DOI: 10.5433/1981-8920.2021v26n2p433

ORGANIZAÇÃO DE CONHECIMENTO BASEADA EM ONTOLOGIAS: UM ESTUDO DE CASO SOBRE OS DESAFIOS DA CONCEITUALIZAÇÃO NO DOMÍNIO DA ENERGIA ELÉTRICA

ONTOLOGY-BASED KNOWLEDGE ORGANIZATION: A CASE STUDY ABOUT THE CHALLENGES OF THE CONCEPTUALIZATION IN THE DOMAIN OF POWER SUPPLY

Jeanne Louize Emygdio[a]
Lívia Marangon Duffles Teixeira[b]
Cristiano Moreira[c]
Murillo Lima Modesto[d]
Maurício Barcellos Almeida[e]

RESUMO
Objetivo: O presente artigo objetiva descrever um estudo de caso de conceitualização no setor elétrico, em que cerca de 1.200 termos técnicos de engenharia foram extraídos para uma ontologia. **Metodologia**: Para tal, delimita-se o arcabouço teórico, apresenta-se o contexto do projeto, apresentam-se fragmentos do conjunto de termos candidatos, descreve-se a aquisição de conhecimento, bem como técnicas para definir termos e derivar categorias. **Resultados**: A pesquisa resultou no registro de técnicas aplicadas aos termos candidatos, no entendimento das peculiaridades do domínio e, à medida do possível em um estudo de caso, na obtenção de experiência e boas práticas para uso em domínios similares em áreas estratégicas da sociedade. **Conclusões**: evidencia-se a necessidade de uma metodologia de construção de ontologias voltada aos profissionais da Ciência da Informação, o que inclui a fase de conceitualização.

Descritores: Organização do Conhecimento. Ontologias. Conceitualização.

[a] Doutoranda do Programa de Pós-Graduação em Gestão e Organização do Conhecimento da Universidade Federal de Minas Gerais (UFMG). Email: jeanne.emygdio@gmail.com

[b] Doutora em Gestão e Organização do Conhecimento pela Universidade Federal de Minas Gerais (UFMG). Email: liviamarangon@gmail.com

[c] Doutorando do Programa de Pós-Graduação em Gestão e Organização do Conhecimento da Universidade Federal de Minas Gerais (UFMG). Email: cristianomoreirasilva@hotmail.com

[d] Mestre em Gestão e Organização do Conhecimento pela Universidade Federal de Minas Gerais (UFMG). Email: murilloef@gmail.com

[e] Doutor em Ciência da Informação. Docente da Escola de Ciência da Informação da Universidade Federal de Minas Gerais. Email: mba@eci.ufmg.br.

Representação do conhecimento. Setor elétrico.

1 INTRODUÇÃO

O setor elétrico consiste, basicamente, no fornecimento de energia elétrica para a sociedade através dos processos de geração, transmissão e distribuição da energia. No contexto brasileiro ele é regulado pela Agência Nacional de Energia Elétrica (ANEEL), que concede os direitos de exploração além de fiscalizar e auditar os serviços prestados pelas empresas. Por regulação, as empresas de energia são gestoras dos recursos que pertencem à União (como os bens e as instalações relacionadas - que serão a partir desse ponto denominados “ativos”) e não deles proprietárias, sendo remuneradas tanto pela manutenção desses recursos em operação quanto por novos investimentos relacionados aos mesmos.

Essa remuneração, entretanto, se dá pelas informações relativas ao ciclo de vida dos ativos que são apresentadas ao órgão regulador em periodicidade e método estabelecido. Entretanto, esse conjunto de informações sobre o mesmo ativo é produzido por diferentes frentes que sobre ele atuam, como engenharia de planejamento e de projeto, aquisição, manutenção, contabilidade dentre outras. Esse cenário, generalizado em empresas não só do setor elétrico, é caracterizado pela presença de silos de informações em diferentes formatos, diferentes tecnologias e muitas vezes em diferentes soluções proprietárias, fazendo uso de diferentes terminologias e padrões de interoperabilidade. Nesse sentido, a integração e a possibilidade de prover interoperabilidade (tanto técnica quanto semântica) entre esses sistemas fica comprometida e impacta, sobremaneira, no acesso e na qualidade dos registros de informações que são disponibilizados à ANEEL.

O objetivo deste artigo consiste em evidenciar os preceitos teóricos e as experiências vividas na etapa de Conceitualização (fundamental em diferentes métodos de construção de ontologias) em um projeto de Pesquisa e Desenvolvimento (P&D) em empresa de grande porte do setor elétrico brasileiro. Para tal, desenvolve-se um background teórico acerca da etapa e seus fundamentos, discorre-se sobre o setor elétrico brasileiro e apresenta-se a

metodologia aplicada, os resultados obtidos e as considerações sobre os mesmos. O projeto foi concebido no âmbito da Ciência da Informação e contempla, em linhas gerais, o desenvolvimento de uma ontologia para fomentar a interoperabilidade semântica e garantir a qualidade, rastreabilidade e auditabilidade das informações sobre os ativos. Espera-se ao final, além do artefato ontológico, contribuir não apenas para a empresa em que o P&D se desenvolve, mas para todo o setor elétrico, que carece de representação de conhecimento do setor uniformizando termos, definições e relações entre as entidades (sejam essas ativos ou processos).

O artigo está organizado como segue: a primeira e segunda seções apresentam o embasamento teórico sobre os principais temas explorados no estudo; a terceira seção expõe o contexto onde o estudo prático se desenvolve; a quarta seção refere-se à metodologia adotada pela equipe e os principais resultados obtidos; na quinta seção são tecidas as discussões sobre os resultados obtidos e, por fim, na sexta seção, são apresentadas as considerações finais sobre os achados e as contribuições da pesquisa.

2 BACKGROUND

As ontologias são uma área de pesquisa que pode ser analisada sob três perspectivas, a saber, a Filosofia, a Ciência da Computação (CC) e a Ciência da Informação (CI). Na Filosofia a pesquisa se dedica ao estudo das entidades que existem na realidade e das relações entre as mesmas. Na CC as ontologias se configuram em artefatos de software em linguagem formal para a aplicação em Inteligência Artificial (IA) e para a modelagem de sistemas de informação. Já na CI elas são definidas como estruturas de representação do conhecimento por meio da organização das entidades que existem por meio de suas categorias, suas classes e respectivas relações.

Diferentes tipos de metodologias contemplam o desenvolvimento de ontologias, tais como a *Methontology* (GÓMEZ-PEREZ et al., 1996; FERNANDÉZ et al., 1997), a *NeOn Methodology* (SUÁREZ-FIGUEROA et al., 2008); a TOVE (GRUNINGER e FOX, 1995) e a *OntoForInfoScience* (MENDONÇA, 2015). Elas se diferenciam, basicamente, pela complexidade das

descrições de suas etapas, dentre as quais podem-se citar a Aquisição do Conhecimento (AC) e a Conceitualização. O desenvolvimento bem planejado e bem conduzido destas etapas confere à ontologia maior qualidade e aderência à realidade. Não consiste um objetivo desse trabalho desenvolver em profundidade cada metodologia e cada uma de suas etapas e, sim, destacar a etapa de Conceitualização, que ocorre na sequência da AC, motivos pelos quais a AC merece uma breve explanação.

Na CI a AC consiste nos atos consecutivos de: i) levantamento e análise de materiais de referência; ii) identificação dos especialistas de domínio; e iii) na aplicação de técnicas colaborativas de extração do conhecimento dos especialistas. Já na Conceitualização, etapa subsequente à AC, todo o conhecimento extraído e registrado por meio das definições dos termos é organizado em classes da ontologia. Essa estruturação do conhecimento permitirá a atribuição de relações, de propriedades e de restrições no artefato ontológico. Assim sendo, a Conceitualização será tão bem sucedida quanto a AC também o for.

Nesta seção serão apresentados conteúdos pertinentes ao objeto de estudo, tais como os fundamentos de categorização, as técnicas para a derivação das categorias em ontologias e um aprofundamento teórico na etapa de Conceitualização.

2.1 FUNDAMENTOS DE CATEGORIZAÇÃO

Antes de começar a simplesmente enumerar categorias em um domínio do conhecimento, é preciso abordar questões metodológicas para a criação de um sistema de categorias. Nesse sentido, uma questão é determinar um critério de adequação para um sistema de categorias de forma que seja útil na tarefa de descrever a realidade.

Para Aristóteles, as categorias realistas são universais instanciados por entidades reais do mundo e não categorias da mente ou da linguagem. Enquanto para Aristóteles a linguagem é apenas como uma pista para distinguir categorias, Kant usa formas de pensamento como pista para definir julgamentos que podem ser feitos sobre objetos. Ao invés de seguir nessa forma mediada, os filósofos

contemporâneos tentam adquirir um sistema de categorias de forma mais direta, pelo estudo do mundo. Johansson[6], por exemplo, apresenta um projeto de categorias realistas que dizem respeito a aspectos reais do ser, distingue categorias pelo método de abstrações sucessivas e obtém uma lista de nove categorias genéricas com algumas subdivisões. Chisholm[7], por sua vez, preconiza um sistema de categorias em que a categoria mais genérica é denominada *entia*. Hoffman[8] e Rosenkrantz[9] sugerem uma árvore com uma entidade no topo, mas como uma divisão diferenciada no primeiro nível.

Revelada a existência de tantos sistemas de categorização desde a Antiguidade, como se pode acreditar que existe um que é o mais adequado? Como tais sistemas ajudam na tarefa de entender a realidade? Torna-se necessário elencar critérios para avaliar esses diferentes sistemas de categorização. Um sistema desse tipo deve proporcionar, pelo menos, um arcabouço para estudar "o que existe" de forma sistemática e responder especificamente sobre os tipos de coisas que existem. Nas próximas seções serão apresentadas as técnicas de derivação de categorias aplicáveis às ontologias e, com fins comparativos e em função de futuras análises e discussões, discute-se brevemente a derivação de categorias na biblioteconomia.

2.2 TÉCNICAS PARA DERIVAÇÃO DE CATEGORIAS EM ONTOLOGIAS

Apresentam-se aqui três propostas contemporâneas de métodos para derivar categorias, os quais podem atender a critérios de adequação: i) Método da negação de propriedades; ii) Método da detecção de absurdos; e iii) Método dos critérios de identidade e existência.

O método da negação de propriedades, como o próprio nome indica, consiste em distinguir categorias pelo fato de uma entidade ter ou não ter certa característica. A questão mais difícil nesse método consiste em descobrir quais categorias usar e, na sequência, garantir as propriedades

[6] Ingvar Johansson (1943 - ?), filósofo sueco
[7] Roderick Milton Chisholm (1916 – 1999), filósofo norte-americano
[8] Joshua Hoffman (xx – ?), filósofo norte-americano
[9] Gary Rosenkrantz (1951 – ?), filósofo norte-americano

mínimas para certos tipos de entidades. Dois princípios lógicos amplamente aceitos garantem uma divisão entre grupos exaustivos e mutuamente exclusivos:

Princípio da não contradição: para qualquer sentença P, não é o caso que (p AND não P), garante que cada categoria é mutuamente exclusiva e que cada entidade pertence a no máximo uma categoria;

Lei do terceiro excluído: para qualquer indivíduo x e predicado F, (Fx OR não (Fx)), garante que cada entidade encontra lugar em pelo menos uma categoria, de forma que as categorias distintas são exaustivas.

O método da detecção de absurdos é um método fenomenológico que remonta a Husserl[10]. As diferenças entre as categorias de significado podem ser discriminadas pela identificação de casos no qual obtém-se um absurdo, ou algo sem sentido, ao substituir um termo por outro. Entretanto, o absurdo não é suficiente para fornecer um método seguro para distinguir categorias conceituais.

Já o método dos critérios de identidade e existência faz uso da estratégia de Husserl para distinguir categorias gramaticais e, em seguida, usa tais diferenças para verificação das condições de identidade e existência como forma de distinguir categorias ontológicas. Esse método tem raízes em Frege[11] e Russell, que consideravam a categoria dos nomes composta não apenas pelo o que normalmente se chama de nomes próprios, mas também por descrições definidas. As descrições definidas permitem a seleção de um único indivíduo, por exemplo, a pessoa mais alta na sala, a partir das descrições de um orador. Os referentes no mundo são então determinados como os objetos que satisfazem tais descrições.

Para Frege, termos de categorias gramaticais diferentes, por exemplo, substantivos e adjetivos, possuem um critério de aplicação, ou seja, um tipo de critério associado que indica quando o termo pode ser aplicado. Assim, nomes são feitos distintos de outros termos porque vêm associados a certo critério de

[10] Edmund Gustav Albrecht Husserl (1859 – 1938), filósofo alemão.

[11] Friedrich Ludwig Gottlob Frege (1848 – 1925), filósofo alemão.

aplicação. Esse método, contemporaneamente, foi incrementado por Dummett[12], quem sugeriu que a categoria de nomes poderia ser selecionada pelo fato de que nomes são associados à critérios de identidade.

Os critérios de aplicação e os critérios de identidade para nomes são fornecidos ao se associar o nome com uma variedade geral de coisa a ser nomeada pelo termo, por exemplo, um artefato, um evento, uma coleção de partículas etc. Essas variedades são nomeáveis por termos sortais genéricos. Um *sortal* fornece critério para o que pode ser contado, por exemplo, “maçã” é um *sortal* na sentença “Eu desejo duas maçãs”, enquanto “água” não é um *sortal* na sentença “Eu desejo água”. Cada nome pode ser associado a um *sortal* básico, de forma que a entidade nomeada pode existir tanto quanto o termo *sortal* continua aplicável a ela. Uma vez identificada a categoria gramatical do nome, pode-se continuar a distinguir diferentes categorias de nomes correlatos, e assim, distinguir categorias ontológicas.

A sugestão de Dummett é que dois termos, bem como os objetos que esses termos nomeiam, são da mesma categoria se eles estão associados ao mesmo critério de identidade. Talvez o caminho mais seguro seja levar em consideração ambas as condições, a saber, as condições de identidade e as condições de aplicação para os termos sortais associados quando na formulação de categorias. Sortais altamente genéricos podem então ser identificados com categorias. Por esse método, pode-se distinguir a categoria “animais” da categoria “coleção de partículas”, uma vez que essas entidades têm diferentes condições de identidade: um animal pode sobreviver à perda de algumas partículas, mas uma coleção de partículas, por definição, não pode.

Usando esse método pode-se resolver várias questões de cunho filosófico, mas a questão de interesse aqui é a questão metodológica de como extrair categorias. A abordagem de Dummett compartilha com a abordagem aristotélica o ponto de vista de que se pode extrair categorias ontológicas por meio de considerações da linguagem – nesse caso, primeiro pela referência à categoria gramatical de nomes e depois pelas condições de aplicação e de

[12] Michael Anthony Eardley Dummett (1925 – 2011), filósofo britânico

identidade associadas como diferentes sortais. A preocupação clássica com essa abordagem é a questão de saber se as categorias da linguagem realmente têm algo a dizer sobre as categorias de entidades reais do mundo. Entretanto, a própria abordagem já fornece respostas sobre essa preocupação: na visão de Frege, não se pode selecionar objetos exceto pela pressuposição de algum *sortal* geral. Assim, não pode-se esperar as categorias a partir do mundo simplesmente ao considerar objetos em primeiro lugar, sem qualquer pressuposição categorial, e então discernir as categorias as quais eles pertencem.

Essa abordagem também proporciona uma resposta fundamentada para a questão sobre qual seria a forma adequada para um sistema de categorias, mais especificamente, se deveria existir ou não uma única categoria de mais alto nível acima de todas as outras. Outro resultado de distinguir categorias a partir dessa forma é a garantia de que as categorias obtidas são mutuamente exclusivas. Com relação ao critério de adequação, é possível garantir que as categorias que foram feitas distintas são exaustivas? Existe um sentido no qual a condição é atendida e um sentido que não depende de como se interpreta o requisito da exaustividade. A questão é metodológica.

2.3 Técnicas para Derivação de Categorias na Biblioteconomia

As práticas de classificação em biblioteconomia são, basicamente, atribuídas por meio da relação entre o sujeito classificador e o mundo. A classificação de um assunto para organização física de bibliotecas e para a construção de linguagens documentárias – taxonomias, tesauros e vocabulários controlados – é fundamentada em teorias como a Teoria da Classificação Facetada de Ranganathan, a Teoria do Conceito (DAHLBERG, 1978) e a Teoria da Terminologia (WÜSTER, 1998; CABRÉ, 1993).

Como todas tomam como ponto de partida os tipos de coisas existentes, os conceitos pessoais e a linguagem humana, desenvolvem categorias da mente ou da própria linguagem sem relacionar os seus tipos aos universais referentes do mundo real. Essas abordagens podem gerar equívocos na construção de ontologias. Casos práticos desses equívocos são encontrados em diferentes

terminologias.

Por exemplo, *National Cancer Institute Thesauros* (NCIt)[13], a classe "Área Geográfica", contém termos como "América do Sul", "Porta", Playground, "Lar", "União Europeia", "Brasil", "Caixa de correio", "Igreja", "Quarto alugado", "Grupos", dentre tantos outros[14]. Se considerado, minimamente, que uma "Área Geográfica" é em um "plano horizontal de uma parte da superfície terrestre", algumas declarações perdem sentido, por exemplo: "Igreja é_um Área Geográfica", "Porta tipo_de Área Geográfica"; inferências como "a União Europeia é um espaço físico tanto quanto o território de seus países associados" ou "Maria reside em uma Caixa de Correio" são logicamente incorretas e mesmo falsas no mundo real.

Já na Classificação Internacional de Doenças (CID), observa-se, dentre diversos outros casos, a categoria "A06 Amebíase" com 10 subcategorias subjacentes. Ao analisá-las, dessa forma, existe um único tipo de diagnóstico – *Shiguelose* – declarado inúmeras vezes indicando suas lesões e locais de manifestação, não caracterizando subtipos e não definindo a doença. A CID parece uma terminologia que não encontra justificativa em princípios classificatórios e só encontra respaldo dentro da própria instituição que a cria e mantém.

Essas abordagens, portanto, resultam na criação de categorias ambíguas, na possibilidade do mesmo tipo pertencer a diferentes categorias e no estabelecimento de relações equivocadas entre uma categoria e seus tipos. São às vezes organizadas em lógica informal contendo erros fundamentais de classificação que impactam tanto na recuperação de documentos quanto em prováveis inferências típicas das ontologias.

2.4 TEORIAS SUBJACENTES

O assunto apresentado até aqui na presente seção como *background* da pesquisa envolve questões filosóficas sobre categorias, assunto seminal da CI, cujas discussões já duram mais de 2000 anos. A abordagem é interdisciplinar e

[13] Disponível em:<https://ncit.nci.nih.gov/ncitbrowser/>
[14] Disponível em: <https://pt.wikipedia.org/wiki/%C3%81rea>

atemporal, de forma que não faz uso exatamente das teorias usualmente adotadas em CI datadas da segunda metade do século XX. A visão professada na teoria das ontologias é, ainda que bastante similar, distinta, pois existem diferentes correntes filosóficas. Isso não implica necessariamente que uma teoria é melhor que a outra, apenas que existem momentos distintos para aplicá-las.

Muita discussão e debate surgem quando se adota uma teoria que não representa exatamente o *mainstream* do campo de pesquisa. Ainda assim, a adoção dessa ou daquela teoria pode variar, mas três modos de abordagem devem estar presentes em qualquer fundamentação: epistemológico, ontológico, metodológico (GUBA, 1990). Como qualquer teoria estabelece os procedimentos adotados na pesquisa científica, o presente artigo destaca a abordagem ontológica, a qual requer um grau de precisão presente apenas em dispositivos lógicos. Essas são necessidades reais ao se construir ontologias, artefatos voltados para inferências automáticas e redução da ambiguidade.

Por tudo isso, acredita-se que está além das possibilidades de um único artigo aprofundar-se na discussão de teorias subjacentes a diferentes correntes. Esse é um trabalho para uma série de artigos ou livro. Por isso, as seções teóricas até aqui apenas sintetizaram os pontos da discussão, sinalizando balizas conceituais adotadas, sem entrar na comparação de teorias. Além disso, a série de artigos mencionada tem sido disseminada no âmbito de publicações de CI já há mais de 10 anos. Para uma publicação recente que vai diretamente ao ponto, confrontando teorias, sugere-se ao leitor consultar ALMEIDA e TEIXEIRA (2020). Para mais sobre o fundamentos da teoria da Ontologia Aplicada, como já mencionado um tema recorrente em publicações da CI, sugere-se a série ALMEIDA e BARBOSA (2009), ALMEIDA, SOUZA e FONSECA (2011); ALMEIDA (2013); ALMEIDA e FARINELLI (2017); ALMEIDA, RIBEIRO e BARCELOS (2020); PINTO e ALMEIDA (2020); MACHADO, ALMEIDA e SOUZA (2020). Os interessados podem ainda ver uma extensão mais detalhada dessa fundamentação em ALMEIDA (2020).

2.5 A ETAPA DE CONCEITUALIZAÇÃO

Diante de toda a fundamentação apresentada para a criação de uma estrutura de categorias, pode-se afirmar, por fim, que a etapa de Conceitualização diz respeito à criação de um modelo conceitual em categorias. Ao criar uma estrutura taxonômica dos termos, agrupar aqueles com algum tipo de similaridade por meio de um critério estabelecido e relacioná-los, a representação do conhecimento emerge. Trata-se, portanto, de uma atividade predominantemente intelectual, onde os processos de abstração e de organização de termos são essenciais na criação de ontologias, sob a unificação de perspectivas para a abordagem ontológica (TEIXEIRA, 2009).

Para o escopo deste artigo, apenas a etapa de Conceitualização será detalhada, sendo possível no entanto o entendimento da metodologia na íntegra a partir da pesquisa original e publicações derivadas (MENDONÇA, 2015; MENDONÇA; ALMEIDA, 2016; MENDONÇA; SOARES, 2017). Assim, a etapa de Conceitualização será descrita por meio de duas metodologias, a *Methontology* e a *OntoForInfoScience*. A seleção se justifica pela maturidade e ampla utilização da primeira no desenvolvimento de ontologias em diferentes áreas do conhecimento e, a segunda, pelo fato de ter sido desenvolvida especificamente para os profissionais da Ciência da Informação (embora ainda careça de ajustes e melhorias em função desses profissionais serem oriundos de diferentes formações).

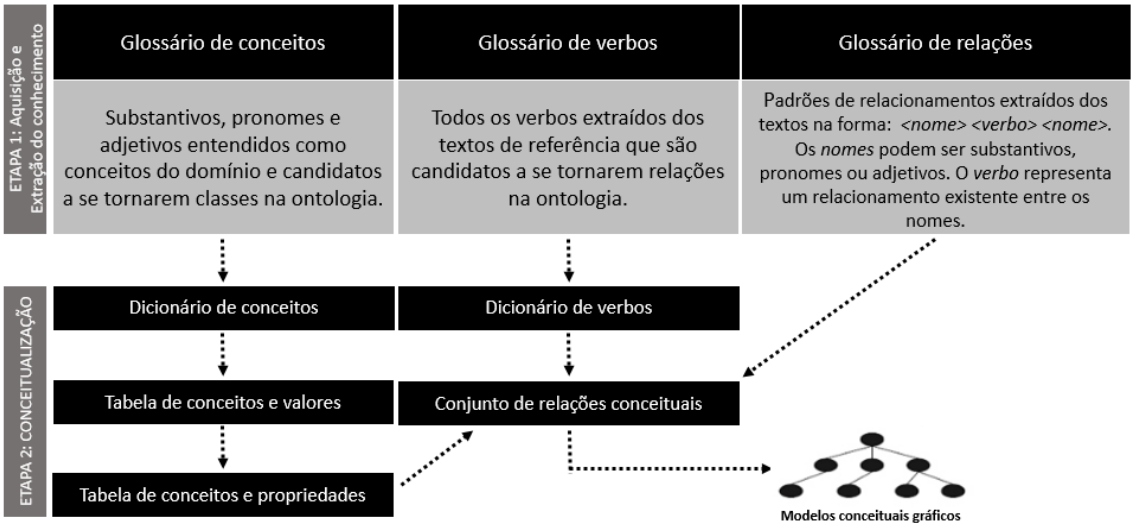
A metodologia *Methontology* foi desenvolvida na construção de uma ontologia no domínio da bioquímica (FERNANDEZ-LOPEZ et al 1997; GÓMEZ-PEREZ et al., 1996) que passou a ser amplamente utilizada e bem aceita. Segundo Almeida (2006), a *Methontology* é a metodologia mais madura para construção de ontologias. O seu ciclo de desenvolvimento contempla seis atividades, a saber: i) aquisição de conhecimento; ii) construção do documento de especificação de requisitos; iii) conceitualização da ontologia; iv) implementação da ontologia; v) avaliação durante cada fase do ciclo de desenvolvimento; vi) documentação após cada fase. A etapa de Conceitualização, foco deste estudo, é realizada em 10

tarefas, quais sejam: i) criar um glossário de termos; ii) construir uma taxonomia de conceitos; iii) construir um diagrama de relações binárias ad hoc; iv) construir um dicionário de conceitos; v) descrever relações binárias ad hoc; vi) descrever atributos das instâncias; vii) descrever atributos de classes; viii) descrever constantes; ix) descrever axiomas e, x) descrever regras e instâncias.

A metodologia *OntoForInfoScience* (MENDONÇA, 2015) foi desenvolvida utilizando etapas de outros métodos e buscando suprir suas limitações, além do foco na sua aplicação no contexto da Ciência da Informação - motivo pelo qual é mais detalhadamente apresentada. A metodologia compreende 9 etapas, a saber: i) avaliação da necessidade da ontologia; ii) especificação da ontologia; iii) aquisição e extração do conhecimento; iv) conceitualização; v) fundamentação ontológica; vi) formalização da ontologia; vii) avaliação da ontologia; viii) documentação da ontologia e; ix) disponibilização da ontologia. A etapa de Conceitualização compreende atividades essenciais à definição dos termos que deverão ser utilizados para representar as classes na ontologia, devendo-se manter o cuidado para que tal representação seja o mais aderente possível à realidade do domínio em estudo.

Para tanto, os insumos obtidos na etapa de AC são submetidos a atividades de identificação, análise e negociação entre profissionais da CI e especialistas do domínio, buscando-se a descoberta precisa dos elementos essenciais à construção de um modelo conceitual capaz de atender às especificações da ontologia. Desta forma, a qualidade do resultado obtido nesta fase representa um fator de alto impacto sobre o sucesso da ontologia, motivo pelo qual recomenda-se a adoção de interações colaborativas entre os atores como fomento à elaboração consensual do referido modelo. Os insumos necessários para a realização da etapa de conceitualização são ilustrados na Figura 1.

Figura 1 - Insumos para a Conceitualização e produtos derivados



Fonte: Adaptado de Mendonça (2015).

Conforme mostra a Figura 1, seis atividades caracterizam a etapa de conceitualização, a saber: i) construção do dicionário de conceitos; ii) construção da tabela de conceitos e valores; iii) construção da tabela de conceitos e propriedades; iv) construção do dicionário de verbos; v) construção do conjunto de relações conceituais e, vi) elaboração de modelos conceituais gráficos. As atividades serão descritas e exemplificadas na sequência.

Atividade 1 – Construção do Dicionário de Conceitos – é obtida a partir do acréscimo de uma definição para cada conceito listado no Glossário de Conceitos. As definições são extraídas de fontes de informação do domínio em estudo e relacionadas a este. Durante o processo de atribuir definições aos conceitos é comum a identificação de sinônimos, tendo em vista o tratamento semântico que é realizado nesta etapa. Os sinônimos identificados deverão ser removidos do Dicionário de Conceitos e anotados ao lado do termo preferencial.

Atividade 2 – Construção da Tabela de Conceitos e Valores – objetiva estabelecer os valores possíveis para cada conceito existente no Dicionário de Conceitos. Compreende-se como valores as ocorrências de um determinado conceito no domínio ou casos reais destes. Os valores possíveis equivalem aos particulares de um determinado universal, candidatos a se tornarem instâncias das classes da ontologia, ou mesmo as próprias classes, tendo em vista o grau de especialização que se deseja alcançar com a ontologia para o domínio em

questão. Vale destacar que esta decisão deverá ser tomada na etapa de formalização da ontologia quando serão aplicados os princípios ontológicos norteadores. Caso a opção seja pela instanciação, estas devem atender às restrições da classe (propriedades de dados).

A título de exemplificação, consideremos a construção de uma ontologia dos Tipos de Instalações presentes no setor elétrico. Os termos a serem considerados foram extraídos do documento de referência do setor, o Manual de Controle Patrimonial do Setor Elétrico (MCPSE). São eles: Geração, Usina e Subestações. Como valores possíveis de cada conceito teríamos: Geração = {Solar, Eólica}, Usina = {Hidroelétrica, Térmica a Vapor - Carvão} e Subestações = {Subestações em tensão menor ou igual a 13,8kV, Subestações em tensão maior que 138kV e menor ou igual a 230kV}. A Tabela de Conceitos e Valores é, portanto, uma evolução do Dicionário de Conceitos a partir do acréscimo dos valores possíveis do domínio. Na Tabela 1 ilustra-se um fragmento da Tabela de Conceitos e Valores do domínio de energia.

Tabela 1 - Fragmento da Tabela de Conceitos e Valores

ID	Conceito	Sinônimo	Definição	Valores
1	Usina	Eólica, Hidroelétrica, Solar e Térmica	É um recurso de manufatura que suporta a função de [transformar energia potencial em energia cinética] onde o processo de [...] realiza a função de [transformar energia potencial em energia cinética] e é um [...].	Hidroelétrica, Térmica a Combustão – Óleo e Gás, Térmica a Vapor – Carvão, Térmica a Vapor – Gás, Térmica a Vapor – Diesel/Querosene/Óleo, Térmica a Vapor – Biomassa, Térmica a Vapor – Nuclear, Solar, Eólica.

Fonte: Elaborado pelos autores.

Atividade 3 – Construção da Tabela de Conceitos e Propriedades – a Tabela de Conceitos e Valores recebe um acréscimo de propriedades específicas para cada conceito, transformando-se na Tabela de Conceitos e Propriedades. Propriedades são características que irão possibilitar, na etapa de formalização, o alcance da distinção precisa entre os conceitos. Enquanto o aspecto ontológico não é analisado, permite-se a definição das propriedades de maneira descritiva em formato textual. Pela mesma razão, propriedades como tipo de valores dos conceitos (numérico, *string* e *booleano*), domínio e faixa e a aridade deverão ser definidas apenas na etapa de formalização a partir de um editor de ontologias.

A identificação de propriedades específicas para cada conceito remete à identificação da "*differentia*" que era realizada por Aristóteles, no sentido de se conhecer a essência de cada ser em estudo, para que a taxonomia de categorias

ontológicas voltadas à classificação destes seres pudesse ser construída com maior aderência possível à realidade. Especificamente, o que se busca, é a representação fiel de um domínio para que as classes sejam futuramente identificadas de forma precisa, possibilitando responder às questões de competência que justificaram a construção da ontologia. Recomenda-se que a Tabela de Conceitos e Propriedades seja validada pelos especialistas de domínio para que sua utilização seja profícua na fase de desenvolvimento da ontologia.

Atividade 4 – Construção do Dicionário de Verbos – consiste em construir o respectivo dicionário a partir do acréscimo de uma definição (em linguagem natural) do significado de cada verbo previamente organizado no Glossário de Verbos. Assim como ocorre com os conceitos, ao se atribuir uma definição para cada verbo torna-se possível a identificação de verbos sinônimos (rótulos diferentes e sentido idêntico). Neste caso, recomenda-se a representação de uma única entrada no Dicionário de Verbos, com o nome preferencial, seguido da anotação dos seus sinônimos. Vale destacar que, o Glossário de Verbos "é o parâmetro de entrada para definição de relações conceituais no domínio" (MENDONÇA, 2015, p.204).

Atividade 5 – Construção do Conjunto de Relações Conceituais – o conjunto de relações começa a ser obtido durante a construção do Glossário de Verbos, sendo refinado a partir da construção do Dicionário de Verbos. No entanto, apenas a partir de sua correlação com o Dicionário de Conceitos do domínio, acompanhado de uma análise individual para validação das relações é que este instrumento se torna válido para uso na ontologia. Segundo Mendonça (2015) a descoberta de novas relações pode se dar a partir de consultas aos insumos da etapa de conceitualização e às fontes de informação adotadas como referência.

Algumas premissas básicas são enumeradas pelo autor para a construção do conjunto de relações conceituais, a saber: i) um item na forma <nome1><verbo><nome2> do Glossário de relações apenas poderá compor o conjunto de relações se: i.i) <nome1> e <nome2> devem existir na Tabela de Conceitos e Propriedades, seja como rótulos preferenciais ou anotados como sinônimos, situação na qual deverão ser substituídos pelo rótulo preferencial na relação que será estabelecida; i.ii) <verbo> deve existir no Dicionário de Verbos, seja como rótulo preferencial ou anotado como sinônimo, situação na qual deverá

ser substituído pelo verbo preferencial na relação que está sendo estabelecida.

Atividade 6 – Elaboração dos modelos conceituais gráficos – é uma atividade que demanda a presença de especialistas para elaboração colaborativa do modelo ou validação. Além disso, uma ferramenta tecnológica para modelagem deve ser adotada, tendo em vista que a descrição textual de tais relações não são capazes de representar os relacionamentos no contexto (MENDONÇA, 2015). São consideradas estruturas gráficas de representação válidas para esta etapa os grafos conceituais, as taxonomias, as hierarquias dentre outros. Caso a construção desta modelagem seja prevista para realização colaborativa, os especialistas de domínio devem ter acesso aos conjuntos terminológicos previamente desenvolvidos, além de receberem treinamento para o uso dos recursos tecnológicos.

Não consiste um objetivo deste trabalho avaliar em que medida qual processo da Conceitualização de qual metodologia é mais bem aceita. Considera-se, entretanto, que os profissionais da CI que não advém da área de tecnologia, encontram dificuldades em ambas metodologias apresentadas, embora a segunda foque nesse público mais abrangente. Também não consiste um objetivo julgar o uso das palavras “termo” e “conceito” nas metodologias, motivo pelo qual devem ser entendidas enquanto referência às entidades do mundo real que elas representam.

3 CONTEXTO

A energia elétrica é um bem de consumo de grande importância para a sociedade e constitui um insumo necessário para as novas tecnologias, desde a sua descoberta e efetivo desenvolvimento no século XIX nos Estados Unidos e Europa. O setor elétrico brasileiro é um dos maiores do mundo e distinto de outros países, uma vez que é basicamente fundamentado na geração hidroelétrica.

3.1 Breve Histórico do Setor Elétrico Brasileiro

A energia elétrica começou a ser explorada no Brasil no final do século XIX, quando as empresas de capital nacional e estrangeiro iniciaram seus empreendimentos atendendo a um mercado consumidor restrito, suprido inicialmente pelo setor público, como o de transporte e atividades fabris

(BRANCO, 1996; GOMES e VIEIRA, 2009). A revolução de 1930 e a chegada de Getúlio Vargas em 1934 à presidência contribuíram para o maior controle do setor pela União. Criado em 1939, o Conselho Nacional de Águas e Energia Elétrica (CNAEE), subordinado diretamente à Presidência da República, ficou responsável por todos os assuntos importantes do setor elétrico. No período pós guerra, foi criada a Eletrobrás, quando um novo período se iniciou no setor elétrico brasileiro, tendo apoio tanto político como econômico. Neste período foram também criadas as Centrais Elétricas de Minas Gerais (CEMIG), o Ministério das Minas e Energias (MME) e o Departamento Nacional de Águas e Energia Elétrica (DNAEE). Assim, o setor elétrico foi contemplado com diversas ações que fizeram crescer o volume de recursos financeiros a ele direcionado e investido.

Nos anos 80 o governo brasileiro foi obrigado a reduzir fortemente os investimentos no setor elétrico, o que levou a uma defasagem nos recursos. A crise chegou ao seu ápice nos anos 90 por conta do endividamento das empresas federais e, neste período, foi apresentado como solução o processo de privatização. A partir de 1996 deu-se início ao processo de privatização e a criação do Operador Nacional do Sistema Elétrico (ONS), com a finalidade de realizar operações interligadas dos sistemas elétricos nacionais.

O Estado, a partir de 2004, implantou um novo modelo de investimento onde retomou a responsabilidade pelo planejamento e desenvolvimento do setor. Neste momento ocorre a desverticalização do setor e a criação da Agência Nacional de Energia Elétrica (ANEEL). Entre as principais atribuições da ANEEL se destacam: i) a iniciativa de criação e monitoramento dos requisitos técnicos que deverão ser observados com o objetivo de garantir a qualidade dos serviços oferecidos aos usuários; ii) solicitar, quando achar necessário, a abertura de processos licitatórios paras as concessões de geração, transmissão e distribuição de energia elétrica; iii) garantir que todas as concessionárias tenham as mesmas condições de participação no mercado, ou seja, não ocorra o privilégio de uma em detrimento de outra; iv) estabelecer os critérios adotados para a gestão de custo de transmissão; v) realização de maneira periódica da avaliação e determinação das tarifas que deverão ser praticadas no varejo (FERREIRA, 2000).

Várias iniciativas em direção à padronização técnica e administrativa do setor

têm sido implementadas pela ANEEL, como o Manual de Contabilidade do Setor Elétrico (MCSE), Manual de Controle Patrimonial do Setor Elétrico (MCPSE) e os Procedimentos de Distribuição (PRODIST). O conjunto de padronização tem o objetivo de melhorar a fiscalização dos serviços prestados pelas concessionárias e garantir a auditabilidade dos processos regulados pelo órgão, como os contábeis. Entretanto, tais iniciativas, além de apresentarem divergências terminológicas entre si, também apresentam características impeditivas ao estabelecimento da interoperabilidade semântica entre os sistemas de uma empresa de energia.

3.2 Demanda Técnica-Tecnológica do Setor Elétrico Brasileiro

Mesmo diante as diferentes iniciativas que buscam, de alguma forma, prover algum tipo de padrão que vise normatizar auditorias e fiscalizações do órgão regulador, o que se observa são padrões díspares entre si. Dentre as iniciativas técnicas e administrativas existentes, são encontradas para uma mesma entidade diferentes terminologias e até mesmo propriedades. Nesse sentido, o Centro de Gestão e Estudos Estratégicos (CGEE), em 2017, realizou um estudo por meio de um projeto de P&D cujo foco esteve na necessidade de reestruturações essenciais para a distribuição de energia elétrica.

O referido estudo (CGEE, 2017) traz à tona objetivos específicos a curto prazo (2017-2020), médio prazo (2020-2030) e longo prazo (2030-2050) para o negócio de distribuição. Onze macrotemáticas foram abordadas, sendo elas: i) automação da rede; ii) medição avançada; iii) geração distribuída e micro-redes; iv) mobilidade elétrica; v) qualidade da energia elétrica; vi) infraestrutura de proteção, automação e controle; vii) operação e manutenção; viii) tecnologia da informação e comunicação; ix) segurança cibernética; x) compartilhamento de serviços no contexto das cidades inteligentes; e xi) subestação e equipamentos e redes de distribuição aéreas e subterrâneas. Para todas elas, apontam-se a visão de futuro, a evolução da maturidade tecnológica e as rotas priorizadas.

A questão técnica da interoperabilidade é identificada em 5 delas, entre requisitos da visão de futuro e da evolução da maturidade tecnológica. São elas: i) medição avançada (interoperabilidade entre os equipamentos); ii) Automação da rede (arquiteturas para a integração de dados); iii)

Compartilhamento de serviços no contexto das cidades inteligentes (integração de tecnologias, sistemas e modelos de negócio compartilhados voltados a cidades inteligentes; integração de tecnologias e sistemas; garantia da interoperabilidade; interoperabilidade entre elementos do sistema); iv) Tecnologia da informação e comunicação (compartilhamento de informações; desenvolvimento de tecnologias de inteligência artificial; internet das coisas (IoT)); e v) Operação e manutenção (internet das coisas (IoT); padronização de sistemas de comunicação; redes inteligentes).

Considerando toda a descrição do cenário tecnológico do setor elétrico em CGEE (2017), é possível afirmar que a heterogeneidade em uma infraestrutura tecnológica sob diferentes padrões terminológicos e de interoperabilidade fomenta inconsistências e assim, a mitigação dessas barreiras é um pré-requisito para tais evoluções pretendidas para o setor. A inviabilidade de troca de informações entre os sistemas afeta a qualidade das informações tanto em demandas regulatórias quanto para gestão e gerenciamento, assim como para a tomada de decisões.

3.3 CONTEXTO DA PESQUISA

A presente pesquisa se desenvolve em um projeto de P&D na Companhia Energética de Minas Gerais (CEMIG), financiado pela ANEEL. O proceder metodológico do Projeto é baseado em preceitos da Ciência da Informação e apresenta dados empíricos da prática de conceitualização, típica em construção de ontologias.

O desenvolvimento de um artefato ontológico foi proposto como solução à necessidade de se prover interoperabilidade semântica entre os sistemas de informações relacionados aos ativos da União sob a tutela da CEMIG. As informações corretas, bem geridas e repassadas à ANEEL em formato e periodicidade regulados, impactam positivamente na remuneração obtida pela Empresa em função da manutenção dos ativos e dos investimentos realizados sobre os mesmos. Portanto, o requisito “qualidade da informação” é fundamental nesse processo. Nesse ínterim, o P&D propõe como solução o desenvolvimento de uma ontologia baseada nas entidades que constituem os ativos da CEMIG sob o ponto de vista contábil. Nesse sentido, o escopo do artefato ontológico é delimitado pelas entidades contidas no

MCPSE. Almeja-se enquanto resultado do projeto o fomento às condições de se prover interoperabilidade semântica entre os sistemas da CEMIG e, em última instância, para aplicação em outras empresas do setor elétrico e na própria ANEEL.

4 METODOLOGIA E RESULTADOS

Considerando que a fase de conceitualização no processo de desenvolvimento de uma ontologia se refere à geração dos modelos conceituais por meio de um vocabulário comum compartilhado de uma domínio, e que, para a sua realização, os passos anteriores são essenciais, foram estabelecidos e realizados previamente a este momento: i) aquisição do conhecimento: pesquisa às fontes de informação - cujo objetivo consistiu em inventariar e avaliar o conhecimento registrado relacionado ao Projeto; ii) entendimento da área do domínio e delimitação dos termos - cujo objetivo consistiu em construir um entendimento coletivo acerca dos documentos de padronização terminológica e processual do setor; e iii) aquisição do conhecimento: entrevistas com especialistas - cujo objetivo consistiu em obter dos especialistas do setor elétrico os seus conhecimentos sobre os termos a serem definidos.

A etapa do Projeto aqui descrita, referente à etapa de Conceitualização na construção de ontologias, se classifica como: i) exploratória, pois envolve o registro de conhecimentos de entidades do setor elétrico; e ii) relato de caso, uma vez que descreve, empiricamente, os passos adotados para tal. Esta seção descreve os passos adotados na pesquisa aplicada.

4.1 Procedimentos Preliminares Realizados

Considerando todo conjunto de procedimentos que envolvem a construção de um artefato ontológico e os impactos dos procedimentos anteriores ao desenvolvimento da Conceitualização, considera-se relevante apresentar uma breve visão dos resultados obtidos até então.

4.2 Levantamento e Análise de Material de Referência

Inicialmente foram realizadas pesquisas com o objetivo de identificar

normas, procedimentos e padronizações da ANEEL que influenciavam os processos dentro da organização. Na busca pelo entendimento do setor elétrico foram realizadas reuniões preliminares com os envolvidos nas rotinas operacionais da concessionária assim como pesquisas cujos objetivos consistiam em identificar terminologias e ontologias do contexto, textos científicos relacionados, documentos normativos do setor e documentos regulatórios vigentes. Todo o processo de análise realizado culminou na definição do escopo do artefato ontológico: o MCPSE. Uma vez definido como escopo do Projeto, considerando a sua amplitude terminológica e o prazo de desenvolvimento do Projeto, delimitou-se um recorte da Distribuição que contempla as entidades mais importantes na remuneração para a CEMIG.

4.3 Identificação de Especialistas de Domínio

Foram identificados e convidados a participar, três especialistas de domínio (um da CEMIG e dois da executora). Entendeu-se que, em função da disponibilidade dos mesmos e da localização geográfica distribuída, a objetividade para a extração do conhecimento era um requisito essencial.

4.4 Aplicação de Técnicas Colaborativas

Foram realizadas entrevistas presenciais e por videoconferência aos três diferentes especialistas por meio de três pesquisadores em CI. Somaram-se 13 entrevistas do tipo estruturadas, com aproximadamente 2:30h de duração cada. Foram apresentados aos especialistas os termos das entidades extraídas do recorte estabelecido no MCPSE para que ambos pudessem se dedicar de acordo com suas experiências e vivências. O acompanhamento pessoal foi importante para que ocorresse o mínimo possível de ambiguidade e para que, na sequência, os termos pudessem ser melhor definidos.

4.5 Pré Conceitualização

Uma análise preliminar do resultado obtido indicou a não uniformidade

terminológica em relação aos processos em que as funções identificadas atuavam. Este *gap* impactou diretamente na construção das definições dos termos pelos especialistas da CI. Nesse sentido, procedeu-se à busca pelo site da ANEEL, da CEMIG e de empresas correlatas. Na indisponibilidade de respostas, foi necessário entrevistar especialistas para se obter um rol de processos padronizados. Tal busca não alcançou os resultados pretendidos, uma vez que essa padronização de processos de negócios não existe para a geração, para a transmissão e para a distribuição.

Nesse sentido, partiu-se em busca de insumos que pudessem subsidiar esse levantamento. Justificado por seus conteúdos e objetivos, optou-se pelo uso dos 11 Procedimentos de Distribuição (PRODIST)[15] da ANEEL e das 26 Normas de Distribuição (ND)[16] da CEMIG, uma vez que ambos padronizam as atividades técnicas da Distribuição. Após estudo e leitura dos documentos, foi gerada uma tabela que subsidiou o desenvolvimento das definições. Foram identificados 10 macroprocessos e 51 processos relacionados.

Sete pesquisadores foram treinados para gerar as definições em padrão ontológico. Cada pesquisador recebeu aproximadamente 50 termos para serem definidos, além dos insumos para desenvolverem seus trabalhos, que são o conhecimento registrado dos especialistas de domínio e a tabela de processos desenvolvida. Todas as quase 400 definições geradas foram revisadas, sendo adequadas ou modificadas, segundo o padrão ontológico estabelecido. Das definições foram identificadas 296 funções que, após extração daquelas sinônimas foram também padronizadas. Das 221 funções resultantes extraíram-se 93 verbos. Nesse sentido, obteve-se nesta etapa tanto o glossário de termos definidos quanto o glossário de verbos.

4.6 Procedimentos da Conceitualização Adotados

Considerando que o projeto se baseia em princípios da CI para a sua realização e é conduzido por profissionais multidisciplinares. Os passos a seguir

[15] Disponível em: https://www.aneel.gov.br/prodist

[16] Disponível em: http://www.cemig.com.br/pt-br/atendimento/Clientes/Paginas/norma_tecnica.aspx

foram propostos, estabelecidos e realizados, motivo pelo qual o resultado da sua aplicação é apresentado.

4.6.1 Passo 1: elaborar o dicionário de termos

A primeira etapa adotada pelo Projeto consistiu em revisar e organizar alfabeticamente os termos que representam as entidades definidas em linguagem natural na pré-conceitualização. A figura 2 apresenta fragmento do resultado obtido nesse passo:

Figura 2 - Fragmento do dicionário de termos

TERMO	DEFINIÇÃO
Banco De Capacitores	[banco de capacitores] é um agregado de objeto [capacitor] no processo industrial de [Mensurar qualidade da energia elétrica].
Cabo Pára-raios	[cabo pára-raios] é um equipamento industrial que suporta a função de [proteger linhas de transmissão contra descargas elétricas atmosféricas] onde os processos de [Fornecimento trifásico de energia elétrica em média tensão AND T] realiza a função de [proteger linhas de transmissão contra descargas elétricas atmosféricas] e é um processo industrial.
Caixa De Automação Submersível	[caixa de automação submersível] é um [unidade terminal remota] que realiza a função de [AND é submersível].

Fonte: elaborado pelos autores

4.6.2 Passo 2: elaborar o dicionário de verbos

A segunda etapa adotada pelo Projeto consistiu em organizar os verbos extraídos das funções das entidades identificadas nas definições dos termos. A figura 3 apresenta fragmento do resultado obtido nesse passo:

Figura 3: Fragmento do dicionário de verbos

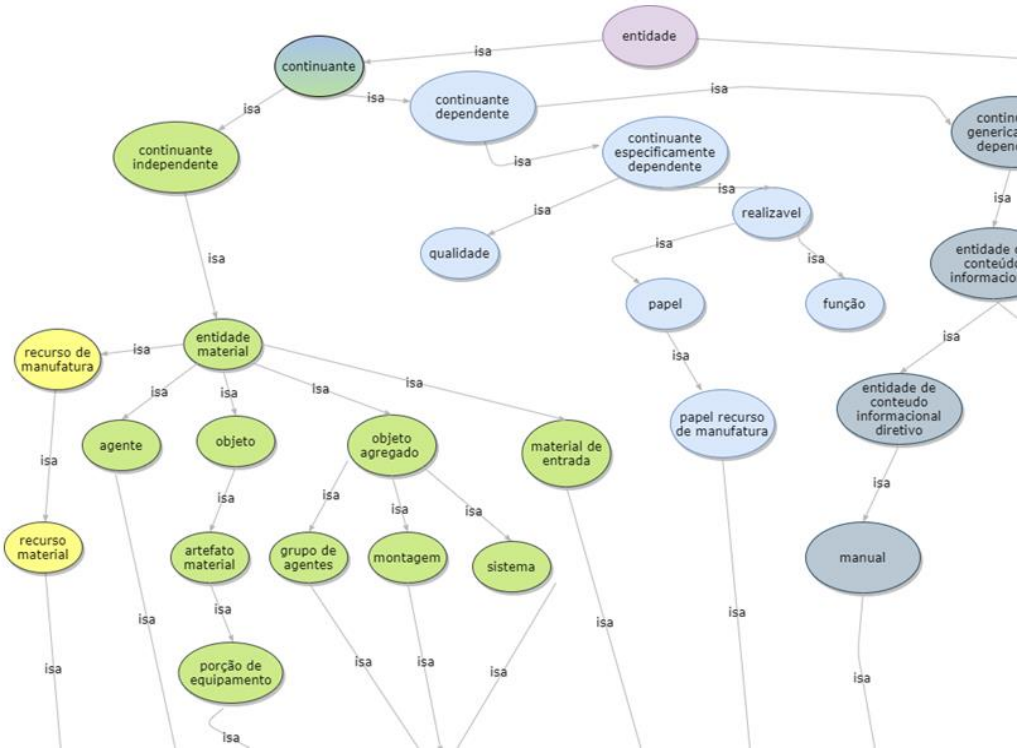
VERBO	DEFINIÇÃO	FUNÇÃO RELACIONADA
abrigar	acolher e proteger de intempérie ou agente agressivo.	abrigar cabo guarda
		abrigar banco de capacitores
		abrigar chave seccionadora
		abrigar disjuntor
		abrigar dispositivo
		abrigar lâmpada
		abrigar sistemas eletrônicos
abrir	efetuar a desobstrução.	abrir contatos
absorver	incorporar ou embeber-se da substância.	absorver a luz solar
		absorver sobretensão atmosférica

Fonte: elaborado pelos autores

4.6.3 Passo 3: elaborar o modelo conceitual

A terceira etapa adotada pelo Projeto consistiu em desenvolver o modelo conceitual do domínio do conhecimento. Foi previsto em etapas anteriores da metodologia que as ontologias a serem reutilizadas seriam a *Basic Formal Ontology* (BFO - ontologia de fundamentação neutra), a *Industrial Ontologies Foundry* (IOF - ontologia de fundamentação industrial) e a *Information Artifact Ontology* (IAO - ontologia de nível médio, neutra, para a representação de tipos de entidades de conteúdo de informação). A figura 4 apresenta o resultado obtido nesse passo:

Figura 4: Fragmento do modelo conceitual da ontologia



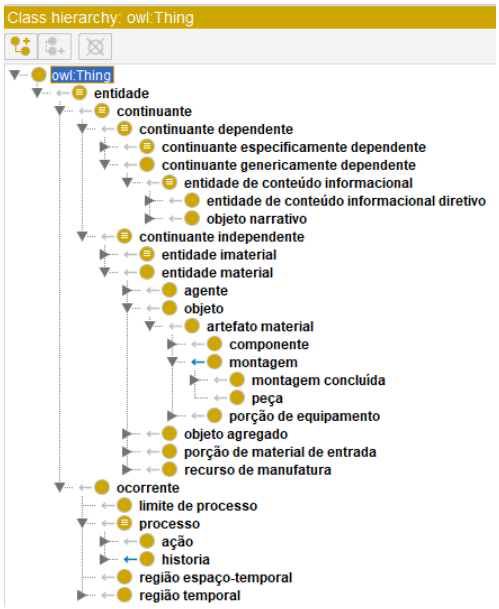
Fonte: elaborado pelos autores

4.6.4 Passo 4: Organizar a taxonomia

A quarta etapa adotada pelo Projeto consistiu no desenvolvimento da taxonomia de termos que representam as entidades e processos do setor elétrico. As ontologias a serem reutilizadas devem ser importadas para um editor de ontologias. As classes de entrada de termos representativos das entidades e processos do setor elétrico devem ser definidas e então populadas. A figura 5

apresenta o resultado obtido da integração das ontologias nesse passo:

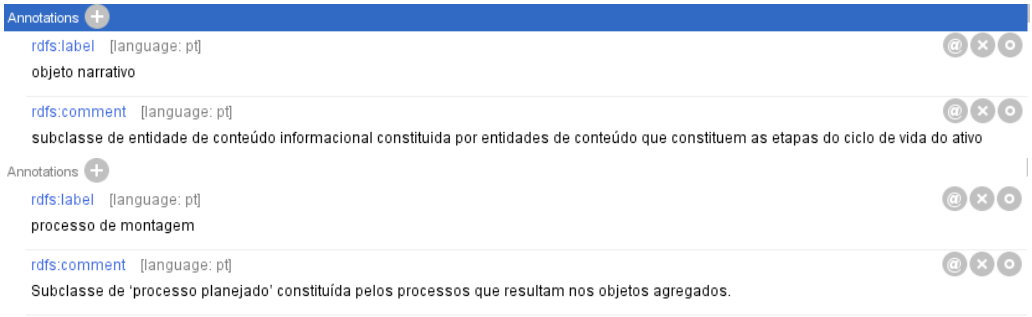
Figura 5 - Fragmento da taxonomia da ontologia



Fonte: elaborado pelos autores

A figura 6 apresenta fragmento das definições das classes de entrada de entidades da ontologia:

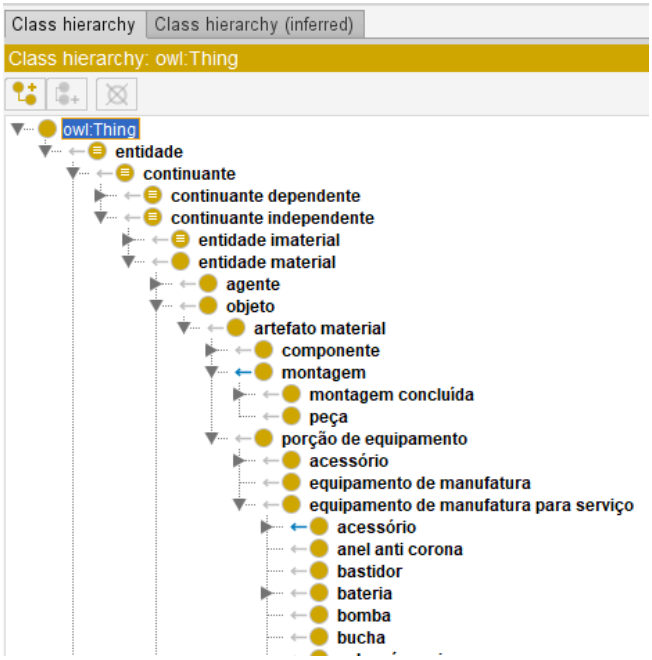
Figura 6 - Fragmentos de definições de classes de entradas de entidades da ontologia



Fonte: elaborado pelos autores

A figura 7 apresenta fragmento das classes de entrada de entidades da ontologia populadas pelos termos definidos:

Figura 7 - Fragmento de classes de entradas populadas



Fonte: elaborado pelos autores

4.6.5 Passo 5: definição das relações semânticas

A terceira etapa adotada pelo Projeto consistiu em definir as relações semânticas que serão atribuídas entre as classes e as categorias de termos que representam as entidades do mundo real. Em função do tempo de execução do Projeto em relação ao universo do escopo, optou-se por manter a granularidade das categorias, mantendo as relações herdadas pela reutilização das ontologias.

5 DISCUSSÃO

A etapa de conceitualização, foco deste trabalho, é aqui apresentada como um resultado acumulado das fases anteriores, motivo pelo qual as mesmas foram brevemente descritas. No entanto, todos os resultados consistiram em insumos para que a conceitualização se desenvolvesse de forma bem fundamentada. Foram eles: i) as fontes de informação selecionadas; ii) a tabela de registros, populada pelas entrevistas com especialistas, iii) a tabela de processos; iv) o glossário de termos e iv) o glossário de verbos.

No passo 4 da metodologia reside o cerne da etapa de Conceitualização:

a categorização dos termos que representam as entidades e os processos do setor elétrico. Essa atividade é puramente intelectual e desenvolvida por meio de critérios estabelecidos em ontologias de fundamentação e de nível médio.

Vale a pena destacar que, à medida que os termos vão sendo definidos e revisados, à medida que estes vão constituindo o corpo de um dicionário (termos e verbos) e à medida que entende-se melhor os processos em que as entidades representadas pelos termos atuam, uma visão de categorização dos mesmos surge naturalmente. Evidenciam-se diferenças entre grandes grupos de termos por critérios da mente, inicialmente. Entretanto, ao definir as categorias de entradas de termos, esta divisão inicial é revista e fortemente alterada em função de critérios ontológicos.

Condições para os termos constituírem as classes de categorias de equipamentos, agregado de equipamentos, conjunto de equipamentos e parte de equipamentos, por exemplo, são então baseadas em critérios ontológicos. Da mesma forma, questões terminológicas que requerem atenção se tornam evidentes. Um bom exemplo é "*monitor de computador"* e o "*monitor de variáveis"*, que, por critério da mente ou terminológico poderiam constituir *tipo_de* uma mesma coisa. Entretanto, um *monitor de computador* é um [monitor que realiza a função de exibir as informações do computador] e o *monitor de variáveis* é um [equipamento industrial que suporta a função de monitorar variáveis de pressão e monitorar variáveis de temperatura onde o processo de Mensurar qualidade da energia elétrica e Medir grandezas elétricas do sistema de distribuição realiza a função de monitorar variáveis de pressão e monitorar variáveis de temperatura e é um processo industrial].

Uma questão relevante para os resultados obtidos nesta etapa de Conceitualização reside no fato da utilização inédita da IOF. A IOF é uma ontologia de fundamentação, de cunho industrial, desenvolvida de forma agregada à BFO. Seu objetivo consiste na fundamentação de módulos de ontologias de qualidade para a interoperabilidade semântica em domínios industriais, como o setor elétrico. Da junção das ontologias propostas, 21 categorias se estabeleceram enquanto entrada de entidades, com a geração de definições claras e objetivas sobre os critérios e as condições para que entidades possam estar nelas alocadas. Estas categorias específicas são apresentadas no

Quadro I, acompanhadas de suas respectivas definições.

Quadro 1 - Categorias de entradas da ONTOMCPSE

Entrada	Definição
Ação	Subclasse de 'processo' constituída por atos que culminam nas entidades de conteúdo das etapas do ciclo de vida do ativo.
Acessório	Subclasse de 'porção de equipamento' constituída por equipamentos que acondicionam ou suportam materiais.
Agente	Subclasse de 'entidade material', constituída pelos executores das ações que culminam nas entidades de conteúdo das etapas do ciclo de vida do ativo.
Componente	Subclasse de 'artefato material' constituída por entidades-partes de sistemas.
Equipamento de manufatura para serviço	Subclasse de 'porção de equipamento' constituída por entidades portadoras de função na execução dos processos de serviço.
Especificação da operação	Subclasse de 'especificação do plano' constituída por documentos declarativos das etapas do ciclo de vida do ativo.
Função do negócio	Subclasse de 'função' constituída pelas funções dos documentos das etapas do ciclo de vida do ativo.
Função de processo de serviço	Subclasse de 'função' constituída pelas funções das entidades portadoras de função na execução dos processos de serviço.
História	Subclasse de 'processo' é constituído pelo processo de criar o dossiê do ativo.
Montagem concluída	Subclasse de 'montagem' constituída por entidades que são montadas com um objetivo específico.
Obra civil	Subclasse de 'recurso material' constituída por entidades de infraestrutura resultantes de um processo de construção.
Objeto narrativo	Subclasse de 'entidade de conteúdo informacional' constituída por entidades de conteúdo que das etapas do ciclo de vida do ativo.
Papel de recurso de manufatura	Subclasse de ‘papel’ constituída pelos papéis inerentes aos recursos de manufatura.
Porção de material de entrada	Subclasse de ‘entidade material’ constituída por entidades fracionáveis, não contáveis unitariamente.
Processo de montagem	Subclasse de ‘processo planejado’ constituída pelos processos que resultam nos objetos agregados.
Processo de negócio	Subclasse de ‘processo planejado’ constituída pelos processos relativos a cada etapa do ciclo de vida de um ativo.
Processo de produção do serviço	Subclasse de ‘processo planejado’ constituída pelos processos por meio dos quais os tipos de artefatos materiais realizam suas respectivas funções.

Qualidade	Subclasse de 'continuante especificamente dependente' constituída por entidades que representam as características de tensão relativas aos artefatos materiais.
Recurso de manufatura	Subclasse de 'entidade material' constituída por pelos tipos de instalações e seus tipos de recursos materiais relacionados.
Recurso material	Subclasse de 'recurso de manufatura' constituída por entidades e obras civis localizadas nos tipos de instalações.
Sistema	Subclasse de 'objeto agregado' constituída por entidades tipos de sistema.

Fonte: Elaborado pelos autores.

Ao se representar cada entidade de acordo com a sua essência, tem-se, por exemplo, uma distinção fundamental entre pessoas e os papéis que lhe são socialmente atribuídos dentro de um domínio, os quais geralmente são tidos como uma só entidade em diversos modelos conceituais não fundamentados ontologicamente. Nos manuais de energia elétrica encontramos descritos em linguagem natural os diversos agentes que participam do processo de distribuição de energia, como os consumidores finais, fornecedores, funcionários internos, funcionários de campo, por exemplo. A classificação ontológica nos permite compreender que um comprador é na realidade uma entidade material que tem um papel de comprador, a se realizar durante a participação desta entidade material em um ato de compra. Desta forma as entidades espaciais, temporais e sociais são adequadamente compreendidas em sua essência e classificadas corretamente no âmbito ontológico, reduzindo ambiguidades seminais que tendem a gerar complexidade durante as etapas posteriores de raciocínio automático para obtenção de novos conhecimentos sobre o conhecimento preliminarmente representado.

Uma discussão um pouco além das fronteiras do projeto, leva a questões caras à Ciência da Informação, como descrito em Almeida (2020). Mesmo que "conceitualizar", termo usado nos primórdios da pesquisa em ontologias não tenha definição clara, esse termo se mantém até hoje e a dificuldade em conceitualizar é a dificuldade em classificar. Para entender essas dificuldades, considere-se que alguém julga que uma rosa é vermelha. Para fazer esse julgamento, usa a categoria "vermelho" e a atribui a rosa, sendo que "rosa"

também é uma categoria. A rosa não é apenas vermelha, mas também uma flor, um objeto físico, um produto para venda, etc. O que torna possível que uma entidade participe de várias categorias na organização em formato hierárquico.

Algumas vezes categorias podem parecer um artifício teórico sem utilidade porque uma rosa é uma planta e é algo para vender, mas nem todas as plantas são coisas para vender e nem todas as coisas para vender são plantas. Essas duas formas de classificar – item para venda e planta – se sobrepõem, mas não existe relação hierárquica entre elas. Uma solução sugere que "item para venda" não é categoria natural, porque nada em sua natureza intrínseca é algo para a venda. A rosa pertence a essa categoria porque as pessoas tomaram decisões, ao contrário do fato de que a rosa, por natureza, é uma planta. Assim, preserva-se a ordenação hierárquica ao não incluir categorias que dependem de demarcação humana.

Então, caberia confrontar categorias das criadas pelas pessoas e as naturais. Existe um sentido em que as categorias são construídas por pessoas, pois não se pode pensar em uma categoria sem pensar nessa categoria de fato. Entretanto, fica a questão de se as categorias produzidas por pessoas correspondem a divisões no mundo ou essas categorias impõem divisões ao mundo. Que o mundo natural impõe divisões as pessoas não resta dúvida, mas parece claro que as pessoas também impõem divisões ao mundo. São problemas em constante discussão, que estão presentes mesmo quando se abordam questões práticas.

6 CONSIDERAÇÕES FINAIS

Os resultados obtidos sustentam premissas advogadas por princípios das ontologias. Critérios da mente ou critérios puramente terminológicos não são capazes de refletir a realidade de um domínio do conhecimento tal como os critérios ontológicos são. Para fins de se estabelecer a interoperabilidade semântica entre sistemas, possibilitar o raciocínio automatizado e obter uma recuperação da informação com alta atinência à demanda, é necessário sustentar os modelos conceituais sobre a perspectiva unificada de entendimento e descrição do mundo.

Ainda assim, na prática, o que se observa são modelos desenvolvidos sob a perspectiva do desenvolvedor (seja pessoa ou instituição) ou do banco de dados. Essa necessidade de quebra de paradigmas deve ser destacada em

função da alta propagação de métodos ágeis para desenvolvimento de software, onde menos tempo e dedicação são alocados para essas etapas em projetos. Os métodos ágeis privilegiam a colaboração ativa dos membros de uma equipe de desenvolvimento de sistemas para otimizar entregas incrementais de etapas previamente acordadas em um projeto. Já a construção de ontologias, em contrapartida, requer uma dedicação intelectual em processo diferente do que ocorre no desenvolvimento de sistemas. Tal situação ocorre pelo fato de que o conhecimento especializado nem sempre é obtido por acordos de projetos e raramente se encontra registrado em bases de conhecimento organizacionais. Soma-se a este fato a necessidade de uniformização semântica deste conhecimento, resultante de profundas reflexões filosóficas para se identificar os elementos ontológicos do discurso. Até o momento não há registros de uso destes métodos em projetos como o que aqui se apresenta, e por este motivo o consideramos pouco aderente a este tipo de modelagem.

Para os profissionais da CI fica ainda um *gap* técnico. Evidencia-se a necessidade de uma metodologia menos voltada aos profissionais da Ciência da Computação. A *OntoForInfoScience* busca mitigar essa deficiência mas ainda carece de melhorias nesse sentido. Não se trata, porém, de substituir competências e, sim, descrever o método de forma multidisciplinar. A etapa de Conceitualização consiste na representação taxonômica do conhecimento de um domínio, justificando sua necessidade de melhor descrição.

AGRADECIMENTO

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) por subsidiar a participação da pesquisadora Jeanne Louize Emygdio.

REFERÊNCIAS

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA - ANEEL. MCPSE: **manual de controle patrimonial do setor elétrico**. 2014. Disponível em: http://www.aneel.gov.br/documents/656815/14887121/MANUAL+DE+CONTROLE+PATRIMONIAL+DO+SETOR+EL%C3%89TRICO+-+MCPSE/3308b7e2-649e-4cf3-8fff-3e78ddeba98b. Acesso em: 17 abr. 2019.

ALMEIDA, M.B. **Um modelo baseado em ontologias para a representação da memória organizacional**. 2006. 321p. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

ALMEIDA, M.B. **Ontologia em Ciência da Informação***:* Teoria e Método. Curitiba: CRV, 2020, 377 p. (Coleção Representação do Conhecimento em Ciência da Informação: Volume I), ISBN 978-65-5578-679-8 (volume 1, físico).

ALMEIDA, M. B. **Provocações acadêmicas**: ontologias, tesauros, documentos, conteúdo de documentos, e… unicórnios, 2016. Disponível em: http://mba.eci.ufmg.br/downloads/MBAatSeminario_MHTX_2016.pdf. Acesso em: 20 mar. 2012.

ALMEIDA, M.B.; BARBOSA, R. R. Ontologies in knowledge management support – a case study. **Journal of American Society of Information Science and Technology**. vol. 60, n.10, p. 2032-2047, 2009.

ALMEIDA, M. B.; FARINELLI, F. Ontologies for the representation of electronic medical records: The obstetric and neonatal ontology. **Journal of the Association for Information Science and Technology**, v. 68, n. 11, p. 2529–2542, 2017. doi:10.1002/asi.23900.

ALMEIDA, M.B.; RIBEIRO, E.F.; BARCELOS, R. Toward a Document-Centered Ontological Theory for Information Architecture in Corporations. **Journal of the Association for Information Science and Technology***, v*. 71, n. 11, p. 1308-1326, 2020. DOI: 10.1002/asi.24337, EUA.

ALMEIDA, M. B; SOUZA, R. R.; FONSECA, F. Semantics in the Semantic Web: a critical evaluation. **Knowledge Organization Journal**, vol. 38, n.3, pg. 187-203, 2011.

BRANCO, E. C. A reestruturação e a modernização do setor elétrico brasileiro. **Nova Economia**, [S. l.], v. 6, n. 1, 1996.

DAHLBERG, I. Teoria do Conceito. **Ciência da Informação**, Brasília, DF, v.7, n.2, p.101-107, 1978.

CABRÉ, M. T. **La terminología**: teoría, metodología, aplicaciones. Barcelona: Antártida, 1993.

FERNÁNDEZ, M.; GÓMEZ-PÉREZ, A.; JURISTO, N. METHONTOLOGY: From ontological art towards ontological engineering. **AAAI Technical Report** SS-97-06, 1997.

FERREIRA, C. K. L. **Privatização do setor elétrico no Brasil**. 2000.

GOMES, J. P. P.; VIEIRA, M. M. F. O campo da energia elétrica no Brasil de 1880 a 2002. **Revista de Administração Pública**, [S. l.], v. 43, n. 2, p. 295–321, 2009.

GÓMEZ-PÉREZ, A.; FERNÁNDEZ, M.; VICENTE, A. J. Towards a method to conceptualize domain ontologies. In: **ECAI WORKSHOP ON ONTOLOGICAL ENGINEERING**, 1996, Budapest. Disponível em: http://citeseer.ist.psu.edu/483876.html. Acesso em: 13 de março de 2014.

GONZÁLEZ, J. A. M. Evolução ontológica das linguagens documentárias. Relato de uma experiência de curso organizado conjuntamente para o DT/SIBI-USP e o PPGCI/ECA. InCID: **Revista de Ciência da Informação e Documentação**, 2 (1), 143-164.

GRUNINGER, M. et al. **Ontology summit 2007 - ontology, taxonomy, folksonomy**: Understanding the distinctions. Applied Ontology, 2008.

GUBA, Egon. **The paradigm dialog**. Newbury Park: Sage Publications, 1990.

LEITE, A. L. da S.; DE CASTRO, N. J. Crescimento e estruturação das firmas: a formação dos conglomerados do setor elétrico brasileiro. **REGE - Revista de Gestão***,* [S. l.], v. 21, n. 3, p. 343–359, 2014.

MACHADO, L.M.O.; ALMEIDA, M.B.; SOUZA, R.R. What researchers are currently saying about ontologies: a review on recent Web of Science articles. **Knowledge Organization Journal**, v. 47 n. 3, 2020.

MENDONÇA, F. M. **Ontoforinfoscience**: metodologia para construção de ontologias pelos cientistas da informação: uma aplicação prática no desenvolvimento da ontologia sobre componentes do sangue humano (HEMONTO). Tese (doutorado) - Universidade Federal de Minas Gerais, Escola de Ciência da Informação. Belo Horizonte, 2015.

MENDONÇA, F. M.; ALMEIDA, M. B. Ontoforinfoscience: A Detailed Methodology For Construction of Ontologies and its Application In The Blood Domain. **Brazilian Journal of Information Science**, v. 10, n. 1, 1 mar. 2016.

MENDONÇA, F. M.; SOARES, A. L. Construindo ontologias com a metodologia ontofo- rinfoscience: uma abordagem detalhada das atividades do desenvolvimento ontológico. **Ciência da Informação**, v. 46, n. 1, p. 43–59, 2017.

PESSANHA, C. P.; COELHO, K. C.; ALMEIDA, M. B. Uma investigação filosófico-legal para atos dos documentos. **Anais do 17 Encontro Nacional de Pesquisa em Ciência da Informação***:* descobrimentos da ciência da informação. Desafios da multi, inter e transdisciplinaridade (mit), 2016.

PINTO, J.A.; ALMEIDA, M.B. An Applied Ontology-Oriented Study Case to Distinguish Public and Private Institutions Through Their Documents. **Knowledge Organization**, v. 47, n. 6, 2020, p. 582-592.

PIRES, J. C. L. **Desafios da reestruturação do setor elétrico brasileiro**. Rio de Janeiro: Banco Nacional de Desenvolvimento Econômico e Social, 2000. 45 p. (Textos para discussão; 76).

PRADO, O. Agências reguladoras e transparência: a disponibilização de informações pela Aneel. **Revista de Administração Pública**, v. 40, n. 4, p. 631-646, 2006.

SOUSA, S. de. **CDU**: como entender e utilizar a 2º edição-padrão internacional em língua portuguesa. Brasília: Thesaurus, 2009. 163 p.

SUÁREZ-FIGUEROA, M. C.; AGUADO de CEA, G.; BUIL C.; DELLSCHFT, K.; FERNANDEZ-LOPEZ, M.; GARCIA, A; GOMEZ-PEREZ, A.; HERRERO, G.; MONTIELPONSODA, E.; SABOU, M.; VILLAZON-TERRAZAS, B.; YUFEI, Z. NeOn D5.4.1. **Néon Methodology for Building Contextualized Ontology Networks**. NeOn project. http://www.neon-project.org. February 2008.

TEIXEIRA, L. M. D. **Conceitualização na construção de ontologias**: relações semânticas no âmbito do Blood Project. Dissertação (Mestrado). Escola de Ciência da Informação, Universidade Federal de Minas Gerais, 2009.

WÜSTER, E. **Introducción a la teoría general de la terminologia y a la lexicografia terminológica***.* Barcelona: Universitat Pompeu Fabra, 1998.

W3C. *OWL 2 Web* **Ontology Language Document Overview***:* second edition. Disponível em: <https://www.w3.org/TR/2012/REC-owl2-overview-20121211/>. Acesso em: 25 mar. 2019.

ONTOLOGY-BASED KNOWLEDGE ORGANIZATION: A CASE STUDY ABOUT THE CHALLENGES OF THE CONCEPTUALIZATION IN THE DOMAIN OF POWER SUPPLY

ABSTRACT

Objective The goal of this article is to describe a conceptualization case study in the power supply sector, in which around 1,200 technical engineering terms were extracted to build an ontology. **Methodology**: For this purpose, we here first delimit the theoretical framework, then present the context of the project, in addition to list the set of candidate terms. We also describe some of knowledge acquisition techniques, as well as discuss techniques for defining terms and deriving categories. **Results**: Our research resulted in the record of techniques applied to the candidate terms, in the understanding of the domain peculiarities and, as far as possible in a case study, in obtaining experience and good practices to be applied in similar domains in strategic areas of society. **Conclusions**: it is clear the need for a methodology for building ontologies aimed at Information Science professionals, which includes the conceptualization phase.

Descriptors: Knowledge Organization. Ontologies. Conceptualization. Knowledge Representation. Electric sector.

ORGANIZACIÓN DEL CONOCIMIENTO BASADA EN LA

ONTOLOGÍA: UN ESTUDIO DE CASO SOBRE LOS DESAFÍOS DE LA CONCEPTUALIZACIÓN EN EL ÁMBITO DE ENERGIA ELECTRICA.

RESUMEN
Objetivo: Este artículo tiene como objetivo describir un estudio de caso de conceptualización en el sector eléctrico, en el que se extrajeron alrededor de 1.200 términos de ingeniería técnica para una ontología. **Metodología**: para este fin, se delimita el marco teórico, se presenta el contexto del proyecto, se presentan fragmentos del conjunto de términos candidatos, se describe la adquisición de conocimiento, así como técnicas para definir términos y derivar categorías. **Resultados**: La investigación resultó en el registro de técnicas aplicadas a los términos candidatos, en la comprensión de las peculiaridades del dominio y, en la medida de lo posible en un estudio de caso, en la obtención de experiencia y buenas prácticas para su uso en dominios similares en áreas estratégicas de la sociedad. **Conclusiones**: existe la necesidad de una metodología para construir ontologías dirigidas a los profesionales de las Ciencias de la Información, que incluya la fase de conceptualización.

Descriptores: Organización del conocimiento. Ontologías. Conceptualización. Representación del conocimiento. Sector eléctrico.

Recebido em: 19.06.2020
Aceito em: 22.03.2021