



UNIVERSIDADE FEDERAL DE PERNAMBUCO
GRADUAÇÃO EM SISTEMAS DE INFORMAÇÃO
CENTRO DE INFORMÁTICA
2015.2



**OpenCInProj: Publicação e Consumo de Dados Abertos
Conectados sobre os Projetos de Cooperação do CIn/UFPE**

PROPOSTA DE TRABALHO DE GRADUAÇÃO

Aluno: Karine Galdino Maia Gomes (kgmg@cin.ufpe.br)

Orientador: Bernadette Farias Loscio (bfl@cin.ufpe.br)

Recife, 29 de Setembro de 2015

1. Contextualização

A Web está cada vez mais presente no cotidiano das pessoas, seu uso varia desde simples buscas por informações até realização de compras. Devido a esse crescimento, a quantidade de dados disponíveis na Web vem aumentando exponencialmente e estruturá-los vem se tornando um desafio. A partir desse contexto, a Web vem passando por algumas mudanças e evoluções para ter um desempenho que atenda as necessidades atuais.

A Web surgiu por volta da década de 90 e tinha como objetivo principal a leitura de informações. Segundo Berners-Lee, sua primeira implementação, a Web 1.0, era considerada "Web read-only". Basicamente, a Web permitia apenas a busca e visualização de informações. Com posteriores avanços, a Web 1.0 passou a ser considerada como "estática" [1].

A segunda implementação da Web, a Web 2.0, surgiu com o objetivo de permitir que os usuários deixassem de ser apenas espectadores e passassem a contribuir com a geração de conteúdo na Internet. Essa versão da Web é definida como "dinâmica" e é utilizada, atualmente, como meio de interação e comunicação global. Um dos grandes desafios da Web 2.0 é lidar com o grande volume de dados disponíveis, além de não possuir tecnologias que filtrem as informações. A partir deste desafio, surge a Web 3.0.

A ideia de Web Semântica, ou Web 3.0, surgiu em 2001, e tinha como conceito permitir que os dados fossem estruturados de forma que seres humanos e máquinas entendessem as informações disponíveis na Web. A partir destas características, a Web 3.0 é definida como "participativa". Diferente da Web 2.0 que faz conexão apenas entre pessoas, a Web semântica conecta dados. Esta implementação tem como objetivo estruturar os conteúdos disponíveis na Web de forma mais inteligente. Para isto, foi necessário a criação de tecnologias que dessem suporte à Web Semântica, como RDF (Resource Description Framework), SPARQL (SPARQL Protocol and RDF Query Language) e OWL (Web Ontology Language) [2]. Estas linguagens colaboram com a proposta da Web Semântica de estender a Web de documentos para uma Web de Dados [3].

O Centro de Informática da Universidade Federal de Pernambuco possui informações do próprio centro, porém o conteúdo encontra-se de forma desorganizada, dificultando a extração de informação dos dados. Portanto, a partir disso, o CIn criou um ambiente de dados, conhecido como OpenCIn, onde é possível extrair conhecimentos diversos sobre os docentes, como por exemplo suas áreas de interesse ou até mesmo participações em eventos. Neste projeto foi imprescindível o conhecimento sobre Web Semântica e ontologia [4], como também

o conceito de dados abertos, que tem por definição que qualquer pessoa pode livremente usar, reutilizar e redistribuir dados, estando sujeito a, no máximo, a exigência de creditar a sua autoria e compartilhar pela mesma licença [5].

2. Objetivos

O objetivo deste trabalho é criar um conjunto de dados abertos conectados sobre os projetos de cooperação do Centro de Informática da UFPE. Portanto, será preciso criar um dataset que se conecte com os demais dados do OpenCIn para possibilitar a identificação de parcerias, como também mapear competências dos docentes do CIn. Para isso, será feita uma extensão da ontologia para permitir a descrição dos dados relativos aos projetos.

Além disso, uma aplicação será desenvolvida para facilitar a análise e visualização das informações extraídas a partir do conjunto de dados proposto.

3. Cronograma

A tabela 1 apresenta o cronograma do Trabalho de Graduação:

Atividade	Setembro				Outubro				Novembro				Dezembro				Janeiro					
Elaboração da Proposta e revisão da literatura	X	X	X	X	X																	
Extensão da ontologia						X	X	X	X	X												
Povoamento do dataset										X	X	X										
Publicação do dataset OpenCInProj e implementação da aplicação de análise dos dados.											X	X	X	X	X							
Elaboração do relatório final														X	X	X	X					
Defesa																					X	

Tabela 1- Cronograma de Trabalho de Graduação

4. Possíveis avaliadores

Os possíveis avaliadores para o trabalho de graduação serão:

- Bernadette Farias Lóscio
- Carla Taciana Silva

5. Referências

[1] W3C BRASIL, **Internet e Web: passado, presente e futuro**, BRASIL, 2008. Disponível em: <<http://www.w3c.br/palestras/internet-web-jun-jul-2008/internetWeb-02.html/>>. Acesso em: 26 de setembro 2015.

[2] W3C BRASIL – **Web Semântica**, BRASIL, 2011. Disponível em: <<http://www.w3c.br/Padroes/WebSemantica/>>. Acesso em: 25 de setembro 2015.

[3] MENDES, Graycielli; XAVES, Ana Carolina; TEIXEIRA, Jean; GOMES, Fernanda. **Web Semântica**. Resende, RJ. Disponível em: <<http://w.aedb.br/seacIV/SI/Poster/websemantica.pdf/>>. Acesso em: 29 de setembro 2015.

[4] LIMA, Júnior César de; CARVALHO, Cedric Luiz de. **Uma Visão da Web Semântica**. Disponível em: <http://www.portal.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_001-04.pdf/>. Acesso em: 29 de setembro 2015.

[5] PORTAL BRASILEIRO DE DADOS ABERTOS, **O que são dados abertos?**. Disponível em: <<http://dados.gov.br/dados-abertos/>>. Acesso em: 24 de setembro 2015.

6. Assinaturas

Karine Galdino Maia Gomes
Orientando

Bernadette Farias Loscio
Orientador