

A SEMANTIC PROPOSAL TO FEDERATE SPARQL QUERIES ON EKGS

XIII Encontro de Pesquisa e Pós-Graduação

Tulio Vidal Rolim, Vania Maria Ponte Vidal

Um Grafo de Conhecimento Corporativo (EKG) é compreendido como uma tripla $\{OD, V, L\}$: onde $OD = \text{Ontologia de Domínio}$; $V = \text{Conjunto de Visões Locais} \{Ovi, Mvi\}$, onde Ovi é um conjunto de Ontologias Locais de cada fonte de dados S e Mvi é um mapeamento entre Ov e S . $L = \text{Conjunto de Visões de Ligações entre } V's$. A proposta pode ser realizada através de duas abordagens distintas: i) Federada sob Endpoints e ii) Federada sob Repositórios, neste trabalho focou-se na proposta Federada sob Endpoints. Como forma de validar a proposta foi realizado um estudo de caso envolvendo dados da REDESimples (REDESIM) e da Receita Federal do Brasil (RFB) com ênfase em encontrar uma resposta para a seguinte pergunta: “Existem empresas na RFB com sócios distintos na REDESIM?”. Para responder a questão são realizados os seguintes passos: i) uma consulta SPARQL Q é processada através do Semagrow um Federador SPARQL F' responsável por fazer a seleção das fontes $S_1 \dots S_n$ aptas a responderem um triple pattern t , onde cada fonte é um Endpoint SPARQL $E_1 \dots E_n$, sendo a RFB compreendida como S_1 acessível em E_1 e a REDESIM como S_2 sendo acessível E_2 ; ii) A consulta Q é dividida em sub-consultas $Q_K \dots Q_{Kn}$ e executadas em E_1 e E_2 ; iii) Para S_1 o armazenamento é relacional necessitando-se da utilização do Ontop para realização da virtualização através de um wrapper que obtém uma sub-consulta Q_K e faz a reescrita em uma consulta SQL Q_Q sob S_1 . Para S_2 no formato RDF os dados são armazenados no Triplestore (GraphDB) e acessados diretamente através de E_2 ; iv) Após processar cada sub-consulta $Q_K \dots Q_{Kn}$ o resultado é retornado ao Semagrow que faz a junção dos resultados e retorna diretamente ao usuário. O presente trabalho expõe um relato de proposta para realização de consultas em EKGs híbridos utilizando a abordagem sob Endpoints SPARQL. Como trabalhos futuros pretende-se apoiar a proposta através de uma ferramenta para configuração e integração semi-automática entre os componentes da proposta.

Palavras-chave: Enterprise Knowledge Graphs. Federate Query. Ontology. Semantic.