

Projeto Eletromemória:
História da Energia Elétrica no Estado de São Paulo
(1890-2005)

O vocabulário controlado do projeto
Eletromemória

Vânia Mara Alves Lima
Maria de Fátima Gonçalves Moreira Tálamo
Marilda Lopes Ginez de Lara
Cristina Hilsdorf
Eduardo de Abreu de Jesus

São Paulo
2009

O vocabulário controlado do projeto Eletromemória

Vânia Mara Alves Lima¹
Maria de Fátima Gonçalves Moreira Tálamo²
Marilda Lopes Ginez de Lara³
Cristina Hilsdorf⁴
Eduardo de Abreu de Jesus⁵

Introdução

Identificar, pesquisar, diagnosticar e referenciar o acervo documental do setor energético paulista é o objetivo do Projeto Eletromemória. Este acervo relacionado à implantação e ao desenvolvimento da geração, transmissão e distribuição da energia elétrica no Estado de São Paulo, no período de 1890 a 2005, encontra-se atualmente sob a guarda da Fundação Patrimônio Histórico da Energia e Saneamento e de diferentes concessionárias do setor elétrico, como AES Eletropaulo, AES Tietê, CESP, ISA CTEEP e Duke Energy.

Por esse motivo, o acervo tem recebido tratamento documental diferenciado, não só nas diversas instituições, mas também dentro de uma única instituição, devido aos diferentes tipos documentais, como livros, fotografias, peças de museu etc. conservados, o que dificulta a recuperação dos documentos e a conseqüente produção historiográfica sobre o assunto.

Na Fundação destacam-se atualmente as atividades de estruturação de três bases de dados para acesso, controle e gerenciamento do patrimônio arquivístico, bibliográfico e museológico da Fundação: Enerdoc, Enerbiblio e Enermuseu. Ocorre que cada uma dessas bases de dados possui uma tabela de assuntos diferente para representação e recuperação dos documentos inseridos no sistema.

¹ Prof^a Dr^a do Departamento de Biblioteconomia e Documentação da Escola de Comunicação e Artes da Universidade de São Paulo. Contato: vamaal@usp.br.

² Prof^a Dr^a do Departamento de Biblioteconomia e Documentação da Escola de Comunicação e Artes da Universidade de São Paulo. Contato: mfgmtala@usp.br.

³ Prof^a Dr^a do Departamento de Biblioteconomia e Documentação da Escola de Comunicação e Artes da Universidade de São Paulo. Contato: larama@usp.br.

⁴ Historiadora formada pela Universidade de São Paulo, atualmente é coordenadora de projetos arquivísticos da Fundação Energia e Saneamento. Contato: crishb@usp.br.

⁵ Graduando no curso de Biblioteconomia pela Universidade de São Paulo. Contato: edubibliotec@yahoo.com.br.

Para tornar público seus conteúdos informacionais, o grupo de documentação do Projeto Eletromemória está elaborando um vocabulário controlado para representar, tratar e recuperar a área de energia dos acervos citados. O controle terminológico pretendido tem por objetivo garantir a apresentação dos termos e a condição de sua interpretação sob o ponto de vista do universo focalizado.

A seguir apresentamos os pressupostos teórico-metodológicos, o método e os instrumentos utilizados até o momento para o desenvolvimento do trabalho proposto.

Pressupostos teórico-metodológicos

O conceito de vocabulário controlado associa-se ao de *tesauro* documentário. Esse termo, tomado de empréstimo do *Thesaurus de Roget*, no original, “*Thesaurus of English words and phrases*”, dicionário que apresenta as palavras a partir da associação de idéias, publicado pela primeira vez em Londres em 1852, designa um instrumento em que os descritores (termos preferenciais que representam documentariamente noções ou conceitos) se relacionam a partir de relações hierárquicas, associativas e de equivalência sinonímica ou quase-sinonímica. Esse aporte, no entanto, não se realiza totalmente se não forem bem determinados os domínios para os quais os termos são utilizados. Nesse sentido, a Documentação realiza uma interlocução com a Terminologia, campo de estudos que se preocupa com a organização de sistemas de conceitos em áreas de especialidade.

Para a Terminologia, uma linguagem de especialidade (*Language for Special Purposes – LSP*) pode se referir a um domínio do conhecimento – por exemplo, o domínio da Física, ou uma área de atividade – a Publicidade, a Olaria, o Comércio Exterior. Não importa, portanto, o grau de especialização do domínio ou área, mas seu vocabulário como uma linguagem com certas características identificadoras, dentro da linguagem geral.

Como em qualquer área, também no caso da Energia Elétrica existe um conjunto de termos que têm significados particulares, se analisados comparativamente a outras áreas. No entanto, a simples listagem dos termos não é suficiente para que se possa inferir o seu significado. Recorremos, nesse ponto, à **Lógica**, à **Lingüística**, à **Terminologia** e à **Documentação**, articulando as propostas de cada campo para usar a noção de relação

como meio de, simultaneamente, conferir uma estrutura ao conjunto de termos e dar-lhes um significado razoavelmente preciso e compartilhado dentro de um domínio.

A Lógica fornece o aparato para compreender e proceder às conjunções e disjunções entre os termos, mas não tem vocabulário para tornar possível decidir em qual domínio o termo deve ser interpretado. A Lingüística – e a Semântica, em particular a estrutural – discute abstratamente as questões das relações entre expressões na língua geral. A Terminologia fornece princípios metodológicos para delimitar áreas de atividade e domínios do saber. A Documentação, por fim, operacionaliza os conhecimentos acima para produzir instrumentos que permitam representar e recuperar documentos.

O princípio subjacente à proposta adotada é o de que a organização estrutural dos termos, ancorada na rede de relações lógico-semânticas definidas a partir das terminologias dos domínios que compõem o vocabulário, é uma das garantias de controle da significação do vocabulário, uma vez que incide sobre a forma/conteúdo dos descritores, diferentemente do que ocorre com uma lista exclusivamente alfabética. Ao lado das relações lógico-semânticas entre as unidades preferenciais – os descritores –, outra se impõe, respondendo pelo princípio de economia, próprio de todo vocabulário, exercitando duas funções decorrentes da noção de equivalência. A primeira delas encontra-se associada ao princípio de economia que rege os vocabulários: unidades lingüísticas com significados próximos (sinônimos, quase sinônimos, termos específicos) passam a integrar um único descritor, reduzindo o número de unidades a serem relacionadas lógica e semanticamente. Esse recurso ao mesmo tempo possibilita o acesso ao sistema porque estabelece uma rede, entre os descritores e não descritores, necessária para o arranjo entre a linguagem comum e as terminologias.

Procedimento metodológico

Em um primeiro momento reunimos os arquivos de descritores denominados Enerbiblio e Enerweb e após análise inicial, foram identificados os seguintes grupos de descritores de acordo com o Quadro 1:

Quadro 1: Número de descritores por grupo

Grupos	Número de descritores
Assuntos	7.458
Instituições	942
Nomes próprios	683
Locais Geográficos	995
TOTAL	10.048

O Grupo “Assuntos” representa o acervo das bibliotecas, museus e arquivos da Fundação e das empresas de energia elétrica que foram analisados e reunidos em planilhas do Excel por grandes áreas do conhecimento, totalizando 65 domínios, conforme Quadro 2, apresentado a seguir:

Quadro 2: Número de descritores por área do conhecimento

Áreas do Conhecimento	Número de descritores
ADMINISTRAÇÃO	185
AGRICULTURA	143
ARQUITETURA	22
ARQUITETURA PAISAGÍSTICA	25
ARTES	60
ARTES GRÁFICAS	11
BIOLOGIA	15
BOTÂNICA	69
CIÊNCIA	11
CIÊNCIA DA COMPUTAÇÃO	239
CIÊNCIA DA INFORMAÇÃO	95
CIÊNCIAS AUXILIARES DA HISTÓRIA	6
CIÊNCIAS SOCIAIS	794
COMBUSTÍVEIS	157
COSMOLOGIA	13
ECONOMIA	332
ECONOMIA DOMÉSTICA	9
EDUCAÇÃO FÍSICA E ESPORTES	53
ELETRÔNICA	46
ENERGIA	266
ENERGIA ELÉTRICA	1.100
ENGENHARIA	104
ENGENHARIA CIVIL	149
ENGENHARIA HIDRÁULICA	122
ENGENHARIA NAVAL	8
FILOSOFIA	32
FÍSICA	39
FOTOGRAFIA	13
GEOFÍSICA	1

GEOGRAFIA	19
GEOLOGIA	88
GEOTECNIA	87
GESTÃO AMBIENTAL	245
HISTÓRIA	46
HISTÓRIA DO BRASIL	50
INSTRUMENTOS	42
LAZER	22
LÍNGUA E LITERATURA	106
MÁQUINAS	56
MATEMÁTICA	125
MECANICA DOS FLUIDOS	20
MEDICINA	88
METALURGIA	46
METEOROLOGIA	17
MUSEOLOGIA	7
MÚSICA	13
NORMAS	6
PROFISSÃO	1
PATRIMÔNIO	5
PESCA	42
PESQUISA	9
PLANEJAMENTO URBANO	17
PSICOLOGIA	163
PUBLICIDADE	27
QUÍMICA	50
RELIGIÃO	34
SAÚDE PÚBLICA	9
TECNOLOGIA	15
TELECOMUNICAÇÕES	60
SEGURANÇA DO TRABALHO	44
TRANSPORTES	79
TURISMO	7
VETERINÁRIA	4
ZOOLOGIA	75
ZOOTECNIA	17

Como o trabalho tem por objetivo mapear o acervo histórico do setor elétrico paulista, pela utilização de um Vocabulário Controlado em Energia, selecionamos os descritores das áreas Energia e Energia elétrica, iniciando o trabalho de estruturação do núcleo semântico pelo estabelecimento de relações entre eles, conforme Quadro 3 a seguir:

Quadro 3: Número de descritores das áreas do conhecimento selecionadas: Energia e Energia elétrica

Área do conhecimento	Número de descritores
Energia	266
Energia elétrica	1.110

Como primeiro exercício para o estabelecimento das relações lógico-semânticas entre os descritores selecionados, escolhemos o descritor BARRAGEM e utilizamos como fonte de referência o Vocabulário Controlado da Agência Nacional de Energia Elétrica, disponível em <http://www.aneel.gov.br/biblioteca/vocabulario.cfm>.

No Quadro 4, abaixo, aparecem em azul os descritores identificados na nossa lista de descritores e também no Vocabulário da ANEEL. Em vermelho aparecem termos ligados ao nosso descritor, mas que aparecem apenas no Vocabulário da ANEEL. Em preto o descritor que está na nossa lista, sobre o qual não encontramos referência para ser hierarquizado:

Quadro 4: Exemplo dos descritores e suas relações com o descritor Barragem

TERMO	UP	TG	TE	TR
BARRAGEM		CONSTRUÇÃO HIDRÁULICA	BARRAGEM DE ATERRO; BARRAGEM DE CONCRETO; BARRAGEM MÓVEL	
BARRAGEM DE ATERRO		BARRAGEM	BARRAGEM DE ENROCAMENTO; BARRAGEM DE REJEITO; BARRAGEM DE TERRA; BARRAGEM DE TERRA- ENROCAMENTO	ATERRO HIDRÁULICO
BARRAGEM DE CONCRETO	BARRAGEM DE CIMENTO	BARRAGEM	BARRAGEM DE CONTRAFORTE; BARRAGEM DE GRAVIDADE; BARRAGEM EM ABÓBADA	
BARRAGEM DE CONTRAFORTE		BARRAGEM DE CONCRETO		
BARRAGEM DE ENROCAMENTO		BARRAGEM DE ATERRO		
BARRAGEM DE GRAVIDADE		BARRAGEM DE CONCRETO		

BARRAGEM DE REJEITO		BARRAGEM DE ATERRO		
BARRAGEM DE TERRA		BARRAGEM DE ATERRO		
BARRAGEM DE TERRA-ENROCAMENTO		BARRAGEM DE ATERRO		
BARRAGEM INFLÁVEL				
BARRAGEM MÓVEL		BARRAGEM		

Procuramos visualizar o produto final pretendido na forma hierárquica a partir do termo mais genérico para o mais específico e incluindo-se as remissivas, indicadas pela sigla UP (Usado para):

CONSTRUÇÃO HIDRÁULICA

UP OBRA HIDRÁULICA

BARRAGEM

BARRAGEM DE ATERRO

BARRAGEM DE ENROCAMENTO

BARRAGEM DE REJEITO

BARRAGEM DE TERRA

BARRAGEM DE TERRA-ENROCAMENTO

BARRAGEM DE CONCRETO

UP BARRAGEM DE CIMENTO

BARRAGEM DE CONTRAFORTE

BARRAGEM DE GRAVIDADE

BARRAGEM MÓVEL

A seguir, procuramos também visualizar o produto final em sua forma alfabética, organizando os descritores dessa maneira e indicando as relações estabelecidas entre eles pelos códigos: TG, TE, TR onde:

TG = Termo genérico

TE = Termo específico

TR = Termo relacionado

UP = Usado para

BARRAGEM

TG CONSTRUÇÃO HIDRÁULICA

TE BARRAGEM DE ATERRO

BARRAGEM DE CONCRETO

BARRAGEM MÓVEL

BARRAGEM DE ATERRO

TG BARRAGEM
TE BARRAGEM DE ENROCAMENTO
BARRAGEM DE REJEITO
BARRAGEM DE TERRA
BARRAGEM DE TERRA-ENROCAMENTO

BARRAGEM DE CONCRETO

UP BARRAGEM DE CIMENTO
TG BARRAGEM
TE BARRAGEM DE CONTRAFORTE
BARRAGEM DE GRAVIDADE

BARRAGEM DE ENROCAMENTO

TG BARRAGEM DE ATERRO

BARRAGEM DE REJEITO

TG BARRAGEM DE ATERRO

BARRAGEM DE TERRA

TG BARRAGEM DE ATERRO

BARRAGEM DE TERRA-ENROCAMENTO

TG BARRAGEM DE ATERRO

BARRAGEM INFLÁVEL

BARRAGEM MÓVEL

TG BARRAGEM

CONSTRUÇÃO HIDRÁULICA

UP OBRA HIDRÁULICA
TE BARRAGEM

Devido ao grande número de dados a serem manipulados, e para que não ocorra a omissão de algum descritor, a partir do exercício realizado optou-se por inserir esses descritores e suas relações no *software* TheW32.

O *software* TheW32

O *software* TheW32 foi selecionado para a estruturação e relacionamento dos descritores por se tratar de *software* didático e gratuito. Ele é um programa em 32 bits para Windows desenvolvido por Timothy Craven, professor da “The University of Western

Ontário” para desenvolvimento e manutenção de *tesauros*, disponível para acesso em <http://publish.uwo.ca/~craven/thew32/thew32.htm>. O TheW32 permite a criação de uma base de dados em que os descritores são inseridos e organizados de acordo com as relações hierárquicas, associativas e de sinonímia existentes entre eles, de modo a serem consultados para a indexação, além de gerar o vocabulário controlado impresso em ordem hierárquica e alfabética.

Em uma primeira etapa os 266 descritores do domínio Energia foram inseridos em ordem alfabética no TheW32 e analisado um a um, ou seja, cada descritor tem sido pesquisado em outros instrumentos de controle terminológico, como o Vocabulário Controlado da Aneel, de forma a estabelecer as relações hierárquicas, de equivalência e associativas entre eles.

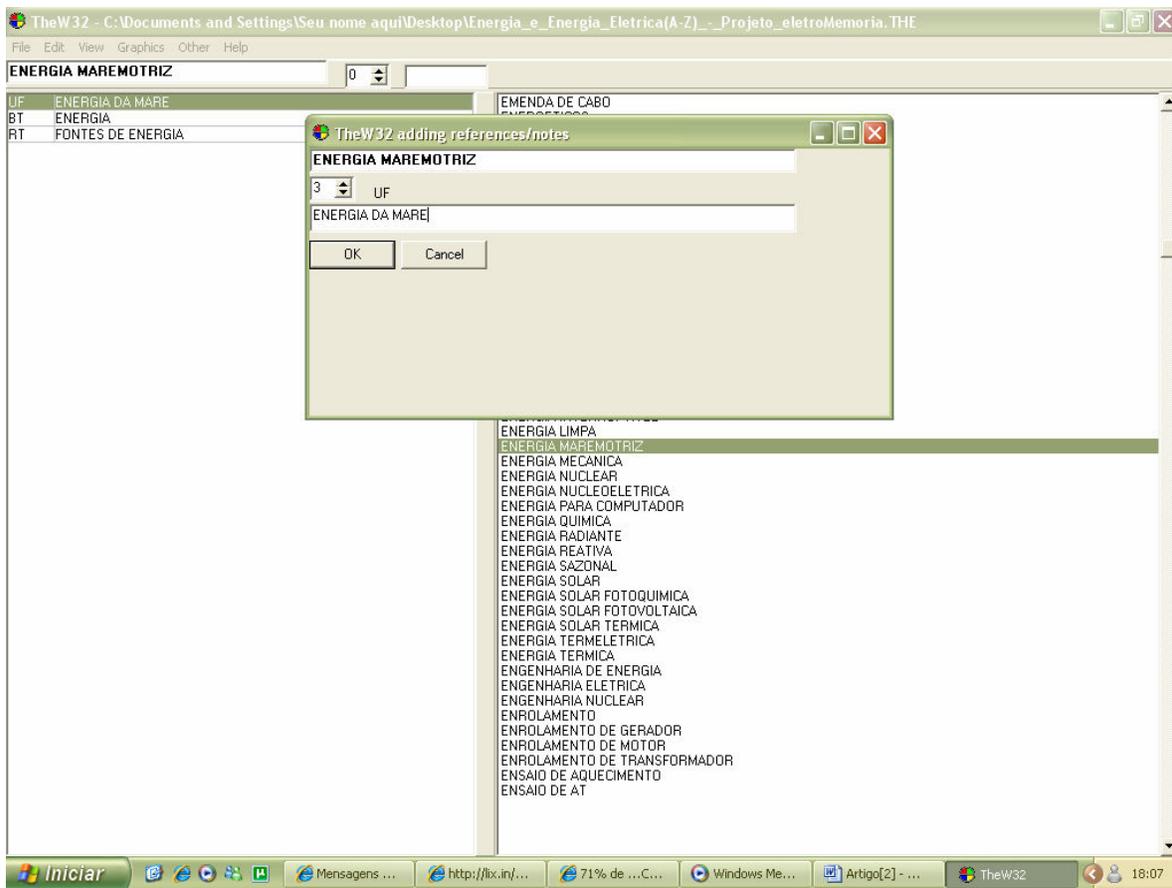
Dessa maneira foram inseridas remissivas para esses descritores, isto é, termos equivalentes, sinônimos, os quais também se constituirão em pontos de acesso de um futuro sistema para recuperação de informação.

Foram identificados outros 40 assuntos em relação hierárquica com os descritores do domínio energia, os quais também foram incluídos na base.

Os descritores que suscitam dúvidas foram separados em uma planilha e devem ser objeto de consulta a especialistas da área e, até o momento, somam 17 descritores.

Até o momento foram inseridos 1.100 descritores do domínio Energia Elétrica, e ao analisarmos os dados obtidos por meio de relatórios gerados pelo TheW32 identificamos que para um total de 1.445 descritores ocorreram mais 496 inserções devido às relações estabelecidas entre eles. Apresentamos a seguir três exemplos de telas do TheW32:

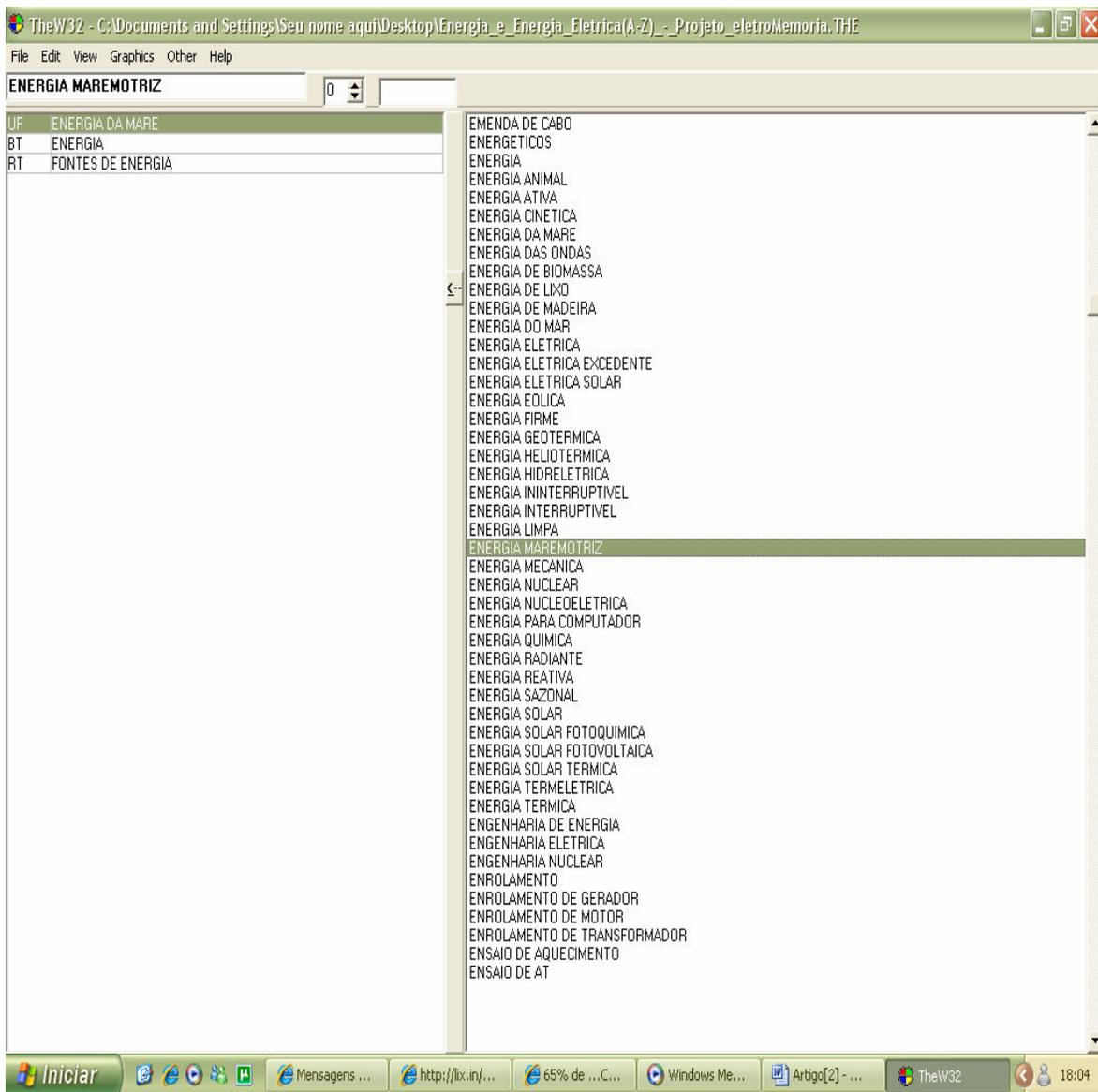
Na primeira tela, vemos um exemplo de inserção, em que o descritor ENERGIA MAREMOTRIZ é USADO PARA a remissiva ENERGIA DA MARÉ.



Tela 1: Inclusão de descritor no TheW32

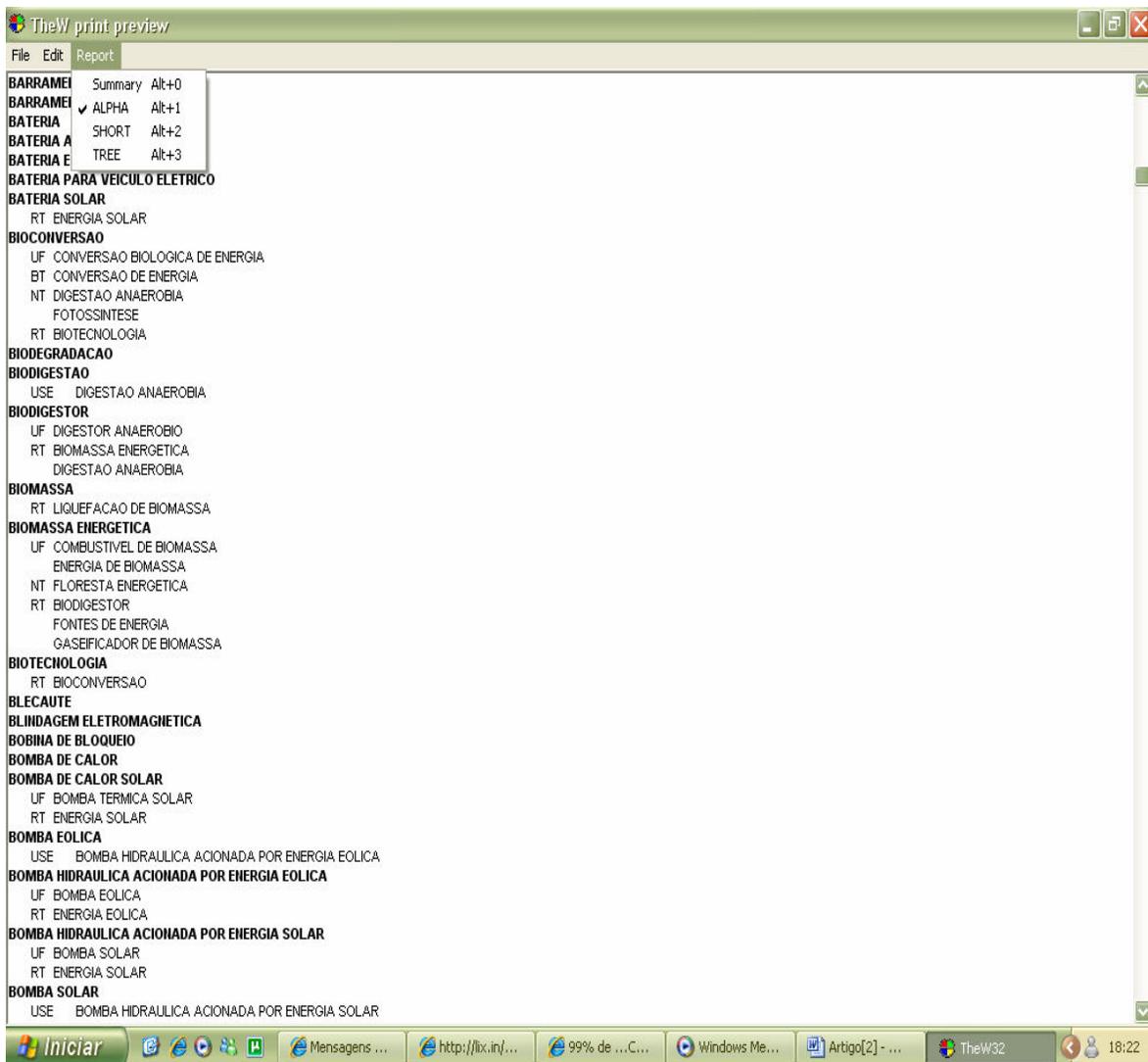
Na segunda tela, vemos como o *layout* do *software* TheW32 apresenta essas relações:

À direita da tela 2 aparece a lista em ordem alfabética dos descritores e remissivas de Energia e Energia Elétrica, e à esquerda o descritor selecionado e suas relações feitas abaixo, TG (BT); UF (Usado Para) e TR (RT):



Tela 2: Lista de descritores em ordem alfabética no TheW32

No último exemplo, a tela 3 mostra o *layout* de apresentação que podemos obter do *software*, como produto final: ALPHA; SHORT e TREE. Optamos por visualizar o produto final em ordem alfabética (ALPHA), em que as relações de todas as maneiras são preservadas:



Tela 3: Lista de descritores e suas relações em ordem alfabética no TheW32

Considerações finais

A partir da elaboração de um Vocabulário Controlado de Energia será possível auxiliar na constituição da terminologia para coleta e interpretação dos dados para o mapeamento da memória e do acervo histórico do setor elétrico paulista, atribuindo consistência conceitual aos questionários e roteiros de observação que subsidiarão o diagnóstico e definir os parâmetros para a adequação da base de dados à função de recuperação da informação.

Sob o ponto de vista da comunicação, o sistema de informação que tem como componente nuclear o vocabulário controlado que vem sendo desenvolvido apresenta-se

como um processo de difusão que integra a ação comunicativa geral. De fato, o processo documentário não cria as mensagens originais, manipula-as estabelecendo mediações para a efetiva transmissão e circulação das mensagens originais. Nesse sentido, o Vocabulário Controlado confere ao sistema de informação que será erigido pela instituição as seguintes características que o identificam como sistema de recuperação da informação: efetividade – desempenho ótimo da tarefa; benefício – o ganho nos processos de circulação da memória da instituição; e eficiência – equilíbrio entre custo e benefício.

REFERÊNCIAS BIBLIOGRÁFICAS

BARROS, L. A. **Curso básico de terminologia**. São Paulo: EDUSP, 2004.

CABRÉ, M.T. **La terminología**: representación y comunicación. Barcelona: IULA-UPF, 1999.

CINTRA, A. M. M. et al. Linguagens documentárias e terminologia. **Cadernos de Terminologia**. São Paulo, nº 1, pp. 17-22, 1996.

CINTRA, A. M. M. et al. **Para entender as linguagens documentárias**. 2ª ed. rev. ampl. São Paulo: Polis, 2002.

DICIONÁRIO Michaelis. Disponível em:

<<http://michaelis.uol.com.br/moderno/portugues/index.php>> – acesso em 01/07/2008.

DICIONÁRIO Prático de Energia Elétrica São Paulo: Duke Energy Brasil. Disponível em: <<http://www.duke-energy.com.br/negocios/dicionario.asp>> – acesso a partir de 01/07/2008.

PROJETO Eletromemória. Disponível em:

<http://www.fphesp.org.br/projeto_fapesp/projeto_fapesp.html>. Acesso em 26/11/2008.

THEW32. Disponível em <<http://publish.uwo.ca/~craven/thew32/thew32.htm>>. Acesso 01/07/2008.

VOCABULÁRIO Controlado da Aneel. Disponível em:

<<http://www.aneel.gov.br/biblioteca/vocabulario.cfm>>. Acesso em 01/07/2008.

VOCABULÁRIO Controlado do SIBI/USP. Disponível em:

<<http://143.107.73.99/Vocab/Sibix652.dll>> – acesso em 01/07/2008.