



Modelagem de Ontologia para Apoiar Expansão de Consultas ao Acervo de Bibliotecas

JOÃO PAULO VOLPATE¹, CRISTIANE AKEMI YAGUINUMA²

¹Graduando em Tecnologia em Análise e Desenvolvimento de Sistemas, Bolsista PIBITI/CNPQ, IFSP, Câmpus Araraquara, jpaulov@outlook.com

²Docente, IFSP, Câmpus Araraquara, cristiane.yaguinuma@ifsp.edu.br

Área de conhecimento: Sistemas de Informação – 1.03.03.04-9

RESUMO: O sistema de busca via palavra-chave vem se provando cada vez mais ineficiente e ineficaz, dificultando e em muitos casos impossibilitando a conclusão das pesquisas. Devido a problemas como esses a ontologia vem ganhando destaque, por trabalhar utilizando o campo semântico contido no termo de busca, é capaz de trazer resultados mais ligados ao termo pesquisado. O objetivo desse projeto foi desenvolver uma ontologia para auxiliar a expansão semântica de consultas permitindo a realização de consultas no acervo de bibliotecas. A ontologia construída foi avaliada por bibliotecários a partir da aplicação de questionários. Os resultados obtidos indicam que o uso da ontologia e suas inferências permite retornar resultados mais significativos que a abordagem com base em palavras-chave.

PALAVRAS-CHAVE: biblioteca digital; engenharia de ontologias; imprecisão; ontologia; representação de conhecimento.

INTRODUÇÃO

Sistemas de busca refinados vem se tornando cada dia mais importantes diante da vasta gama de informações pesquisadas. Diversas vezes o erro de digitação ou a falta de contexto na busca são suficientes para retornar resultados insuficientes.

Tais falhas e limitações são tratadas com o uso da ontologia, que permite retornar respostas significativas por associar os termos pesquisados ao campo semântico modelado nela. Segundo Gruber (2009), ontologia é um conjunto de primitivas representacionais que podemos utilizar para modelar um domínio de conhecimento, essas primitivas são os atributos, indivíduos, conceitos, relacionamentos e suas restrições.

Problemas como esses ocorrem com frequência em buscas realizadas no contexto de bibliotecas e, portanto, o objetivo desse projeto foi construir uma ontologia para possibilitar consultas semânticas ao acervo de bibliotecas, reduzindo a quantidade de resultados indesejados nas buscas.

METODOLOGIA

Para a construção da ontologia foram utilizados os procedimentos da Metodologia 101 (NOY; MCGUINNESS, 2001). O primeiro passo é a busca por ontologias já consolidadas. Foi então utilizada a metodologia ACM para modelar o conhecimento na ontologia existente.

O sistema ACM de classificação (ACM, 2016) é uma referência internacional de vocabulário estruturado. Após a seleção dos livros foi elaborada uma planilha com os dados que seriam utilizados na ontologia, como assunto, ISBN, publicação, autores, de um total de 57 livros na área de informática

A próxima etapa da metodologia foi expandir a taxonomia ACM a partir da coleta de requisitos com os especialistas, sendo feita também a análise de dados do acervo. Assim, a Biblioteca do IFSP disponibilizou os dados de acervos dos livros, composto por ISBN, assuntos (palavra-chave), título e autores.

Com essa base de dados estabelecida foi iniciado o processo de construção da ontologia com os campos de conhecimento da ACM. Foram então definidos os rótulos do tipo *preflabel*, que correspondem às palavras-chave escritas da forma mais adequada, e os rótulos do tipo *altlabel*, que são as variações mais possíveis de um mesmo termo.

Definida a base de conhecimento na ontologia foi então iniciado o processo de associação dos livros às categorias de livros modeladas na ontologia, esse processo foi elaborado com o estudo junto a especialistas da área para determinar a melhor classificação do livro dentro das classes da ontologia.

Para a construção da ontologia foi utilizada a ferramenta Protégé (KNUBLAUCH et al., 2004), que permite modelar classes, propriedades, indivíduos e relacionamentos de um domínio. Além disso, essa ferramenta pode ser utilizada com um raciocinador. Um raciocinador utiliza a Lógica de Descrição implementada na ontologia nos permitindo fazer inferências sobre o determinado campo de conhecimento. Essa técnica permite deduzir informações implícitas a partir dos elementos modelados na ontologia.

RESULTADOS E DISCUSSÃO

A partir da Metodologia 101 foi construída a ontologia no domínio de informática, tendo como base a taxonomia proposta pela ACM, e nas palavras-chave obtidas com os dados do acervo da biblioteca, como demonstrado na Figura 1.

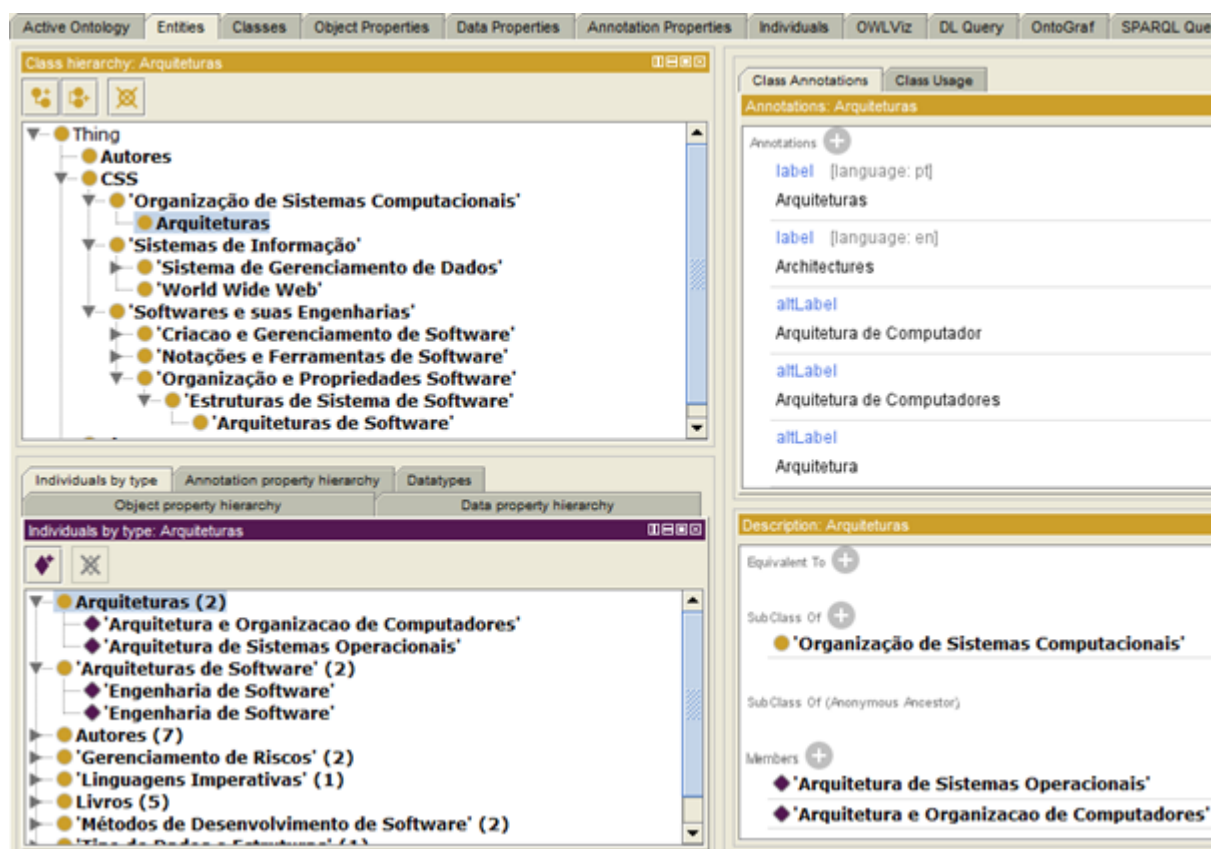


Figura 1 – Protégé – Trecho da Ontologia construída.

Fonte: Elaborada pelo autor.

Na Figura 1 são demonstrados os aspectos da ontologia, a hierarquia de classes (a *Class Hierarchy*) as anotações da classe (*Class Annotations*), os indivíduos e as descrições da classe. Uma das principais características de uma ontologia é a hierarquia de classes, abrangendo o conteúdo da área de conhecimento retratada em níveis de especificação. Na ontologia modelada nesse projeto foi utilizada a classificação de conhecimento da ACM (ACM, 2016) para elaboração desse sistema de classes. Na aba de descrição de classes são definidas suas propriedades hierárquicas e os indivíduos pertencentes a essa mesma classe, existem casos onde são inferidos membros pelo raciocinador.

Ao lado da aba de classes ficam as anotações do tipo *preflabel* e *altlabel* referentes a classe selecionada. As formas de escrita dos *labels* foram feitas de acordo com as fornecidas pela ACM e também foram internacionalizadas para português e inglês (nativos da classificação ACM). Com a determinação de *altlabels* nas classes é possível criar uma herança de *labels* de forma que um livro das subclasses como

“Arquitetura de Software” herdará também as *labels* presentes nas superclasses dela, como “Estruturas de Sistema de Software”. Com isso é possível evitar a inserção exaustiva de *tags* comuns como Informática, que está presente em todos os livros dessa ontologia.

O raciocinador da Protégé analisa a ontologia através das propriedades modeladas nas classes para fazer as inferências. Também são utilizadas as propriedades de objeto e de indivíduos para encontrar relacionamentos entre indivíduos pelo raciocinador. Por meio desses recursos a ontologia auxilia o sistema de busca a desviar de erros de escrita e problemas de ambiguidade. Nota-se na Figura 1 que existem duas classes com nomes semelhantes, “Arquitetura” e “Arquitetura de Software”, em um sistema de busca comum o termo de busca Arquitetura traria livros de ambas as classes sem distinção nenhuma de contexto, na ontologia existem informações suficientes para tratar a ambiguidade em uma ferramenta de expansão de consulta, através do uso de API's que trabalham com a plataforma OWL.

Ao finalizar esse primeiro modelo da ontologia foi realizada uma reunião com os bibliotecários do IFSP – Araraquara, em que foi apresentada e explicada a ontologia, demonstrando as vantagens dessa abordagem de pesquisa. Ao final da reunião foi proposto um questionário abordando a experiência de busca com a ontologia, o nível de satisfação com as inferências feitas e se a implantação de uma ontologia seria vantajosa para as buscas ou se o sistema atual (buscas por palavra-chave) é o suficiente.

Os resultados mostraram que os funcionários responsáveis pela biblioteca ficaram satisfeitos com a ontologia, com ressalvas para diminuição de falhas que as inferências proporcionam e a precisão nos resultados.

CONCLUSÕES

Como citado no artigo, o sistema de busca tradicional pode ser insuficiente para o trabalho cotidiano de alguns usuários. Em sistemas como o de busca em acervos de biblioteca, o uso da ontologia pode ser de grande ajuda auxiliando na obtenção de resultados mais significativos.

A partir de estudos de artigos científicos, reuniões com especialistas da Biblioteca do Instituto, e estudos a respeito da construção de ontologias foi criada uma ontologia para as consultas nos livros da Biblioteca. Segundo os bibliotecários os resultados do projeto foram muito satisfatórios, uma das sugestões foi justamente a implantação desse sistema de buscas nos diversos campus do Instituto, padronizando o sistema de busca, evitando os erros e divergências de informações que existe atualmente.

O projeto será analisado por especialistas das áreas de atuação abordadas na ontologia, a fim de refinar o conhecimento representado.

Em projetos futuros a ontologia será utilizada para expandir consultas por palavras-chave e tratamento de ambiguidade, para avaliar os resultados de expansão de consultas baseadas na ontologia elaborada.

AGRADECIMENTOS

Os autores agradecem ao Programa PIBITI-CNPq – Edital 281/2015 – pelo auxílio financeiro a este trabalho de pesquisa, e a Biblioteca do IFSP Campus Araraquara pelo apoio em reuniões e material de busca.

REFERÊNCIAS

ACM Computing Classification System. Disponível em: <http://www.acm.org/about/class/2012>. Acesso 10 de junho de 2016.

GRUBER, T. R. Ontology. In: LIU, L.; OZSU, M. T. (Ed.). **Encyclopedia of Database Systems**. New York, NY, USA: Springer, 2009. p. 1963-1965.

KNUBLAUCH, H. et al. The Protégé OWL Plugin: An open development environment for semantic web applications. In: INTERNATIONAL SEMANTIC WEB CONFERENCE. **Proceedings...** Hiroshima, Japan: Springer, 2004. p. 229-243.

NOY, N. F.; MCGUINNESS, D. L. **Ontology Development 101: A Guide to Creating Your First Ontology**. Stanford University. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880. 2001.