Título: **CONSULTAS SOBRE FONTES DE DADOS LIGADOS**

BASEADAS EM RECONHECIMENTO DE ENTIDADES

NOMEADAS

Horário: **10:00h**

Data: **16/12/19**

Local: **Bloco 952 - Sala de Seminários**

Resumo:

A Web evoluiu de uma rede de documentos interligados a uma onde tanto documentos e dados estão ligados, resultando no que é comumente conhecido como a Web de Dados, que inclui uma grande variedade de dados, normalmente publicados no formato RDF, sobre múltiplos domínios. Métodos intuitivos de acessar os dados RDF possuem grande importância, uma vez que a abordagem padrão seria executar uma consulta em SPARQL. Entretanto, isso pode ser muito difícil para usuários não-técnicos. Neste trabalho, abordamos o problema de question answering sobre bases RDF. Dada uma busca em linguagem natural ou em palavras-chaves, nosso objetivo é traduzi-la em uma consulta formal em SPARQL que capture a informação necessitada. Nós propomos duas abordagens baseadas em esquema para buscar sobre dados RDF sem nenhum conhecimento prévio da ontologia: Von-QBE e Von-QBNER. Isso é diferente do estado da arte uma vez que suas abordagens são baseadas nas instâncias de dados. Entretanto, isso pode ser infactível em cenários de Big Data, onde os dados são demasiados grandes e requerem muitos recursos computacionais para manter a base em memória. Também, muitas dessas soluções requerem que a base esteja triplificada, o que pode ser uma tarefa difícil em bases de dados legado. Por esta razão, Von-QBE utiliza apenas o esquema da base RDF para responder a busca do usuário. Entretanto, a busca do usuário pode conter informações sobre as instâncias de dados, que não vai corresponder, sintaticamente, a nenhum conceito ou propriedade no esquema da ontologia. Por exemplo, a busca “filmes com Angelina Jolie”. Considere que o esquema apresente apenas os conceitos Filme e Atriz, e a propriedade estrelando, que relaciona os dois conceitos. Se utilizarmos apenas o esquema da ontologia, apenas Filme será identificado na busca. Von-QBNER resolve essa limitação identificando as instâncias envolvidas na busca e o conceito a que correspondem no esquema utilizando modelos de Named Entity Recognition (NER). Os resultados são promissores para alguns conjuntos de dados reais avaliados, considerando que apenas o esquema da ontologia foi utilizado para gerar as consultas em SPARQL.

Banca:

- Prof. Dr. José Antônio Fernandes de Macedo (MDCC/UFC - Orientador)
- Profª. Drª. Ticiana Linhares Coelho da Silva (UFC - Coorientador)
- Prof. Dr. César Lincoln Cavalcante Mattos (MDCC/UFC)
- Prof. Dr. Marco Antonio Casanova (PUC/RJ)