

ESTUDO SOBRE ESTRUTURA DEFINITÓRIA PARA DESENVOLVIMENTO DE ONTOLOGIAS

Lígia Café *
Fernanda Mendes **

RESUMO

Investiga-se a identificação da tipologia de relacionamentos semânticos existentes no domínio de Análise do Ciclo de Vida de Produtos (ACV) do Projeto de Inventário do Ciclo de Vida para a Competitividade Ambiental da Indústria Brasileira, gerenciado pelo Instituto Brasileiro de Ciência e Tecnologia (IBICT). Para tanto, constrói-se uma estrutura definitória com base na metodologia para desenvolvimento de ontologia descrita por Bräscher (2007) e nas orientações descritas por Uschold (1996) e Uschold e King (1995) e Uschold et al. (1996). Testa-se a validade desta estrutura por meio da aplicação do método experimental num *corpus* formado por 231 termos pertencentes a este domínio com o intuito de contribuir para o desenvolvimento de ontologias, bem como glossários e dicionários especializados. Pela adoção da estrutura em questão, é possível identificar a categoria, tipo, definição, fonte, propriedade, nota, sinônimo, exemplo e termo limite dos termos coletados. Observou-se na pesquisa, que a estrutura segue um padrão de classificação bastante flexível, podendo ser aplicada a qualquer área do conhecimento. Conclui-se que a estrutura definitória contempla a complexidade das relações semânticas envolvidas na conceitualização, sendo a mesma válida para a construção de ontologias.

Palavras-chave: Estrutura definitória, Análise do ciclo de vida (ACV), Ontologia, Web semântica.

* Docente do Programa de Pós-Graduação em Ciência da Informação da UFSC. Mestre em Biblioteconomia e Documentação e Doutora em Linguística

** Graduada em Letras-Português e Literatura Brasileira (UFSC), Mestranda em Linguística na UFSC

I INTRODUÇÃO

Os desenvolvimentos mais recentes da World Wide Web levaram ao surgimento da Web semântica, uma extensão da Web atual em que os computadores entendem os significados das “coisas”, oferecendo respostas às questões dos usuários de forma automática. Segundo Berners-Lee et al. (2001), “A Web Semântica vai trazer estrutura ao conteúdo significativo das páginas na Web, criando um ambiente em que os agentes de software [...] possam prontamente executar tarefas sofisticadas para os usuários”. Para que este fenômeno seja possível, é fundamental que haja uma representação do conhecimento do mundo real.

Esta representação utiliza linguagens específicas para formar as bases de conhecimento cujo conteúdo deve ter significados bem definidos, uma vez que as ambigüidades semânticas e lexicais são características intrínsecas da língua, porém prejudiciais a recuperação da informação. No contexto da Web Semântica, estas linguagens são as ontologias compostas por uma taxonomia e um conjunto de regras de inferência que, de forma conjunta, possibilitam que o conhecimento seja usado automaticamente por diferentes aplicações.

No plano da significação, Guarino e Giaretta (1995) apresentam diversos sentidos para o termo ontologia: disciplina filosófica, sistema conceitual informal, consideração semântica

formal, especificação de uma conceitualização, representação de um sistema conceitual por meio da teoria lógica, vocabulário utilizado por uma teoria lógica e especificação em nível meta de uma teoria lógica. Uschold (1996), baseando-se nas noções definidas por Guarino e Giaretta (1995), ressalta a importância da idéia de conceitualização no sentido de visão de mundo que corresponde “a uma maneira de pensar sobre algum domínio [...], tipicamente concebida e/ou expressa como um conjunto de conceitos (p.ex. entidades, atributos, processos), suas definições e seus interrelacionamentos” (USCHOLD, 1996, p.3). Para este autor, uma ontologia é uma representação de alguma parte de uma conceitualização. É também nesse sentido que Gruber (1993) já havia definido ontologia referindo-se a uma especificação de uma conceitualização.

Considerando que uma ontologia inclui um conjunto de termos e conceitos, Uschold (1996) alerta para o que ele considera a principal atividade na sua elaboração: “a produção de definições”. Tais definições devem ser construídas com base em uma estrutura definitória que contemple a complexidade das relações semânticas envolvidas na conceitualização. Esta é a questão que envolve esta pesquisa, focando como unidade de análise a ontologia a ser desenvolvida na área de Análise do Ciclo de Vida de Produtos (ACV) do Projeto de Inventário do Ciclo de Vida para a Competitividade Ambiental da Indústria Brasileira, gerenciado pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).

Segundo Rey (1988, p.14), existem diferentes abordagens para determinar os diversos tipos de definições. Para este autor, podemos nos referir à definição do tipo descritiva que “respeita o uso na comunidade que a emprega” ou a construtiva “cujo objetivo é de construir cientificamente ‘uma língua bem feita’ [...] ou, no plano estético, de dar um sentido mais puro as palavras da tribo”. No que se refere ao seu objeto de análise, a definição é domínio da lexicografia quando o seu o foco são as palavras, enquanto que, no âmbito das línguas de especialidade (aquelas das áreas do conhecimento), a definição é objeto da terminografia, cujo interesse são os termos. De acordo com Bessé (1998, p. 255), “para que uma palavra tenha o direito ao título de termo, será necessário que ela possa, como elemento de

um conjunto (uma terminologia), ser distinguida das outras. A única maneira para expressar este sistema de distinções recíprocas é a operação dita definição”. Sendo o universo desta pesquisa delimitado pela terminologia pertencente ao domínio do Ciclo de Vida dos Produtos para a Competitividade Ambiental da Indústria Brasileira, estamos no contexto terminológico e, portanto, a estrutura definitória reflete as características deste tipo de definição. Desta forma, a definição terminológica segue um caminho onomasiológico, partindo do conceito ao termo; é sincrônica; possui um usuário especializado em uma área do conhecimento; caracteriza-se por ter um texto dominado pelo aspecto semântico incluindo uma rede de remissivas e um contexto com a função de identificar, completar e atestar características conceituais do termo; e representa a diversidade semântica e lexical da área. Por fim, como ressaltam Krieger e Finatto (2004, p. 175), “a definição [terminográfica] e seu entorno são modelados por exigências do acesso à informação e condicionados por circunstâncias comunicativas e socioculturais particulares das diferentes áreas do conhecimento”.

Este estudo insere-se na linha de pesquisa Fluxos de Informação do Programa de Pós-Graduação do Departamento de Ciência da Informação e na linha de pesquisa Comunicação Científica do Grupo de Pesquisa Núcleo de Estudos em Informação e Mediações Comunicacionais Contemporâneas¹, na medida em que tem por produto uma estrutura definitória que apóie o desenvolvimento de ontologias, instrumentos aplicados a Web Semântica essenciais a transferência de informações e intercâmbio entre especialistas de diferentes disciplinas.

O principal resultado deste trabalho é o padrão de estrutura definitória para elaboração de ontologia. Apesar de este padrão ser elaborado especificamente no âmbito do Projeto de Inventário do Ciclo de Vida para a Competitividade Ambiental da Indústria Brasileira, ele pretende ser aplicável na construção de outras ontologias assim como ser útil na elaboração de glossário e dicionários especializados.

Quanto ao aspecto de acesso à informação, a adoção do esquema definitório levará a

¹ Diretório de Grupos de Pesquisa no Brasil - CNPq

construção de um modelo de representação do conhecimento consensual e compartilhado da área de Análise do Ciclo de Vida (ACV), o que garantirá o uso de uma linguagem comum na representação e recuperação da informação. Segundo Chehebe (1997, p.10), Análise do Ciclo de Vida (ACV) “é a técnica para avaliação dos aspectos ambientais e dos impactos potenciais associados a um produto, compreendendo etapas que vão desde a retirada da natureza das matérias-primas elementares que entram no sistema produtivo (berço) à disposição do produto final (túmulo)[...]”. A modelagem do conhecimento em ACV, construída com base em uma estrutura definitória que reflita a diversidade lexico-conceitual da área, contribui para a troca de experiências entre especialistas favorecendo o desenvolvimento da área.

Além disso, a aplicação, na organização da informação, de ontologias derivadas deste padrão definitório poderá favorecer o acesso eficiente aos conteúdos e bons resultados na transferência da informação entre os especialistas da Indústria Brasileira, preocupados com a redução dos impactos ambientais causados pela indústria.

2 PROCEDIMENTOS METODOLÓGICOS

A presente pesquisa seguiu o método experimental aplicado no total de 231 termos pertencentes ao domínio da ACV, adotando os princípios da socioterminologia que prevê o estudo lexical descritivo de uma área do conhecimento levando em consideração o uso

social da língua. Desta forma, os termos, suas variantes lexicais e semânticas foram analisados no meio social ao qual pertencem. Portanto, foram consideradas as comunicações escritas dos especialistas do campo da ACV.

Tomando como ponto de partida a “Metodologia para desenvolvimento de ontologia no âmbito do Projeto de Inventário do Ciclo de Vida para a Competitividade Ambiental da Indústria Brasileira”, elaborada por Bräscher (2007), a estrutura definitória foi determinada com base nas orientações descritas por Uschold (1996), Uschold e King (1995) e Uschold et al. (1996), assim como nos tipos de relacionamentos existentes em ontologias e possíveis de serem aplicados no *software* Protégé, indicado para ser utilizado na construção desta ontologia. Estes relacionamentos são formados por *classes* e *subclasses*, que descrevem conceitos mais gerais e mais específicos organizados de forma taxonômica, refletindo a área objeto da ontologia; *funções* e *propriedades* que descrevem características e atributos dos conceitos, das classes e das instâncias, assim como as restrições também denominadas de *facetas*.

Uma vez determinada a estrutura da definição, esta foi aplicada aos conceitos básicos sobre ACV, contando para isto com o apoio de especialistas e fontes da área.

Com a intenção de sistematizar os itens da estrutura definitória, foi elaborada uma base de dados no programa *Access* contendo dez aspectos, acerca dos quais se discorrerá pontualmente logo após a apresentação dos mesmos na Figura 1 abaixo.

Figura 1: Formulário do Access.:

- 1 Termo: espaço reservado para a entrada do termo. Por exemplo: Acidificação, Entrada, Unidade de processo.
- 2 Categoria: espaço que indica a categoria a qual o termo pertence. Por exemplo: Objetivo e escopo, Interpretação, Análise do inventário do ciclo de vida.
- 3 Tipo de termo: espaço que indica qual o tipo do termo. Por exemplo: Entidade, Entidade Agente e Processo.
- 4 Definição/ Contexto: espaço destinado para a definição ou contexto do termo.
- 5 Fonte da definição²: espaço reservado para a fonte bibliográfica da qual se extraiu a definição ou contexto do termo. Por exemplo: CHE: 65; 14001: 02; 14041: 01³.
- 6 Propriedades: espaço garantido para as propriedades do termo.
- 6.1. ACV propriedade: espaço destinado às propriedades extraídas da definição ou contexto do termo.
- 6.2. ACV natureza da propriedade: espaço reservado à natureza das propriedades extraídas do termo. Por exemplo: É um, Tem por função, Tem por restrição.
- 7 Notas: espaço destinado às notas relativas ao termo.
- 8 Exemplos: espaço reservado aos exemplos do termo.
- 9 Sinônimos: espaço garantido aos sinônimos do termo.
- 10 Termos limites: espaço reservado aos termos limites do termo em questão.

É importante ressaltar que a base de dados foi configurada desta maneira tendo em vista o *software* Protégé, como foi acima tratado, e que, por essa razão, há uma relação íntima entre os

espaços destinados na base e no *software*. Por exemplo, o espaço destinado à categoria, na base de dados, corresponde ao espaço reservado à classe no Protégé; bem como o espaço reservado ao tipo de termo, na base de dados, é equivalente ao espaço destinado às funções no *software* e o espaço reservado às propriedades, notas, exemplos, sinônimos e termos limites, na base de dados, corresponde ao espaço destinado às propriedades no Protégé.

3 ANÁLISE DOS DADOS

Assim como foi acima citado, foram coletados 231 termos pertencentes ao domínio da ACV. Destes, alguns foram enviados por um especialista de renome da área⁴ e outros foram coletados diretamente das fontes. Neste último caso, utilizou-se o livro de Chehebe (1997) e as normas NBR ISO 14001, 14011, 14012, 14040, 14041, 14042 e 14043, ISO/TR 14025 e 14049 e ISO/TS 14048. Além disso, é importante frisar que desses 231 termos, não foi possível encontrar a definição de 53; logo, foram definidos, de fato, 178 termos, ou seja, 77,05% do total de termos.

Em relação às categorias, foram adotadas, a princípio, as quatro definidas na Metodologia (Bräscher, 2007):

- 1) Objetivo e escopo;
- 2) Análise do inventário do ciclo de vida;
- 3) Avaliação do impacto do ciclo de vida e
- 4) Interpretação do ciclo de vida.

Entretanto, estas não se apresentaram suficientes para classificar todos os termos, até mesmo porque muitos termos foram coletados além das normas NBR ISO 14040, 14041, 14042 e 14043, que são as que tratam, especificamente, dessas quatro categorias.

Logo, foram sugeridas outras duas categorias: 1) Geral e 2) Relatório, análise e revisão crítica; sendo que a primeira trata, mormente, dos termos coletados das normas NBR ISO 14001, 14011, 14012 e ISO/TS 14048; enquanto a segunda categoria trata, principalmente, dos termos coletados no último capítulo do livro de Chehebe (1997). Essas duas novas categorias estão identificadas da seguinte maneira na base

² Está organizado da seguinte maneira: referência bibliográfica abreviada: n° da página, n° da página; referência bibliográfica abreviada: n° da página; e assim por diante.

³ As referências bibliográficas foram abreviadas com o intuito de fazer com que ficasse mais simples no formulário do Access. A correspondência é a seguinte (para a referência completa, consultar a seção de Referências ao final do artigo):

CHE = CHEHEBE, José Ribamar Brasil.

14001 = ABNT - NBR. ISO 14001.

14011 = ABNT - NBR. ISO 14011.

14012 = ABNT - NBR. ISO 14012.

14025 = ISO/TR 14025.

14040 = ABNT - NBR. ISO 14040.

14041 = ABNT - NBR. ISO 14041.

14042 = ABNT - NBR. ISO 14042.

14043 = ABNT - NBR. ISO 14043.

14048 = ISO/TS 14048.

⁴ Professor Dr. Gil Anderi da Silva (USP).

de dados: i) S/C (sem categoria) Geral e ii) S/C (sem categoria) RAR (relatório, análise e revisão crítica).

Para uma análise quantitativa, é interessante perceber que S/C Geral foi a categoria que classificou a maior parte dos termos, totalizando 53 (29,77%), em segundo lugar ficou Avaliação do impacto do ciclo de vida, que classificou 42 termos (23,59%), em terceiro lugar Análise do inventário do ciclo de vida, que classificou 38 termos (21,34%), em quarto lugar Objetivo e escopo, que classificou 25 termos (14,04%), em quinto lugar Interpretação do ciclo de vida, que classificou 13 termos (7,3%) e, por último, em sexto lugar ficou S/C Relatório, análise e revisão crítica, que classificou 07 termos (3,93%).

Acredita-se ser importante frisar que é verdade afirmar que a maioria dos termos categorizados como Avaliação de impacto do ciclo de vida, por exemplo, encontra-se na norma NBR ISO 14042; entretanto não é verdade afirmar que, por exemplo, um termo retirado da norma NBR ISO 14043 seja, necessariamente, categorizado como pertencente a Interpretação do ciclo de vida, apenas pelo fato de que foi encontrado dentro desta norma; é necessário perceber como esse termo ocorre dentro do texto e ao que ele se refere, pois as quatro (ou seis) categorias, apesar de estarem separadas neste estudo, por fins didáticos, são interligadas na prática.

Quanto ao tipo de termo, foram especificados, *a priori*, três:

- 1) Entidade
- 2) Entidade agente,
- 3) Processo.

Segundo Uschold *et al.* (1996, p. 12) “uma ontologia para empresa é composta de um conjunto de ENTIDADES e um conjunto de RELAÇÕES entre ENTIDADES. ENTIDADES podem exercer FUNÇÕES nas RELAÇÕES.” Eles ainda determinam que “Algumas FUNÇÕES nas RELAÇÕES são especiais no sentido de que o papel exercido por elas englobam ações e cognições [...] ENTIDADES que exercem estas FUNÇÕES são chamadas de ATORES. (USCHOLD *et al.*, 1996, p. 14).

A releitura do que é dito por estes autores para os propósitos desta pesquisa nos levou a

determinar a tipologia citada acima, sendo que Entidade Agente, por possuir a característica de agentividade, equivale para este estudo a denominação ATORES estabelecida para Uschold *et al.* (1996).

A definição do tipo de termo Processo se fundamentou na declaração de Uschold *et al.* (1996) para quem as ontologias de atividades, planejamento e processo apenas representam estes últimos, acrescentando que os processos contêm um histórico de atividades que devem ser modeladas na ontologia.

Quanto as FUNÇÕES e RELAÇÕES mencionadas por Uschold *et al.* (1996), estas foram traduzidas nos campos 6.1 ACV propriedade e 6.2 ACV natureza da propriedade respectivamente. Estes campos serão detalhados mais adiante.

Sob o ponto de vista quantitativo, o tipo de termo mais produtivo da amostra foi o denominado Entidade contando com 100 itens (56,17%). Em seguida encontram-se 67 Processos (37,64%) e 11 Entidades Agentes (6,17%).

Referente ao campo definição/contexto, preferiu-se construí-lo de maneira que abarcasse o maior número de informações possíveis, para que assim se pudesse tirar dele todas as propriedades e ainda os outros atributos. Ou seja, optou-se por criar uma definição bastante densa, e dela retirar as propriedades do termo e também seus outros atributos, como notas, exemplos, sinônimos e termos limites. Por isso, a denominação definição/contexto.

Pode-se dizer que definição/contexto é a “chave mestra” dentro do formulário, pois é dela que se retiram todas as informações e, sem ela não há como preencher os outros campos, principalmente os campos relativos a propriedade e natureza da propriedade. Isso pode ser confirmado ao se notar que os termos cujas definições não foram encontradas, não possuem quaisquer outros atributos preenchidos.

A respeito das fontes da definição/contexto, como dito anteriormente, foram usadas Chehebe (1997) e as normas NBR ISO e ISO/TR e ISO/TS. Como explicado nas notas de rodapé 2 e 3, as referências bibliográficas foram abreviadas (ver correspondência com a referência na nota 3) com o intuito de simplificar o registro no formulário do *Access* e estão organizadas do seguinte modo: referência bibliográfica abreviada

(dois pontos): número da página (vírgula), número da página (ponto-e-vírgula); referência bibliográfica abreviada (dois pontos): número da página (vírgula), número da página (ponto-e-vírgula); e assim por diante.

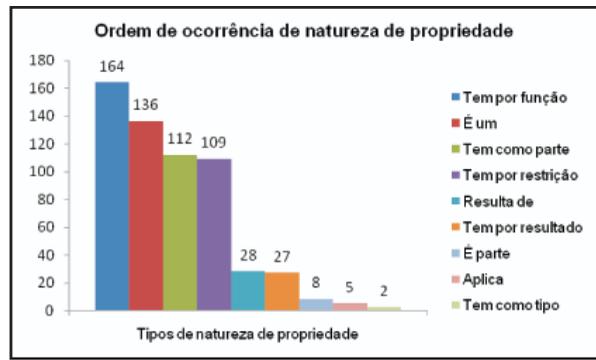
Relativamente às propriedades pode-se dizer que esse espaço foi subdividido em duas seções (6.1 e 6.2 na seção anterior): 1) ACV propriedade e 2) ACV natureza da propriedade. A primeira seção preocupa-se em coletar as propriedades do termo contidas na definição/contexto do mesmo, enquanto a preocupação da segunda é classificar essa propriedade coletada em relação a sua natureza semântica. Dessa forma, nota-se que há uma relação biunívoca entre essas duas seções, isto é, para cada propriedade coletada há apenas uma “etiqueta” semântica; e para cada “etiqueta” semântica há apenas uma propriedade coletada.

Assim, foram encontrados nove tipos de naturezas de propriedades:

- 1) Tem por função, aquele que tem a propriedade de desempenhar alguma função ou atividade;
- 2) É um, aquele que tem a propriedade de ser um tipo de algo mais genérico a ele ou de ter alguma característica ou atributo a mais de seu genérico;
- 3) Tem como parte, aquele que tem a propriedade de ter subdivisões ou subpartes;
- 4) Resulta de, aquele que tem a propriedade de resultar de uma ação ou uma atividade;
- 5) Tem por resultado, aquele que tem a propriedade de possuir um resultado ou de resultar em algo;
- 6) Tem por restrição, aquele que tem a propriedade de não ter algum tipo de propriedade;
- 7) É parte, aquele que tem a propriedade de ser uma subdivisão ou fazer parte de algo;
- 8) Tem como tipo, aquele que possui a sua espécie em sua definição; e
- 9) Aplica, aquele que expressa em que o conceito é aplicado. Elas classificaram um total de 591 propriedades coletadas - ou seja, uma média de 3,32 propriedades coletadas de cada termo - cada natureza semântica classificou uma média de 65,66 propriedades coletadas.

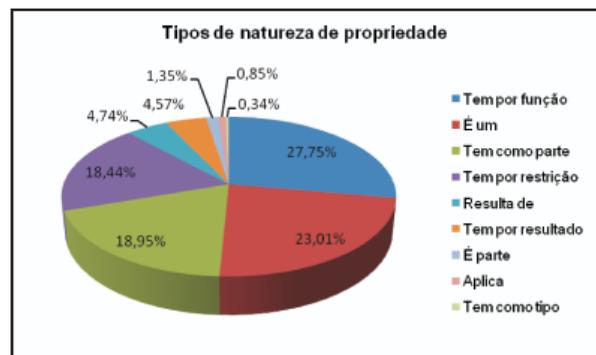
Segue abaixo dois gráficos que ilustram a análise quantitativa dessas naturezas de propriedades. No gráfico 1, é apresentado no eixo y, o número de ocorrências e, no eixo x, os nove tipos de natureza de propriedades, mostrado, em ordem crescente, que natureza de propriedade classificou mais propriedades coletadas; ainda pode-se notar um quadro à direita do gráfico que indica, por meio de cores, qual barra representa cada tipo de natureza de propriedade e a sua respectiva porcentagem.

Gráfico 1 - Ordem dos tipos de natureza de propriedade de ACV



No gráfico 2, ilustram-se os mesmos valores em termos percentuais.

Gráfico 2 - Tipos de natureza de propriedade de ACV



Quanto à ocorrência das naturezas de propriedades em cada tipo de termo, a tabela a seguir mostra que apenas Processo contém todas as nove naturezas. Enquanto que Entidade possui oito naturezas e Entidade Agente apenas 4.

Tabela 1 - Relação entre Tipo de termo e Natureza da Propriedade

ACV_DEF_TIPO_TERMOS	ACV_NATUREZA_PROPRIEDADE
Entidade	Aplica É parte É um Resulta de Tem com parte Tem por função Tem por restrição Tem por resultado
Entidade Agente	É um Resulta de Tem por função Tem por resultado
Processo	Aplica É parte É um Resulta de Tem como parte Tem como tipo Tem por função Tem por restrição Tem por resultado

As tabelas 2, 3 e 4 ilustram exemplos de aplicação de algumas dessas naturezas por tipo de termo.

Tabela 2 - Exemplo de aplicação de natureza no termo *Alocação baseada em propriedade física*

Alocação baseada em propriedade física		
Categoria	Análise do Inventário	
Tipo de termo	Processo	
Propriedade/Natureza	Objetiva manter a alocação estável dada uma determinada tecnologia	Tem por função
	Funciona bem quando existe uma forte correlação positiva entre as propriedades físicas escolhidas e o valor dos co-produtos	Tem por restrição
	A vantagem é sua capacidade de fornecer resultados consistentes	Tem por resultado
	A alocação por massa é um tipo de alocação física	Tem como tipo

Tabela 3 - Exemplo de aplicação de natureza no termo *Auditor ambiental*

Auditor ambiental		
Categoria	S/G Geral	
Tipo de termo	Entidade Agente	
Propriedade/Natureza	Pessoa qualificada	É um
	Executar auditorias ambientais	Tem por restrição

Tabela 4 - Exemplo de aplicação de natureza no termo *Energia de processo*

Energia de processo		
Categoria	S/G Geral	
Tipo de termo	Entidade	
Propriedade/Natureza	Energia requerida por um processo elementar para operar o processo ou equipamento dentro do processo	É um
	Excluindo entradas de energia para a produção e distribuição desta energia	Tem por restrição

Com respeito às notas, algumas foram coletadas diretamente da fonte já especificamente reconhecidas como notas, mais precisamente das normas ISO, já outras foram retiradas da definição/contexto do termo. Quanto ao segundo tipo de coleta, considerou-se que a parte mais “irrelevante”, ou seja, a parte menos central da definição do termo, que estivesse ali mais a título de curiosidade, fosse a nota.

Pode-se dizer que 46,62% dos termos as possuem. Ou seja, dos 178 termos que poderiam ter este atributo, pelo fato de que também têm a definição/contexto preenchida, 83 as têm de fato, enquanto 95 termos não as possuem.

Referente aos exemplos, alguns foram coletados diretamente da fonte já especificamente reconhecidos como exemplos, alguns nas normas ISO e outros em Chehebe (1997); já outros poucos foram criados ao longo da pesquisa.

Pode-se dizer que 21,34% dos termos os possuem. Isto é, dos 178 termos que poderiam ter este atributo, pelo fato de que também têm a definição/contexto preenchida, 38 os têm de fato, enquanto 140 termos não os possuem.

Em relação ao conceito de sinônimo, registra-se no Dicionário Houaiss (HOUAISS, 2001, p.2581) ser “palavra que tem com outra uma semelhança de significação que permite que uma seja escolhida pela outra em alguns contextos, sem alterar a significação literal da sentença”.

Alguns termos sinônimos foram encontrados dentro da mesma fonte, e utilizados nesta como tal, como por exemplo, o caso de *Atribuição de peso* (31) e *Atribuição de valor* (32) coletados de Chehebe (1997, p. 81). Outros foram notados por terem sido usados em fontes diferentes, com o mesmo sentido, por exemplo, o trio *Processo elementar* (181), *Processo unitário* (183) e *Unidade de processo* (225), sendo que o primeiro

foi coletado da norma NBR ISO 14041 (p. 03), o segundo foi enviado por um especialista da área⁵ e o terceiro foi coletado de Chehebe (1997, p. 30) e da norma NBR ISO 14040 (p. 04, 06).

Foram encontrados sete pares e um trio, consistindo em 17 (dezessete) termos (9,55% do total de termos possíveis), como mostram as duas tabelas abaixo.

Tabela 5 - Pares de sinônimos encontrados nos dados coletados.

Nº do registro ⁶	Termo	Nº do registro	Termo
11	Alocação por massa	12	Alocação por peso
31	Atribuição de peso	32	Atribuição de valor
14	Ambiente do sistema	148	Meio ambiente
108	Fator de caracterização	109	Fator de equivalência
120	Folha de coleta	191	Questionário
125	Fronteira do sistema	142	Limite do sistema
162	Parte interessada	217	Terceira parte

Tabela 6 - Trio de sinônimos encontrado nos dados coletados.

Nº do registro ⁷	Termo	Nº do registro	Termo	Nº do registro	Termo
181	Processo elementar	183	Processo unitário	225	Unidade de processo

Um caso interessante dentro da Tabela 1 é o de *peso*, pois trata-se de uma ambigüidade lexical. Ou seja, a palavra *peso*, no registro nº 12, significa o mesmo que *massa* (peso em grammas, quilogramas, etc.); enquanto, no registro nº 31, quer dizer o mesmo que *valor*.

Relativamente aos termos limites, decidiu-se que seriam termos que estivessem relacionados com o “termo-título” do registro da base de dados do *Access*, porém que não entrassem para

5 Professor Dr. Gil Anderi da Silva (USP).

6 Número de registro na base de dados do *Access*.

7 Idem nota de rodapé anterior.

a categoria de sinônimos e fizessem parte tanto dessa área do conhecimento, quanto de outra. No entanto, pelo fato de o estudo – e os termos também – estar profundamente mergulhado dentro do assunto de ACV, não há ocorrência de nenhum termo limite.

4 CONCLUSÃO

Como proposta de desenvolvimento de uma estrutura definitória para a construção de ontologias, aplicou-se o método experimental no total de 231 termos pertencentes ao domínio da ACV, com o intuito de contemplar a complexidade das relações semânticas envolvidas na conceitualização.

Conforme os resultados atingidos pela pesquisa, pode-se afirmar que a estrutura definitória desenvolvida, principal resultado desta pesquisa, é válida para ser aplicada nesta área do conhecimento, pois contemplou todos os termos coletados, independentemente de seu

tipo, categoria, propriedades e outros atributos. Então, conclui-se que apesar de este padrão ser elaborado especificamente no âmbito do Projeto do Inventário do Ciclo de Vida para a Competitividade Ambiental da Indústria Brasileira, é construído para que sua adoção seja possível no desenvolvimento de outras ontologias assim como seja útil na elaboração de glossário e dicionários especializados.

É importante ressaltar que esta pesquisa não está concluída, logo, fica como sugestão para trabalhos futuros dentro deste campo, a interação com especialistas da área com o intuito de investigar alguns (inter-)relacionamentos dentro da estrutura definitória.

Fica aqui como proposta a pesquisa, principalmente, dentro das propriedades, em relação ao que há de comum entre as propriedades de ACV, coletadas da definição do termo e sua “etiqueta” semântica, ou ainda entre os próprios termos e a “etiqueta” semântica que recebe na natureza das suas propriedades.

A STUDY OF A DEFINING STRUCTURE FOR DEVELOPMENT OF ONTOLOGIES

ABSTRACT

This study examines the identification of the typology of semantic relationships that exist in the Product Life Cycle Assessment (LCA) domain of the *Projeto de Inventário do Ciclo de Vida para a Competitividade Ambiental da Indústria Brasileira* [Life Cycle Inventory Project for Environmental Competitiveness in Brazilian Industry] managed by the *Instituto Brasileiro de Ciência e Tecnologia* (IBICT) [Brazilian Institute of Science and Technology]. To this end it creates a defining structure based on the methodology for ontology development described by Bräscher (2007), and in guidelines described by Uschold (1996), Uschold & King (1995) and Uschold et al. (1996). With the intention of contributing to the development of ontologies, as well as glossaries and specialized dictionaries, it tests the validity of this structure by applying the experimental method to a corpus of 231 terms belonging to the domain. By adopting the structure in question it is possible to identify the category, type, definition, source, property, note, synonym, example and limit term for the collected terms. It was noted in the study that the structure follows a pattern of a highly flexible classification, thus enabling it to be applied to any area of knowledge. It was concluded that the defining structure comprehends the complexity of the semantic relations involved in the conceptualization, and is therefore valid for the construction of ontologies.

Keywords: Defining structure, Life cycle Assessment, (LCA), Semantic Web.

Artigo recebido em 02/04/2009 e aceito para publicação em 24/06/2009

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR ISO 14001**: Sistemas da gestão ambiental: requisitos com orientações para uso. 2. ed. Rio de Janeiro, 2004.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR ISO 14011**: Diretrizes para auditoria ambiental: procedimentos de auditoria: auditoria de sistemas de gestão ambiental. Rio de Janeiro, s.d.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR ISO 14012**: Diretrizes para auditoria ambiental: critérios de qualificação para auditores ambientais. Rio de Janeiro, s.d.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR ISO 14040**: Gestão ambiental: avaliação do ciclo de vida: princípios e estrutura. Rio de Janeiro, 2001.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR ISO 14041**: Gestão ambiental: avaliação do ciclo de vida: definição de objetivo e escopo e análise do inventário. Rio de Janeiro, 2004.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR ISO 14042**: Gestão ambiental: avaliação do ciclo de vida: avaliação do impacto do ciclo de vida. Rio de Janeiro, 2004.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR ISO 14043**: Gestão ambiental: avaliação do ciclo de vida: interpretação do ciclo de vida. Rio de Janeiro, 2005.

BERNERS-LEE, Tim; HENDLER, James and LASSILA, Ora. **The semantic web**: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American, 17 de maio 2001. Disponível em: <<http://www.sciam.com>>. Acesso em: 06 de maio de 2007.

BESSÉ, Bruno de. La définition terminologique. In: Colloque la Définition, 1988, Paris. **Anais...** Canadá: Librairie Larousse, 1990, p. 252 - 260.

BRÄSCHER, Marisa. **Metodologia para desenvolvimento de ontologia no âmbito do projeto de Inventário do Ciclo de Vida para**

a Competitividade Ambiental da Indústria Brasileira. Brasília, IBICT, 2007. 7 p.

CHEHEBE, José Ribamar Brasil. **Análise do ciclo de vida de produtos**: ferramenta gerencial da ISO 14000. Rio de Janeiro: Qualitymark: CNI, 1997.

GUARINO, Nicola; GIARETTA, Pierdaniele. Ontologies and knowledge bases: towards a terminological classification. In: MARS, N. (Ed.). **Towards a very large knowledge bases**: knowledge building and knowledge sharing. [S.I]: IOS Press, 1995, p. 25 - 32.

HOUAISS, Antônio. **Dicionário Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva, 2001.

INTERNATIONAL STANDARD ORGANIZATION. **ISO/TR 14025**: Environmental labels and declarations: type III environmental declarations. Genebra, 2000.

INTERNATIONAL STANDARD ORGANIZATION. **ISO/TS 14048**: Environmental management: life cycle assesement: data documentation format. Genebra, 2002.

INTERNATIONAL STANDARD ORGANIZATION. **ISO/TR 14049**: Environmental management: life cycle assessment: examples of application of ISO 14041 to goal and scope definition and inventory analysis. Genebra, 2000.

REY, Alain. Polysémie du terme definition. In: Colloque la Définition, 1988, Paris. **Anais...** Canadá: Librairie Larousse, 1990, p. 13 - 22.

USCHOLD, Mike. Building ontologies: towards a unified methodology. In: Annual Conference of the British Computer Society Specialist Group of Expert Systems, 16., 1996, Cambridge, UK. **Anais...** Cambridge, UK: University of Edinburg, 1996. p.1-17.

USCHOLD, Mike; KING, Martin. Towards a methodology for building ontologies. In: Workshop on Basic Ontological Issues in Knowledge Sharing, 1995, Edinburg, UK. **Anais...** Edinburg, UK: University of Edinburg, 1995, p.1-13.

USCHOLD, Mike et al. **The enterprise ontology**. Edinburg: University of Edinburg, 1996. 55 p.