XXVII Congresso de Iniciação Científica Unicamp
16 a 18 de outubro de 2019 - Campinas | Brasil

Gerenciamento semi-automático de marcações semânticas em páginas web

Giovanna Batalha*, Julio Cesar dos Reis

Resumo

Marcações semânticas em páginas Web são relevantes para apoiar a organização e busca de conteúdo. O modelo de dados RDF desenvolve um papel central para a publicação de dados estruturados na Web. Este projeto visa o estudo das técnicas e modelos para anotação semântica de conteúdos Web com uso de RDF-a. Será apresentado uma síntese da literatura no tema e resultados iniciais da concepção do método para o gerenciamento semi-automático de marcações em páginas Web.

Palavras-chave:

Marcações semânticas, Ontologias, RDFa.

Introdução

Páginas Web desenvolvidas em Hypertext Markup Language (HTML) requerem meios de tornar à semântica da informação do conteúdo explícita para máquinas. O modelo Resource Description Framework (RDF)[1] de dados estruturados permite descrever de forma padronizada dados interoperáveis na Web. Nesse contexto, RDFa[2] permite associar marcações RDF em documentos HTML.

Esta pesquisa objetiva investigar uma técnica semi-automática para, dado um documento HTML e uma ontologia, analisar esse documento visando identificar conceitos presentes na ontologia. Dessa maneira, sugerir marcações RDF e embutí-las na página utilizando a tecnologia RDFa. Este relatório apresenta resultados iniciais rumo ao desenvolvimento de uma ferramenta que seja capaz de realizar estas atividades propostas.

Resultados e Discussão

Esta pesquisa propõe o estudo e desenvolvimento de uma ferramenta para auxiliar a identificação de conceitos descritos em ontologias de um domínio, que resultarão em marcações RDFa em um documento HTML.

Em um segundo passo, propomos monitorar a evolução da ontologia a fim de sugerir atualizações e ajustes nas marcações já definidas via à ferramenta com base na evolução dos conceitos da ontologia. Nosso objetivo é facilitar o gerenciamento das marcações pelos usuários finais. No contexto desta investigação, propomos um processo para tais atividades dentro da ferramenta:

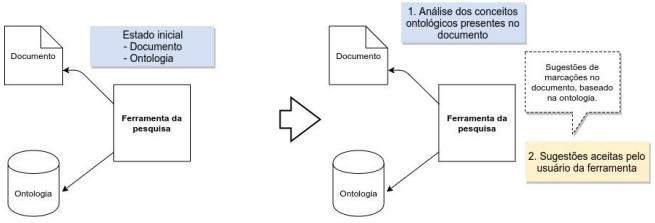

Imagem 1.Etapa em que a aplicação recebe os dados e sugere as primeiras marcações no documento HTML.

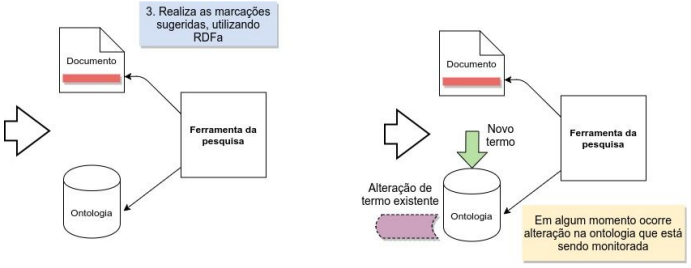

Imagem 2. Etapa em que a aplicação realizada as marcações em RDFa no documento.

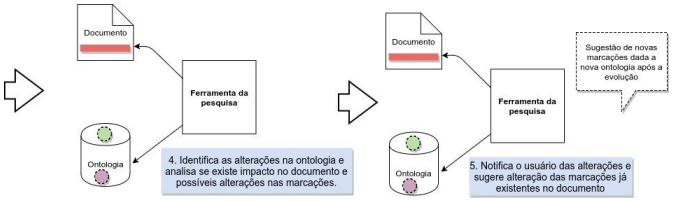

Imagem 3. Etapa do fluxo em que a aplicação de software reage à evolução na ontologia.

Conclusões

Ao manter os documentos HTML associados a conceitos de ontologias é possível expressar conhecimentos e evoluções dos conceitos nos dados Web. Isso é bom na busca e manipulação de dados tornando a informação mais proveitosa por diferentes aplicações.

Para este fim, foi concebido uma prototipação, para validar a possibilidade de se realizar as atividades de marcação semântica e monitoramento da evolução ontológica. De acordo com os protótipos implementados será possível validar a efetividade das soluções e vislumbrar o aperfeiçoamento da ferramenta. A avaliação envolve analisar o uso das funcionalidades da ferramenta por potenciais usuários.

[1] https://www.w3.org/RDF/
[2] https://www.w3.org/TR/rdfa-primer/

(cc) BY-NC-ND