

Modelagem de relações conceituais para a área nuclear

Luana Farias Sales^{1,2}, Luís Fernando Sayão², Dilza Fonseca da Motta³

¹Programa de Pós-Graduação em Ciência da Informação – Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e Universidade Federal do Rio de Janeiro (UFRJ)

²Comissão Nacional de Energia Nuclear (CNEN)

³Financiadora de Estudos e Projetos (FINEP)

lsales@ien.gov.br; lsayao@cnen.gov.br;
dilzafmotta@yahoo.com.br

Abstract. *The nuclear energy area is a complex domain involving a large number of disciplines, concepts and relations. Despite its long tradition in organizing, processing and dissemination of information, reflected in important databases and international information systems, in recent decades this field has not evolved satisfactorily regarding the development of tools to standardize terminology and, consequently, concerning the extension of the theoretical and methodological framework for conceptual modeling. This fact creates an obstacle in the development of more sophisticated information systems. Starting from the systematization of ontic conceptual relations in the domain of the nuclear area, this paper presents a new way to conceptual modeling, anchored in the theoretical basis of information science. The proposed model, based on a classification principles, seeks to combine categorical and formal relations, to reaching the triadic model of relations for the nuclear area.*

Resumo. *A área de Energia Nuclear é um domínio complexo que envolve um grande número de disciplinas, conceitos e relações. Apesar da sua longa tradição na organização, no tratamento e na disseminação de informação, refletido em importantes bases de dados e sistemas internacionais de informação, nas últimas décadas não evoluiu a contento no que diz respeito à elaboração de instrumentos de padronização terminológica e, conseqüentemente, no que diz respeito à ampliação de arcabouço teórico-metodológico para modelagem conceitual. Esse fato cria um obstáculo no desenvolvimento de sistemas de informação mais sofisticados. Partindo da sistematização das relações conceituais ônticas existentes no domínio da área nuclear, o presente trabalho visa apresentar um novo caminho para a modelagem conceitual, ancorado na base teórica da Ciência da Informação. A modelagem proposta, baseada em princípios classificatórios, procura combinar relações categoriais e formais, chegando ao modelo triádico de relações para a área de ciências nucleares.*

1. Considerações iniciais

A necessidade de elaborar modelos deriva inicialmente da dificuldade de o homem entender a complexidade da realidade do universo que o envolve. Assim, em uma primeira instância, o ser humano elabora modelos para compreender o mundo ou simplesmente uma questão no mundo; estabelecer padrões de comunicação entre ele e outros seres e representar de forma simplificada um objeto ou uma situação no mundo.

Os modelos podem ser construídos “por meio de formalismos matemáticos, fenomenológicos ou conceituais” e permitem “testar hipóteses, tirar conclusões, caminhar no sentido da generalização e da particularização, através de processos de indução e têm sempre vida provisória”. (SAYÃO, 2001, p.83).

Os modelos conceituais são construídos a partir de abstrações que especificam relacionamentos entre conceitos, trabalhando semelhanças, diferenças e outras associações de significado.

Na esfera da modelagem conceitual, identificar os tipos de relações presentes em um domínio pode ser considerada uma etapa da modelagem conceitual deste domínio, pois são as relações que ligam os conceitos a outros conceitos, permitindo evidenciar a abstração de uma dada realidade.

O problema que se coloca, no entanto, é que as relações não-hierárquicas, também chamadas de associativas ou ônticas se manifestam diferentemente de acordo com a área, variando de acordo com o objetivo e com a questão que se visa modelar. Acredita-se que uma possível solução para este problema esteja na proposta de um método para modelagem de relações conceituais que possa ser generalizável a qualquer domínio. O presente trabalho tem por objetivo apresentar um método para modelagem de relações não-hierárquicas que foi aplicado, neste momento, na área de ciências nucleares.

2. A área de Energia Nuclear e a necessidade de modelos conceituais

No domínio da Ciência da Computação, a modelagem conceitual é um estágio anterior ao desenvolvimento do sistema. Nesta área, a elaboração de modelos conceituais fornece subsídios para construção de sistemas eficazes aos seus propósitos. Já na Ciência da Informação, os modelos são construídos para servirem de instrumentos padronizadores de informações, tornando a recuperação e a comunicação mais precisas.

A área de Energia Nuclear tem uma longa tradição na organização, tratamento e na disseminação da informação. O International Nuclear System (INIS), órgão subordinado à Agência Internacional de Energia Atômica, da ONU, deu prosseguimento à política de valorização da informação nuclear, como insumo estratégico para o desenvolvimento das aplicações pacíficas da energia nuclear.

No que tange a elaboração de regras e padrões, o INIS criou um conjunto de ferramentas terminológicas, incluindo classificação e tesauro. O desenvolvimento do tesauro foi derivado de um modelo gráfico chamado INIS Terminology Charts, que exibiu representações gráficas dos conceitos envolvidos e as suas relações, incluindo a intensidade de cada relação e o seu tipo (INIS, 1970).

As iniciativas de construtos terminológicos do INIS foram muito interessantes e consideradas avançadas para os padrões da época. No entanto, apesar dos indícios da

necessidade desses instrumentos acompanharem evolutivamente a produção do conhecimento e as mudanças temporais, a área de Energia Nuclear, nos últimos anos, se manteve indiferente em relação à atualização metodológica de seus instrumentos. Enquanto na tabela de classificação houve uma diminuição considerável do número de áreas abrangidas, o tesouro vem se mantendo estático no que tange a inserção de novos termos e novas relações.

Mesmo com a descontinuidade dessas ferramentas conceituais, elas continuam sendo de fundamental importância para essa área, principalmente devido ao fato de sempre ter utilizado, para compreensão de seus fenômenos, modelos e generalizações como uma ponte entre os níveis de observação e o teórico. Nas últimas décadas, o uso massivo de técnicas de simulação em computadores para realização de experimentos na área, mais uma vez, coloca em voga a necessidade dos modelos. Por fim, a intensa geração de dados, fruto do uso de *software*, leva à necessidade de curadoria desses dados para fins de recuperação e reuso e, neste caso, são os modelos que dão estrutura e significado aos dados. Esses fatores evidenciam a necessidade de modelos conceituais e consequentemente da modelagem das relações conceituais.

3. A modelagem das relações da área de Energia Nuclear

A modelagem das relações da área de Energia Nuclear foi realizada a partir da base teórica da Ciência da Informação para construção de linguagens ou modelos (CAMPOS, 2001b). Essa escolha se justifica por ter esta área abordagens teórico-metodológicas consolidadas para construção de instrumentos de representação da informação e do conhecimento. Conforme abordagem de Campos (2001a), a união da tríade teórica formada pela Teoria da Classificação Facetada (RANGANATHAN, 1967), Teoria Geral da Terminologia (WÜSTER, 1981) e Teoria do Conceito (DAHLBERG, 1978) vem atendendo não apenas a elaboração de linguagens, mas também toda e qualquer necessidade de classificação, sistematização, mapeamento de domínio e modelagem conceitual.

Em trabalho apresentado anteriormente (SALES, 2008), um modelo triádico de relações foi proposto para reunir as relações da Ciência da Informação (duplas de categorias, ex: coisa-material, as quais são aqui chamadas de **Relações Categoriais**) e também as da Ciência da Computação (**Relações formais**, ex: *has_material*, que revelam a forma como um conceito se relaciona com outro). Este modelo é composto da seguinte forma: <Relação categorial1 – relação formal – Relação categorial2> = (<rc1- rf- rc2>). Ex: <Processo-results-Entidade>

Para se chegar ao modelo triádico de relações foi utilizado no processo de modelagem um método que reuniu aspectos conceituais, lógico-classificatórios e matemáticos. Este método está sendo chamado *a priori* de “Método relacional-categorial”, pois visa estabelecer relações a partir da combinação das categorias existentes no domínio mapeado. No que tange aos aspectos conceituais, o método se valeu de abordagens advindas da Teoria do Conceito e da Terminologia para a identificação de definições terminológicas consistentes para os termos que fizeram parte do corpus selecionado, bem como para a intervenção nessas definições, quando necessário. Quanto aos aspectos lógico -classificatórios, o método se valeu da Teoria da Classificação Facetada para identificação das categorias e classificação dos termos identificados para compor o corpus. Com relação aos aspectos matemáticos, o método

utilizou-se de arranjo em uma análise combinatória para compor os pares de relações categoriais.

O método abrangeu as seguintes etapas: 1)Análise das definições. 2) Categorização dos termos; 3)Análise combinatória das categorias; 4)Retirada de assertivas ¹das definições no modelo <rc1-rf-rc2>; 5)Categorização das assertivas e identificação das relações formais; 6)Sistematização das relações formais possíveis para cada par de relações categoriais.

A primeira etapa do método foi a análise do Glossário Nuclear da CNEN (2011), composto por 107 termos, que se configuraram como corpus. Suas definições bem construídas auxiliaram na categorização dos termos e, conseqüentemente, no estabelecimento das relações entre os termos.

A segunda etapa do método foi baseada na Teoria da Classificação Facetada de Ranganathan que sugere cinco categorias fundamentais representadas pela sigla PMEST, cujas iniciais significam: Personalidade, Matéria, Energia, Espaço e Tempo². Os termos foram separados em categorias baseadas nestas últimas, adaptadas para a área em análise.

Chegou-se então a sete categorias: Entidade, Equipamento, Propriedade, Matéria, Processo, Espaço, Tempo. Na categoria Entidade foram agrupados todos os objetos, indivíduos, agentes e produtos. Na categoria Equipamento foram considerados os instrumentos, as ferramentas, as máquinas ou outros artefatos, aparelhos, utensílios, dispositivos e aparatos experimentais que sirvam para executar um processo e/ou gerar um produto. Os conceitos referentes às características, atributos, medidas, dimensões e outras qualidades dos objetos, das entidades, dos processos, do tempo, do espaço etc. foram incluídos na categoria Propriedade. Os conceitos referentes aos materiais constituintes e substâncias foram categorizados em Matéria. Ações, operações e fenômenos em geral foram agregados na categoria Processo. Os conceitos referentes às áreas, locais ou regiões foram agrupados na categoria Espaço. Por fim, na categoria Tempo foram incluídos os conceitos cuja noção de tempo é determinante. Veja o quadro 1.

Quadro 1 - Exemplo de categorização dos conceitos constantes nos glossários

ENTIDADE	EQUIPAMENTO	PROPRIEDADE	MATÉRIA	PROCESSO	ESPAÇO	TEMPO
Átomo	Acelerador	Radioatividade	Elemento combustível	Radioterapia	Área controlada	Meia vida

¹ Assertivas são afirmações ou proposições que podem ser atestadas como verdadeiras. Neste caso, as assertivas são modeladas por triplas que vem a ser predicados binários.

²Para Ranganathan, o PMEST pode ser explicado da seguinte forma: Na categoria **Tempo** estão as idéias isoladas de tempo, na categoria **Espaço** estão aquelas referentes ao local de pertencimento de um determinado objeto, seja ele indivíduo, coisa, fenômeno, entre outras entidades. Na categoria **Energia** estão as idéias de processo, ação ou fenômeno. Na categoria **Matéria** - suas manifestações são de duas espécies - Material e Propriedade, que são partes intrínsecas de um objeto ou processo. Na categoria **Personalidade**, Ranganathan considera através de um método residual, tudo aquilo que não cabe nas outras categorias.

A terceira etapa constou da análise combinatória das categorias identificadas anteriormente. Como na Ciência da Informação as relações se manifestam entre duplas de categorias, combinar as categorias que mapeiam uma área é uma forma de identificar, no nível mais genérico possível, as possibilidades de relações a serem encontradas em um domínio. A partir da análise combinatória chegou-se a 49 pares de relações categoriais. Para citar algumas: entidade-entidade, entidade-propriedade, entidade-matéria, entidade-equipamento, entidade-processo, entidade-tempo, entidade-espaço, entre outras.

Para encontrar as relações formais, partiu-se então para a quarta etapa do método que constou da retirada de assertivas no modelo relação categorial1-relação formal-relação categorial2 <rc1-rf-rc2>, a partir da análise das definições. Conforme primeira coluna do quadro 2:

Quadro 2 - Identificação e categorização das assertivas

ASSERTIVA IDENTIFICADA	CATEGORIZAÇÃO DAS ASSERTIVAS
Área controlada- controla -exposições à radiação	Espaço - controla - processo
Área controlada -previne - disseminação e contaminação radioativa	Espaço - previne- processo

A quinta etapa constou da categorização das assertivas, reescrevendo-as de acordo com as categorias identificadas na etapa 2. O objetivo foi generalizar, no grau máximo, as relações categoriais, deixando em um nível mais específico apenas as relações formais, já que a ideia era identificar todas as possibilidades de relações formais do domínio escolhido. Esta etapa pode ser observada na segunda coluna do quadro 3.

Na sexta etapa, foi realizada a sistematização das relações formais possíveis para cada par de relações categoriais. Assim, chegou-se a 19 pares de categorias e 39 relações formais. No quadro 3 pode ser visualizada uma pequena amostra das relações identificadas.

Quadro 3 - Sistematização das relações

RELAÇÃO CATEGORIAL	RELAÇÃO FORMAL	MODELO TRIÁDICO
ENTIDADE - PROCESSO	emite	entidade - emite- processo
ENTIDADE - ENTIDADE	parte_de contém transforma	entidade - parte_de - entidade entidade - contém - entidade entidade - transforma - entidade
ENTIDADE - PROCESSO	recebe sofre	entidade - recebe - processo entidade - sofre - processo
ENTIDADE - PROPRIEDADE	representa	entidade - representa - propriedade
EQUIPAMENTO - ENTIDADE	acelera produz aumenta	equipamento - acelera - entidade equipamento - produz - entidade equipamento - aumenta - entidade
EQUIPAMENTO - MATÉRIA	contém reutiliza utiliza	equipamento - contém - matéria equipamento - reutiliza - matéria equipamento - utiliza - matéria

5. Considerações finais

O presente trabalho é fruto de estudos que caminham em direção à busca dos fundamentos de modelagem. O assunto em questão é abordado aqui como um ponto de

interseção entre a Ciência da Informação e a Ciência da Computação, pois reúne relações interessantes para as duas áreas. A modelagem conceitual requer cuidado especial com conceitos e também com as relações entre esses conceitos. Desta forma, este trabalho contemplou apenas um desses aspectos: Relações conceituais óticas para modelagem.

Este trabalho é parte de um projeto maior que está em andamento, que envolve curadoria digital de dados de pesquisas nucleares, o que pressupõe a necessidade de modelo conceitual e conseqüentemente uma modelagem das relações. Entende-se que a modelagem das relações conceituais será de fundamental importância para ligar os dados de pesquisa aos documentos que os geraram. Sendo assim, pretende-se também, em momento oportuno, realizar um experimento utilizando o modelo de relações identificadas para, em uma publicação ampliada, ter os dados científicos ligados às publicações de forma consistente.

6. Referências

- CAMPOS, Maria Luiza de Almeida. **A organização de unidades do conhecimento em hiperdocumentos: um modelo conceitual como um espaço comunicacional para a realização de autoria**. Rio de Janeiro: IBICT/UFRJ, 2001b. Tese de Doutorado.
- CAMPOS, Maria Luiza de Almeida. **Linguagem documentária: teorias que fundamentam sua elaboração**. Niterói: EdUFF, 2001a.
- CNEN. **Glossário de termos usados em Energia Nuclear**. 2011. Disponível em: <www.cnen.gov.br/noticias/documentos/glossario_tecnico.pdf>. Acesso em 26 jun. 2012.
- DAHLBERG, I. **Optical structures and universal classification**. Bangalore: Sarada Ranganathan Endowment, 1978.
- GUEDES, V; BORSCHIVER, S. **Bibliometria: uma ferramenta estatística para a gestão da informação e do conhecimento, em sistemas de informação, de comunicação e de avaliação científica e tecnológica**, 2005. Disponível em: <<http://dici.ibict.br/archive/00000508/01/VaniaLSGuedes.pdf>>. Acesso em: 12 jul 2012.
- INIS: **Terminology charts**. Viena: IAEA, 1970.
- RANGANATHAN, S.R. **Prolegomena to library classification**. Bombay: Asia Publishing House, 1967.
- SALES, Luana Farias. Modelo triádico de relações para aplicação em ontologias. Seminário de pesquisa em ontologia no Brasil. In: SEMINÁRIO BRASILEIRO DE ONTOLOGIAS, 1., 11 e 12 de Julho 2008, Niterói, Rio de Janeiro. **Anais...** Niterói: UFF, 2008. Disponível em: <<http://www.uff.br/ontologia/artigos/13.pdf>> Acesso em: 26 jun 2012.
- SAYÃO, Luis Fernando. Modelos teóricos em Ciência da Informação: abstração e método científico. **Ciência da Informação**, Brasília: IBICT, v. 30, n. 1, p. 82-91, jan./abr. 2001.
- WÜSTER, E. L'étude scientifique générale de laterminologie, zone frontalière entre la linguistique, la logique, l'ontologie, L'informatique et les sciences des chose. In: RONDEAU, G. ; FELBER, E. (Org.). **Textes choisis de terminologie**. Québec: GIRSERM, 1981. V.I : fondéments théoriques de la terminologie, p. 57-114.