

Descoberta de Conhecimento em Textos de Sumários de Relatórios de Impacto Ambiental: Análise Exploratória

Cláudia Viviane Viegas¹, Roseli Búrigo¹, José Leomar Todesco²,
Fernando Álvaro Ostruni Gauthier², Paulo Maurício Selig²

Resumo

Os Relatórios de Impacto Ambiental (RIMAs) são documentos destinados a resumir resultados de Estudos de Impacto Ambiental (EIAs), que são documentos complexos exigidos para a identificação e o dimensionamento de efeitos de empreendimentos com significativo potencial de alteração ambiental, e para a apresentação de alternativas – sobretudo tecnológicas e locacionais – e medidas mitigadoras. Este artigo tem por objetivo a análise de textos de quatro sumários de RIMAs do setor hidrelétrico, elaborados entre 1997 e 2005. Para isto, são utilizadas técnicas de KDT – Descoberta de Conhecimento em Textos – seguindo a abordagem proativa, que se inicia sem hipóteses predefinidas. É elaborado um tesauro a partir do Termo de Referência (TR) – documento dado por órgãos ambientais para guiar a construção de EIAs-RIMAs. Os termos do tesauro são comparados aos dos sumários dos RIMAs, após estes passarem por pré-processamento de dados. São recuperadas palavras e expressões-chave nos sumários, as quais são avaliadas quanto à sua representatividade relativa (*weighting*). Os resultados da comparação do tesauro com os termos obtidos dos sumários sugerem que as estruturas dos TRs devem priorizar determinados aspectos de análise nos EIAs-RIMAs, como a questão locacional, os impactos, as alternativas ambientais e as medidas de mitigação e compensação. Alternativas tecnológicas são pouco citadas nos sumários.

Palavras-chave: Descoberta de conhecimento. Textos. Relatórios de Impacto Ambiental.

¹ Doutorandas e ² Professores Doutores do Programa de Pós-graduação em Engenharia e Gestão do Conhecimento (EGC) - Universidade Federal de Santa Catarina (UFSC) - Florianópolis/SC - Brasil.

Abstract

Environmental Impact Reports (EIRs) summarise Environmental Impact Assessment (EIAs), that are complex documents requested for identification and measurement of the effects related to activities that present highlighted potential of environmental changing, and also requested for alternative presentation – mainly technologic and of placement –, and mitigation of these activities. This paper aims to analyse texts of four EIRs' summaries from hydroelectric sector, elaborated between 1997 and 2005. Knowledge Discovery from Texts (KDT) technics are applied – following a proactive approach, without predefined hypothesis. A thesaurus is elaborated from the Reference Term (RT) – document gave by environmental authorities in order to guide EIA-EIRs construction. Thesaurus' terms are compared with that one of EIRs summaries, after these last one have been submitted to pre processing data. Key words and expressions are recovered in summaries, and evaluated according their relative representativeness (weightening). Comparison's results between thesaurus and terms gotten from summaries suggest that EIRs structures must consider priority some aspects of EIAs-EIRs analysis, as placement, impacts, environmental alternatives, and mitigation and compensation issues. Technologic alternatives are less mentioned in summaries.

Keywords: Knowledge discovery. Texts. Environmental impact reports.

1. Introdução: RIMAs e KDT – Uma Análise Oportuna

Relatórios de Impacto Ambiental (RIMAs) são documentos que resumem Estudos de Impacto Ambiental (EIAs), os quais, por sua vez, são processos sistemáticos que envolvem coleta de informações, bem como estruturação e análises dessas informações, de acordo com diretrizes, procedimentos e técnicas consagrados, e análise prévia das conseqüências de ações relativas à instalação, modificação ou ampliação de empreendimentos que apresentem significativos impactos – alterações – ao ambiente natural ou criado pelo homem (ZILBERMAN, 1995).

Embora estejam previstos na Constituição e em outras leis, os EIAs-RIMAs foram, na prática, introduzidos no Brasil por meio da Resolução do Conselho Nacional do Meio Ambiente (Conama) 001/1986 (BRASIL, 1986). Esta resolução serve de base para os órgãos ambientais licenciadores emitirem o chamado Termo de Referência (TR), documento para guiar a elaboração desses estudos. Zilberman (1995) destaca cinco etapas genéricas de um EIA-RIMA, sendo as duas últimas a sumarização das três iniciais. Por tal razão, são consideradas, neste estudo, apenas três:

Etapa I: dimensionamento ou identificação – Inclui informações e análises sobre características locais, recursos tecnológicos e financeiros disponíveis para o controle dos efeitos ambientais do projeto, contexto socioeconômico, objetivos da política de uso e ocupação do solo e legislação em vigor. Inclui ainda a identificação dos possíveis impactos ambientais do projeto, a magnitude desses impactos e suas alternativas.

Etapa II: diagnóstico ambiental – Nesta etapa, é estimada a importância de cada impacto identificado e dimensionado anteriormente. São analisados os meios físico, biológico (ou biótico) e socioeconômico.

Etapa III: prognóstico dos impactos – Nesta etapa, são identificados e analisados os efeitos ambientais possíveis do empreendimento e as possibilidades tecnológicas e econômicas de prevenção

e controle, mitigação e reparo dos efeitos. São identificados e analisados os efeitos ambientais potenciais e é escolhida a alternativa que será a base para o EIA.

Os RIMAs apresentados à avaliação pública são muitas vezes alvo de críticas por sua incompletude ou superficialidade (MOREIRA, 2002). A análise desses documentos, com a utilização de técnicas de mineração de textos, apresenta-se como uma alternativa de investigação de seus pontos fortes e fracos. O processo de mineração de textos, também conhecido como KDT, ou *Knowledge Discovery from Textual Databases*, é uma evolução natural do processo de *Knowledge Discovery Data* (KDD), ou descoberta de conhecimento em bases de dados. Trata-se de “(...) uma evolução natural da recuperação de informações” (WIVES, 2002, p.89). Para Lopes (2004, p.18), KDT ou *text mining* “(...) refere-se ao processo de extrair padrões interessantes e não-triviais ou conhecimento a partir de documentos em textos não-estruturados”. Ou seja, é “(...) um conjunto de técnicas e processos que se prestam a descobrir conhecimento inovador nos textos” (LOPES, 2004, p. 18).

Neste estudo, utilizam-se técnicas de KDT para a análise de sumários de RIMAs do setor hidrelétrico, com vistas à descoberta de padrões que possam elucidar pontos fortes e fracos desses documentos, em sua correlação com palavras e expressões-chave de um glossário construído a partir das principais etapas de elaboração dos EIAs-RIMAs. A investigação é baseada em proposta de Loh et al. (2000), que tratam o processo de descoberta de conhecimento em textos segundo uma abordagem proativa, ou seja, “que se inicia sem hipóteses predefinidas e é baseada numa ação efetiva do pesquisador” (LOH et al., 2000, p.143). Na abordagem proativa, “o usuário não tem uma necessidade ou problema consciente – ele quer encontrar conhecimento novo e útil, mas não sabe o que está procurando ou de que precisa, nem sabe por onde começar” (LOH et al., 2000, p. 143).

No item 2, são apresentadas as bases conceituais, o objetivo, a justificativa e a relevância do estudo.

No item 3, são descritos o método e as restrições da investigação. Os itens 4 e 5 apresentam, respectivamente, resultados e discussão. Em seguida (item 6), delineiam-se algumas conclusões e sugestões.

2. Recuperação de Informações: Técnicas no Contexto dos RIMAs

Para melhor compreensão dos conteúdos de sumários de RIMAs, estudos da área de Recuperação de Informações (RI) podem ser bastante valiosos. RI é “(...) uma atividade que envolve aspectos de descrição de informação (indexação, padronização) e sua especificação para busca, além de qualquer técnica, sistema ou máquina empregada para realizar ou auxiliar essas tarefas” (WIVES, 2002, p. 23). Conforme Saracevic (1995), RI é o processo ou método em que um possível usuário de informação pode converter sua necessidade de informação numa lista real de citações de documentos armazenados que contenham informações úteis a ele. Buckland (1997, apud WIVES, 2002) destaca que “documento” tem vários significados: pode ser qualquer expressão do pensamento humano, ou base material que estenda nosso conhecimento e seja disponível para estudo ou comparação, e ainda qualquer fonte de informação, em formato material, capaz de ser usada para referência como estudo ou como uma autoridade. No caso em questão, documentos são considerados os sumários textuais de RIMAs.

No processo de recuperação de informações, existe o problema do vocabulário, ou seja, muitas vezes o usuário consegue descrever corretamente o que busca, mas o sistema não recupera as mesmas palavras que ele descreveu. Além disto, há palavras com o mesmo significado das que são buscadas, as quais nem sempre são recuperadas. E ainda existem palavras com a mesma forma das que são buscadas, porém com significado diverso. Por isto, duas questões são de máxima relevância: a construção de um glossário comum (*thesaurus*) e a contextualização da busca dentro de um domínio de conhecimento.

Na RI, o primeiro passo é a indexação, ou seja, a seleção de palavras relevantes em um documento, o que pode ser feito por meio da técnica de vocabulário controlado. A indexação tem por objetivo construir pontos de acesso para um documento. Ela é realizada pela identificação de palavras ou expressões-chave (WIVES, 2002).

Um *thesaurus* – ou dicionário de sinônimos, também chamado tesouro – pode ser empregado para auxiliar na normalização do vocabulário (JESUS, 2002). Conforme WIVES (2002, p. 66), um *thesaurus* é “uma ferramenta similar a um dicionário, só que, ao invés de informar o significado das palavras, informa o relacionamento entre elas”. Essas relações podem ser do tipo palavras mais abrangentes *versus* palavras mais específicas, ou então palavras padrão e lista de palavras com significados comuns.

A construção de um tesouro de palavras e expressões chave a partir de etapas de elaboração do EIA-RIMA, seguindo o roteiro de Zilberman (1995), constituiu um dos primeiros passos na relação entre RIMAs e KDT proposta neste trabalho. Esta construção foi necessária para o posterior processo de identificação de informações relevantes, denominado *matching* ou casamento. Tal processo identifica a similaridade entre as informações que são relevantes à consulta do usuário e as informações armazenadas no sistema (WIVES, 2002).

A abordagem contextual complementa a idéia de construção de um glossário. Ela parte do princípio de que todo documento possui um contexto, pois a pessoa que escreve um texto o faz a partir de um assunto específico, utilizando expressões encadeadas, que têm sentido dentro de um contexto ou situação (WIVES, 2002). O modelo contextual “(...) não elimina o problema do vocabulário, mas pode minimizá-lo, se o conjunto de palavras usado na descrição for bem escolhido” (WIVES, 2002, p. 46). No modelo contextual, o ideal é que um especialista selecione os termos mais relevantes do contexto e adicione sinônimos e palavras específicas. Assim, com base na experiência de leitura de RIMAs e TRs, trabalhou-se na contextualização do glossário, ampliando-o a partir das palavras e expressões originais, dadas nas etapas indicadas por Zilberman (1995). Ao tratar contextualmente a criação do tesouro, partiu-se do princípio do conhecimento prático, ou seja, de que “os sistemas de informação devem ser modelados conforme as necessidades de informação dos usuários, levando-se em conta seu padrões de comportamento de busca de informações” (FERNEDA, 2003, p. 18).

Para o tratamento dos sumários dos RIMAs, lançou-se mão de técnicas de análise semântica, como *stopwords* e *stemming* ou lematização. As *stopwords* são as palavras não-relevantes, como preposições, conjunções, pronomes e outras que não têm significado no contexto buscado. Conforme Lopes (2004, p.21), trata-se de “palavras que não têm conteúdo semântico significante no contexto em que ela existe e são palavras consideradas não relevantes na análise de textos”. Já a normalização morfológica, chamada lematização ou *stemming*, considera como relevante o radical das palavras, sem levar em conta as desinências (terminações). “Com essa técnica, o usuário não precisa preocupar-

se com a forma ortográfica com a qual uma palavra foi escrita no texto. Assim, uma idéia, independente de ter sido escrita através de seu substantivo, adjetivo ou verbo, é identificada por um mesmo (e único radical” (WIVES, 2002, p. 53).

Uma vez realizado o tratamento dos textos dos sumários, com apoio de ferramentas como *stopwords* e *stemming*, ponderou-se a freqüência de ocorrência dos termos combinantes (*matching*) com os do tesouro. Isto significa a relevância de cada palavra ou expressão-chave, a qual é dada pela freqüência relativa das palavras indexadas, ou seja, pelo número de vezes em que elas aparecem em relação ao número de palavras do documento. O processo de atribuição de pesos neste contexto, segundo Lopes (2004), chama-se *weighing*.

Para a compreensão do significado dos pesos, utilizou-se então a técnica de *clustering*. O modelo de aglomerados ou *clustering* consiste em identificar a quantidade de palavras similares e freqüentes contidas em documentos. A hipótese de *cluster* ou agrupamento é aquela segundo a qual os objetos semelhantes e relevantes a um mesmo assunto tendem a permanecer em um mesmo grupo e possuem atributos em comum (WIVES, 2002). O modelo de *cluster* aumenta a qualidade dos resultados, pois retorna todo o grupo coeso relevante à consulta e reduz o tempo de processamento, uma vez que grupos semelhantes tendem a ser armazenados em um mesmo bloco do dispositivo de armazenamento.

Ao invés de trabalhar-se com hipóteses de investigação, empregou-se a abordagem proativa de aquisição de informações, delineando-se um estudo exploratório, o qual “(...) serve para detectar problemas potenciais ou oportunidades” (CHOUHRY e SAMPLER, 1997, apud LOH et al., 2000, p. 146). As necessidades da abordagem proativa são classificadas como “dinâmicas e abrangentes” (LOH et al., 2000, p. 145) porque podem mudar durante o processo e porque o usuário não sabe exatamente o que está buscando. “É o caso típico de quando se quer monitorar alguma situação ou se encontrar algo de interessante que possa levar a investigações posteriores” (LOH et al., 2000, p. 147).

A relevância desta abordagem e do estudo está na necessidade de fundamentação de críticas generalizadas, feitas especialmente quanto à estrutura dos EIAs-RIMAs (MOREIRA, 2002; ROHDE, 2002; STRINGUINI, 2002), bem como na ausência de uma análise menos subjetiva dos mesmos. A aplicação de técnicas de KDT a RIMAs é relevante também porque pode contribuir para a maior concisão e adequação desses relatórios às exigências de órgãos ambientais, possibilitando, ainda, a comparação preliminar entre diferentes RIMAs de um mesmo setor, a fim de buscar-se uma estrutura mais uniforme para tais documentos.

3. Método e Restrições

O método do estudo seguiu técnicas de KDT descritas na seção 2. A partir do roteiro genérico de elaboração de EIAs-RIMAs descrito por Zilberman (1995), foram listadas palavras e expressões-chave para as três primeiras etapas da estrutura desses documentos, uma vez que as duas últimas apenas sumarizam as anteriores, não acrescentando novidades. Esta lista foi tratada contextualmente, acrescentando-se sinônimos de palavras e expressões já citadas por Zilberman (1995). Deste modo, obteve-se um tesouro.

Foram selecionados como objetos de análise, aleatoriamente, quatro sumários de RIMAs do setor hidrelétrico, elaborados por empresas especializadas no ramo, entre os anos de 1997 e 2005 –

das usinas Paulistas (MG/GO) (MONTANO e PITHAN, 2005), Ipueiras (SP) (JENSEN, 2005), Jirau (RO) (CAMPOS, 2005) e Barra Grande (SC/RS) (CASTRO, 1997). Cada sumário passou por tratamento semântico, com uso de técnicas de *stopwords* e *stemming*. Assim, após esse tratamento, obteve-se uma lista de palavras e expressões-chave de cada sumário, sendo tais palavras hierarquizadas e ponderadas, com um peso, segundo a frequência de sua ocorrência em cada sumário. Os resultados foram confrontados com a lista de palavras e expressões-chave das etapas do tesouro e comparados entre si. A partir das palavras e expressões-chave de cada sumário, e das palavras e expressões do tesouro, foi produzido um quadro com palavras e expressões comuns aos sumários. Selecionaram-se palavras e expressões com maior peso e com ocorrência no tesouro, as quais foram representadas em modelos espacial e de *clustering*. Foi também construído um *cluster* para a representação genérica do melhor modelo possível de estrutura de sumário de RIMA com base na ponderação de pesos de termos e correspondência deles com o tesouro, para os casos estudados. O estudo foi limitado por sua representatividade de análise, uma vez que baseou-se em apenas quatro casos de estudo, com um número de palavras entre 119 e 354 por sumário. Em razão disto, a aplicação de algoritmos específicos de KDT foi substituída por cálculos de proporcionalidade simples na ponderação dos pesos de palavras e expressões analisadas.

4. Resultados

4.1 Construção do Tesouro

Do roteiro de elaboração de EIAs-RIMAs sugerido por Zilberman (1995) enquanto TR, consideraram-se as seguintes etapas com os respectivos grupos de palavras e expressões-chave agrupados por análise contextual:

Etapa I

A- localização, local(is), alternativa(s) locacional(is), área, área de influência, área(s) influenciada(s), área afetada, região, região de influência, onde;

B- recursos tecnológicos; tecnologia;

C- recursos financeiros; financiamento;

D- contexto socioeconômico, aspecto(s) socioeconômico(s), socioeconomia;

E- política de uso do solo; uso do solo;

F- legislação, lei(s), resolução(ões), aspectos legais.

Etapa II

G- diagnóstico ambiental

H- impacto(s) ambiental(is), alterações ambientais, mudanças ambientais

I- meio físico;

J- meio biológico, meio biótico;

K- meio físico-biótico, meios físico e biótico;

L- meio socioeconômico, aspecto(s) socioeconômico(s);

M- dimensão dos impactos;
N- importância dos impactos.

Etapa III

O- prognóstico dos impactos, prognóstico(s) ambiental(is);
P- efeitos ambientais;
Q- alternativas (ambientais), planos (ambientais), projetos (ambientais), programas (ambientais);
R- alternativas tecnológicas;
S- alternativas econômicas;
T- mitigação, medidas, medidas atenuadoras, medidas compensatórias, medidas compensadoras, medidas corretivas, compensação, compensar, correção, corrigir, reparo;
U- controle, monitoramento;
V- prevenção.

4.2 Tratamento Semântico dos Sumários

Uma vez obtido esse tesauro, foram sistematizados os dados do tratamento semântico dos textos dos quatro sumários, os quais foram submetidos às técnicas *stopword* e *stemming*. A ponderação dos pesos relativos dos termos que combinam (*match*) com os do tesauro resulta no Quadro 1. Já o Quadro 2 apresenta aspectos de combinação (*matching*) e ponderação (*weighting*) da análise realizada.

Grupos de palavras ou Expressões chave do Tesauro	Paulistas		Ipueiras		Jirau		Barra Grande	
	Nº ocor.	Peso	Nº ocor.	Peso	Nº ocor.	Peso	Nº ocor.	Peso
A	5	1,4	10	3,1	4	3,4	6	3,6
B	-	-	-	-	-	-	-	-
C	-	-	-	-	-	-	-	-
D	-	-	-	-	-	-	-	-
E	-	-	-	-	-	-	-	-
F	-	-	2	0,6	-	-	1	0,6
G	-	-	1	0,3	-	-	1	0,6
H	4	1,1	6	1,9	2	1,7	1	0,6
I	-	-	1	0,3	-	-	2	1,2
J	-	-	-	-	-	-	2	1,2
K	1	0,3	1	0,3	-	-	1	0,6
L	-	-	2	0,6	-	-	3	1,8
M	-	-	-	-	-	-	-	-
N	-	-	-	-	-	-	-	-
O	-	-	1	0,3	-	-	-	-
P	-	-	-	-	-	-	-	-
Q	22	6,2	4	1,3	1	0,8	1	0,6
R	-	-	1	0,3	-	-	-	-
S	-	-	-	-	-	-	-	-
T	2	0,6	2	0,6	3	2,5	1	0,6
U	5	1,4	-	-	-	-	-	-
V	-	-	-	-	-	-	-	-

Quadro 1: Resultados da classificação semântica dos termos dos sumários dos RIMAs ponderados e comparados com os termos do tesauro.

Fonte: Viegas et al. (2006)

Sumário	Nº total de palavras	Nº palavras/expressões que combinam com o Tesauro	Percentual
Barra Grande	169	19	11,2
Paulistas	354	39	11,0
Ipueiras	320	31	9,7
Jirau	119	10	8,4

Quadro 2: Aspectos de *matching* e *weighting* da análise.

Fonte: Viegas et al. (2006).

5. Análise dos Resultados e Discussão

5.1 Palavras ou Expressões Mais Comuns e Representação Gráfica

As palavras ou expressões-chave mais comuns encontradas pertencem aos grupos A, H, Q, T. Observa-se que elas representam todas as etapas descritas por Zilberman (1995): A (I), H (II) e Q e T (III). Segundo os resultados obtidos, os itens mais importantes dos sumários são a localização, o impacto, as alternativas, os planos, projetos ou programas ambientais e as medidas de mitigação. Esses itens podem ser representados graficamente, por clusterização, de acordo com a frequência de sua ocorrência nos sumários, conforme a Figura 2.

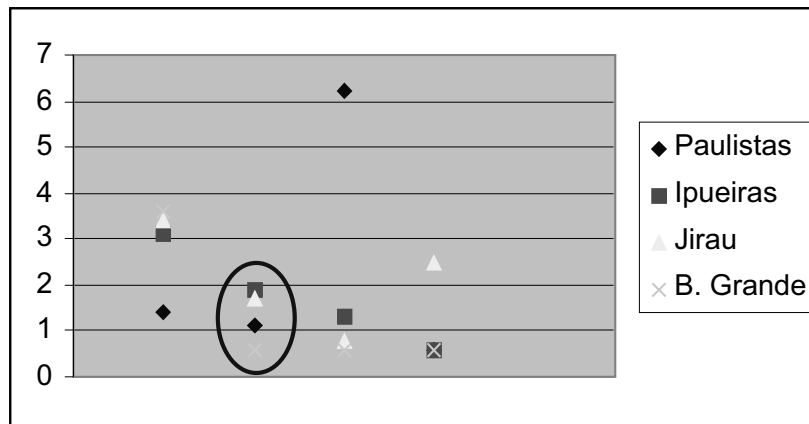


Figura 2: Representação por *clustering* dos pesos dos grupos de palavras e expressões chave mais ranqueadas em cada sumário*.

Fonte: Viegas et al. (2006).

* A elipse mostra o *clustering* das palavras e expressões do grupo H, o que apresenta maior incidência de concentração.

5.2 Relevância: Número de Palavras ou Expressões de cada Sumário em Relação ao Total de Palavras do Tesouro

Considerando-se o número absoluto de palavras ou expressões-chave de cada sumário e o número de palavras ou expressões de cada um deles que combinam com os da lista do tesouro adotado, constatou-se que o sumário do RIMA da usina de Barra Grande é o mais adequado, pois contém a maior proporção relativa de palavras ou expressões-chave (11,2%) que combinam com as do tesouro. Em seguida aparecem, respectivamente, os sumários dos RIMAs das usinas Paulistas (11%), Ipueiras (9,7%) e Jirau (8,4%).

5.3 Número de Grupos de Palavras ou Expressões Seleccionadas em cada Sumário em Relação ao Tesouro

No que diz respeito ao número de grupos de palavras ou expressões-chave comuns entre cada sumário e o tesouro, verificou-se que o sumário de Ipueiras apresenta a maior representatividade: 11 palavras ou 50% do total listado no tesouro. Em seguida, aparecem, por ordem, Barra Grande (45,4%), Paulistas (27,2%) e Jirau (18,1%).

Assim, pôde-se concluir que o sumário menos coincidente com o Termo de Referência, tanto no critério por palavras ou expressões quanto no critério por grupos, é o de Jirau, e os mais adequados são os de Barra Grande, pelo primeiro critério (5.2), e Ipueiras, pelo segundo (5.3).

6. Conclusões e Sugestões para Estudos Futuros

As análises realizadas permitiram concluir que os itens mais importantes dos sumários, comparativamente ao tesouro construído, são a localização, o impacto, as alternativas, os planos,

projetos ou programas ambientais e as medidas de mitigação. Pelos critérios de frequência de palavra ou expressão do tesouro em cada sumário e de grupos de palavras – foram listados 22 –, os RIMAs com sumários mais adequados são os de Barra Grande e Ipueiras, e o menos adequado é o de Jirau. Tais conclusões, embora relativas apenas a sumários com um número relativamente pequeno de palavras – entre 119 e 354 –, evidenciam em quais assuntos/tópicos e em quais tipos de RIMAs devem se fixar estudos futuros de análise semântica de textos de RIMAs, os quais considerem o documento integral.

Embora sem abordagem intencionalmente calcada em hipóteses – por estar baseada na orientação proativa de investigação de KDT (LOH et al., 2000) –, o estudo possibilitou a constatação de que muitos itens elencados em Termos de Referência (TR) – representados, neste caso, por um glossário ou tesouro construído a partir de um TR genérico – não têm eco na elaboração prática de sumários de RIMAs. Ou seja, palavras ou expressões-chave consideradas essenciais em um TR “ideal” não são citadas logo à primeira vista no sumário desses documentos. Tal pode ser constatado em relação aos grupos de termos dos itens B (recursos tecnológicos), C (recursos financeiros), D (contexto socioeconômico), E (uso do solo), M (dimensão dos impactos), N (importância dos impactos), P (efeitos ambientais, que podem ser positivos ou negativos), S (alternativas econômicas) e V (prevenção de danos). Dessa lista, pôde-se concluir que os sumários analisados não se preocupam em transparecer questões tecnológicas e econômicas, por exemplo, ou de dimensionamento e relevância dos impactos ambientais.

Embora quaisquer conclusões do estudo possam ser derrubadas por uma análise mais amíúde, que considere maior número de documentos – e que, portanto, tenha maior representatividade –, os resultados obtidos permitem, no mínimo, configurar uma idéia preliminar da caracterização semântica da estrutura de RIMAs, servindo para dar início a estudos de maior profundidade, aos quais se possam aplicar algoritmos de KDT para se obterem resultados sobre características já “pré-mineradas” num tesouro.

Referências

BRASIL (1986). Resolução CONAMA 001/1986. **Diário Oficial da União**: Brasília, 17 de fevereiro de 1986.

CAMPOS, P.M.P. (org.) (2005) - **Usinas hidrelétricas de Santo Antônio e Jirau – RIMA**. Furnas e Odebrecht, Rio de Janeiro (RJ), 82p.

CASTRO, T.L.C. (org.) (1997) - **UHE Barra Grande – Relatório de Impacto ao Meio Ambiente**. Sumário. Engevix, São Paulo (SP), 59p.

FERNEDA, E. (2003) - **Recuperação de Informação**: Análise sobre a Contribuição da Ciência da Computação para a Ciência da Informação. Escola de Comunicação e Artes da Universidade de São Paulo/ USP (Tese de Doutorado), São Paulo (SP).

JENSEN, P.D. (org.) (2005). **Usina Hidrelétrica Ipueiras – Relatório de Impacto Ambiental – RIMA**. Rede Ipueiras Empresas de Energia Elétrica e Themag Engenharia. São Paulo (SP), 97p.

JESUS, J.B.M.M. (2002) **Tesouro**: Um Instrumento de Representação do Conhecimento em Sistemas de Recuperação da Informação. Anais do XII Seminário Nacional de Bibliotecas Universitárias, Recife (PE).

LOPES, M.C.S. (2004). **Mineração de Dados Textuais Utilizando Técnicas de Clustering para o Idioma Português**. Programa de Pós-graduação em Engenharia Civil da Universidade Federal do Rio de Janeiro/ UFRJ (Tese de Doutorado), 191 p. Rio de Janeiro (RJ).

LOH, S.; WIVES, L.K.; OLIVEIRA, J.P.M. (2000) – Descoberta proativa de conhecimento em coleções textuais: iniciando sem hipóteses. In: **IV Oficina de Inteligência Artificial, 2000**, Pelotas, RS. p. 143-154. Disponível em <<http://www.inf.ufrgs.br/~palazzo/OAI/00%20OIA.pdf>>. Acesso em: 20/04/2006.

MONTANO, C.F.B.; PITHAN, R.O. (org.) (2005) - **Relatório de Impacto Ambiental – RIMA – AHE Paulistas**, Rio São Marcos (GO/MG). Biodinâmica Engenharia. Rio de Janeiro (RJ), 54p.

MOREIRA, RUY (2002). Para que o EIA-RIMA Quase Vinte Anos Depois? In: VERDUM, Roberto e MEDEIROS, Rosa M. (org.). **RIMA – Relatório de Impacto Ambiental**. Ed. UFRGS (4ª edição): Porto Alegre, p.11-21.

ROHDE, Geraldo M. (2002). Estudos de Impacto Ambiental: A Situação Brasileira em 2000. In: VERDUM, Roberto e MEDEIROS, Rosa M. (org.). **RIMA – Relatório de Impacto Ambiental**. Ed. UFRGS (4ª edição): Porto Alegre, p. 41-65.

SARACEVIC, T. Evaluation of Evaluation in Information Retrieval. In: **Conference on Research and Development in Information Retrieval**. 18th Annual International SIGIR, Seattle, USA. Proceedings, USA: ACM Press, 1995, p. 137-146.

STRINGUINI, Millos A. (2002). A Importância da Atuação dos Profissionais, a Metodologia Desenvolvida nas Áreas Específicas e a Articulação Final do RIMA. In: VERDUM, Roberto e MEDEIROS, Rosa M. (org.). **RIMA – Relatório de Impacto Ambiental**. Ed. UFRGS (4ª edição): Porto Alegre, p.66-74.

WIVES, L.K. (2002). **Tecnologias de Descoberta de Conhecimento em Textos Aplicadas à Inteligência Competitiva**. Programa de Pós-graduação em Computação (Exame de Qualificação), 116 p. Porto Alegre (RS), Universidade Federal do Rio Grande do Sul (UFRGS).

ZILBERMAN, Isaac. (1995). **Conceitos e Metodologias para Estudos de Impacto Ambiental**. Ed. Ulbra: Canoas (RS).

