apenas o profissional da informação/auditor está autorizado a fazer alterações e exclusões de termos. O profissional de informação/auditor é quem valida o BTA quanto à atualização dos dados, identificação de informações inconsistentes e modifica o status para *Validado* dos termos registrados e auditados, como mostra a (Anexo 4).

O BTA é uma ferramenta que propicia a atualização e o registro de novos termos, de forma dinâmica e continuada, tornando a atividade de controle terminológico eficaz e menos penosa. Porém vale lembrar que, quando a operacionalização do sistema é descentralizada, faz-se necessário instituir um comitê para avaliar os termos novos propostos, bem como deliberar sobre o uso de termos livres.

Considera-se que a integração de bases de autoridades com bases de recursos informacionais traz muitas vantagens, tanto para o profissional de informação como para o usuário final. Para o catalogador, oferece maior facilidade operacional, reduz possibilidades de erro, elimina redundâncias e inconsistências na descrição do recurso e, para o usuário, enriquece a busca e recuperação da informação.

O desenvolvimento do serviço web para o BTA (Berenguel *et al.*, 2008) deverá ser detalhado em trabalhos futuros. Pode-se adiantar, entretanto, que essa implementação do BTA tem respondido positivamente, interagindo com outros servidores de serviços web.

## CONSIDERAÇÕES FINAIS

A representação descritiva e temática são os alicerces do tratamento de recursos de informação eletrônicos no Sistema Agência. Ambas as atividades integram o conjunto de metodologias de organização da informação elaboradas pela Embrapa. A catalogação e a indexação de recursos eletrônicos são partes integrantes dessa política de organização e tratamento da informação estabelecida para o Sistema Agência. Requerem a sistematização e a observância de procedimentos padronizados e universais, bem como a existência de ferramenta informatizada de controle de vocabulário.

Consideram-se fatores importantes na implantação das atividades de representação descritiva e temática: o estabelecimento de regras mínimas; a adoção de padrões e formatos internacionais; a utilização de códigos e tesouros; a elaboração de manuais; a capacitação profissional e o desenvolvimento de ferramentas e aplicações para dar suporte às tais atividades. A reunião desses fatores tem contribuído para a construção de uma política eficaz de tratamento da informação na Embrapa.

## REFERÊNCIAS

AGUIAR, F.L. *O controle de vocabulário como dispositivo metodológico para a organização, tratamento e recuperação da informação documentária*. 2008. 267 f. Dissertação (Mestrado em Ciência da Informação) - Pontifícia Universidade Católica de Campinas, Campinas, 2008.

ALVES, M.D.R.; SOUZA, M.I.F.; SANTOS, A.D. *Manual de representação descritiva de recursos eletrônicos no Sistema Agência*. Campinas: Embrapa Informática Agropecuária, 2008. (Embrapa Informática Agropecuária. Documentos, 81). Disponível em: <<http://www.cnptia.embrapa.br/files/doc81.pdf>>. Acesso em: 5 out. 2009.

BERENGUEL, A.L.A.; QUEIROS, L.R.; SOUZA, M.I.F.; ALVES, M.D.R. Arquitetura AAA em sistemas web baseados em REST. *Global Science and Technology*, v.1, n.1, p.1-7, 2008. Disponível em: <<http://rioverde.ifgoiano.edu.br/periodicos/index.php/gst/article/view/10/3>>. Acesso em: 6 out. 2009.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Biblioteca Nacional de Agricultura. *Thesagro: thesaurus agrícola nacional*. Brasília, DF, 2006. Disponível em: <[http://www.agricultura.gov.br/porta1/page?\\_pageid=33,959135&\\_dad=porta1&\\_schema=PORTAL](http://www.agricultura.gov.br/porta1/page?_pageid=33,959135&_dad=porta1&_schema=PORTAL)>. Acesso em: 5 out. 2009.

CATARINO, M.E.; BAPTISTA, A.A. Folksonomia: um novo conceito para a organização dos recursos digitais na web.

*DataGramaZero - Revista de Ciência da Informação*, v.8, n.3, 2007. Disponível em: <[http://www.datagramazero.org.br/jun07/Art\\_04.htm](http://www.datagramazero.org.br/jun07/Art_04.htm)>. Acesso em: 5 out. 2009.

CESARINO, M.A.N.; PINTO, M.C.F. Cabeçalho de assunto como linguagem de indexação. *Revista da Escola de Biblioteconomia UFMG*, v. 7, n. 2, p.268-288, 1978.

COBBE, R.V. et al. *Recomendações para a elaboração do conteúdo Agência de Informação Embrapa*. Campinas: Embrapa Informática Agropecuária, 2007. 39p.

CÓDIGO de Catalogação Anglo-americano. 2. ed. rev. São Paulo: FEBAB, 2004.

DUBLIN CORE METADATA INITIATIVE. *Dublin Core metadata element set, version 1.1: reference description*. [S.l.], 2008. Available from: <<http://dublincore.org/documents/dces/>>. Cited: 20 Oct. 2009.

EMBRAPA. *V Plano Diretor da Embrapa: 2008-2011-2023*. Brasília: EMBRAPA, 2008. 44 p.

EMBRAPA INFORMÁTICA AGROPECUÁRIA. *Agência de Informação Embrapa: ambiente web para transferência de tecnologia*. Campinas: EMBRAPA, 2005.

EVANGELISTA, S. HiperEditor: editor árvore hiperbólico. In: REDE agrolive. *AgroLivre*: rede de software livre para agropecuária. Campinas: Embrapa Informática Agropecuária, 2005a. Disponível em: <<http://repositorio.agrolivre.gov.br/projects/hipereditor/>>. Acesso em: 20 out. 2009.

EVANGELISTA, S. HiperNavegador - navegador hiperbólico. In: REDE agrolive. *AgroLivre*: rede de software livre para agropecuária. Campinas: Embrapa Informática Agropecuária, 2005b. Disponível em: <<http://repositorio.agrolivre.gov.br/projects/hipernavegador/>>. Acesso em: 20 out. 2009.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. *Home page for ISO/IEC 11179*: information technology, metadata registries. 2<sup>nd</sup> ed. Washington, DC: ISO/IEC, 2004. Available from: <<http://metadata-standards.org/11179/>>. Cited: 25 Jun. 2009.

LANCASTER, F.W. *Indexação e resumos: teoria e prática*. Brasília: Briquet de Lemos, 1993.

NATIONAL AGRICULTURAL LIBRARY (Estados Unidos). *Agricultural thesaurus and glossary*. 2008. Available from: <<http://agclass.nal.usda.gov/agt/agt.shtml>>. Cited: 30 Oct. 2009.

NUNES, C.O.I. Algumas considerações acerca da ausência de políticas de indexação em bibliotecas brasileiras. *Biblos*, v. 16, p. 55-61, 2004. Disponível em: <<http://www.seer.furg.br/ojs/index.php/biblos/article/view/412/92>>. Acesso em: 30 out. 2009.

SANTOS, A.D.; et al. *Manual de usuário do Sistema Gestor de Conteúdo da Agência de Informação Embrapa - versão 2*: módulo administrativo. Campinas: Embrapa Informática Agropecuária, 2007a. v.1. Disponível em: <<http://agenciahome.cnptia.embrapa.br/documentos/docsBibliografia/modulo1.pdf>>. Acesso em: 30 out. 2009.

SANTOS, A.D.; et al. *Manual de usuário do Sistema Gestor de Conteúdo da Agência de Informação Embrapa - versão 2*: módulo de manipulação de árvore do conhecimento. Campi-

nas: Embrapa Informática Agropecuária, 2007b. v. 2. Disponível em: <<http://agenciahome.cnptia.embrapa.br/documentos/docsBibliografia/modulo2.pdf>>. Acesso em: 30 out. 2009.

SANTOS, A.D.; et al. *Manual de usuário do Sistema Gestor de Conteúdo da Agência de Informação Embrapa - versão 2*: módulo de recursos e banco de termos autorizados. Campinas: Embrapa Informática Agropecuária, 2007c. v. 3. Disponível em: <<http://agenciahome.cnptia.embrapa.br/documentos/docsBibliografia/modulo3.pdf>>. Acesso em: 30 out. 2009.

SOUZA, C.F. *Vocabulário controlado em direito do trabalho: o caso da base de dados "REV"*. 2004. 109 f. Trabalho de Conclusão de Curso - Pós-Graduação em Organização do Conhecimento para a Recuperação da Informação, Universidade Federal do Estado do Rio de Janeiro.

SOUZA, M.I.F.; ALVES, M.D.R. *Atuação do bibliotecário no Sistema Agência de Informação Embrapa: perfil profissional*. Campinas: Embrapa Informática Agropecuária, 2001. 4 p. (Embrapa Informática Agropecuária. Comunicado Técnico, 7). Disponível em: <<http://www.cnptia.embrapa.br/files/COMNICADOTEC7.pdf>>. Acesso em: 5 out. 2009.

SOUZA, M.I.F.; ALVES, M.D.R. Representação descritiva e temática de recursos de informação no Sistema Agência Embrapa: uso do padrão Dublin Core. *Revista Digital de Biblioteconomia e Ciência da Informação*, v. 7, n. 1, p. 208-223, 2009. Disponível em: <<http://www.infoteca.cnptia.embrapa.br/handle/item/13656>>. Acesso em: 30 out. 2009.


SOUZA, M.I.F.; et al. *Agência de Informação Embrapa: uma aplicação para a organização e gestão do conhecimento*. In: WORKSHOP DE BIBLIOTECAS DIGITAIS, 2., 2006, Florianópolis. *Anais...* Florianópolis: Sociedade Brasileira de Computação, 2006. p. 51-56.

WEIBEL, S.; et al. RFC 2413 – *Dublin Core Metadata for resource discovery*. 1998. Available from: <<http://www.ietf.org/rfc/rfc2413.txt>>. Cited: 12 Oct. 2009.

## ANEXO 1

### CONTEÚDO DE NÓ (NÍVEL MAIS GENÉRICO) E SUBNÓS (NÍVEIS MAIS ESPECÍFICOS)

Ministério da Agricultura, Pecuária e Abastecimento
Destaques do Governo



Agência de Informação Embrapa
Cana-de-açúcar

Página Inicial
Fale Conosco
Ajuda

Pré-produção Socioeconomia

Socioeconomia

Nó

Autor(es): Carlos Eduardo Freitas Vianna

Buscar

Busca Avançada

Hiperbólica

Pasta de documentos

- Adicionar
- Visualizar

Informes

- Agritempo
- Embrapa Informática Agropecuária
- IEA - Banco de Dados de Bioenergia
- Mapa - Estatísticas
- Ridesa
- Unica

Subnós

O Brasil é o maior produtor mundial de cana-de-açúcar, uma das mais importantes matérias-primas para a fabricação de etanol, o que coloca o País numa posição privilegiada no mercado mundial de biocombustíveis.

No ano de 2005, o governo brasileiro lançou o Plano Nacional de Agroenergia, em conjunto com a iniciativa privada e instituições públicas de pesquisa, com o intuito de direcionar e criar diretrizes para a política de produção e comercialização de agroenergia para o período de 2006 a 2011.

Com o programa de agroenergia, o agronegócio brasileiro tem a possibilidade de aumentar sua importância no cenário mundial.

Serão abordados, nos próximos textos, aspectos de mercado da cana-de-açúcar, estatísticas de produção, políticas e legislação do setor sucroalcooleiro.

Políticas e legislação

Este tópico trata do histórico das principais políticas e leis aplicadas à produção e comercialização de cana-de-açúcar no Brasil

Estatísticas

Este tópico aborda as potencialidades da cana-de-açúcar, suas vantagens competitivas nos âmbitos regional, nacional e internacional e, ainda, apresenta uma comparação das produções canavieiras no plano mundial e nacional

Mercado

Este tópico apresenta características do agronegócio envolvendo a cana-de-açúcar. O Agronegócio conta com significativa importância em nível nacional, visto que o Brasil lidera os segmentos de cana, açúcar e álcool

Informações Complementares:

Uma análise da cadeia produtiva de cana-de-açúcar na Região Norte Fluminense

Introdução; Panorama do setor sucroalcooleiro nacional; O setor sucroalcooleiro no Brasil e sua relação com o Estado; Resumo histórico do setor canavieiro brasileiro e a intervenção do Estado; Resumo histórico da agroindústria no Estado do Rio de Janeiro; Panorama do setor sucroalcooleiro no Estado

Mais Detalhes

Balanço nacional da cana-de-açúcar e agroenergia

Cana-de-açúcar; Álcool; Açúcar; Oleaginosas; Soja; Dendê; Girassol; Algodão; Amendoim; Mamona; Outras oleaginosas; Leilões e biodiesel

Mais Detalhes

Açúcar e álcool no Brasil

O presente documento traz o mapa das regiões produtoras de açúcar e álcool no Brasil. Traz, também, os principais Estados produtores de cana-de-açúcar, e apresenta, ainda, o panorama do setor sucroalcooleiro no País e os comparativos das safras de produção de cana-de-açúcar, açúcar e álcool nos últimos anos

Mais Detalhes

Recursos de informação

Empresa Brasileira de Pesquisa Agropecuária - Embrapa

Todos os direitos reservados, conforme Lei nº 9.610.

Política de Privacidade, sac@embrapa.br

2005-2007

Topo

Página Inicial

Voltar

Embrapa

Parque Estação Biológica - PqEB s/nº,

Brasília, DF - Brasil - CEP 70770-901

Fone: (61)3448-4433 - Fax: (61)3347-1041

## ANEXO 2

## ELEMENTOS DUBLIN CORE

| Título (obrigatório)   |   |
|------------------------|---|
| Definição              | Nome dado ao recurso por meio do qual ele é formalmente conhecido; pode ser palavra, frase, caractere, ou grupo de caracteres.  |
| Qualificadores e valor | <b>Modificador:</b> Nenhum, Principal, Alternativo, Traduzido, Série<br>O Modificador Nenhum é o valor padrão para Título.  |
| Criador                |   |
| Definição              | Pessoa, entidade ou serviço originalmente responsável pela criação intelectual ou artística de um recurso.  |
| Qualificadores e valor | <b>Modificador:</b> Nenhum, Nome pessoal, Endereço pessoal, Afiliação, Nome corporativo, Endereço corporativo<br>O Modificador Nenhum é o valor padrão para o Criador.                  |
| Descrição              |   |
| Definição              | É um relato do conteúdo do recurso.   |
| Qualificadores e valor | <b>Modificador:</b> Resumo, Sumário, Texto livre  |
| Assunto (obrigatório)  |   |
| Definição              | Designa o tema do conteúdo do recurso, o que permite qualificar e agregar valor à informação.   |
| Qualificadores e valor | Subelemento Palavra-chave<br><b>Esquema:</b> THESAGRO, NAL Agricultural Thesaurus, Termos livres, Outros<br>Subelemento Categoria<br><b>Esquema:</b> NAL Agricultural Thesaurus, Outros |
| Cobertura              |   |
| Definição              | Extensão ou alvo do conteúdo do recurso.  |
| Qualificadores e valor | <b>Esquema:</b> Texto livre, Date and Times Formats, Getty Thesaurus of Geographical Names<br><b>Modificador:</b> Espacial, Temporal  |
| Publicador             |   |
| Definição              | Pessoa ou entidade responsável por tornar o recurso disponível na forma em que se apresenta.  |
| Qualificadores e valor | <b>Modificador:</b> Nenhum, Nome pessoal, Endereço pessoal, Afiliação, Nome corporativo, Endereço corporativo<br>O Modificador Nenhum é o valor padrão para o Publicador.               |

| Colaborador            |  |
|------------------------|--|
| Definição              | Pessoa ou entidade responsável pela contribuição intelectual do recurso.   |
| Qualificadores e valor | <b>Modificador:</b> Nenhum, Nome pessoal, Endereço pessoal, Afiliação, Nome corporativo, Endereço corporativo<br>O Modificador Nenhum é o valor padrão para o Colaborador. |

| Tipo                   |  |
|------------------------|--|
| Definição              | A natureza ou gênero do recurso.   |
| Qualificadores e valor | <b>Modificador:</b> Texto, Imagem, Som, Dados, Software, Interativo, Evento, Objeto físico, Coleção, Modelo, Participante, Lugar, Serviço, Outros. |

| Identificador          |  |
|------------------------|--|
| Definição              | Uma referência não-ambígua para o recurso em um dado contexto. |
| Qualificadores e valor | <b>Esquema:</b> Nenhum, URI, ISBN, ISSN, FPI                   |

| Data                   |   |
|------------------------|---|
| Definição              | Uma data associada a um evento no ciclo de vida do recurso. Essa data está relacionada com a criação ou disponibilidade do recurso.   |
| Qualificadores e valor | <b>Esquema:</b> Nenhum, AAAA-MM-DD, AAAA-MM, AAAA<br><b>Modificador:</b> Data de criação, Data da última modificação, Data aproximada |

| Formato                |   |
|------------------------|---|
| Definição              | A manifestação física ou digital do recurso.  |
| Qualificadores e valor | <b>Modificador:</b> Texto, Aplicação, Imagem, Multipartes, Áudio, Vídeo, Modelos, Mensagem, Outro |

| Idioma                 |                                   |
|------------------------|-----------------------------------|
| Definição              | O idioma do conteúdo do recurso.  |
| Qualificadores e valor | <b>Esquema:</b> RFC 1766 e ISO639 |

| Fonte                  |  |
|------------------------|--|
| Definição              | Uma referência a um recurso do qual o presente recurso é derivado. |
| Qualificadores e valor | Texto livre, URI, ISBN, ISSN, FPI                                  |

| Relação                |  |
|------------------------|--|
| Definição              | Uma referência para um recurso relacionado.  |
| Qualificadores e valor | <b>Esquema:</b> Texto livre, URI, ISBN, ISSN, FPI<br><b>Modificador:</b> Parte de, Possui parte, Versão de, Possui versão, Formato de, Possui formato, Referências, Referenciado por, Baseado em, Base para, Requer, Requerido por, Continua em, Continuação de. |

| Direitos               |   |
|------------------------|---|
| Definição              | Informação sobre direitos assegurados dentro e sobre o recurso. |
| Qualificadores e valor | <b>Esquema:</b> Texto livre, URI                                |

## ANEXO 3


### BTA: FORMULÁRIO NOVO TERMO (INSERIR TERMO)

The screenshot displays the 'Gestor de Conteúdo' web application. At the top, the logo 'Gestor de Conteúdo Versão 2.0' is visible, along with user information: 'malves', 'Equipe Editorial: Cana-de-Açúcar', and 'Perfil: Auditor'. Below this is a navigation bar with links: 'Usuário', 'Equipe Editorial', 'Árvore', 'Nó', 'Recurso', 'Banco de Termos', 'Ajuda', and 'Sair'. The 'Banco de Termos' menu is open, showing options: 'Novo termo', 'Alterar termo(s)', 'Excluir termo(s)', and 'Consultar termo(s)'. The main content area is titled 'Banco de Termos > Novo termo' and contains a form with the following fields:

- Termo:\*** (Text input field)
- Idioma:** (Dropdown menu with 'Nenhum' selected)
- Tipo de termo:\*** (Dropdown menu with 'Escolha' selected)
- Esquema:** (Dropdown menu with 'Escolha um tipo de termo' selected)
- Fonte:** (Text input field)
- Definição:** (Text input field)
- Notas:** (Text input field)
- Status da autorização:** (Dropdown menu with 'Não Autorizado' selected)
- Status do registro:** (Dropdown menu with 'Validado' selected)
- Data de criação:** (Text input field with value '28/10/2009')
- Data da última modificação:** (Text input field with value '28/10/2009')

At the bottom left, there is a note: '\*Campo obrigatório'. At the bottom right, there is a 'Salvar' button.

## TERMO REGISTRADO E AUDITADO



**Gestor de Conteúdo**  
Versão 2.0

Sistema Gestor de Conteúdo da Agência de Informação Embrapa

malves  
Equipe Editorial: Cana-de-Açúcar  
Perfil: Profissional de Informação

Usuário   Equipe Editorial   Árvore   Nó   Recurso   Banco de Termos   Ajuda   Sair

Banco de Termos > Consulta termo(s) > Dados do termo

Tipo de termo:  
Assunto Palavra Chave

Termo:

|                         |                              |
|-------------------------|------------------------------|
| Cerrado                 |                              |
| Idioma                  | Português                    |
| Registro                | 00000189                     |
| Autorização             | Autorizado                   |
| Catalogador Responsável | Marcia Izabel Fugisawa Souza |

Categoria:

|                        |  |
|------------------------|--|
| NAL Agrícola Thesaurus |  |
| Termo                  | P Recursos Naturais, Terra e Ciências Ambientais |
| Idioma                 | Português  |

Definição:

Notas:

Estado do registro:  
Validado

Datas:

|                    |            |
|--------------------|------------|
| Criação            | 24/09/2007 |
| Última modificação | 11/07/2008 |

Relacionamentos:

|             |                    |
|-------------|--------------------|
| Usado Para: | Cerradão<br>Savana |
|-------------|--------------------|