

QDAontology – Abordagem para o Desenvolvimento de Ontologias em e-Science

Daniele Palazzi^{1,2}, Fernanda Campos^{1,2}, Regina Braga^{1,2}, Elaine Coimbra³

¹Núcleo de Pesquisa em Qualidade de Software

²Mestrado em Modelagem Computacional

³Departamento de Parasitologia, Microbiologia e Imunologia

Universidade Federal de Juiz de Fora (UFJF)

36036-900 – Juiz de Fora – MG – Brasil

daniele.palazzi@ice.ufjf.br; {fernanda.campos, regina.braga, elaine.coimbra}@ufjf.edu.br

Abstract. *This paper describes QDAontology - Quality Driven Approach for e-Science Ontologies, with six stages. Each stage is composed of activities, in each activity artifacts are generated and participants are related to process stages and activities. Quality characteristics were defined. During implementation process the ontology must be then integrated to other ontologies through mechanisms of ontological correspondence. A case study in the biological domain, Human Disease, illustrates the proposal.*

Resumo. *Este artigo descreve a abordagem QDAontology - Quality Driven Approach for e-Science Ontologie, com seis etapas. Cada etapa é constituída de atividades, em cada atividade são gerados artefatos e os participantes estão relacionados com as etapas e atividades do processo. Características de Qualidade foram definidas. Durante o processo de desenvolvimento a ontologia deve ser integrada a outras ontologias, através de mecanismos de correspondência ontológica. Um estudo de caso no domínio biológico de Doenças Humanas ilustra a proposta.*

1. Introdução

Nas aplicações científicas as ontologias, em geral, são usadas em conjunto com outras ontologias já existentes, portanto devem ser interoperáveis. Elas também se caracterizam por serem desenvolvidas por equipes multidisciplinares, onde os conteudistas pertencem à área do domínio da aplicação e os construtores são Engenheiros Ontológicos oriundos da Ciência da Computação. Descrevemos neste artigo a abordagem QDAontology - Quality Driven Approach for e-Science Ontologies, para a construção de ontologias no domínio da e-Science e composta por seis etapas [Palazzi, D. C. 2010]. Ilustramos a aplicação na área de Biologia. O presente trabalho está organizado em três seções. Na seção 2 descrevemos a QDAontology e ilustramos com os artefatos da Celo-Human Disease e na seção 3 estão as considerações finais.

2. QDAontology – Uma Abordagem para o Desenvolvimento de Ontologias em e-Science

A abordagem propõe um processo de engenharia ontológica com todos os seus componentes. Para cada etapa o modelo descreve (i) participantes, (ii) atividades, (iii)

artefatos (iv) características de qualidade. Cada etapa é constituída de atividades, em cada atividade são gerados artefatos e os participantes estão relacionados com as etapas e atividades. Para cada etapa foram definidas características de qualidade. A partir da evolução dos artefatos ocorre o desenvolvimento da ontologia. São seis etapas e cada ciclo das etapas termina com uma versão da ontologia. A Figura 1 ilustra os componentes da proposta.

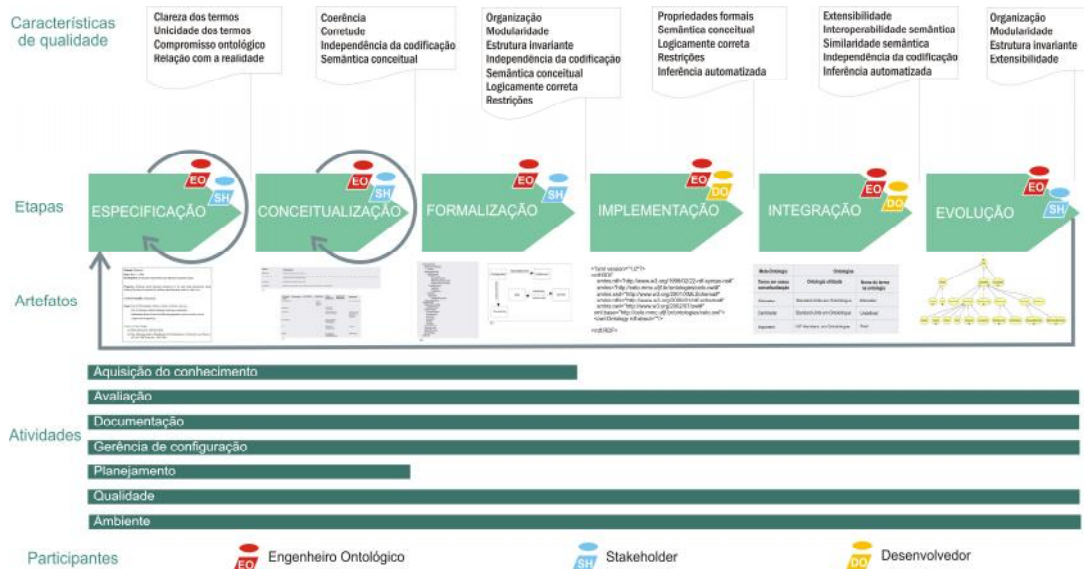


Figura 1. Componentes da abordagem QDAontology.

Para validar e avaliar a abordagem QDAontology foi utilizada a ontologia CelO [Matos, E. *et al.* 2009], resultando na CelO-Human Disease, pela existência de outras ontologias no domínio da área da biologia, como as do OBO Foundry [OBO Foundry 2010], e pela disponibilidade de especialistas. O estudo de caso foi apoiado pelo mecanismo de interoperabilidade entre ontologias denominado alinhamento.

ETAPA 1 - ESPECIFICAÇÃO

Tem como proposta elaborar um documento que identifica os objetivos, especifica os usuários, o conjunto de termos a serem representados, as características e as granularidades e também um conjunto de perguntas de competência para delimitar o escopo e o propósito do domínio. O quadro 1 ilustra o documento gerado.

ETAPA 2 - CONCEITUALIZAÇÃO

Organiza e estrutura o conhecimento como modelos significativos do conhecimento. Como artefatos gerados têm-se glossário e o dicionário de termos (Figura 2).

ETAPA 3 - FORMALIZAÇÃO

Transforma o modelo conceitual em um modelo formal ou semi-computável, possibilita a transformação dos requisitos e a definição dos termos em um projeto, sugerindo uma arquitetura e adaptando o projeto para o ambiente de implementação (Figura 3).

ETAPA 4 - IMPLEMENTAÇÃO

Constrói modelos computáveis utilizando ferramentas para o desenvolvimento de ontologias. Define a organização formal dos termos, conceitos e relacionamentos de forma a permitir inferência automatizada, através da definição de regras e restrições. O resultado é a ontologia codificada em uma linguagem formal.

Quadro 1. Especificação da ontologia CeIO -Human Disease.

Domínio	<i>Human Disease</i>
Data	2009
Desenvolvida por	Elaine Coimbra
Objetivo	<i>Expand the domain represented in the ontolgy CeIO from infectious human disease.</i>
Nível de formalidade	<i>Semiformal</i>
Escopo	<i>List of 48 concepts: Acanthamoeba infection, African trypanosomiasis, Amebiasis, Angiostrongyliasis, Anisakiasis, Ascariasis, Babesiosis, Bacterial Diseases, Balantidiasis, Capillariasis, Cell, Chagas Disease, Clonorchiasis, Cryptosporidiosis, Cutaneous Larva Migrans, Cyclosporiasis, Cysticercosis, Diphyllorhynchiasis, Disease, Dracunculiasis, Ectoparasite, Enterobiasis, Fascioliasis, Fungal Diseases, Giardiasis, Helminths, Hookworm Infection, Human Disease, Hydatidosis, Hyme nolepiasis, Infectious Disease, Visceral Larva Migrans, Leishmaniasis, Malaria, Microsporidiosis, Naegleria Infection, Onchocerciasis, Paragonimiasis, Parasitic Disease, Protozoa, Schistosomiasis, Strongyloidiasis, Taeniasis, Toxoplasmosis, Trichinellosis, Trichomoniasis, Trichuriasis, Viral Disease.</i> <i>List of 12 properties: isDisease, isDiseaseCausedBy, isParasitic, isTypeOf, hasCause, hasDefinition, hasDiagnostic, hasHelminths, hasPrevention, hasPrognostic, hasProtozoa,, hasTransmission.</i>
Fontes de conhecimento	http://www.cdc.gov/ncidod/dpd/parasites/ http://www.who.int/topics/en/ http://www.dpd.cdc.gov/dpd/Default.htm 6 stakeholders

Concept name	Synonyms	Class attributes	Relations
Taeniasis	Tapeworm infection Terriase	-	hasCause hasDefinition hasDiagnostic hasHelminths hasPrevention hasPrognostic hasTransmission
Parasitic_Disease	Doença parasitária	Ectoparasite Helminths Protozoa	hasParasitic

Figura 2. Dicionário de conceitos.

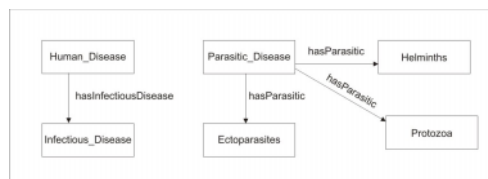


Figura 3. Diagrama de relacionamento.

ETAPA 5 - INTEGRAÇÃO

Considera a reutilização e o compartilhamento da ontologia desenvolvida com outras ontologias do mesmo domínio. A interoperabilidade entre ontologias é um fator essencial e a realização dessa etapa garante o compartilhamento e troca de informações entre aplicações. Técnicas de correspondência ontológica devem ser usadas, como mapeamento, alinhamento, combinação e integração. O artefato produzido nesta etapa é um documento contendo as informações essenciais relacionadas à correspondência entre as ontologias (quadro 2).

Quadro 2. Documento de Integração da Celo -Human Disease com a OBO Foundry.

Meta-Ontologia	Ontologias	
Termo na CeIO – Human Disease	Ontologia utilizada	Nome do termo na ontologia utilizada
<i>Parasitic_Disease</i>	DOID in OBO	<i>parasitic disease</i>
<i>Giardiasis</i>	DOID in OBO	<i>giardiasis</i>
<i>Babesiosis</i>	DOID in OBO	<i>babesiosis</i>
<i>Microsporidiosis</i>	DOID in OBO	<i>microsporidiosis</i>
<i>Trichomoniasis</i>	DOID in OBO	<i>trichomoniasis</i>

ETAPA 6 - EVOLUÇÃO

Tem por objetivo apoiar o enriquecimento do conhecimento, acrescentando novos ou expandindo conceitos. Ao atingirmos essa etapa o artefato gerado é uma versão final da ontologia evoluída (Figura 4).

Em relação a outras metodologias relacionadas na literatura na QDAontology há preocupação com a qualidade da ontologia como um todo, diversos atributos de qualidade foram definidos em todas as etapas, da aquisição de conhecimento à documentação final. Devido à importância da Integração para aplicações científicas a

mesma constitui uma etapa da abordagem. A QDAontology define os participantes associados a cada etapa. Apesar de ter sido elaborada para a criação de ontologias científicas, pode ser aplicada, sem comprometimento, a todos os domínios.

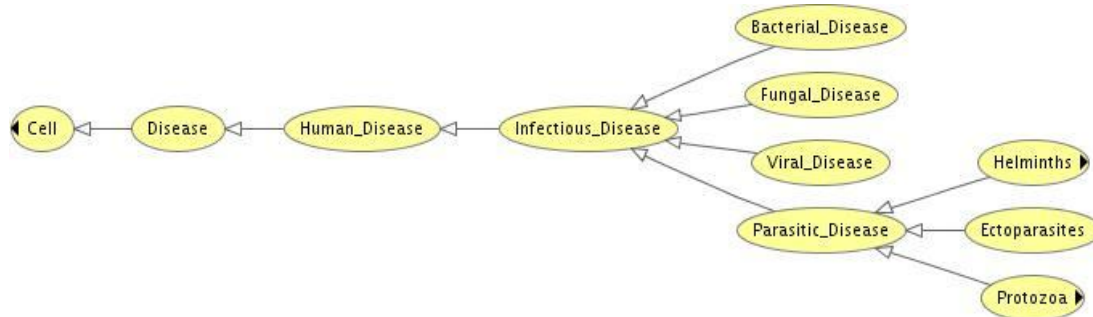


Figura 4 – Parte da ontologia CelO–Human Disease.

3. Considerações Finais

A Engenharia Ontológica é uma área de pesquisa recente e ainda não existem metodologias para o seu desenvolvimento que sejam largamente utilizadas e aceitas como padrão. Este trabalho propôs a abordagem QDAontology, para o desenvolvimento de ontologias em projetos de e-Science, com as etapas necessárias para desenvolvimento, juntamente com um mecanismo que promova a recuperação e integração das informações presentes em outras ontologias do mesmo domínio ou de domínios complementares. A abordagem proposta possibilitou a evolução da ontologia CelO em CelO-Human Disease, cujos artefatos ilustraram esse artigo.

A abordagem se mostrou adequada para projetos de e-Science, caracterizados pela atuação dos especialistas do domínio nas etapas iniciais de extração e formalização do conhecimento, deixando para os profissionais da Ciência da Computação a construção da ontologia e sua interoperabilidade com outras complementares. A aderência dos artefatos às características de qualidade garantiu uma versão final evoluída e pronta para ser utilizada em aplicações científicas onde a troca semântica de informações é necessária. No contexto de infraestruturas para e-Science podemos destacar como trabalhos futuros o estudo de técnicas de alinhamento entre ontologias que melhorem a busca por serviços Web semânticos e a especificação de métricas em relação à atividade de qualidade.

Referências

- Matos, E. E. S., Campos, F., Braga, R., Palazzi, D. CelOWS: a ontology based framework for the provision of semantic web service related to biological models. *Journal of Biomedical Informatics*, 2009.
- OBO Foundry. 2010. “The Open Biomedical Ontologies”. Disponível em <http://www.obofoundry.org/>.
- Palazzi, D. C. QDAontology – Abordagem para o desenvolvimento de ontologias em e-Science: um estudo de caso em biologia. Juiz de Fora, 2010. Dissertação (Mestrado em Modelagem Computacional) - Universidade Federal de Juiz de Fora. Juiz de Fora, MG, Brasil.