

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



FEUP

Plataforma Web para Disponibilização de Serviços Geoespaciais

Soraia Filipa Tavares Ferreira

Mestrado Integrado em Engenharia Informática e Computação

Orientador: António Fernando Vasconcelos Cunha Castro Coelho (PhD)

Julho de 2012

Plataforma Web para Disponibilização de Serviços Geoespaciais

Soraia Filipa Tavares Ferreira

Mestrado Integrado em Engenharia Informática e Computação

Aprovado em provas públicas pelo Júri:

Presidente: Jorge Manuel Gomes Barbosa (PhD)

Vogal Externo: Paulo Miguel de Jesus Dias (PhD)

Orientador: António Fernando Vasconcelos Cunha Castro Coelho (PhD)

13 de Julho de 2012

Resumo

Com o aumento de recursos partilhados na Web, o número de utilizadores dos SIG Web tem vindo igualmente a aumentar, existindo cada vez mais necessidades para as suas aplicações. Os geoportais um exemplo típico de uma aplicação SIG Web, funcionando como pontos de acesso a coleções de informação e serviços georreferenciados. Através deles é possível usufruir de um ambiente Web que permite agregar e partilhar conteúdo e criar entendimento. Mas tipicamente os geoportais não têm em consideração a semântica das palavras, o contexto, as ou ambiguidades da língua natural. Uma pesquisa com base em palavras-chave normalmente não é suficiente para satisfazer as necessidades do utilizador. Após a análise de alguns casos de estudo, é possível constatar que a interface gráfica de um geoportal não reflete de forma intuitiva (ou não apresenta) a semântica dos conteúdos. Por estas razões, o potencial de exploração da interface gráfica com o utilizador é elevado, carecendo de soluções a vários níveis.

Contudo, a atual organização da informação geográfica em catálogos de metadados é incapaz de capturar a semântica dos dados. Como tal, é proposto o conceito de mapas semanticamente enriquecidos que se baseiam no princípio da Web Semântica, assumindo assim que grande parte da informação publicada na Web tende a ser cada vez mais descrita por modelos enriquecidos a nível semântico. Através da metodologia proposta, este conhecimento pode ser explorado na apresentação de informação sobre mapas.

Esta dissertação tem como principal objetivo a conceção de uma plataforma Web para a disponibilização de serviços geoespaciais, com interface gráfica intuitiva baseada em ontologias e facetas que explore ao máximo a componente semântica da informação.

Para tal, foi desenvolvida a metodologia para geoportais semanticamente enriquecidos (baseada na ontologia *SMaps*). São também propostas duas abordagens semanticamente enriquecidas à interface com o utilizador e foi implementado um protótipo para teste e validação da metodologia.

Abstract

With the increase of shared resources on the Web, the number of Web GIS users has also been increasing and becoming more demanding. Geoportals are a typical example of a Web GIS application, which act as access points to collections of georeferenced information and services, featuring the aggregation and sharing of knowledge. Typically, geoportals don't take into account the semantics of terms, the context of data, the ambiguities of natural language, and cultural or linguistic differences. In this sense, queries based solely on keyword analysis are insufficient to cope with the needs of the user.

After the analysis of several case studies, it is possible to conclude that the semantics of the contents presented in the graphical user interface (GUI) of a geoportal is generally missing or not intuitively interpreted by the user. The potential of GUI exploration is very interesting and still lacks solutions to several aspects.

Since the common organization of georeferenced information in catalogs of metadata is unable to capture the semantics of the data, a methodology for deploying semantically enriched geospatial Web services is proposed. The methodology follows the principles of Semantic Web, assuming that a large part of the information published on the Web tends to be more and more described by semantically enriched models. With the proposed methodology, the semantics present in this data can be explored for the visualization of information on maps.

The main goal of this dissertation is the conception of a Web platform for deployment of geospatial services, with a semantically enriched graphical user interface based on ontologies and facets. To achieve this goal, a methodology for semantic enriched geoportals, based on the *SMaps* ontology, was conceived. Additionally, a prototype following this methodology was implemented.

Agradecimentos

Um agradecimento distinto ao meu orientador, Prof. António Coelho, pela sua dedicação e prontidão durante todo o período da dissertação e à nobre instituição que a acolheu, a Faculdade de Engenharia da Universidade do Porto.

Gostava também de agradecer ao meu namorado, Nuno Luz, por manter a minha sanidade mental nos momentos mais críticos e por assumir a pesada tarefa de rever esta dissertação; ao meu pai António Ferreira, à minha mãe Rosa Tavares, pelo seu apoio incondicional.

E por fim mas não por último, gostava de agradecer a todos os meus amigos pelos bons momentos longe do trabalho, especialmente à Sónia Roque por todos os anos de amizade.

Soraia Filipa Tavares Ferreira

Índice

Introdução.....	16
1.1 Enquadramento e Motivações.....	17
1.2 Problema	18
1.3 Objetivos.....	19
1.4 Organização do Relatório	20
SIG Web e Geoportais	21
2.1 Definição de SIG Web.....	22
2.2 Geoportais.....	22
2.3 Exemplos de geoportais	23
2.3.1 MapQuest.....	23
2.3.2 INSPIRE Geoportal.....	25
2.3.3 DouroValley.....	27
2.3.4 NYC Pictures	28
2.4 Resumo	30
Ontologias e Web Semântica.....	31
3.1 A Web Semântica	31
3.2 Ontologias.....	33
3.2.1 Resource Description Framework, RDF Schema e SPARQL Protocol.....	35
3.2.2 Web Ontology Language	36
3.3 Pesquisa por Facetas	38
Metodologia para Geoportais Semanticamente Enriquecidos.....	41
4.1 Ontologia Semantic Maps.....	42
4.2 Ontologias e Informação Geoespacial	44
4.3 Interface com o Utilizador	45
4.3.1 Apresentação Externa de Conteúdos.....	46
4.3.2 Apresentação Interna de Conteúdos.....	49
4.4 Framework <i>SMaps</i>	51
Implementação e Validação.....	53

5.1	Desafios de Implementação	53
5.2	Protótipo Implementado	55
Conclusões e Trabalho Futuro		61
6.1	Conclusões	61
6.2	Trabalho Futuro	62
Referências.....		63
Manchester Syntax da Ontologia <i>SMaps</i>		66

Lista de Figuras

Figura 1 – Geoportál SIGAVE.	18
Figura 2 – Interface Gráfica do MapQuest.	24
Figura 3 – Interface Gráfica do INSPIRE Geoportál.	25
Figura 4 – Interface Gráfica de filtragem de informação.	26
Figura 5 – Interface Gráfica do DouroValley.	27
Figura 6 – Vista Detalhada de um ponto de interesse.	28
Figura 7 – Interface gráfica do NYC Pictures.	29
Figura 8 – Vista detalhada de uma imagem.	30
Figura 9 – <i>Continuum</i> Semântico (Uschold, 2003).	32
Figura 10 – Pilha de tecnologias da Web Semântica.	32
Figura 11 – Conjunto parcial de ontologias extensíveis e interligadas atualmente na Web Semântica.	35
Figura 12 – Principais entidades e relações a ontologia <i>SMaps</i> .	42
Figura 13 – Padrão de representação de listas em OWL (Drummond, Rector, & Stevens, 2006a)(Drummond, Rector & Stevens 2006b).	44
Figura 14 – Ilustração parcial da ontologia <i>Rest</i> .	45
Figura 15 – Apresentação externa de conteúdos de acordo com a ontologia <i>Rest</i> .	47
Figura 16 – Pesquisa por facetas de acordo com a taxonomia de restaurantes da ontologia <i>Rest</i> .	48
Figura 17 – Apresentação interna de conteúdos de acordo com a ontologia <i>Rest</i> .	50
Figura 18 – Arquitetura da Framework <i>SMaps</i> .	51
Figura 19 – Arquitetura do protótipo implementado.	55
Figura 20 – Diagrama UML parcial do <i>OpenLayers</i> após a implementação do protótipo <i>SMaps</i> .	56
Figura 21 – Interface do protótipo: ecrã onde todos os objetos com caracterização são apresentados.	57
Figura 22 – Interface do protótipo: ecrã com filtragem de objetos por hotéis do Porto.	58
Figura 23 – Interface do protótipo implementado. Filtragem de objetos por hotéis do Porto com 4 ou mais estrelas, juntamente com apresentação dos restaurantes associados ao hotel selecionado.	59

Abreviaturas e Símbolos

AMAVE	Associação de Municípios do Vale do Ave
API	Application Programming Interface
GUI	Graphical User Interface
HTML	HyperText Markup Language
HTTP	Hypertext Transfer Protocol
JSON	JavaScript Object Notation
OGC	Open Geospatial Consortium
<i>Open Source</i>	Código Aberto
OWL	Ontology Web Language
RDF	Resource Description Framework
RDF-S	Resource Description Framework Schema
SGBD	Sistema Gestor de Base de Dados
SIG	Sistema de Informação Geográfica
SPARQL	SPARQL Protocol and RDF Query Language
SVG	Scalable Vector Graphics
URI	Uniform resource identifier
USGS	United States Geological Survey
W3C	World Wide Web Consortium
XML	Extensible Markup Language

Capítulo 1

Introdução

Os produtos de base geográfica e de gestão de bases de dados espaciais desempenham uma influência consagrada nos focos de decisão da Administração pública e das políticas do território, ao nível local, regional e central (Chapin, 2003). Existe a necessidade explícita de criar, distribuir e utilizar informação geográfica via Web de forma fiável e atual, para que todo o trabalho colaborativo entre órgãos ao nível local, regional e central seja integrado e proveitoso.

Paralelamente, a procura e a utilização deste tipo de informação estão a mudar, sendo cada vez mais as necessidades dos utilizadores e das aplicações. Como consequência, uma orientação de catálogo e de preenchimento de metadados, que reduz por vezes erradamente, um Sistema de Informação Geográfica (SIG) a um simples visualizador de conteúdos deixou de ser suficiente.

Novas maneiras interativas de usar mapas Web e um conjunto de novos serviços que suportem essa interatividade são imprescindíveis de forma a dar resposta às necessidades emergentes relacionadas com a tomada de decisão no âmbito das políticas de território.

Os SIG, anteriormente definidos por mapas estáticos locais, com baixo nível de interatividade e com origens proprietárias (SIG pagos), definem-se atualmente cada vez mais através de mapas dinâmicos, muito interativos, distribuídos via Web com integração de soluções *Open Source* e de empresas privadas (SIG livres). Considerando o estado atual de desenvolvimento das ferramentas *Open Source* é possível obter funcionalidades e resultados equivalentes aos de uma solução constituída por *software* proprietário (Song et al. 2009) (Deshpande & Riehle 2008) (Steiniger & Hunter 2012).

Tendo como base estas questões, segue-se uma descrição dos motivos que levaram à realização deste trabalho (secção 1.1), do problema (secção 1.2) e dos objetivos (secção 1.3). A última secção (secção 1.4) deste capítulo descreve a organização do relatório.

1.1 Enquadramento e Motivações

A missão das entidades municipais é disponibilizar os serviços necessários para assegurar a saúde, segurança, e bem-estar dos seus residentes. Um SIG é uma poderosa ferramenta para suceder nesta missão, pois consegue revelar informação importante para a otimização do processo de tomada de decisão. Por exemplo, os registos sobre agricultura podem indicar quanto pesticida foi aplicado a uma parcela de terra e intersetando esses registos com os registos de hidrografia é possível prever os efeitos nefastos dos pesticidas nas águas da região.

Nas diversas áreas de aplicação de um SIG, está presente a Gestão de Planos Municipais de Ordenamento de Território e esta é uma área que beneficia, em grande parte, das funcionalidades que um SIG tem para oferecer (por exemplo, na análise de vizinhança e/ou sobreposição de camadas). Com os SIG é possível selecionar/inquirir e visualizar/interpretar dados de um modo fácil e consensualmente compreendido, identificando relações e padrões espaciais que podem ajudar a resolver problemas, contribuindo sem margem de dúvidas para a qualidade e eficácia na tomada de decisão (Oliveira et al., 2012).

Todas estas mais-valias, entre outras, potenciam o forte investimento por parte de entidades públicas na conceção e manutenção de SIG (como é o caso do portal SIGAVE), tanto a nível regional como a nível municipal (Sieber, 2008).

No entanto, existem ainda muitos desafios adversos presentes na implementação de SIG. Entre estes desafios encontram-se (Nikolaos, Kostas, & Michail, 2005):

- A morosidade dos processos manuais/semiautomáticos de configuração dos serviços;
- A integração de diferentes fontes de informação, diversos formatos e terminologia para representar conceitos geoespaciais;
- A filtragem incremental e dinâmica da informação apresentada;
- A falta de semântica nas fontes de informação e na apresentação da informação sobre os mapas.

Por conseguinte, existe a necessidade de criar uma solução integrada com componentes base genéricos que permitam: i) acelerar todo o processo em causa, ii) melhorar a gestão de recursos humanos/temporais e iii) potenciar melhores resultados e o desenvolvimento de novas funcionalidades de valor acrescentado para o utilizador.

Nos últimos anos, tem-se verificado um crescente esforço na resolução dos problemas inerentes aos três primeiros desafios previamente listados (Alameh, 2003) (Jones & Purves, 2007) (OGC, 2012). Mais concretamente, a configuração dos serviços SIG é cada vez mais simples e automatizada, existindo também diversos *standards* para a troca de informação e ferramentas para a sua integração. Relativamente à filtragem incremental e dinâmica da

informação apresentada, a pesquisa por facetas¹ tem-se revelado uma solução viável particularmente na área dos SIG Web (Gaetan, Saldaño, Buccella, & Cechich, 2010).

Apesar de tudo, os SIG atuais ainda não estão preparados para lidar com a componente semântica da informação, que geralmente é facilmente compreendida por humanos mas não pelas máquinas. No caso dos SIG, a apresentação gráfica da informação sobre os mapas acaba muitas vezes por ofuscar esta semântica, limitando a potencialidade de análise. As ontologias e a Web Semântica têm tido um forte impacto na investigação realizada no âmbito dos SIG Web nos últimos cinco anos e proporcionam as metodologias e tecnologias adequadas para a criação desta componente semântica.

1.2 Problema

O Instituto de Engenharia de Sistemas e Computadores do Porto (INESC Porto) - associação privada sem fins lucrativos, reconhecida como instituição de utilidade pública, tem vindo a participar e a desenvolver ativamente vários projetos no âmbito dos SIG, como é o caso do portal centrado na promoção turística da região do Douro – DouroValley, ou do portal da região e municípios do Vale do Ave – SIGAVE, com enfoque no apoio às atividades de gestão e desenvolvimento do território da região do Vale do Ave.

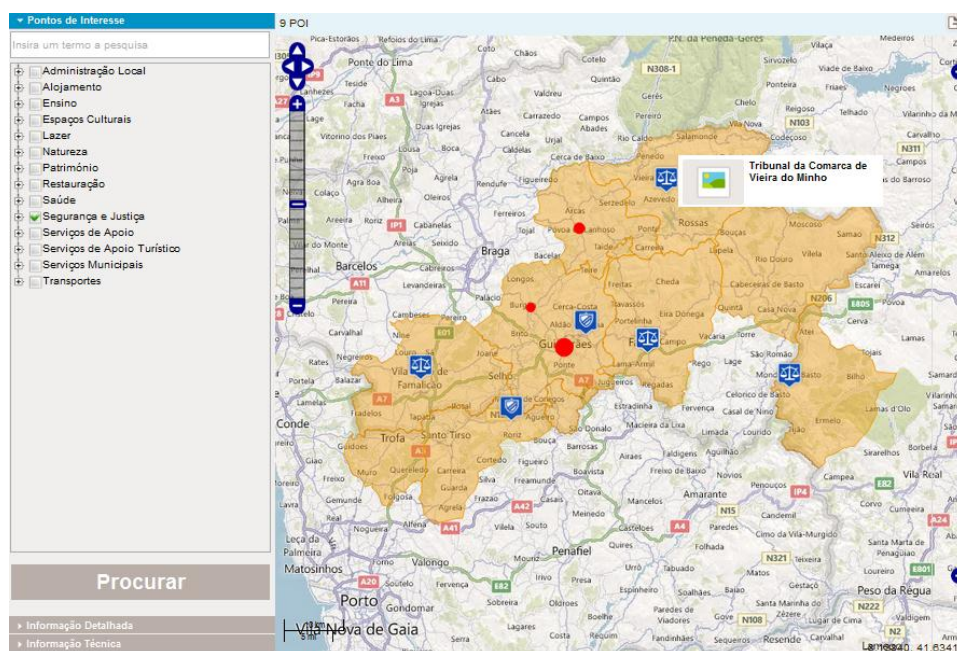


Figura 1 – Geoportal SIGAVE.

¹¹ Do anglo-saxónico *faceted search*.

O SIGAVE foi implementado em 2009 para a gestão centralizada da infraestrutura de informação geográfica (com dimensão regional) e para o suporte de atividades de gestão e desenvolvimento do território da região do Vale do Ave.

No seu estado atual, a plataforma integrada no Portal Regional do Ave, disponibiliza um conjunto muito limitado de funcionalidades geoespaciais interativas (como a maioria dos SIG disponibilizados na Web), estando também em falta a comunicação dinâmica e bidirecional, entre a Associação de Municípios do Vale do Ave (AMAVE), órgão central e os municípios associados.

Os dois maiores desafios científicos prendem-se com a questão da interoperabilidade entre sistemas (comunicação dinâmica e bidirecional) e a interação espacial sobre o mapa (aumento da oferta de serviços). Em particular, a interação espacial sobre o mapa é uma área que carece de novas soluções inovadoras (Turner & Forrest 2008), e que pode beneficiar largamente da componente semântica da informação.

1.3 Objetivos

Esta dissertação tem como principal objetivo a conceção de uma plataforma Web para a disponibilização de serviços geoespaciais, com interface gráfica intuitiva baseada em ontologias e facetas que explore ao máximo a componente semântica da informação. A plataforma deve possuir uma arquitetura genérica e flexível de forma a ser reutilizada no desenvolvimento de futuros SIG e baseada em ferramentas *Open Source*.

Neste sentido, e tendo em conta todos os desafios e necessidades previamente descritas, os objetivos são:

- A conceção de uma arquitetura de um SIG Web flexível, orientada a tecnologias *Open Source* e aos *standards* da Web Semântica, assegurando a interoperabilidade;
- A análise do impacto da componente semântica da informação na sua visualização sobre um mapa;
- A especificação de objetos gráficos extensíveis, e respetivo comportamento (interação), que poderão ser associados a diferentes conceitos presentes num modelo de dados ou ontologia;
- A implementação de um protótipo, baseado no caso de estudo SIGAVE.

1.4 Organização do Relatório

O presente documento divide-se em seis capítulos: Introdução, SIG Web e Geoportais, Ontologias e Web Semântica, Metodologia para Geoportais Semanticamente Enriquecidos, Implementação e Validação, e Conclusões e Trabalho Futuro.

No primeiro capítulo de introdução foi apresentado o tema do projeto, referindo o seu enquadramento e motivações, problema, objetivos e a devida organização do relatório.

No segundo capítulo é apresentado o estado da arte referente aos SIG Web e Geoportais. É apresentada a definição de SIG Web e de geoportal, bem como a importância do segundo para a área de SIG Web. São também analisados alguns casos de estudo de geoportais considerados relevantes e é feita uma análise das suas características.

No terceiro capítulo é apresentado o estado da arte referente às Ontologias e Web Semântica. As ontologias são descritas dentro do contexto da Web Semântica e também são enunciadas algumas linguagens utilizadas para a sua especificação.

O quarto capítulo expõe a solução proposta. Descreve a ontologia (*SMaps*) criada para o propósito, a relação das ontologias e da informação espacial, a pesquisa por facetas, a interface (externa e interna) com o utilizador e a arquitetura da *framework* desenvolvida com base na metodologia concebida para a implementação de geoportais semanticamente enriquecidos.

O quinto capítulo descreve a implementação da metodologia mencionada no capítulo anterior e refere quais os principais desafios em termos de tecnologias utilizadas, o protótipo implementado e o teste e validação de resultados.

Finalmente, o sexto capítulo apresenta as conclusões sobre o trabalho realizado e ideias que podem ser implementadas futuramente.

Capítulo 2

SIG Web e Geoportais

O advento da Internet e da Web ajudou a pavimentar a autoestrada da informação, permitindo uma sociedade baseada em informação e mudando a maneira como vivemos e trabalhamos. Da emergência e evolução da Internet e da Web e a fusão dos SIG, surge o conceito de SIG Web que veio mudar a maneira como a informação geoespacial é adquirida, transmitida, publicada, partilhada e visualizada.

Os SIG tornaram-se aplicações distribuídas via Internet que motivaram, tanto o setor privado como o público, a tirar partido dos seus benefícios na Web (Tait, 2005). Localizar amigos, detetar pontos de acesso Wi-Fi, colocar etiquetas em fotos pessoais ou marcar localizações de novos eventos sociais são todos exemplos de aplicação de SIG Web utilizados atualmente.

Um SIG é um sistema de *hardware*, *software*, e um conjunto de procedimentos que captura, armazena, edita, manipula, gere, analisa, partilha, e visualiza informação georreferenciada (Bhat, Shah, & Ahmad, 2011). A tecnologia SIG já existia antes do aparecimento da Internet e da Web, e as suas capacidades e complexidade vão muito para além do simples mapeamento. Oferece um conjunto rico de funções analíticas que podem revelar relações escondidas, padrões, e tendências que não são aparentes, permitindo às pessoas pensar espacialmente para resolver problemas e tomar decisões inteligentes. Durante décadas, os profissionais de SIG usaram esta tecnologia para integrar, analisar, e visualizar informação geográfica e conhecimento, levando a um abundante número de aplicações SIG beneficiando inúmeras áreas de conhecimento. Mas o acesso aos SIG tem sido limitado a um número relativamente pequeno de profissionais e só agora com a emergência dos SIG Web o seu poder está a ser desbloqueado para uma audiência mais ampla. Os SIG Web não são apenas mais acessíveis, mas também mais flexíveis através da sua implementação utilizando APIs Web que facilitam a integração com outros sistemas de informação.

2.1 Definição de SIG Web

Os SIG Web têm evoluído rapidamente desde 1993 e são na sua essência sistemas de informação distribuídos. A sua natureza distribuída tem grandes implicações na sua habilidade de usar, gerir e partilhar dados georreferenciados. Estas implicações criam barreiras de partilha de dados que impedem que múltiplos SIGs cooperarem entre si.

Um SIG Web é qualquer SIG que use tecnologias Web para comunicar entre os seus componentes (*hardware, software*, dados e utilizadores), envolvendo no mínimo duas camadas: servidor e cliente. Todos os componentes podem estar no mesmo computador, mas são na realidade componentes separados e estão geralmente distribuídos na Web.

Com a partilha de recursos na Web, os SIG tradicionais expandiram as suas funcionalidades, tirando partido p. ex., de mapas base servidos pela *United States Geological Survey* (USGS) ou pela *Google*, e permitindo que utilizadores de um lado do globo acedam a um servidor localizado no outro lado do globo.

Esta natureza global dos SIG Web aumenta o seu número de utilizadores (muitas vezes simultâneos) face ao SIG tradicional, normalmente utilizado apenas por um utilizador de cada vez, mas também requer maior desempenho e escalabilidade.

Como a maioria dos “clientes” de SIG Web são *browsers*, baseados em *Javascript* e HTML, o suporte de diferentes sistemas operativos/plataformas é garantido. Ao contrário dos SIG tradicionais, destinados a utilizadores profissionais com anos de treino, os SIG Web destinam-se a uma audiência mais ampla (que pode não ter formação em SIG) que espera que a interface seja fácil de utilizar, à semelhança de uma página Web regular (simples e intuitiva) (Carver, Evans, & Kingston, 2000). Com o aumento de utilizadores dos SIG Web, existem cada vez mais necessidades e portanto multiplicam-se as suas aplicações (como p. ex., mostrar a taxa de natalidade em cada região).

Um caso concreto de um tipo de aplicação SIG Web são os geoportais, que funcionam como pontos de acesso a coleções de informação e serviços. Através deles é possível usufruir de um ambiente Web que permite agregar e partilhar conteúdo e criar entendimento (Maguire & Longley, 2005).

2.2 Geoportais

Um portal Web, ou simplesmente portal, é uma página Web que funciona como um ponto de acesso à Internet, fornecendo conteúdo útil e diversas hiperligações para outras páginas Web². O seu grande sucesso está correlacionado com o aumento de conteúdo disponibilizado na

² <http://dictionary.reference.com>

Web, doutra forma inacessível ao utilizador devido ao crescimento exponencial do volume de informação.

Um geoportal é um portal que fornece um único ponto de acesso à informação geoespacial, a serviços Web, e a outros recursos geoespaciais relacionados (Tait, 2005). Os utilizadores podem navegar e pesquisar através de geoportais para encontrar recursos geoespaciais relevantes e avaliar a qualidade dos mesmos.

Os geoportais são diferentes dos portais convencionais, pelo facto de se especializarem na descoberta de conteúdo geoespacial. A informação geográfica é de vital relevância na tomada de boas decisões ao nível local, regional, central e global. É crucial por abordar problemas sociais, económicos e ambientais de grande importância. Os geoportais facilitam a partilha de informação, eliminam a coleta de dados duplicados e suportam uma tomada de decisão mais racional. À medida que o volume de conteúdo geoespacial aumenta, também a importância dos geoportais aumentará. Mas o seu verdadeiro potencial, de expandir o uso de informação geográfica, verifica-se através do forte investimento governamental e da participação pública (Sieber, 2008).

Entre os muitos novos desafios do desenvolvimento de geoportais, está o suporte à pesquisa semântica e multilíngue, que tornam mais fácil encontrar recursos adequados às pesquisas. A pesquisa semântica refere-se à pesquisa baseada no significado das palavras em vez da ortografia e a pesquisa multilíngue permite que seja feita uma correspondência do mesmo conceito entre várias línguas. Tipicamente os geoportais não têm em consideração a significação, o contexto, as ambiguidades da língua natural e as diferenças entre várias línguas. Uma pesquisa apenas por palavras-chave, sem a consideração da semântica e da língua, normalmente não é suficiente para satisfazer as necessidades do utilizador.

2.3 Exemplos de geoportais

Recentemente, os geoportais regionais ou temáticos têm proliferado. Apesar da elevada oferta, foram apenas aqui considerados os mais relevantes no contexto desta dissertação. O método usado é sobretudo empírico, devido ao facto de a maior parte da infraestrutura física e lógica dos geoportais não se encontrar pública.

2.3.1 MapQuest

A MapQuest é uma aplicação pioneira no que diz respeito a mapas interativos na Web. Criada em 1994 e adquirida pela AOL em 2000, foi lançada em 1996 na Internet, mudando a maneira como as pessoas obtêm direções de trânsito para sempre. Atualmente é uma aplicação bastante popular que disponibiliza mapas, imagens de satélite, informação de condução, pontos

de interesse e rotas. A empresa MapQuest, Inc³ fornece mapas e direções com grande precisão, a milhões de pessoas em vários países, e está empenhada em tornar fácil a pesquisa de informação local.

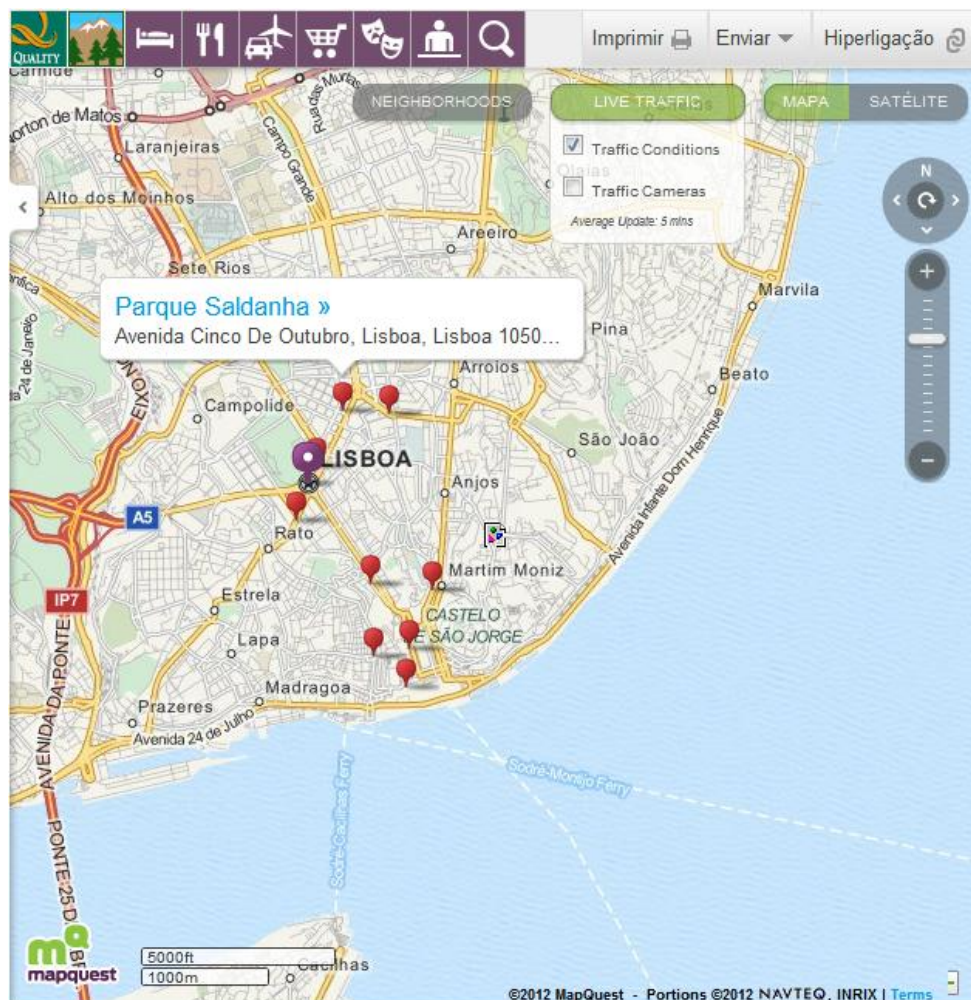


Figura 2 – Interface Gráfica do MapQuest.

Em 2010 a aplicação foi lançada com um nova interface gráfica simplificada, e foi anunciado que todos os dados seriam provenientes do projeto *OpenStreetMap*⁴ (projeto colaborativo para criar um mapa livre e editável do mundo).

Entre os seus pontos fortes estão a indicação de direções (de condução) precisas e a localização de postos de abastecimento próximos com o preço mais barato, sendo que o segundo apenas se encontra disponível nos Estados Unidos da América.

³ <http://www.mapquest.com/>

⁴ <http://www.openstreetmap.org/>

Possui a grande desvantagem de não conseguir lidar bem com moradas ou nomes de negócio incompletos (ao contrário de aplicações como o *Google Maps*⁵ que conseguem interpretar facilmente informação incompleta), não retornando qualquer resultado.

A MapQuest conseguiu capitalizar o interesse da comunidade científica para localizar e mapear locais na Web. À medida que a Web foi amadurecendo, muitas outras organizações começaram a considerar o uso de informação geográfica nas suas páginas Web.

2.3.2 INSPIRE Geoportal

A diretiva INSPIRE⁶ visa a criação de uma infraestrutura de dados espaciais europeia, de modo a permitir a partilha de informação geoespacial de cariz ambiental entre organizações públicas. De forma semelhante, visa facilitar o acesso dessa mesma informação ao público em geral. É uma iniciativa bastante importante no que diz respeito à criação de novas políticas de fronteiras, e portanto inclui informação geoespacial bastante extensiva.

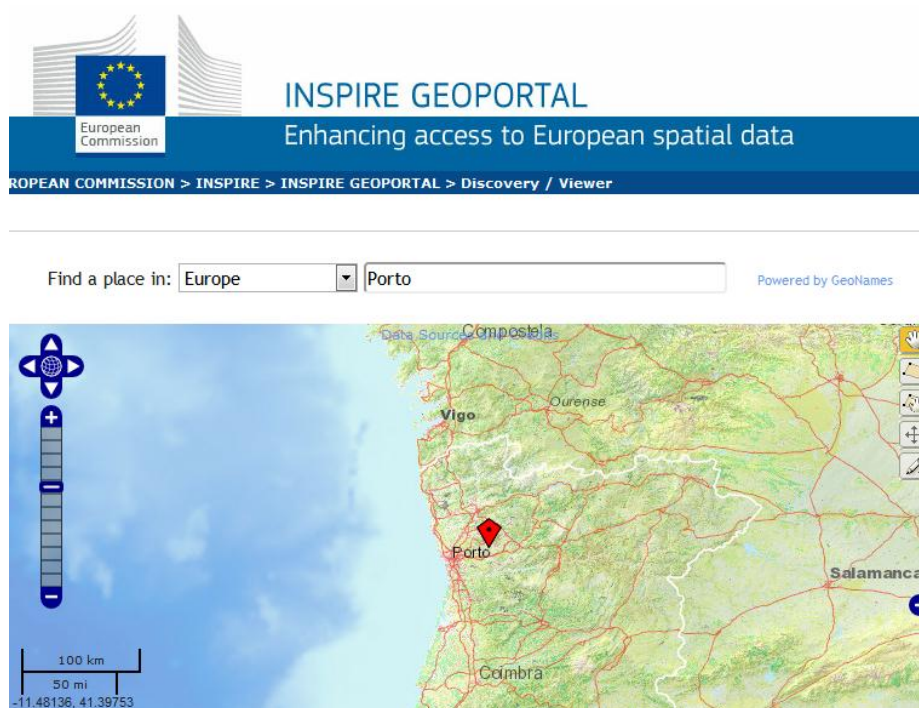


Figura 3 – Interface Gráfica do INSPIRE Geoportal.

⁵ <https://maps.Google.com/>

⁶ <http://inspire.ec.europa.eu/index.cfm>

A criação do INSPIRE Geoportal fornece os meios de pesquisa e os serviços geoespaciais necessários para a visualização de dados geoespaciais de todos os estados-membros. Publicado em 2011 pela primeira vez, disponibiliza serviços de descoberta e visualização de informação geoespacial, mas estão previstas funcionalidades adicionais em lançamentos futuros. Apesar da sua atual orientação de catálogo (de metainformação), disponibiliza vários critérios de filtragem de informação que permitem obter melhores resultados. Principalmente pela pesquisa ter em consideração, além de todos os critérios especificados, a localização geográfica (permite pesquisar dentro duma determinada área, designada *Bounding Box*).

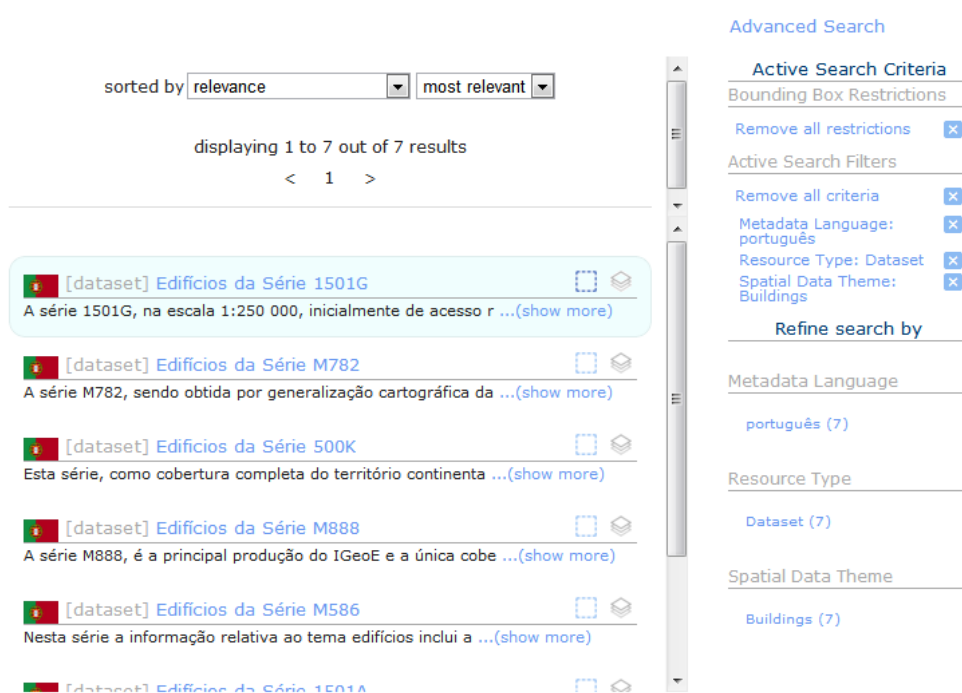


Figura 4 – Interface Gráfica de filtragem de informação.

A descoberta de recursos adequados depende inteiramente da correspondência sintática do texto inserido pelo utilizador, o que torna o processo de recuperação de informação complicado. Sobretudo devido a questões sintáticas, esquemáticas e semânticas (Heydari, Mansourian, & Taleai, 1991). De forma a ultrapassar estes desafios, são cada vez mais utilizados vocabulários estruturados ou *thesaurus* (Fugazza & Luraschi, 2012) e está em curso a construção de uma camada de abstração no topo da infraestrutura INSPIRE. Essa camada visa tornar os recursos INSPIRE acessíveis e detetáveis a partir do paradigma *Linked Open Data* (LOD) que permite interligar e interrogar dados de diferentes fontes, suportando a interligação e pesquisa semântica (através de vocabulários comuns).

2.3.3 DouroValley

O DouroValley⁷ é um geoportal que promove e divulga o turismo na região do Douro. Todo o conteúdo é georreferenciado e organizado em categorias, de acordo com o domínio do turismo (p. ex. Património, Alojamento e Restauração). O DouroValley permite efetuar pesquisas em qualquer parte do mapa, visualizar pontos de interesse, criar itinerários flexíveis e consultar eventos que ocorrem na região. O geoportal tem ainda uma vertente comercial e de intermediação de serviços, incluindo possibilidade de reserva de alojamento. Sem dúvida uma mais-valia para o geoportal.

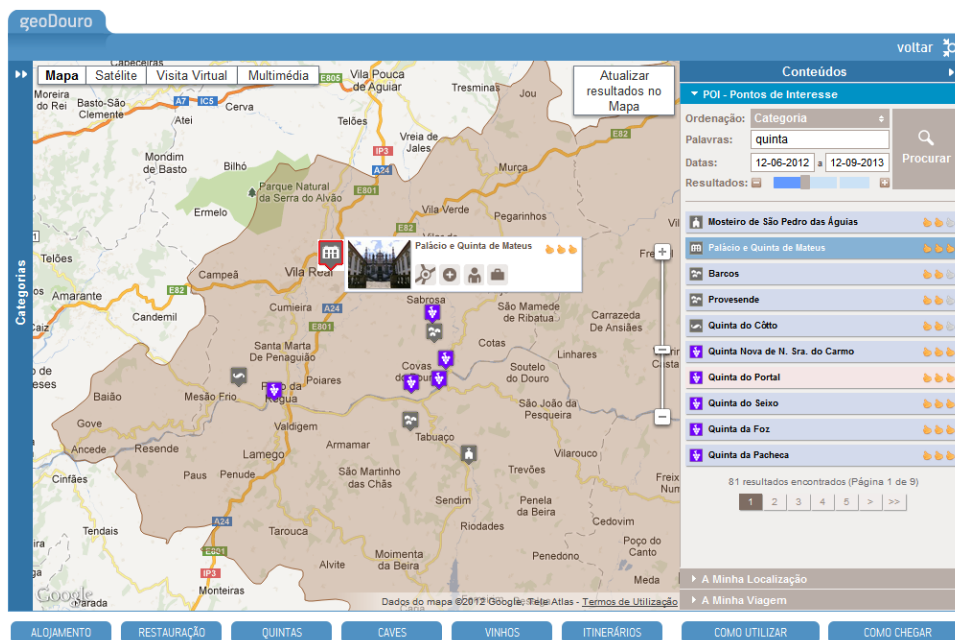


Figura 5 – Interface Gráfica do DouroValley.

No que diz respeito à visualização da informação detalhada de um ponto de interesse em concreto, esta é externa ao mapa e é feita através de um janela que surge ao clicar no ponto. Nessa janela é possível consultar detalhes como uma breve descrição, pontos de interesse associados, a morada, as coordenadas de longitude e latitude, os contatos, etc.

Para consultar os detalhes de um ponto de interesse relacionado, basta clicar no nome do ponto de interesse relacionado e surge uma janela semelhante à que o utilizador se encontra (figura 6) mas com os detalhes do ponto de interesse relacionado.

⁷ <http://dourovalley.eu/>



Figura 6 – Vista Detalhada de um ponto de interesse.

Futuramente seria interessante que a visualização de pontos relacionados fosse feita dentro do mapa, de modo a garantir o enfoque no mapa e a tirar partido da componente geoespacial da informação. Isto evitaria uma quebra no fluxo de navegação realizada pelo utilizador, levando a uma consequente melhoria na descoberta de novo conteúdo com representação geoespacial.

2.3.4 NYC Pictures

A NYC Pictures⁸ é uma página Web que disponibiliza uma coleção de imagens da cidade de Nova Iorque. É possível pesquisar por imagem, tópico ou área (componente geoespacial). O que torna tão interessante esta aplicação é o facto de tirar partido da semântica dos dados. Toda a informação é proveniente de fontes de dados estruturadas como a DBpedia⁹, a Freebase¹⁰, entre outras.

⁸ <http://ny-pictures.com/nyc/photo/>

⁹ http://en.wikipedia.org/wiki/Main_Page

¹⁰ <http://www.freebase.com/>

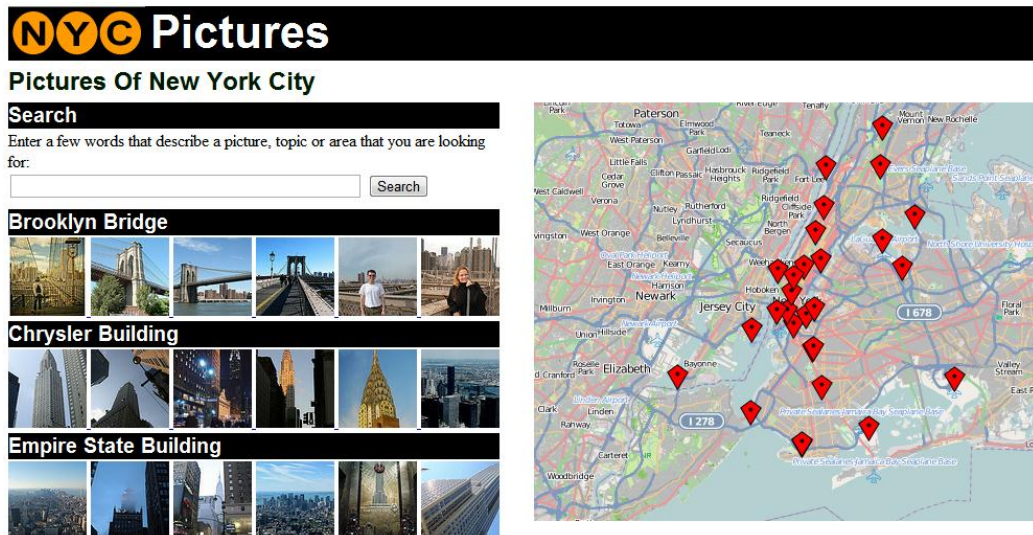


Figura 7 – Interface gráfica do NYC Pictures.

A base de conhecimento da DBpedia, é suportada por uma comunidade que faz um esforço por extrair informação estruturada da Wikipedia¹¹ e disponibilizar esta informação na Web. A DBpedia permite realizar pesquisas sofisticadas à Wikipedia, e permite ligar outros conjuntos de dados (p. ex o flickr¹² ou o Freebase) à Wikipedia. O objetivo consiste em inspirar novos mecanismos de navegação, interligação e melhorar a enciclopédia em si.

Ao pesquisar, p. ex. pela Estátua da Liberdade, é apresentada uma breve descrição textual sobre a mesma (oriunda da Wikipedia), a sua localização geográfica assinalada por um marcador¹³ no mapa (com recurso ao mapa base do OpenStreetMap) e o conjunto de imagens disponíveis sobre aquele tópico. O utilizador pode ainda consultar mais informações noutras fontes de dados (Freebase p. ex.) se assim o entender. Um tópico é entendido como um monumento, uma loja, um hotel, um nome de uma rua ou um lugar de interesse público.

¹¹ http://en.wikipedia.org/wiki/Main_Page

¹² <http://www.flickr.com/>

¹³ Do anglo-saxónico *marker*.

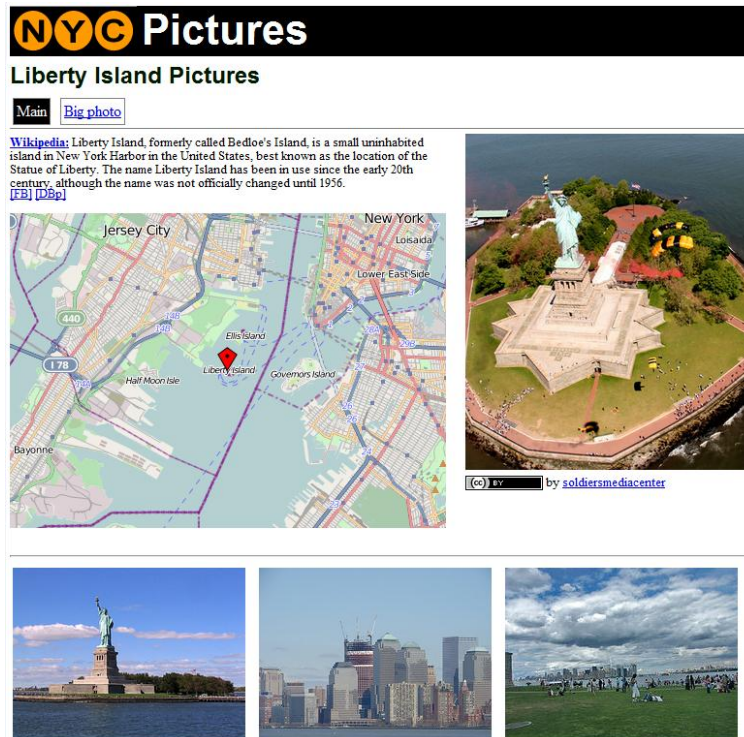


Figura 8 – Vista detalhada de uma imagem.

Apesar de simples na sua interface e sem grandes métodos de navegação, demonstra uma evidente preocupação em resolver problemas como: (i) a falta de interoperabilidade entre diferentes fontes (através da utilização de ontologias que são uma componente chave na abordagem Semântica, pois permitem uma especificação formal de conceitos e relações num determinado domínio, facilitando a interoperabilidade e reutilização entre aplicações SIG e permitem novos mecanismos de análise) e (ii) a falta de qualidade dos resultados obtidos (considerando a semântica de palavras-chave utilizadas para realizar a pesquisa, é possível aprimorar o conjunto de resultados).

2.4 Resumo

Neste capítulo analisaram-se os conceitos sobre SIG Web e Geoportais. Em relação aos SIG, verificou-se que são uma poderosa ferramenta, no processamento de dados georreferenciados e permitem auxiliar na tomada de decisão, em diversos domínios de aplicação. Foi feita também uma revisão sobre a proliferação dos geoportais, tendo sido concluído que são pontos de acesso por excelência, que permitem agregar e partilhar conteúdo e criar entendimento na Web.

Por último, estudaram-se alguns casos práticos relacionados com as temáticas propostas nesta dissertação.

Capítulo 3

Ontologias e Web Semântica

Neste capítulo, são apresentados conceitos considerados relevantes para a contextualização e compreensão das Ontologias e Web Semântica com vista à utilização e exploração dos mesmos em SIG Web.

3.1 A Web Semântica

Um dos principais obstáculos em fornecer melhor apoio aos utilizadores da Web consiste no significado do conteúdo da Web não ser acessível às máquinas (Berners-Lee, 1998). Existem ferramentas que conseguem restaurar textos, dividi-los em partes, verificar a ortografia, ou contar as palavras, mas relativamente a interpretar frases e extrair informação útil para o utilizador, as capacidades do *software* corrente são ainda bastante limitadas. Isto deve-se, principalmente, aos mecanismos de representação de informação *online* terem surgido com uma orientação muito humana (pouco formal) e pouco virada para as máquinas. Na figura 9 encontra-se ilustrado o *continuum* semântico, demonstrando que existem diferentes formas de representação de conhecimento: desde as implícitas às formais. À medida que nos deslocamos no *continuum* (da esquerda para a direita), diminui a ambiguidade e aumenta a robustez e interoperabilidade.

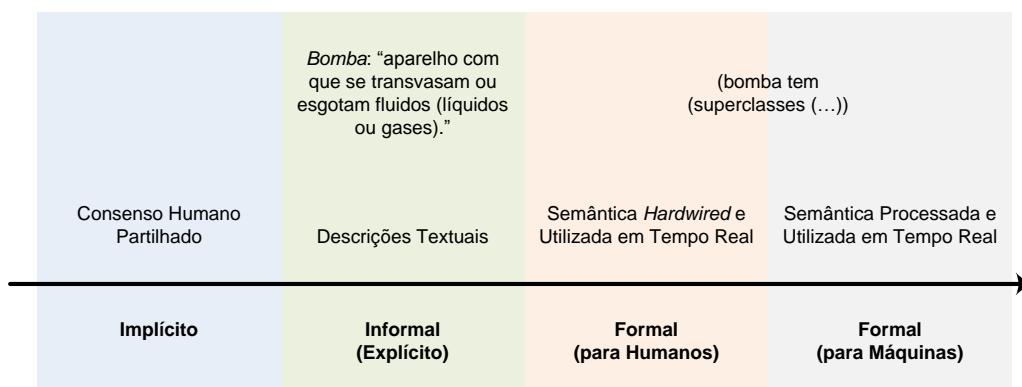


Figura 9 – Continuum Semântico (Uschold, 2003).

A Web Semântica tenciona criar um meio de suporte universal à partilha de informação na Web, através da (i) especificação de relações entre conteúdos publicados em diferentes páginas Web. Esta partilha é também caracterizada pela (ii) introdução de significação (semântica) nos dados existentes nos documentos de forma a permitir que as máquinas compreendam e interpretem o significado dos mesmos (Berners-Lee, Hendler, & Lassila, 2006). Enquanto o primeiro ponto (i) se foca na criação de uma rede global de informação distribuída, atualmente mais conhecida por LOD, o segundo (ii) assenta fortemente sobre a utilização de ontologias que permitem uma descrição formal da informação encontrada na LOD.

As ontologias permitem resolver mais facilmente alguns problemas de integração e interoperabilidade na Web como é o caso da desambiguação de pesquisas (por exemplo, palavras com múltiplos sentidos). Em sistemas heterogêneos e distribuídos, do qual a Web é o caso mais paradigmático e relevante, dada a sua importância e dimensão, a integração de informação é uma componente fundamental para a interoperabilidade entre aplicações.

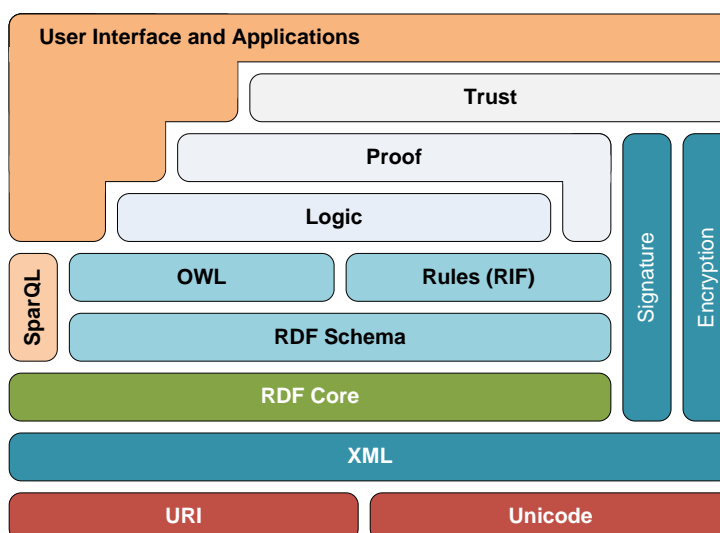


Figura 10 – Pilha de tecnologias da Web Semântica.

A Web Semântica é suportada pelo *W3C Semantic Web Activity*, um corpo internacional de padronização para a Internet.

Na figura 10 estão representadas as principais camadas da visão da Semântica Web, na qual se encontra a *Resource Description Framework* (RDF). No nível mais baixo encontra-se o XML, que permite escrever documentos Web estruturados com um vocabulário definido pelo utilizador e é particularmente adequada para enviar documentos através da Web, ou seja, é basicamente um formato de serialização. No nível imediatamente superior surge o RDF que, apesar do seu modelo de dados não depender do XML, tem uma sintaxe baseada no mesmo.

3.2 Ontologias

Etimologicamente, ontologia significa "ciência do ser" ou "o estudo metafísico da natureza do ser e existência" e lida com a natureza e organização da realidade, tendo sido inicialmente estudado por Platão e Aristóteles (Silva, 2004).

No âmbito das ciências da computação, trata-se de uma especificação explícita e formal de uma conceptualização partilhada (Gruber, 2009). Por outras palavras, trata-se de uma rigorosa e exaustiva organização do conhecimento de um domínio, podendo ser estruturado de forma hierárquica ou em rede, e que contém todas as entidades e relações relevantes.

Uma ontologia define um conjunto de primitivas representativas que permitem moldar um domínio de conhecimento. As primitivas representativas são tipicamente classes (ou conjuntos), atributos (ou propriedades) e relacionamentos (ou relações entre membros das classes). As definições das primitivas representativas incluem informação acerca do seu significado e restrições na sua aplicação lógica consistente.

No contexto dos sistemas de base de dados, a ontologia pode ser vista como um nível de abstração dos modelos de dados, análogos a modelos hierárquicos e relacionais, mas cuja intenção é modelar o conhecimento acerca de indivíduos, os seus atributos, e os seus relacionamentos com outros indivíduos. As ontologias fazem parte do que se denomina por bases de conhecimento¹⁴, que são repositórios semânticos (como p. ex. o OWLIM¹⁵) cujo conteúdo é passível de ser processado por máquinas¹⁶ e usam algoritmos de inferência para deduzir "novo" conhecimento (implícito). Os repositórios semânticos são motores, similares aos SGBD, que permitem armazenar, pesquisar e gerir informação estruturada. Com a grande diferença, de usarem ontologias como esquema semântico. Isto permite-lhes automaticamente utilizar técnicas de inferência sobre o conhecimento e funcionar com modelos de dados flexíveis, como p. ex. os grafos.

¹⁴ Do anglo-saxónico *knowledge bases*.

¹⁵ <http://www.ontotext.com/owlim/>

¹⁶ Do anglo-saxónico *machine-readable*.

As ontologias são tipicamente especificadas em linguagens que permitem abstração das estruturas de dados e das estratégias de implementação; na prática, as linguagens das ontologias estão mais próximas do poder de expressão da lógica de primeira ordem do que outras linguagens usadas na modelação de bases de dados.

Por esta razão, é dito que as ontologias estão ao nível "semântico", enquanto os esquemas de base de dados são modelos de dados no nível "lógico" ou "físico".

Devido à sua independência relativamente a modelos de dados de nível inferior, as ontologias são usadas para integrar bases de dados heterogêneas, permitindo a interoperabilidade entre múltiplos sistemas, e especificando interfaces a serviços independentes, baseados em conhecimento. As ontologias também podem funcionar como componente de bases de conhecimento.

Dado um domínio, a sua ontologia forma o núcleo de qualquer sistema de representação do conhecimento desse domínio. Sem ontologias, ou uma conceptualização que sirva de base ao conhecimento, não existe um vocabulário para a representação desse conhecimento.

As ontologias fazem parte dos modelos padrão recomendados pela W3C para a Web Semântica, nos quais são utilizadas para especificar vocabulários conceptuais padrão para troca de dados entre sistemas, fornecimento de serviços para resposta a pesquisas, publicação de bases de conhecimento reutilizáveis, e oferecer serviços para facilitar a interoperabilidade através de múltiplos, sistemas heterogêneos e bases de dados. Atualmente, existem diversas ontologias na Web que descrevem diferentes dimensões (p. ex. tempo, espaço) e domínios de conhecimento (p. ex. turismo, geografia). Com o aparecimento de múltiplas ontologias para o mesmo domínio de conhecimento, tem sido feito um esforço no sentido de definir *standards* (ver figura 11).

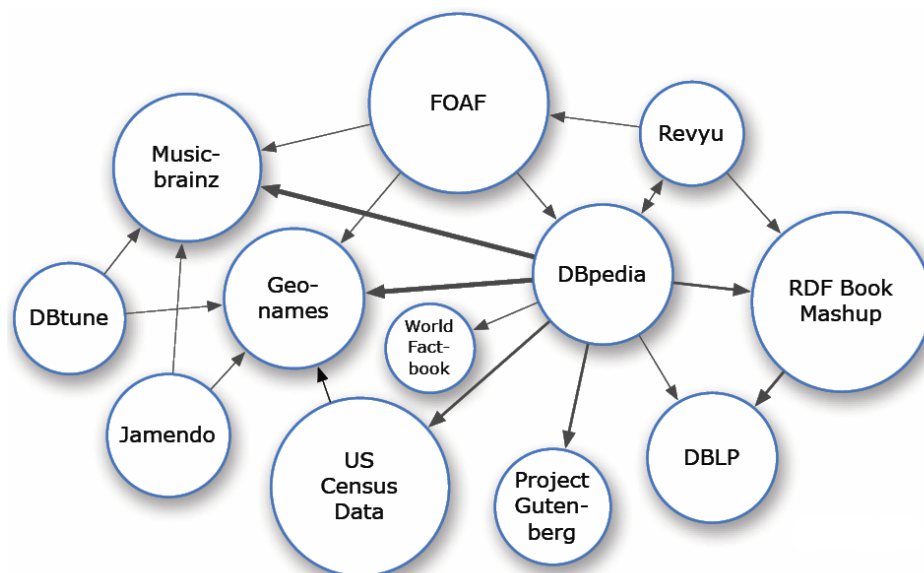


Figura 11 – Conjunto parcial de ontologias extensíveis e interligadas atualmente na Web Semântica.

A necessidade de vocabulário controlado e de consenso partilhado sugere que as ontologias irão desempenhar um papel chave na gestão da informação com contexto geográfico (Scharl, 2007). O poder dos SIG advém da habilidade de relacionar informação diferente no contexto espacial e chegar a uma conclusão sobre essas relações. Através da utilização de ontologias, é possível apresentar informação proveniente de qualquer lugar na Web Semântica, tirando partido de funcionalidades como a inferência de novos factos e a pesquisa por facetas (Vallea & Qasima, 2010).

As linguagens mais populares para a criação de ontologias baseiam-se na lógica descritiva. Esta representa um conjunto de famílias (p. ex. ALC, SHIQ, etc.) com diferentes expressividades, sempre inferiores à expressividade da lógica de primeira ordem. Atualmente, as linguagens mais utilizadas são a RDF, a *Resource Description Framework Schema* (RDF-S) e a mais expressiva *Web Ontology Language* (OWL), ambas *standards* da W3C.

3.2.1 RDF, RDF-S e SPARQL

A linguagem RDF é em tudo similar à modelação conceptual do tipo Entidade-Relação ou diagrama de classes, já que é baseada na criação de instruções sobre recursos (recursos Web em particular) na forma sujeito-predicado-objeto. Nesses triplos o sujeito referencia o recurso, e o predicado (também designado por propriedade do triplo) referencia os aspetos do recurso,

expressando a relação entre o sujeito e o objeto. Por exemplo, para representar "a relva tem a cor verde" num triplo RDF, o sujeito será a "relva", o predicado o "tem" e o objeto "cor verde".

A RDF permite que *software* automatizado possa guardar, trocar e usar informação distribuída pela Web. Outra das características da linguagem é a simplicidade, permitindo modelar conceitos abstratos e díspares, o que levou a um aumento da sua utilização em aplicações de gestão do conhecimento, não relacionadas com a Web Semântica.

Por um lado, a RDF pode ser encarada como um mecanismo para a criação da LOD, ou seja permite a representação e o relacionamento de dados distribuídos. A RDF-S, por outro lado, permite a definição do vocabulário sobre o qual atua a RDF; mais concretamente, a especificação da ontologia. A RDF-S é uma linguagem de representação de conhecimento extensível, que fornece elementos básicos para a descrição de ontologias, previamente intitulados de vocabulários RDF, sendo uma recomendação da W3C desde Fevereiro de 2004. Muitos dos elementos que compõem a RDF-S fazem parte da linguagem OWL (detalhada na secção 3.2.2), bastante mais expressiva que a RDF-S.

Mas a RDF possui também várias linguagens de pesquisa, sendo que a mais predominante é a *SPARQL Protocol and RDF Query Language* (SPARQL), que permite recuperar dados em formato RDF (triplos) com base em padrões e filtros de pesquisa. As pesquisas SPARQL são executadas em conjuntos de dados RDF (grafos RDF) e aceites através de múltiplos pontos de acesso SPARQL que retornam a junção de todos os resultados via HTTP. Os resultados das pesquisas SPARQL podem ser retornados numa variedade de formatos, tais como o XML, JSON, RDF ou HTML. Tanto o protocolo, como a linguagem SPARQL são uma recomendação da W3C desde Janeiro de 2008 e são considerados como uma das tecnologias chave da Web semântica.

Mais recentemente em 2011, o Open Geospatial Consortium (OGC) criou o *standard* GeoSPARQL, cujo objetivo é suportar a representação e pesquisa de informação geoespacial na Web Semântica. O GeoSPARQL define um vocabulário para representar os dados em RDF, e define uma extensão ao protocolo e linguagem SPARQL para processar informação geográfica. Suporta classes como a *SpatialObject*, a *Feature* e a *Geometry*, relações (como p. ex. a relação *is3D* da classe *Geometry* que reflete se a geometria é representada em 3 dimensões), entre as quais relações topológicas (relações *within*, *overlaps* e *disjoint*).

3.2.2 Web Ontology Language

A OWL é uma linguagem usada para a especificação e instanciação de ontologias na Web e criada de acordo com os desafios impostos pela Web Semântica. Construída com base na RDF-S e na sua predecessora DARPA *Agent Markup Language + Ontology Inference Layer* (DAML+OIL¹⁷), possui um vocabulário mais vasto e uma maior expressividade que a RDF-S. É uma recomendação da W3C e tem vindo a evoluir em termos de expressividade, acompanhando a evolução dos motores de inferência (como p. ex. o FaCT++¹⁸ ou o Pellet¹⁹) sobre lógica descritiva. Algumas das principais motivações para a utilização da OWL são:

- A OWL explora a capacidade do XML para a criação de esquemas estruturados e customizados;
- O XML só por si, não permite a imposição de restrições semânticas sobre o significado dos documentos;
- A OWL possui um vocabulário extenso para a descrição de propriedades e classes: relações entre classes (p. ex. se são disjuntivas), cardinalidade, equivalência, tipificação enriquecida de propriedades, caracterização de propriedades (p. ex. simetria), enumeração de classes, entre outras.

A linguagem OWL é considerada bastante relevante para a implementação da Web Semântica. Com o decorrer do tempo, tem vindo a ocupar um papel cada vez mais importante num grande número de aplicações, e tem sido objeto de pesquisa para ferramentas e técnicas de inferência.

A nova versão da OWL, a OWL 2²⁰, introduziu uma maior expressividade, incluindo também diferentes perfis (sub-linguagens) com diferentes expressividades, orientados a diferentes requisitos:

- OWL 2 EL – permite a utilização de algoritmos de complexidade temporal polinomial para as tarefas de raciocínio *standard* (ideal para ontologias de grande dimensão);
- OWL 2 QL – permite a resposta a consultas que envolvem operações de conjunção com uma complexidade espacial logarítmica (ideal para aplicações com ontologias leves mas com uma grande quantidade de instâncias acessíveis através de consultas relacionais);
- OWL 2 RL – permite a utilização de algoritmos de complexidade temporal polinomial com regras (ideal para aplicações com ontologias leves mas com uma grande quantidade de instâncias acessíveis sob a forma de triplos).

¹⁷ <http://pt.wikipedia.org/wiki/DAML%2BOIL>

¹⁸ <http://owl.man.ac.uk/factplusplus/>

¹⁹ <http://clarkparsia.com/pellet/>

²⁰ <http://www.w3.org/TR/owl2-overview/>

Apesar de não ser uma linguagem de fácil leitura humana, ferramentas como o editor de ontologias Protégé²¹, permitem a criação e manipulação gráfica de ontologias OWL, utilizando diferentes formatos como o XML e N3²² (N-Triple).

3.3 Pesquisa por Facetas

A classificação com base em hierarquias tem sido aplicada ao longo dos anos com sucesso para a organização e categorização de informação. Contudo, esta classificação começou por ser feita com base numa única hierarquia de conceitos. A ideia da classificação com base em múltiplas hierarquias surgiu na década de 1930, no entanto, só foi realmente aplicada por investigadores nas décadas de 1980 e 1990 (Makela, Hyvonen, & Sidoroff, 2005). Atualmente, a classificação poli-hierárquica de itens tem sido alvo de crescente interesse devido ao aparecimento da Web Semântica e à respetiva utilização de ontologias para descrever e modelar conhecimento (Hildebrand & Ossenbruggen, 2006).

A pesquisa por facetas consiste na utilização da classificação poli-hierárquica para assistir o utilizador na procura e filtragem de informação. Cada hierarquia (ou taxonomia) ortogonal corresponde a uma faceta ou vista, sendo o papel do utilizador selecionar diferentes categorias das facetas de modo a filtrar a informação (Baeza-Yates & Ribeiro-Neto, 1999) (Suominen, Viljanen, & Hyvänen, 2007). A interface com o utilizador é gerada automaticamente com base nas facetas e de acordo com o nível conceptual em que o mesmo se encontra. Por exemplo, informação relativa a pontos de interesse turísticos pode ser classificada com base nas dimensões tempo (por ex., ontologias OWL-Time e SUMO) e espaço (por ex. ontologias GeoNames, Basic Geo Vocabulary e SUMO). Paralelamente, existe ainda a classificação que diz respeito ao domínio de conhecimento do turismo. Cada uma destas classificações ou facetas pode estar descrita em diferentes ontologias. A expressividade inerente às ontologias permite criar automaticamente interfaces de pesquisa por facetas, em que cada instância é apresentada de acordo com a sua descrição conceptual. As taxonomias utilizadas na filtragem são geralmente definidas por relações ontológicas, como sub-classe-de ou parte-de.

A pesquisa por facetas é particularmente interessante na área dos SIG Web, visto que na maior parte dos casos o utilizador acede à informação através da utilização de filtros de pesquisa. Todos estes fatores têm levado a que os resultados obtidos no seguimento da Web Semântica, nos últimos cinco anos, tenham tido uma forte influência na investigação associada aos SIG Web.

²¹ <http://protege.stanford.edu/>

²² <http://www.w3.org/TR/rdf-testcases/#ntriples/>

Capítulo 4

Metodologia para Geoportais Semanticamente Enriquecidos

Um geoportal permite aos utilizadores aceder a informação através de um mapa e seleccionar áreas de interesse a explorar. Contudo, a actual organização da informação geográfica em catálogos de metadados é incapaz de capturar a semântica dos dados descritos. Como consequência, os utilizadores muitas vezes não conseguem encontrar recursos geoespaciais de interesse quando procuram por conteúdo na Web.

Os mapas semanticamente enriquecidos (conceito proposto neste capítulo) baseiam-se no princípio da Web Semântica, de que os conteúdos publicados na Web devem ser acompanhados de significado/sentido perceptível tanto pelos humanos como pelos computadores. Assumindo assim que grande parte da informação publicada na Web tende a ser cada vez mais descrita por modelos enriquecidos a nível semântico (p. ex., ontologias), este conhecimento pode agora ser aproveitado e explorado na apresentação de informação sobre mapas. Um exemplo concreto são os restaurantes que se relacionam entre si (p. ex. pelo mesmo proprietário, mesmo tipo de cozinha, etc.). O potencial de exploração gráfica deste tipo de relações é bastante interessante mas carece de soluções.

Neste capítulo é apresentada a ontologia *Semantic Maps (SMaps)* (ver secção 4.1), que constitui a base para o conceito de mapas semanticamente enriquecidos posteriormente detalhado na secção 4.2. Na secção 4.3 são propostas duas abordagens semanticamente enriquecidas para a interface com o utilizador. Finalmente, é apresentada uma arquitetura de alto nível (ver secção 4.5) para a implementação de SIG Web baseados no conceito proposto.

4.1 Ontologia Semantic Maps

A ontologia *SMaps* representa um vocabulário base para a descrição de conceitos espaciais. Esta ontologia absorve a informação consolidada na ontologia *Basic Geo*²³ (WGS84 lat/long), que inclui a classe *Point*.

Visto que a ontologia *Basic Geo* é relativamente simples, e não satisfaz a maior parte das necessidades de um SIG, foi feita uma análise das *features*²⁴ propostas na implementação *standard* do *Open Geographic Consortium* (OGC). Com base nesta análise, para além da classe *Point*, foram acrescentadas as classes *LineString*, *Polygon*, *MultiPoint*, *MultiLineString*, *MultiPolygon* e *GeomCollection*.

Paralelamente, a ontologia foi criada tendo em conta a complexidade computacional relativamente a processos de inferência (expressividade da ontologia).

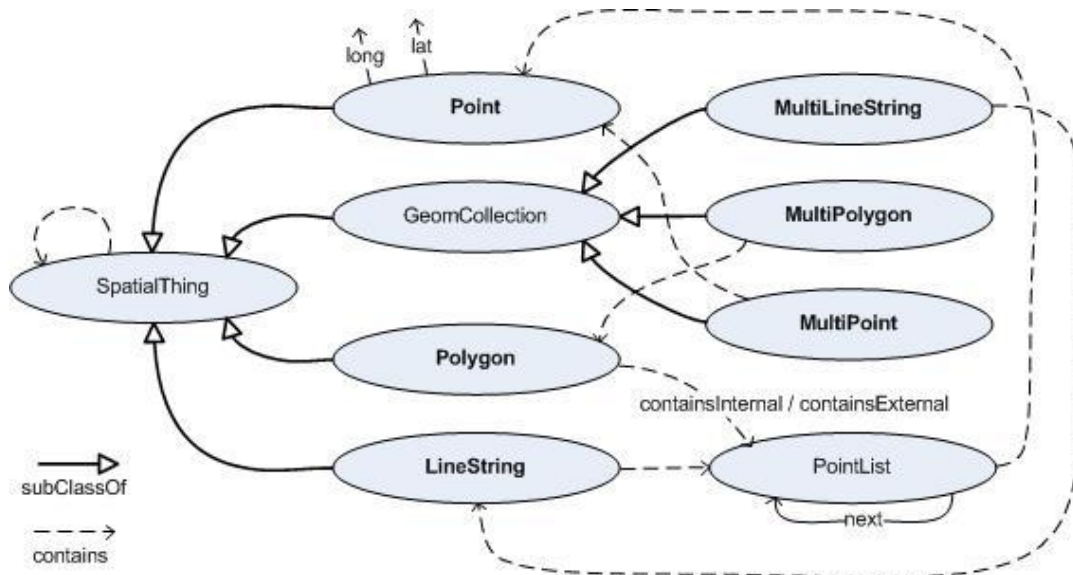


Figura 12 – Principais entidades e relações a ontologia *SMaps*.

Assim como acontece na ontologia *Basic Geo*, a *SMaps* disponibiliza duas formas de representação de geometrias complexas:

- Através da propriedade *contains*, que permite a definição, p. ex. de um polígono, por meio de um conjunto de instâncias da classe *Point*;

²³ <http://www.w3.org/2003/01/geo/>

²⁴ Neste contexto entende-se *feature* como qualquer entidade com representação física, p. ex. uma linha ou um polígono.

- Através de propriedades *datatype*²⁵, que permitem a definição, p.ex. de um polígono, por meio de uma cadeia de caracteres que indica um conjunto de pontos.

Existem duas possíveis abordagens à criação de ontologias: centrada nas propriedades e centrada em lógica descritiva. A primeira tem sido bastante utilizada na Web Semântica devido à sua simplicidade e fácil compreensão por parte de peritos não especializados em engenharia do conhecimento. No entanto, a abordagem centrada nas propriedades pode levar a maus resultados a longo prazo no que diz respeito a processos de inferência e a reutilização de propriedades.

Numa abordagem centrada nas propriedades, o domínio (indivíduos sobre os quais a propriedade vai atuar) e o contradomínio (gama de valores possíveis da propriedade) de cada propriedade é definido, restringindo a utilização da propriedade. Isto implica que, para uma propriedade X com domínio A e contradomínio B num processo de inferência tradicional, todas as instâncias descritas pela propriedade X , sejam implicitamente classificadas como sendo do tipo A . O mesmo acontece para o contradomínio. No caso da ontologia *SMaps*, restringir a propriedade *contains* de forma a que o seu domínio fosse *GeomCollection* e o seu contradomínio *SpatialThing*, iria impedir que a mesma fosse utilizada para a definição de listas (p. ex. a *PointList*).

Sendo assim, foi adotada uma abordagem centrada em lógica descritiva, onde o foco incide sobre as classes e não sobre as propriedades. Desta forma, as classes são definidas através de axiomas em lógica descritiva. Por exemplo, a classe *PointList* é sub-classe das classes de objetos que respeitam os seguintes axiomas, respetivamente: $\exists \textit{contains.Point}$ e $\exists \textit{next.PointList}$, em que o quantificador existencial significa que para as instâncias de *PointList* tem que existir pelo menos um *Point* relacionado através da propriedade *contains*, e uma *PointList* relacionada através da propriedade *next*.

A definição de sequências em OWL não é de todo trivial, tendo sido utilizado um padrão de lista ligada apresentado e descrito em diversas publicações (Drummond, Rector & Stevens 2006a) (Drummond, Rector & Stevens 2006b). Este padrão de recursividade envolve a criação da classe *OWLList*, definida através das propriedades *contains (head)* e *next (tail)*. A propriedade *contains* é utilizada para definir um elemento da lista, sendo a propriedade *next* utilizada para indicar outro elemento do tipo *OWLList*. O resultado para listas do tipo *PointList* (uma sub-classe de *OWLList*) encontra-se ilustrado no exemplo apresentado na figura 13.

²⁵ Propriedades *datatype* são propriedades cujo contradomínio é um tipo de dados primitivo (p. ex. string ou integer).

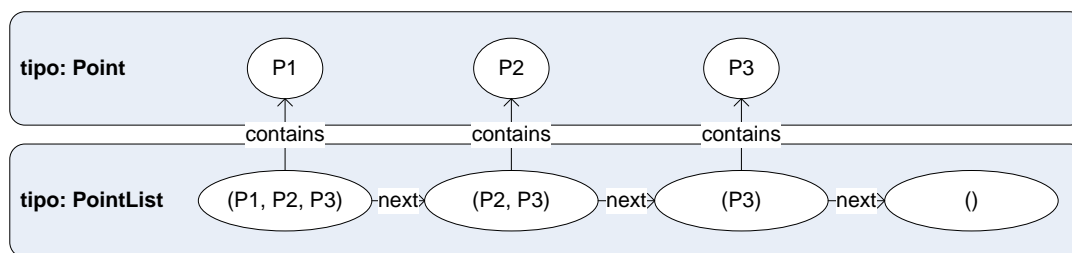


Figura 13 – Padrão de representação de listas em OWL (Drummond, Rector, & Stevens, 2006c)(Drummond, Rector & Stevens 2006b).

A utilização da ontologia *SMaps* permite a representação de informação geográfica com base nos *standards* da Web Semântica (como o RDF e OWL). A sua implementação foi feita utilizando o editor de ontologias *Protégé*, através do qual foi obtido o *Manchester Syntax Render* encontrado no anexo A.

4.2 Ontologias e Informação Geoespacial

Numa ontologia, as classes são definidas através de axiomas que permitem a sua caracterização com base em relações cujo domínio e contradomínio pode (ou não) ser restrito. Esta representação de conhecimento forma uma rede que interliga diferentes conceitos, permitindo uma navegação que em muito se assemelha a percorrer os nós de um grafo. Os axiomas através dos quais são definidas as classes (que geralmente incidem sobre propriedades), são de particular interesse neste tipo de navegação, permitindo (para além de navegação por facetas) a fácil extração de instâncias relacionadas através dos valores das suas propriedades.

Na figura 14, apresentada uma representação parcial de uma ontologia que importa a ontologia *SMaps*, e que descreve o domínio da restauração: a ontologia *Rest*. Na ontologia *Rest*, existem as classes *Restaurant*, *Country* e *Hotel*. A classe *Restaurant*, em particular, é caracterizada pelas propriedades *name*, *lon*, *lat*, *website*, *amountOfSeats*, *cuisine* (com contradomínio *Country*, restrito por axioma sobre a classe) e *hotel* (com contradomínio *Hotel*, também restrito por um axioma sobre a classe).

A existência de conceitos com caracterização espacial implica muitas vezes a existência de uma representação visual sobre o mapa. Este é o caso das classes *Restaurant* e *Hotel* (representadas por pontos), e *Country* (representada por polígonos). Neste tipo de situações, os valores da propriedade *cuisine* da classe *Restaurant* deixam de ser apenas “valores” sem semântica a apresentar (possivelmente) numa caixa de texto, e passam a ser reconhecidos como instâncias de *Country* cuja semântica se encontra na ontologia *Rest*. Isto permite que *Country*

seja reconhecido como um conceito espacial e apresentado sobre o mapa juntamente com o restaurante.

Para que uma classe seja reconhecida como conceito espacial, é necessário que exista um vocabulário base que faça a descrição de conceitos como *Point* e *Polygon*. Esta caracterização espacial é dada pela ontologia *SMaps*.

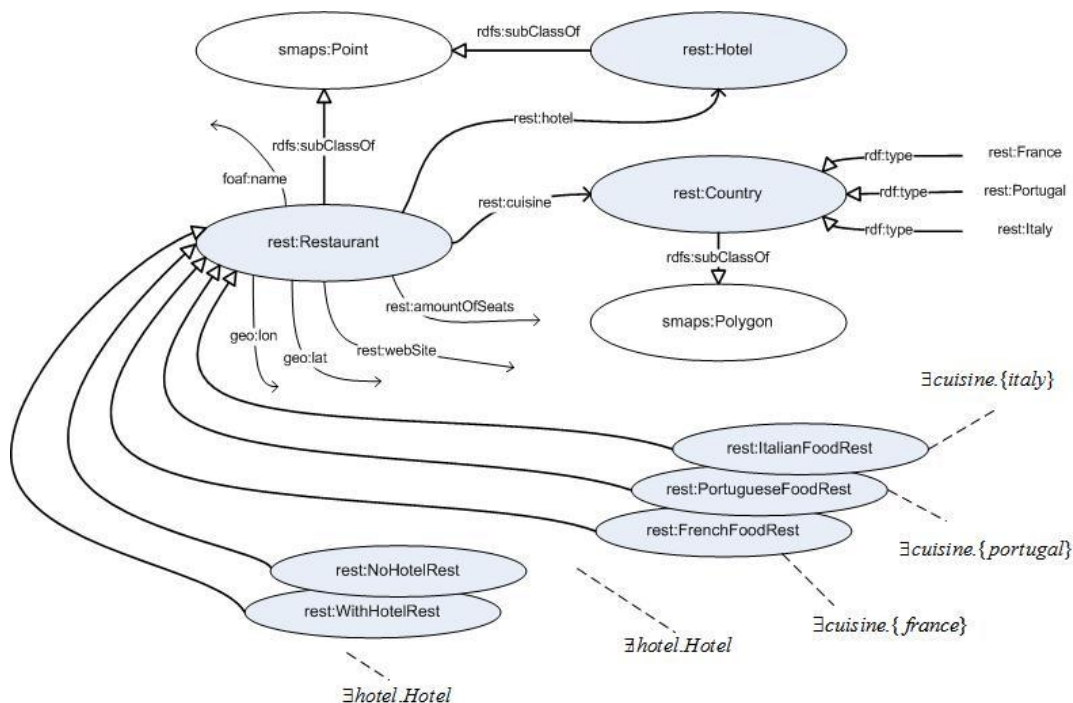


Figura 14 – Ilustração parcial da ontologia Rest.

A representação gráfica enriquecida dos conceitos espaciais permite que o utilizador consiga visualizar e absorver mais facilmente informação de *features* relacionadas. Adicionalmente, devido à forma como é estruturada a informação (do tipo grafo), é possível criar interfaces que permitem uma fácil e intuitiva navegação pelos conceitos e instâncias.

4.3 Interface com o Utilizador

A interface com o utilizador deve tirar partido da componente semântica presente na arquitetura proposta. Tendo em conta esta arquitetura, duas possíveis estratégias para a apresentação de informação podem ser adotadas:

- A apresentação externa dos conteúdos, em que a informação relativamente a uma *feature* é apresentada fora da área do mapa;

- A apresentação interna dos conteúdos, em que a informação relativamente a uma *feature* é apresentada sobre o mapa, através de objetos gráficos e estilos (p. ex. em SVG) associados aos conceitos ontológicos (p. ex. *Polygon*, *Restaurant* e/ou *Country*).

Independentemente do tipo de estratégia de apresentação adotada, são especialmente focados alguns aspetos tanto ao nível dos dados como dos metadados. Relativamente aos dados, é dada particular importância à:

- Navegação através da informação utilizando o grafo de triplos, permitindo a apresentação de instâncias relacionadas;
- Visualização da metainformação da instância selecionada e respetivas propriedades.

No que diz respeito aos metadados, os seguintes aspetos são tidos em conta:

- Pesquisa por facetar;
- Visualização das restrições semânticas de uma determinada classe;
- Visualização de estatísticas para cada classe, de uma perspectiva centrada nas propriedades.

4.3.1 Apresentação Externa de Conteúdos

A apresentação externa de conteúdos permite libertar a área do mapa para a visualização exclusiva de *features* (geometrias). Apesar de implicar o desvio da atenção do utilizador para um (possível) painel lateral, esta abordagem pode ser vantajosa em situações de grande densidade de *features*.

Um geoportal permite a pesquisa e apresentação de conteúdos com caracterização espacial. Neste sentido, qualquer instância da classe *SpatialThing* (incluindo subclasses) pode ser pesquisada e visualizada. De acordo com a ontologia *Rest*, estas instâncias são os restaurantes (da classe *Restaurant*), os hotéis (da classe *Hotel*) e os países (da classe *Country*). Os atributos apresentados para cada instância irão depender dos axiomas encontrados na ontologia, e da informação encontrada no grafo de triplos. Na figura 15 é mostrado um exemplo de apresentação externa de conteúdos caracterizados com a ontologia *Rest*. O restaurante “Bife de Ferro”, cujos detalhes se encontram ao lado direito do mapa, é caracterizado pelas propriedades (com etiquetas em português e inglês) *name*, *cuisine*, *hotel*, *amountOfSeats*, entre outras. Cada uma destas propriedades possui valores com diferentes tipos de dados, o que pode alterar a forma como são apresentadas.



Figura 15 – Apresentação externa de conteúdos de acordo com a ontologia *Rest*.

Tendo em conta uma melhor visualização dos detalhes das instâncias, foram identificados quatro tipos de contradomínio para as propriedades:

- **Conjuntos complexos** (p. ex. classes), como é o caso da propriedade *hotel* e *cuisine* para a classe *Restaurant*;
- **Conjuntos primitivos** (p. ex. uma enumeração de cadeias de caracteres);
- **Conjuntos indefinidos** (não é possível saber a extensão do conjunto), como é o caso da propriedade *name* para a classe *Restaurant*;
- **Intervalos numéricos** (p. ex. inteiros), como é o caso da propriedade *amountOfSeats* para a classe *Restaurant*.

Para cada um destes tipos de contradomínio, podem ser adotadas diferentes estratégias de visualização. No caso dos intervalos numéricos, pode ser utilizada uma representação gráfica que permita a visualização do valor mínimo, médio e máximo para todas as instâncias encontradas num raio de X quilómetros. Para conjuntos primitivos pode ser sugerida a quantidade de instâncias associadas a um determinado elemento do conjunto. Já para os conjuntos indefinidos, o mesmo perde o significado. Relativamente a conjuntos complexos, sempre que o contradomínio for do tipo *SpatialThing*, devem ser permitidas:

- A navegação direta para o valor da propriedade (navegação pelo grafo de triplos);
- A apresentação sobre o mapa das instâncias relacionadas.

Tanto para propriedades cujo contradomínio representa um conjunto primitivos como para propriedades cujo contradomínio representa um conjunto complexos, deve ser possível a apresentação de instâncias relacionadas. Duas instâncias são relacionadas perante uma determinada propriedade, se a interseção dos seus valores para a propriedade não for um conjunto vazio. Por exemplo, dois restaurantes de cozinha portuguesa estão relacionados pela propriedade *cuisine*, visto possuírem um valor em comum: Portugal. Na figura 15, o ícone com um cadeado apresentado ao lado esquerdo da propriedade *hotel* permite mostrar/esconder restaurantes relacionados através da propriedade *hotel*.

Adotando uma abordagem focada nos metadados é ainda possível apresentar estatísticas como:

- O número de instâncias para uma determinada classe;
- A quantidade mínima, média e máxima de vezes que uma propriedade é definida para instâncias de uma determinada classe, de acordo também com as restrições de cardinalidade aplicadas.

Estas estatísticas podem ser restringidas a um determinado raio espacial de pesquisa.

Relativamente às facetas para a pesquisa e filtragem de conteúdos (como apresentado na figura 16), estas podem ser encontradas na ontologia de forma (i) explícita ou (ii) implícita.



Figura 16 – Pesquisa por facetas de acordo com a taxonomia de restaurantes da ontologia *Rest*.

A forma (i) explícita implica que as facetas apresentadas tenham uma correspondência direta com classes declaradas na ontologia. Tendo como referência a ontologia apresentada na figura 14, a figura 16 ilustra uma forma de apresentação explícita de facetas.

A forma (ii) implícita implica uma análise prévia da ontologia e das suas instâncias. Um exemplo é a análise dos axiomas que definem cada uma das classes. No caso da ontologia apresentada na figura 14, a existência do axioma $\exists cuisine.Country$ para a classe *Restaurant* e das instâncias de *Country* (*Portugal*, *Italy* e *France*), podem levar à apresentação implícita de três facetas sem a necessidade de existirem as classes *ItalianFoodRest*, *PortugueseFoodRest* e *FrenchFoodRest*.

Esta análise pode ainda ser feita de forma a orientar a apresentação de facetas de acordo com o perfil do utilizador. Isto é especialmente importante num ambiente geoespacial onde grandes quantidades de informação, provenientes de diferentes fontes de informação com diferentes ontologias, são apresentadas. Desta forma, por exemplo, para um utilizador que tenha particular interesse em automóveis, podem ser apresentadas em primeiro lugar facetas relacionadas com conceitos do domínio dos automóveis.

4.3.2 Apresentação Interna de Conteúdos

A apresentação interna de conteúdos permite que o utilizador se foque na informação geoespacial, levando a uma rápida interpretação e correlação dos dados apresentados. Visto que a interação e apresentação de informação sobre o mapa ocupa espaço, ofuscando muitas vezes outras *features*, esta estratégia de apresentação de conteúdos é ideal para ambientes em que existe pouca densidade de *features*.

À semelhança do que acontece com a apresentação externa de conteúdos, o objetivo é tirar partido da semântica encontrada nas ontologias. No entanto, existem alguns cuidados adicionais a ter relativamente à qualidade e usabilidade da interface, que incidem principalmente na ofuscação de *features*. Neste sentido, e também por motivos de flexibilidade, é proposto o conceito de objeto gráfico.

Um objeto gráfico é uma representação gráfica sobre o mapa, dinâmica e extensível, de um determinado conceito espacial (por ex., *Restaurant*). Para além de permitir uma fácil identificação do tipo de *feature* apresentada, um objeto gráfico permite também a visualização dos atributos da instância (propriedades associadas ao conceito/classe) e a navegação pelo grafo de triplos.

A caracterização dinâmica dos objetos gráficos advém do facto de serem uma ferramenta de interação com o utilizador, sujeita a diferentes estados e contendo diferentes tipos de ações. Por outro lado, a sua caracterização extensível deriva do facto de poderem ser reutilizados para a criação de novos objetos gráficos.

Os objetos gráficos são associados a uma classe na ontologia, sendo utilizados posteriormente para a representação de qualquer instância dessa classe ou de uma das suas subclasses (se esta não tiver outro objeto gráfico explicitamente associado) sobre o mapa.

Na interface apresentada na figura 17, a *feature* (instância de *Restaurant*) com o nome “Bife de Ferro” encontra-se selecionada. Esta *feature* é representada por um objeto gráfico que está associado à classe *Restaurant*. Este objeto gráfico estende o objeto gráfico associado à classe *SpatialThing*, sendo apenas alterado o ícone para um mais adequado à classe *Restaurant*.

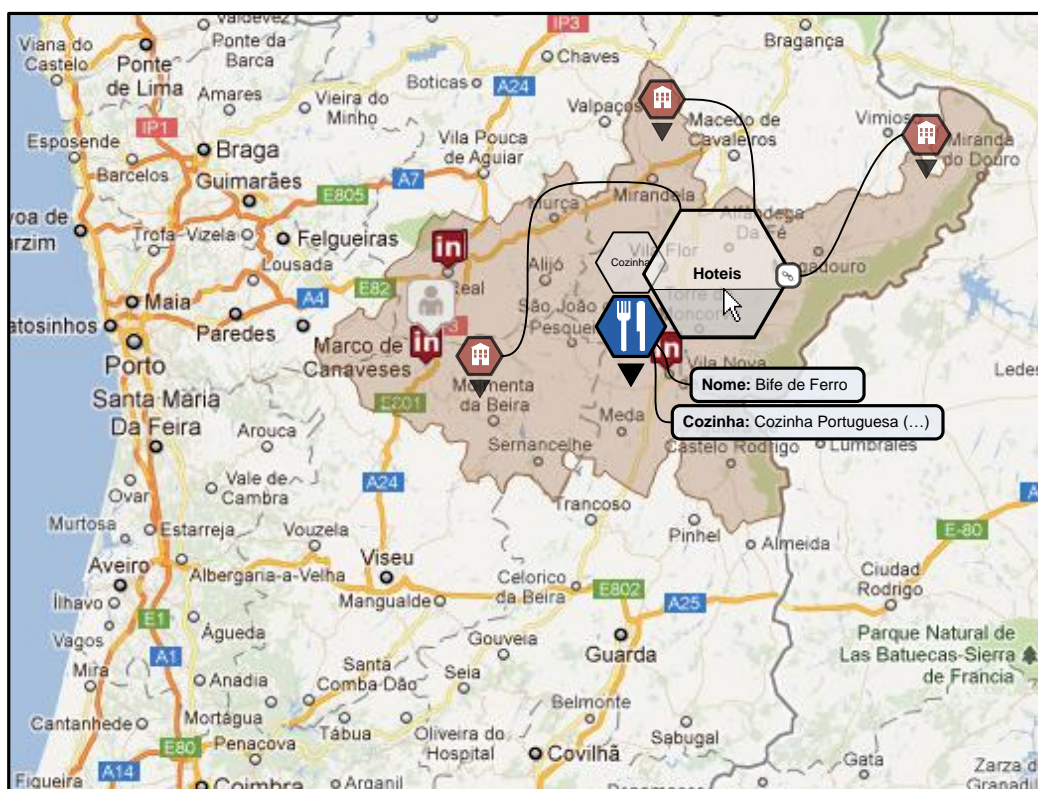


Figura 17 – Apresentação interna de conteúdos de acordo com a ontologia *Rest*.

O objeto gráfico da classe *SpatialThing* é constituído por vários hexágonos. O hexágono central possui o ícone representativo da classe. Os hexágonos periféricos representam, cada um deles, uma propriedade associada à classe cujo contradomínio é um conceito espacial (ou seja, subclasse de *SpatialThing*). Para todos os outros tipos de contradomínio é apresentada uma caixa de texto.

O preenchimento dos hexágonos periféricos pode também ser utilizado para a apresentação de estatísticas relativamente à quantidade mínima e máxima de instâncias associadas através da propriedade selecionada. Por exemplo, na figura 17 é possível ver que o preenchimento do hexágono “Hotéis” indica o número de hotéis associados ao restaurante “Bife de Ferro” (neste

caso em concreto são três) por comparação ao número mínimo e máximo de hotéis associados a instâncias de *Restaurant*.

4.4 Framework *SMaps*

De acordo com a metodologia previamente descrita, é proposta uma *framework*, intitulada *SMaps*, para a implementação de geoportais seguindo os princípios da Web Semântica. A arquitetura desta *framework* possui diversas camadas e encontra-se ilustrada na figura 18.

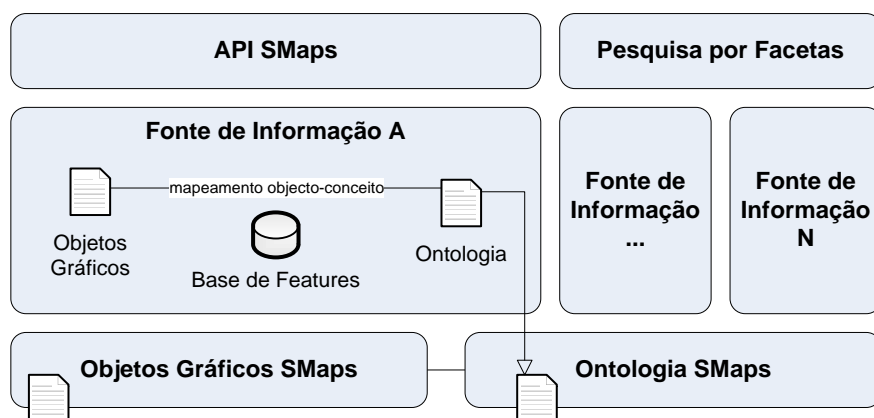


Figura 18 – Arquitetura da *framework SMaps*.

Na base da arquitetura encontram-se a ontologia e os objetos gráficos *SMaps*. Esta base é estendida pela camada que se segue, que especifica o(s) domínio(s), e que representa as diversas fontes de informação georreferenciada. Estas fontes de informação podem constituir serviços de pesquisa na Web Semântica (como pontos de acesso SPARQL) ou referências para ficheiros RDF, RDF-S e OWL. No topo da arquitetura existe a API *SMaps* (para a apresentação do mapa e da informação georreferenciada) e os algoritmos de pesquisa por facetas, que extraem facetas com base nas ontologias que descrevem a informação proveniente das diversas fontes que coexistem na camada inferior.

A implementação de cada uma destas camadas, juntamente com os desafios inerentes à mesma, é detalhada no próximo capítulo.

Capítulo 5

Implementação e Validação

Para a implementação da Framework, baseada na metodologia previamente apresentada, foi feita uma análise das APIs Web e fontes de informação atualmente disponíveis. Apesar de algumas APIs, como é o caso da *OpenLayers*²⁶, serem amplamente utilizadas e já com alguma maturidade no seu desenvolvimento, existem diversos desafios inerentes a esta implementação. Neste capítulo será feita uma avaliação das APIs atuais tendo em conta estes desafios encontrados durante a implementação do protótipo.

5.1 Desafios de Implementação

Existem várias APIs que permitem a utilização de mapas em ambiente Web, como é o caso da API *Google Maps* (da *Google*), da API *Bing Maps*²⁷ (da *Microsoft*), *MapQuest*, *Leaflet*²⁸ (da *CloudMade*) e *OpenLayers* (do projeto *OSGeo*²⁹). Tanto a API do *Google Maps* como Tanto a API do *Bing Maps*, são bibliotecas proprietárias que permitem apenas a utilização exclusiva de mosaicos³⁰ da *Google* e da *Microsoft*, deixando assim de fora outros fornecedores de mapas (como é o caso do *OpenStreetMap*) e tendem cada vez mais para um modelo de negócio de pesquisa paga³¹ (preço em função do número de visualizações por dia).

Por outro lado, a *Leaflet* e a *OpenLayers*, são bibliotecas genéricas e *Open Source* que visam o desenvolvimento colaborativo de *software* geoespacial, e promovem o seu uso

²⁶ <http://OpenLayers.org/>

²⁷ <http://www.bing.com/maps/>

²⁸ <http://leaflet.cloudmade.com/>

²⁹ <http://www.osgeo.org/>

³⁰ Do anglo-saxónico *tiles*.

³¹ Do anglo-saxónico *paid search*.

generalizado. Permitem a utilização de qualquer tipo de mosaicos e camadas de informação provenientes de diversas fontes.

Por estas razões, e devido à sua maturidade, popularidade e maior quantidade de em relação à Leaflet, foi escolhida a biblioteca *OpenLayers* para a implementação da Framework *SMaps*. A biblioteca *OpenLayers* (baseada nos *standards*³² da OGC) possui uma documentação bastante vasta e rivaliza com os seus concorrentes comerciais, com a grande vantagem de ser possível ver como é que o código funciona, os testes unitários, etc. (isto significa que facilmente se consegue construir novas funcionalidades por cima da API). É altamente extensível (implementada usando o paradigma orientado a objetos) e completamente livre de dependências do lado do servidor, simplificando o desenvolvimento das aplicações “cliente” muito mais fácil.

Apesar dos seus muitos pontos fortes, a *OpenLayers* apresentou os seguintes desafios a nível de implementação:

- Pouco suporte a HTML5 Canvas. O modo de apresentação da informação geoespacial incide principalmente na utilização de SVG;
- Falta de suporte às tecnologias da Web Semântica (por ex. *parsers* de RDF e OWL).

Para que a Framework *SMaps* seja concretizada como proposto, as seguintes funcionalidades têm de ser implementadas por cima da biblioteca *OpenLayers*:

- *Parsers* RDF, RDF-S e OWL. Estas linguagens podem estar codificadas em diferentes formatos como o XML, N3, entre outros. A informação pode também estar acessível através de um ponto de acesso SPARQL, podendo ser obtida em tempo real através de pesquisas;
- Uma linguagem própria (que amplie uma linguagem de representação de vetorial, como a SVG) que permita a especificação de objetos gráficos e a sua associação a conceitos ontológicos;
- *Parser* para a linguagem de especificação de objetos gráficos;
- Núcleo com algoritmos que interpretam a informação proveniente dos *parsers*, gerindo a interação do utilizador e a informação apresentada;
- Algoritmos para filtragem e apresentação de facetas:
 - De forma explícita e implícita;
 - De acordo com o histórico de navegação do utilizador;
 - De acordo com o perfil do utilizador.

³² <http://www.opengeospatial.org/standards/>

5.2 Protótipo Implementado

A implementação do protótipo *SMaps* possui parte das funcionalidades previamente mencionadas. A implementação do mesmo exigiu uma análise prévia da biblioteca *OpenLayers* e resultou na arquitetura apresentada na figura 19.

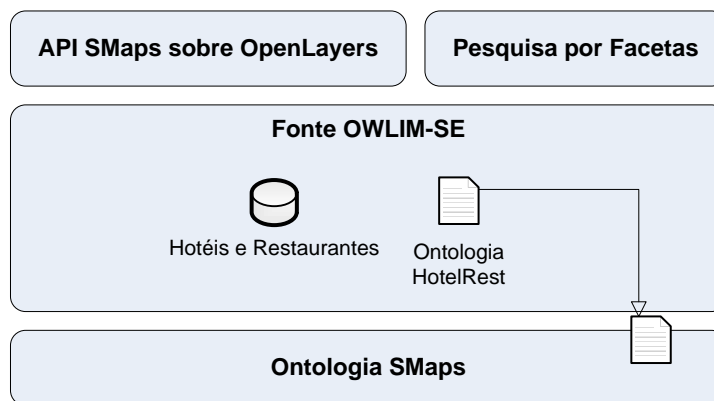


Figura 19 – Arquitetura do protótipo implementado.

O protótipo implementado é constituído por:

- Ontologia *SMaps*;
- Acesso a bases de conhecimento através de pontos de acesso SPARQL (como p. ex. a biblioteca Hercules³³);
- Uma base de conhecimento (em OWLIM-SE, um repositório semântico) com concelhos de Portugal e pontos de interesse do *Google Places*³⁴, descritos utilizando uma pequena ontologia (semelhante à apresentada nos exemplos utilizados no capítulo 4) baseada na ontologia *SMaps*;
- Apresentação externa de conteúdos;
- Pesquisa por facetas explícita.

O protótipo criado suporta apenas a estratégia de apresentação externa de conteúdos. Para a implementação desta estratégia, foi feita uma análise extensiva da biblioteca *OpenLayers*, de modo a incorporar as novas funcionalidades de forma transparente. A arquitetura estudada encontra-se ilustrada na figura 20.

³³ <http://www.arielworks.net/works/codeyard/hercules/demo/index.html>

³⁴ <http://www.Google.com/places/>

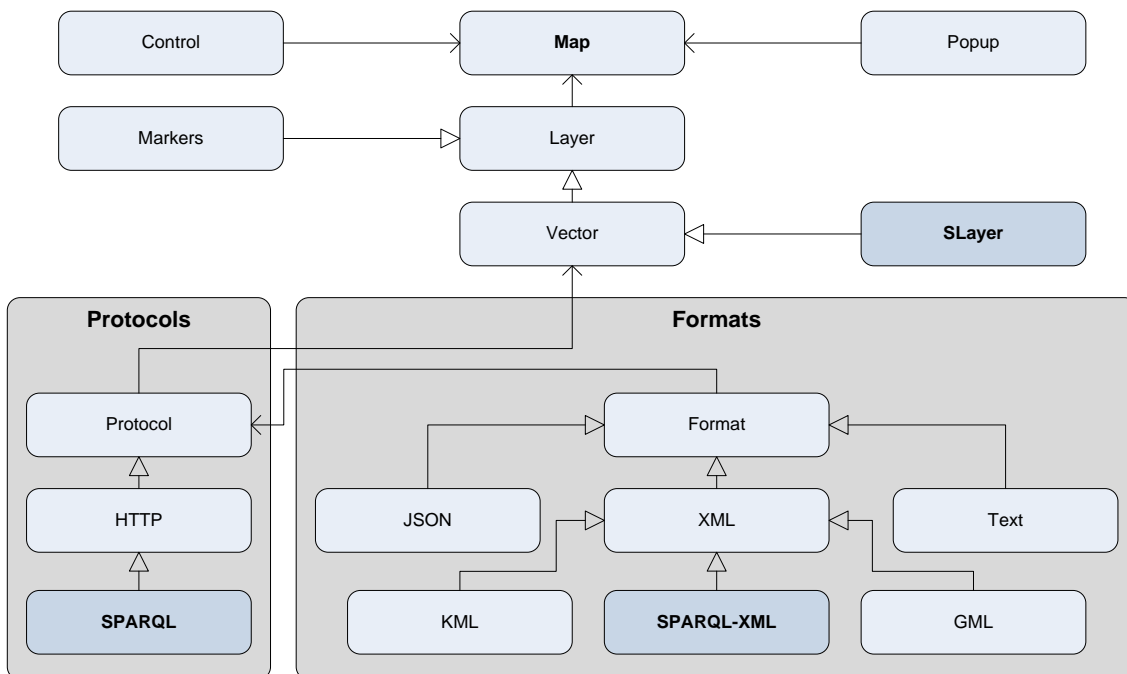


Figura 20 – Diagrama UML parcial do *OpenLayers* após a implementação do protótipo *SMaps*.

Após a implementação do protótipo, a biblioteca *OpenLayers* foi incrementada com novos componentes que se encontram ilustrados a azul na figura 20.

Para carregar informação da base de conhecimento através do ponto de acesso SPARQL, é necessário criar um *layer* do tipo *Vector* ou *SLayer*. Estes *layers* utilizam posteriormente um protocolo do tipo SPARQL para se ligar ao ponto de acesso.

Finalmente, para cada pesquisa SPARQL efetuada, a resposta será interpretada pelo formato SPARQL-XML que irá gerar instâncias de objetos geométricos compreendidas pelo *OpenLayers*, de acordo com a ontologia *SMaps*.

No que diz respeito à filtragem/pesquisa por facetas, estas são obtidas e apresentadas através do objeto *SLayer*, estando estas dependentes da pesquisa efetuada. Quando uma nova pesquisa é feita (por exemplo quando o utilizador clica numa das facetas), uma nova pesquisa é efetuada. Esta pesquisa poderá fazer com que as facetas apresentadas sejam alteradas, ou seja, se ajustem ao contexto atual.

As figuras 21, 22 e 23 apresentam a interface atual do protótipo. A interface apresenta dois tipos de navegação: por facetas (por baixo) e através do grafo (à direita). À direita do mapa, é possível navegar pelo grafo de objetos com caracterização geoespacial, através de cliques nas propriedades (p. ex., restaurantes, hotéis) que associam o objeto selecionado a outros objetos geoespaciais.

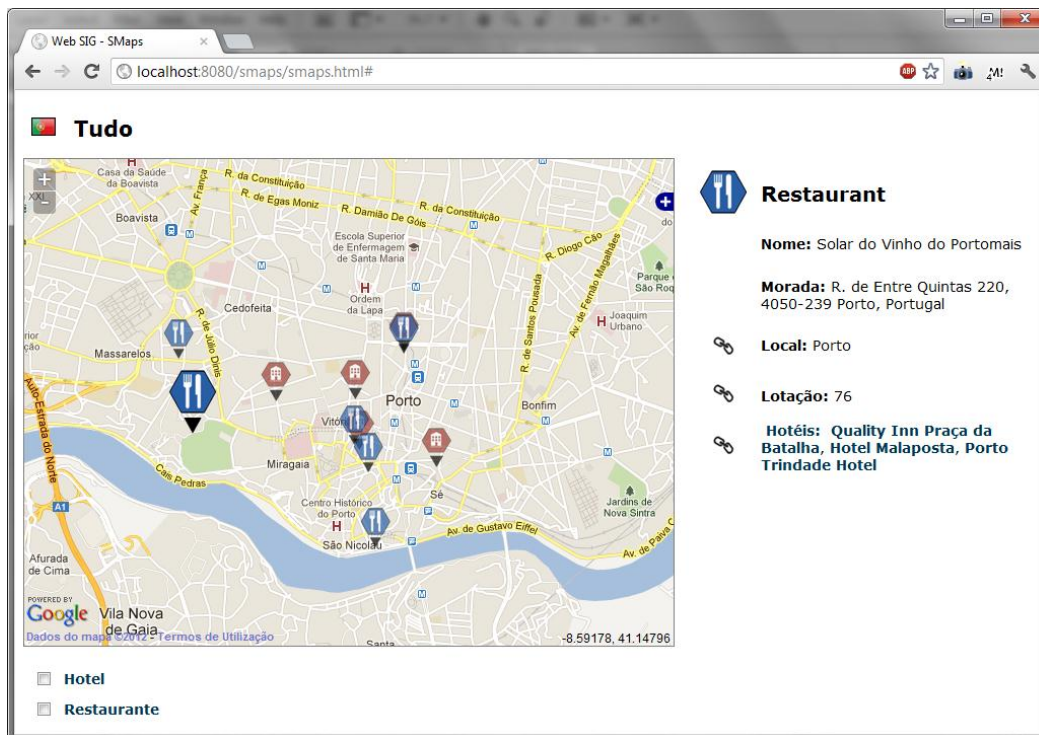


Figura 21 – Interface do protótipo: ecrã onde todos os objetos com caracterização são apresentados.

Relativamente à navegação por facetas, as facetas são atualizadas de acordo com o contexto em que o utilizador se encontra. Quando uma faceta é selecionada, as outras não podem colidir com o conjunto de facetas selecionado. Por exemplo, quando o utilizador selecionou a faceta “Hotel 4 ou + Estrelas”, a faceta “Hotel 4- Estrelas” desapareceu visto originarem em classes disjuntas na ontologia de hotéis e restaurantes (ver figura 23).

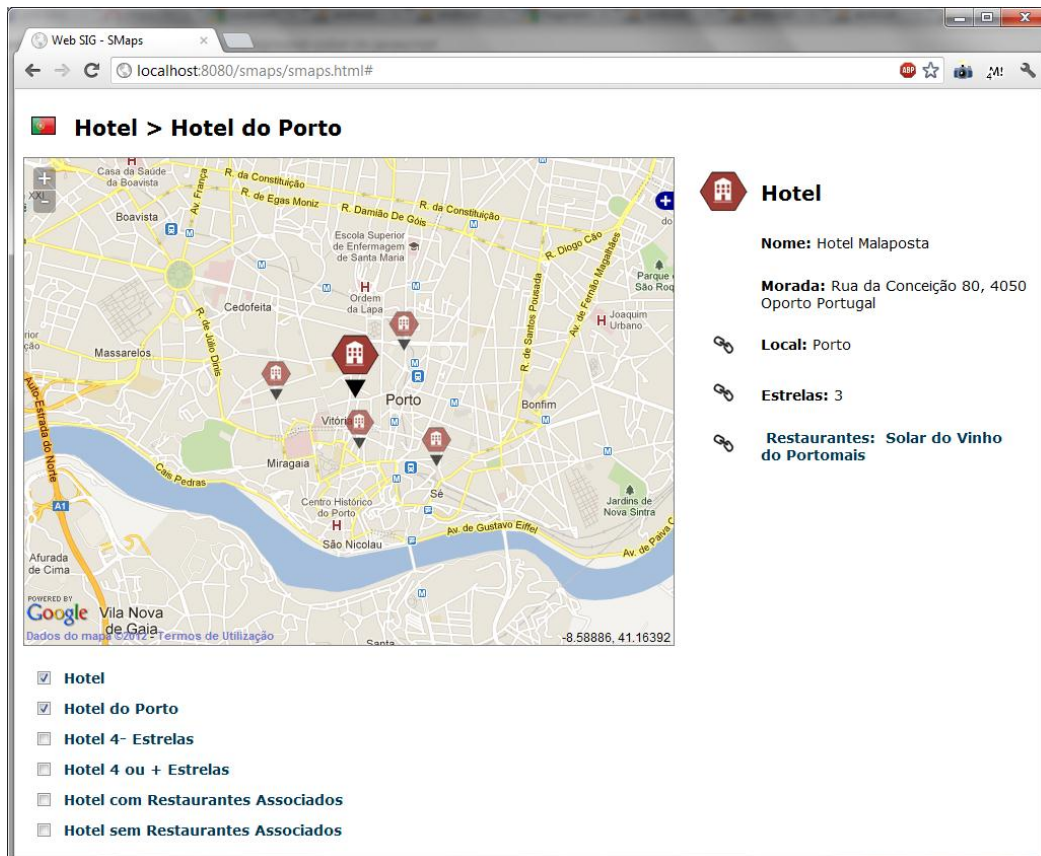


Figura 22 – Interface do protótipo: ecrã com filtragem de objetos por hotéis do Porto.

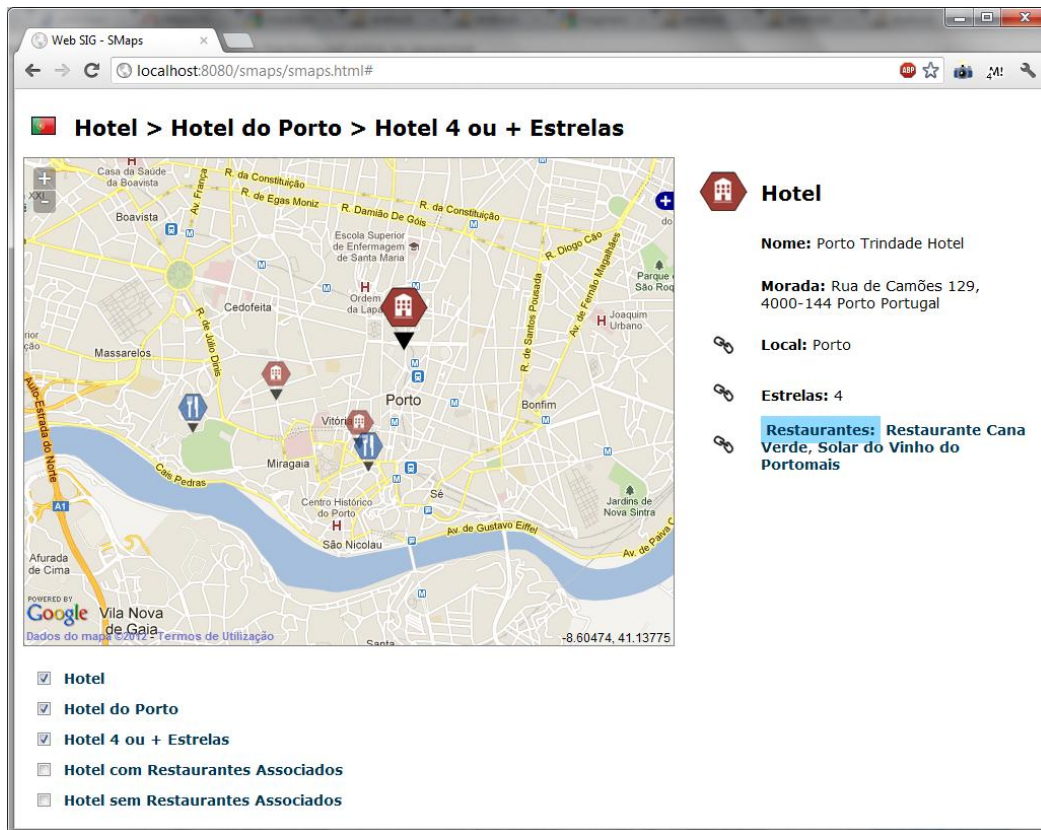


Figura 23 – Interface do protótipo implementado. Filtragem de objetos por hotéis do Porto com 4 ou mais estrelas, juntamente com apresentação dos restaurantes associados ao hotel selecionado.

Capítulo 6

Conclusões e Trabalho Futuro

Neste capítulo são apresentadas as conclusões relativamente à metodologia proposta, juntamente com as perspectivas de trabalho futuro. O capítulo termina com uma apreciação final sobre a dissertação.

6.1 Conclusões

A principal contribuição desta dissertação é a metodologia para geoportais semanticamente enriquecidos. Esta metodologia baseia-se na organização de recursos geográficos a nível semântico através de ontologias geográficas apropriadas, e exploração desta organização através da interface gráfica do geoportal.

Apesar de as dificuldades associadas à utilização de diversas tecnologias ainda com pouca maturidade, a metodologia proposta cumpre com os objetivos inicialmente enumerados, tendo para tal sido feito um estudo do atual do estado da arte a dois níveis: o dos SIG e o da Web Semântica. Após examinar vários geoportais em desenvolvimento e já em produção, chegou-se à conclusão de que o número daqueles que tiram partido da semântica é muito reduzido. Este estudo culminou com o desenvolvimento de um protótipo para disponibilização de serviços geoespaciais com base nos princípios e tecnologias da Web semântica.

Concluindo, esta dissertação contribui com os seguintes aspetos:

- Análise do estado da arte relativamente a geoportais e serviços geoespaciais;
- Análise do estado da arte relativamente à Web semântica, onde foram avaliados vários princípios e tecnologias posteriormente aplicadas na elaboração da metodologia e na implementação do protótipo;

- Definição de uma metodologia para geoportais e serviços geoespaciais semanticamente enriquecidos;
- Implementação de um protótipo que explora a metodologia proposta.

Com a crescente utilização de tecnologias da Web Semântica, esta metodologia desempenhará um papel importante na implementação de geoportais. Estas tecnologias irão proporcionar uma mudança de paradigma para um semanticamente enriquecido que irá ajudar o problema de interoperabilidade e heterogeneidade da informação. Paralelamente, irão também promover o aumento de qualidade da informação e a existência de bases de conhecimento de grande relevância.

Sendo assim, é visionada a aplicação da metodologia e *framework* desenvolvidas em diferentes domínios de conhecimento, tirando partido da interoperabilidade entre diferentes sistemas com diferentes bases de conhecimento.

6.2 Trabalho Futuro

Relativamente a trabalho futuro é necessário continuar a implementação do protótipo, mais concretamente, a estratégia de apresentação interna juntamente com os objetos gráficos extensíveis e a respetiva linguagem de especificação (para a criação desses mesmos objetos).

No que diz respeito à avaliação da *framework SMaps*, devido à simplicidade da interface do protótipo, não foi considerado que a elaboração dos testes nesta fase leva-se à recolha de resultados de qualidade no que diz respeito à importância e utilidade da *framework SMaps*. Neste sentido, como trabalho futuro, serão elaborados três inquéritos tendo em conta cada um dos seguintes aspetos:

- Usabilidade da interface;
- Facilidade de navegação (pelo grafo);
- Reconhecimento semântico (capacidade do utilizador reconhecer a semântica da informação na interface).

Depois de obter todos os resultados do anterior inquérito seria interessante, com base no trabalho efetuado nesta dissertação, elaborar um artigo científico para publicação.

Existem muitas perspetivas de desenvolvimento para a metodologia de geoportais semanticamente enriquecidos, a partir do atual trabalho. É de particular interesse a exploração das abordagens semanticamente enriquecidas de interface com o utilizador em vários domínios de aplicação, e a continuação do desenvolvimento do protótipo de forma a tirar o máximo proveito da semântica que se encontra por detrás dos serviços geoespaciais disponibilizados.

Referências

- Alameh, N. (2003). Service Chaining of Interoperable Geographic Information Web Services. *Evolution*.
- Baeza-Yates, R., & Ribeiro-Neto, B. (1999). User Interfaces and Visualization. *Modern Information Retrieval* (p. 65). ACM Press.
- Berners-Lee, T. (1998). Semantic Web Road. Retrieved from <http://www.w3.org/DesignIssues/Semantic.html>
- Berners-Lee, T., Hendler, J., & Lassila, O. (2006). The Semantic Web. *Scientific American Magazine*, 184, 5.
- Bhat, M., Shah, R., & Ahmad, B. (2011). Cloud Computing: A solution to Geographical Information Systems (GIS). *International*, 3(2), 594-600. Retrieved from <http://serviceorientedarchitecturesoa.net/goto/http://www.enggjournals.com/ijcse/doc/IJCS E11-03-02-006.pdf>
- Carver, S., Evans, A., & Kingston, R. (2000). Accessing Geographical Information Systems over the World Wide Web: Improving public participation in environmental decision-making. *Information Infrastructure and*, 6, 157-170. Retrieved from <http://iospress.metapress.com/index/n02wp95pd749tj2d.pdf>
- Chapin, T. S. (2003). Revolutionizing the core: GIS in the planning curriculum. *Environment and Planning B: Planning and Design*, 30(4), 565-573. doi:10.1068/b12993
- Drummond, N., Rector, A., & Stevens, R. (2006a). Putting OWL in order: Patterns for sequences in OWL. *OWLED ESWC* (pp. 1-10).
- Drummond, N., Rector, A., & Stevens, R. (2006b). Sequences in Protégé OWL. *9th Intl. Protégé Conference* (pp. 3-6).
- Drummond, N., Rector, A., & Stevens, R. (2006c). Putting OWL in order: Patterns for sequences in OWL. *OWLED ESWC* (pp. 1-10). Retrieved from <http://www.cs.man.ac.uk/~drummond/publications/OWLListsPaper/owl-lists-iswc.pdf>
- Drummond, N., Rector, A., & Stevens, R. (2006d). Sequences in Protégé OWL. *9th Intl. Protégé Conference* (pp. 3-6). Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Sequences+in+Protégé+OWL#0>
- Fugazza, C., & Luraschi, G. (2012). Semantics-Aware Indexing of Geospatial Resources Based on Multilingual Thesauri : Methodology and Preliminary. *International Journal of Spatial Data Infrastructures Research*, 7, 16-37. doi:10.2902/1725-0463.2012.07.art2

- Gaetan, G., Saldaño, V., Buccella, A., & Cechich, A. (2010). Fifth International Conference on Software Engineering Advances A Domain-Oriented Approach for GIS Component Selection. *A Domain-Oriented Approach for GIS Component Selection*. doi:10.1109/ICSEA.2010.22
- Gruber, T. (2009). Definition of Ontology. *of Database Systems*. Springer-Verlag. Retrieved from <http://tomgruber.org/writing/ontology-definition-2007.htm>
- Heydari, N., Mansourian, A., & Taleai, M. (1991). Ontology-based GIS web service for increasing semantic interoperability among organizations involving drilling in city of Tehran. *gsdi.org*. Retrieved from <http://www.gsdi.org/gsdi11/papers/pdf/163.pdf>
- Hildebrand, M., & Ossenbruggen, J. V. (2006). /facet: A browser for heterogeneous semantic web repositories. *The Semantic Web-ISWC*, 272-285. Springer-Verlag. Retrieved from <http://www.springerlink.com/index/d13t53k0022072t5.pdf>
- Jones, C. B., & Purves, R. S. (2007). Geographical information retrieval. *International Journal of Geographical Information Science*, 22(3), 219-228. doi:10.1080/13658810701626343
- Maguire, D. J., & Longley, P. A. (2005). The emergence of geoportals and their role in spatial data infrastructures. *Computers, Environment and Urban Systems*, 29(1), 3-14. doi:10.1016/j.compenvurbsys.2004.05.012
- Makela, E., Hyvonen, E., & Sidoroff, T. (2005). View-Based User Interfaces for Information Retrieval. *Information Retrieval*.
- Nikolaos, A., Kostas, K., & Michail, V. (2005). The emerge of semantic geoportals. *On the Move to Meaningful*, 1127-1136. Retrieved from <http://www.springerlink.com/index/92BN4U2EBVYRUB0E.pdf>
- OGC. (2012). Programs. Retrieved from <http://www.opengeospatial.org/projects>
- Oliveira, L., Rocha, A., Coelho, A., Rodrigues, A., Sousa, M., & Silva, D. (2012). Implementação de uma Infraestrutura de Dados Espaciais regional baseada em software livre. *7th Iberian Conference on Information Systems and Technologies* (p. 4).
- Scharl, A. (2007). Media Platforms for Managing Geotagged Knowledge Repositories. *Towards the Geospatial Web* (pp. 3-14).
- Sieber, R. (2008). Public Participation Geographic Information Systems: A Literature Review and Framework. *Information Systems*, (January 2012), 37-41.
- Silva, N. (2004). *MULTI-DIMENSIONAL SERVICE-ORIENTED ONTOLOGY MAPPING*. Universidade de Trás-os-Montes e Alto Douro.
- Suominen, O., Viljanen, K., & Hyvänen, E. (2007). User-centric faceted search for semantic portals. In E. Franconi, M. Kifer, & W. May (Eds.), *The Semantic Web: Research and Applications* (pp. 356-370). Springer. Retrieved from <http://www.springerlink.com/index/H8479834KN50N502.pdf>

- Tait, M. G. (2005). Implementing geoportals: applications of distributed GIS. *Computers, Environment and Urban Systems*, 29(1), 33-47. doi:10.1016/j.compenvurbsys.2004.05.011
- Turner, A., & Forrest, B. (2008). *Where 2.0: The State of the Geospatial Web*. Shakespeare (p. 55).
- Uschold, M. (2003). Where are the Semantics in the Semantic Web? *AI Magazine*, 24(3), 25-36. Retrieved from <http://www.aaai.org/ojs/index.php/aimagazine/article/viewArticle/1716>
- Vallea, E. D., & Qasima, H. (2010). Towards treating GIS as virtual RDF graphs. Retrieved from http://www.isprs.org/proceedings/XXXVIII/4-W13/ID_36.pdf

Anexo A

Manchester Syntax da Ontologia *SMaps*

Prefix: rdf: <<http://www.w3.org/1999/02/22-rdf-syntax-ns#>>

Prefix: : <<http://www.semanticmaps.pt/ontologies/2012/6/SMaps.owl#>>

Prefix: xsd: <<http://www.w3.org/2001/XMLSchema#>>

Prefix: xml: <<http://www.w3.org/XML/1998/namespace>>

Prefix: rdfs: <<http://www.w3.org/2000/01/rdf-schema#>>

Prefix: owl: <<http://www.w3.org/2002/07/owl#>>

Ontology: <<http://www.semanticmaps.pt/ontologies/2012/6/SMaps.owl>>

ObjectProperty: containsInterior

SubPropertyOf:

contains

ObjectProperty: containsExterior

SubPropertyOf:

contains

ObjectProperty: contains

ObjectProperty: next

DataProperty: *PointListInternal*

SubPropertyOf:

PointList

DataProperty: point

DataProperty: *PointList*

DataProperty: *PointListExternal*

SubPropertyOf:

PointList

DataProperty: long

DataProperty: lat

Class: *PointList*

Annotations:

 rdfs:label "Lista de Pontos"@pt,

 rdfs:comment "A class used to represent sequences of *PointList* instances. See OWLList class comments for more information on OWL Lists.",

 rdfs:label "Point List"

SubClassOf:

 next some *PointList*,

 OWLList,

 contains some Point

Class: MultiPolygon

Annotations:

 rdfs:label "MultiPolygon",

 rdfs:comment "A collection of polygons.",

 rdfs:label "Conjunto de Polígonos"@pt

SubClassOf:

 (contains some Polygon) or (*PointList* some rdfs:Literal),

 GeomCollection

Class: MultiLineString

Annotations:

rdfs:label "Conjunto de Linhas"@pt,
rdfs:label "MultiLineString",
rdfs:comment "A collection of line strings."

SubClassOf:

GeomCollection,
(contains some LineString) or (*PointList* some rdfs:Literal)

Class: MultiPoint

Annotations:

rdfs:comment "A collection of two-dimensional points.",
rdfs:label "MultiPoint",
rdfs:label "Conjunto de Pontos"@pt

SubClassOf:

GeomCollection,
(contains some Point) or (point some rdfs:Literal)

Class: Polygon

Annotations:

rdfs:label "Polygon",
rdfs:comment "A polygon specified through multiple sequences of points. This sequence can be set through the containsInternal/containsExternal property, using instances of *PointList* (with enriched semantics), or through the *PointListInternal/PointListExternal* datatype property, using simple strings.",

rdfs:label "Polígono"@pt

SubClassOf:

(contains some *PointList*) or (*PointList* some rdfs:Literal),

SpatialThing

Class: LineString

Annotations:

rdfs:label "Linha"@pt,

rdfs:comment "A line string specified through a sequence of points. This sequence can be set through the contains property, using instances of *PointList* (with enriched semantics), or through the *PointList* datatype property, using a simple string.",

 rdfs:label "LineString"

SubClassOf:

 (contains exactly 1 *PointList*) or (*PointList* exactly 1 rdfs:Literal),

 SpatialThing

Class: GeomCollection

Annotations:

 rdfs:label "Conjunto de Geometrias"@pt,

 rdfs:label "Geometry Collection",

 rdfs:comment "A collection of spatial things."

SubClassOf:

 contains some SpatialThing,

 SpatialThing

Class: SpatialThing

Annotations:

 rdfs:comment "Any spatial thing. This class represents the set of instances that contain a spatial representation; can be visualized over a map.",

 rdfs:label "Spatial Thing"

DisjointWith:

 OWLList

Class: OWLList

Annotations:

 rdfs:comment "A class used to represent sequences in OWL. Usage is similar to those of linked lists, where there is a head (specified through the contains property) and a tail (specified through the next property).",

 rdfs:label "OWL List",

 rdfs:label "Lista OWL"@pt

SubClassOf:

next some owl:Thing,
contains some owl:Thing

DisjointWith:
SpatialThing

Class: Point

Annotations:

rdfs:label "Ponto"@pt,
rdfs:comment "A two-dimensional point in space defined through latitude and
longitude.",

rdfs:label "Point"

SubClassOf:

lat exactly 1 xsd:decimal,
SpatialThing,
long exactly 1 xsd:decimal