

MARCOS VINÍCIUS MONTANARI

**USO COMBINADO DE EDITOR DE METADADOS E ÁRVORE
HIPERBÓLICA PARA AUXÍLIO NA RECUPERAÇÃO DE DADOS EM
INFRAESTRUTURAS DE DADOS ESPACIAIS: CASO DE ESTUDO DA
IDE-CEMIG**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência da Computação, para obtenção do título de *Magister Scientiae*.

**VIÇOSA
MINAS GERAIS - BRASIL
2016**

**Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa**

T

M764u
2016

Montanari, Marcos Vinícius, 1989-
 Uso combinado de editor de metadados e árvore hiperbólica
para auxílio na recuperação de dados em infraestruturas de dados
espaciais : caso de estudo da IDE-CEMIG / Marcos Vinícius
Montanari. – Viçosa, MG, 2016.
 viii, 36f. : il. (algumas color.) ; 29 cm.

Inclui apêndice.

Orientador: Jugurta Lisboa Filho.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f.30-31.

1. Metadados. 2. Mineração de dados (Computação).
3. Indexação. 4. Sistemas de recuperação de dados. 5. Sistemas
de informação geográfica. 6. Estrutura de dados (Computação).
7. Pesquisa espacial. I. Universidade Federal de Viçosa.
Departamento de Informática. Programa de Pós-graduação em
Ciência da Computação. II. Título.

CDD 22. ed. 005.74

MARCOS VINÍCIUS MONTANARI

**USO COMBINADO DE EDITOR DE METADADOS E ÁRVORE
HIPERBÓLICA PARA AUXÍLIO NA RECUPERAÇÃO DE DADOS EM
INFRAESTRUTURAS DE DADOS ESPACIAIS: CASO DE ESTUDO DA
IDE-CEMIG**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência da Computação, para obtenção do título de *Magister Scientiae*.

APROVADA: .

Mauro Nacif Rocha

Alexandra Moreira

Jugurta Lisboa Filho
(Orientador)

AGRADECIMENTOS

Agradeço primeiramente a Deus, pela força, saúde e por ter permitido chegar até onde cheguei. A meus pais, Paulo e Cornélia, e minha irmã, Shimeny, por todo amor, esforço e apoio durante esta Caminhada.

Agradeço a minha namorada, Giovanna, pelo amor, pela paciência, pela dedicação e pelo incentivo. Ela que também faz parte dessa conquista.

Ao professor e orientador Jugurta Lisboa Filho, pela paciência, elogios e críticas. A todos os professores do Departamento de Informática da UFV.

A todos os colegas, do mestrado, pela ajuda, companheirismo e por todos os momentos que dividimos nesses anos de trabalho. Ao Vitor Dias pela ajuda no desenvolvimento desse projeto.

Agradeço a doutora Alexandra Moreira pela imensa ajuda e colaboração e a todos os componentes do projeto IDE-CEMIG pelo apoio.

Agradeço a Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG, pelo financiamento parcial do projeto.

SUMÁRIO

LISTA DE FIGURAS.....	v
LISTA DE TABELAS	vi
RESUMO.....	vii
ABSTRACT	viii
1 INTRODUÇÃO	1
1.1 O Problema e sua Importância.....	2
1.2 Objetivos.....	3
1.3 Estrutura da Dissertação	3
2 REFERENCIAL TEÓRICO	4
2.1 Infraestrutura de Dados Espaciais	4
2.2 Metadados.....	5
2.3 Mineração de Textos e a Extração dos Termos.....	9
2.3.1 Etapas do Processo de Mineração de Textos.....	9
2.3.1.1 Identificação do Problema.....	10
2.3.1.2 Pré-processamento	10
2.3.1.3 Extração de Padrões	11
2.3.1.4 Pós-Processamento.....	11
2.3.1.5 Utilização do Conhecimento	12
2.4 Recuperação e Visualização de Informação	12
2.4.1 Árvore Hiperbólica (<i>Hyperbolic Tree</i>).....	13
3 RESULTADOS E DISCUSSÕES	15
3.1 edpMGB: UM EDITOR DE METADADOS PARA O PERFIL MGB.....	15
3.2 Desenvolvimento da Árvore Hiperbólica baseada no Sistema Elétrico	21
3.3 Indexando os Metadados na Árvore Hiperbólica	27
4 CONCLUSÕES.....	29
BIBLIOGRAFIA	30

APÊNDICE A32

LISTA DE FIGURAS

Figura 2-1 - Natureza e as relações entre os componentes de uma IDE (CROMPVOETS e STUIVER, 2005).....	5
Figura 2-2 - Perfil de metadados de uma comunidade (CONCAR, 2009).....	7
Figura 2-3 - Etapas do processo de Mineração de Textos (REZENDE et al., 2003).....	10
Figura 2-4 - Representação do processo de recuperação de informação (FERNEDA, 2003)..	13
Figura 2-5 - Exemplo de Árvore Hiperbólica (FREITAS et al. 2001).....	14
Figura 3-1 - Tela Principal do edpMGB.....	17
Figura 3-2 - Caixa de alerta ao usuário sobre elementos obrigatórios	18
Figura 3-3 - Resultado com erros encontrados pelo serviço de validação	19
Figura 3-4 - Tela de geração do arquivo XML.....	20
Figura 3-5 - Árvore Hiperbólica do Sistema Elétrico.....	22
Figura 3-6 - Subárvore Geração	22
Figura 3-7 - Dados Ambientais	23
Figura 3-8 - Distribuição	23
Figura 3-9 - Imagens Georreferenciadas	24
Figura 3-10 - Infraestrutura	24
Figura 3-11 - Instalações CEMIG	24
Figura 3-12 - Organização Político-Administrativo.....	25
Figura 3-13 - Recursos Hídricos.....	25
Figura 3-14 - Referências Cartográficas.....	26
Figura 3-15 - Malha Distribuição	26
Figura 3-16 - Transmissão.....	26
Figura 3-17 - Painel de botões do editor edpMGB.....	27
Figura 3-18 - Quadro de atributos	27
Figura 3-19 - Menu Ajuda	28

LISTA DE TABELAS

Tabela 2-1 - Entidades e elementos do núcleo de metadados do perfil MGB sumarizado 8

RESUMO

MONTANARI, Marcos Vinícius. M. Sc., Universidade Federal de Viçosa, março de 2016. **Uso combinado de editor de metadados e árvore hiperbólica para auxílio na recuperação de dados em infraestruturas de dados espaciais: caso de estudo da IDE-CEMIG.** Orientador: Jugurta Lisboa Filho.

O conjunto de informações utilizado para documentar e organizar dados, com o objetivo de minimizar sua redundância e facilitar sua manutenção e obtenção, é denominado metadado. Um mesmo dado acaba sendo, muitas vezes, produzido por diversos produtores de forma isolada. Para tentar evitar a duplicidade de ações e o desperdício de recursos na obtenção de dados espaciais, o governo brasileiro criou a Infraestrutura Nacional de Dados Espaciais (INDE). A INDE tem como objetivo catalogar, integrar e harmonizar os dados geoespaciais produzidos e mantidos pelas diversas instituições governamentais, visando facilitar sua localização, exploração e acesso por qualquer usuário ligado à Internet. Para definir o conjunto estruturado de elementos básicos que retratam as características dos produtos geoespaciais brasileiros, garantindo sua identificação, avaliação e utilização consistente, a Comissão Nacional de Cartografia (CONCAR) criou o Perfil de Metadados Geoespaciais do Brasil (Perfil MGB). Para pesquisar informações dentro de uma Infraestrutura de Dados Espaciais (IDE) é necessário fazer a busca utilizando uma ou mais das seguintes alternativas: palavras-chave; coordenadas espaciais; classificação temática ou período de tempo. Entretanto, muitos usuários podem apresentar dificuldades na busca de dados geoespaciais através de termos específicos, por não terem conhecimento sobre o assunto ou quais critérios deverão ser utilizados na pesquisa. Este trabalho propõe a utilização de uma árvore hiperbólica de termos para a indexação dos metadados, facilitando sua recuperação. Após a indexação, o usuário pode navegar pelos nós da árvore e realizar buscas pelos metadados relacionados com os termos pesquisados. Para ajudar na elaboração de metadados utilizando o perfil MGB foi desenvolvido o edpMGB, que consiste em um editor de metadados classificado como um software livre e está disponibilizado na Web seguindo o modelo de Software como Serviço (SaaS). Este editor foi desenvolvido no âmbito do projeto de pesquisa e desenvolvimento (P&D) “GeoPortal Cemig – SIG corporativo baseado em IDE”, projeto cujo objetivo é a implantação de uma IDE Corporativa para a Companhia Energética de Minas Gerais. Os metadados criados por esse editor podem ser validados e indexados aos nós de uma árvore hiperbólica criada para o setor elétrico.

ABSTRACT

MONTANARI, Marcos Vinícius. M. Sc., Universidade Federal de Viçosa, March of 2016. **Use of thesaurus for help on the data retrieval in spatial data infrastructure: SDI-Cemig study case.** Adviser: Jugurta Lisboa Filho.

The Information set used to document and organize data, with the objective of minimize its redundancy and obtainment, is called metadata. Different producers can produce a same data many times in an isolated way. To avoid the duplication of efforts and waste of resources, the Brazilian government has created the National Spatial Data Infrastructure (Infraestrutura Nacional de Dados Espaciais - INDE). The INDE aims to catalogue, integrating and harmonizing geospatial data produced and hold by several government institutions, aiming to facilitate its location, exploration and access by any user from Internet. To define the structured set of basic elements that portrays the characteristics of the Brazilian geospatial products, the National Commission of Cartography (Comissão Nacional de Cartografia - CONCAR) has created the Geospatial Metadata Profile of the Brazil (Perfil de Metadados Geoespaciais do Brasil – MGB Profile). To search information in a Spatial Data Infrastructure (SDI), it is necessary to search using one or more the following alternatives: keywords, spatial coordinates, thematic classification or periods of time. However, the untrained users may show difficulties in search geospatial data through specific terms, because the user may not have knowledge about the subjects and which criteria will be used in the search. This work proposes the use of a hyperbolic tree of terms to index metadata, helping its retrieval. After the indexing, the user can navigate through the tree's nodes and search by the metadata related to the terms searched. To help in the metadata creation using the MGB Profile, was developed the edpMGB, which consists a metadata editor, classified as a software open-source, and it is available in the Internet following the model Software as a Service (SaaS). The edpMGB was developed in the research and development project “Geoportal Cemig – SIG corporativo baseado em IDE”, which objective is the implantation of a corporate SDI for the Companhia Energética de Minas Gerais (Cemig). The metadata create by the editor can be validated and indexed to the hyperbolic tree's nodes, created by the electric system.

1 INTRODUÇÃO

Existe o interesse em saber onde certo fato ou fenômeno ocorreu. Isso é próprio do ser humano. Mas esse interesse pela localização geográfica se tornou uma questão estratégica para empresas e governos. Possuir uma informação geográfica precisa significa, para uma empresa vantagem no mercado, e para um governo um meio de planejar suas ações (SOARES et al., 2011).

Para se obter informações geoespaciais é necessário ter acesso e conseguir analisar dados geoespaciais que sejam tão precisos quanto à informação que se deseja. A quantidade de dados geoespaciais disponível na Internet é imensa, esses dados são obtidos utilizando diferentes meios de coleta como receptores GPS e imagens aéreas, com diferentes níveis de precisão. Mas os dados existentes estão espalhados por diversas fontes, em diferentes formatos, o que dificulta a sua busca. Se esses dados foram obtidos por diferentes instituições e estão armazenados em diferentes bancos de dados, com diferentes esquemas, algo precisará ser feito para que a prefeitura consiga obter a informação geoespacial desejada (SOARES et al., 2011).

O compartilhamento de conjuntos de dados geográficos de origens diversas tem enorme importância. Entretanto, segundo Davis e Alves (2005), essa não é a realidade atual, pois os produtores de dados que pretendem compartilhar dados, normalmente se deparam com obstáculos referentes aos formatos dos dados, armazenamento, parâmetros de qualidade, limitações de conteúdo, parâmetros de projeção cartográfica e, até mesmo, estruturas de dados.

Na Tecnologia da Informação e Comunicação a redundância de informações e a falta de padronização dos dados é muito comum. Um mesmo dado é produzido, utilizado e armazenado, em formatos e padrões diferentes, sem levar em consideração a sua necessidade e a sua demanda (DORNELLES e IESCHECK, 2013).

Com o objetivo de evitar a duplicidade de ações e o desperdício de recursos na obtenção de dados espaciais, o Governo Brasileiro iniciou em 2003 estudos visando a integração e a reutilização de dados geoespaciais produzidos pelos diversos órgãos da administração pública federal. Em 2008, através do Decreto 6.666 de 27/11, foi instituída a Infraestrutura Nacional de Dados Espaciais (INDE) (BRASIL, 2008).

O objetivo da INDE é “catalogar, integrar e documentar os dados geoespaciais, produzidos e mantidos pelas diversas instituições governamentais, visando facilitar sua localização, exploração e acesso por qualquer usuário ligado à Internet” (CONCAR, 2009).

Para Leme (2006) a utilização de metadados geográficos é uma das formas mais utilizadas na hora de documentar os dados geográficos, uma vez que são elaborados de acordo com padrões. Tais padrões consistem em um conjunto de elementos que permitem a descrição textual do dado geográfico de forma previamente estabelecida.

Os padrões de metadados geográficos mais conhecidos foram definidos pelo *Federal Geographic Data Committee* (FGDC) e pela *International Organization of Standards* (ISO). Esses padrões internacionais atendem a diversos tipos de usuários, abrangendo a variabilidade da informação geográfica (PRADO et al., 2010).

Para Soares et al. (2011) a criação do *Open Geospatial Consortium* (OGC), e iniciativas de implantação de Infraestruturas de Dados Espaciais (IDE) como, por exemplo, a INDE do Brasil, a NSDI (*National Spatial Data Infrastructure*) dos Estados Unidos e a INSPIRE (*Infrastructure for Spatial Information in the European Community*) da Europa auxiliam na integração e compartilhamento de dados geoespaciais.

A fim de atender à demanda de padronização de metadados iniciada com a INDE, a Comissão Nacional de Cartografia (CONCAR) elaborou o Perfil de Metadados Geoespaciais do Brasil (Perfil MGB), baseado na norma ISO 19115:2003. (ISO, 2003).

Com base na norma ISO 19139:2007 foram definidos esquemas XML para a materialização e codificação da ISO 19115:2003 em arquivo. Sendo o perfil MGB baseado na norma ISO 19115:2003, seus metadados devem obedecer também aos padrões de materialização definidos na norma ISO 19139:2007 aumentando a interoperabilidade entre sistemas que utilizam o perfil como base (PASCOAL et al. 2013).

1.1 O Problema e sua Importância

A Companhia Energética de Minas Gerais (Cemig) é um dos mais sólidos e importantes grupos do segmento de energia elétrica do Brasil, participando em mais de 200 empresas, além de consórcios e fundos de participação. Companhia de capital aberto controlada pelo Governo do Estado de Minas Gerais, possui 114 mil acionistas em 44 países¹. Suas ações são negociadas nas Bolsas de Valores de São Paulo, Nova York e Madri. A Cemig busca desenvolver uma IDE, denominada IDE-Cemig, para padronizar os processos que utilizem dados geoespaciais na empresa e auxiliar o compartilhamento e descoberta desses dados pelos seus funcionários.

De acordo com Hochmair (2005), a pesquisa dentro de uma IDE baseia-se principalmente em buscas que usam palavras-chave, coordenadas espaciais, classificação

¹ http://www.cemig.com.br/pt-br/a_cemig/quem_somos/Paginas/default.aspx

temática ou temporal. Com o resultado da pesquisa, o usuário pode acessar a fonte dos dados para visualizá-los ou obtê-los ou refazer a pesquisa. Essa abordagem apresenta dificuldades relacionadas principalmente aos usuários, que podem não saber quais palavras-chave usar ou definir quantos critérios serão utilizados. Essas limitações no processo de recuperação da informação podem prejudicar buscas mais complexas.

1.2 Objetivos

Este trabalho tem como objetivo propor a utilização combinada de um editor de metadados e árvore hiperbólica de modo a auxiliar na recuperação dos dados geoespaciais dentro de uma infraestrutura de dados espaciais para o setor elétrico. Para isso, pretende-se, especificamente:

- Desenvolver um Editor de Metadados para o Perfil MGB que valide os metadados produzidos;
- Estabelecer quais termos serão utilizados para fazer parte da árvore hiperbólica baseada no setor elétrico, através do processo de Mineração de Textos;
- Criar um método de indexação dos metadados geográficos produzidos no editor com os termos da árvore hiperbólica.

1.3 Estrutura da Dissertação

O restante da dissertação está organizado como segue:

- Capítulo 2. Apresenta o referencial teórico e o resultado da pesquisa sobre Infraestrutura de Dados Espaciais, Metadados, Mineração de Textos e a Extração dos Termos e suas etapas e Árvores hiperbólicas;
- Capítulo 3. Apresenta uma proposta de sistema para recuperação de informação em catálogo de metadados baseados no Perfil MGB. Trata da criação de uma Árvore Hiperbólica baseada no sistema elétrico e mostra a indexação dos metadados com os termos da árvore criada;
- Capítulo 4. Apresenta as conclusões e possíveis desdobramentos desta dissertação;
- Ao final do documento, é apresentado um apêndice contendo os elementos extraídos para a construção da árvore hiperbólica.

2 REFERENCIAL TEÓRICO

Neste capítulo são apresentados os conceitos utilizados nesta dissertação. A seção 2.1 descreve o conceito de Infraestrutura de Dados Espaciais, seus possíveis componentes e quais desses componentes serão considerados neste trabalho. A seção 2.2 apresenta conceitos relacionados a Metadados, descrevendo os seus usos e aplicações. A seção 2.3 descreve o processo de Mineração de Textos e a Extração dos Termos e a seção 2.5 mostra a Recuperação e Visualização de Informação.

2.1 Infraestrutura de Dados Espaciais

De acordo com Nebert (2004) o termo "Infraestrutura de Dados Espaciais" (IDE) refere-se à coleção de base relevante de tecnologias, políticas e arranjos institucionais que facilitam a disponibilidade e o acesso a dados espaciais. Ela fornece uma base para a descoberta de dados geográficos, além da avaliação e aplicação para os usuários e prestadores de serviços em todos os níveis de governo, do setor comercial, de organizações sem fins lucrativos, universidades e dos cidadãos em geral. Além disso ela faz a hospedagem dos dados geográficos, atributos, documentação (metadados), e alguns métodos para fornecer o acesso a esses dados geográficos. Para ser considerada funcional a IDE deve incluir os acordos organizacionais necessárias para coordená-la e administrá-la em uma escala local, regional, nacional e/ou global, proporcionando o ambiente ideal para interligar aplicativos a dados.

Segundo Crompvoets (2004) os principais componentes de uma IDE são:

- Plataforma institucional que está relacionada com as políticas e os acordos administrativos na implementação dos padrões e dos dados;
- Padrões técnicos que definem as características técnicas dos principais dados;
- Rede de acesso que são os meios que tornam os dados acessíveis aos usuários;
- Dados que são produzidos na plataforma institucional e que obedecem aos padrões técnicos;
- Pessoas, envolvendo os usuários, os produtores de dados e qualquer agente que seja uma mais-valia no processo de desenvolvimento das IDE.

A Figura 2.1 mostra a relação entre os elementos de uma IDE considerando as pessoas e os dados como uma categoria, e as redes de acesso, a política e as normas técnicas como sendo outra categoria. A relação entre as categorias é muito dinâmica, devido às mudanças que

ocorrem nas comunidades e suas necessidades, bem como sua exigência para diferentes conjuntos de dados.

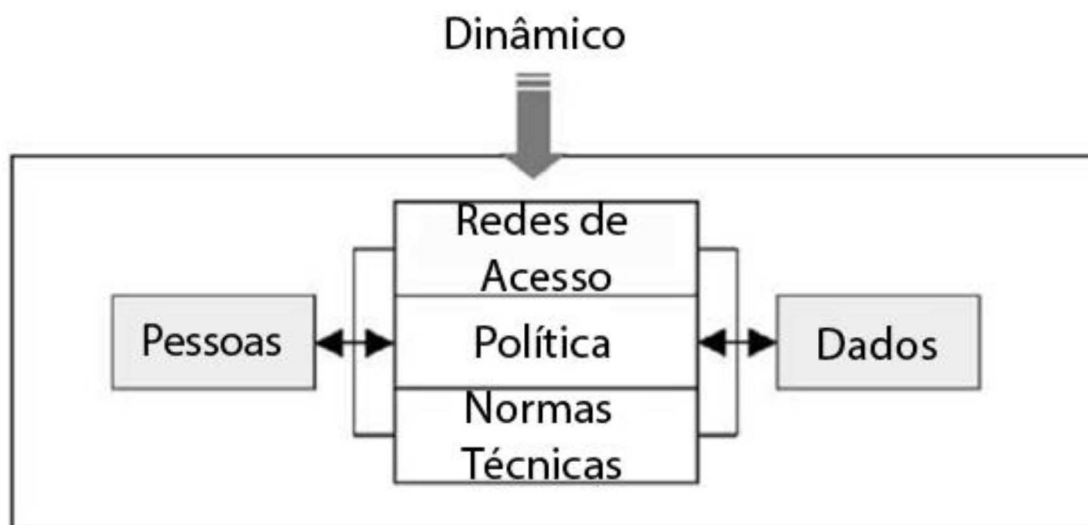


Figura 2-1 - Natureza e as relações entre os componentes de uma IDE (CROMPVOETS e STUIVER, 2005)

A importância da IDE se traduz pelo potencial de compartilhamento das informações geográficas, integrando diversos tipos de usuários. Tem por objetivo diminuir a falta de interoperabilidade entre diferentes sistemas criando uma padronização dos dados. É um conceito simples, mas existe uma dificuldade de implantação devida às diferenças tecnológicas, políticas e culturais dos produtores de dados geográficos (NAKAMURA e FILHO, 2013).

2.2 Metadados

O crescente desenvolvimento na produção de dados espaciais aumenta a necessidade de documentá-los visando seu uso em futuros estudos. Um dado com seu contexto (documentação e metadados) se transforma em informação, todavia, sem estes detalhes praticamente se torna uma informação sem valor (NERBERT, 2004).

Os metadados representam as informações sobre o dado e qualquer descrição sobre quem, como, quando, porque, produziu um dado são exemplos destes. Segundo o Decreto Lei nº 6.666 de 27 de novembro de 2008, instituída no âmbito do Poder Executivo Federal, metadados representa o:

“Conjunto de informações descritivas sobre os dados, incluindo as características de seu levantamento, produção, qualidade e estrutura de armazenamento, essenciais para promover a sua documentação, integração e disponibilização, bem como possibilitar sua busca e exploração”.

A utilização de metadados está relacionada as necessidades que uma organização tem de conhecer melhor os dados armazenados com mais detalhes. A catalogação dos dados facilitará a sua utilização atendendo as necessidades dos usuários. Sem essa documentação eficiente, os usuários têm mais dificuldade para localizar os dados necessários para as suas aplicações. O objetivo dessa forma de representação dos dados é colaborar na orientação, no desenvolvimento e descrição dos documentos eletrônicos, criando padrões, produção e manipulação da descrição por metadados (SOUZA, 1997).

De acordo com Brasil (2008) cada órgão e entidade do poder público federal deve compartilhar os metadados geoespaciais produzidos. Para que isso aconteça é preciso utilizar padrões na hora de criar os metadados, uma vez que estão sendo gerados por órgãos diferentes, porém seguindo a mesma norma, para garantir a interoperabilidade.

Com isso a Comissão Nacional de Cartografia (CONCAR), em 2009, estabeleceu o Perfil de Metadados Geoespaciais do Brasil (Perfil MGB) com a finalidade de estabelecer um padrão nacional utilizado na criação de metadados geoespaciais. Foi desenvolvido por vários órgãos produtores de dados geoespaciais do Brasil. O perfil foi baseado na norma internacional ISO 19115. Porém esse perfil não define como estruturar esses metadados em arquivos eletrônicos, parte que é gerida pela norma ISO 19139:2007, que tem por objetivo materializar os conceitos da ISO 19115:2003 em arquivo, codificando esses metadados em esquemas XML. Antes da criação do Perfil MGB alguns perfis já estabelecidos, também baseados na norma ISO 19115, foram analisados como por exemplo, os perfis: MIG – Metadados de Informação Geográfica (Portugal); NEM – Núcleo Espanhol de Metadados (Espanha); NAP – North American Profile (EUA/Canadá); LAMP – *Latin American Metadata Profile* (proposto para América Latina) (PASCOAL et al. 2013).

Um perfil de metadados é um conjunto de elementos que atendem as necessidades de uma determinada comunidade e garante sua identificação. Possivelmente, o conjunto de metadados definido pelo padrão pode não ser suficiente para atender as necessidades de determinada comunidade, assim, o perfil pode permitir a utilização de elementos adicionais para atender o usuário. A Figura 2.2 mostra a relação existente entre os metadados estabelecidos pelo padrão, os componentes do núcleo, o perfil de uma comunidade e a extensão feita para esse perfil (CONCAR, 2009).

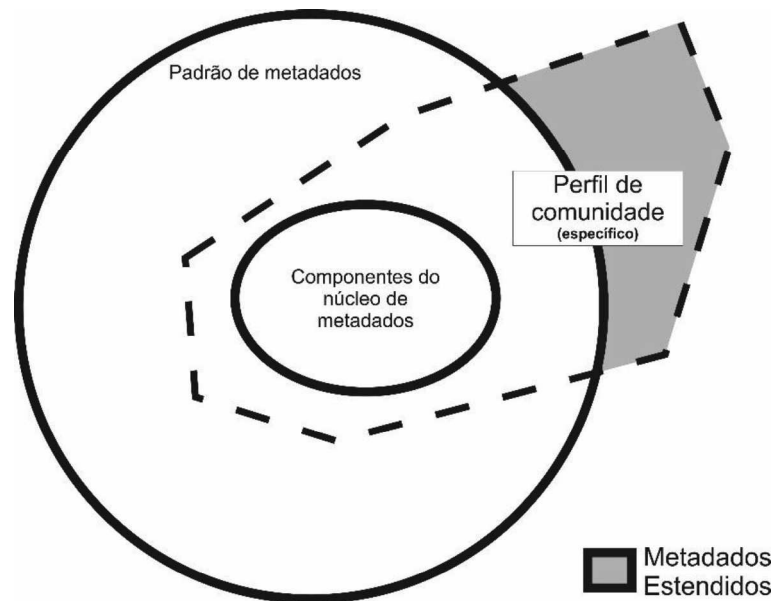


Figura 2-2 - Perfil de metadados de uma comunidade (CONCAR, 2009)

De acordo com a Concar (2009) o perfil MGB possui uma versão denominada perfil sumarizado, que representa o conjunto mínimo de elementos que os metadados geoespaciais produzidos devem ter. O perfil sumarizado possui 23 elementos, ilustrados na Tabela 2.1. O usuário tem a opção de escolher entre a versão completa, com 82 elementos, ou a versão sumarizada. Os elementos do Perfil MGB estão organizados nas seguintes seções:

- Seção 1 (Identificação) - fornece informações básicas sobre o conjunto de dados geoespaciais, tais como título, data, responsável e resumo sobre o dado;
- Seção 2 (Identificação do conjunto de dados geográficos - CDG) - refere-se às informações necessárias à identificação e avaliação de um CDG. Essa seção caracteriza o tipo de representação espacial, escala, idioma, extensão, dentre outros;
- Seção 3 (Informação de restrição) - disponibiliza informações relativas a restrições de acesso e uso, sendo composta por duas entidades, a primeira referente às restrições legais e a segunda às restrições de segurança;
- Seção 4 (Qualidade) - permite uma avaliação sobre a qualidade do conjunto de dados, informando o nível hierárquico, linhagem e o relatório sobre o dado;
- Seção 5 (Informação de manutenção) - informa a frequência de manutenção e atualização;
- Seção 6 (Informação de representação espacial) - descreve os mecanismos usados para representar a informação espacial (matricial ou vetorial);

- Seção 7 (Sistema de referência) - informação acerca do sistema de referência, incluindo o sistema de coordenadas e o referencial geodésico do conjunto de dados espaciais;
- Seção 8 (Informação do conteúdo) - descreve o catálogo de feições e o conteúdo dos dados matriciais;
- Seção 9 (Distribuição) - reporta informações relacionadas ao distribuidor e às alternativas para obtenção dos dados geográficos;
- Seção 10 (Metadados) - seção responsável por informações sobre os próprios metadados. Inclui o responsável pela elaboração do metadado, a data de criação, norma utilizada, dentre outros.

Tabela 2-1 - Entidades e elementos do núcleo de metadados do perfil MGB sumarizado

Entidade/Elemento	Condição	Entidade/Elemento	Condição
1. Título	obrigatório	13. Sistema de referência	obrigatório
2. Data	obrigatório	14. Linhagem	opcional
3. Responsável	obrigatório	15. Acesso online	opcional
4. Extensão Geográfica	condicional	16. Identificador metadados	opcional
5. Idioma	obrigatório	17. Nome padrão de metadados	opcional
6. Código de caracteres	condicional	18. Versão da norma de metadados	opcional
7. Categoria temática	obrigatório	19. Idioma do metadados	condicional
8. Resolução espacial	opcional	20. Código de caracteres do metadado	condicional
9. Resumo	obrigatório	21. Responsável pelos metadados	obrigatório
10. Formato de distribuição	obrigatório	22. Data dos metadados	obrigatório
11. Extensão temporal e altimétrica	opcional	23. Status	obrigatório
12. Tipo de representação espacial	opcional		

Fonte: CONCAR (2009).

A ISO 19115:2003 utiliza a *Unified Modeling Language* (UML) para representar as seções dos metadados. Possui cerca de 400 elementos sendo oito obrigatórios para perfis derivados deste padrão. É um padrão internacional adotado por diversos países, tais como Portugal, Espanha, Estados Unidos, Canadá e Peru. Esta norma permite definir perfis e extensões para campos específicos de aplicações (CONCAR, 2009).

2.3 Mineração de Textos e a Extração dos Termos

A quantidade de informação adicionada anualmente no meio digital é muito grande. Em se tratando de informações textuais, a grande maioria da informação é representada por documentos textuais, considerando que essa é a forma mais natural de armazenar informações. A maioria desses documentos são armazenados em meios digitais e acaba sendo lançada diariamente na Internet, produzindo grandes coleções de dados, como relatórios, especificações de produtos, resumos, notas e publicações eletrônicas (HAN et al., 2011).

Em um documento nem todos os itens são importantes para retratar o conteúdo. Sendo assim, é necessário a utilização de técnicas para lidar com essa grande quantidade de documentos. Essas técnicas devem transformar as informações de forma automática ou semiautomática em conhecimento útil e organizado. Para isso usa-se o conceito de Mineração de Textos. As técnicas que podem ser usadas para a extração de termos são técnicas de simplificação de termos que reduzem a forma como os termos dos documentos são apresentados. Uma dessas técnicas, a radicalização (“stemmização” ou *stemming*), reduz os termos às suas formas inflexionáveis e as vezes reduz às suas derivações. Outra técnica utilizada na mineração de textos é a lematização (ou redução à forma canônica) que tem por objetivo agrupar as variantes de um termo em um único lema. E, por fim, tem-se a técnica de substantivação (ou “nominalização”), na qual são derivados substantivos de palavras de outras categorias morfológicas (CONRADO, 2009).

2.3.1 Etapas do Processo de Mineração de Textos

Para Rezende et al. (2003) o processo de Mineração de Textos é dividido em cinco grandes etapas: Identificação do Problema, Pré-processamento, Extração de Padrões, Pós-Processamento e Utilização do Conhecimento, conforme ilustrado na Figura 2.3.



Figura 2-3 - Etapas do processo de Mineração de Textos (REZENDE et al., 2003)

2.3.1.1 Identificação do Problema

A Identificação do Problema é onde o especialista do domínio identifica o problema, o conjunto de textos que será utilizado no processo, a existência de algum conhecimento prévio que possa ser utilizado e se define quais os resultados esperados. É uma etapa que fornece informações necessárias a todas as outras etapas do processo por isso representa grande importância (CONRADO, 2009).

2.3.1.2 Pré-processamento

A segunda etapa difere o processo de Mineração de Dados e o de Mineração de Textos. Em ambos os casos os dados não estão sempre em um formato adequado para a extração dos padrões. Para a mineração de textos é necessário que os documentos estejam em um formato que possa ser utilizado por algoritmos de extração de conhecimento, além de tratar, limpar e reduzir o volume de textos (BATISTA, 2003).

Os dados de entrada, que correspondem ao conjunto de documentos de interesse, podem estar em diferentes formatos, sendo necessário padronizar as formas em que se encontram. Logo depois, elimina-se as *stopwords*, que são aquelas palavras que nada acrescentam à representatividade dos termos ou que sozinhas não significam nada, como artigos, pronomes e

advérbios. Esse processo ajuda a reduzir a dimensão final do vocabulário eliminando termos não significativos para a análise (MOURA, 2006).

A quantidade de termos presentes nos textos utilizados após a remoção das *stopwords* pode ainda ser muito grande, assim, na extração dos termos, é necessário a escolha da melhor técnica possível para garantir uma maior qualidade nos termos extraídos. Entre as técnicas mais utilizadas para este fim, encontram-se a radicalização, a lematização e a substantivação (BATISTA, 2003).

2.3.1.3 Extração de Padrões

Na etapa de Extração de Padrões são definidas as tarefas a serem realizadas de acordo com o processo. Essas tarefas podem ser preditivas ou descritivas. As tarefas preditivas realizam a generalização de exemplos ou experiências passadas com respostas conhecidas já as tarefas descritivas consistem na identificação de componentes da coleção de textos, que são exemplos não rotulados (CONRADO, 2009).

A qualidade dos resultados depende principalmente do trabalho realizado nesta etapa, mais especificamente nas tarefas de seleção e extração dos termos. O termo, também chamado de característica ou atributo, pode ser uma palavra simples ou composta. Quando o termo é composto por apenas uma palavra, este é denominado de unigrama ou termo simples, e quando é composto por mais de uma palavra, é chamado de n-grama (termo composto ou combinação) (MOURA, 2006).

Nesta etapa os resultados são avaliados e validados. Esses resultados genéricos podem mostrar se há necessidade de repetir algum dos passos anteriores (CONRADO, 2009).

2.3.1.4 Pós-Processamento

Nessa parte do processo de mineração de texto, o especialista do domínio e o minerador de textos devem trabalhar em conjunto para responder a questões como:

- De que maneira o conhecimento do especialista difere do obtido;
- Validação dos resultados obtidos;
- Identificação da adequação de procedimentos nas etapas anteriores para tentar melhorar os resultados;
- De que maneira os resultados obtidos devem ser utilizados.

2.3.1.5 Utilização do Conhecimento

Na última etapa do processo de Mineração os resultados já estão validados e prontos para o uso. Dessa forma, o conhecimento extraído pode ser aplicado para apoiar algum processo de tomada de decisão, conforme objetivo pré-estabelecido na etapa de Identificação do Problema (CONRADO, 2009).

2.4 Recuperação e Visualização de Informação

De acordo com Figuerido (2006) a recuperação da informação é um recurso de comunicação entre um emissor e um receptor onde há troca de informação. O homem, ao fazer uma pergunta ao sistema funciona é considerado um emissor e o sistema um receptor. Quando o sistema devolve uma resposta ele passa a ser o emissor e o homem o receptor. Essa relação só se torna possível através do uso da linguagem. Sendo assim, o estudo do processo de recuperação envolve conhecimentos lógicos, tecnológicos e linguísticos.

Com o aumento da produção de informação no decorrer dos anos foram desenvolvidos meios para facilitar o processo de recuperação da informação auxiliando as necessidades que os usuários de bibliotecas, tradicionais ou digitais, apresentavam. Uma dessas técnicas para auxiliar a recuperação é denominado índice, que é uma coleção de termos que indicam o local onde a informação de interesse do usuário possa ser localizada. Estes termos são organizados para facilitar sua busca (FRAKES e BAEZA-YATES, 1992).

Com o aumento da informação, a complexidade dos objetos armazenados também aumentou e a grande quantidade de dados passaram a exigir técnicas de recuperação cada vez mais aprimorados. Diante desta nova realidade, o processo de recuperação da informação se apresenta como uma área cada vez mais importante (CARDOSO, 2004).

Para Frakes e Baeza-Yates (1992), o processo de indexação relaciona a criação de estruturas de dados associados à parte textual dos documentos como as estruturas de arranjos de sufixos e arquivos invertidos, por exemplo. Estas estruturas podem conter dados sobre as características dos termos na coleção de documentos, tais como a quantidade de vezes que cada termo aparece em determinado documento.

Os sistemas de recuperação de informação apresentam o conteúdo encontrado no *corpus* possibilitando a escolha dos itens que atendam a necessidade do usuário, através da uma expressão de busca. Uma representação simplificada do processo de recuperação de informação é apresentada na Figura 2.4 (FERNEDA, 2003).

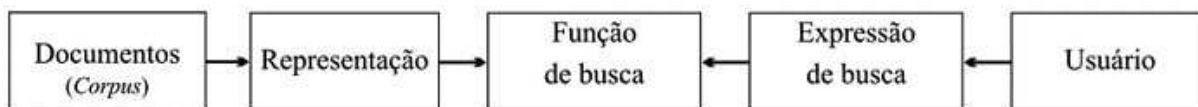


Figura 2-4 - Representação do processo de recuperação de informação (FERNEDA, 2003)

Dentro do processo de recuperação de informação existe a função de busca, que faz comparações entre a expressão de busca fornecida pelo usuário e os documentos que fazem parte do *corpus* e retorna os itens que fornecem a informação que o usuário procura. Porém, mesmo que um termo que faz parte da expressão de busca apareça na representação de um documento não significa que esse documento seja útil para o usuário (FERNEDA, 2003).

Visualizar informação consiste em utilizar recursos computacionais para auxiliar a análise e compreensão de um conjunto de dados. Esse auxílio é fornecido, na sua grande maioria, pela utilização das formas visuais permitindo que o usuário manipule como os dados são visualizados (FREITAS et al. 2001).

Esse tipo de visualização estuda os meios de transformar dados abstratos em imagens reais ou mentalmente visíveis facilitando a sua compreensão e ajudar na obtenção de novas informações contidas nesses dados. Esse processo auxilia o entendimento de um assunto que sem visualização seria difícil de ser entendido (NASCIMENTO e FERREIRA, 2005).

As técnicas de visualização representam graficamente os dados facilitando sua compreensão. Essas representações podem ser distribuídas em três classes: unidimensional, bidimensional ou tridimensional, de acordo com a dimensão do espaço onde os elementos geométricos utilizados estejam situados (FREITAS et al. 2001).

2.4.1 Árvore Hiperbólica (*Hyperbolic Tree*)

Árvore Hiperbólica (Figura 2.5) é um recurso de visualização gráfica, que facilita a navegação em hierarquias de dados muito grandes. Nesse tipo de representação a base é a forma visual de uma estrutura de árvore dentro de um espaço hiperbólico delimitado por uma região circular. O centro do círculo é onde fica o foco da observação e quanto maior a distância do centro maior é o número de informações apresentados e as regiões próximas à borda do círculo apresentam uma grande densidade de informações (FREITAS et al. 2001).

Esta técnica é bastante utilizada em mapas de sites e como ferramenta de navegação em diversas hierarquias. Nela pode ser atribuído aos nós, *links* para outras páginas. Também é possível personalizar as cores dos nós (EVANGELISTA, 2007).

Este tipo de visualização foi escolhido para este trabalho por se tratar de uma visualização de fácil entendimento sobre como está estruturada a hierarquia de conceitos e por ser uma técnica ideal para visualização e navegação de grandes hierarquias.

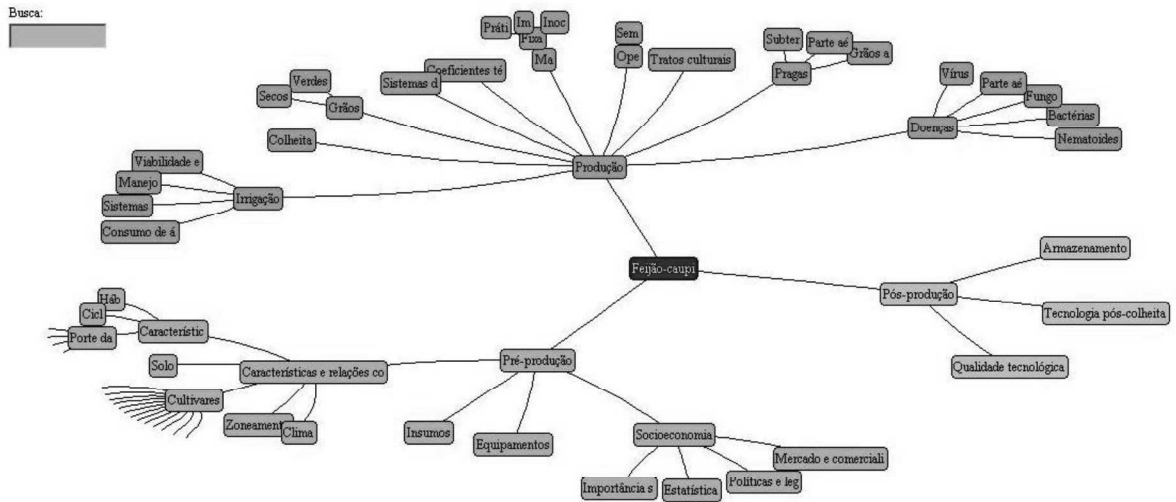


Figura 2-5 - Exemplo de Árvore Hiperbólica (FREITAS et al. 2001)

3 RESULTADOS E DISCUSSÕES

Neste capítulo são apresentados os principais resultados obtidos nesse trabalho. A seção 3.1 descreve o funcionamento do editor edpMGB que foi desenvolvido para a criação e edição de metadados no perfil MGB. A seção 3.2 apresenta a Árvore hiperbólica desenvolvida para o setor elétrico baseada em documentos da ANEEL. A seção 3.3 detalha o processo de indexação dos metadados produzidos pelo editor edpMGB com os nós da Árvore Hiperbólica do setor elétrico.

3.1 edpMGB: UM EDITOR DE METADADOS PARA O PERFIL MGB

A documentação de metadados no Perfil MGB é uma tarefa importante para que estes possam ser facilmente compartilhados e reutilizados. Até o momento não há nenhum editor específico para o Perfil MGB, sendo assim, cada conjunto de elementos que compõe um metadado do Perfil MGB é especificado de acordo com as preferências de seus autores (PEIXOTO et al. 2014).

O processo de preenchimento de metadados é uma tarefa árdua, como qualquer outra catalogação de produto. O uso de softwares para a geração de arquivos de metadados geográficos em diversos formatos e que atendam diferentes normas tem sido proposto para melhorar a documentação do dado geográfico. Dentre estas ferramentas pode-se destacar as de uso gratuito, como o CatMDEdit e o Geonetwork, ambas com formulários para preenchimento dos metadados segundo um perfil definido e com possibilidade de extração automática de algumas características do conjunto de dados.

O CatMDEdit é uma ferramenta de edição e visualização de metadados em vários padrões, que facilita a documentação de recursos, principalmente de informações geográficas. Foi desenvolvido pelo Instituto Geográfico Nacional da Espanha (IGN), juntamente com o *Advanced Information Systems Group* (IAAA) da Universidade de Zaragoza, com o apoio técnico do GeoSpatiumLab (GSL). A ferramenta foi implementada em JAVA e possui características importantes para a documentação de metadados. Entre elas pode-se citar: é um sistema multiplataforma (operando em Linux e Windows); multilíngue (espanhol, inglês, francês, alemão, polonês, português e tcheco); de código aberto, com suporte para extração e geração automática de metadados de arquivos (Shapefile, DGN, ECW, FICC, GeoTIFF, GIF/GFW, JPG/JGW, PNG/PGW); conversão de padrões de metadados e personalização para

gerar novos padrões e perfis de metadados de modo a atender todos os tipos de dados geográficos (CATMDEDIT, 2015).

O Geonetwork é um ambiente padronizado e descentralizado com base em sistema de catálogo para facilitar o acesso, recuperação, atualização e gerenciamento de dados geoespaciais. Fornece um ambiente completo com editor e catálogo de metadados com funções de pesquisa. Possui também um visualizador interativo de mapas na Web utilizando Web Map Service. Atualmente é usado em inúmeras iniciativas de IDE em diversos países. Dentre seus principais recursos, destacam-se: (1) suporte nativo aos padrões de metadados ISO19115, ISO 19139, FGDC e Dublin Core, além de poder configurar um novo perfil de um padrão de metadados; (2) sincronização de metadados entre catálogos distribuídos; (3) gerenciamento de usuários e controle de acesso personalizado; (4) catalogação e acesso a diversos tipos de dados e documentos (upload/download); (5) interface com suporte multilíngue; e (6) importação de metadados no padrão ISO19115 para um perfil de metadados configurado no Geonetwork. Este sistema é livre e de código aberto, facilitando sua evolução e customização por desenvolvedores de IDE (GEONETWORK, 2015).

O editor edpMGB foi construído para suprir a falta de uma ferramenta específica para o padrão MGB, visando os usuários brasileiros da área de informação espacial. Foi desenvolvido baseado no trabalho de Peixoto et al. (2014) cuja a ferramenta desenvolvida possibilita a criação e edição de metadados para documentar padrões de análise.

O edpMGB é um software livre desenvolvido para ser disponibilizado na nuvem via Web seguindo o modelo de Software como Serviço (SaaS). Sendo assim, ele pode ser acessado de qualquer local, necessitando apenas de um navegador Web. Algumas das funcionalidades deste editor de metadados são também disponibilizadas através de Serviços Web, permitindo que outros sistemas possam fazer uso de algumas das funcionalidades desta ferramenta.

Embora os editores CatMDEdit e Geonetwork possam ser configurados para editar metadados conforme o Perfil MGB, são sistemas que requerem instalação em um servidor ou estação de trabalho, que nem sempre estão disponíveis para uso imediato. O fato do edpMGB ser desenvolvido para Web contorna esse problema.

Por meio do edpMGB o usuário pode criar, editar e salvar um metadado na forma de um arquivo .xml seguindo padrões internacionais podendo ser utilizado em várias ferramentas da área geoespacial. O arquivo XML, antes de ser gerado, tem de ser validado em conformidade com o XML *shema* do Perfil MGB.

Este editor foi desenvolvido no âmbito do projeto de pesquisa e desenvolvimento (P&D) “GeoPortal Cemig – SIG corporativo baseado em IDE”, projeto sendo desenvolvido para auxiliar na implantação de uma IDE corporativa para a Cemig. Um dos componentes do GeoPortal Cemig é o catálogo de metadados geoespaciais. Portanto, os metadados são documentados por meio do edpMGB, integrando a IDE-Cemig à Infraestrutura Nacional de Dados Espaciais (INDE).

A Figura 3.1 ilustra a tela principal do edpMGB. Ao lado esquerdo tem a árvore de navegação pelas seções e elementos do perfil. Nesta parte do sistema o usuário tem a opção de trabalhar com as versões sumarizada e completa apresentadas pelo perfil MGB. Na área direita-central está disposta a tela com os campos do perfil MGB divididos em dez painéis representando as seções, que podem ser acessadas utilizando as setas verdes para a direita e para a esquerda.



Figura 3-1 - Tela Principal do edpMGB

A Figura 3.2 (a) mostra as regras para obrigatoriedade, ocorrência e tipo de valor de um elemento. Já a Figura 3.2 (b) apresenta um botão de ajuda referente a cada um dos elementos.



Figura 3-2 - Caixa de alerta ao usuário sobre elementos obrigatórios

Na parte inferior da tela principal existe o painel de botões do editor. O botão “Abrir”, dá acesso a caixa de diálogo onde o usuário poderá carregar um arquivo XML para ser editado. Cada elemento do perfil será carregado em sua respectiva caixa de texto.

O botão “Limpar” tem a função de limpar a caixa de texto do elemento que está aberto e o “Limpar Tudo”, limpa todas as caixas de texto do metadado que está sendo editado. O botão “Validar” permite realizar uma das principais funcionalidades do sistema, que é a validar se o metadado está em conformidade com o esquema XML do Perfil MGB. Ao clicar no botão o sistema verifica os dados inseridos, podendo exibir, por exemplo, uma caixa de diálogo como ilustrado na Figura 3.3, informando que o metadado preenchido não está respeitando o Perfil MGB. Na caixa de diálogo é exibido um log de erros alertando o usuário e mostrando quais campos não estão de acordo com as regras do perfil. O usuário tem a opção de continuar editando o metadado ou armazená-lo em sua máquina mesmo não estando validado pelo Perfil MGB. Ao gerar um metadado não validado, esse recebe uma marcação informando sua não conformidade com o Perfil MGB.

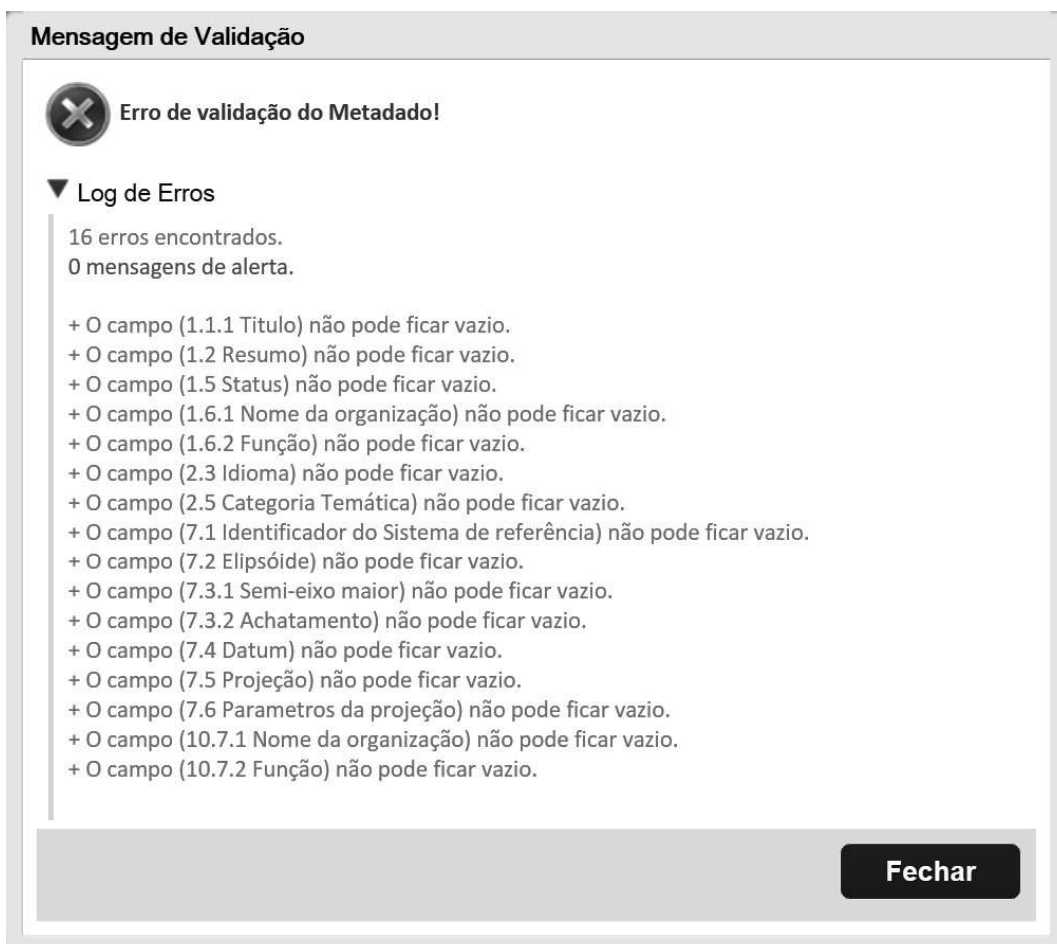


Figura 3-3 - Resultado com erros encontrados pelo serviço de validação

Um método `gerarXMLScript()` recebe como parâmetro os elementos do perfil MGB, que são validados pelo método `validarMGB()`, e em seguida o primeiro retorna o script XML na forma de uma *string*. Já o método `validarMGB()` valida o metadado de acordo com o perfil MGB, recebendo como parâmetro seus elementos e retornando uma lista com as mensagens de erros encontrados no metadado ou uma lista vazia caso esteja respeitando o perfil.

Para Pascoal et al. (2013), a maioria dos metadados fornecidos por produtores de dados nacionais não respeitam totalmente as regras do perfil e isso é um grande problema pois compromete a interoperabilidade entre sistemas que utilizam esse mesmo perfil. Mesmo assim, a impossibilidade de salvar um metadado não validado pode gerar um problema para os usuários, eventualmente pela ausência da informação sobre os elementos do metadado. Sendo assim o usuário tem a opção de salvar os metadados mesmo sem estar em conformidade com o Perfil MGB, e se preciso, voltar a carregá-lo na ferramenta e editá-lo novamente. O botão “XML” salva o metadado em um arquivo `.xml` (Figura 3.4), contendo todas as informações inseridas nos campos. O arquivo `.xml` é salvo no computador do usuário. Portanto, este editor

de metadados com seu serviço de validação de conformidade com um XML *schema* é uma importante contribuição para a evolução da INDE.

```
<?xml version="1.0" encoding="UTF-8"?>
<gmd:MD_Metadata xmlns:gmd="http://www.isotc211.org/2005/gmd"
xmlns:gco="http://www.isotc211.org/2005/gco"
xmlns:geonet="http://www.fao.org/geonetwork"
xmlns:gml="http://www.opengis.net/gml"
xmlns:gts="http://www.isotc211.org/2005/gts"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
>
  <gmd:identificationInfo>
    <gmd:MD_DataIdentification>
      <gmd:citation>
        <gmd:CI_Citation>
          <gmd:title>
            <gco:CharacterString>edpMGB</gco:CharacterString>
          </gmd:title>
          <gmd:date>
            <gmd:CI_Date>
              <gmd:date/>
              <gmd:dateType/>
            </gmd:CI_Date>
          </gmd:date>
        </gmd:CI_Citation>
      </gmd:citation>
      <gmd:abstract/>
      <gmd:status/>
      <gmd:pointOfContact>
        <gmd:CI_IndividualName/>
        <gmd:CI_OrganisationName/>
        <gmd:CI_Role/>
        <gmd:CI_Contact>
          <gmd:address>
            <gmd:CI_Address>
              <gmd:city/>
              <gmd:administrativeArea/>
              <gmd:postalCode/>
              <gmd:country/>
              <gmd:electronicMailAddress/>
            </gmd:CI_Address>
          </gmd:address>
        </gmd:CI_Contact>
      </gmd:pointOfContact>
    </gmd:MD_DataIdentification>
  </gmd:identificationInfo>
</gmd:MD_Metadata>
```

Figura 3-4 - Tela de geração do arquivo XML

Como as funções de validar e gerar o XML foram desenvolvidas também como serviços Web, outros desenvolvedores poderão criar aplicações (ex.: gerenciador de catálogo de metadados) que utilizem os serviços implementados na ferramenta remotamente, por meio do protocolo SOAP, o que aumenta o alcance de uso dessa ferramenta.

Ao clicar no botão “Indexar” o editor realiza uma nova validação do metadado, após realizar essa validação, um arquivo XML é gerado contendo a estrutura de uma árvore hiperbólica criada para o setor elétrico. Essa árvore possui nós que são representados por termos referentes ao setor elétrico. Logo em seguida, o editor varre essa árvore buscando os nós que estão presentes na descrição desse metadado. Quando a busca retorna um resultado verdadeiro, um *link* para esse metadado é armazenado dentro desse nó, dessa forma o usuário pode acessar os metadados que tem relação com os termos da árvore.

3.2 Desenvolvimento da Árvore Hiperbólica baseada no Sistema Elétrico

Para o projeto IDE-Cemig, os termos foram extraídos com o auxílio da ferramenta $E\chi ATO_{LP}$ – Extrator Automático de Termos para Ontologias em Língua Portuguesa (Lopes et al. 2009), desenvolvida para extrair termos de *corpus* anotados linguisticamente. O $E\chi ATO_{LP}$ é uma ferramenta que recebe um *corpus* e extrai automaticamente todos os sintagmas nominais (SN) deste texto classificando-os segundo o número de palavras. Os sintagmas extraídos são salvos em listas que podem conter tanto os SN na sua forma original no texto, como em sua forma canônica. A ferramenta ainda oferece algumas opções de manipulação usuais para listas de termos como a aplicação de pontos de corte, comparação de listas e cálculo de medidas usuais de precisão e abrangência (Lopes et al. 2009).

Um *corpus* de domínio é um conjunto de textos sobre um campo de conhecimento específico que pode ser utilizado para caracterizar esse domínio. Nesse caso, foi utilizada como base de teste o glossário de termos da Agência Nacional de Energia Elétrica (ANEEL)².

A ferramenta $E\chi ATO_{LP}$ extraiu, do *corpus* selecionado, 2688 termos. Foi feita uma análise dos termos extraídos e logo em seguida foi realizado um processo de limpeza desse conjunto de termos. Após o processo de limpeza dos termos foram acrescentados novos termos, referentes ao setor elétrico, que estão presentes na modelagem do banco de dados utilizado pela CEMIG. O Apêndice A apresenta uma amostra dos termos identificados. Para este trabalho apenas os termos com maior frequência foram utilizados, totalizando 284 termos. Os termos que apresentam apenas uma ocorrência foram omitidos por questão de espaço

Logo após a definição da lista de termos foi criada a árvore hiperbólica utilizando o *Hyper Tree Studio*³, ferramenta de código aberto baseada na biblioteca *Hyperbolic Tree*. A Figura 3.5 mostra árvore hiperbólica, tendo como elemento central o termo “Setor Elétrico”, em verde, a partir do qual o usuário pode navegar para as subárvores, em amarelo, como “Distribuição”, “Transmissão” e “Geração”.

² <http://www.aneel.gov.br/biblioteca/glossario.cfm>

³ <http://sourceforge.net/projects/hypertreestudio/>

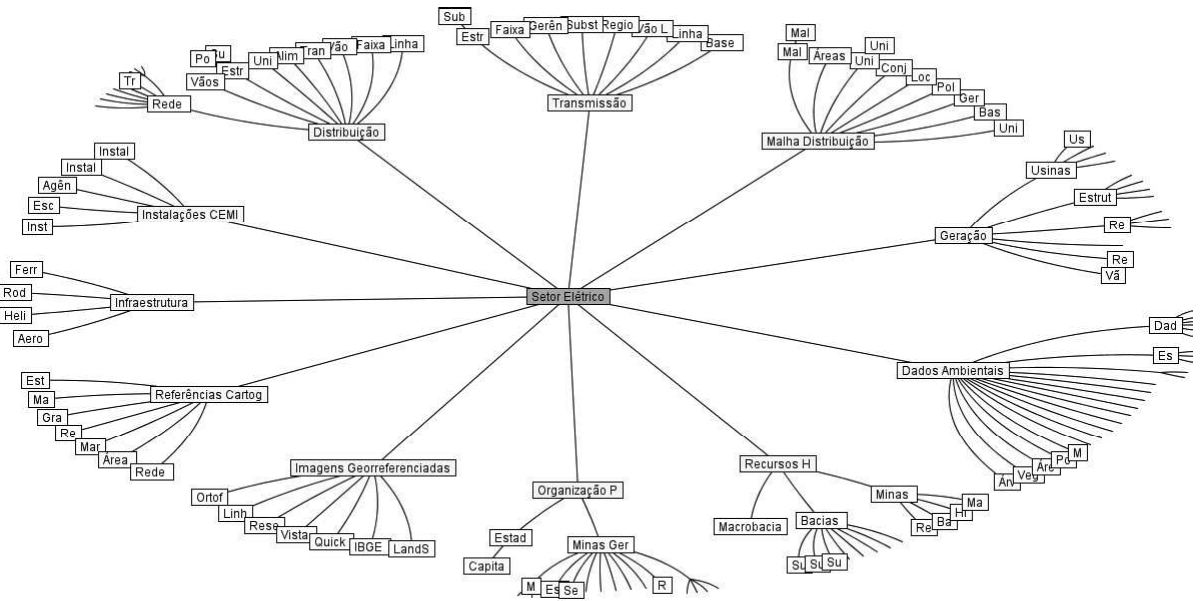


Figura 3-5 - Árvore Hiperbólica do Sistema Elétrico

As subárvores estão organizadas de acordo com a relação que os termos possuem. A Figura 3.6 apresenta a subárvore *Geração* com todos os itens relacionados a área de geração de energia elétrica, como as Usinas e seus tipos e os Reservatórios.

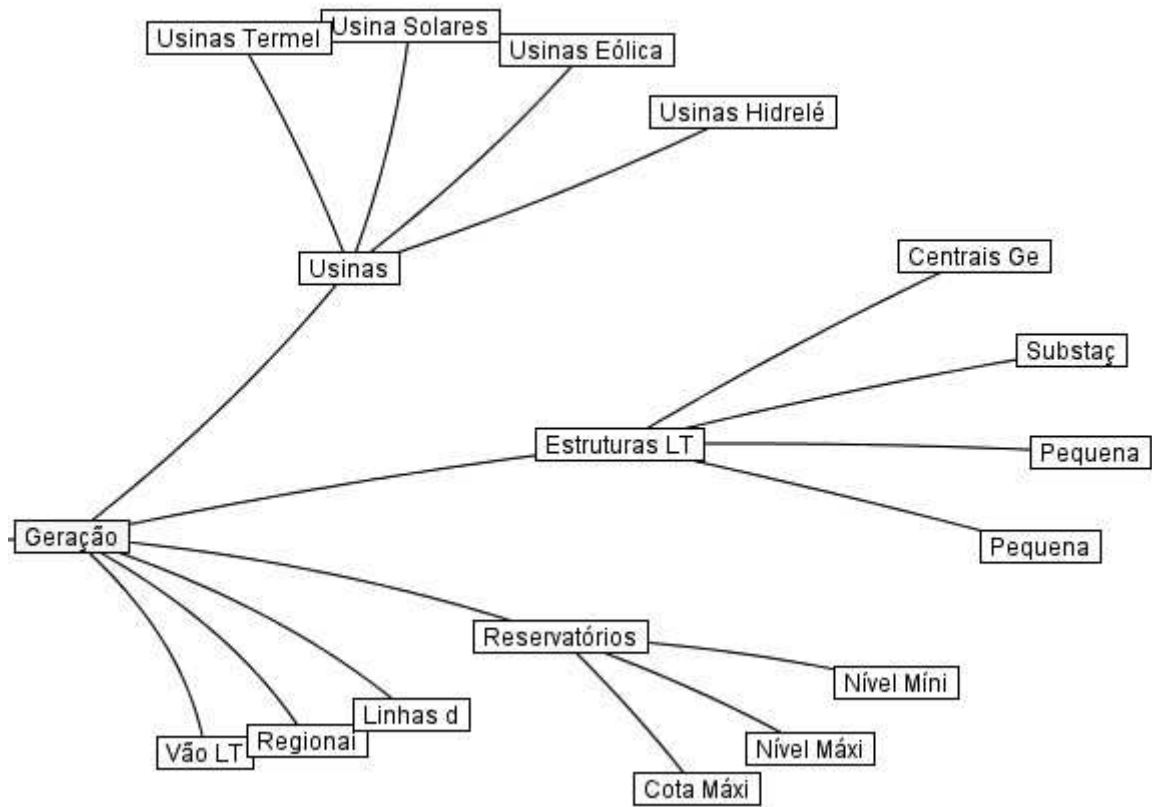


Figura 3-6 - Subárvore Geração

A Figura 3.7 mostra os Dados Ambientais apresentando a relação entre os termos que fazem parte dessa subárvore como Dados Ambientais Históricos e Relevos. Dados Ambientais Históricos engloba termos que representam dados ambientais que são armazenados ao longo do tempo. Estes dados são utilizados para auxiliar no processo de tomada de decisão.

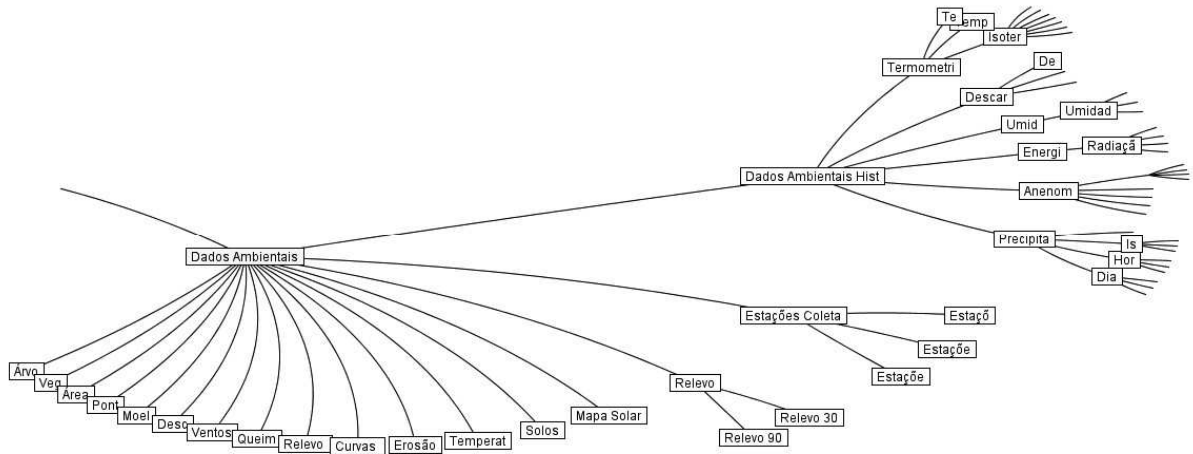


Figura 3-7 - Dados Ambientais

A subárvore da Figura 3.8 apresenta os termos, e suas relações, que fazem parte da Distribuição de energia elétrica, como Transformadores, Rede Subterrânea e Vãos.

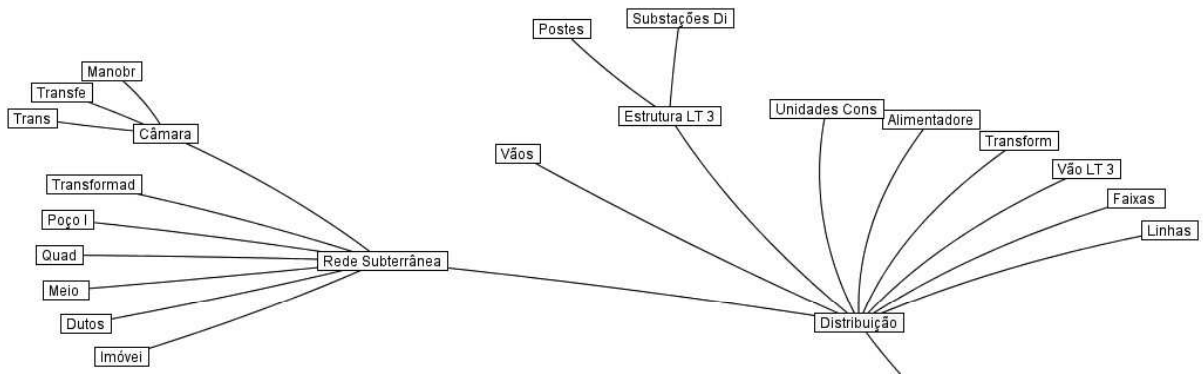


Figura 3-8 - Distribuição

A Figura 3.9 destaca os termos ligados às Imagens Georreferenciadas, como as Ortofotos.

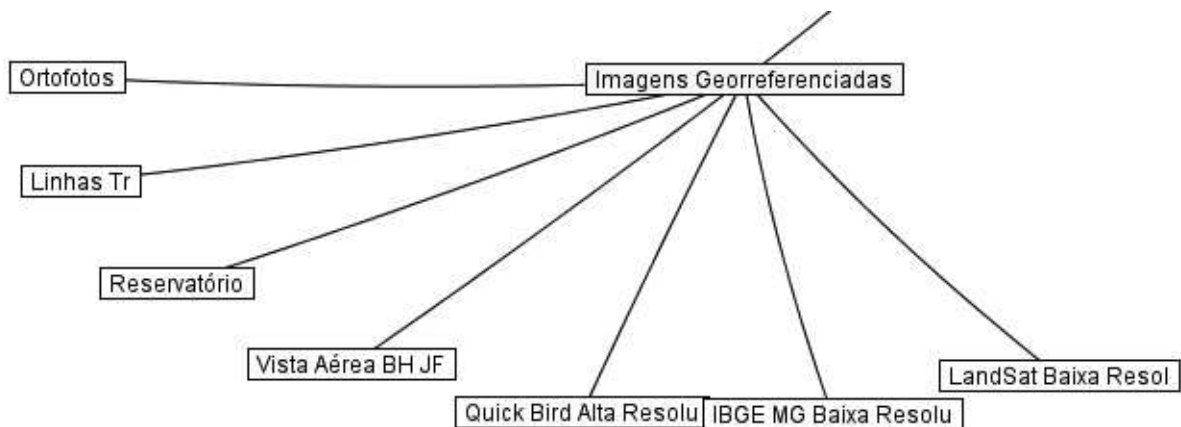


Figura 3-9 - Imagens Georreferenciadas

Os itens que compõe a Infraestrutura no setor elétrico estão apresentados na Figura 3.10. A Cemig mantém os dados dos aeroportos, helipontos, ferrovias e rodovias do estado de Minas Gerais.

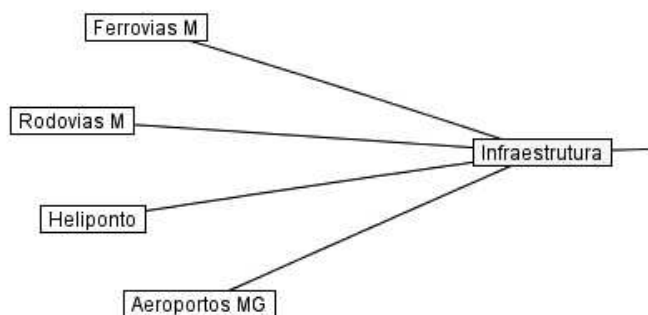


Figura 3-10 - Infraestrutura

A Figura 3.11 mostra os componentes, e suas relações, das Instalações que fazem parte da CEMIG.

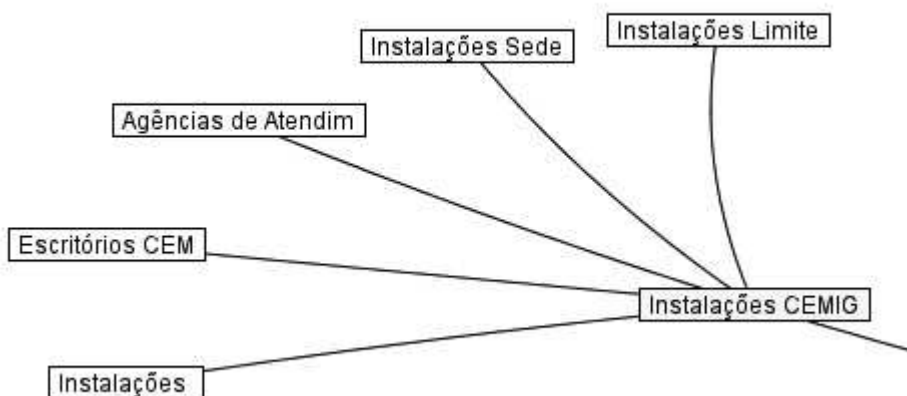


Figura 3-11 - Instalações CEMIG

A Organização Político-Administrativa é composta por termos como, Minas Gerais, Estado Minas, Municípios, Regiões entre outros (Figura 3.12).

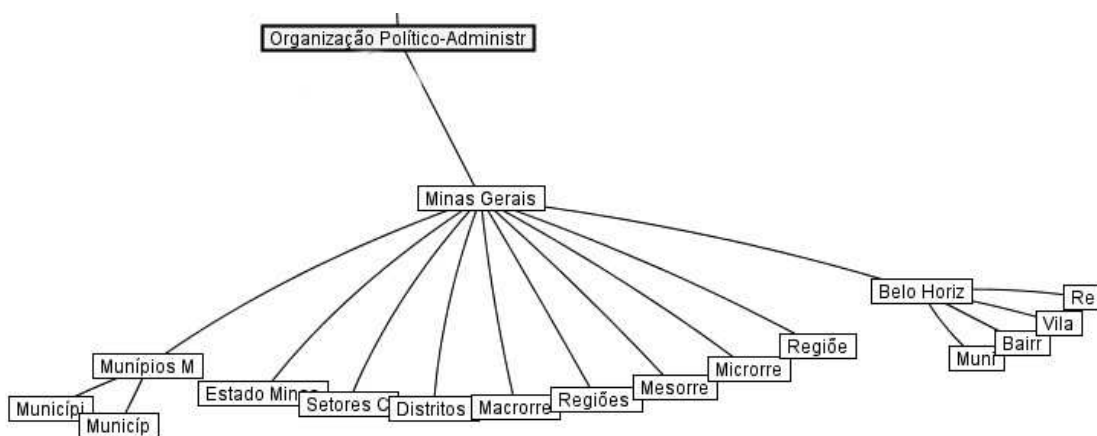


Figura 3-12 - Organização Político-Administrativa

Os Recursos Hídricos possuem Termos como Minas Gerias, Hidrografia, Bacias Hidrográficas entre outros. A Figura 3.13 apresenta esses termos e a relação entre eles. Estes termos apresentam tipos de hidrografia, e representam os reservatórios mantidos e utilizados pela Cemig.

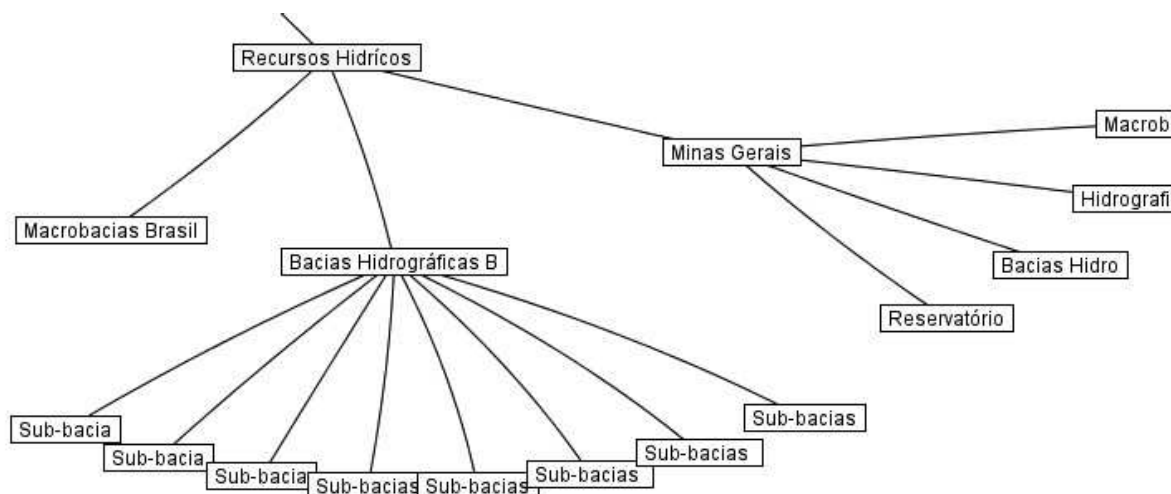


Figura 3-13 - Recursos Hídricos

A subárvore da Figura 3.14 mostra os elementos e relações sobre as Referências Cartográficas, que ajudam na criação dos mapas utilizados pela Cemig, contendo, por exemplo, as linhas de latitude e longitude, Rede Altimétrica e os marcos geodésicos.

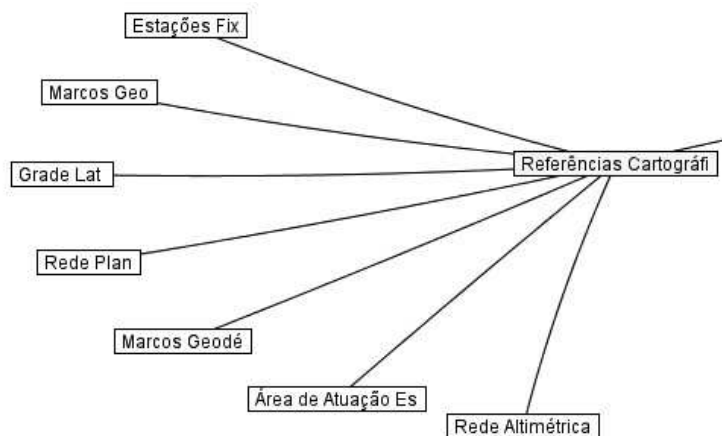


Figura 3-14 - Referências Cartográficas

A Malha Distribuição (Figura 3.15) é composta por termos como Malha Regional, Conjuntos ANEEL, Locais CEMIG, entre outros, e seus relacionamentos.

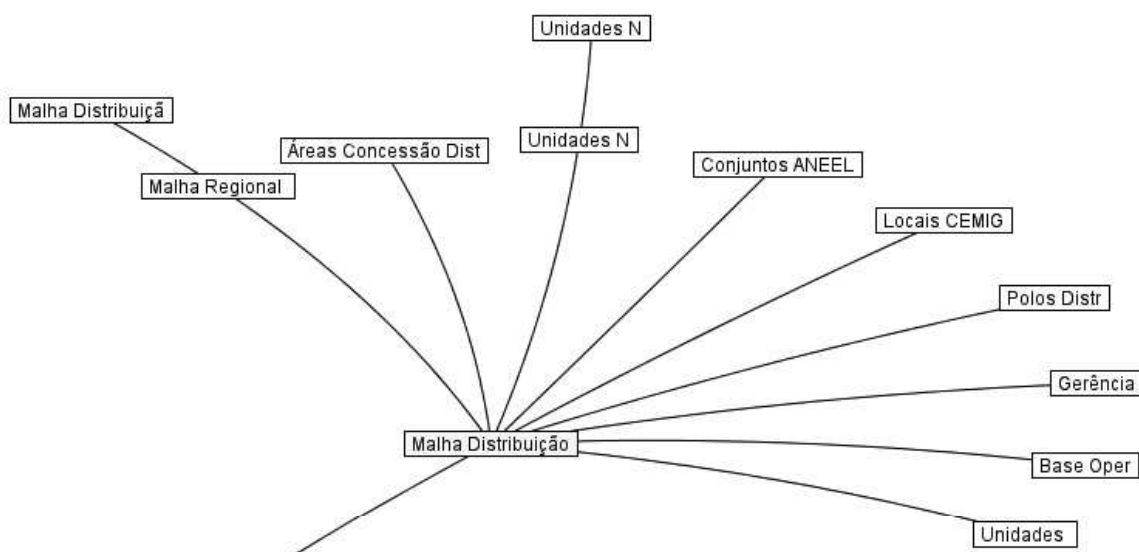


Figura 3-15 - Malha Distribuição

A subárvore na Figura 3.16 mostra os termos relacionados a área de Transmissão de energia elétrica, como Linhas de Transmissão e Subestações de Transmissão.

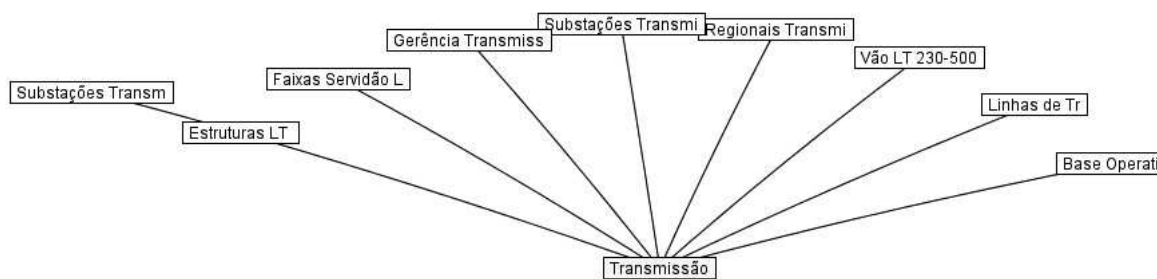


Figura 3-16 - Transmissão

3.3 Indexando os Metadados na Árvore Hiperbólica

O editor edpMGB apresenta a função de indexar os metadados produzidos e validados por ele na estrutura da árvore hiperbólica criada para o sistema elétrico. Ao clicar no botão indexar (Figura 3.17), o editor verifica se os metadados estão validados e logo em seguida os armazena. Após armazenar o metadado, o editor cria um endereço indicando onde esse metadado poderá ser acessado futuramente.

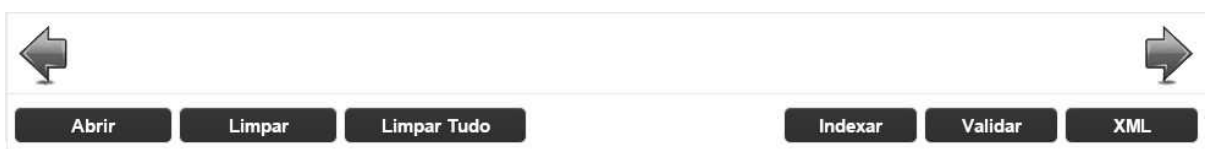


Figura 3-17 - Painel de botões do editor edpMGB

Com esse endereço criado, o processo de indexação realmente se inicia. O Editor utiliza os termos digitados no campo “título” do metadado e logo em seguida faz uma busca através dos nós da árvore hiperbólica. Ao encontrar um nó que possui um ou mais termos que estão no campo “título”, o editor armazena dentro desse nó o endereço onde o metadado está armazenado. Por fim, o editor exporta esta árvore em formato XML já com os endereços dos metadados armazenados. A Figura 3.18 exemplifica esse processo. Ao acessar o nó com o termo “Usinas”, por exemplo, o usuário irá encontrar, na janela Atributos, o campo Metadados que armazena o link para os metadados que estão relacionados a esse termo específico.

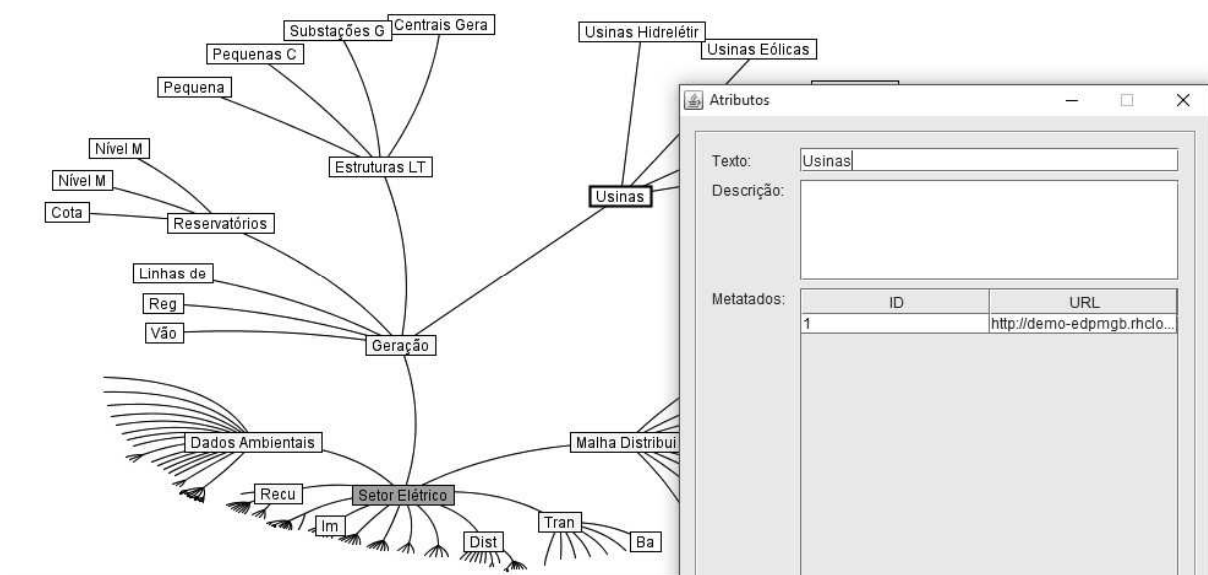


Figura 3-18 - Quadro de atributos

A medida que os usuários acessam o editor e geram os metadados, os links ficam armazenados podendo ser visualizados a qualquer momento. Essa visualização é feita através do programa *Hyper Tree Studio* que está disponível para download no menu Ajuda do editor, como mostra Figura 3.19.



Figura 3-19 - Menu Ajuda

Através desse menu o usuário também poderá fazer o download da Árvore de termos, em formato XML, com todos os links dos metadados validados e armazenados.

4 CONCLUSÕES

A partir dos estudos e pesquisas realizadas, um sistema para recuperação de informação em catálogos de metadados para o setor elétrico foi especificado. Este sistema descreve um mecanismo de busca que faz uso de árvore hiperbólica. Com isso, o objetivo geral, proposto nessa pesquisa, foi atingido.

O editor edpMGB foi desenvolvido especificamente para a criação de metadados geoespaciais de acordo com o Perfil MGB. Além de utilizar um software desenvolvido em seu próprio idioma, o usuário tem a facilidade de acesso ao sistema sem a necessidade de instalação em seu próprio computador.

A funcionalidade de validar se o metadado está em conformidade com o padrão definido pelo Perfil MGB auxilia na produção de metadados com melhor qualidade, mais completos e corretos.

Após realizar o levantamento bibliográfico e definir os documentos que fariam parte do *corpus*, foi realizado o processamento desse texto tendo como resultado o conjunto de termos que foram organizados em forma de árvore hiperbólica facilitando o processo de indexação e busca.

Ao fazer a busca por um termo dentro da árvore hiperbólica, o usuário tem a opção de navegar pelos termos que estão relacionados ao termo pesquisado. Sendo assim, caso o usuário não tenha conhecimento prévio sobre o termo pesquisado os termos que estão interligados podem trazer o resultado esperado para a pesquisa.

No Capítulo 3 está descrito o módulo de indexação presente no editor edpMGB. Essa indexação permite recuperar os metadados de uma IDE organizados dentro dos nós da árvore hiperbólica. A viabilidade do sistema desenvolvido foi verificada por meio do desenvolvimento de um caso de estudo, no contexto de uma IDE do setor elétrico, utilizando a base de dados da CEMIG.

Como sugestão de trabalhos futuros pode-se apontar a expansão e evolução das especificações do editor, a fim de que seja oferecido um mecanismo de criação, edição e busca cada vez mais completo de metadados no Perfil MGB. Além disso, melhorar cada vez mais o processo de recuperação de informação em um catálogo de metadados dentro do setor elétrico. Outra extensão é fazer a integração do editor com módulos de extração automática utilizando o retângulo envolvente dos dados geoespaciais e o tratamento de coleções de metadados fortemente relacionadas.

BIBLIOGRAFIA

- BATISTA, G. E. A. P. A. **Pré-processamento de dados em aprendizado de máquina supervisionado**. 2003. Tese de Doutorado. Universidade de São Paulo, São Carlos, SP, Brasil.
- BRASIL. Decreto Presidencial nº 6.666, de 27 de novembro de 2008. Institui no âmbito do Poder Executivo federal a Infra-Estrutura Nacional de Dados Espaciais – INDE, e dá outras providências. Diário Oficial, Brasília, DF, 27 nov. 2008. p. 57. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2008/Decreto/D6666.htm>. Acesso em: 15 set. 2015.
- CATMDEDIT. CatMDEdit OpenSource Project. 2012. Disponível em: <<http://catmdedit.sourceforge.net/index.html>>. Acesso em 04 ago. 2015.
- CARDOSO, O. N. P. Recuperação de Informação. **INFOCOMP Journal of Computer Science**, v. 2, n. 1, p. 33-38, 2004.
- CONRADO, M. S. **O efeito do uso de diferentes formas de extração de termos na compreensibilidade e representatividade dos termos em coleções textuais na língua portuguesa**. 2009. Tese de Doutorado. Universidade de São Paulo, São Paulo, SP, Brasil.
- CROMPVOETS, J., e STUIVER, J. Worldwide impact assessment of geoportals. **AGILE Conference Proceedings**, Lisboa, pp. 565-568, 2005.
- DAVIS, C. A. e ALVES, L. L. Local spatial data infrastructures based on a service oriented architecture. In: Simpósio Brasileiro de Geoinformática, 7., 2005, Campos do Jordão. **Anais...** Campos do Jordão, p. 30-48, 2005.
- DORNELLES, M. A. e IESCHECK, A. L. Análise da aplicabilidade da Infraestrutura Nacional de Dados Espaciais (INDE) para dados vetoriais em escalas grandes. **Boletim de Ciências Geodésicas**, Curitiba, v. 19, no 4, p.667-686, 2013.
- EVANGELISTA. S. R. M., Manual do HiperEditor e do HiperNavegador. **Embrapa Informática Agropecuária**, Campinas, 2007.
- FERNEDA, E. **Recuperação de Informação: Análise sobre a contribuição da Ciência da Computação para a Ciência da Informação**. 2003. Tese de Doutorado. Universidade São Paulo, São Paulo, SP, Brasil.
- FIGUEREIDO, D. E. A. **Recuperação da informação: uma análise sobre os sistemas de busca da web**. 2006. Monografia de Graduação. Universidade de Brasília, Brasília, DF, Brasil.
- FRAKES, W. B. e BAEZA-YATES, R. **Information Retrieval Data Structures & Algorithms**. Prentice Hall, p. 464, 1992.
- FREITAS, C. M. D. S., CHUBACHI, O. M., LUZZARDI, P. R. G. e CAVA, R. A. Introdução à visualização de informações. **Revista de informática teórica e aplicada**, Porto Alegre, vol. 8, n. 2, p. 143-158, 2001.
- HAN, J., KAMBER, M. e PEI, J. **Data mining: concepts and techniques: concepts and techniques**. 3. ed. Morgan Kaufmann, 2011.

- HOCHMAIR, H. Ontology matching for spatial data retrieval from Internet portals. In: First International Conference, 2005, Cidade do México. **Proceedings...** Cidade do México, p. 166-182, 2005.
- ISO 19115:2003. **Geographic information - Metadata**. Disponível em http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020, 2003. Acesso em: 15 out. 2015.
- LOPES, L., FERNANDES, P., VIEIRA, R., e FEDRIZZI, G. ExATOLP—An Automatic Tool for Term Extraction from Portuguese Language Corpora. In: 4th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics (LTC), 2009, Poznań, Polônia. **Proceedings...** Poznań: Faculty of Mathematics and Computer Science of Adam Mickiewicz University, 2009.
- MOURA, M. F. **Uma abordagem para a construção e atualização de taxonomias de tópicos a partir de coleções de textos dinâmicas**. 2006. Monografia de Qualificação de Doutorado. Universidade de São Paulo, São Carlos, SP, Brasil.
- NAKAMURA, E. T., e FILHO, A. P. Q. Infraestrutura De Dados Espaciais: Exemplo Do Parque Estadual De Intervalos-Sp. **Revista Brasileira de Cartografia**, n. 64 p. 723-735, 2012.
- NASCIMENTO, H. A. e FERREIRA, C. B. Visualização de Informações—uma abordagem prática. In: Congresso da Sociedade Brasileira de Computação, 25., 2005, São Leopoldo. **Anais...** São Leopoldo: UNISINOS, 2005.
- NEBERT, D. D. Technical Working Group Chair GSDI. Developing Spatial Data Infrastructures: The SDI Cookbook. V.2. **GSDI – Global Spatial Data Infrastructure**. Disponível em: <http://www.gsdi.org/docs2004/Cookbook/cookbookV2.0.pdf>. Acesso em: 27 de outubro 2015. 2004.
- PASCOAL, A. P., CARVALHO, R. B. e XAVIER, E. M. A. Materialização do Perfil de Metadados Geoespaciais do Brasil em esquema XML derivado da ISO 19139. In: Simpósio Brasileiro de Sensoriamento Remoto, 16., 2013, Foz do Iguaçu, **Anais...** São José dos Campos: INPE, p. 2441- 2448, 2013.
- PEIXOTO, D. A., DA MATTA VEGI, L. F. e LISBOA-FILHO, J. Um Editor de Metadados para Documentar Padrões de Análise em uma Infraestrutura de Reuso. **iSys-Revista Brasileira de Sistemas de Informação**, v. 7, n. 4, p. 23-42, 2014.
- PRADO, B. R.; HAYAKAWA, E. H.; BERTANI, T. C.; SILVA, G. B. S.; PEREIRA, G.; SHIMABUKURO, Y. E. Padrões para metadados geográficos digitais: modelo ISO19115:2003 e modelo FGDC. **Revista Brasileira de Cartografia**, n. 62, v. 1, p. 33-41, 2010.
- REZENDE, S. O.; PUGLIESI, J. B.; MELANDA, E. A.; PAULA, M. F. Mineração de dados. In: Rezende, S. O. **Sistemas Inteligentes: Fundamentos e Aplicações**, 1 ed, c. 12, p. 307-335, 2003.
- SOARES, P. G., TANAKA, A. e BAIÃO, F. Estudo dos Principais Conceitos sobre Integração de Dados Geoespaciais. **RelaTe-DIA**, v. 4, n. 1, 2011.
- SOUZA, T. B., CATARINO, M. E. e SANTOS, P. C. Metadados: catalogando dados na Internet. **Transifirmação**, v. 9, n. 2, p. 93-105, 1997.

APÊNDICE A

Este apêndice apresenta os termos extraídos do glossário de termos da ANEEL pela ferramenta EχATOLP, além dos termos presentes na modelagem do banco de dados utilizado pela CEMIG.

Termos	Frequência	Termos	Frequência
Energia	182	Dia	10
Energia elétrica	144	Fornecimento	10
Sistema	57	Geradores	10
Distribuidora	49	Permissionária	10
Unidade	46	Tarifa	10
Serviço	44	Etapa	9
Aneel	42	Informações	9
Unidade consumidora	36	Interesse	9
Empresa	33	Prestação de serviço	9
Consumidor	32	Termo	9
Produto	32	Usuário	9
Leilão	28	Caso	8
Lei	26	CCEE	8
Instalação	21	Critério	8
Tensão	20	Editais	8
Valor	20	Energia de reserva	8
Decreto	19	Entidade	8
Equipamento	19	Fornecimento de energia	8
Agente	18	Ocorrência	8
Concessionária	18	Sociedade	8
Rede	18	Solicitante	8
Atividade	17	Tempo	8
Fonte	17	Unidade geradora	8
Empreendimento	16	Centrais geradoras	7
Sistema elétrico	16	Classe	7
Usinas	16	Consumidor livre	7
Condições	15	Garantia	7
Distribuição de energia	15	Linha	7
Poder concedente	15	Mercado	7
Projeto	15	Perda	7
Quantidade	15	Prazo	7
Acessante	14	Primeira fase	7
Geração	14	Primeira fase de Leilão	7
Potência	14	Rodada	7
Contrato	13	Total de chamadas	7
Distribuição de energia elétrica	13	Ano	6
Operação	13	Comprador	6
Período	13	Conforme	6
Sistema de distribuição	13	Consumo de energia	6
Ativos	12	Consumo de energia elétrica	6
Carga	12	Demanda de potência	6
Serviço público	12	Disponibilidade	6
Bens	11	Finalidade	6
Centrais	11	Geração de energia elétrica	6
Chamada	11	Horas de utilização de dia	6
Concessão	11	Instalação de transmissão	6
Forma	11	Legislação	6
Quantidade total	11	Partes	6
Concessionária	10	Pessoa jurídica	6
Processo	6	Distribuição	4
Produção	6	Empreendimento hidrelétrico	4

Propriedade	6	Entidade executora	4
Rede elétrica	6	Estudos	4
Resolução	6	Faixas de tensão	4
Serviço público de energia	6	Garantia financeira	4
Sistemática	6	Gás natural	4
Utilização de dia	6	Importação de energia elétrica	4
Amostra	5	Instalação elétrica	4
Arts	5	Interesse restrito	4
Campo	5	Intervalo de tempo	4
Consumo	5	Lance	4
Custo	5	Liquidação	4
Estrutura	5	Meio	4
Gás	5	Modalidade	4
Geração de energia	5	Modalidade de concorrência	4
Iluminação	5	MW	4
Liquidação financeira	5	MW médio	4
Mme	5	Normas	4
Montante de energia	5	Objeto	4
Organizacional	5	Padrão	4
Participação	5	País	4
Período de apuração	5	Prazo determinado	4
Poder	5	Preço de lance	4
Poder público	5	Proponente	4
Ponto de conexão	5	Requisitos	4
Potência instalada	5	Residencial baixa renda	4
Prestação	5	Rodadas uniformes	4
Proponente vendedor	5	Serviço público de distribuição de energia elétrica	4
Sobras	5	Serviços auxiliares	4
Sobras líquidas	5	Setor	4
Subclasse	5	SIN	4
Subestação	5	Sobras líquidas contratuais	4
União	5	Subclasse rural	4
Uso	5	Tensão igual	4
Utilidade	5	Tensão nominal de operação	4
Acessada	4	Termos percentuais	4
Agência	4	Valor eficaz de tensão sistema	4
Agentes de distribuição	4	Vida	4
Área	4	Ação	3
Atendimento	4	Acesso	3
Autoprodutor	4	Agente de ccee	3
Capacidade	4	Agente de Liquidação	3
Consumidor final	4	Agropecuária	3
Conta	4	Alta precisão	3
Contratação de energia	4	Aproveitamento	3
Contratos iniciais	4	Atendente	3
Controle	4	Atividades descentralizadas	3
Avaliação	3	Medição	3
Brasil	3	Medidor	3
Campo elétrico	3	Modo	3
Capital	3	Montante de energia elétrica	3
Característica	3	Necessidade	3
Caráter	3	Obra	3
Chamadas recebidas	3	Oferta de referência	3
Compra de energia elétrica em que	3	Ons	3
Conforme regulamentação de ANEEL	3	Perdas elétricas	3
Consumidor cativo	3	Pessoa física	3
Contratação	3	Plano diretor	3
Contratação de energia de reserva	3	Ponto	3
Coordenadas	3	Produção independente	3
Corrente	3	Produto de fonte	3
Critérios técnicos	3	Projeto básico	3
Data	3	Proposta de preço	3

Desligamento	3	Rede básica	3
Determinação de oferta de referência	3	Regime operativo médio	3
Difusão	3	Reservatório	3
Duração	3	Respectivos fatores de ponderação	3
Efeitos	3	Revisões	3
Eletrificação rural	3	Semana	3
Empresa cooperada	3	Serviço público de distribuição de energia	3
Energia elétrica em que	3	Setor elétrico	3
Energia livre	3	Sistema de acompanhamento de informações de mercado	3
Entrada em operação	3	Situação	3
Estados	3	Supridora	3
Existência	3	Suprimento	3
Fonte hidroelétrica	3	Tarifas diferenciadas de consumo de energia	3
Forma de energia	3	Tempo especificado	3
Fornecimento de energia elétrica	3	Tensão de sistema	3
Função	3	Tensão evento	3
Garantia física	3	Titular	3
Geração de energia elétrica a gás na tural	3	Transformação	3
Iluminação pública	3	Transmissão	3
Implantação	3	Transmissão de energia elétrica	3
Importação de energia	3	Uso de bem público	3
Indisponibilidade	3	Uso de sistema de distribuição	3
Instituição	3	Uso exclusivo	3
Intervalo de tempo especificado	3	Valor eficaz	3
Lei municipal	3	Vida útil	3
Acionamentos	2	Cláusulas	2
Acordo de Acionistas	2	Cogeração	2
Adequação a padrões	2	Combustíveis	2
Adequação a padrões técnicos	2	Componente	2
Agente comprador	2	Componente fundamental	2
Agente custodiante	2	Componentes de sistema	2
Agente importador	2	Compra	2
Agente responsável	2	Conceito	2
Agentes de distribuição de energia	2	Concessionárias de distribuição	2
Água	2	Concessionárias de serviço público de energia elétrica	2
Alta tensão de distribuição	2	Concessionárias de transmissão	2
Altitude	2	Condição anormal de operação	2
Ambiente de contratação	2	Condição de operação	2
Âmbito de Empresa	2	Condições estabelecidas	2
Aparelho	2	Condutores	2
Aparelho de aquecimento de água	2	Conforme definido	2
Aplicação de tarifas diferenciadas de consumo de energia	2	Conjunto	2
Área rural	2	Conjunto de unidades	2
Associadas	2	Conjunto de unidades consumidoras	2
Atendimento de demanda de produto	2	Consequência de contingências	2
Atendimento humano	2	Contingências severas	2
Atividade agropecuária	2	Contratação livre	2
Auditagem de dados	2	Contrato de concessão	2
Autorização de ANEEL	2	Controle dispositivos	2
Autorização de Poder Concedente	2	Convenção de Comercialização	2
Bem público	2	Convenção do mercado	2
Bens reversíveis	2	Corte de carga	2
Capacidade de transmissão	2	Critério estabelecido	2
Capacidade instalada	2	Curto prazo	2
Capital próprio aplicado	2	Custos de uso	2
Carga instalada em unidade	2	Cvu não nulo	2
Carga instalada em unidade consumidora	2	Dac sistema telefônico programável que recebe chamadas	2
Caso sob regimes de autoprodução ou de produção	2	Dados	2

Caso sob regimes de autoprodução ou de produção independente	2	Data de início de vigência de revisão tarifária periódica	2
Central termelétrica	2	Débitos	2
Central termelétrica cogeneradora	2	Definição de solicitante de resolução	2
Centro de operação	2	Demanda de produto	2
		Depreciação acumulada	2
Descontinuidade de distribuição de energia elétrica	2	Estabelecido	2
Desempenho	2	Estrutura física	2
Desenvolvimento de atividade	2	Estrutura física específica	2
Detalhamento de sistemática	2	Etapa inicial	2
Deteção de condição anormal de operação	2	Etapa uniforme	2
Determinada área	2	Eventos	2
Determinada metodologia de auditoria de dados	2	Exploração de serviço público de distribuição de energia	2
Determinado conjunto de unidades	2	Exploração de serviços de energia	2
Determinado conjunto de unidades consumidoras	2	Exploração de serviços de energia elétrica	2
Determinado local	2	Faixas de tensão críticas	2
Dimensões para avaliar tal difusão	2	Fator de potência	2
Direitos	2	Finalidade de preservar integridade de equipamentos e linhas de transmissão	2
Disponibilidade de energia	2	Fins	2
Disponibilidade de energia elétrica	2	Fiscalização de aneel	2
Disponibilidade termelétrica	2	Fiscalização de serviços e instalações de energia	2
Distribuição de acessada	2	Fonte hidroelétrica caso quantidade total ofertada em primeira fase de leilão	2
Duração relativa de leituras de tensão	2	Fonte primária	2
Eficiência	2	Força	2
Eficiência energética	2	Forma periódica	2
Eletrificação	2	Frequência	2
Empreendimento a gás	2	Funcionamento	2
Empreendimento eólico	2	Geração de energia elétrica a gás	2
Empresa de auditoria	2	Geração distribuída	2
Empresa declarante	2	Geração hidrelétrica	2
Empresa proponente	2	Gerente de Programa	2
Empresas de setor de energia elétrica	2	Gerente de projeto	2
Encargo	2	Gestão associada de serviços	2
Energia ativa	2	Gestão associada de serviços públicos	2
Energia contratada	2	Grupo	2
Energia elétrica a gás	2	Grupo de controle	2
Energia elétrica a gás natural	2	Ibge	2
Energia elétrica ativa	2	Inciso	2
Energia elétrica em Brasil	2	Início de vigência de Revisão Tarifária Periódica	2
Instalações de conexão	2	Período de observação definido	2
Instalações de energia	2	Planejamento de operação de sistema de distribuição	2
Instalações de energia elétrica	2	Ponto de acesso	2
Instalações de geração	2	Ponto de rede	2
Instalações próprias	2	Ponto de rede existente	2
Integridade de equipamentos e linhas de transmissão	2	População	2
Interrupções	2	Potência ativa	2
Interrupções ocorridas	2	Prestação de serviço público	2
Intervalos	2	Prestação de serviço público de energia	2
Investimento	2	Prestação de serviço público de fornecimento de energia	2
Lance único	2	Prestação de serviço público de fornecimento de energia elétrica	2
Legislação vigente	2	Procedimentos	2

Ligação	2	Procedimentos de comercialização	2
Liquidação financeira relativa a contratação de energia	2	Procedimentos de comercialização específicos	2
Liquidação financeira relativa a contratação de energia de reserva	2	Procedimentos de Rede	2
Logradouros públicos	2	Proposta	2
Lucros	2	Público de Energia Elétrica	2
Manobra	2	Quadro efetivo de empresa	2
Manual de Operacionalização do Programa Luz	2	Quantidade de energia	2
Média tensão	2	Quantidade de energia elétrica	2
Montante de uso	2	Reajuste	2
Não conformidades encontradas	2	Realização de atividades	2
Natural ciclo	2	Receita	2
Necessária	2	Rede de distribuição	2
Necessário para atendimento de demanda de produto	2	Redução de consumo de energia	2
Necessidade total de energia de produto	2	Redução de consumo de energia elétrica	2
Necessidades de corte de carga	2	Referência valor	2
Novas vias	2	Referência valor usado	2
Objetiva	2	Regime	2
Objeto de autorização	2	Regime operativo	2
Objeto regular	2	Regimes de autoprodução ou de produção	2
Obrigações	2	Regimes de autoprodução ou de produção independente	2
Ocorrência emergencial	2	Registro	2
Operação comercial	2	Regulação	2
Operação de sistema	2	Regulamentos	2
Regulamentos aplicáveis	2	Uar	2
Requisitos dispostos em arts	2	Usina acionada por energia	2
Respectivas operações	2	Usina hidrelétrica	2
Respectivos encargos	2	Uso de bem	2
Ressarcimento	2	Uso exclusivo de autoprodutor	2
Revisão tarifária periódica	2	Valor de componente	2
Rodada subsequente	2	Valor de componente fundamental	2
Seguintes requisitos	2	Valor eficaz de tensão	2
Seguintes subclasses	2	Tarifas diferenciadas de consumo de energia elétrica	2
Segunda fase	2	Tempo inferior	2
Serviços de energia elétrica	2	Tensão crítica	2
Setor de energia elétrica	2	Tensão de curta duração	2
Signatários de Acordo de Acionistas	2	Tensão de referência	2
Sistema de distribuição de energia	2	Terminais de linha	2
Sistema de distribuição de energia elétrica	2	Termos líquidos	2
Sistema de Medição	2	Terra	2
Sistema de transmissão de conexão	2	Terreno	2
Sistema elétrico brasileiro	2	Titular de autorização expedida por ANEEL	2
Sistema elétrico de distribuição	2	Titular de unidade	2
Sistema UTM	2		
Somatório	2		
Submissão de lances	2		
Superfície	2		