

Extreme Spots

João Aguiar ¹, Pedro Carmona ²

jaguiar@student.dei.uc.pt, pcarmona@student.dei.uc.pt

Resumo: Serviços de pesquisa de pontos de interesses são actualmente uma ferramenta essencial para planear uma viagem. Mas muitas vezes eles pecam pela falta de informação. Este trabalho pretende utilizar a web semantica para realizar uma ferramenta deste tipo uma vez que poderá ser a solução para este problema, permitindo relacionar informação com toda a web. Alem disso foi possível criar um artefacto para o sistema iOS.

Palavras-chave: Ontologias, RDF, Triple-Store, POIs.

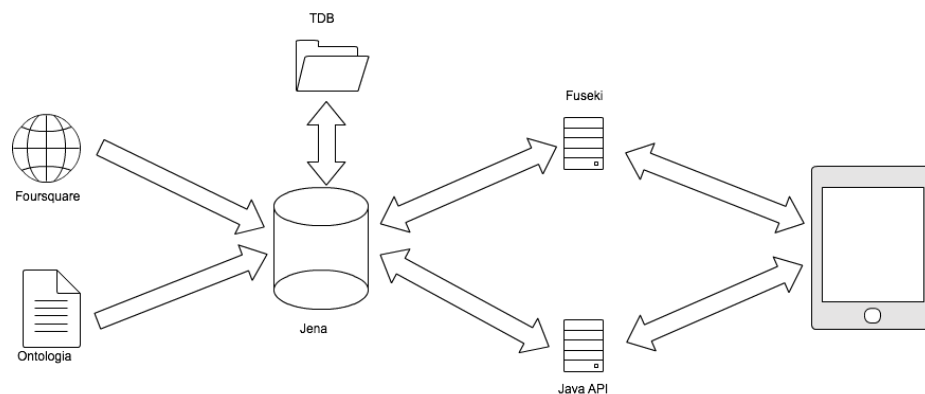
1. Introdução

A web semântica é uma abordagem da ciência da computação que pretende implementar um modelo de dados, onde cada objecto possui um conjunto de conceitos que lhe possibilita relacionar com outros objectos ou literais, através de predicados.

O projecto ExtremeSpots consiste no desenvolvimento de uma aplicação orientada para a pesquisa de pontos de interesse, usando esta abordagem.

2. Arquitetura

A arquitetura por nós desenvolvida, está de forma sucinta no figura a baixo.



A nossa fonte de informação foi o Foursquare, aonde vamos buscar os dados de todos os POIs de Sidney Australia. Desenvolvemos uma ontologia de forma a se adequar aos dados. Em seguida os dados são inseridos no Triplestore Database (TDB) para posteriormente efectuarmos a pesquisa e recomendação semântica e o browsing. Para o browsing usamos a framework Fuseki, que nos permite efectuar queries em SPARQL. Para a pesquisa e recomendação tivemos que criar a nossa própria API. Em seguida iremos aprofundar todos estes tópicos.

2. Escolha da Fonte de Informação

2.1. Escolha da Fonte de informação

Para a recolha de dados que posteriormente são utilizados para popular o triple store foi-nos permitido comparar as duas APIs, tendo em vista os dados que poderíamos retirar de ambas. A API da google oferecia mais pontos de interesse. Por outro lado, a API do foursquare disponibiliza Venues (locais) com descrição do local, o que nos permite melhorar a pesquisa semântica. O foursquare, devido à forma como os locais estão organizados (em árvore), também nos permitia criar uma ontologia mais rica/mais ramificada.

Assim decidimos implementar a API do foursquare.

2.2. Desafios e Soluções

Em ambas as APIs experimentadas existe um limite de requests que pode ser feito por minuto. No caso do Foursquare, este limite é de 80 requests por minuto.

Outra limitação é o número de elementos da resposta é limitado, por isso foi adoptado o seguinte procedimento:

Recolha de elementos com uma pesquisa num raio muito pequeno. Variar o ponto central de maneira a tentar obter o maior número de elementos. O elemento devolvido neste tipo de request é reduzido, por isso é necessário criar outra thread para a obtenção dos dados completos. Os elementos são guardados numa base de dados No-SQL.

Obtenção dos dados completos. É realizada uma query à API com o identificador do ponto de interesse. Com a resposta do Foursquare são substituídos e adicionados campos ao elemento correspondente.

No total foram recolhidos 15.000 pontos de interesse.

3. Ontologia

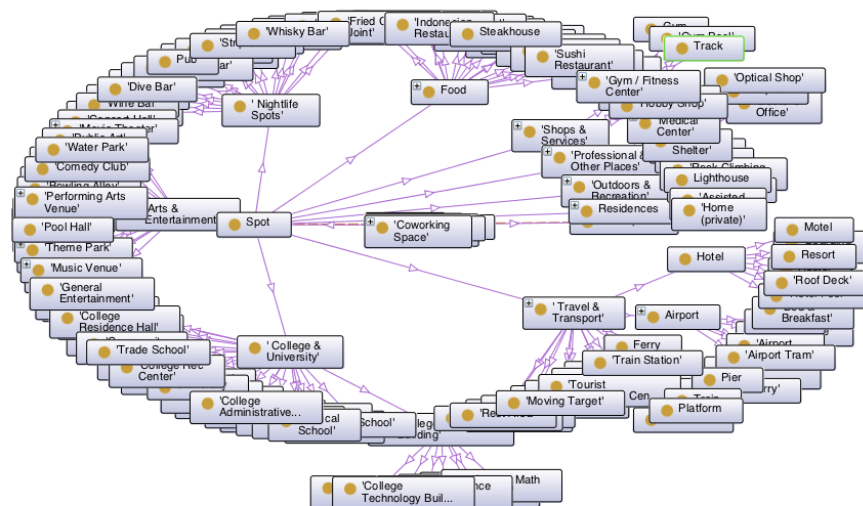
3.1. Criação da Ontologia

O principal objectivo da ontologia é definir um vocabulário para um determinado domínio. No caso prático do ExtremeSpots é possível estabelecer triplos que são pertencentes ao domínio dos pontos de interesse. Com base na estrutura dos pontos de interesse (venues) e da camada de categorias destes no serviço foursquare, foi nos permitido criar uma estrutura em árvore, que pode ser observada na figura do ponto 2.

Existem duas classes distintas, a class spot, que engloba todas as classes de pontos de interesse, e a classe description. Os pontos de interesse podem ser caracterizados pela sua posição geográfica, preço, rating, contacto e morada. Estas são as suas propriedades gerais. Uma classe do tipo spot pode vir a assumir muitas facetas como bares, residências, restaurantes ou piscinas.

Para a criação da ontologia foi utilizado o Protege 4.3. Uma aplicação muito simples que nos permitiu criar a ontologia com recurso a uma interface gráfica. Apesar de ser uma boa ferramenta, sentimos que ainda está muito incompleta, dificultando imenso a nossa tarefa.

3.2 Estrutura da ontologia.



Como podemos observar existem duas classes “mães”, A “Spot” e a “Description”. Da primeira derivam 9 classes, dessas novas ainda mais se multiplicam, tendo a nossa ontologia a ter 4 níveis de ramificação.

3.3. Propriedades

A classe “Spot” têm 8 propriedades:

- name
- contact
- id
- ltg (latitude)
- lng (longitude)
- price
- rating

Criamos também uma ”object propertie” para relacionar a “Description” ao “Spot”

5. Triplestore

Nesta fase já tínhamos a ontologia criada e os dados do foursquare todos armazenados.

Criamos os triplos criamos com recurso JENA, através de uma aplicação em java por nós criada, que transforma os dados da base de dados em RDFs, e por fim inseri-los no Triplestore, para guardarmos os dados de forma permanente.

6. SPARQL

Após termos os dados todos guardados na forma de triplos, necessitamos de fazer pesquisas á nossa base de dados. Mas uma vez que estávamos a desenvolver uma aplicação para iOS, necessitávamos também de uma framework que nos permitisse fazer estes pedidos através de um “request”.

Descobrimos assim o Fuseki. Está framework corre em cima do apache, que nos permitia ter uma pagina web aonde poderíamos fazer os nossas “queries” em SPARQL.

Começamos depois a tentar perceber a forma como é que este encapsulava os pedidos no url, até que conseguimos implementa-lo no iOS.

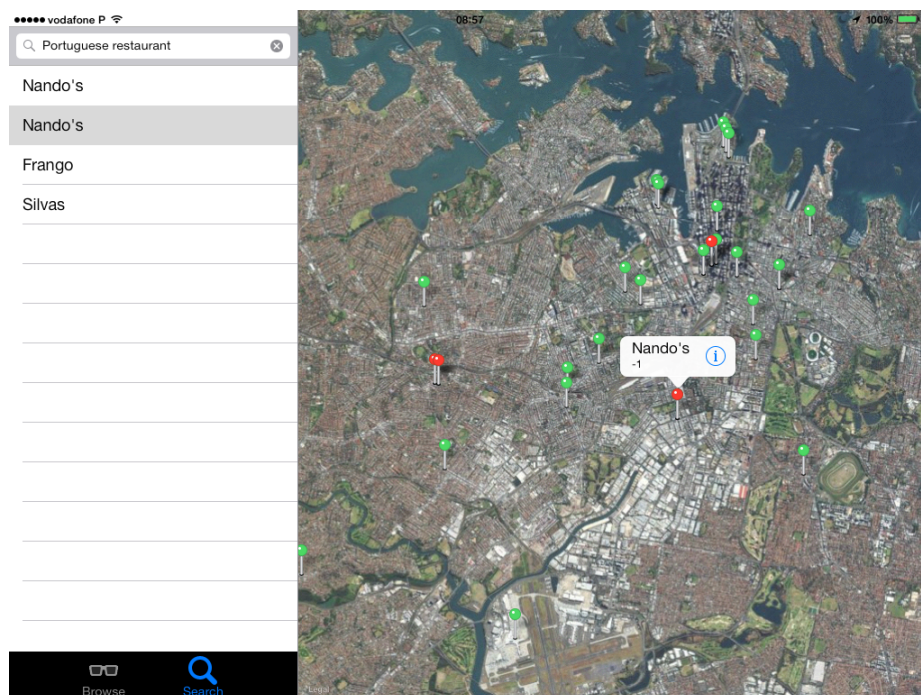
Está ferramenta foi extremamente util nesta fase, pois permitiu-nos poupar imenso tempo ao evitar que criássemos a nossa API para este efeito.

7. Recomendação e Pesquisa Semântica

Esta foi a parte que se tornou mais complexa.

Quando introduzimos uma query, vamos tentar identificar quais são as classes, as instancias e as propriedades. Após esta identificação fazemos o nosso json request através da playframework,

Após isto, baseado nas classes identificadas, vamos sugerir outras sitios que o utilizador poderá estar interessado, por exemplo, caso pesquisemos “portuguese restaurants” ele devolve-nos vários outros restaurantes. Como podemos ver na figura a baixo.



8. Conclusões

Este trabalho, apesar de árduo foi muito interessante do nosso ponto de vista. Ensinou-nos novas maneira de guardar e relacionar dados. Mostrou-nos as potencialidades da web semântica e como por ela poderá passar o futuro da web.

Em relação ao trabalho, apesar de ser um trabalho para a cadeira de web semântica pensamos que o resultado final foi bastante positivo, conseguindo realizar, os pressupostos iniciais.