

Universidade Federal de Campina Grande  
Centro de Engenharia Elétrica e Informática  
Coordenação de Pós-Graduação em Ciência da Computação

Anotação Automática de Dados Geográficos  
Baseada em Bancos de Dados Abertos e Interligados

Hamon Barros Henriques  
(Mestrando)

Cláudio de Souza Baptista, Ph. D.  
(Orientador)

Campina Grande, Paraíba, Brasil

© Hamon Barros Henriques, 31/08/2015

Universidade Federal de Campina Grande  
Centro de Engenharia Elétrica e Informática  
Coordenação de Pós-Graduação em Ciência da Computação

Anotação Automática de Dados Geográficos  
Baseada em Bancos de Dados Abertos e Interligados

Hamon Barros Henriques

Dissertação submetida à Coordenação do Curso de Pós-Graduação em  
Ciência da Computação da Universidade Federal de Campina Grande -  
Campus I como parte dos requisitos necessários para obtenção do grau  
de Mestre em Ciência da Computação.

Área de Concentração: Ciência da Computação  
Linha de Pesquisa: Sistemas de Informações Geográficas

Cláudio de Souza Baptista, Ph. D.  
(Orientador)

Campina Grande, Paraíba, Brasil  
©Hamon Barros Henriques, 31/08/2015

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCG

H519a      Henriques, Hamon Barros.  
              Anotação automática de dados geográficos baseada em bancos de dados abertos e interligados / Hamon Barros Henriques. – Campina Grande, 2015.  
              107 f. : color.

              Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática.

              "Orientação: Ph. D. Cláudio de Souza Baptista".  
              Referências.

              1. Anotação Semântica. 2. Ontologia. 3. Camada. 4. Atributo.  
              I. Baptista, Cláudio de Souza. II. Título.

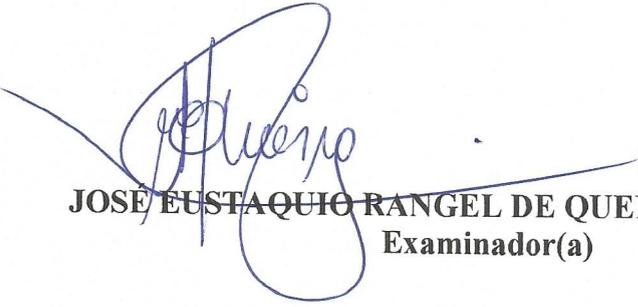
CDU 004.65(043)

**"ANOTAÇÃO AUTOMÁTICA DE DADOS GEOGRÁFICOS BASEADA EM BANCOS DE  
DADOS ABERTOS E INTERLIGADOS"**

**HAMON BARROS HENRIQUES**

**DISSERTAÇÃO APROVADA EM 31/08/2015**

*Claudio de Souza Baptista*  
**CLAUDIO DE SOUZA BAPTISTA, PhD., UFCG**  
**Orientador(a)**

  
**JOSE EUSTAQUIO RANGEL DE QUEIROZ, D.Sc, UFCG**  
**Examinador(a)**

*Hugo Feitosa de Figueiredo*  
**HUGO FEITOSA DE FIGUEIREDO, Dr., IFPB**  
**Examinador(a)**

*Fabio Gomes de Andrade*  
**FABIO GOMES DE ANDRADE, D.Sc., IFPB**  
**Examinador(a)**

**CAMPINA GRANDE - PB**

## Resumo

Recentemente, infraestruturas de dados espaciais (IDE) têm se popularizado como uma importante solução para facilitar a interoperabilidade de dados geográficos oferecidos por diferentes organizações. Um importante desafio que precisa ser superado por estas infraestruturas consiste em permitir que seus clientes possam localizar facilmente os dados e serviços que se encontram disponíveis. Atualmente, esta tarefa é implementada a partir de serviços de catálogo. Embora tais serviços tenham representado um importante avanço para a recuperação de dados geográficos, estes ainda possuem limitações importantes. Algumas destas limitações surgem porque os serviços de catálogo resolvem suas consultas com base nas informações contidas em seus registros de metadados, que normalmente descrevem as características do serviço como um todo. Além disso, muitos catálogos atuais resolvem consultas com restrições temáticas apenas com base em palavras-chaves, e não possuem meios formais para descrever a semântica dos recursos disponíveis. Para resolver a falta de semântica, esta dissertação apresenta uma solução para a anotação semântica automática das camadas e dos seus respectivos atributos disponibilizados em uma IDE. Com isso, motores de busca, que utilizam ontologias como insumo para a resolução de suas consultas, irão encontrar os dados geográficos que estão relacionados semanticamente a um determinado tema pesquisado. Também foi descrita nesta pesquisa uma avaliação do desempenho da solução proposta sobre uma amostra de serviços *Web Feature Service*.

## **Abstract**

Recently, Spatial Data Infrastructure (SDI) has become popular as an important solution for easing the interoperability of geographic data offered by different organizations. An important challenge that must be overcome by such infrastructures consists in allowing their users to easily locate the available data and services. Presently, this task is implemented by means of catalog services. Although such services represent an important advance for retrieval of geographic data, they still have serious limitations. Some of these limitations arise because the catalog service resolves their queries based on information contained in their metadata records, which normally describes the characteristics of the service as a whole. In addition, many current catalogs solve queries with thematic restrictions based only on keywords, and have no formal means for describing the semantics of available resources. To resolve the lack of semantics, this dissertation presents a solution for automatic semantic annotation of feature types and their attributes available in an IDE. With this, search engines, which use ontologies as input for solving their queries will find the geographic data that are semantically related to a particular topic searched. Also has described in this research an evaluation of the performance of the proposed solution on a sample of Web Feature Service services.

## **Agradecimentos**

Agradeço primeiramente a Deus por ter me dado força e fé para vencer mais um objetivo na minha vida.

Aos meus familiares, em especial, meu pai, Adenilson, pelo exemplo de ser humano, no qual me espelho e me orgulho; a minha mãe, Edilena, pelo esforço e dedicação na educação de todos os seus filhos; a minha esposa, Daiana, pelo amor, dedicação, incentivo e paciência durante o período desta pesquisa. Aos meus irmãos, Hamana, Halline, Harisson e Amanda, pelas fontes de motivação, fraternidade e alegria que me proporcionam.

Ao meu orientador, Cláudio Baptista, por ter aceitado meu convite e me orientado neste programa, transmitindo sua experiência e conhecimento. Ao professor Fabio Andrade, que não mediu esforços para repassar todo o seu conhecimento na área. Ambos proporcionaram aprendizados que levarei para minha vida pessoal e profissional.

Aos meus colegas de mestrado, Arthur Marques, Anderson Felinto, George Alves e Ruan Pierre, pela ajuda e compartilhamento de conhecimento neste período.

Aos meus amigos e ex-colegas de trabalho do IFPB, Fernando Gurjão, Marx Medeiros e Marco Cunha, pelo apoio e incentivo para que eu conseguisse conciliar o trabalho e as disciplinas do mestrado.

À COPIN, pela oportunidade de cursar o mestrado e pelo atendimento aos meus pedidos de prorrogação. Afinal, não é nada fácil conciliar trabalho com mestrado e manter os prazos.

A todas as pessoas que contribuíram para a conclusão desta pesquisa.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Objetivos da Pesquisa . . . . .	5
1.1.1	Geral: . . . . .	5
1.1.2	Específicos: . . . . .	5
1.2	Estrutura da Dissertação . . . . .	5
<b>2</b>	<b>Fundamentação Teórica</b>	<b>6</b>
2.1	Web Semântica . . . . .	6
2.2	Ontologia . . . . .	9
2.3	Serviços OGC . . . . .	13
2.3.1	CSW ( <i>Catalog Service for the Web</i> ) . . . . .	15
2.3.2	WMS ( <i>Web Map Service</i> ) . . . . .	19
2.3.3	WFS ( <i>Web Feature Service</i> ) . . . . .	21
2.4	IDE (Infraestrutura de Dados Espaciais) . . . . .	24
2.5	Anotação Semântica . . . . .	25
2.6	Linked Open Data . . . . .	26
2.7	DBpedia . . . . .	29
2.8	Ontologia YAGO . . . . .	31
2.9	Considerações Finais do Capítulo . . . . .	32
<b>3</b>	<b>Pesquisas Relacionadas</b>	<b>33</b>
3.1	Yingjie et al. . . . .	33
3.2	Oliveira et al. . . . .	34
3.3	Fugazza et al. . . . .	35

---

3.4	Gimenez et al. . . . .	36
3.5	Borzacchiello e Craglia . . . . .	37
3.6	Lutz et al. . . . .	37
3.7	Janowicz et al. . . . .	38
3.8	Li et al. . . . .	39
3.9	Macário et al. . . . .	41
3.10	Batcheller e Reitsma . . . . .	42
3.11	Lutz e Kolas . . . . .	43
3.12	Andrade et al. . . . .	44
3.13	Considerações Finais do Capítulo . . . . .	44
<b>4</b>	<b>Abordagem Proposta</b>	<b>46</b>
4.1	Modelagem de Dados . . . . .	46
4.1.1	Esquema Conceitual . . . . .	46
4.1.2	Coleta dos Dados . . . . .	50
4.2	Anotação Semântica no Nível de Camadas . . . . .	53
4.2.1	Geração do saco de palavras das informações textuais das camadas	54
4.2.2	Recuperação de classes e categorias associadas à camada na DBpedia	55
4.2.3	Seleção das classes e categorias similares à camada . . . . .	57
4.2.4	Seleção dos conceitos relacionados semanticamente com a camada .	58
4.3	Anotação Semântica no Nível de Atributos . . . . .	59
4.3.1	Seleção dos atributos relacionados com a camada . . . . .	60
4.3.2	Seleção dos conceitos relacionados semanticamente com cada atributo	60
4.3.3	Persistência do mapeamento entre o atributo e os conceitos . . . . .	60
4.4	Refinamento da Anotação Semântica . . . . .	61
4.5	Considerações Finais do Capítulo . . . . .	62
<b>5</b>	<b>Avaliação Experimental</b>	<b>64</b>
5.1	Visão Geral da Avaliação Experimental . . . . .	64
5.2	Geração do Banco de Dados . . . . .	65
5.3	Anotação Semântica no Nível de Camadas . . . . .	66
5.4	Anotação Semântica no Nível de Atributos . . . . .	71

---

5.5	Considerações Finais do Capítulo . . . . .	74
<b>6</b>	<b>Considerações Finais</b>	<b>76</b>
6.1	Contribuições . . . . .	77
6.2	Resultados Obtidos . . . . .	78
6.3	Sugestões para Pesquisas Futuras . . . . .	78
	<b>Referências Bibliográficas</b>	<b>80</b>
<b>A</b>	<b>Linha de Base para a Anotação no Nível de Camadas</b>	<b>86</b>
<b>B</b>	<b>Linha de Base para a Anotação no Nível de Atributos</b>	<b>94</b>
<b>C</b>	<b>Resultado da Anotação no Nível de Camadas</b>	<b>99</b>
<b>D</b>	<b>Resultado da Anotação no Nível de Atributos</b>	<b>103</b>
<b>E</b>	<b>URL dos Serviços da Amostra</b>	<b>107</b>

# Lista de Símbolos

CSDI	<i>Canadas Spatial Data Infrastructure</i>
CSW	<i>Catalog Service for the Web</i>
DAML	<i>DARPA Agent Markup Language</i>
DL	<i>Description Logic</i>
DTD	<i>Document Type Definition</i>
ebRIM	<i>ebXML Registry Information Model</i>
GML	<i>Geography Markup Language</i>
GEOSS	<i>Global Earth Observation System of Systems</i>
INDE	<i>Infraestrutura Nacional de Dados Espaciais</i>
IDE	<i>Infrastrukturade Dados Espaciais</i>
IR	<i>Information Retrieval</i>
ISO	<i>International Organization for Standardization</i>
LOD	<i>Linked Open Data</i>
NSDI	<i>National Spatial Data Infrastructure</i>
OGC	<i>Open Geospatial Consortium</i>
OIL	<i>Ontology Inference Layer</i>
OWL	<i>Web Ontology Language</i>
OWL-S	<i>Web Ontology Language for Services</i>
RDF	<i>Resource Description Framework</i>
RIF	<i>Rule Interchange Format</i>
RuleML	<i>Rule Markup Language</i>
SDI	<i>Spatial Data Infratructures</i>
SVG	<i>Scalable Vector Graphics</i>
SWRL	<i>Semantic Web Rule Language</i>

URI	<i>Uniform Resource Identifier</i>
URL	<i>Uniform Resource Locator</i>
XML	<i>Extensible Markup Language</i>
WCS	<i>Web Coverage Service</i>
WFS	<i>Web Feature Service</i>
WFS-T	<i>Web Feature Service Transactional</i>
WMS	<i>Web Map Service</i>
W3C	<i>World Wide Web Consortium</i>

# Lista de Figuras

2.1	Camadas de Web Semântica. Fonte: Fensel et al. (2003) . . . . .	7
2.2	Organização dos serviços OGC. . . . .	14
2.3	Exemplo da operação <i>GetCapabilities</i> do serviço CSW. . . . .	15
2.4	Exemplo da operação <i>GetRecords</i> do CSW. . . . .	16
2.5	Exemplo da operação <i>GetRecordById</i> do serviço CSW. . . . .	18
2.6	Exemplo da operação <i>GetCapabilities</i> do serviço WMS. . . . .	19
2.7	Exemplo da operação <i>GetMap</i> do serviço WMS. . . . .	20
2.8	Mapa gerado pela operação <i>GetMap</i> do WMS. . . . .	21
2.9	Exemplo da operação <i>GetCapabilities</i> do serviço WFS. . . . .	22
2.10	Exemplo da operação <i>DescribeFeatureType</i> do serviço WFS. . . . .	23
2.11	Dados publicados no projeto LOD até setembro de 2011. . . . .	27
2.12	Exemplo de conceito disponibilizado pela <i>DBpedia</i> . . . . .	29
2.13	Exemplo de requisição ao serviço de busca da <i>DBpedia</i> . . . . .	30
4.1	Esquema Conceitual da Abordagem Proposta . . . . .	47
4.2	Modelagem Lógica Relacional da Abordagem Proposta . . . . .	48
4.3	Fluxograma de alimentação da tabela “ <i>Serviço</i> ” . . . . .	50
4.4	Fluxograma de alimentação das tabelas “ <i>Camada</i> ” e “ <i>Atributo</i> ” . . . . .	51
4.5	Fluxograma de alimentação da tabela “ <i>Conceito</i> ” . . . . .	52
4.6	Fluxograma de Alimentação da Tabela “ <i>Anotação Camada</i> ” . . . . .	52
4.7	Fluxograma de Alimentação da Tabela “ <i>Anotação Atributo</i> ” . . . . .	53
4.8	Exemplo de requisição ao serviço de busca da <i>DBpedia</i> . . . . .	56
5.1	Exemplo de um registro de <i>log</i> na anotação de camadas . . . . .	68
5.2	Resultado da precisão na anotação de camadas . . . . .	68

---

5.3	Resultado da cobertura na anotação de camadas . . . . .	69
5.4	Quantidade de anotações X tipo de anotação . . . . .	70
5.5	Exemplo de um registro de <i>log</i> na anotação de atributos . . . . .	72
5.6	Resultado da precisão na anotação de atributos . . . . .	73
5.7	Resultado da cobertura na anotação de atributos . . . . .	74

# Lista de Tabelas

3.1	Comparação entre as pesquisas apresentados. . . . .	45
4.1	Comparação entre as pesquisas apresentadas e a pesquisa desta dissertação. . . . .	63
A.1	Linha de Base para a anotação no nível de camadas. . . . .	86
B.1	Linha de base para a anotação no nível de atributos. . . . .	94
C.1	Resultado da anotação semântica da linha de base das camadas . . . . .	99
D.1	Resultado da anotação semântica da linha de base dos atributos . . . . .	103
E.1	URL dos serviços da amostra . . . . .	107

# Lista de Códigos Fonte

2.1	Exemplo de uso do XML . . . . .	8
2.2	Exemplo de utilização de OWL . . . . .	12
2.3	Metadados do CSW do <i>British Geological Survey</i> . . . . .	15
2.4	Um dos registros retornado pela operação <i>GetRecords</i> do CSW. . . . .	17
2.5	Resultado da operação <i>GetRecordById</i> do serviço CSW. . . . .	18
2.6	Descrição de uma camada no serviço WMS. . . . .	20
2.7	Descrição de uma camada em um serviço WFS. . . . .	22
2.8	Descrição do esquema de uma camada no serviço WFS. . . . .	23
2.9	Exemplo de registro retornado pelo serviço de busca da <i>DBpedia</i> . . . . .	30
4.1	Exemplo de registro retornado pelo serviço de busca da <i>DBpedia</i> . . . . .	56

# Capítulo 1

## Introdução

O crescente aumento do uso de dados geográficos para os mais variados tipos de necessidades (gerenciamento de catástrofes, monitoramento de bacias hidrográficas, monitoramento de florestas, distâncias entre determinadas localidades, entre outras) proporcionou o surgimento de diversos sistemas de informações geográficas (SIG). As aplicações desenvolvidas eram heterogêneas, uma vez que os dados e aplicações produzidos eram voltados para atender apenas a demanda das organizações. Esta heterogeneidade levava a problemas de interoperabilidade dos dados, tornando difícil para uma organização reutilizar dados geográficos que já tinham sido produzidos por outras agências.

O problema da heterogeneidade dos dados geográficos tem sido amenizado com os padrões especificados pelo *Open Geospatial Consortium* [1], que é um consórcio composto por mais de quatrocentas organizações de vários tipos, como empresas privadas, organizações governamentais, organizações sem fins lucrativos e instituições de pesquisa. O OGC tem como um dos seus principais objetivos a especificação de padrões públicos e abertos para o domínio geoespacial, almejando com isso, o aumento da interoperabilidade entre os dados e aplicações existentes. Dentre os padrões propostos pelo OGC, está um conjunto de serviços que possibilitam o acesso uniforme a dados geográficos em diferentes formatos. Alguns destes serviços disponibilizam dados geográficos na forma de camadas, também chamadas de *feature types*.

Os serviços propostos pelo OGC facilitaram o acesso aos dados geográficos disponibilizados por diferentes organizações, porém, não resolveram o problema da recuperação e integração destes dados. Isto porque os padrões propostos estão voltados apenas para as

questões sintáticas referentes ao acesso, não especificando nenhuma regra para descrever a semântica dos dados providos pelo serviço. Para tentar resolver questões pendentes relativas à interoperabilidade de dados e serviços, foram desenvolvidas as Infraestruturas de Dados Espaciais (IDE). Essas infraestruturas são compostas por padrões, políticas e tecnologias para promover o acesso de dados geográficos oferecidos por diferentes organizações [2]. Vários estudos sobre IDE têm sido apresentados na literatura como, por exemplo, as pesquisas de Cooper et al.[3] e Oliveira e Filho [4].

Para melhorar a recuperação de dados geográficos, as IDE atuais normalmente oferecem um serviço de catálogo, que é utilizado pelos provedores de dados geográficos e pelos clientes. Os provedores disponibilizam no serviço de catálogo os serviços que provêm os seus conjuntos de dados. Por outro lado, os clientes consultam o serviço de catálogo para localizar os dados geográficos de que necessitam. Embora tenham facilitado a localização de dados geográficos, os serviços de catálogo atuais ainda possuem algumas limitações. Uma destas limitações acontece porque a maioria desses catálogos resolve consultas com restrições temáticas (a partir de um determinado tema como relevo, por exemplo) apenas com base em palavras-chaves. Esta característica faz com que recursos relevantes que são descritos com termos relacionados ao tema procurado sejam descartados durante o processo de recuperação da informação, reduzindo assim, a cobertura destas consultas. Por exemplo, suponha-se uma consulta na qual o cliente do serviço de catálogo realiza uma consulta pelo tema “*Corpos Hídricos*”. Nesse caso, um serviço descrito pelas palavras-chaves açudes, rios e lagos é considerado irrelevante, uma vez que este não possui a palavra-chave usada na consulta em sua descrição.

Uma forma de superar esta limitação consiste no uso de **ontologias** [5] para enriquecer a descrição da semântica dos dados disponíveis e das consultas dos clientes. **Ontologia** pode ser definida como uma especificação formal e explícita de uma conceitualização compartilhada [6]. Tal conceitualização é obtida mediante uma linguagem formal, possibilitando a compreensão tanto por humanos quanto por aplicações de *software*. O uso de **ontologias** permite que motores de busca possam compreender a semântica dos dados oferecidos e das consultas dos clientes, o que permite melhorar de forma significativa a qualidade dos resultados das consultas. Esta melhoria é possível porque os motores de busca podem compreender os relacionamentos semânticos existentes entre o tema requisitado pelo cliente e o

tema usado para a descrição do serviço. Entretanto, este tipo de solução requer que a descrição dos dados disponíveis seja associada a conceitos definidos em **ontologias**, mediante um processo chamado de **anotação semântica**. Tomando o exemplo anterior para ilustração de uma **ontologia**, os conceitos *rios*, *lagos*, *açudes* e *corpos hídricos* estariam relacionados a um conceito maior denominado *hidrologia*, em que este conceito abrangeria os demais conceitos. Dessa forma, se o cliente do serviço de catálogo procurar por mapas de açudes, será possível recuperar serviços que possuem em suas informações textuais temas como *corpos hídricos*, *rios* e *lagos*. Neste caso, o conceito “açudes” está semanticamente relacionado com os demais conceitos na **ontologia** utilizada durante a consulta temática. Sendo assim, o uso de **ontologias** para descrever a semântica das camadas oferecidos pelo serviço melhora a qualidade dos resultados das consultas com restrições temáticas.

Uma possível solução para este problema seria delegar esta tarefa para os provedores de dados geográficos. Nesse tipo de solução, durante o cadastro de um novo serviço em um serviço de catálogo, os provedores teriam que realizar o processo de **anotação semântica** manualmente. Tal solução seria inviável, uma vez que a **anotação semântica** é uma tarefa normalmente custosa e o volume de dados para ser anotado normalmente é grande. Ademais, os serviços que já estão disponibilizados precisariam ser atualizados, o que seria uma tarefa bastante árdua. A limitação deste tipo de solução faz surgir a necessidade de uma estratégia que consiga realizar a **anotação semântica** dos dados geográficos de forma automatizada ou semiautomatizada.

Outra questão importante relacionada à **anotação semântica** de dados geográficos diz respeito ao nível de detalhe em que os dados são anotados. Em várias pesquisas existentes na literatura, a **anotação semântica** feita é realizada em nível de camadas, sem considerar a semântica dos seus atributos. Tal limitação inviabiliza a recuperação de camadas com um nível de detalhamento maior. Por exemplo, suponha-se que o cliente esteja interessado em informações sobre as extensões dos rios de uma determinada região. Se a **anotação semântica** for realizada apenas em nível de camadas, o motor de busca apresenta para o cliente todos as camadas que oferecem dados sobre rios. Feito isso, o usuário precisa identificar, dentre todos os recursos retornados, aqueles que oferecem informação sobre a extensão, que é o objetivo da sua consulta. Esta tarefa pode ser custosa se a quantidade de recursos retornados para o cliente for grande.

---

A **anotação semântica** no nível de atributos de uma camada é possível apenas para camadas oferecidas mediante serviços WFS, pois apenas estes proporcionam a recuperação dos atributos e características existentes em uma determinada camada ou de uma lista de camadas disponíveis. Para facilitar o entendimento, suponha-se um serviço WFS que contém dados geográficos sobre a rodovia BR-101, que percorre boa parte do litoral brasileiro. Este serviço disponibiliza a camada *municípios*. Esta camada possui, por sua vez, atributos como nome, população, UF, extensão e ano de emancipação. Com a anotação no nível de atributo, é possível realizar a **anotação semântica** de cada um dos atributos desta camada e, com isso, realizar consultas mais detalhadas sobre o mesmo. Um exemplo de consulta possível seria identificar entre os municípios que são cruzados pela rodovia BR-101, os que possuem mais de cinquenta mil habitantes e que foram emancipados após o ano 1980.

Nesta pesquisa, propõe-se uma **anotação semântica** automática de camadas que utiliza **ontologias** atreladas aos *links* semânticos propostos pelo *Linked Open Data* [7]. Com isso, promove-se a interoperabilidade e o reuso dos dados geográficos. A proposta de LOD proporciona um grande potencial ao conectar recursos informacionais mediante *links* semânticos que são compreensíveis também por máquinas. Dessa forma, tais *links* podem ser processados por programas de forma a explorar e enriquecer cognitivamente o significado da ligação entre dois recursos.

Para validar a solução proposta por esta pesquisa foi utilizada uma amostra de serviços WFS que contém informações textuais (nome, descrição e palavras-chaves) cadastradas na língua inglesa. Esta restrição foi necessária para garantir a viabilidade da solução no idioma universal. Porém, a solução também pode ser expandida para outros idiomas. Para cada serviço, usado no processo de validação, foram identificadas as suas camadas com os seus respectivos atributos. Após a coleta dos dados geográficos, o algoritmo de **anotação semântica** foi executado, identificando conceitos relacionados a cada um deles. Em seguida, o desempenho da **anotação semântica** foi mensurado utilizando duas métricas bem conceituadas na avaliação de desempenho dos sistemas de recuperação da informação: cobertura e precisão.

## 1.1 Objetivos da Pesquisa

### 1.1.1 Geral:

Desenvolver uma solução para realizar a **anotação semântica** de forma automática das camadas e seus respectivos atributos disponibilizados em uma IDE.

### 1.1.2 Específicos:

- Propor uma estratégia para a anotação automática de dados espaciais em nível de camadas que utiliza o LOD e a *Wikipedia* para incorporar semântica entre os conceitos da **ontologia** e os dados espaciais;
- Propor uma estratégia para a anotação automática de dados espaciais em nível de atributos que utiliza os nomes dos atributos para selecionar conceitos relevantes da **ontologia**;

Diante do que foi exposto neste capítulo foram levantadas as seguintes questões de pesquisa:

- É possível identificar características semânticas nas informações textuais disponibilizadas nos serviços e suas camadas?
- É possível anotar semanticamente camadas de forma automática?
- É possível anotar semanticamente atributos de uma camada de forma automática?

## 1.2 Estrutura da Dissertação

Esta dissertação é composta por seis capítulos. No Capítulo 2, são apresentados os principais conceitos e tecnologias que fundamentam esta pesquisa. No Capítulo 3, são discutidas as principais pesquisas relacionadas existentes na literatura. No Capítulo 4, é descrita a abordagem proposta para a resolução do problema, apresentando a modelagem de dados utilizada e os processos de anotações semânticas nos níveis de camadas e de atributo. No Capítulo 5, é descrita a avaliação experimental conduzida para a validação. Finalmente, no Capítulo 6, são apresentadas as conclusões e propostos temas derivados para pesquisas futuras.

# Capítulo 2

## Fundamentação Teórica

Neste capítulo, são abordados os principais conceitos e tecnologias utilizados nesta dissertação. O objetivo do capítulo é fornecer uma discussão inicial sobre cada tema relacionado à dissertação. Entretanto, para cada tema discutido, o capítulo descreve referências para locais onde o leitor poderá encontrar uma discussão detalhada do tema. Inicialmente, é apresentada a conceitualização de **ontologia**. Em seguida, os seguintes temas são discutidos: *web* semântica, serviços *Open Geospatial Consortium* (OGC), Infraestrutura de Dados Espaciais (IDE) e **Anotação Semântica**. Por fim, é feita uma apresentação do *Linked Open Data* (LOD).

### 2.1 Web Semântica

O conceito da *web* semântica foi apresentado inicialmente por Lee et al. [8] em seu artigo “*The Semantic Web*”. Em sua pesquisa, os autores pretendiam fazer com que agentes de *software* conseguissem entender os dados publicados na *web* e realizar conexões semânticas entre eles, com o objetivo de facilitar a recuperação e o uso dos mesmos. Para que este entendimento fosse possível pelos agentes de *software*, os autores propuseram o uso de ontologias e a representação do conhecimento para proporcionar o enriquecimento semântico dos dados publicados. Fensel et al. [9] propuseram uma arquitetura para proporcionar a implementação da *web* semântica. Tal arquitetura é composta por um modelo de sete camadas que geram uma arquitetura completa e estruturada para a representação da *web* semântica. O formato de estruturação e a disposição destas camadas pode ser visualizado na Figura 2.1 e, na sequência, serão apresentadas brevemente cada uma destas camadas.

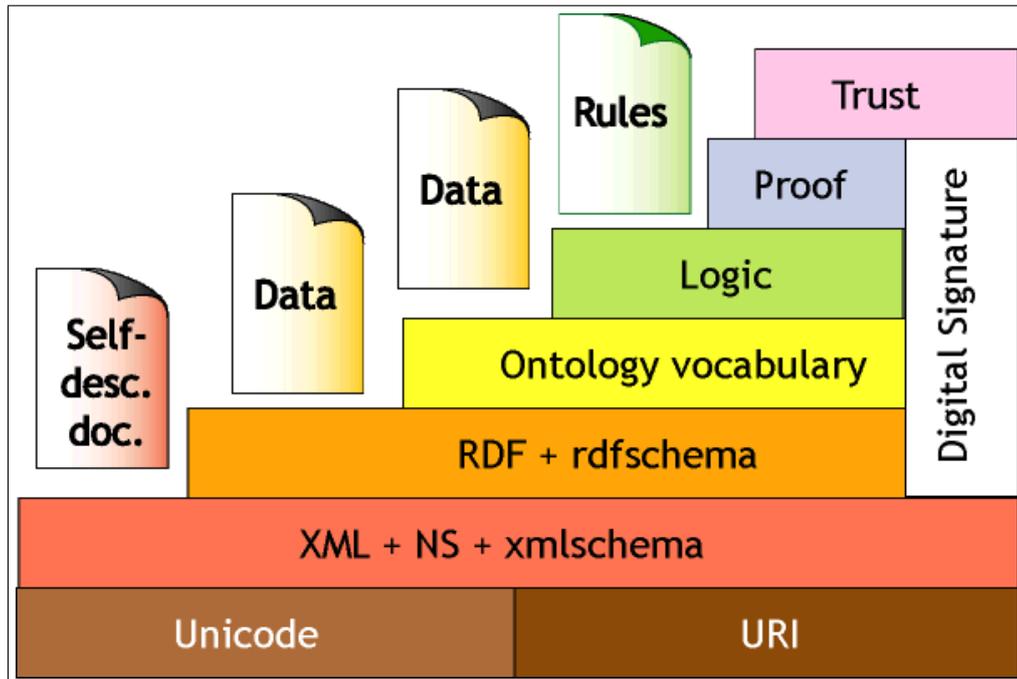


Figura 2.1: Camadas de Web Semântica. Fonte: Fensel et al. (2003)

A primeira camada trata das informações publicadas na *web*, na qual cada dado ou conteúdo publicado contém uma *Uniform Resource Identifier (URI)*. Esta *URI* é utilizada para identificar unicamente os recursos publicados na *web*.

No contexto desta pesquisa, as *URI* identificarão os possíveis componentes de uma **ontologia**: conceitos, propriedades e relações entre os conceitos. Um exemplo de uma *URI* é a seguinte: <http://dbpedia.org/ontology/River>. Ela representa o conceito *rio* na **ontologia** mantida pela *DBpedia*<sup>1</sup>.

A segunda camada é composta pela linguagem *Extensible Markup Language (XML)*. Ela é composta por regras simples de sintaxe e pode ser utilizada em várias aplicações. Uma das principais atribuições do XML é o provimento de troca de informações entre aplicações na *web*. É possível fazer uso desta linguagem a partir de uma estrutura pré-formatada ou de um documento de forma livre. Para o uso a partir de um documento pré-formatado, há duas possibilidades. A primeira é mediante uma estrutura definida por um documento XML descrito na linguagem XML Schema. A outra possibilidade é fazendo uso de um *Document Type Definition (DTD)*, no qual são definidas a gramática utilizada na estruturação dos elementos e a ordem em que estes elementos são distribuídos no documento.

<sup>1</sup><http://dbpedia.org/>

No Código Fonte 2.1, ilustra-se um exemplo da modelagem de um rio na linguagem XML. Trata-se da representação do rio Paraíba e contém as informações da nascente, extensão e dos municípios que são cruzados por ele. É possível observar simplicidade para representação dos dados, além da semântica existente nas marcações do documento.

#### Código Fonte 2.1: Exemplo de uso do XML

```
<?xml version="1.0" encoding="UTF-8"?>
<rio>
  <nome>Paraíba</nome>
  <nascente>Serra de Jabitacá</nascente>
  <extesao>380</extesao>
  <estados>
    <estado>
      <uf>PB</uf>
      <municipios>
        <municipio>Monteiro</municipio>
        <municipio>Campina Grande</municipio>
        <municipio>...</municipio>
        <municipio>João Pessoa</municipio>
        <municipio>Santa Rita</municipio>
        <municipio>Bayeux</municipio>
        <municipio>Cabedelo</municipio>
        <municipio>Lucena</municipio>
      </municipios>
    </estado>
  </estados>
</rio>
```

A sintaxe do XML serviu de base para o desenvolvimento de outros padrões. Um desses padrões é a linguagem *Resource Description Framework* (RDF) [10], que representa a terceira camada. A linguagem nos permite definir os relacionamentos semânticos existentes entre as informações publicadas na *web*. RDF é a padronização proposta pelo W3C para criar um modelo simples de dados para representar formalmente a semântica existente entre os recursos disponibilizados na *web*.

As construções RDF possuem três componentes básicos: sujeito, predicado e objeto. O sujeito representa o recurso da declaração, o predicado representa a propriedade ou característica de um recurso e o objeto representa o valor de uma propriedade. Estes componentes são descritos em um arquivo XML e ordenados da seguinte forma: <sujeito><predicado><objeto>. Os recursos representados em RDF são identificados por uma URI. Um exemplo de uma tripla RDF é representado a seguir, a qual está informando que o conceito *rio* é um tipo do conceito *classe*.

```
<http://dbpedia.org/ontology/River>
```

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

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

A camada seguinte faz uso de **ontologias** para melhorar a descrição e descoberta de informações. Conforme já foi mencionado, foram desenvolvidas algumas linguagens para a definição de **ontologias**, tais como DAML, OIL, Description Logic e OWL.

A camada de lógica possibilita a inferência de conhecimento derivado de conceitos e regras definidos. Algumas linguagens foram desenvolvidas para a descrição de regras da *web* semântica: *Rule Markup Language* [11] (RuleML), *Semantic Web Rule Language* [12] (SWRL) e *Rule Interchange Format* [13] (RIF).

Finalmente, as camadas de prova e confiança tratam da validação do conhecimento obtido na camada anterior, verificando a confiabilidade e a qualidade da informação obtida.

## 2.2 Ontologia

Segundo Gruber [5], **ontologia** é a especificação de uma conceitualização, isto é, a descrição dos conceitos e os relacionamentos entre os mesmos. Esta definição considera que uma **ontologia** é apenas um conjunto de conceitos e definições. Guarino [6], por sua vez, define **ontologia** como sendo uma especificação formal e explícita de uma conceitualização compartilhada. Em outras palavras, esta definição pode ser formatada como sendo a especificação de conceitos e propriedades de algum domínio de forma explícita, que seja legível por computadores e de conhecimento consensual. Tanto Gruber [5], quanto Guarino [6] afirmam que o objetivo principal das **ontologias** é possibilitar o compartilhamento e reutilização do conhecimento.

Dessa forma, é possível criar **ontologias** para os mais variados domínios e, com isso, possibilitar a localização e o compartilhamento de informações na *web* de forma mais rápida e precisa. Uma vez criada uma **ontologia** para um determinado domínio, ela pode ser utilizada por ferramentas de busca para prover a melhoria na qualidade das consultas sobre aquele domínio. Segundo Gruber [5], há dois tipos de conhecimentos que podem ser utilizados pelos motores de busca que utilizam **ontologias** para a resolução de consultas: intensional e extensional. O primeiro representa o conhecimento geral sobre o domínio do problema mediante análise de classes que compõem a **ontologia**, enquanto que o segundo

especifica um problema particular a partir da análise das instâncias associadas à **ontologia**. Levando-se em consideração o domínio de hidrologia como exemplo, poderíamos ter a definição intensional de corpos hídricos como sendo caracterizada pelas propriedades, conceitos e relacionamentos entre conceitos. Como, extensão, volume de água, margem, correnteza, queda d'água. Por outro lado, a definição extensional de corpos hídricos seriam instâncias do mesmo: rio Paraíba, açude de Boqueirão, cataratas do Iguaçu, oceano Atlântico, entre outros.

Segundo Tello [14], os seguintes componentes fazem parte da estrutura de uma **ontologia**:

- **Conceitos:** são as ideias básicas que são formalizadas. Podem ser classes de objetos, métodos, planos, estratégias, processos de raciocínio, dentre outros;
- **Propriedades:** representam um conjunto de atributos que caracterizam um conceito;
- **Relações:** representam a interação entre os conceitos de um domínio. Exemplos de relações incluem: especialização, generalização, meronímia, entre outras;
- **Funções:** é um caso específico de relacionamento entre conceitos em que um conjunto de conceitos tem uma relação única com outro elemento. Um exemplo é a função *ser\_filho\_biológico*, onde um conceito *pessoa* está relacionado com um conceito *homem* e um conceito *mulher*;
- **Instâncias:** representam objetos de um conceito. São as concretizações de um conceito;
- **Axiomas:** são teoremas declarados nas relações que devem atender os elementos da **ontologia**. Este é um dos pontos mais fortes do conceito de ontologia, pois possibilita inferir novos conhecimentos com base nos teoremas declarados.

Linguagens como RDF-Schema [15], OIL (*Ontology Inference Layer*) [16] e DAML (*DARPA Agent Markup Language*) [17] foram criadas para permitir a criação de **ontologias**. Entretanto, as duas linguagens que estão sendo mais utilizadas em pesquisas recentes são a *Description Logic* [18] e a OWL (*Web Ontology Language*) [19].

*Description Logic* (DL) é uma linguagem baseada em lógica para representação do conhecimento. Usada para descrever um domínio de maneira formal e estruturada. O domínio é representado pela descrição de seus conceitos, que são relacionados mediante a semântica formal baseada em lógica. Esta linguagem possibilita o uso de operadores tais como conjunção, negação, restrição existencial e quantificadores existenciais. A DL é composta por duas partes: *T-Box* e *A-Box*.

A *T-Box* define os conceitos que compõem a hierarquia da **ontologia** a partir de regras. Seguem alguns exemplos de conceitos representados por *T-Box*:

- $Rio \equiv CorpoHidrico \sqcap \exists desembocaNo.CorpoHidrico$ : Este termo afirma que o conceito *rio* é um corpo hídrico e desemboca em outro corpo hídrico;
- $Rio \sqsubseteq RioPerene$ : Este termo afirma que o conceito *rio perene* está contido no conceito *rio*.

A *A-Box*, por sua vez, contém as instâncias pertencentes à **ontologia** e é chamada de *A-Box*. Como exemplos de instâncias representadas por *A-Box* podem-se listar os seguintes:

- $RioPerene(saofrancisco)$ : Este termo está instanciando o rio São Francisco como um rio perene;
- $desembocaNo(saofrancisco, Mar)$ : Este termo está instanciando o relacionamento de que o rio São Francisco desemboca no mar.

A linguagem OWL foi padronizada pelo *Word Wide Web Consortium* (W3C) para a criação e instanciação de **ontologias**. Ela provê a descrição das entidades de um domínio, suas propriedades e os relacionamentos existentes entre elas. Baseada em XML e RDF, OWL faz uso de esquemas de ambas as linguagens para definir as estruturas de um determinado domínio. Esta estruturação possibilita o “aprendizado” computacional eficiente e performático. Segundo Harmelen e McGuinness [19], a OWL se subdivide em três tipos, de acordo com o nível de expressividade:

- **OWL Lite**: é o tipo de OWL com restrições simples, em que todos os relacionamentos entre as entidades possuem cardinalidade 0 ou 1;

- **OWL DL:** é o tipo de OWL um pouco mais expressivo, que mantém a completude computacional (todas as conclusões são garantidamente computáveis) e a decidibilidade (todas as decisões terminarão em um tempo finito);
- **OWL Full:** é o tipo de OWL que possui o maior nível de expressividade e é compatível com RDF. Uma **ontologia** representada por este tipo de OWL pode aumentar a expressividade de um vocabulário pré-definido em RDF ou OWL. Porém, não há garantias de que o que ela representa seja computacionalmente decidível.

No Código Fonte 2.2, mostra-se um exemplo de uma **ontologia** descrita na linguagem OWL no domínio de corpos hídricos. São criados os relacionamentos existentes entre a classe *BodyOfWater* e as suas subclasses *Lake*, *Sea* e *Stream*, além do relacionamento da classe *Stream* com a subclasse *River*.

#### Código Fonte 2.2: Exemplo de utilização de OWL

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE rdf:RDF [
  <!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
  <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
  <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
  <!ENTITY owl "http://www.w3.org/2002/07/owl#" >
  <!ENTITY ns_ontology "http://dbpedia.org/ontology#" >]>

<rdf:RDF
  xmlns:xsd="&xsd;"
  xmlns:rdf="&rdf;"
  xmlns:rdfs="&rdfs;"
  xmlns:owl="&owl;"
  xmlns:ns_ontology="&ns_ontology;">

  <owl:Ontology rdf:about="&ns_ontology;BodyOfWaterOntology">
    <rdfs:label>Classificação de corpos hídricos</rdfs:label>
    <rdfs:comment>
      São apresentados os conceitos e subconceitos
      relacionados
      com corpos hídricos.
    </rdfs:comment>
  </owl:Ontology>

  <owl:Class rdf:about="&ns_ontology;BodyOfWater">
    <rdfs:label>Corpos Hídricos</rdfs:label>
    <rdfs:comment>
      Conceito que engloba todas as representações de corpos
      hídricos.
    </rdfs:comment>
  </owl:Class>
```

```

<owl:Class rdf:about="&ns_ontology;Lake">
  <rdfs:subClasseOf rdf:resource="&ns_ontology;BodyOfWater" />
  <rdfs:label>Lagos</rdfs:label>
  <rdfs:comment>
    Conceito que engloba todas as representações de lagos.
  </rdfs:comment>
</owl:Class>

<owl:Class rdf:about="&ns_ontology;Sea">
  <rdfs:subClasseOf rdf:resource="&ns_ontology;BodyOfWater" />
  <rdfs:label>Oceanos</rdfs:label>
  <rdfs:comment>
    Conceito que engloba todas as representações de oceanos
  </rdfs:comment>
</owl:Class>

<owl:Class rdf:about="&ns_ontology;Stream">
  <rdfs:subClasseOf rdf:resource="&ns_ontology;BodyOfWater" />
  <rdfs:label>Águas Correntes</rdfs:label>
  <rdfs:comment>
    Conceito que engloba todas as representações de águas
    correntes.
  </rdfs:comment>
</owl:Class>

<owl:Class rdf:about="&ns_ontology;River">
  <rdfs:subClasseOf rdf:resource="&ns_ontology;Stream" />
  <rdfs:label>Rios</rdfs:label>
  <rdfs:comment>
    Conceito que engloba todas as representações de rios.
  </rdfs:comment>
</owl:Class>
</rdf:RDF>

```

---

## 2.3 Serviços OGC

O *Open Geospatial Consortium* (OGC) é um consórcio composto por organizações de vários tipos, como empresas privadas, organizações governamentais, organizações não lucrativas e instituições de pesquisa. Ele tem como um de seus principais objetivos a especificação de padrões públicos e abertos para o domínio geoespacial, almejando com isso, o aumento da interoperabilidade entre os dados geográficos e as aplicações de SIG existentes. Dentre os padrões propostos pelo OGC, pode-se destacar os serviços que possibilitam o acesso a dados geográficos em diferentes formatos. Três importantes exemplos destes serviços são o *Web Map Service* [20] (WMS), que possibilita o acesso a camadas de mapas geo-referenciados em

formatos como PNG, JPEG, GIF e SVG, o *Web Feature Service* [21] (WFS), que possibilita o acesso e a manipulação de dados geográficos das feições em formatos como GML e *shapefile*, e o *Catalog Service for the Web* [22] (CSW), que possibilita anunciar dados e serviços geográficos, por parte dos provedores, e recuperar os dados geográficos disponibilizados, por parte dos clientes.

Os serviços especificados pelo OGC possibilitam a recuperação de dados geográficos fornecidos por um determinado provedor a partir de uma interface padronizada, sem a necessidade de se conhecer os detalhes sobre o seu armazenamento. Os dados recuperados a partir de alguns destes serviços são conjuntos de camadas presentes no conjunto de dados oferecidos por um provedor.

Os dados geográficos são disponibilizados por alguns serviços padronizados pelo OGC na forma de camadas. Cada camada fornece informações sobre um determinado tema. Por exemplo, suponha-se um serviço WMS que provê dados sobre o relevo do Brasil, conforme mostrado na Figura 2.2.

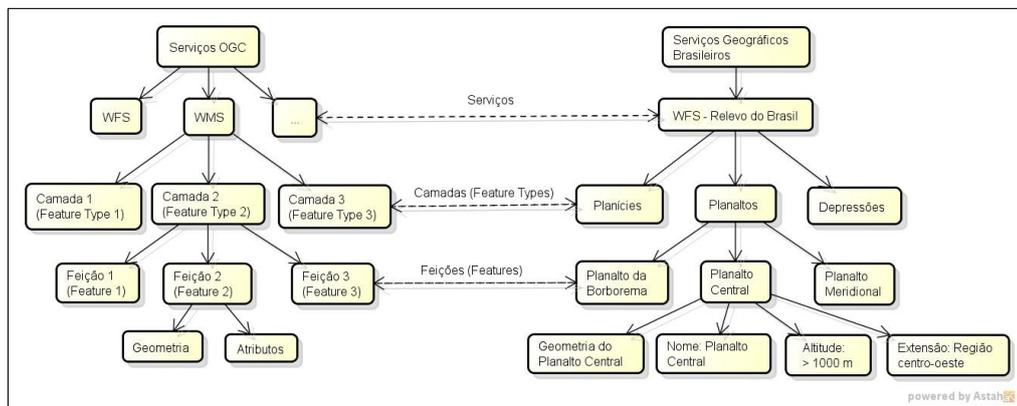


Figura 2.2: Organização dos serviços OGC.

Ele pode ser composto por camadas tais como planaltos, planícies, depressões, serras, etc. Cada camada de um serviço possui um conjunto de dados geo-referenciados sobre o seu tema. Estes dados são denominados feições (*feature*). Por exemplo, uma camada que provê informações sobre planaltos pode oferecer feições sobre o Planalto Meridional, o Planalto Central, o Planalto da Borborema, etc. Por fim, cada feição é composta pela sua geometria, que descreve a sua localização, e um conjunto de atributos, que a caracterizam. Um exemplo que se pode apresentar da feição Planalto Central é a sua localização, mediante sua geometria e seus atributos, tais como o nome, a altitude e a extensão.

### 2.3.1 CSW (*Catalog Service for the Web*)

Os serviços de catálogo (CSW) padronizados pelo OGC visam a facilitar a descoberta dos dados e serviços publicados pelos provedores de informações geográficas. Dessa forma, os clientes que consomem este tipo de informação podem consultar o catálogo para encontrar dados e serviços de seu interesse a partir de palavras chaves, por exemplo. Para publicar o seu conjunto de dados em um serviço de catálogo, o provedor precisa fornecer uma série de metadados acerca dos dados a serem publicados. Estes metadados serão utilizados nas consultas feitas pelos clientes para localizar os possíveis serviços de interesse.

O padrão CSW possui algumas operações principais como: *GetCapabilities*, *GetRecords* e *GetRecordById*. Outras operações são: *DescribeRecord*, *GetDomain* e *Transaction*.

A operação *GetCapabilities* possibilita ao cliente conhecer as funcionalidades do serviço de catálogo. A resposta de uma requisição deste tipo é um documento XML contendo várias informações sobre o serviço de catálogo. Algumas informações retornadas são título, descrição textual, palavras-chave, operações implementadas pelo serviço, dentre outras.

Na Figura 2.3, ilustra-se uma requisição a uma operação *GetCapabilities* de um serviço CSW provido pelo *British Geological Survey*<sup>2</sup>. Os parâmetros que foram repassados na requisição são, respectivamente, o tipo do serviço, a operação a ser executada e a versão do serviço.

```
http://metadata.bgs.ac.uk/geonetwork/srv/en/csw?  
service=CSW  
&request=GetCapabilities  
&version=2.0.2
```

Figura 2.3: Exemplo da operação *GetCapabilities* do serviço CSW.

No Código Fonte 2.3, apresenta-se um trecho da resposta da operação que contém os metadados identificadores de serviço de catálogo.

#### Código Fonte 2.3: Metadados do CSW do *British Geological Survey*.

```
<?xml version="1.0" encoding="UTF-8"?>  
<ows:ServiceIdentification>  
  <ows:Title>  
    CSW service for BGS metadata  
  </ows:Title>  
  <ows:Abstract>
```

<sup>2</sup><http://www.bgs.ac.uk/>

```

    This service provides access to BGS ISO19115:2003 metadata
  </ows:Abstract>
  <ows:Keywords>
    <ows:Keyword>NERC_DDC</ows:Keyword>
    <ows:Keyword>Geology</ows:Keyword>
    <ows:Keyword>Geophysics</ows:Keyword>
    <ows:Keyword>Boreholes</ows:Keyword>
    <ows:Keyword>Maps</ows:Keyword>
    <ows:Keyword>Digital maps</ows:Keyword>
    <ows:Keyword>Geochemistry</ows:Keyword>
    <ows:Keyword>Geophysical surveys</ows:Keyword>
    <ows:Keyword>Data</ows:Keyword>
    <ows:Keyword>GIS</ows:Keyword>
    <ows:Type>theme</ows:Type>
  </ows:Keywords>
  <ows:ServiceType>CSW</ows:ServiceType>
  <ows:ServiceTypeVersion>2.0.2</ows:ServiceTypeVersion>
  <ows:Fees>none</ows:Fees>
  <ows:AccessConstraints>none</ows:AccessConstraints>
</ows:ServiceIdentification>

```

A operação *GetRecords* possibilita ao cliente recuperar os registros de metadados disponibilizados pelos serviços. Para isso, é necessário inserir uma série de parâmetros na requisição desta operação. A Figura 2.4 apresenta uma requisição de uma operação *GetRecords* ao serviço CSW provido pelo *British Geological Survey*. Os parâmetros que foram passados na requisição representam, respectivamente, a operação a ser executada, o tipo do serviço, a versão do serviço, o tipo da resposta (*results*, *hits* ou *validates*), a quantidade máxima de registros retornado pela requisição, o tipo do esquema a ser usado para descrever a resposta, a linguagem utilizada no filtro da pesquisa, a versão desta linguagem e o próprio filtro da pesquisa.

```

http://metadata.bgs.ac.uk/geonetwork/srv/en/csw?request=GetRecords
&service=CSW
&version=2.0.2
&resultType=results
&maxRecords=10
&typeName=csw:Record
&constraintLanguage=CQL_TEXT
&constraint_language_version=1.1.0
&constraint=AnyText+LIKE+%27%25water%25%27

```

Figura 2.4: Exemplo da operação *GetRecords* do CSW.

No Código Fonte 2.4, apresenta-se um dos registros retornados pela requisição da operação *GetRecords* feita ao serviço CSW provido pelo *British Geological Survey*. Para cada registro retornado é possível verificar informações como o identificador, o título, o tipo, a descrição textual, dentre outras.

---

**Código Fonte 2.4: Um dos registros retornado pela operação *GetRecords* do CSW.**


---

```

<csw:SummaryRecord xmlns:dct="http://purl.org/dc/terms/"
  xmlns:geonet="http://www.fao.org/geonetwork" xmlns:dc="http://purl.
    org/dc/elements/1.1/">
  <dc:identifier>
    c42c40be-da98-45ce-e044-0003ba9b0d98
  </dc:identifier>
  <dc:title>
    BGS 1:625 000 Scale Digital Hydrogeological Data (OGC WxS
      INSPIRE)
  </dc:title>
  <dc:type>service</dc:type>
  <dc:subject>Geology</dc:subject>
  <dc:subject>Digital maps</dc:subject>
  <dc:subject>Geology</dc:subject>
  <dc:subject>Aquifers</dc:subject>
  <dc:subject>Hydrogeology</dc:subject>
  <dc:subject>infoMapAccessService</dc:subject>
  <dc:subject>NERC_DDC</dc:subject>
  <dc:format>image/svg+xml</dc:format>
  <dc:format>image/gif</dc:format>
  <dc:format>image/png</dc:format>
  <dc:format>image/tiff</dc:format>
  <dc:format>image/jpeg</dc:format>
  <dc:format>image/bmp</dc:format>
  <dct:abstract>
    The hydrogeological map indicates aquifer potential in
    generalised terms using a threefold division of geological
    formations: those in which intergranular flow in the
    saturated zone is dominant, those in which flow is
    controlled by fissures or discontinuities and less
    permeable formations including aquifers concealed at depth
    beneath covering layers. Highly productive aquifers are
    distinguished from those that are only of local importance
    or have no significant groundwater. Within each of these
    classes the strata are grouped together according to age
    or lithology. The 1:625 000 scale data may be used as a
    guide to the aquifers at a regional or national level, but
    should not be relied on for local information.
  </dct:abstract>
</csw:SummaryRecord>

```

---

A operação *GetRecordById* possibilita ao cliente recuperar um ou mais registros de metadados disponibilizados pelo serviço a partir da identificação dos mesmos. Para utilizar esta operação, é necessário se ter a informação dos ids referentes aos registros de interesse. Na Figura 2.5 apresenta-se uma requisição de uma operação *GetRecordById* ao serviço CSW provido pelo *British Geological Survey*. Os parâmetros que foram passados na requisição são, respectivamente, a operação a ser executada, o tipo do serviço, a versão do serviço e o *id* que identifica o registro a ser recuperado.

```
http://metadata.bgs.ac.uk/geonetwork/srv/en/csw?request=GetRecordById
&service=CSW
&version=2.0.2
&id=b0496466-f09a-6d51-e044-0003ba9b0d98
```

Figura 2.5: Exemplo da operação *GetRecordById* do serviço CSW.

No Código Fonte 2.5, apresenta-se o retorno da requisição da operação *GetRecordById* ao serviço CSW provido pelo *British Geological Survey*. Assim como na operação *GetRecords*, os registros retornados pela operação *GetRecordById* também disponibilizam informações como o identificador, o título, a descrição textual, o tipo, entre outras.

#### Código Fonte 2.5: Resultado da operação *GetRecordById* do serviço CSW.

```
<csw:GetRecordByIdResponse>
  <csw:SummaryRecord xmlns:dct="http://purl.org/dc/terms/"
    xmlns:geonet="http://www.fao.org/geonetwork" xmlns:dc="http
      ://purl.org/dc/elements/1.1/">
    <dc:identifier>
      b0496466-f09a-6d51-e044-0003ba9b0d98
    </dc:identifier>
    <dc:title>
      BGS GeoIndex - Hazards data theme (OGC WxS INSPIRE)
    </dc:title>
    <dc:type>service</dc:type>
    <dc:subject>Geology</dc:subject>
    <dc:subject>Landslides</dc:subject>
    <dc:subject>Natural hazards</dc:subject>
    <dc:subject>Geology</dc:subject>
    <dc:subject>Water flooding</dc:subject>
    <dc:subject>Earthquakes</dc:subject>
    <dc:subject>Digital maps</dc:subject>
    <dc:subject>infoMapAccessService</dc:subject>
    <dc:subject>NERC_DDC</dc:subject>
    <dc:format>image/svg+xml</dc:format>
    <dc:format>image/gif</dc:format>
    <dc:format>image/png</dc:format>
    <dc:format>image/tiff</dc:format>
    <dc:format>image/jpeg</dc:format>
    <dc:format>image/bmp</dc:format>
    <dct:abstract>
      Data from the British Geological Survey's GeoIndex
      Hazards theme are made available for viewing here.
      GeoIndex is a website that allows users to search
      for information about BGS data collections covering
      the UK and other areas world wide. Access is free,
      the interface is easy to use, and it has been
      developed to enable users to check coverage of
      different types of data and find out some background
      information about the data. More detailed
      information can be obtained by further enquiry via
      the web site: www.bgs.ac.uk/geoindex.
    </dct:abstract>
```

```
</csw:SummaryRecord>  
</csw:GetRecordByIdResponse>
```

---

A operação *DescribeRecord* retorna informações que descrevem os registros com base no modelo de informação suportado pelo serviço de catálogo. A operação *GetDomain* fornece uma interface para retornar todos os valores possíveis para uma determinada propriedade dos registros providos pelo catálogo ou dos parâmetros usados nas consultas dos clientes. A operação *Transaction*, por sua vez, possibilita o cadastro de novos registros e a alteração ou remoção de registros existentes no catálogo.

### 2.3.2 WMS (Web Map Service)

O serviço WMS padronizado possibilita a recuperação de imagens de mapas georeferenciados mediante uma interface simples. Estes mapas são gerados por um servidor de mapas que utiliza os dados armazenados em um Sistema de Informação Geográfica (SIG). A requisição que é feita ao servidor é constituída por parâmetros como uma lista de camadas que devem ser recuperados e a área de interesse. O servidor responde com uma imagem no formato PNG, JPEG, etc.

O padrão WMS possui operações obrigatórias e opcionais. As operações obrigatórias são *GetCapabilities* e *GetMap*. Algumas operações opcionais são, dentre outras, *GetFeatureInfo*, *DescribeLayer* e *GetLegendGraphic*.

A operação *GetCapabilities* possibilita ao cliente conhecer os metadados do serviço, isto é, as informações gerais sobre o seu funcionamento e as camadas disponíveis. A resposta de uma requisição deste tipo é um documento XML contendo as configurações e funcionalidades do serviço, bem como a descrição das camadas disponíveis.

Na Figura 2.6, ilustra-se uma requisição a uma operação *GetCapabilities* de um serviço WMS fornecido pela INDE. Os parâmetros que foram repassados na requisição representam, respectivamente, o tipo do serviço, a operação a ser executada e a versão do serviço.

```
http://www.geoservicos.inde.gov.br/geoserver/ows?  
service=WMS  
&request=GetCapabilities  
&version=1.0.0
```

Figura 2.6: Exemplo da operação *GetCapabilities* do serviço WMS.

O Código Fonte 2.6 corresponde à descrição de uma das camadas descritas pelo documento retornado pela operação *GetCapabilities* ilustrada na Figura 2.6. Nesta descrição, é possível observar informações como nome, título, descrição textual, palavras-chaves e extensão geográfica.

#### Código Fonte 2.6: Descrição de uma camada no serviço WMS.

```
<Layer queryable="1">
  <Name>MPOG:BASE_SPI_pol</Name>
  <Title>Municípios PAC</Title>
  <Abstract />
  <KeywordList />
  <SRS>EPSG:4326</SRS>
  <LatLonBoundingBox minx="-73.991" miny="-33.752"
    maxx="-32.392" maxy="5.272" />
  <BoundingBox SRS="EPSG:4326" minx="-73.991" miny="-33.752"
    maxx="-32.392" maxy="5.272" />
  <Style>
    <Name>BASE_SPI_pol</Name>
    <Title />
    <Abstract />
    <LegendURL width="20" height="20">
      <Format>image/png</Format>
      <OnlineResource xmlns:xlink="http://www.w3.org/1999/
        xlink"
        xlink:type="simple"
        xlink:href="http://www.geoservicos.inde.gov.br
          :80/geoserver/wms?request=GetLegendGraphic&
          ;format=image%2Fpng&width=20&height
          =20&layer=BASE_SPI_pol" />
    </LegendURL>
  </Style>
</Layer>
```

Após recuperar as camadas disponibilizadas pelo serviço, o cliente pode escolher as camadas de seu interesse e recuperá-las mediante operação *GetMap*. Na figura 2.7, ilustra-se um exemplo de requisição desta operação.

```
http://www.geoservicos.inde.gov.br/geoserver/ows?
service=WMS
&version=1.0.0
&REQUEST=GetMap
&LAYERS=MPOG:BASE_SPI_pol
&CRS=EPSG:4326
&BBOX=-73.991,-33.752,-32.392,5.272
&HEIGHT=500&
&WIDTH=800&
&FORMAT=image/jpeg
```

Figura 2.7: Exemplo da operação *GetMap* do serviço WMS.

Os parâmetros que foram passados na requisição representam, respectivamente, o tipo do

serviço, a versão do serviço, a operação a ser executada, o nome da camada a ser recuperada, o sistema de coordenadas que deve ser adotado, as coordenadas da extensão geográfica, a altura, a largura e o formato da imagem a ser gerada. As informações necessárias para especificar valores para os parâmetros, tais como camadas e estilos podem ser obtidos a partir do documento de funcionalidades.

Como resposta a requisição da operação *GetMap*, é gerada uma imagem contendo todas as camadas solicitadas. Na Figura 2.8, apresenta-se o mapa obtido a partir da requisição ilustrada na Figura 2.7.

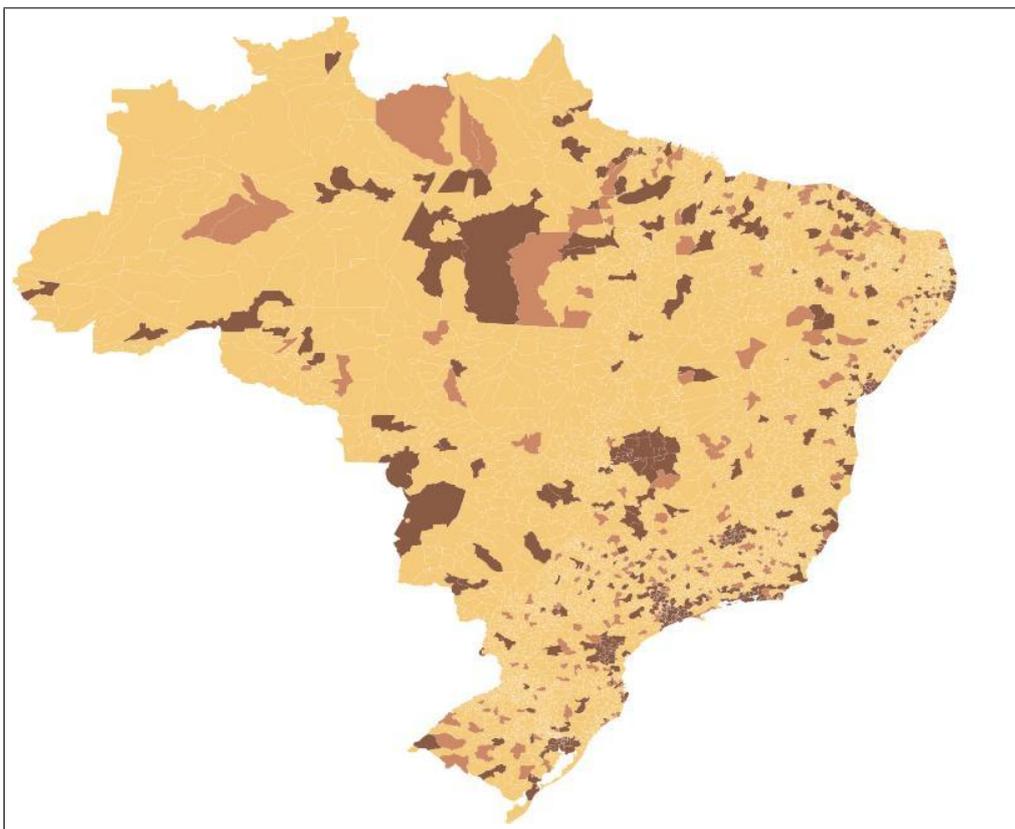


Figura 2.8: Mapa gerado pela operação *GetMap* do WMS.

### 2.3.3 WFS (*Web Feature Service*)

O serviço WFS possibilita a recuperação de camadas a partir da *web* nos formatos GML, que é uma linguagem baseada em XML para descrever as características geográficas, e *shapefiles*. Este serviço permite abrir os dados geográficos, possibilitando ao cliente recuperá-los e determinar como visualizá-los. Também é possível a atualização destes dados junto ao ser-

vidor. Porém, as operações que permitem esse tipo de comportamento só estão disponíveis para clientes autorizados.

Os serviços WFS podem ser classificados de duas formas: WFS básico, que disponibiliza operações de consulta e recuperação dos dados geográficos, e o *Web Feature Service Transactional* (WFS-T), que, além das operações possíveis do WFS básico, também oferece operações de inclusão, alteração e remoção.

As operações do serviço WFS básico são: *GetCapabilities*, *DescribeFeatureType*, *GetFeature*, *LockFeature* e *GetGMLObject*. O WFS-T disponibiliza, além destas operações, a operação *Transaction*.

A operação *GetCapabilities* retorna quais operações e serviços são disponibilizados pelo servidor. O resultado desta operação é um documento XML com as informações gerais do serviço e o conjunto de camadas providas pelo serviço. Na Figura 2.9, ilustra-se um exemplo de uma operação *GetCapabilities* de um serviço WFS disponibilizado pela INDE.

```
http://www.geoservicos.inde.gov.br/geoserver/ows?
service=WFS
&request=GetCapabilities
&version=1.0.0
```

Figura 2.9: Exemplo da operação *GetCapabilities* do serviço WFS.

No Código Fonte 2.7, apresenta-se a descrição de uma das camadas retornadas pela operação ilustrada na Figura 2.9.

Código Fonte 2.7: Descrição de uma camada em um serviço WFS.

```
<FeatureType>
  <Name>MPOG:Agua_em_areas_urbanas</Name>
  <Title>Agua_em_areas_urbanas</Title>
  <Abstract />
  <Keywords />
  <SRS>EPSG:4674</SRS>
  <LatLongBoundingBox minx="-73.38535999999999" miny="-34.07286"
    maxx="-34.333639999999995" maxy="4.97886" />
</FeatureType>
```

Depois de selecionadas as camadas oferecidas pelo serviço, o cliente pode solicitar mais detalhes sobre a estrutura de um deles, como, por exemplo, sua lista de atributos com seus respectivos tipos. Isto é feito a partir da operação *DescribeFeatureType*. Os parâmetros desta operação são os seguintes: tipo do serviço, operação, versão e a lista dos nomes das camadas

que devem ser apresentadas. Na Figura 2.10, apresenta-se um exemplo da requisição desta operação.

```
http://www.geoservicos.inde.gov.br/geoserver/ows?
service=WFS
&request=DescribeFeatureType
&version=1.0.0
&typename=MPOG:Agua_em_areas_urbanas
```

Figura 2.10: Exemplo da operação *DescribeFeatureType* do serviço WFS.

O resultado da operação apresentada na Figura 2.10 é um documento XML com as descrições dos esquemas das camadas solicitadas. No Código Fonte 2.8, ilustra-se o esquema recebido.

Código Fonte 2.8: Descrição do esquema de uma camada no serviço WFS.

```
<xsd:schema
  elementFormDefault="qualified"
  targetNamespace="www.planejamento.gov.br/">
<xsd:import
  namespace="http://www.opengis.net/gml"
  schemaLocation="http://www.geoservicos.inde.gov.br:80/geoserver/
    schemas/gml/2.1.2/feature.xsd" />
<xsd:complexType name="Agua_em_areas_urbanasType">
  <xsd:complexContent>
    <xsd:extension base="gml:AbstractFeatureType">
      <xsd:sequence>
        <xsd:element maxOccurs="1" minOccurs="0" name="ID"
          nillable="true" type="xsd:int" />
        <xsd:element maxOccurs="1" minOccurs="0" name="
          EMPREENDIM" nillable="true" type="xsd:string" />
        <xsd:element maxOccurs="1" minOccurs="0" name="SUBEIXO"
          nillable="true" type="xsd:string" />
        <xsd:element maxOccurs="1" minOccurs="0" name="TIPO"
          nillable="true" type="xsd:string" />
        <xsd:element maxOccurs="1" minOccurs="0" name="
          VALOR_2011" nillable="true" type="xsd:string" />
        <xsd:element maxOccurs="1" minOccurs="0" name="
          INV_11_14" nillable="true" type="xsd:double" />
        <xsd:element maxOccurs="1" minOccurs="0" name="
          VALOR_POS_" nillable="true" type="xsd:string" />
        <xsd:element maxOccurs="1" minOccurs="0" name="
          INV_POS_14" nillable="true" type="xsd:double" />
        <xsd:element maxOccurs="1" minOccurs="0" name="
          INVESTIMEN" nillable="true" type="xsd:string" />
        <xsd:element maxOccurs="1" minOccurs="0" name="INVEST_T
          " nillable="true" type="xsd:double" />
        <xsd:element maxOccurs="1" minOccurs="0" name="
          INVESTIM_2" nillable="true" type="xsd:string" />
        <xsd:element maxOccurs="1" minOccurs="0" name="UF"
          nillable="true" type="xsd:string" />
      </xsd:sequence>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

```

<xsd:element maxOccurs="1" minOccurs="0" name="
  MUNICIPIOS" nillable="true" type="xsd:string" />
<xsd:element maxOccurs="1" minOccurs="0" name="
  EXECUTORES" nillable="true" type="xsd:string" />
<xsd:element maxOccurs="1" minOccurs="0" name="ORGAO"
  nillable="true" type="xsd:string" />
<xsd:element maxOccurs="1" minOccurs="0" name="ESTAGIO"
  nillable="true" type="xsd:string" />
<xsd:element maxOccurs="1" minOccurs="0" name="
  DATA_CICLO" nillable="true" type="xsd:string" />
<xsd:element maxOccurs="1" minOccurs="0" name="
  DATA_SELEC" nillable="true" type="xsd:string" />
<xsd:element maxOccurs="1" minOccurs="0" name="
  DATA_CONCL" nillable="true" type="xsd:string" />
<xsd:element maxOccurs="1" minOccurs="0" name="
  EMBLEMATIC" nillable="true" type="xsd:string" />
<xsd:element maxOccurs="1" minOccurs="0" name="
  OBSERVACAO" nillable="true" type="xsd:string" />
<xsd:element maxOccurs="1" minOccurs="0" name="LAT1"
  nillable="true" type="xsd:string" />
<xsd:element maxOccurs="1" minOccurs="0" name="LON1"
  nillable="true" type="xsd:string" />
<xsd:element maxOccurs="1" minOccurs="0" name="
  GEOCODIGO" nillable="true" type="xsd:string" />
<xsd:element maxOccurs="1" minOccurs="0" name="the_geom
  " nillable="true" type="gml:PointPropertyType" />
</xsd:sequence>
</xsd:extension>
</xsd:complexContent>
</xsd:complexType>
<xsd:element
  name="Agua_em_areas_urbanas"
  substitutionGroup="gml:_Feature"
  type="MPOG:Agua_em_areas_urbanasType" />
</xsd:schema>

```

A operação *GetFeature* é utilizada para recuperar os objetos associados as camadas solicitadas. Ele tem como resultado um documento GML com as feições de cada camada. Por outro lado, a operação *Transaction* possibilita criar, atualizar e excluir camadas. A operação *LockFeature* não permite que uma camada seja modificada. Ela é extremamente útil durante as operações de *Transaction*. Por fim, a operação *GetGMLObject* recupera informações de camadas no formato GML.

## 2.4 IDE (Infraestrutura de Dados Espaciais)

IDE (Infraestrutura de Dados Espaciais) é um sistema de informação que integra uma coleção de tecnologias, políticas e acordos institucionais para possibilitar a disponibilidade e o acesso

aos dados geográficos [2]. Com base nesta integração, as IDE têm se popularizado como uma importante solução para permitir a interoperabilidade de dados geoespaciais disponibilizados por diferentes organizações.

Dessa forma, as IDE precisam fornecer, além de um mecanismo para cadastro de registros de dados geográficos, meios para possibilitar que as informações cadastradas sejam recuperadas de forma conveniente. Um padrão que atende a este requisito e que é utilizado atualmente nas IDE é o serviço de catálogo CSW (seção 2.3.1). Este serviço possibilita que os provedores disponibilizem seus conjuntos de dados mediante cadastro dos seus serviços, enquanto que os clientes consultam o serviço de catálogo para localizar os dados geográficos de que necessitam.

Os dados geográficos providos por uma IDE também podem ser cadastrados a partir de formulários disponibilizados em portais geográficos. Uma vez cadastrados, os metadados são formatados no padrão de metadados adotados pela IDE e armazenados. Este método provê acessibilidade aos provedores que não estão familiarizados com a padronização do serviço de catálogo. De forma semelhante, as IDE também podem disponibilizar interfaces gráficas para realizar a consulta dos registros geográficos.

Atualmente, existem várias iniciativas para o desenvolvimento de IDE no mundo. Dentre elas, está a *National Spatial Data Infrastructure* [23] (NSDI), nos Estados Unidos, a *Canadas Spatial Data Infrastructure* [24] (CSDI), no Canadá, a *Global Earth Observation System of Systems* [25] (GEOSS), que propõe uma IDE de nível global e a *Infraestrutura Nacional de Dados Espaciais* [26] (INDE), que está sendo desenvolvida no Brasil.

## 2.5 Anotação Semântica

A **anotação semântica** no domínio de dados geográficos consiste em realizar o mapeamento entre o dado geográfico e os conceitos existentes em uma **ontologia** que estão semanticamente relacionados com aquele dado geográfico. Este procedimento facilita a descoberta de dados geográficos semanticamente relevantes que estão disponibilizados em fontes distintas, proporcionando assim, interoperabilidade.

A **anotação semântica** de dados geográficos pode ser realizada de três formas:

- **Manual:** é a associação de conceitos de um domínio a dados geográficos de forma

manual, ou seja, um humano liga um determinado dado geográfico aos conceitos que ele acredita haver semântica entre eles. Esta forma de anotação é assertiva, porém subjetiva. Ela é recomendada para a anotação de poucos dados;

- **Semiautomática:** é a associação de conceitos a dados geográficos de forma sugestiva, ou seja, um humano terá alguns conceitos elegíveis automaticamente para decidir se há semântica entre eles e o dado geográfico em questão. Esta forma de anotação também é assertiva e subjetiva, porém pode ser aplicada a um volume maior de dados geográficos devido a automatização da seleção dos conceitos.
- **Automática:** é a associação de conceitos a dados geográficos de forma automatizada, sem a intervenção humana. Esta forma de anotação não é tão assertiva, pois é uma máquina que decide sobre a semântica entre um conceito e um dado geográfico. Porém, ela é objetiva, pois a máquina decide de forma igualitária sobre a semântica entre um conceito e um dado geográfico. Ela é recomendada para grandes volumes de dados geográficos.

Os dados geográficos que são anotados nesta pesquisa são as camadas e seus atributos disponibilizados em alguns serviços OGC das IDE e a **anotação semântica** proposta é realizada de forma automatizada, ou seja, sem a intervenção humana.

## 2.6 **Linked Open Data**

Com a crescente disponibilização de dados sendo publicados na *web*, se faz necessária a organização destas informações com o intuito de facilitar as suas descobertas. Para isso, Tim Berners-Lee propõe *Linked Open Data* (LOD) como um conjunto de diretrizes informando como se publicar dados na *web* e, ao mesmo tempo, criar relacionamentos semânticos entre eles. A idéia principal do LOD é possibilitar a identificação única dos dados disponibilizados na *web* e realizar relacionamentos entre eles. A seguir, as diretrizes propostas pelo autor serão apresentadas. O conceito do LOD resulta em uma grande nuvem de dados interligados [27]. Na Figura 2.11, apresenta-se o conjunto de dados publicados pelo projeto Dados Abertos Interligados até setembro de 2011.

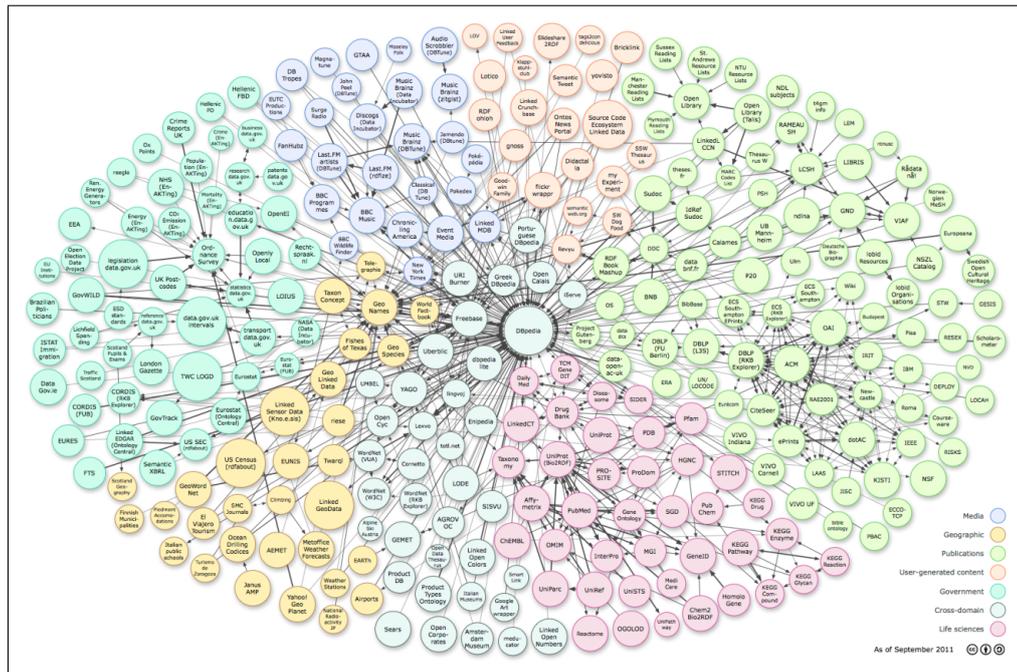


Figura 2.11: Dados publicados no projeto LOD até setembro de 2011.

Cada nó pertencente ao diagrama representa um conjunto de dados e as arestas entre os nós são os relacionamentos existentes entre eles. A coloração representa o domínio a que os conjuntos de dados pertencem.

Segundo Berners-Lee [28], quatro regras são necessárias para o funcionamento do LOD:

1. Usar URI para nomear dados;
2. Usar URI HTTP para que estes dados possam ser referenciados e verificados tanto por pessoas quanto por agentes de *software*;
3. Disponibilizar conjuntos de dados que se relacionam com os dados publicado na *web*, utilizando, para isso, a padronização RDF e/ou os *endpoints* SPARQL;
4. Incluir *links* para outras coisas relacionadas (usando suas URI) quando publicar dados na *web*.

Mesmo as regras sendo simples, vários dados que foram publicados após a iniciativa não tinham conexões relevantes com outros dados. Com isso, Berners-Lee [29] simplificou ainda mais as regras, ficando da seguinte forma:

1. Todo tipo de conceito de algo, agora terá um nome que começa com HTTP, ou seja, os conceitos e os relacionamentos entre eles estarão acessíveis na *web*. O conceito *cidade* está acessível no domínio da *DBpedia* a partir da URI <http://dbpedia.org/ontology/City>. As *URL* tornaram-se populares como sendo endereços onde podem ser encontrados documentos e outros recursos existentes na *web*. Por outro lado, as URI são formas mais genéricas de identificação de qualquer entidade existente na *web*, seja ela um documento ou uma descrição desse documento;
2. Quando uma URI for recuperada, informações relevantes para as pessoas que procuram por aquela URI também deverão ser trazidas. Após acessar a URI do recurso *Rio de Janeiro* ([http://dbpedia.org/resource/Rio\\_de\\_Janeiro](http://dbpedia.org/resource/Rio_de_Janeiro)), outras informações relevantes sobre ele são trazidas, tais como os pontos turísticos, os bairros e outras localizações que fazem parte da cidade do Rio de Janeiro;
3. Quando recuperar alguma informação, ela não deve ser a altura, o peso ou a data de nascimento de alguém: tem que ter relacionamentos. Quando houver relacionamento, o outro recurso que está sendo relacionada, deve também ter um nome começando com HTTP. Ao acessar o recurso *Rio de Janeiro*, por exemplo, é possível observar os vários relacionamentos existentes entre ele e outros recursos e conceitos disponíveis na *web*. Um dos relacionamentos existentes é a relação *país* (<http://dbpedia.org/ontology/country>) que liga o recurso *Rio de Janeiro* ao recurso *Brasil* (<http://dbpedia.org/resource/Brazil>).

Utilizando-se estas regras, a publicação de dados na *web* deixa de ser da forma tradicional, com disponibilização de dados heterogêneos e com pouca preocupação com sua integração semântica, e passa a ser de forma interligada com outros dados presentes na *web*. Estas ligações são feitas de maneira expressa mediante utilização do padrão RDF, possibilitando a ligação dos documentos publicados e suas informações.

Diante do exposto, para se compartilhar dados na *web* seguindo a iniciativa dos dados abertos interligados, estes devem estar relacionados com outros dados que possuam alguma relevância.

## 2.7 DBpedia

A *DBpedia* é um projeto colaborativo que visa extrair as informações estruturadas da *Wikipedia* e publicá-las na *web*, mantendo as ligações já existentes com os outros recursos e criando novas ligações com outras fontes de informações heterogêneas.

A base de conhecimento da *DBpedia* é totalmente aberta e pode ser utilizada para os mais variados fins. Atualmente, a versão da *DBpedia* em inglês descreve mais de quatro milhões e duzentos mil conceitos que estão classificados em **ontologias**. Também são oferecidas versões em outros 125 idiomas. O conjunto de dados completo é constituído por trinta e oito milhões de conceitos e resumos em 125 línguas diferentes<sup>3</sup>.

O conceito *cidade* é um exemplo de um conceito disponibilizado na *DBpedia*. Na Figura 2.12, apresenta-se a página do conceito *cidade* acessado a partir da URI <http://dbpedia.org/ontology/City>. Algumas informações adicionais ao conceito são trazidas nesta página mediante relacionamentos, tais como um resumo, o nome e os conceitos mais genéricos que o englobam.

**About: city**  
An Entity of Type : [Class](#), from Named Graph : <http://dbpedia.org/resource/classes#>, within Data Space : [dbpedia.org](#)

a relatively large and permanent settlement, particularly a large urban settlement

Property	Value
<a href="#">rdf:type</a>	<a href="#">owl:Class</a>
<a href="#">rdfs:comment</a>	<a href="#">a relatively large and permanent settlement, particularly a large urban settlement</a>
<a href="#">rdfs:isDefinedBy</a>	<a href="http://dbpedia.org/ontology/">http://dbpedia.org/ontology/</a>
<a href="#">rdfs:label</a>	<a href="#">city</a>
<a href="#">rdfs:subClassOf</a>	<a href="#">dbpedia-owl:Settlement</a>
<a href="#">owl:equivalentClass</a>	<a href="http://schema.org/City">http://schema.org/City</a>
<a href="#">wdrs:describedby</a>	<a href="#">dbpedia-owl:data/definitions.xml</a> <a href="#">dbpedia-owl:data/definitions.ttl</a> <a href="#">dbpedia-owl:data/definitions.jsonld</a>
<a href="#">prov:wasDerivedFrom</a>	<a href="http://mappings.dbpedia.org/index.php/OntologyClass:City">http://mappings.dbpedia.org/index.php/OntologyClass:City</a>
<a href="#">is http://open.vocab.org/terms/defines of</a>	<a href="http://dbpedia.org/ontology/">http://dbpedia.org/ontology/</a>
<a href="#">is http://open.vocab.org/terms/describes of</a>	<a href="#">dbpedia-owl:data/definitions.xml</a> <a href="#">dbpedia-owl:data/definitions.ttl</a> <a href="#">dbpedia-owl:data/definitions.jsonld</a>
<a href="#">is rdfs:domain of</a>	<a href="#">dbpedia-owl:CrownDependency</a>
<a href="#">is rdfs:range of</a>	<a href="#">dbpedia-owl:capital</a> <a href="#">dbpedia-owl:city</a> <a href="#">dbpedia-owl:city</a> <a href="#">dbpedia-owl:foundationPlace</a> <a href="#">dbpedia-owl:locationCity</a> <a href="#">dbpedia-owl:belwayCity</a> <a href="#">dbpedia-owl:administrativeHeadCity</a> <a href="#">dbpedia-owl:currentCity</a>
<a href="#">is rdfs:subClassOf of</a>	<a href="#">dbpedia-owl:Capital</a> <a href="#">dbpedia-owl:CapitalOfRegion</a>

Browse using: [OpenLink Data Explorer](#) | [OpenLink Faceted Browser](#) | Raw Data in: [CSV](#) | [RDF](#) ( [N-Triples](#) [N3](#) [Turtle](#) [JSON XML](#) ) | [OData](#) ( [Atom](#) [JSON](#) ) | [Microdata](#) ( [JSON HTML](#) ) | [JSON-LD](#) | [About](#)

This content was extracted from [Wikipedia](#) and is licensed under the [Creative Commons Attribution-ShareAlike 3.0 Unported License](#)

Figura 2.12: Exemplo de conceito disponibilizado pela *DBpedia*.

A *DBpedia* também disponibiliza um serviço denominado *DBpedia Lookup*<sup>4</sup> para recuperar URI de conceitos mediante palavras-chaves. Os resultados são ranqueados de acordo com o número de apontamentos de outras páginas da *Wikipedia*. Ou seja, quanto maior a

<sup>3</sup><http://wiki.dbpedia.org/about>

<sup>4</sup><http://wiki.dbpedia.org/projects/dbpedia-lookup>

quantidade de apontamentos da *Wikipedia* para um conceito, maior a chance deste aparecer entre os primeiros resultados da busca do serviço.

Na Figura 2.13, apresenta-se um exemplo de requisição ao serviço de busca para uma camada que possui o título “*Conservation Areas*”. Quando o texto passado para o serviço de busca for composto por duas ou mais palavras, estas precisam ser separadas pelo caractere “+”.

```
http://lookup.dbpedia.org/api/search.asmx/PrefixSearch?
QueryString=Conservation+Areas
```

Figura 2.13: Exemplo de requisição ao serviço de busca da *DBpedia*

O resultado de uma requisição ao serviço de busca da *DBpedia* é composto pelos registros correspondentes ao texto passado para a operação *QueryString*. Os registros contêm informações como o rótulo, a URI e a descrição do recurso, além das classes e categorias que estão semanticamente relacionadas com ele. No Código Fonte 2.9, apresenta-se um exemplo de um registro retornado pela requisição ilustrada na Figura 2.13. Neste exemplo, não existe nenhuma classe associada ao recurso, porém, há uma categoria correspondente.

#### Código Fonte 2.9: Exemplo de registro retornado pelo serviço de busca da *DBpedia*

```
<Result>
  <Label>Conservation area</Label>
  <URI>http://dbpedia.org/resource/Conservation_area</URI>
  <Description>
    A conservation areas is a tract of land that has been awarded
      protected status in order to ensure that natural features
      , cultural heritage or biota are safeguarded. A
      conservation area may be a nature reserve, a park, a land
      reclamation project, or other area.
  </Description>
  <Classes />
  <Categories>
    <Category>
      <Label>Conservation areas</Label>
      <URI>
        http://dbpedia.org/resource/Category:
          Conservation_areas
      </URI>
    </Category>
  </Categories>
  <Templates />
  <Redirects />
  <Refcount>514</Refcount>
</Result>
```

## 2.8 Ontologia YAGO

*YAGO*<sup>5</sup> é uma **ontologia** desenvolvida em conjunto pelo *DBWeb group* da *Télécom Paris-Tech University*. Trata-se de uma base de conhecimento semântico derivada das informações extraídas das seguintes fontes: *Wikipedia*, *Wordnet*<sup>6</sup> e *Geonames*<sup>7</sup>.

*YAGO* é bastante expressiva, contendo mais de dez milhões de entidades e mais de 120 milhões de relacionamentos. Também possui uma precisão de 95% avaliada manualmente e todas suas relações são anotadas com seu valor de confiança.

A **ontologia** pode ser explorada utilizando-se duas ferramentas de consulta disponibilizadas: *Browse Yago* e *Query YAGO*. Esta última possibilita que a consulta seja realizada utilizando-se as dimensões temporal e espacial.

Como a **ontologia** é de uso livre, são disponibilizados vários arquivos para *download*. É possível baixar toda a **ontologia** ou partes dela. As partes são subdivididas em nove grupos de temas:

- **TAXONOMIA**: Todos os tipos de entidades e a estrutura de classes da *YAGO*. Além das definições formais dos relacionamentos;
- **SIMPLETAX**: Taxonomia mais simples de *YAGO*;
- **CORE**: Os relacionamentos centrais da *YAGO*, tais como os relacionamentos entre as entidades e os relacionamentos que contêm literais;
- **GEONAMES**: Entidades geográficas, classes retiradas do *GeoNames*;
- **META**: Relacionamentos com escopo temporal e espacial juntamente com fontes de extração e estatísticas sobre estes;
- **MULTILINGUAL**: Os nomes multilíngues para as entidades;
- **LINK**: A conexão da *YAGO* para *Wordnet*, *DBPedia*, etc;
- **WIKIPEDIA**: atributos, modelos, fontes, etc dos *infoboxes* multilíngues para os *infoboxes* da *Wikipedia*;

---

<sup>5</sup><http://www.mpi-inf.mpg.de/departments/databases-and-information-systems/research/yago-naga/yago/>

<sup>6</sup><http://wordnet.princeton.edu/>

<sup>7</sup><http://www.geonames.org/>

- **OUTROS:** Características diversas da *YAGO*, tais como *links* de entrada e saída da *Wikipedia*, dados do *GeoNames*, etc.

## 2.9 Considerações Finais do Capítulo

Neste capítulo, foram apresentados alguns conceitos e tecnologias utilizados nesta pesquisa, os quais servirão de base para melhor entendimento da abordagem proposta. O Capítulo 4 fará menção aos conceitos e tecnologias discutidos neste capítulo ao apresentar a ferramenta de **anotação semântica** proposta nesta pesquisa. O próximo capítulo irá apresentar algumas pesquisas relacionadas com esta dissertação.

# Capítulo 3

## Pesquisas Relacionadas

O crescente aumento de dados geográficos disponibilizados na *web* tem demandado o surgimento de novas formas de recuperação desse tipo de informação. Boa parte das pesquisas desenvolvidas propõe a utilização de **ontologias** para associar os dados geográficos disponibilizados com os conceitos providos pelas **ontologias**. A seguir, são apresentadas as pesquisas relacionadas com o estudo apresentado nessa dissertação.

### 3.1 Yingjie et al. [30]

Yingjie et al. [30] propuseram uma solução que possibilita a recuperação de mapas que estão relacionados semanticamente com um determinado tema e o suporte a consultas customizadas. Para isso, os autores apresentaram uma rotina de enriquecimento semântico dos mapas disponibilizados pelo *ArcGIS*<sup>1</sup> com conceitos e entidades nomeadas cujos relacionamentos podem ser legíveis por máquina. Com isso, uma amostra de metadados do *ArcGIS* foi convertida em *Linked Data*. Eles também desenvolveram uma ferramenta de busca semântica que utiliza ontologia na resolução das consultas. Esta ferramenta extrai dinamicamente os termos semânticos e as entidades geográficas correspondentes aos termos que compõem o tema pesquisado. A extração dos termos semânticos e das entidades geográficas são tratadas de forma diferente. Além disso, os autores discutiram duas possíveis consultas flexíveis a partir da solução proposta e mostraram que o conhecimento pode ser descoberto a partir dos metadados interligados. Eles também implementaram um protótipo de portal *web* baseado

---

<sup>1</sup><http://www.arcgis.com/>

em *Linked Data* para o *ArcGIS* que utiliza as soluções apresentadas e define uma prova de conceito para a abordagem proposta por esta pesquisa.

O estudo realizado pelos autores está utilizando um processo de **anotação semântica** de forma automatizada dos dados geográficos disponibilizados no portal geográfico *ArcGIS*. Para isso, eles utilizam dois serviços orientados a *Linked Data* e que estão semanticamente habilitados. Um deles é o *DBpedia Spotlight* que rotula os termos existentes nas informações textuais dos dados geográficos com entradas correspondentes aos conceitos da *DBpedia*. O outro é o *OpenCalais* que extrai e categoriza os termos das informações textuais com conceitos mais amplos. O processo de anotação proposto está limitado aos conceitos da *DBpedia* fornecidos pelo serviço *DBpedia Spotlight* e aos conceitos fornecidos pelo serviço *OpenCalais*. Enquanto que o processo de anotação semântica proposto pela pesquisa desta dissertação possibilita associar os dados geográficos disponibilizados em uma IDE aos conceitos existentes em qualquer ontologia. Sendo necessário para isso, a carga dos conceitos da ontologia na modelagem da solução proposta. Observou-se também que o estudo desenvolvido pelos autores realiza **anotação semântica** apenas no nível de camadas. O processo de anotação proposto pela pesquisa desta dissertação associa os conceitos de uma determinada ontologia tanto para as camadas quanto para seus respectivos atributos, quando existir.

## 3.2 Oliveira et al. [31]

O estudo realizado por Oliveira et al. [31] propõe o desenvolvimento de uma Infraestrutura de Dados Espaciais acessível a partir de um portal geográfico com suporte a diferentes idiomas. Esta IDE possibilita o acesso a informações espaciais sensíveis ao contexto em que o usuário está inserido, ou seja, informações sobre os eventos relacionados com a Copa do Mundo de 2014 e sobre a infraestrutura de alojamentos, alimentação e transporte disponíveis. Também é possível que os usuários possam contribuir voluntariamente com novos dados e informações espaciais sobre eventos, instalações e serviços relacionados à Copa do Mundo, marcando a posição do mesmo no mapa e adicionando detalhes sobre estes. As informações voluntárias são posteriormente validadas e disponibilizadas por uma equipe técnica, com o objetivo de aumentar a base de dados geográfica da Copa do Mundo de 2014.

Embora a IDE proposta pelos autores esteja aplicada ao domínio da Copa do Mundo de

2014, ela também pode ser reutilizada em outros eventos internacionais, como por exemplo, os jogos Olímpicos.

A pesquisa apresentada pelos autores desenvolveu a sua própria ontologia aplicada ao domínio da Copa do Mundo. Os relacionamentos da ontologia desenvolvida foram traduzidos na implementação de um modelo conceitual. Este modelo inclui as principais classes da ontologia de domínio e suas relações hierárquicas e associativas com suas respectivas multiplicidades.

O estudo realizado pelos autores utilizam um catálogo de metadados para gerenciar o conjunto de descrições de dados disponibilizados pela IDE. Conseqüentemente, facilitando a recuperação das informações espaciais providas. Por outro lado, a pesquisa realizada nesta dissertação propõe um processo de **anotação semântica** de forma automática que irá associar os conceitos de uma ontologia aos dados disponibilizados em uma IDE. Uma vez anotados, os dados geográficos são enriquecidos semanticamente, possibilitando uma melhoria na recuperação dos mesmos em motores de busca que utilizam conceitos ontológicos para resolução de consultas temáticas.

### 3.3 Fugazza et al. [32]

Fugazza et al. [32] propõem a construção de uma Infraestrutura de Dados Espaciais para compartilhar dados geográficos no projeto *RITMARE*<sup>2</sup>, que é um projeto mantido pelo governo italiano aplicado à marinha. A IDE proposta é do tipo descentralizada, ou seja, os dados e metadados disponibilizados são armazenados em provedores de dados distintos. O projeto envolve um conjunto heterogêneo de fornecedores de dados (organizações públicas de pesquisa e consórcios inter-universitários), bem como uma variedade de partes interessadas (administração pública, empresas privadas e cidadãos).

As entidades utilizadas para anotar os dados geográficos mantidos pela IDE são extraídas de três componentes: as estruturas de dados que modelam os pesquisadores e as instituições envolvidas no projeto; uma base de conhecimento que inclui uma lista de códigos multilíngue; e representações de feições geográficas extraídas de estruturas de dados *Gazeteers*. As entidades extraídas são mantidas seguindo o padrão RDF e armazenadas em uma instância

---

<sup>2</sup><http://www.ritmare.it/>

do Servidor Universal Virtuoso<sup>3</sup>.

Quanto ao enriquecimento dos registros dos metadados, os autores propõem a anotação destes itens com as entidades RDF mantidas no Virtuoso. Estas entidades são provenientes de fontes da nuvem LOD e possibilitam referenciar as estruturas de Dados Interligados correspondentes. O estudo proposto nesta dissertação propõe um processo de anotação dos dados geográficos disponibilizados em uma IDE que utiliza a nuvem LOD da *DBpedia* como ferramenta de apoio para encontrar os conceitos de uma determinada ontologia que tem ligação semântica com estes dados.

### 3.4 Gimenez et al. [33]

A pesquisa realizada por Gimenez et al. [33] foca no desenvolvimento de uma ontologia geográfica para consolidar os aspectos de metadados espaciais, características espaciais e nomes geográficos. A ontologia proposta serve de base para a integração semântica dos dados espaciais disponibilizados pela Infraestrutura Nacional de Dados Espaciais (INDE). A INDE é uma IDE de âmbito nacional com foco na disseminação da geoinformação do contexto brasileiro.

A fim de proporcionar a integração semântica de bases de conhecimento distintas que contêm dados geográficos, a ontologia geográfica proposta é composta por três sub-ontologias: ontologia dos nomes geográficos ("Identificação Geográfica"), ontologia do perfil de Metadados Geográficos Brasileiro (MGB) ("Metadados Geográficos") e entidades geográficas com base em objetos geográficos.

Os autores propõem uma ontologia para proporcionar o enriquecimento semântico dos dados geográficos disponibilizados pela INDE. O estudo realizado nesta dissertação poderia complementar a pesquisa dos autores, realizando a **anotação semântica** de forma automatizada dos dados geográficos da INDE com os conceitos da ontologia proposta por eles.

---

<sup>3</sup><http://virtuoso.openlinksw.com/>

### **3.5 Borzacchiello e Craglia [34]**

O estudo apresentado por Borzacchiello e Craglia [34] propõe a avaliação dos benefícios frente aos investimentos realizados no desenvolvimento de Infraestruturas de Dados Espaciais. Tendo em vista que a maioria dos provedores de informação geográfica são as organizações públicas e que a literatura não tem muitas evidências sobre o assunto. Os autores realizaram uma pesquisa para comparar, em termos de tempo e custo, os serviços digitais com os serviços tradicionais para o cadastramento de terras nos países da União Europeia. Os objetivos da pesquisa realizada foi estudar se há benefícios decorrentes da utilização dos serviços de governo eletrônico, em comparação com serviços tradicionais e apresentar uma metodologia para estimar estes benefícios.

Ao final da pesquisa, verificou-se que a alternativa digital tem um impacto positivo tanto em seus usuários diretos, quanto nos usuários que escolhem a abordagem baseada no papel tradicional. Isso porque ao utilizarem a alternativa digital, há uma redução significativa de filas nos escritórios de cadatramento.

Esta pesquisa está relacionada com o estudo apresentado nesta dissertação porque está reforçando com uma avaliação experimental que o desenvolvimento da IDE traz benefícios e redução de custos para as organizações provedoras de dados geográficos. Como o estudo apresentado nesta dissertação propõe um enriquecimento semântico dos dados disponibilizados por IDE, a pesquisa realizada pelos autores reafirma a viabilidade desse estudo, pois apresenta resultados que favorecem o ganho real com a utilização de serviços digitais para atender as necessidades da população em geral.

### **3.6 Lutz et al. [35]**

Lutz et al. [35] tiveram como objetivo melhorar o problema da heterogeneidade semântica existente em três níveis de uma IDE: metadados, esquema e conteúdo dos dados. A heterogeneidade semântica existente no nível de metadados impede a descoberta de dados geográficos. Por outro lado, no nível de esquema, ela impede a recuperação dessas informações. Por fim, no nível do conteúdo dos dados, a falta de homogeneidade semântica inviabiliza a interpretação, integração e mudança de dados geográficos. Visando melhorar

estes pontos de impedimento nos três níveis das IDE, foi proposta uma metodologia baseada em **ontologias** e raciocínio lógico.

As **ontologias** utilizadas na abordagem proposta pelos autores são expressas utilizando *Description Logic* (DL), subconjunto da lógica de primeira ordem. Para organizar os conceitos da **ontologia** em uma hierarquia, os autores utilizaram a ideia do *subsumption reasoning*, que define um conceito mais geral do que outro. Os esquemas das fontes de dados são anotados com os conceitos da **ontologia** a partir do *registration mapping* introduzido por Bowers et al [36]. A sua ideia principal é separar as descrições dos conceitos de aplicação, dos detalhes estruturais das camadas. Para utilizar *registration mapping* na anotação das fontes dos dados é necessário criar regras DL que definem as associações entre dados das fontes com os conceitos da **ontologia**.

A abordagem proposta pelos autores tem como pré-requisito que haja um vocabulário compartilhado para o domínio de interesse. Além disso, as fontes de informações precisam estar anotadas semanticamente com **ontologias** de aplicação, utilizando para isso *registration mappings*.

Assim como na pesquisa atual, este estudo também faz uso de **ontologias** para proporcionar a redução de heterogeneidade semântica entre os conceitos existentes na consulta e nas fontes de informações. Entretanto, este estudo toma como premissa que as fontes de informações já estão anotadas semanticamente, enquanto a pesquisa atual propõe uma solução para automatizar o processo de **anotação semântica**. Quanto à definição da **ontologia**, a pesquisa desta dissertação não exige uma tecnologia específica para expressá-la, enquanto que este estudo necessita que a **ontologia** seja expressa em DL.

### 3.7 Janowicz et al. [37]

Janowicz et al. [37] propuseram uma camada semântica transparente e compartilhada para IDE, que integra serviços *reasoning* que são os mecanismos de inferência ligando os componentes-chaves da *web* semântica com os padrões especificados pelo *Open Geospatial Consortium* (OGC). Para isso, são criados perfis de serviços existentes implementando um mapeamento transparente entre o OGC e **ontologias**, combinando IDE com dados vinculados registrados em uma **ontologia**. Os autores apresentaram três objetivos a serem in-

vestigados: ligar codificações de dados e protocolos de serviço a especificações formais armazenadas nas **ontologias**, gerenciar e manter as **ontologias** utilizadas e incorporar serviços *reasoning* ao processo de recuperação dos dados. O primeiro desafio foi superado com a **anotação semântica**, que é realizada mediante uma API de código-aberto denominada *Sapience*<sup>4</sup>. Esta ferramenta possibilita a **anotação semântica** dos metadados de um *web service* a partir de sua conexão com a API. Para resolver os demais desafios, foi proposto o agrupamento das funcionalidades das classes de conformidade em dois serviços: o *Web Ontology Service* (WOS), para gerenciar e acessar **ontologias**, e o *Web Reasoning Service* (WRS), para fornecer funcionalidade de raciocínio dentro da IDE. Os dois serviços são encapsulados em uma única camada. O WOS é um catálogo de suporte de recuperação de informação semanticamente habilitado que encapsula **ontologias**. O WRS encapsula as funcionalidades de raciocínio fazendo a ponte para os motores de inferências entre os componentes chaves da *web* semântica e o mundo OGC.

Este estudo faz uso de **ontologias** para melhorar a resolução de consultas semânticas, usando **anotação semântica** automática a partir da API *Sapience*. Esta API recebe um documento XML serializado com as informações a serem anotadas e devolve o documento atualizado com os conceitos associados da **ontologia** registrada na ferramenta. Contudo, os autores não apresentaram uma análise de desempenho da API utilizada para justificar seu uso. A pesquisa desta dissertação, por sua vez, está propondo uma nova abordagem para gerar as anotações semânticas de forma automatizada com embasamento experimental. Esta abordagem utiliza a *Wikipedia* e a *DBpedia* como ferramentas de apoio para recuperar os conceitos ontológicos semanticamente equivalentes aos dados geográficos disponibilizados por uma IDE.

### 3.8 Li et al. [38]

Li et al. [38] propõem o desenvolvimento de uma infraestrutura de dados espaciais na região multinacional do Ártico para possibilitar o compartilhamento e integração de dados, informações e serviços. Para construção da IDE, três pontos são propostos: (1) uma abordagem híbrida para descobrir serviços de forma eficiente em catálogos *web* distribuídos; (2)

<sup>4</sup><https://code.google.com/p/sapience-api/>

uma base de conhecimento de domínio para modelar os relacionamentos semânticos entre dados e serviços científicos; e (3) um mecanismo de raciocínio lógico inteligente para semiautomatizar a seleção de serviços. A IDE proposta contém cinco componentes-chaves: os repositórios de dados, padrões abertos (WMS, WFS, WCS, etc.), a comunicação entre os dados produzidos e os dados consumidos, políticas e pessoas.

A abordagem híbrida proposta constrói uma ponte que se conecta com os serviços de catálogo distribuídos para descobrir recursos registrados existentes, empregando um *crawler* para localizar dados dispersos na *Internet* e que não estejam registrados nos catálogos.

Na construção da base de conhecimento para modelar a hidrologia do Ártico, foi desenvolvida uma **ontologia** para retirar a ambiguidade existente em diferentes terminologias mediante definições de conceituação explícita. Também foi provida uma ferramenta de busca inteligente para seleção de recursos em um serviço habilitado semanticamente.

O raciocínio semântico é o componente central de um motor de busca semântica. O motor de inferência adotado foi o *framework Jena*<sup>5</sup>, que usa os seguintes procedimentos de processamento: (1) a **ontologia** é carregada na memória ou persistida na base do *framework*; (2) as consultas baseadas em DL são transformadas dentro da formalização das consultas *SPARQL* [39]; (3) a partir do *Jena*, as subconsultas são conduzidas e os resultados são recuperados; (4) os resultados das consultas são combinados para obter informações expandidas e mais específicas.

Para a resolução de ambiguidades entre diferentes terminologias, os autores usaram **ontologias** associadas ao motor de inferência do *framework Jena*, definindo explicitamente uma conceituação. Como a IDE proposta está relacionada com o domínio da Hidrologia foi construída uma **ontologia** para atendê-la.

Com relação à **anotação semântica** dos dados, os autores propuseram uma abordagem híbrida que coleta os serviços disponibilizados em serviços de catálogo ou dispersados na *Internet* sob demanda. À medida que os dados vão sendo encontrados, seus metadados são inseridos no repositório virtual, devidamente classificados com os temas definidos na consulta realizada para descoberta dos mesmos. Este estudo está condicionado apenas ao escopo dos dados e serviços que contém informação sobre a região do Ártico. A pesquisa desta dissertação propõe anotar semanticamente dados geográficos de qualquer domínio,

---

<sup>5</sup><https://jena.apache.org/>

bastando para isso, que a **ontologia** utilizada seja abrangente suficiente.

### 3.9 Macário et al. [40]

O problema abordado por Macário et al. [40] é a ausência de soluções eficientes que permitam o armazenamento das análises e interpretações de dados geográficos realizadas por especialistas. Esta ausência acarretava problemas como retrabalho e dificuldades de compartilhamento de informação. Para resolver este problema, os autores especificaram um *framework* que realiza **anotação semântica** semiautomática de dados geoespaciais, possibilitando um entendimento comum dos conceitos utilizados. Para isso, *workflows* são criados por especialistas de domínio, de forma que cada um seja especializado para a **anotação semântica** de um determinado tipo de informação. Quando um novo recurso é cadastrado no *framework*, o processo de anotação se inicia com a definição de um *workflow* de anotação para o tipo de dado geográfico do recurso cadastrado. Cada *workflow* contém informações sobre o esquema de anotação que será utilizado, as **ontologias** que descreverão os dados e as operações a serem executadas. Ao final da execução de um *workflow*, as anotações semânticas são armazenadas como triplas *Resource Description Framework* [10] (RDF) em uma base de dados XML, ficando disponíveis para consultas utilizando declarações *XQuery*, por exemplo.

Neste estudo, propõe-se a automatização de **anotação semântica** de recursos geográficos com conceitos providos por uma **ontologia**. Esse processo de anotação é possível para os cenários cujos tipos de informação já possuem *workflow* cadastrado. Quando o tipo de informação é desconhecido, é necessário criar um novo *workflow* para atendê-lo ou realizar a anotação de forma manual. Além disso, os autores fazem uso de *workflows* definidos manualmente por especialistas para realizarem a **anotação semântica** dos recursos geográficos, o que torna esse processo custoso. A **anotação semântica** proposta na pesquisa desta dissertação não necessita da avaliação de um especialista, pois utiliza o modelo vetorial para resolver questões de similaridade entre os conceitos de uma **ontologia** e os dados geográficos a serem anotados, sendo o processo de anotação executado automaticamente.

### 3.10 Batcheller e Reitsma [40]

Batcheller e Reitsma [41] focam na ambiguidade semântica existente no conteúdo das consultas e dos dados geográficos. O crescimento destes dados na *web* demanda a existência de ferramentas que facilitem a descoberta dos mesmos. Os autores propõem, para isso, a utilização dos métodos emergentes da *web* semântica. Eles apresentam uma abordagem para expressar e utilizar a semântica no nível de camadas. Tal abordagem foi construída com base em componentes de código aberto e é direcionada para um conjunto de dados geográficos legados. Dessa forma, o uso da semântica formal não requer a transformação dos dados existentes. Inicialmente, o *framework* proposto exige a modelagem do domínio do discurso utilizando **ontologias** a partir de classes OWL. Para isso, foi utilizada a plataforma D2RQ, responsável por mapear as tabelas relacionais com os respectivos atributos dos dados dos sistemas legados em classes OWL e suas respectivas propriedades. Ao requisitar uma consulta ao *framework*, o usuário seleciona qual o conceito da **ontologia** que representa o tipo da informação que ele está procurando, ao que é gerada uma consulta na linguagem *SPARQL*, que é enviada para a ferramenta. A ferramenta, a partir da classe OWL selecionada, consegue identificar as tabelas que contêm as informações desejadas pelo usuário. Os registros das tabelas mapeadas são recuperados e convertidos para triplas RDF. Em seguida, é usado um *reasoning* para selecionar as instâncias que satisfazem as restrições definidas na consulta do usuário. Por fim, as instâncias selecionadas são apresentadas.

Este estudo está voltado para os bancos de dados georreferenciados existentes em sistemas legados e o processo de conversão dos seus dados para triplas RDF será sob demanda, quando forem sendo solicitados. A **anotação semântica** proposta pelos autores utiliza um arquivo de mapeamento entre os conceitos definidos na **ontologia** e as camadas cadastradas no banco de dados. A construção deste arquivo é produzida de forma manual, tornando esse processo uma tarefa custosa. A pesquisa desta dissertação, por sua vez, propõe uma **anotação semântica** de forma automática que não necessita da intervenção humana para resolver os conceitos de uma **ontologia** que são equivalentes a determinados dados geográficos.

Outro ponto em comum entre ambas as pesquisas é que elas propõem a **anotação semântica** no nível de atributos dos dados geográficos, sendo que os autores anotam os atributos das tabelas relacionais dos sistemas legados fazendo uso da plataforma D2RQ, enquanto que

a pesquisa desta dissertação extrai os atributos das camadas disponibilizadas pelos serviços de uma IDE e realiza o mapeamento com os conceitos de uma **ontologia**.

### 3.11 Lutz e Kolas [42]

Lutz e Kolas [42] focam no problema da impossibilidade de se executar consultas expressivas nos serviços de catálogos de uma IDE, além da falta de interoperabilidade entre fontes de informações distintas. Muitas vezes, para se resolver uma consulta, é necessário executar a consulta em mais de uma fonte, o que pode acarretar problemas de heterogeneidade. Para solucionar este problema, os autores propuseram uma metodologia que usa regras para a descoberta de fontes de dados e, baseado nos dados descobertos, responder às questões dos usuários nos serviços de catálogos das IDE.

A metodologia desenvolvida permite inferir relacionamentos entre dados individuais e a combinação deles em diferentes fontes. Tal metodologia é dividida em duas fases: a construção da base de conhecimento e a inferência. Para a fase da construção da base de conhecimento, os autores desenvolveram dois algoritmos que são utilizados para a descoberta de fontes de dados e, conseqüentemente, o preenchimento de uma base de conhecimento. Durante a fase de inferência, regras de domínio são usadas para inferir fatos previamente desconhecidos que podem ser usados para responder a pergunta do usuário. Como os esquemas das fontes de dados distribuídos são mapeados para uma **ontologia** de domínio comum, o usuário não precisa saber ou entender o esquema local, mas pode representar sua consulta e receber a resposta no vocabulário de domínio.

Como as **ontologias** são expressas por regras, a **anotação semântica** das fontes de dados é realizada mediante regras de mapeamento. Essas regras consistem em definir como as camadas de uma IDE podem ser mapeadas para a **ontologia** utilizada. Com isso, as fontes de dados precisam manter um conjunto de regras de mapeamento de forma manual. A pesquisa desta dissertação, por outro lado, não precisa manter nenhuma regra de mapeamento para executar o processo de **anotação semântica** proposto. Necessita apenas das informações textuais que já estão disponíveis nos serviços e nas próprias camadas. Além disso, o processo proposto é executado de forma automática.

### 3.12 Andrade et al. [43]

Andrade et al. [43] propõem um *framework* baseado em informações em nível de camadas e **ontologias** para melhorar a recuperação de dados geográficos oferecidos por uma IDE, superando algumas das limitações dos serviços de catálogos atuais. Para isso, eles usam **anotação semântica** de forma automática e técnicas de recuperação de informação clássica. As páginas da *Wikipedia* são utilizadas durante a **anotação semântica** dos dados geográficos objetivando a solução de quatro problemáticas semânticas: (1) resolução de sinônimos; (2) resolução de acrônimos; (3) categorização; e (4) identificação de palavras flexionadas e relacionadas. O *framework* também propõe um motor de busca que explora os relacionamentos espaciais, temáticos e temporais entre as camadas oferecidas por um serviço para gerar o *ranking* dos resultados recuperados em suas consultas.

Este estudo está vinculado à pesquisa apresentada nesta dissertação, pois os autores tentam melhorar o desempenho de consultas com restrição temática fazendo uso de **anotação semântica** de forma automatizada. A diferença entre estes estudos é a forma como são selecionados os conceitos relevantes a uma determinada camada. Os autores utilizam a *Wikipedia* como ferramenta de apoio, enquanto a pesquisa desta dissertação utiliza a *DBpedia* e a *Wikipedia*. O uso da *DBpedia* é um dos estudos futuros apresentado pelos autores. Outro ponto que a pesquisa atual agrega ao estudo dos autores é a **anotação semântica** automática utilizando **ontologias** aliadas aos *links* semânticos do LOD. Por fim, um outro ponto de relevância da pesquisa desta dissertação em relação ao estudo dos autores é a apresentação da **anotação semântica** de forma automatizada também no nível dos atributos das camadas.

### 3.13 Considerações Finais do Capítulo

Este capítulo apresentou algumas pesquisas relacionadas ao escopo dessa pesquisa, que é a **anotação semântica** automática de dados geográficos. Com o que foi apresentado, é possível verificar que todas as pesquisas revisadas utilizaram **ontologias** para melhorar a recuperação de dados e serviços geográficos. Além disso, alguns estudos também usaram **anotação semântica** dos dados geográficos de forma automatizada. Os estudos supracitados não resolvem completamente os problemas apresentados nesta pesquisa, demonstrando

assim relevância e contribuição para o estado da arte. Como forma de sumarizar a relação dos problemas apresentados nesta pesquisa com os estudos apresentados anteriormente, na Tabela 3.1, compara-se algumas questões envolvidas.

Os significados das colunas da Tabela 3.1 são apresentados abaixo:

- Pesquisa: refere-se as pesquisas apresentadas neste capítulo.
- C1 (Uso de Ontologias): verifica se a pesquisa faz uso de ontologias na solução proposta;
- C2 (Anotação semântica automática): identifica se a solução proposta pela pesquisa realiza anotação semântica de forma automática;
- C3 (Anotação semântica em nível de atributos): identifica se a solução proposta pela pesquisa realiza anotação semântica no nível de atributos;
- C4 (Uso de LOD na resolução de consultas): identifica se a solução proposta pela pesquisa faz uso de LOD na resolução de consultas a objetos geográficos.

Tabela 3.1: Comparação entre as pesquisas apresentados.

<b>Pesquisa</b>	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>C4</b>
<b>Yingjie et al.</b>	sim	sim	não	sim
<b>Oliveira et al.</b>	sim	não	não	não
<b>Fugazza et al.</b>	sim	sim	não	sim
<b>Gimenez et al.</b>	sim	não	não	não
<b>Borzacchiello e Craglia</b>	não	não	não	não
<b>Lutz et al.</b>	sim	não	não	não
<b>Janowicz et al.</b>	sim	sim	não	não
<b>Li et al.</b>	sim	sim	não	não
<b>Macário et al.</b>	sim	não	não	não
<b>Batcheller e Reitsma</b>	sim	não	sim	não
<b>Lutz e Kolas</b>	sim	não	não	não
<b>Andrade et al.</b>	sim	sim	não	não

No próximo capítulo será apresentada a abordagem proposta por esta pesquisa.

# Capítulo 4

## Abordagem Proposta

Neste capítulo, apresenta-se um novo método de **anotação semântica** de camadas oferecidas por serviços de dados geográficos disponibilizados em serviços de catálogos (CSW) de Infraestruturas de Dados Espaciais (IDE). Além disso, o método proposto permite a **anotação semântica** dos atributos de cada camada, proporcionando a melhoria na recuperação destes dados.

Inicialmente, o capítulo apresenta a modelagem de dados proposta para o método de **anotação semântica**. Em seguida, são apresentadas as estratégias utilizadas para anotar semanticamente as camadas e os seus respectivos atributos.

### 4.1 Modelagem de Dados

Nesta subseção, apresentam-se a composição do esquema da modelagem de dados proposta nesta pesquisa e a forma como os dados foram coletados para povoarem as entidades existentes na modelagem.

#### 4.1.1 Esquema Conceitual

A modelagem de dados proposta é composta por seis entidades que serão preenchidas em momentos distintos. As entidades são listadas a seguir juntamente com uma breve descrição de cada uma:

- **Serviço:** esta entidade contém as informações textuais referentes aos serviços disponibilizados em um serviço de catálogo de uma IDE;
- **Camada:** esta entidade contém as informações textuais das camadas disponibilizadas pelas entidades *Serviço* que estão associadas com a mesma;
- **Atributo:** esta entidade contém as informações referentes aos atributos das entidades *Camada* que estão associadas com a mesma;
- **Conceito:** esta entidade contém as informações dos conceitos cadastrados em uma **ontologia** e os seus respectivos rótulos (*labels*);
- **Anotação Camada:** esta entidade contém o mapeamento semântico entre as entidades *Camada* e *Conceito*;
- **Anotação Atributo:** esta entidade contém o mapeamento semântico entre as entidades *Atributo* e *Conceito*.

Na Figura 4.1, apresenta-se o esquema conceitual da modelagem de dados proposta por esta pesquisa e que contempla todas as entidades listadas anteriormente. Cada entidade “*Serviço*” está relacionada com várias entidades “*Camada*” que, por sua vez, estão relacionadas com várias entidades “*Atributo*”. Este último relacionamento é possível apenas quando os serviços forem do tipo WFS. A entidade “*Conceito*”, por sua vez, se relaciona com a entidade “*Camada*” a partir da entidade “*Anotação Camada*” e com a entidade “*Atributo*” a partir da entidade “*Anotação Atributo*”.

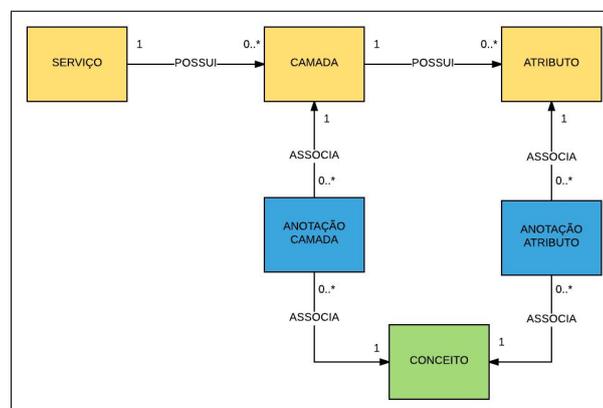


Figura 4.1: Esquema Conceitual da Abordagem Proposta

Na Figura 4.2, apresenta-se a modelagem lógica relacional com as respectivas tabelas. Como as tabelas serão povoadas em momentos distintos, estas foram agrupadas em três classificações: serviços, **ontologia** e **anotação semântica**, conforme legenda.

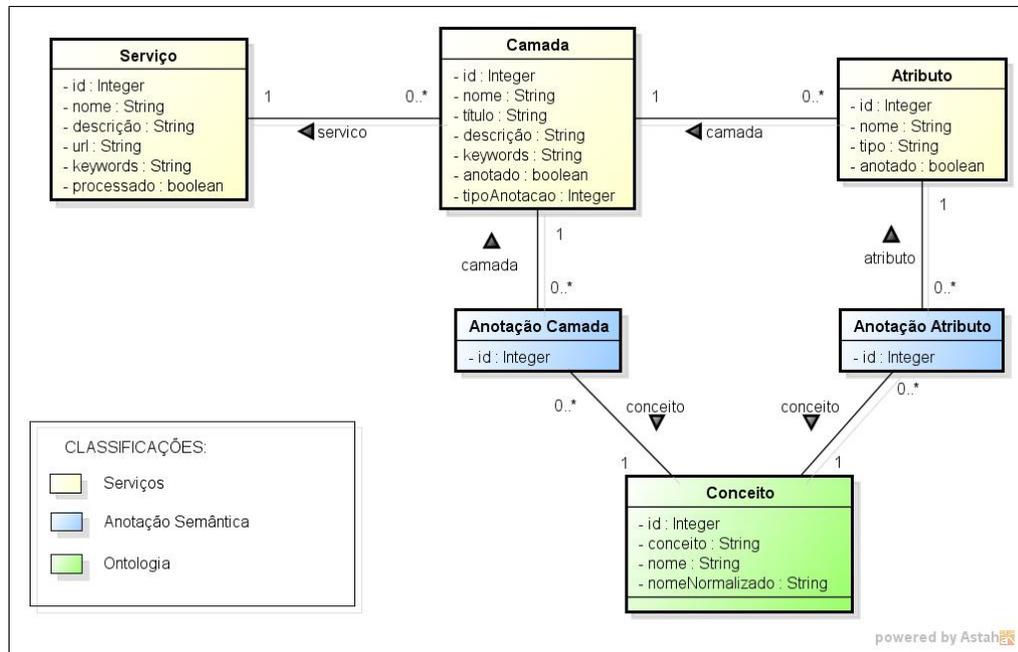


Figura 4.2: Modelagem Lógica Relacional da Abordagem Proposta

A seguir, são apresentadas as tabelas que compõem a modelagem lógica relacional da abordagem proposta. As primeiras tabelas a serem definidas são as que fazem parte dos serviços. Esta classificação é composta pelas tabelas seguintes:

- **Serviço:** Essa tabela contém informações sobre os serviços identificados no catálogo de serviços de uma IDE. Cada registro da tabela armazena o nome, a descrição, a *URL* e as palavras-chaves de um serviço. Ela também contém uma *tag* para identificar se o processo de anotação já processou os dados geográficos referentes àquele serviço. Esta *tag* foi necessária para inclusão de um mecanismo de recuperação a falha que possibilita a reexecução do processo de anotação, caso ocorra algum problema em sua execução;
- **Camada:** Esta tabela contém informações sobre cada uma das camadas disponibilizadas pelos serviços persistidos na tabela “*Serviço*”. Cada registro da tabela irá armazenar o nome, o título, a descrição e as palavras-chaves de uma determinada camada. Ela também contém mais dois atributos: “*anotado*” e “*tipoAnotacao*”. A *tag* “*anotado*”

é marcada sempre que o processo de **anotação semântica** conseguir anotar a camada registrada. Por outro lado, o atributo “*tipoAnotacao*” é preenchido com o valor “1”, quando a camada é anotada mediante a estratégia de LOD, ou com o valor “2”, quando a camada é anotada a partir da estratégia de casamento de *strings*. Caso a camada não seja anotada, o atributo “*tipoAnotacao*” não será preenchido;

- **Atributo:** Esta tabela é alimentada apenas quando o serviço é do tipo WFS, pois apenas este tipo de serviço disponibiliza as informações sobre os atributos das camadas. Cada registro da tabela irá armazenar o nome e o tipo do atributo. A tabela também contém a *tag* “*anotado*” que será marcada sempre que o processo de **anotação semântica** conseguir anotar o atributo registrado. Tanto na tabela “*Camada*” como na tabela “*Atributo*”, a *tag* “*anotado*” também tem a característica de recuperação a falha. Pois, quando ocorrer algum problema durante o processo de **anotação semântica**, as camadas e os seus atributos que já foram anotados não serão mais processados em uma segunda execução do processo. Tendo em vista que este irá utilizar apenas as camadas e os atributos que ainda não tiveram a sua *tag* “*anotado*” marcada.

A classificação da **ontologia** é composta apenas pela tabela “*Conceito*” que armazena as *URL* e os nomes dos conceitos de uma determinada **ontologia**. Além disso, também é armazenado o nome do conceito após um tratamento textual. Tal tratamento consiste em separar palavras compostas divididas por letras maiúsculas (ex.: PalavraComposta) e transformar todas as letras do nome do conceito para minúsculo. A persistência desta informação facilita a busca dos conceitos que são equivalentes a um determinado tema. Tomando-se como exemplo um conceito cujo nome seja “*BodyOfWater*” e o tema pesquisado seja “*body of water*”, esse conceito não seria retornado se não houvesse uma outra opção contendo o nome com o tratamento textual semelhante ao tema pesquisado. As *URL*, os nomes dos conceitos e os nomes dos conceitos pré-processados correspondem aos atributos “*conceito*”, “*nome*” e “*nomeNormalizado*”, respectivamente, na tabela “*Conceito*”.

A classificação da **anotação semântica** é composta pelas tabelas “*Anotação Camada*” e “*Anotação Atributo*”. A primeira contém as relacionamentos entre as camadas armazenadas na tabela “*Camada*” e os conceitos cadastrados na tabela “*Conceito*”. Por outro lado, a segunda contém os relacionamentos entre os atributos armazenados na tabela “*Atributo*” e

os conceitos cadastrados na tabela “*Conceito*”. Estes relacionamentos são resultantes do processo de **anotação semântica** proposto nesta pesquisa.

### 4.1.2 Coleta dos Dados

A seguir, é apresentada a ordem de povoamento das tabelas de cada classificação e a forma como os dados são coletados para povoá-las.

As tabelas dos serviços são as primeiras que precisam ser povoadas. Elas são alimentadas com as informações textuais disponíveis nos serviços OGC providos por IDE. A primeira tabela que precisa ser povoada é a tabela “*Serviço*”. Esta tabela é alimentada com dados de serviço. Esses dados são obtidos junto ao serviço de catálogo da IDE, seguindo o fluxograma apresentado na Figura 4.3.

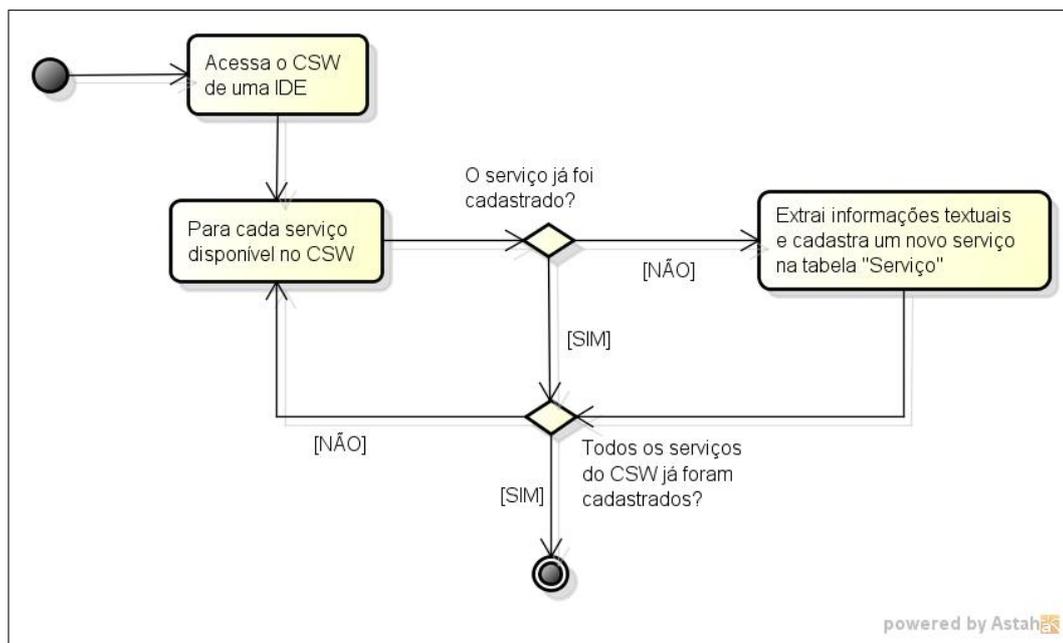


Figura 4.3: Fluxograma de alimentação da tabela “*Serviço*”

Inicialmente, o serviço de catálogo da IDE é acessado e para cada um dos serviços disponibilizados, são extraídas suas informações textuais e armazenadas nessa tabela. Cada registro contém informações sobre um serviço identificado no catálogo. Informações adicionais sobre o serviço, como o nome, a descrição, a *URL* e as palavras-chaves são extraídas a partir dos seus respectivos documentos de funcionalidades, que podem ser obtidos mediante a invocação da operação *GetCapabilities*.

A próxima tabela povoada é a “*Camada*”. Ela contém informações sobre cada uma das camadas disponibilizadas pelos serviços persistidos na tabela “*Serviço*”, conforme fluxograma apresentado na Figura 4.4.

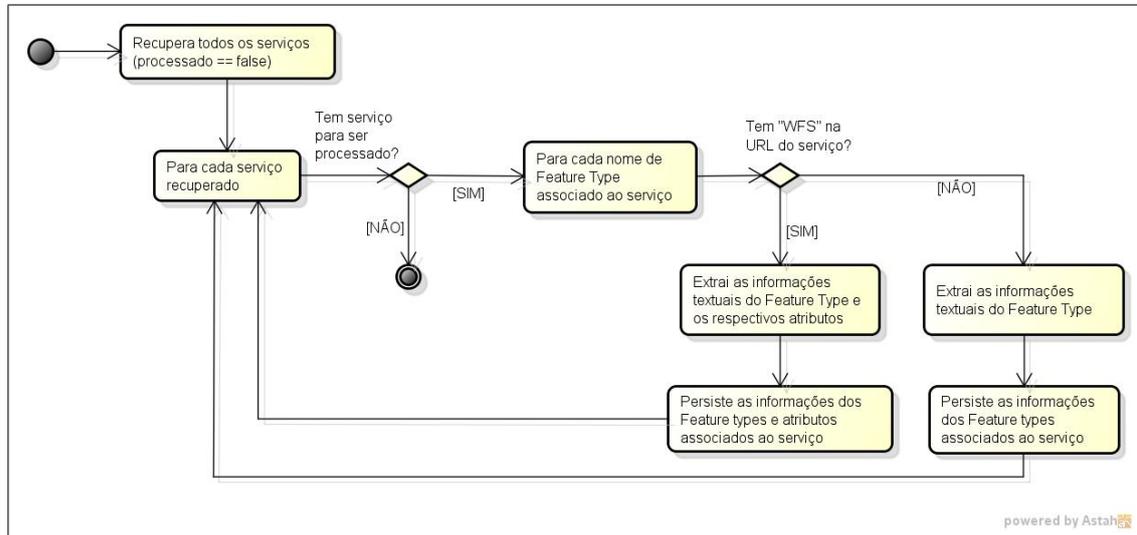


Figura 4.4: Fluxograma de alimentação das tabelas “*Camada*” e “*Atributo*”

O processo de preenchimento desta tabela consiste na extração das informações textuais como nome, título, descrição e palavras-chaves existentes em cada uma das camadas disponibilizadas em um serviço. Estas informações também são encontradas no documento de funcionalidades do serviço.

Após o armazenamento das informações da camada, os dados dos seus respectivos atributos, quando existirem, são persistidos na tabela “*Atributo*”, conforme o fluxograma apresentado na Figura 4.4. Esta tabela é alimentada apenas quando o serviço é do tipo WFS, pois apenas este tipo de serviço disponibiliza as informações sobre os atributos das camadas. As informações do nome e tipo do atributo são extraídas de cada um dos atributos disponibilizados em uma camada. A extração é realizada a partir da operação *DescribeFeatureType* existente nos serviços WFS, que possibilita a recuperação dos atributos de uma determinada camada.

A tabela da classificação da **Ontologia** também pode ser povoada inicialmente, pois a ordem de persistência dela e das tabelas dos serviços não importa. O procedimento para povoar esta tabela consiste em extrair os conceitos existentes em uma **ontologia** e persistí-los na tabela “*Conceito*”. Na Figura 4.5, apresenta-se o fluxograma de alimentação desta tabela.

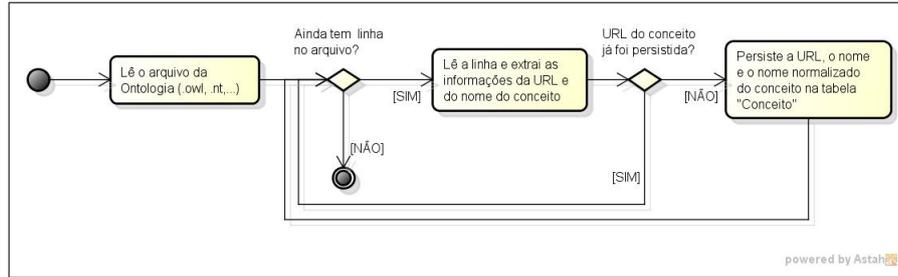


Figura 4.5: Fluxograma de alimentação da tabela “*Conceito*”

Inicialmente é lido o arquivo de extração de uma determinada **ontologia** (estes arquivos geralmente possuem a extensão OWL). Para cada linha do arquivo, são extraídos a *URL* e o nome do conceito. Caso a *URL* ainda não tenha sido persistida na tabela “*Conceito*”, é inserido um novo registro com as informações extraídas e o nome do conceito após o tratamento textual relatado anteriormente.

Para esta pesquisa, decidiu-se utilizar a ontologia *YAGO* [44] para povoar a tabela “*Conceito*”. Tal escolha ocorreu devido à sua grande expressividade, contendo mais de 10 milhões de entidades e mais de 120 milhões de relacionamentos entre elas. Outro fator que levou a escolha foi o seu relacionamento com outros bancos de dados abertos disponibilizados na *Internet*, como a *DBpedia*.

As tabelas da classificação da **anotação semântica** devem ser povoadas após a alimentação das tabelas dos Serviços e da **Ontologia**. Elas usam as informações persistidas nestas duas classificações para realizar o mapeamento dos conceitos das **ontologias** com as camadas existentes no banco dados e, quando possível, com os seus respectivos atributos.

Na Figura 4.6, exibe-se uma visão geral do processo usado para a alimentação da tabela de **anotação semântica** das camadas.

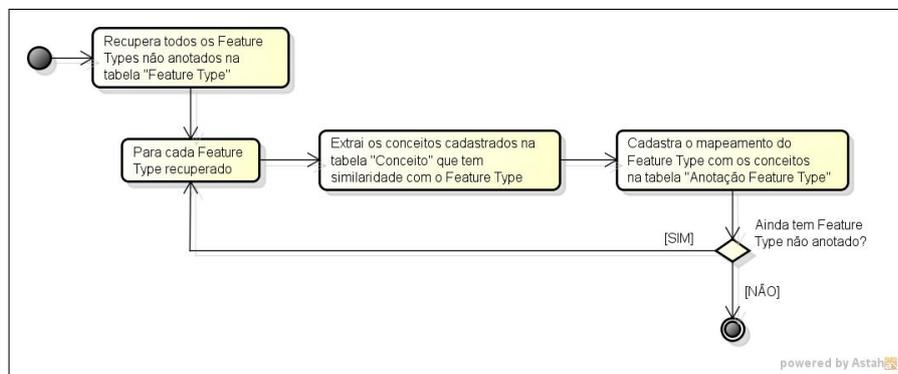


Figura 4.6: Fluxograma de Alimentação da Tabela “*Anotação Camada*”

Inicialmente, são extraídas todas as camadas cadastradas na tabela “*Camada*” que não foram anotadas em uma execução anterior do processo de anotação. Para cada camada recuperada, são identificados os conceitos que se relacionam semanticamente com ela e, em seguida, os relacionamentos identificados são persistidos na tabela “*Anotação Camada*”. O processo de extração dos conceitos semanticamente similares à camada é explicado e detalhado na seção 4.2.

Na Figura 4.7, ilustra-se uma visão geral do processo usado para realizar o preenchimento da tabela de **anotação semântica** dos atributos.

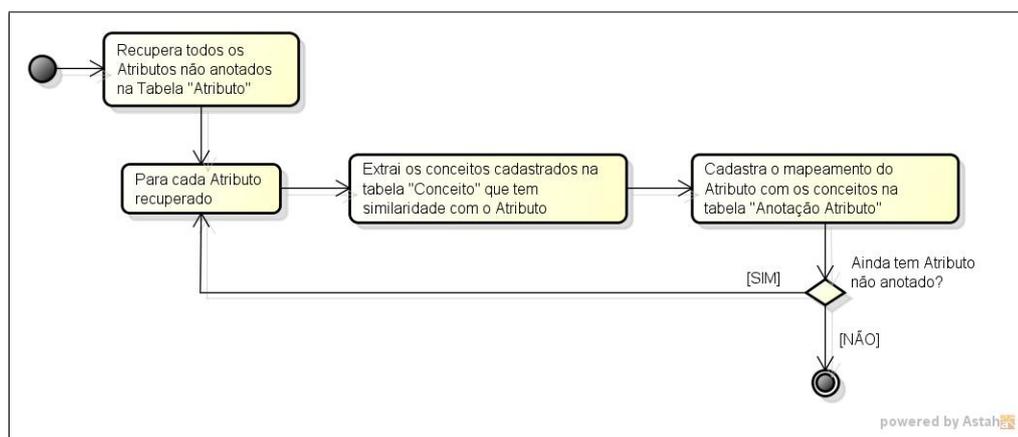


Figura 4.7: Fluxograma de Alimentação da Tabela “*Anotação Atributo*”

Inicialmente, são extraídos todos os atributos cadastrados na tabela “*Atributo*” que estão relacionados com a camada já processada e que não foram anotados em uma execução anterior do processo de anotação. Para cada atributo recuperado, são identificados os conceitos que se relacionam semanticamente com ele e, em seguida, os relacionamentos identificados são persistidos na tabela “*Anotação Atributo*”. Também é importante frisar que o processo de **anotação semântica** de atributos é executado sempre após o término do processo de anotação de uma camada para os atributos do mesmo, quando houver disponibilidade. O processo de **anotação semântica** de atributos é explicado em detalhes na seção 4.3.

## 4.2 Anotação Semântica no Nível de Camadas

O processo de **anotação semântica** das camadas existentes nos serviços OGC consiste em identificar os conceitos de uma **ontologia** que podem ter algum relacionamento semântico

com elas. Uma vez identificados os conceitos, o mapeamento entre eles e a respectiva camada é persistido no banco de dados. Esse processo é realizado de forma automática e é aplicado para todas as camadas cadastradas no banco de dados.

Em linhas gerais, o processo de **anotação semântica** no nível de camadas pode ser dividido em quatro etapas: geração do saco de palavras das informações textuais das camadas; recuperação de classes e categorias associadas à camada na *DBpedia*; seleção das classes e categorias similares à camada; e seleção dos conceitos relacionados semanticamente com a camada. Nas próximas subseções são descritos os detalhes de cada etapa.

### 4.2.1 Geração do saco de palavras das informações textuais das camadas

A **anotação semântica** de uma camada se dá a partir das informações textuais existentes tanto na descrição da camada quanto na descrição do serviço pelo qual o mesmo é disponibilizado. As informações textuais extraídas nas camadas são o nome, o título, a descrição e as palavras-chaves, enquanto nos serviços são extraídas as informações do nome, descrição e palavras-chaves. Tais extrações auxiliarão no processo de anotação.

Uma vez agrupadas, as informações textuais são pré-processadas e inseridas em um modelo de representação textual denominado saco de palavras. Este modelo é definido como sendo uma representação numérica de uma coleção de documentos. O modelo é composto por uma matriz, onde cada linha representa um documento e cada coluna representa um termo. Os valores inseridos nas interseções documento/termo da matriz podem seguir duas abordagens, uma que representa a existência ou não do termo no documento e outra que representa a frequência do termo no documento. Para esta pesquisa, a matriz do saco de palavras contém apenas uma linha, pois se trata de um documento único (as informações textuais da camada) e as interseções documento/termo são preenchidas com valores booleanos representando a existência ou não do termo no documento.

O pré-processamento das informações textuais da camada é inicializado com a tokenização dos seus termos. Esta etapa tem o objetivo de separar os termos existentes nas informações textuais para possibilitar a listagem dos mesmos. À medida que os termos vão sendo separados, é aplicado o filtro de termos não representativos (*stop words*) como pro-

nomes, artigos, advérbios, números, preposições e pontuação. Dessa forma, os termos com estas características são descartados. Se for aplicado um filtro de remoção de *stop words* ao texto “*mapa das cidades alagadas*”, por exemplo, ele resultará em “*mapa cidades alagadas*”, sendo descartada a preposição “*das*”. Outro processamento que é realizado nos termos relevantes é a aplicação do filtro de *stemming*. Esta etapa consiste em reduzir um termo ao seu radical, retirando as suas derivações. Aplicando-se o filtro de *stemming* ao exemplo “*mapa cidades alagadas*” têm-se como resultado “*map cidad lag*”, deixando-se apenas os radicais de cada palavra. Esse processamento se faz necessário para evitar que dois ou mais termos derivados do mesmo radical tenham registros duplicados no saco de palavras. O processo de remoção de *stop words* e a aplicação do filtro de *stemming* foi realizado utilizando-se a ferramenta *Lucene*<sup>1</sup>.

Embora o filtro de *stop of words* tenha sido aplicado às informações textuais das camadas, avaliações experimentais constataram que alguns termos irrelevantes (advérbios, numeral, verbos, etc) não estavam sendo filtrados. Com isso, aplicou-se o processo de seleção dos substantivos e adjetivos do saco de palavras, descartando todas as outras classes gramaticais. Esta etapa foi realizada com o auxílio da ferramenta *Treetagger*<sup>2</sup>.

#### 4.2.2 Recuperação de classes e categorias associadas à camada na DBpedia

Esta etapa do processo de **anotação semântica** de camadas é iniciada com o processo de busca das classes e categorias existentes na *DBpedia* que têm alguma correspondência com o título da camada em questão. O processo inicia-se com a aplicação de um pré-processamento textual no título da camada, com o objetivo de se retirar do título os trechos que não são necessários, sendo considerados apenas os substantivos e adjetivos. As demais classes gramaticais são descartadas, pois os conceitos existentes em uma **ontologia** são nomeados, geralmente, por um substantivo ou um adjetivo. Após a seleção dos substantivos e adjetivos do título da camada, eles são submetidos ao serviço de busca da *DBpedia* (*DBpedia Lookup*). Este serviço possibilita a busca de classes e categorias existentes na *DBpedia* a partir de informações textuais.

---

<sup>1</sup><https://lucene.apache.org/>

<sup>2</sup><http://www.ims.uni-stuttgart.de/projekte/complex/TreeTagger/>

Na Figura 4.8, é apresentado um exemplo de requisição ao serviço de busca para uma camada que possui o título “*Conservation Areas*”. Quando o texto passado para o serviço de busca é composto por duas ou mais palavras, elas precisam ser separadas pelo caractere “+”.

```
http://lookup.dbpedia.org/api/search.asmx/PrefixSearch?
QueryString=Conservation+Areas
```

Figura 4.8: Exemplo de requisição ao serviço de busca da *DBPedia*

O resultado de uma requisição ao serviço de busca da *DBPedia* é composto pelos registros correspondentes ao texto passado para a operação *PrefixSearch*. Os registros contêm informações como o rótulo, a URI e a descrição do recurso, além das classes e categorias que estão semanticamente relacionadas ao mesmo. No Código 4.1, é apresentado um exemplo de um registro retornado pela requisição ilustrada na Figura 4.8. Neste exemplo, não existe nenhuma classe associada ao recurso, porém, há uma categoria correspondente.

#### Código Fonte 4.1: Exemplo de registro retornado pelo serviço de busca da *DBPedia*

```
<Result>
  <Label>Conservation area</Label>
  <URI>http://dbpedia.org/resource/Conservation_area</URI>
  <Description>
    A conservation areas is a tract of land that has been awarded
      protected status in order to ensure that natural features,
      cultural heritage or biota are safeguarded. A conservation area
      may be a nature reserve, a park, a land reclamation project, or
      other area.
  </Description>
  <Classes/>
  <Categories>
  <Category>
    <Label>Conservation areas</Label>
    <URI>http://dbpedia.org/resource/Category:Conservation_areas<
      /URI>
  </Category>
  </Categories>
  <Templates/>
  <Redirects/>
  <Refcount>514</Refcount>
</Result>
```

Caso a busca não retorne nenhum registro para o texto pré-processado do título da camada, esse texto é quebrado em combinações entre os termos vizinhos que o compõe, onde cada combinação gerada é submetida ao serviço *DBpedia Lookup*. Para melhor ilustrar o

algoritmo de quebra das combinações entre as palavras vizinhas do título, será tomado como exemplo o título “*Defined Policy Areas*”. Ele será quebrado nas seguintes combinações: “*Defined Policy*”, “*Policy Areas*”, “*Defined*”, “*Policy*” e “*Areas*”. Cada uma destas combinações é submetida ao serviço de busca e as classes e categorias associadas são recuperadas e agrupadas para o próximo passo do processo de **anotação semântica**.

### 4.2.3 Seleção das classes e categorias similares à camada

Muitas das classes e categorias recuperadas na etapa anterior não têm relação semântica com a camada, ou seja, há muita ambiguidade semântica entre eles. Dessa forma, é necessário aplicar as classes e categorias recuperadas um filtro que selecione apenas os recursos semanticamente relevantes à camada. Para isso, decidiu-se utilizar o modelo vetorial [45] [46], que é um modelo usado na recuperação da informação clássica.

O primeiro passo desta etapa consiste em recuperar as páginas correspondente às classes e categorias da *DBpedia* na *Wikipedia*. Para cada uma das classes e categorias, são obtidas as páginas correspondentes da *Wikipedia* e todo o seu conteúdo textual é confrontado com o saco de palavras da camada em questão, utilizando, para isso, o modelo vetorial que permite calcular a similaridade entre dois documentos a partir das presença ou não dos termos. Para o caso em questão, os dois documentos são o saco de palavras e o conteúdo textual da página da *Wikipedia* referente a uma classe ou a um conceito da *DBpedia*.

Dessa forma, o saco de palavras é representado como um vetor n-dimensional, no qual cada dimensão corresponde a um *token* do saco de palavras. Para dimensão deste vetor é atribuído o valor 1. Por outro lado, o vetor que representa o conteúdo textual da página da *Wikipedia* tem suas dimensões preenchidas com o valor 1, quando o conteúdo textual contém o *token* da dimensão corrente, ou com o valor 0, quando não o contém. Os valores binários foram escolhidos no lugar das métricas *tf* e *idf* (tradicionalmente utilizada no modelo vetorial) porque se percebeu, em experimentos realizados, que as páginas irrelevantes não possuem a maioria dos *tokens* do saco de palavras.

Uma vez que os vetores do saco de palavras e do conteúdo textual da *Wikipedia* são preenchidos, o cosseno do ângulo formado entre os dois vetores é usado para determinar a similaridade de cada classe e categoria com a camada em questão. Na Equação 4.1, apresenta-se a fórmula de extração da similaridade dos dois textos discutidos. Os valores possíveis desta

equação variam entre 0, quando não há qualquer similaridade entre os documentos, e 1, quando os documentos são completamente similares.

$$\text{similaridade}(d_1, d_2) = \frac{\vec{v}(d_1) \cdot \vec{v}(d_2)}{|\vec{v}(d_1)| \cdot |\vec{v}(d_2)|} \quad (4.1)$$

Após o cálculo da similaridade entre o saco de palavras e o conteúdo textual da *Wikipedia*, o valor resultante é submetido a um filtro para que sejam considerados similares apenas os valores que ultrapassarem um determinado limite pré-estabelecido. Este limiar possibilita a filtragem das páginas da *Wikipedia* que são significativamente relevantes com o saco de palavras da camada. Para se chegar ao valor deste limiar foi montado um experimento para calcular a similaridade entre os vetores do saco de palavras e do conteúdo textual da *Wikipedia* a partir da Equação 4.1. O experimento foi executado para várias camadas disponibilizadas nas IDE canadense e GEOSS e extraídas as similaridades para cada uma das páginas da *Wikipedia* correspondentes a uma determinada camada. Os resultados mostraram que o cálculo da similaridade que ultrapassava 0,1 entre os dois vetores, correspondia às páginas da *Wikipedia* que eram semanticamente relevantes à camada em questão. Por fim, as classes e categorias referentes às páginas da *Wikipedia* relevantes são retornadas no final desta etapa.

#### 4.2.4 Seleção dos conceitos relacionados semanticamente com a camada

Esta etapa consiste em recuperar os conceitos da ontologia *YAGO* correspondente às classes e categorias selecionadas na etapa anterior. Após a recuperação dos conceitos, estes são associados à camada em questão. Realizando assim, a **anotação semântica** do mesmo.

Após a seleção das classes e categorias similares à camada em questão, são recuperadas as páginas referentes à cada classe e categoria selecionada na etapa anterior. Para cada página da *DBpedia* acessada, é verificada a existência de alguma *URL* que referencie a ontologia *YAGO* nas propriedades “*owl:sameAs*” e “*rdf:type*”. Existindo alguma referência, as *URL* são armazenadas.

As *URL* da ontologia *YAGO* extraídas das páginas da *DBpedia* das classes e categorias similares à camada são pesquisadas na tabela “*Conceito*” existente na modelagem de dados

proposta. Todos os conceitos recuperados são associados à camada corrente e o mapeamento é persistido na tabela “*Anotação Camada*”. Por fim, a camada tem sua *tag* “*anotado*” marcada e o tipo de anotação atualizado com o valor que representa a anotação realizada a partir de LOD.

Caso não seja encontrada nenhuma referência à ontologia *YAGO* nas páginas das classes e categorias similares à camada, é adotada a estratégia de casamento de *strings* entre os rótulos das classes e categorias similares e os nomes dos conceitos cadastrados na tabela “*Conceito*”. Cada um dos rótulos é pesquisado entre os conceitos cadastrados. Caso não seja encontrado nenhum conceito equivalente ao rótulo e este possua duas ou mais palavras, todas as combinações dos *tokens* também serão pesquisadas junto aos conceitos da **ontologia** materializada na tabela “*Conceito*” da modelagem proposta.

Todos os conceitos recuperados são associados a camada corrente e o mapeamento é persistido na tabela “*Anotação Camada*”. Então, a camada tem sua *tag* “*anotado*” marcada e o tipo de anotação atualizado com o valor que representa a anotação realizada a partir do casamento de *strings*.

Caso a camada corrente não seja relacionada com nenhum conceito após as estratégias de anotação utilizando LOD e casamento de *strings*, este continuará com a sua *tag* “*anotado*” não marcada e o tipo de anotação sem valor associado.

### 4.3 Anotação Semântica no Nível de Atributos

Após o processo de **anotação semântica** de uma determinada camada, caso exista algum atributo associado ao mesmo, é iniciado o processo de **anotação semântica** dos atributos relacionados à camada em questão. O processo de **anotação semântica** no nível de atributos consiste em identificar os conceitos de uma **ontologia** que possuem alguma ligação semântica com eles. Uma vez identificados os conceitos, o mapeamento entre eles e o respectivo atributo é persistido no banco de dados da solução proposta. Esse processo é realizado de forma automática para todos os atributos cadastrados na tabela “*Atributo*” e que estão relacionados com a tabela “*Camada*” da modelagem proposta.

Em linhas gerais, o processo de **anotação semântica** no nível de atributos pode ser dividido em três etapas:

- Seleção dos atributos relacionados com a camada;
- Seleção dos conceitos relacionados semanticamente com cada atributo;
- Persistência do mapeamento entre cada atributo e os respectivos conceitos.

#### 4.3.1 Seleção dos atributos relacionados com a camada

Após a execução do processo de **anotação semântica** de uma determinada camada, seus atributos são recuperados da tabela “*Atributo*” a partir do relacionamento existente entre eles.

Uma vez recuperados os atributos, cada um deles é submetido ao processo de **anotação semântica**. Inicialmente, o nome do atributo recebe dois tratamentos textuais:

- Substituição do caractere “\_” por espaço vazio, caso exista alguma ocorrência no nome do atributo;
- Separação de palavras compostas unificadas por letras maiúsculas, caso exista alguma ocorrência no nome do atributo. Por exemplo, após esse tratamento, o nome de um atributo chamado “*extensãoTerrestre*” é convertido para “*extensão terrestre*”.

#### 4.3.2 Seleção dos conceitos relacionados semanticamente com cada atributo

Após a realização do tratamento textual sob o nome do atributo, este é pesquisado na tabela “*Conceito*” objetivando-se encontrar algum conceito que possua o mesmo nome do atributo.

Caso não seja encontrado nenhum conceito com o mesmo nome do atributo após a realização do tratamento textual e este possua duas ou mais palavras em sua composição, serão extraídas todas as combinações possíveis entre as palavras do nome do atributo. Cada uma das combinações será pesquisada entre os conceitos cadastrados na tabela “*Conceito*”.

#### 4.3.3 Persistência do mapeamento entre o atributo e os conceitos

Uma vez selecionados os conceitos relacionados com o atributo corrente, os mapeamentos entre os mesmos são persistidos na tabela “*Anotação Atributo*” da modelagem proposta. E,

por fim, a tag “*anotado*” do atributo é marcada.

Também é possível que não se encontre nenhum conceito para o atributo em questão. Tal situação pode ocorrer em virtude do nome do atributo não ser entendível (ex.: “ID5MREL”) ou por não existir nenhum conceito com o nome correspondente. Neste caso, a tag “*anotado*” continua não sendo marcada.

Após a finalização desta etapa, o processo de anotação volta para etapa de seleção dos conceitos relacionados, caso ainda existam mais atributos na camada corrente. Alternativamente, o processo de anotação de atributos é finalizado para esta camada.

A fim de melhor ilustrar o processo de anotação dos atributos, suas etapas serão exemplificadas a partir da seleção do atributo de nome “*OFFICIAL\_GRID\_REFERENCE*”. Inicialmente, o seu nome é pré-processado, trocando-se o caractere “\_” pelo caractere vazio, após o que é realizada uma busca na tabela “*Conceito*” pelo nome pré-processado. Como não há nenhum conceito correspondente a ele, os *tokens* que compõe o nome são combinados da seguinte forma: “*OFFICIAL GRID*”, “*GRID REFERENCE*”, “*OFFICIAL*”, “*GRID*” e “*REFERENCE*”. Para cada combinação, é realizada uma busca na tabela “*Conceito*”. As combinações que encontraram conceitos equivalentes foram as compostas por um único *token*. Os conceitos obtidos são persistidos na tabela “*Anotação Atributo*” relacionados ao atributo “*OFFICIAL\_GRID\_REFERENCE*”. Por fim, a tag “*anotado*” do atributo é marcada, caracterizando que o mesmo foi anotado pelo processo de **anotação semântica**.

## 4.4 Refinamento da Anotação Semântica

O processo da **anotação semântica** tanto das camadas quanto dos atributos resulta em muitos conceitos que possuem o mesmo nome. Como por exemplo, a camada “*Allotment Locations*” após o processo de anotação foi associado a vários conceitos que tem o nome “*location*” na ontologia *YAGO*.

Para evitar a replicação de vários conceitos com o mesmo nome na **anotação semântica** das camadas e atributos, foi desenvolvido um tratamento para generalizar os conceitos de mesmo nome que tivessem mais de quatro registros na **anotação semântica**. Tal tratamento consiste em localizar as superclasses dos conceitos que englobam todos ou a maioria dos conceitos de mesmo nome. Decidiu-se aplicar o refinamento para anotações com mais de

quatro conceitos de nomes iguais, pois quando há vários conceitos de mesmo nome em uma ontologia, eles tendem a ser muito especialistas e podem perder a significância semântica com o dado geográfico anotado.

Para aplicar este refinamento, foi montado um experimento para analisar quantitativamente os conceitos distintos que anotavam uma determinada camada ou atributo após o processo de anotação semântica e as superclasses distintas destes conceitos. Observou-se nos resultados que, na grande maioria dos casos, as superclasses dos conceitos convergiam para um conceito generalista. Proporcionando, com isso, uma redução relevante na quantidade de conceitos.

Com este refinamento, houve uma redução significativa da quantidade de conceitos anotados, mantendo-se a significância semântica com os dados geográficos anotados.

## 4.5 Considerações Finais do Capítulo

Neste capítulo, foi apresentada a abordagem proposta nesta pesquisa, que tem como objetivo melhorar a recuperação de dados geográficos mediante a **anotação semântica** automática deste tipo de informação. Inicialmente, foi apresentada a modelagem de dados proposta e a forma como as tabelas devem ser povoadas. Em seguida, descreveu-se como ocorre o processo de **anotação semântica** no nível de camadas e as etapas que compõem esse processo. Por fim, foi apresentado o processo de **anotação semântica** no nível de atributos das camadas.

Após a apresentação da abordagem proposta e considerando os estudos relacionados com esta pesquisa apresentados no capítulo 3 é possível compará-los na tabela 4.1. Ao analisá-la, verifica-se que apenas a solução proposta utiliza as quatro estratégias para possibilitar a melhoria na recuperação de dados geográficos: C1 (uso de ontologias), C2 (anotação semântica automática), C3 (anotação semântica em nível de atributos) e C4 (uso de LOD na resolução de consultas). Esta comparação possibilita afirmar a relevância e viabilidade desta pesquisa em relação aos estudos relacionados apresentados.

Tabela 4.1: Comparação entre as pesquisas apresentadas e a pesquisa desta dissertação.

<b>Pesquisa</b>	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>C4</b>
<b>Yingjie et al.</b>	sim	sim	não	sim
<b>Oliveira et al.</b>	sim	não	não	não
<b>Fugazza et al.</b>	sim	sim	não	sim
<b>Gimenez et al.</b>	sim	não	não	não
<b>Borzacchiello e Craglia</b>	não	não	não	não
<b>Lutz et al.</b>	sim	não	não	não
<b>Janowicz et al.</b>	sim	sim	não	não
<b>Li et al.</b>	sim	sim	não	não
<b>Macário et al.</b>	sim	não	não	não
<b>Batcheller e Reitsma</b>	sim	não	sim	não
<b>Lutz e Kolas</b>	sim	não	não	não
<b>Andrade et al.</b>	sim	sim	não	não
<b>Abordagem Proposta</b>	sim	sim	sim	sim

No próximo capítulo será apresentada a avaliação experimental utilizando-se a abordagem proposta.

# Capítulo 5

## Avaliação Experimental

Neste capítulo, é apresentado o processo de avaliação experimental da abordagem proposta no Capítulo 4. Os experimentos foram elaborados com o objetivo de mensurar a qualidade das anotações semânticas realizadas pela abordagem desenvolvida nesta pesquisa. Inicialmente, apresenta-se a preparação do banco de dados usado para a execução dos testes. Em seguida, são abordadas as estratégias de execução do experimento para as anotações semânticas no nível de camadas e no nível de atributos.

### 5.1 Visão Geral da Avaliação Experimental

A avaliação experimental da solução proposta visa mensurar o seu desempenho sob uma amostra de dados geográficos anotados por ela. Para isso, foi realizada uma carga de dados reais no banco de dados da solução. Esta carga foi extraída dos serviços WFS disponibilizados no apêndice E. A escolha por estes tipos de serviços foi embasada na necessidade de avaliar tanto o desempenho da anotação semântica no nível de camadas quanto no nível de atributos. Uma vez que estes serviços disponibilizam ambos os tipos de dados geográficos. A extração dos dados geográficos destes serviços resultou em 100 camadas e 1.119 atributos. Após a realização da carga, o processo de **anotação semântica** proposto foi executado.

As estratégias de avaliação tanto para **anotação semântica** de camadas quanto de atributos utilizaram as métricas de precisão e cobertura, amplamente difundidas para avaliar o desempenho de sistemas de recuperação da informação. A extração dessas métricas exige a disponibilização prévia dos conceitos relevantes que devem anotar as camadas e os atributos.

O processo de extração das métricas será contemplado nas próximas seções.

## 5.2 Geração do Banco de Dados

A carga dos dados geográficos (camadas e atributos) se faz necessária para produzir insumos que possibilitem a execução do processo de **anotação semântica** proposto e, conseqüentemente, a sua avaliação experimental.

A montagem do banco de dados para a execução do experimento foi realizada em duas etapas. Primeiro, foi realizada a carga dos dados na tabela “*Conceito*” utilizada pela solução proposta. Em seguida, foi realizada a carga das tabelas dos serviços que contêm os dados geográficos a serem anotados no experimento.

A carga da tabela que representa a **ontologia** é necessária para proporcionar o registro dos conceitos definidos em uma **ontologia**. Para realizar esta carga foi feito, inicialmente, o *download* do arquivo em que a ontologia *YAGO* está mapeada em triplas<sup>1</sup>, com os conceitos e os seus respectivos nomes na língua inglesa.

Uma classe *Java* foi desenvolvida para ler o arquivo das triplas e selecionar apenas aquelas que tivessem o relacionamento de *label*. Durante a leitura do arquivo, sempre que algum relacionamento deste tipo era encontrado, um novo registro era inserido na tabela “*Conceito*”, com a URI e o *label* do relacionamento representando, respectivamente, o conceito e o nome do registro. O nome normalizado do registro na tabela “*Conceito*” é preenchido com a informação do *label* após o tratamento textual que consiste em separar palavras compostas divididas por letras maiúsculas (ex.: *BodyOfWater*) e transformar todas as letras do *label* para minúsculo.

Para realizar a carga das tabelas do domínio de serviços, foram utilizados 31 serviços WFS disponibilizados em algumas IDE acessíveis na *web* cujas informações textuais dos dados geográficos são disponibilizadas na língua inglesa. A escolha por serviços de mais de uma IDE foi necessária para reduzir possíveis viés nos dados geográficos que compõem a amostra. O apêndice E contempla todas as *URL* destes serviços. A validação foi feita apenas com serviços WFS porque apenas ele provê tanto as informações das camadas quanto as informações de seus respectivos atributos, que são necessárias para a avaliação da abordagem

---

<sup>1</sup>[http://downloads.dbpedia.org/3.2/en/yago\\_en.nt.bz2](http://downloads.dbpedia.org/3.2/en/yago_en.nt.bz2)

proposta.

Inicialmente, a tabela “*Serviço*” foi preenchida com os dados dos serviços WFS. As informações de cada serviço (nome, descrição e palavras-chaves) foram obtidas mediante seus respectivos documentos de funcionalidades e preenchidas de forma manual.

As cargas das tabelas “*Camada*” e “*Atributo*” foram realizadas a partir de um módulo em *Java* responsável pela leitura dos serviços persistidos e, a partir das suas respectivas *URL*, extrair as informações das camadas e os seus respectivos atributos. Para realizar a leitura destas informações, foi utilizada a API de código aberto *GeoTools*<sup>2</sup>. Obtidas as informações, estas foram persistidas nas tabelas “*Camada*” e “*Atributo*” com os seus respectivos relacionamentos. Após a execução do módulo de carga das tabelas “*Camada*” e “*Atributo*”, as camadas e os atributos da amostra foram persistidos, respectivamente.

Uma vez executada a carga das tabelas dos domínios da **ontologia** e dos serviços, o algoritmo de **anotação semântica** para o nível de camadas e de atributos foi aplicado à mesma. Os resultados das anotações serão detalhados e explicados nas próximas seções.

### 5.3 Anotação Semântica no Nível de Camadas

Para validar os resultados da **anotação semântica** no nível de camadas foram utilizadas as métricas de cobertura e precisão. No contexto do experimento, a cobertura é obtida mediante a proporção entre a quantidade de conceitos relevantes anotados para uma determinada camada e a quantidade de conceitos relevantes existentes na **ontologia** (5.1). Por outro lado, a precisão é obtida a partir da proporção entre a quantidade de conceitos relevantes anotados e a quantidade total de conceitos anotados pela abordagem proposta (5.2).

$$cobertura = \frac{quantidadeDeConceitosRelevantesAnotados}{quantidadeDeConceitosRelevantesExistentes} \quad (5.1)$$

$$precisao = \frac{quantidadeDeConceitosRelevantesAnotados}{quantidadeTotalDeConceitosAnotados} \quad (5.2)$$

Para que houvesse a avaliação das métricas de desempenho, foi necessária a criação de uma linha de base contendo todos os conceitos relevantes para cada uma das camadas usadas

<sup>2</sup><http://www.geotools.org/>

durante a validação. A linha de base foi montada em um arquivo texto contendo 84 camadas e os respectivos *ids* dos conceitos relevantes. Cada linha do arquivo continha o título da camada, seguido do caractere “|” e os *ids* dos conceitos relevantes separados por vírgula. No Apêndice A é ilustrada a composição da linha de base, apresentando o título das camadas e os nomes e as URI dos conceitos relacionados semanticamente com os mesmos.

Após a geração da linha de base e da execução do processo de **anotação semântica**, foram extraídas as métricas de qualidade para cada camada. Para isso, foi desenvolvida uma classe *Java* para realizar o processamento de extração das métricas de precisão e cobertura. O processo lia o arquivo da linha de base e para cada linha recuperava a camada a partir do seu título e as variáveis necessárias para o cálculo das métricas. Estas variáveis eram:

- **Conceitos relevantes existentes na ontologia:** recuperam os conceitos dos *id* registrados na linha de base para a camada;
- **Conceitos selecionados pela abordagem:** recuperam todos os conceitos que foram selecionados para a anotação da camada que estava sendo avaliada;
- **Conceitos relevantes selecionados pela abordagem:** extraem, dentre todos os conceitos que foram selecionados durante a anotação pela abordagem proposta, apenas os que eram relevantes, de acordo com as informações definidas na linha de base.

Após a extração das variáveis supracitadas, as métricas de cobertura e precisão foram calculadas para cada camada usada na validação. Ao final do processamento, algumas informações referentes à extração das métricas foram escritas em um arquivo de *log*. As informações foram estruturadas da seguinte forma para cada resultado das camadas:

- Título da camada e o tipo de anotação a qual foi submetido (via casamento de *strings* ou via LOD);
- Os nomes dos conceitos relevantes e os respectivos *ids*;
- Os nomes dos conceitos selecionados pela abordagem e os respectivos *ids*;
- Os URI dos conceitos relevantes selecionados pela abordagem;
- Os valores obtidos para a precisão e para a cobertura.

Na Figura 5.1, é apresentado o exemplo de um registro de *log* para a camada de título “*Conservation Areas*”, conforme estruturação definida acima.

```

-----
Título: Conservation Areas - Tipo de Anotação: Anotação via Casamento de Strings
Conceitos relevantes: [conservation-207319, conservation-207320]
Conceitos recuperados: [conservation-207319, conservation-207320]
Conceitos relevantes recuperados: [http://dbpedia.org/class/yago/Conservation100819274,
                                  http://dbpedia.org/class/yago/Conservation106104073]

Precisão: 1.0
Cobertura: 1.0
-----

```

Figura 5.1: Exemplo de um registro de *log* na anotação de camadas

A primeira linha corresponde ao título da camada seguido do tipo de anotação a que foi submetida. A segunda, terceira e quarta linhas correspondem, respectivamente, aos conceitos cadastrados na linha de base, aos conceitos selecionados pela abordagem proposta e aos conceitos relevantes que foram recuperados pela abordagem. Como todas as variáveis dos cálculos das métricas de cobertura e precisão têm os mesmos conceitos, tanto uma quanto a outra tiveram os seus valores máximos.

Os resultados das métricas de desempenho para as anotações semânticas das camadas registradas na linha de base são apresentados no Apêndice C. Para uma melhor ilustração destes resultados, foram gerados dois gráficos associando a quantidade de camadas com o valor de cada uma das métricas de desempenho.

Na Figura 5.2, apresentam-se os valores da métrica de precisão para a **anotação semântica** no nível de camadas e a respectiva quantidade de camadas que obtiveram estes valores.



Figura 5.2: Resultado da precisão na anotação de camadas

Os resultados mostram que a métrica 1,00 (100% de precisão) foi extraída para 70 das 84 camadas da linha de base, enquanto que a métrica 0,00 (0% de precisão) foi extraída para apenas 3. A maior parte das camadas atingiu uma precisão de 100%, o que indica que a maioria dos conceitos anotados pela abordagem proposta estava entre os conceitos definidos na linha de base. A média de precisão entre os resultados das anotações semânticas no nível de camadas foi de 76%. Apenas nove das 84 camadas da amostra tiveram precisão inferior a 50%. Os três casos que tiveram a precisão zero foram as camadas de títulos “*Coastal waterbodies*”, “*Licenced IPPC Facilities*” e “*NapierCityCouncilImContours*”. O primeiro caso não foi anotado porque a **ontologia** não provê um conceito com o nome “*waterbody*”, mas contém conceitos com o nome “*body of water*”. Estes eram os conceitos esperados após a anotação, porém, não foi selecionado nenhum. O segundo caso não foi anotado porque ao aplicar o filtro de stemming ao título, não foi possível encontrar nenhuma classe ou categoria relevantes na DBpedia. O terceiro e último caso foi prejudicado pelo trecho “*Im*” no meio de seu título.

Na Figura 5.3, apresenta-se os valores da métrica de cobertura para a **anotação semântica** no nível de camadas e a respectiva quantidade de camadas que obtiveram estes valores.

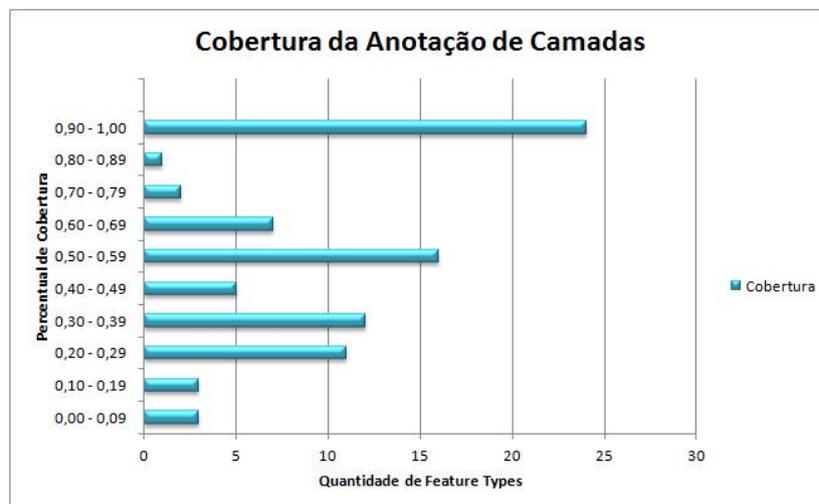


Figura 5.3: Resultado da cobertura na anotação de camadas

Por outro lado, a cobertura não atingiu resultados tão satisfatórios quanto à precisão, pois os resultados variam bastante entre as camadas, chegando a uma média de 50%. A métrica 1,00 (100% de cobertura) foi extraída para 24 das 84 camadas da linha de base, em contrapartida, a métrica 0,50 (50% de cobertura) para 15 e a métrica 0,00 (0% de cobertura) para

3. Isso possibilita aferir que os conceitos anotados pela abordagem proposta correspondem à metade dos conceitos listados na linha de base. Para que a cobertura atinja um percentual elevado, é necessário que todos os conceitos cadastrados na linha de base tenham anotado as camadas. O que se observou no experimento foi que parte desses conceitos não anotaram as camadas, rebaixando a métrica da cobertura.

Com relação à existência de conceitos anotados para uma determinada camada, verificou-se que entre os 84 registros que compõem a linha de base, apenas 3 deixaram de ser anotados. Tais situações ocorreram devido à baixa similaridade entre as informações textuais deles e as páginas da *Wikipedia* das classes e categorias selecionadas, consequentemente, não foi possível selecionar os conceitos semanticamente relevantes.

Na Figura 5.4, apresenta-se a quantidade de anotações realizadas para cada um dos tipos de anotações utilizadas pela abordagem proposta.



Figura 5.4: Quantidade de anotações X tipo de anotação

Com relação ao tipo de anotação realizada sob as camadas, seis foram anotados utilizando a estratégia via LOD e 75 foram anotados utilizando a estratégia de casamento de *strings*. A baixa quantidade de anotações via LOD ocorreu devido à pouca frequência de *links* semânticos entre as classes/categorias selecionadas e a ontologia *YAGO*, ou seja, após a seleção das classes e categorias da *DBpedia* que eram similares às informações textuais da camada, suas páginas não continham nenhum relacionamento com a ontologia *YAGO*.

Outro ponto que pode ser destacado dos resultados do experimento é que a qualidade das anotações realizadas via casamento de *strings* foi superior em relação às anotações via LOD. A precisão média das anotações via LOD (45%) ficou abaixo da média geral das anotações.

Por outro lado, a cobertura média deste tipo de anotação (55%) foi um pouco melhor que a média geral.

Diante dos resultados apresentados, pode-se concluir que a **anotação semântica** realizada no nível de camadas foi bastante satisfatória, uma vez que foram anotadas 96,4% das camadas da linha de base e os conceitos anotados foram relevantes em sua grande maioria.

## 5.4 Anotação Semântica no Nível de Atributos

Para validar os resultados da **anotação semântica** no nível de atributos, também foram utilizadas as métricas de cobertura e precisão. No contexto do experimento, a cobertura era obtida mediante a proporção entre a quantidade de conceitos relevantes anotados para um determinado atributo e a quantidade de conceitos relevantes existentes na **ontologia**. Por outro lado, a precisão é obtida a partir da proporção entre a quantidade de conceitos relevantes anotados e a quantidade total de conceitos anotados pela abordagem proposta.

Assim como no experimento da anotação no nível de camadas, também foi criada uma linha de base contendo todos os conceitos relevantes para cada atributo usado durante o processo de validação. Para sua composição, foram priorizados os atributos com nomes legíveis, tendo em vista que há alguns atributos cadastrados em serviços WFS que não são morfologicamente compreendíveis. Nos atributos carregados no banco de dados utilizado pela avaliação experimental, é possível encontrar os exemplos “*HBNUM*”, “*EMEP\_REF*”, “*AQIHTEXT*” e “*AQIHLASTKNOWN*”. Com isso, a linha de base foi montada em um arquivo de texto contendo 100 atributos e os respectivos *id* dos conceitos relevantes. Cada linha do arquivo era formada pelo nome do atributo, seguido do caractere “|” e os *ids* dos conceitos relevantes separados por vírgula. No Apêndice B, é ilustrada a composição da linha de base apresentando o nome dos atributos e os nomes e as URI dos conceitos relacionados semanticamente com os mesmos.

Após a geração da linha de base, foram extraídas as métricas de qualidade para cada um dos atributos que a compõe. Para isso, foi utilizada a mesma estratégia do experimento da anotação no nível de camadas, recuperando as seguintes variáveis necessárias para o cálculo das métricas:

- **Conceitos relevantes existentes na ontologia:** representam os conceitos dos *id* regis-

trados na linha de base para o atributo;

- **Conceitos selecionados pela abordagem:** representam todos os conceitos que foram anotados para o atributo pela abordagem proposta;
- **Conceitos relevantes selecionados pela abordagem:** extraem, dentre todos os conceitos que foram anotados pela abordagem proposta, apenas os que eram relevantes, de acordo com as informações definidas na linha de base.

Ao final do processamento, algumas informações referentes à extração das métricas também são escritas em um arquivo de *log*, sendo estruturadas da seguinte forma para cada resultado dos atributos:

- Nome do atributo;
- Os nomes dos conceitos relevantes e os respectivos *id*;
- Os nomes dos conceitos selecionados pela abordagem e os respectivos *id*;
- Os URI dos conceitos relevantes selecionados pela abordagem;
- Os valores obtidos para a precisão e para a cobertura.

Na Figura 5.5, é apresentado o exemplo de um registro de *log* para o atributo de nome “CLASSIFICATION”, conforme estruturação descrita acima. A primeira linha corresponde ao nome do atributo. A segunda, terceira e quarta linhas correspondem, respectivamente, ao conceito cadastrado na linha de base, ao conceito selecionado pela abordagem proposta e ao conceito relevante que foi recuperado pela abordagem. Como todas as variáveis dos cálculos das métricas de cobertura e precisão têm os mesmos conceitos, tanto uma quanto a outra tiveram os seus valores máximos.

```
-----  
Atributo: CLASSIFICATION  
Conceitos relevantes: [compartmentalization-169276]  
Conceitos recuperados: [compartmentalization-169276]  
Conceitos relevantes recuperados: [http://dbpedia.org/class/yago/Categorization101012712]  
  
Precisão: 1.0  
Cobertura: 1.0  
-----
```

Figura 5.5: Exemplo de um registro de *log* na anotação de atributos

Os resultados das métricas de desempenho para as anotações semânticas dos atributos registrados na linha de base são apresentados no Apêndice D. Para uma melhor ilustração destes resultados, foram gerados dois gráficos associando a quantidade de atributos com o valor de cada uma das métricas de desempenho.

Na Figura 5.6, são apresentados os valores da métrica de precisão para a **anotação semântica** no nível de atributos e a respectiva quantidade de atributos que obtiveram estes valores. A métrica 1,00 (100% de precisão) foi extraída para 98 dos 100 atributos da linha de base, enquanto que a métrica 0,00 (0% de precisão) foi extraída para apenas 2.

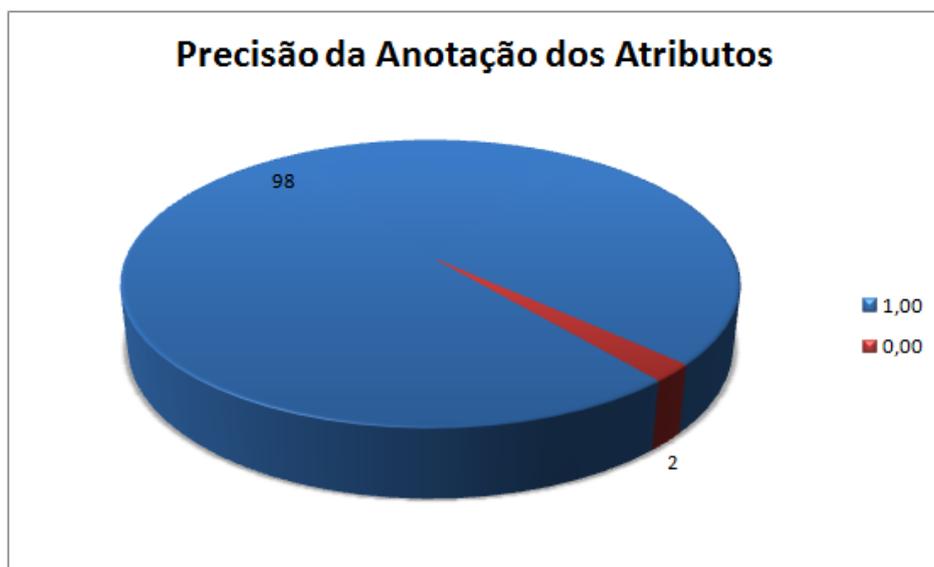


Figura 5.6: Resultado da precisão na anotação de atributos

Assim como aconteceu com os resultados da anotação no nível de camadas, estes resultados mostram que a grande maioria dos atributos atingiu uma precisão de 100%, isso significa dizer que a maioria dos conceitos anotados pela abordagem proposta estavam entre os conceitos da linha de base. A média de precisão entre os resultados das anotações semânticas no nível de atributos foi de 98%, superando a precisão média dos resultados do experimento para anotação no nível de camadas.

Na Figura 5.7, são apresentados os valores da métrica de cobertura para a **anotação semântica** no nível de atributos e a respectiva quantidade de atributos que obtiveram estes valores. A métrica 1,00 (100% de cobertura) foi extraída para 73 dos 100 atributos da linha de base, por outro lado, a métrica 0,50 (50% de cobertura) para 10 e a métrica 0,00 (0% de cobertura) para 2.

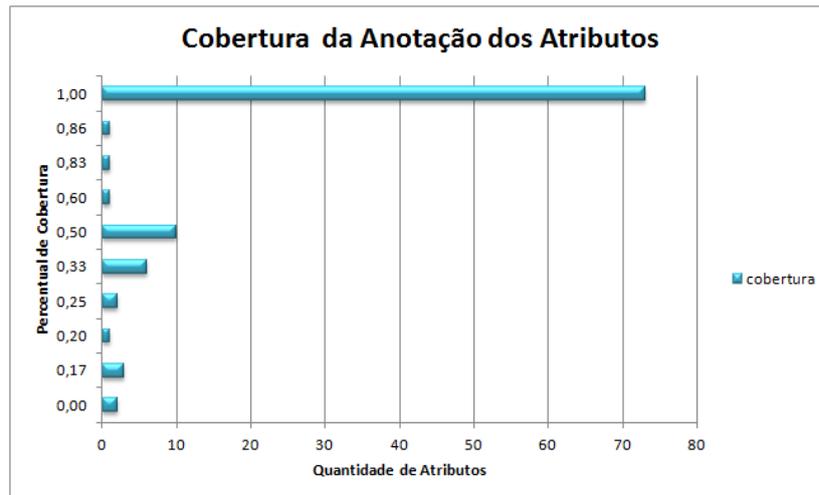


Figura 5.7: Resultado da cobertura na anotação de atributos

A cobertura também atingiu resultados satisfatórios, chegando a uma cobertura média de 76%. Isso possibilita afirmar que os conceitos anotados pela abordagem proposta correspondem a um pouco mais de três quartos dos conceitos listados na linha de base.

Com relação à existência de conceitos anotados para um determinado atributo, verificou-se que todos os 100 registros que compõem a linha de base foram anotados. Também é importante deixar registrado que existe a possibilidade de alguns atributos não serem anotados devido à falta de legibilidade do seu nome. Esta situação é bem comum na nomeação de atributos disponibilizados em serviços WFS. Ao final do processo de anotação, dos 1.119 atributos cadastrados no banco de dados, 728 foram anotados.

Diante dos resultados apresentados, pode-se concluir que a **anotação semântica** realizada no nível de atributos foi bastante satisfatória para a amostra considerada, uma vez que foram anotados 100% dos atributos da linha de base e os conceitos anotados foram relevantes em sua grande maioria.

## 5.5 Considerações Finais do Capítulo

Neste capítulo, apresentou-se a avaliação experimental da abordagem proposta por esta pesquisa. No decorrer do capítulo, foi apresentada a geração do banco de dados para a execução do experimento, seguido dos detalhes do processo de execução do experimento tanto da anotação no nível de camadas, quanto no nível de atributos.

Os resultados da avaliação experimental foram satisfatórios, tendo-se em vista que a maior parte das camadas e atributos foram anotados com conceitos relevantes de acordo com as suas respectivas linhas de base. Os resultados da anotação no nível de camadas computaram uma precisão média de 76%, o que significa dizer que a maior parte dos conceitos selecionados pela anotação no nível de camadas são relevantes para a amostra utilizada. Por outro lado, a cobertura média foi inferior, computando 50%, ou seja, apenas metade dos conceitos que eram previstos para anotarem as camadas da amostra do experimento foram efetivamente selecionados pela solução proposta. Esta última métrica não rebaixa a eficácia da solução proposta para a amostra realizada no experimento, pois de acordo com estes resultados, os conceitos que anotaram as camadas foram bem precisos e atenderam a uma cobertura razoável do que estava previsto na linha de base.

Os resultados da **anotação semântica** no nível da atributos foram bem melhores em relação aos resultados da anotação no nível de camadas. Eles computaram uma precisão média de 98%, o que significa dizer que quase a totalidade dos conceitos selecionados pela anotação são relevantes para a amostra utilizada. A cobertura também foi significativa, computando uma média de 76%, significando que a maior parte dos conceitos que eram previstos para anotarem os atributos da amostra foram efetivamente selecionados pela solução proposta.

No próximo capítulo, serão apresentadas as considerações finais desta dissertação.

# Capítulo 6

## Considerações Finais

Neste capítulo, será apresentada a conclusão desta pesquisa de dissertação. São discutidas as conclusões obtidas a partir de sua avaliação experimental, as contribuições oferecidas pelo desenvolvimento da pesquisa e as publicações obtidas ao longo do estudo. Por fim, são sugeridas as pesquisas futuras deste estudo.

As infraestruturas de dados espaciais têm melhorado bastante a interoperabilidade de dados geográficos providos por organizações distintas mediante a padronização de disponibilização destes dados. Para isso, as IDE normalmente fornecem um serviço de catálogo, que possibilita a recuperação dos seus conjuntos de dados geográficos. Embora tenham trazido melhorias consideráveis na recuperação de dados geográficos, os serviços de catálogo ainda apresentam limitações importantes que dificultam o acesso aos mesmos. Um exemplo destas limitações é o uso de palavras-chave para resolver consultas temáticas, uma vez que esse tipo de estratégia descarta dados geográficos relevantes que são relacionados semanticamente com a palavra-chave usada na consulta, mas que não a possuem em sua descrição.

Para resolver esse problema, o uso de **ontologias** pode proporcionar o enriquecimento semântico dos dados disponibilizados por IDE e os temas consultados pelos clientes. Isso permite melhorar de forma significativa a qualidade das consultas, pois os motores de busca podem compreender os relacionamentos semânticos entre o tema requisitado pelo cliente e o tema usado para descrição do serviço. Entretanto, é necessário que as descrições dos dados disponíveis sejam associadas aos conceitos definidos em **ontologias**, a partir de um processo chamado de **anotação semântica**.

Faz-se necessário anotar semanticamente tanto os dados que já estão publicados, como os

novos dados que vierem a ser disponibilizados. Executar esta tarefa manualmente seria muito custoso, pois o volume de dados disponibilizados normalmente é grande. Para superar este problema, esta pesquisa propôs uma abordagem para a **anotação semântica** automatizada dos dados geográficos disponibilizados por uma IDE. Foram utilizadas estratégias distintas para realizar a **anotação semântica** das camadas e dos atributos. Como as camadas dispõem de mais informações textuais que os atributos, foi possível fazer uso de técnicas de extração de similaridade para encontrar conceitos semanticamente equivalentes em uma **ontologia**. O mecanismo de anotação dos atributos, por sua vez, utilizou o nome do atributo para encontrar os conceitos relevantes mediante o casamento de *strings* com os nomes dos conceitos da **ontologia**. Após a execução da solução proposta, as anotações semânticas geradas ficam disponíveis para resolver consultas temáticas realizadas em motores de busca que fazem uso de **ontologias**.

A avaliação experimental utilizou as métricas de cobertura e precisão para mensurar o desempenho da solução proposta sob uma amostra de dados reais disponibilizados em serviços WFS. Os experimentos realizados mostraram que o processo de **anotação semântica** proposto nesta pesquisa possui um desempenho satisfatório. A grande maioria dos dados geográficos foi anotada e os conceitos que compuseram a anotação foram majoritariamente relevantes em relação as linhas de base construídas.

A avaliação experimental também proporcionou respostas positivas para as questões de pesquisas levantadas na introdução desta pesquisa.

## 6.1 Contribuições

O desenvolvimento desta pesquisa ofereceu contribuições para as comunidades de informação geográfica. Estas contribuições foram:

- Uma nova abordagem para a **anotação semântica** no nível de camadas;
- Uma abordagem para a **anotação semântica** no nível de atributos;

## 6.2 Resultados Obtidos

Ao longo do desenvolvimento desta pesquisa, foram produzidos dois artigos, sendo que um foi publicado e o outro foi submetido. Ambos foram produzidos para conferências internacionais:

- Geoprocessing 2014 [47] (publicado);
- ACM MEDES 2015 [48] (submetido).

## 6.3 Sugestões para Pesquisas Futuras

Para dar continuidade às pesquisas iniciadas nesta dissertação, são apresentadas as seguintes sugestões de pesquisas futuras:

- Aplicar a solução proposta a dados geográficos disponibilizados em outros idiomas, como o português e o espanhol. Com isso, será possível anotar dados geográficos disponibilizados por IDE cujos dados estão descritos nestes idiomas. A exemplo da Infraestrutura de Dados Espaciais Brasileira (INDE);
- Testar outras **ontologias** em que a *DBpedia* tenha uma quantidade maior de *links* para os seus conceitos. Este ponto poderá melhorar os resultados quantitativos das anotações via LOD. Para isso, pode ser desenvolvido um algoritmo para acessar as páginas da *DBpedia* em busca de *links* com outras bases de dados e realizar um *ranking* por quantidade de *links*;
- Investigar técnicas de inteligência artificial para se anotar uma camada a partir de padrões mapeados de acordo a existência ou não de determinados atributos. Este ponto possibilitará sugerir anotações semânticas para camadas que não foram anotadas;
- Desenvolver um motor de busca que utilize **ontologia** para a resolução de consultas. Fazendo uso da **anotação semântica** provida por esta pesquisa e aplicando relações semânticas ao tema consultado, o motor de busca seria uma prova de conceito para avaliar a qualidade dos seus resultados diante de um serviço de catálogo;

- Utilizar outros serviços na avaliação experimental para aumentar a confiabilidade da solução proposta. Aumentar a amostra dos dados geográficos utilizados na avaliação experimental e, conseqüentemente, inserir novos registros nas linhas de base;
- Comparar os resultados da **anotação semântica** de camadas proposta nesta pesquisa com os resultados da **anotação semântica** proposta por Andrade et al. [43]. Esta comparação possibilitará mensurar se a utilização da DBpedia como ferramenta de apoio para realização das anotações semânticas agrega algum valor ao processo; e
- Realizar um experimento para analisar qual é o melhor limiar a ser considerado no refinamento das anotações semânticas dos dados geográficos. Atualmente, esta pesquisa está utilizando o limiar de quatro conceitos com o mesmo nome para aplicar o refinamento.

# Referências Bibliográficas

- [1] LUPP, M. Open geospatial consortium. *Springer*, p. 815, 2008.
- [2] NEBERT, D. Developing spatial data infrastructures: The sdi cookbook. technical report, global spatial data infrastructure, 2004. Available at <http://www.gsdi.org/docs2004/Cookbook/cookbookV2.0.pdf> [Online; accessed 08-jun-2015].
- [3] COOPER, A. K.; MOELLERING, H.; HJELMAGER, J.; RAPANT, P.; DELGADO, T.; LAURENT, D.; COETZEE, S.; DANKO, D. M.; DÜREN, U.; IWANIAK, A.; BRODEUR, J.; ABAD, P.; HUET, M.; RAJABIFARD, A. A spatial data infrastructure model from the computational viewpoint. *International Journal of Geographical Information Science*, v. 27, n. 6, p. 1133–1151, 2013.
- [4] OLIVEIRA, I. L.; FILHO, J. L. A spatial data infrastructure review - sorting the actors and policies from enterprise viewpoint. *International Conference on Enterprise Information Systems*, p. 287–294, 2015.
- [5] GRUBER, T. R. A translation approach to portable ontology specifications. *Knowledge Acquisition*, v. 5, p. 199–220, 1993.
- [6] GUARINO, N. Formal ontology, conceptual analysis and knowledge representation. *International Journal of Human-Computer Studies*, Duluth, MN, USA, v. 43, n. 5-6, p. 625–640, Dec. 1995.
- [7] BIZER, C. The emerging web of linked data. *IEEE Intelligent Systems*, Piscataway, NJ, USA, v. 24, n. 5, p. 87–92, 2009.

- [8] BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web. *Scientific American*, v. 284, n. 5, p. 34–43, May 2001.
- [9] FENSEL, D.; VAN HARMELEN, F.; HORROCKS, I. *Oil and daml + oil: Ontology languages for the semantic web*. John Wiley & Sons, Ltd, 2003. p. 11–31.
- [10] LASSILA, O.; SWICK, R. Resource description framework (RDF). model and syntax specification. Technical report, W3C, 1999. W3C Recommendation. Available at <http://www.w3.org/TR/REC-rdf-syntax> [Online; accessed 08-jun-2015].
- [11] BOLEY, H.; TABET, S.; WAGNER, G. Design rationale for ruleml: A markup language for semantic web rules. *Semantic Web Working Symposium*, p. 381–401, aug 2001.
- [12] HORROCKS, I.; PATEL-SCHNEIDER, P. F.; BOLEY, H.; TABET, S.; GROSOFF, B.; DEAN, M. SWRL: A semantic web rule language combining OWL and RuleML. W3C Member Submission, may 2004. Available at <http://www.w3.org/Submission/SWRL/> [Online; accessed 08-jun-2015].
- [13] KIFER, M.; BOLEY, H. Rif overview. Working group note, W3C, jun 2010. Available at <http://www.w3.org/TR/2010/NOTE-rif-overview-20100622/> [Online; accessed 08-jun-2015].
- [14] TELLO, A. L. Ontologías en la web semántica. *Jornadas de Ingeniería Web*, v. 1, 2001.
- [15] W3C. RDF Vocabulary Description Language 1.0: RDF Schema, 2004. Available at <http://www.w3.org/TR/rdf-schema/> [Online; accessed 08-jun-2015].
- [16] FENSEL, D.; VAN HARMELEN, F.; HORROCKS, I.; MCGUINNESS, D. L.; PATEL-SCHNEIDER, P. F. Oil: an ontology infrastructure for the semantic web. *Intelligent Systems, IEEE [see also IEEE Intelligent Systems and Their Applications]*, v. 16, n. 2, p. 38–45, 2001.
- [17] HENDLER, J.; MCGUINNESS, D. L. The DARPA Agent Markup Language. *IEEE Intelligent Systems*, v. 15, n. 6, p. 67–73, 2000.

- [18] BAADER, F.; CALVANESE, D.; MCGUINNESS, D. L.; NARDI, D.; PATEL-SCHNEIDER, P. F. (Eds.). *The description logic handbook: Theory, implementation, and applications*. New York, NY, USA: Cambridge University Press, 2003.
- [19] MCGUINNESS, D. L.; VAN HARMELEN, F. Owl web ontology language overview. Technical Report REC-owl-features-20040210, W3C, 2004. Available at <http://www.w3.org/TR/owl-features/> [Online; accessed 08-jun-2015].
- [20] Open Geospatial Consortium. Ogc web map service interface. OGC 03-109r1, Version 1.3.0, jan 2004. Available at [http://portal.opengeospatial.org/files/?artifact\\_id=4756](http://portal.opengeospatial.org/files/?artifact_id=4756) [Online; accessed 08-jun-2015].
- [21] Open Geospatial Consortium. Web feature service implementation specification. OGC 04-094, Version 1.1.0, May 2005. Available at [http://portal.opengeospatial.org/files/?artifact\\_id=8339](http://portal.opengeospatial.org/files/?artifact_id=8339) [Online; accessed 08-jun-2015].
- [22] VOGES, U.; SENKLER, K. OpenGIS Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile (OGC 07-045). Technical report, OGC, July 2007. Available at [http://portal.opengeospatial.org/files/?artifact\\_id=21460](http://portal.opengeospatial.org/files/?artifact_id=21460) [Online; accessed 08-jun-2015].
- [23] COMMITTEE, F. G. D. The national spatial data infrastructure, 2005. Available at <http://www.fgdc.gov/library/factsheets/documents/nsdi.pdf> [Online; accessed 08-jun-2015].
- [24] GEOCONNECTIONS. The canadian geospatial data infrastructure, achieving the vision of the cgdi, 2005. Available at [http://ftp2.cits.mcan.gc.ca/pub/geott/ess\\_pubs/288/288846/cgdi\\_ip\\_04\\_e.pdf](http://ftp2.cits.mcan.gc.ca/pub/geott/ess_pubs/288/288846/cgdi_ip_04_e.pdf) [Online; accessed 08-jun-2015].
- [25] ON EARTH OBSERVATIONS, G. Global earth observation system of systems - geoss, 2005. Available at [https://www.earthobservations.org/documents/10-Year Plan Reference Document.pdf](https://www.earthobservations.org/documents/10-Year%20Plan%20Reference%20Document.pdf) [Online; accessed 08-jun-2015].

- [26] DE CARTOGRAFIA, C. N. Plano de ação para implantação da inde - infraestrutura nacional de dados espaciais, 2010. Available at <http://www.concar.gov.br/arquivo/PlanoDeAcaoINDE.pdf> [Online; accessed 08-jun-2015].
- [27] CYGANIAK, R.; JENTZSCH, A. The linking open data cloud diagram, Sept. 2011. Available at <http://lod-cloud.net/> [Online; accessed 08-jun-2015].
- [28] BERNERS-LEE, T. Linked data. 2006. Available at <http://www.w3.org/DesignIssues/LinkedData.html> [Online; accessed 08-jun-2015].
- [29] BERNERS-LEE, T. The Next Web, 2009. Available at [http://www.ted.com/talks/tim\\_berniers\\_lee\\_on\\_the\\_next\\_web.html](http://www.ted.com/talks/tim_berniers_lee_on_the_next_web.html) [Online; accessed 08-jun-2015].
- [30] HU, Y.; JANOWICZ, K.; PRASAD, S.; GAO, S. Enabling semantic search and knowledge discovery for arcgis online: A linked-data-driven approach. *Springer*, p. 107–124, 2015.
- [31] DE OLIVEIRA, W. M.; FILHO, J. L.; OLIVEIRA, A. P.; DE SOUZA, W. D.; VEGI, L. F. M.; OLIVEIRA, J. R. M. Building a thematic spatial data infrastructure and situation-aware for global events. *Journal of Software.*, v. 10, n. 3, p. 304–316, 2015.
- [32] FUGAZZA, C.; PEPE, M.; OGGIONI, A.; PAVESI, F.; CARRARA, P. A holistic, semantics-aware approach to spatial data infrastructures. *International Conference on Data Management Technologies and Applications*, p. 349–356, 2014.
- [33] GIMENEZ, P. J. A.; TANAKA, A. K.; BAIÃO, F. A. A geo-ontology to support the semantic integration of geoinformation from the national spatial data infrastructure. *GeoInfo*, p. 103–114, 2013.
- [34] BORZACCHIELLO, M. T.; CRAGLIA, M. Estimating benefits of spatial data infrastructures: A case study on e-cadastres. *Computers, Environment and Urban Systems*, v. 41, p. 276–288, 2013.

- [35] LUTZ, M.; SPRADO, J.; KLIEN, E.; SCHUBERT, C.; CHRIST, I. Overcoming semantic heterogeneity in spatial data infrastructures. *Computers & Geosciences*, v. 35, n. 4, p. 739–752, 2009.
- [36] BOWERS, S.; LIN, K.; LUDÄSCHER, B. On integrating scientific resources through semantic registration. *Scientific and Statistical Database Management, International Conference*, Los Alamitos, CA, USA, p. 349–352, 2004.
- [37] JANOWICZ, K.; SCHADE, S.; BRÖRING, A.; KEßLER, C.; MAUÉ, P.; STASCH, C. Semantic enablement for spatial data infrastructures. *Transactions in GIS*, v. 14, n. 2, p. 111–129, 2010.
- [38] LI, W.; YANG, C.; NEBERT, D.; RASKIN, R.; HOUSER, P.; WU, H.; LI, Z. Semantic-based web service discovery and chaining for building an arctic spatial data infrastructure. *Computers & Geosciences*, v. 37, n. 11, p. 1752–1762, 2011.
- [39] PRUD’HOMMEAUX, E.; SEABORNE, A. SPARQL Query Language for RDF. W3C Recommendation, 2008. Available at <http://www.w3.org/TR/rdf-sparql-query/> [Online; accessed 08-jun-2015].
- [40] MACARIO, C. G. N.; DE SOUSA, S. R.; MEDEIROS, C. B. Annotating geospatial data based on its semantics. *ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, p. 81–90, 2009.
- [41] BATCHELLER, J. K.; REITSMA, F. Implementing feature level semantics for spatial data discovery: Supporting the reuse of legacy data using open source components. *Computers, Environment and Urban Systems*, v. 34, n. 4, p. 333–344, 2010.
- [42] LUTZ, M.; KOLAS, D. Rule-based discovery in spatial data infrastructures. *Transactions on GIS*, v. 11, n. 3, p. 317–336, June 2007.
- [43] DE ANDRADE, F. G.; DE SOUZA BAPTISTA, C.; JR., C. A. D. Improving geographic information retrieval in spatial data infrastructures. *GeoInformatica*, v. 18, n. 4, p. 793–818, 2014.

- [44] SUCHANEK, F.; KASNECI, G.; WEIKUM, G. YAGO - a large ontology from Wikipedia and WordNet. *Elsevier Journal of Web Semantics*, v. 6, n. 3, p. 203–217, sep 2008.
- [45] SALTON, G.; LESK, M. E. Computer evaluation of indexing and text processing. *Journal of the ACM*, New York, NY, USA, v. 15, n. 1, p. 8–36, Jan. 1968.
- [46] SALTON, G. *The smart retrieval system: experiments in automatic document processing*. Prentice-Hall series in automatic computation. Prentice-Hall, 1971.
- [47] HENRIQUES, H.; ANDRADE, F.; BAPTISTA, C. An evaluation of semantically enriched spatial data infrastructure. *Geoprocessing 2014*, Barcelona, Spain, p. 115–121, Mar. 2014.
- [48] ANDRADE, F.; BAPTISTA, C.; HENRIQUES, H. Semantic annotation of geodata based on linked-open data. Submitted to ACM MEDES, 2015.

# Apêndice A

## Linha de Base para a Anotação no Nível de Camadas

A linha de base para a anotação no nível de camadas é composta de oitenta e quatro títulos de camadas e os respectivos *id* dos conceitos relevantes ligados semanticamente a eles. Para melhor ilustrar a linha de base, serão listados na Tabela A.1 os títulos dos camadas com os nomes e as URI dos conceitos relevantes.

Tabela A.1: Linha de Base para a anotação no nível de camadas.

Título	Nome do Conceito	URI do Conceito
Allotment Locations	allotment	<a href="http://dbpedia.org/class/yago/Allotment113289467">http://dbpedia.org/class/yago/Allotment113289467</a>
	allotment	<a href="http://dbpedia.org/class/yago/Allotment101083645">http://dbpedia.org/class/yago/Allotment101083645</a>
	location	<a href="http://dbpedia.org/class/yago/Location100027167">http://dbpedia.org/class/yago/Location100027167</a>
Community Council Boundaries	community	<a href="http://dbpedia.org/class/yago/Community108223802">http://dbpedia.org/class/yago/Community108223802</a>
	council	<a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a>
	boundary	<a href="http://dbpedia.org/class/yago/WritingSystemsWithoutWordBoundaries">http://dbpedia.org/class/yago/WritingSystemsWithoutWordBoundaries</a>
Conservation Areas	conservation	<a href="http://dbpedia.org/class/yago/Conservation100819274">http://dbpedia.org/class/yago/Conservation100819274</a>
	conservation	<a href="http://dbpedia.org/class/yago/Conservation106104073">http://dbpedia.org/class/yago/Conservation106104073</a>
Core Path Network	midpoint	<a href="http://dbpedia.org/class/yago/Center108521816">http://dbpedia.org/class/yago/Center108521816</a>
	way	<a href="http://dbpedia.org/class/yago/Way100415676">http://dbpedia.org/class/yago/Way100415676</a>
	web	<a href="http://dbpedia.org/class/yago/Network108434259">http://dbpedia.org/class/yago/Network108434259</a>
Defined Policy Areas	policy	<a href="http://dbpedia.org/class/yago/Policy105901508">http://dbpedia.org/class/yago/Policy105901508</a>
	federal agency	<a href="http://dbpedia.org/class/yago/Agency108337324">http://dbpedia.org/class/yago/Agency108337324</a>
	intelligence	<a href="http://dbpedia.org/class/yago/Intelligence108339454">http://dbpedia.org/class/yago/Intelligence108339454</a>
Development Proposals Area	development	<a href="http://dbpedia.org/class/yago/Development100250259">http://dbpedia.org/class/yago/Development100250259</a>
	proposal	<a href="http://dbpedia.org/class/yago/Proposal107162194">http://dbpedia.org/class/yago/Proposal107162194</a>
Development Proposals Line	development	<a href="http://dbpedia.org/class/yago/Development100250259">http://dbpedia.org/class/yago/Development100250259</a>
	proposal	<a href="http://dbpedia.org/class/yago/Proposal107162194">http://dbpedia.org/class/yago/Proposal107162194</a>
Employment Land Audit	land	<a href="http://dbpedia.org/class/yago/Land113250048">http://dbpedia.org/class/yago/Land113250048</a>
	audit	<a href="http://dbpedia.org/class/yago/Audit100141176">http://dbpedia.org/class/yago/Audit100141176</a>
	organ	<a href="http://dbpedia.org/class/yago/Organ105297523">http://dbpedia.org/class/yago/Organ105297523</a>
	organ	<a href="http://dbpedia.org/class/yago/Organ103854065">http://dbpedia.org/class/yago/Organ103854065</a>
Fife Common Good Land	organization	<a href="http://dbpedia.org/class/yago/Organization108008335">http://dbpedia.org/class/yago/Organization108008335</a>
	property	<a href="http://dbpedia.org/class/yago/Property113244109">http://dbpedia.org/class/yago/Property113244109</a>
	fife	<a href="http://dbpedia.org/class/yago/Fife103334492">http://dbpedia.org/class/yago/Fife103334492</a>

Continua na próxima página

<b>Título</b>	<b>Nome do Conceito</b>	<b>URI do Conceito</b>
Fife Common Good Land	good	<a href="http://dbpedia.org/class/yago/Good105159725">http://dbpedia.org/class/yago/Good105159725</a>
	land	<a href="http://dbpedia.org/class/yago/Land113250048">http://dbpedia.org/class/yago/Land113250048</a>
	work	<a href="http://dbpedia.org/class/yago/Oeuvre103841417">http://dbpedia.org/class/yago/Oeuvre103841417</a>
Fife Primary School Catchments	fife	<a href="http://dbpedia.org/class/yago/Fife103334492">http://dbpedia.org/class/yago/Fife103334492</a>
	primary	<a href="http://dbpedia.org/class/yago/Primary109398769">http://dbpedia.org/class/yago/Primary109398769</a>
	school	<a href="http://dbpedia.org/class/yago/School108276720">http://dbpedia.org/class/yago/School108276720</a>
	catchment	<a href="http://dbpedia.org/class/yago/Catchment102983357">http://dbpedia.org/class/yago/Catchment102983357</a>
	primary election area	<a href="http://dbpedia.org/class/yago/Primary100182571">http://dbpedia.org/class/yago/Primary100182571</a> <a href="http://dbpedia.org/class/yago/Area108497294">http://dbpedia.org/class/yago/Area108497294</a>
Fife RC Primary School Catchments	fife	<a href="http://dbpedia.org/class/yago/Fife103334492">http://dbpedia.org/class/yago/Fife103334492</a>
	primary	<a href="http://dbpedia.org/class/yago/Primary109398769">http://dbpedia.org/class/yago/Primary109398769</a>
	school	<a href="http://dbpedia.org/class/yago/School108276720">http://dbpedia.org/class/yago/School108276720</a>
	catchment	<a href="http://dbpedia.org/class/yago/Catchment102983357">http://dbpedia.org/class/yago/Catchment102983357</a>
	primary election	<a href="http://dbpedia.org/class/yago/Primary100182571">http://dbpedia.org/class/yago/Primary100182571</a>
Fife RC Secondary School Catchments	fife	<a href="http://dbpedia.org/class/yago/Fife103334492">http://dbpedia.org/class/yago/Fife103334492</a>
	secondary	<a href="http://dbpedia.org/class/yago/Secondary108429899">http://dbpedia.org/class/yago/Secondary108429899</a>
	school	<a href="http://dbpedia.org/class/yago/School108276720">http://dbpedia.org/class/yago/School108276720</a>
	catchment	<a href="http://dbpedia.org/class/yago/Catchment102983357">http://dbpedia.org/class/yago/Catchment102983357</a>
	teacher	<a href="http://dbpedia.org/class/yago/Teacher110694258">http://dbpedia.org/class/yago/Teacher110694258</a>
Fife Secondary School Catchments	fife	<a href="http://dbpedia.org/class/yago/Fife103334492">http://dbpedia.org/class/yago/Fife103334492</a>
	secondary	<a href="http://dbpedia.org/class/yago/Secondary108429899">http://dbpedia.org/class/yago/Secondary108429899</a>
	school	<a href="http://dbpedia.org/class/yago/School108276720">http://dbpedia.org/class/yago/School108276720</a>
	catchment	<a href="http://dbpedia.org/class/yago/Catchment102983357">http://dbpedia.org/class/yago/Catchment102983357</a>
	teacher	<a href="http://dbpedia.org/class/yago/Teacher110694258">http://dbpedia.org/class/yago/Teacher110694258</a>
	area	<a href="http://dbpedia.org/class/yago/Area108497294">http://dbpedia.org/class/yago/Area108497294</a>
Housing Land Audit 2014	housing	<a href="http://dbpedia.org/class/yago/Housing103546766">http://dbpedia.org/class/yago/Housing103546766</a>
	land	<a href="http://dbpedia.org/class/yago/Land113250048">http://dbpedia.org/class/yago/Land113250048</a>
	audit	<a href="http://dbpedia.org/class/yago/Audit100141176">http://dbpedia.org/class/yago/Audit100141176</a>
	entity	<a href="http://dbpedia.org/class/yago/Entity100001740">http://dbpedia.org/class/yago/Entity100001740</a>
Listed Buildings	edifice	<a href="http://dbpedia.org/class/yago/Building102913152">http://dbpedia.org/class/yago/Building102913152</a>
Local Landscape Areas	local	<a href="http://dbpedia.org/class/yago/Local103680942">http://dbpedia.org/class/yago/Local103680942</a>
	landscape	<a href="http://dbpedia.org/class/yago/Landscape108646902">http://dbpedia.org/class/yago/Landscape108646902</a>
Natural Heritage Sites	natural	<a href="http://dbpedia.org/class/yago/Natural110346392">http://dbpedia.org/class/yago/Natural110346392</a>
	natural	<a href="http://dbpedia.org/class/yago/Natural106867218">http://dbpedia.org/class/yago/Natural106867218</a>
	natural	<a href="http://dbpedia.org/class/yago/Natural101245950">http://dbpedia.org/class/yago/Natural101245950</a>
	heritage	<a href="http://dbpedia.org/class/yago/Heritage105669797">http://dbpedia.org/class/yago/Heritage105669797</a>
	land site	<a href="http://dbpedia.org/class/yago/Site108651247">http://dbpedia.org/class/yago/Site108651247</a>
	group	<a href="http://dbpedia.org/class/yago/Group100031264">http://dbpedia.org/class/yago/Group100031264</a>
	organization	<a href="http://dbpedia.org/class/yago/Organization108008335">http://dbpedia.org/class/yago/Organization108008335</a>
	social group	<a href="http://dbpedia.org/class/yago/SocialGroup107950920">http://dbpedia.org/class/yago/SocialGroup107950920</a>
Polling Districts	poll	<a href="http://dbpedia.org/class/yago/Poll105539012">http://dbpedia.org/class/yago/Poll105539012</a>
	poll	<a href="http://dbpedia.org/class/yago/Poll100653811">http://dbpedia.org/class/yago/Poll100653811</a>
	territory	<a href="http://dbpedia.org/class/yago/District108552138">http://dbpedia.org/class/yago/District108552138</a>
	election	<a href="http://dbpedia.org/class/yago/Elections">http://dbpedia.org/class/yago/Elections</a>
	election	<a href="http://dbpedia.org/class/yago/Election100181781">http://dbpedia.org/class/yago/Election100181781</a>
	vote	<a href="http://dbpedia.org/class/yago/Vote100182213">http://dbpedia.org/class/yago/Vote100182213</a>
	voting system	<a href="http://dbpedia.org/class/yago/VotingSystem105904135">http://dbpedia.org/class/yago/VotingSystem105904135</a>
	legal system	<a href="http://dbpedia.org/class/yago/LegalSystem105903229">http://dbpedia.org/class/yago/LegalSystem105903229</a>
Polling Places	poll	<a href="http://dbpedia.org/class/yago/Poll105539012">http://dbpedia.org/class/yago/Poll105539012</a>
	poll	<a href="http://dbpedia.org/class/yago/Poll100653811">http://dbpedia.org/class/yago/Poll100653811</a>
	land site	<a href="http://dbpedia.org/class/yago/Site108651247">http://dbpedia.org/class/yago/Site108651247</a>
	election	<a href="http://dbpedia.org/class/yago/Elections">http://dbpedia.org/class/yago/Elections</a>
	election	<a href="http://dbpedia.org/class/yago/Election100181781">http://dbpedia.org/class/yago/Election100181781</a> <a href="http://dbpedia.org/class/yago/Vote100182213">http://dbpedia.org/class/yago/Vote100182213</a>
Tree Preservation Order Areas	tree	<a href="http://dbpedia.org/class/yago/Tree113104059">http://dbpedia.org/class/yago/Tree113104059</a>
	preservation	<a href="http://dbpedia.org/class/yago/Preservation114578940">http://dbpedia.org/class/yago/Preservation114578940</a>
	preservation	<a href="http://dbpedia.org/class/yago/Preservation113540610">http://dbpedia.org/class/yago/Preservation113540610</a>
	order	<a href="http://dbpedia.org/class/yago/Order107168623">http://dbpedia.org/class/yago/Order107168623</a>
Tree Preservation Order Points	tree	<a href="http://dbpedia.org/class/yago/Tree113104059">http://dbpedia.org/class/yago/Tree113104059</a>
	preservation	<a href="http://dbpedia.org/class/yago/Preservation113540610">http://dbpedia.org/class/yago/Preservation113540610</a>
	order	<a href="http://dbpedia.org/class/yago/Order107168623">http://dbpedia.org/class/yago/Order107168623</a>

*Continua na próxima página*

<b>Título</b>	<b>Nome do Conceito</b>	<b>URI do Conceito</b>
Tree Preservation Order Points	preservation	<a href="http://dbpedia.org/class/yago/Preservation114578940">http://dbpedia.org/class/yago/Preservation114578940</a>
Vacant And Derelict Land	derelict	<a href="http://dbpedia.org/class/yago/Derelict110006081">http://dbpedia.org/class/yago/Derelict110006081</a>
	derelict	<a href="http://dbpedia.org/class/yago/AbandonedShip102666501">http://dbpedia.org/class/yago/AbandonedShip102666501</a>
	land	<a href="http://dbpedia.org/class/yago/Land113250048">http://dbpedia.org/class/yago/Land113250048</a>
Wind Turbines	rotary engine	<a href="http://dbpedia.org/class/yago/RotaryEngine104110654">http://dbpedia.org/class/yago/RotaryEngine104110654</a>
	turbine	<a href="http://dbpedia.org/class/yago/WindTurbines">http://dbpedia.org/class/yago/WindTurbines</a>
	wind turbine	<a href="http://dbpedia.org/class/yago/WindTurbine104591517">http://dbpedia.org/class/yago/WindTurbine104591517</a>
	device	<a href="http://dbpedia.org/class/yago/Device103183080">http://dbpedia.org/class/yago/Device103183080</a>
	heat engine	<a href="http://dbpedia.org/class/yago/HeatEngine103507963">http://dbpedia.org/class/yago/HeatEngine103507963</a>
	instrumentality	<a href="http://dbpedia.org/class/yago/Instrumentality103575240">http://dbpedia.org/class/yago/Instrumentality103575240</a>
	machine	<a href="http://dbpedia.org/class/yago/Machine103699975">http://dbpedia.org/class/yago/Machine103699975</a>
	engine	<a href="http://dbpedia.org/class/yago/Engine103287733">http://dbpedia.org/class/yago/Engine103287733</a>
	motor	<a href="http://dbpedia.org/class/yago/Motor103789946">http://dbpedia.org/class/yago/Motor103789946</a>
Windfarm Broad Area Of Search	turbine	<a href="http://dbpedia.org/class/yago/Turbine104498523">http://dbpedia.org/class/yago/Turbine104498523</a>
	broad	<a href="http://dbpedia.org/class/yago/NorfolkBroads">http://dbpedia.org/class/yago/NorfolkBroads</a>
	broad	<a href="http://dbpedia.org/class/yago/Broad109875663">http://dbpedia.org/class/yago/Broad109875663</a>
	search	<a href="http://dbpedia.org/class/yago/Search100637354">http://dbpedia.org/class/yago/Search100637354</a>
AIR EMEP Monitoring Sites	search	<a href="http://dbpedia.org/class/yago/Search105770058">http://dbpedia.org/class/yago/Search105770058</a>
	search	<a href="http://dbpedia.org/class/yago/Search100644366">http://dbpedia.org/class/yago/Search100644366</a>
	air	<a href="http://dbpedia.org/class/yago/Air114841267">http://dbpedia.org/class/yago/Air114841267</a>
	air	<a href="http://dbpedia.org/class/yago/Air108653314">http://dbpedia.org/class/yago/Air108653314</a>
	air	<a href="http://dbpedia.org/class/yago/Air104727214">http://dbpedia.org/class/yago/Air104727214</a>
	air	<a href="http://dbpedia.org/class/yago/Air114842703">http://dbpedia.org/class/yago/Air114842703</a>
	monitoring	<a href="http://dbpedia.org/class/yago/Monitoring100880046">http://dbpedia.org/class/yago/Monitoring100880046</a>
Air Monitoring Sites	land site	<a href="http://dbpedia.org/class/yago/Site108651247">http://dbpedia.org/class/yago/Site108651247</a>
	region	<a href="http://dbpedia.org/class/yago/Region108630039">http://dbpedia.org/class/yago/Region108630039</a>
	monitor	<a href="http://dbpedia.org/class/yago/Proctor110478960">http://dbpedia.org/class/yago/Proctor110478960</a>
	air	<a href="http://dbpedia.org/class/yago/Air114841267">http://dbpedia.org/class/yago/Air114841267</a>
	air	<a href="http://dbpedia.org/class/yago/Air108653314">http://dbpedia.org/class/yago/Air108653314</a>
	air	<a href="http://dbpedia.org/class/yago/Air104727214">http://dbpedia.org/class/yago/Air104727214</a>
	air	<a href="http://dbpedia.org/class/yago/Air114842703">http://dbpedia.org/class/yago/Air114842703</a>
Air Zones	monitoring	<a href="http://dbpedia.org/class/yago/Monitoring100880046">http://dbpedia.org/class/yago/Monitoring100880046</a>
	land site	<a href="http://dbpedia.org/class/yago/Site108651247">http://dbpedia.org/class/yago/Site108651247</a>
	region	<a href="http://dbpedia.org/class/yago/Region108630039">http://dbpedia.org/class/yago/Region108630039</a>
	monitor	<a href="http://dbpedia.org/class/yago/Proctor110478960">http://dbpedia.org/class/yago/Proctor110478960</a>
	air	<a href="http://dbpedia.org/class/yago/Air114841267">http://dbpedia.org/class/yago/Air114841267</a>
Air Zones	air	<a href="http://dbpedia.org/class/yago/Air108653314">http://dbpedia.org/class/yago/Air108653314</a>
	air	<a href="http://dbpedia.org/class/yago/Air104727214">http://dbpedia.org/class/yago/Air104727214</a>
	air	<a href="http://dbpedia.org/class/yago/Air114842703">http://dbpedia.org/class/yago/Air114842703</a>
	zone	<a href="http://dbpedia.org/class/yago/Zone108688247">http://dbpedia.org/class/yago/Zone108688247</a>
	region	<a href="http://dbpedia.org/class/yago/Region108630039">http://dbpedia.org/class/yago/Region108630039</a>
Bathing Water Quality	bath	<a href="http://dbpedia.org/class/yago/Bath102806530">http://dbpedia.org/class/yago/Bath102806530</a>
	water	<a href="http://dbpedia.org/class/yago/Water114845743">http://dbpedia.org/class/yago/Water114845743</a>
	quality	<a href="http://dbpedia.org/class/yago/Quality104723816">http://dbpedia.org/class/yago/Quality104723816</a>
	quality	<a href="http://dbpedia.org/class/yago/Quality104728068">http://dbpedia.org/class/yago/Quality104728068</a>
	quality	<a href="http://dbpedia.org/class/yago/Quality113948026">http://dbpedia.org/class/yago/Quality113948026</a>
Coal Restricted Areas	coal	<a href="http://dbpedia.org/class/yago/Coal114814616">http://dbpedia.org/class/yago/Coal114814616</a>
	coal	<a href="http://dbpedia.org/class/yago/Ember109273130">http://dbpedia.org/class/yago/Ember109273130</a>
Coastal waterbodies	body of water	<a href="http://dbpedia.org/class/yago/Middle-earthBodiesOfWater">http://dbpedia.org/class/yago/Middle-earthBodiesOfWater</a>
	body of water	<a href="http://dbpedia.org/class/yago/BodiesOfWater">http://dbpedia.org/class/yago/BodiesOfWater</a>
	body of water	<a href="http://dbpedia.org/class/yago/BodyOfWater109225146">http://dbpedia.org/class/yago/BodyOfWater109225146</a>
EPA Offices	EPA	<a href="http://dbpedia.org/class/yago/EnvironmentalProtectionAgency108122960">http://dbpedia.org/class/yago/EnvironmentalProtectionAgency108122960</a>
	office	<a href="http://dbpedia.org/class/yago/Office103841666">http://dbpedia.org/class/yago/Office103841666</a>
	administration	<a href="http://dbpedia.org/class/yago/Administration101135952">http://dbpedia.org/class/yago/Administration101135952</a>
Geological Linear Features	characteristic	<a href="http://dbpedia.org/class/yago/Feature105849789">http://dbpedia.org/class/yago/Feature105849789</a>
Ground waterbodies	solid ground	<a href="http://dbpedia.org/class/yago/Land109334396">http://dbpedia.org/class/yago/Land109334396</a>
	body of water	<a href="http://dbpedia.org/class/yago/Middle-earthBodiesOfWater">http://dbpedia.org/class/yago/Middle-earthBodiesOfWater</a>
Ground waterbodies	body of water	<a href="http://dbpedia.org/class/yago/BodyOfWater109225146">http://dbpedia.org/class/yago/BodyOfWater109225146</a>
	body of water	<a href="http://dbpedia.org/class/yago/BodiesOfWater">http://dbpedia.org/class/yago/BodiesOfWater</a>
Historical Groundwater Quality	quality	<a href="http://dbpedia.org/class/yago/Quality104723816">http://dbpedia.org/class/yago/Quality104723816</a>
	quality	<a href="http://dbpedia.org/class/yago/Quality104728068">http://dbpedia.org/class/yago/Quality104728068</a>

*Continua na próxima página*

Titulo	Nome do Conceito	URI do Conceito
	quality	<a href="http://dbpedia.org/class/yago/Quality113948026">http://dbpedia.org/class/yago/Quality113948026</a>
Large Towns	large	<a href="http://dbpedia.org/class/yago/Large105096191">http://dbpedia.org/class/yago/Large105096191</a>
	town	<a href="http://dbpedia.org/class/yago/Town108665504">http://dbpedia.org/class/yago/Town108665504</a>
Licenced IPPC Facilities	facility	<a href="http://dbpedia.org/class/yago/Facility103315023">http://dbpedia.org/class/yago/Facility103315023</a>
Licenced Waste Facilities	waste	<a href="http://dbpedia.org/class/yago/Waste114856263">http://dbpedia.org/class/yago/Waste114856263</a>
	waste	<a href="http://dbpedia.org/class/yago/Barren108504594">http://dbpedia.org/class/yago/Barren108504594</a>
	waste	<a href="http://dbpedia.org/class/yago/Waste101252124">http://dbpedia.org/class/yago/Waste101252124</a>
	facility	<a href="http://dbpedia.org/class/yago/Facility103315023">http://dbpedia.org/class/yago/Facility103315023</a>
Linear Mine Features	mine	<a href="http://dbpedia.org/class/yago/Mine103768346">http://dbpedia.org/class/yago/Mine103768346</a>
	characteristic	<a href="http://dbpedia.org/class/yago/Feature105849789">http://dbpedia.org/class/yago/Feature105849789</a>
Linear Waste Features	waste	<a href="http://dbpedia.org/class/yago/Waste114856263">http://dbpedia.org/class/yago/Waste114856263</a>
	waste	<a href="http://dbpedia.org/class/yago/Barren108504594">http://dbpedia.org/class/yago/Barren108504594</a>
	waste	<a href="http://dbpedia.org/class/yago/Waste101252124">http://dbpedia.org/class/yago/Waste101252124</a>
	characteristic	<a href="http://dbpedia.org/class/yago/Feature105849789">http://dbpedia.org/class/yago/Feature105849789</a>
Medium Towns	town	<a href="http://dbpedia.org/class/yago/Town108665504">http://dbpedia.org/class/yago/Town108665504</a>
	medium	<a href="http://dbpedia.org/class/yago/Medium106254669">http://dbpedia.org/class/yago/Medium106254669</a>
Mine District	mine	<a href="http://dbpedia.org/class/yago/Mine103768346">http://dbpedia.org/class/yago/Mine103768346</a>
	territory	<a href="http://dbpedia.org/class/yago/District108552138">http://dbpedia.org/class/yago/District108552138</a>
Mine Site Boundaries	mine	<a href="http://dbpedia.org/class/yago/Mine103768346">http://dbpedia.org/class/yago/Mine103768346</a>
	land site	<a href="http://dbpedia.org/class/yago/Site108651247">http://dbpedia.org/class/yago/Site108651247</a>
	boundary	<a href="http://dbpedia.org/class/yago/WritingSystemsWithoutWordBoundaries">http://dbpedia.org/class/yago/WritingSystemsWithoutWordBoundaries</a>
Mine Site Location	mine	<a href="http://dbpedia.org/class/yago/Mine103768346">http://dbpedia.org/class/yago/Mine103768346</a>
	land site	<a href="http://dbpedia.org/class/yago/Site108651247">http://dbpedia.org/class/yago/Site108651247</a>
	location	<a href="http://dbpedia.org/class/yago/Location100027167">http://dbpedia.org/class/yago/Location100027167</a>
Other Waste Areas	waste	<a href="http://dbpedia.org/class/yago/Waste114856263">http://dbpedia.org/class/yago/Waste114856263</a>
	waste	<a href="http://dbpedia.org/class/yago/Barren108504594">http://dbpedia.org/class/yago/Barren108504594</a>
	waste	<a href="http://dbpedia.org/class/yago/Waste101252124">http://dbpedia.org/class/yago/Waste101252124</a>
Point Mine Features	mine	<a href="http://dbpedia.org/class/yago/Mine103768346">http://dbpedia.org/class/yago/Mine103768346</a>
	characteristic	<a href="http://dbpedia.org/class/yago/Feature105849789">http://dbpedia.org/class/yago/Feature105849789</a>
Point Waste Features	waste	<a href="http://dbpedia.org/class/yago/Waste114856263">http://dbpedia.org/class/yago/Waste114856263</a>
	waste	<a href="http://dbpedia.org/class/yago/Barren108504594">http://dbpedia.org/class/yago/Barren108504594</a>
	waste	<a href="http://dbpedia.org/class/yago/Waste101252124">http://dbpedia.org/class/yago/Waste101252124</a>
	characteristic	<a href="http://dbpedia.org/class/yago/Feature105849789">http://dbpedia.org/class/yago/Feature105849789</a>
Radiation Monitoring	radiation	<a href="http://dbpedia.org/class/yago/Radiation111499284">http://dbpedia.org/class/yago/Radiation111499284</a>
	radiation	<a href="http://dbpedia.org/class/yago/Radiation101252918">http://dbpedia.org/class/yago/Radiation101252918</a>
	radiation	<a href="http://dbpedia.org/class/yago/Radiation107439883">http://dbpedia.org/class/yago/Radiation107439883</a>
	radiation	<a href="http://dbpedia.org/class/yago/Radiation105503555">http://dbpedia.org/class/yago/Radiation105503555</a>
	monitor	<a href="http://dbpedia.org/class/yago/Proctor110478960">http://dbpedia.org/class/yago/Proctor110478960</a>
Radon Map	map	<a href="http://dbpedia.org/class/yago/Map103720163">http://dbpedia.org/class/yago/Map103720163</a>
River Basin District	river	<a href="http://dbpedia.org/class/yago/River109411430">http://dbpedia.org/class/yago/River109411430</a>
	basin	<a href="http://dbpedia.org/class/yago/Basin102801525">http://dbpedia.org/class/yago/Basin102801525</a>
	territory	<a href="http://dbpedia.org/class/yago/District108552138">http://dbpedia.org/class/yago/District108552138</a>
River Catchments	river	<a href="http://dbpedia.org/class/yago/River109411430">http://dbpedia.org/class/yago/River109411430</a>
	catchment	<a href="http://dbpedia.org/class/yago/Catchment102983357">http://dbpedia.org/class/yago/Catchment102983357</a>
River Sub Basins	river	<a href="http://dbpedia.org/class/yago/River109411430">http://dbpedia.org/class/yago/River109411430</a>
	basin	<a href="http://dbpedia.org/class/yago/Basin102801525">http://dbpedia.org/class/yago/Basin102801525</a>
SAC CO Habitats - Area	sac	<a href="http://dbpedia.org/class/yago/Theca113092722">http://dbpedia.org/class/yago/Theca113092722</a>
	sac	<a href="http://dbpedia.org/class/yago/Sac105515670">http://dbpedia.org/class/yago/Sac105515670</a>
	habitat	<a href="http://dbpedia.org/class/yago/Habitat108580583">http://dbpedia.org/class/yago/Habitat108580583</a>
	Co	<a href="http://dbpedia.org/class/yago/Cobalt114635290">http://dbpedia.org/class/yago/Cobalt114635290</a>
	CO	<a href="http://dbpedia.org/class/yago/CarbonMonoxide114797641">http://dbpedia.org/class/yago/CarbonMonoxide114797641</a>
SAC CO Species - Area	CO	<a href="http://dbpedia.org/class/yago/ConscientiousObjector109957013">http://dbpedia.org/class/yago/ConscientiousObjector109957013</a>
	sac	<a href="http://dbpedia.org/class/yago/Theca113092722">http://dbpedia.org/class/yago/Theca113092722</a>
	sac	<a href="http://dbpedia.org/class/yago/Sac105515670">http://dbpedia.org/class/yago/Sac105515670</a>
	species	<a href="http://dbpedia.org/class/yago/Species108110373">http://dbpedia.org/class/yago/Species108110373</a>
SAC CO Species - Area	species	<a href="http://dbpedia.org/class/yago/Species105844881">http://dbpedia.org/class/yago/Species105844881</a>
	Co	<a href="http://dbpedia.org/class/yago/Cobalt114635290">http://dbpedia.org/class/yago/Cobalt114635290</a>
	CO	<a href="http://dbpedia.org/class/yago/CarbonMonoxide114797641">http://dbpedia.org/class/yago/CarbonMonoxide114797641</a>
SPA CO Habitats - Area	CO	<a href="http://dbpedia.org/class/yago/ConscientiousObjector109957013">http://dbpedia.org/class/yago/ConscientiousObjector109957013</a>
	spa	<a href="http://dbpedia.org/class/yago/WateringPlace108678615">http://dbpedia.org/class/yago/WateringPlace108678615</a>
SAC CO Habitats - Area	habitat	<a href="http://dbpedia.org/class/yago/Habitat108580583">http://dbpedia.org/class/yago/Habitat108580583</a>

Continua na próxima página

Titulo	Nome do Conceito	URI do Conceito
SPA CO Habitats - Area	Co	<a href="http://dbpedia.org/class/yago/Cobalt114635290">http://dbpedia.org/class/yago/Cobalt114635290</a>
	CO	<a href="http://dbpedia.org/class/yago/CarbonMonoxide114797641">http://dbpedia.org/class/yago/CarbonMonoxide114797641</a>
	CO	<a href="http://dbpedia.org/class/yago/ConscientiousObjector109957013">http://dbpedia.org/class/yago/ConscientiousObjector109957013</a>
Site Features	land site characteristic	<a href="http://dbpedia.org/class/yago/Site108651247">http://dbpedia.org/class/yago/Site108651247</a> <a href="http://dbpedia.org/class/yago/Feature105849789">http://dbpedia.org/class/yago/Feature105849789</a>
Small Towns	small	<a href="http://dbpedia.org/class/yago/Small105559023">http://dbpedia.org/class/yago/Small105559023</a>
	small	<a href="http://dbpedia.org/class/yago/Small105097081">http://dbpedia.org/class/yago/Small105097081</a>
	town	<a href="http://dbpedia.org/class/yago/Town108665504">http://dbpedia.org/class/yago/Town108665504</a>
Solid Waste Analyses	solid	<a href="http://dbpedia.org/class/yago/Solid115046900">http://dbpedia.org/class/yago/Solid115046900</a>
	waste	<a href="http://dbpedia.org/class/yago/Waste114856263">http://dbpedia.org/class/yago/Waste114856263</a>
	waste	<a href="http://dbpedia.org/class/yago/Barren108504594">http://dbpedia.org/class/yago/Barren108504594</a>
	waste	<a href="http://dbpedia.org/class/yago/Waste101252124">http://dbpedia.org/class/yago/Waste101252124</a>
	utilisation	<a href="http://dbpedia.org/class/yago/Use100947128">http://dbpedia.org/class/yago/Use100947128</a>
	activity	<a href="http://dbpedia.org/class/yago/Activity100407535">http://dbpedia.org/class/yago/Activity100407535</a>
	event	<a href="http://dbpedia.org/class/yago/Event100029378">http://dbpedia.org/class/yago/Event100029378</a>
	technology	<a href="http://dbpedia.org/class/yago/PollutionControlTechnologies">http://dbpedia.org/class/yago/PollutionControlTechnologies</a>
	application	<a href="http://dbpedia.org/class/yago/Application100949134">http://dbpedia.org/class/yago/Application100949134</a>
	engineering	<a href="http://dbpedia.org/class/yago/Technology100949619">http://dbpedia.org/class/yago/Technology100949619</a>
Solid Waste Heaps	solid	<a href="http://dbpedia.org/class/yago/Solid115046900">http://dbpedia.org/class/yago/Solid115046900</a>
	waste	<a href="http://dbpedia.org/class/yago/Waste114856263">http://dbpedia.org/class/yago/Waste114856263</a>
	waste	<a href="http://dbpedia.org/class/yago/Barren108504594">http://dbpedia.org/class/yago/Barren108504594</a>
	waste	<a href="http://dbpedia.org/class/yago/Waste101252124">http://dbpedia.org/class/yago/Waste101252124</a>
Stream Sediment Analyses	stream	<a href="http://dbpedia.org/class/yago/Stream109448361">http://dbpedia.org/class/yago/Stream109448361</a>
	sediment	<a href="http://dbpedia.org/class/yago/Sediments">http://dbpedia.org/class/yago/Sediments</a>
	scheme	<a href="http://dbpedia.org/class/yago/System108435388">http://dbpedia.org/class/yago/System108435388</a>
	computer program	<a href="http://dbpedia.org/class/yago/Program106568978">http://dbpedia.org/class/yago/Program106568978</a>
	software system	<a href="http://dbpedia.org/class/yago/Software106566077">http://dbpedia.org/class/yago/Software106566077</a>
Subsoils	coding system	<a href="http://dbpedia.org/class/yago/CodingSystem106353757">http://dbpedia.org/class/yago/CodingSystem106353757</a>
	subsoil	<a href="http://dbpedia.org/class/yago/Subsoil114695838">http://dbpedia.org/class/yago/Subsoil114695838</a>
Transitional waterbodies	body of water	<a href="http://dbpedia.org/class/yago/Middle-earthBodiesOfWater">http://dbpedia.org/class/yago/Middle-earthBodiesOfWater</a>
	body of water	<a href="http://dbpedia.org/class/yago/BodiesOfWater">http://dbpedia.org/class/yago/BodiesOfWater</a>
	body of water	<a href="http://dbpedia.org/class/yago/BodyOfWater109225146">http://dbpedia.org/class/yago/BodyOfWater109225146</a>
	transit	<a href="http://dbpedia.org/class/yago/AstronomicalTransits">http://dbpedia.org/class/yago/AstronomicalTransits</a>
Water Analyses (Summer)	water	<a href="http://dbpedia.org/class/yago/Water114845743">http://dbpedia.org/class/yago/Water114845743</a>
	summer	<a href="http://dbpedia.org/class/yago/Summer115145012">http://dbpedia.org/class/yago/Summer115145012</a>
	method	<a href="http://dbpedia.org/class/yago/Method105660268">http://dbpedia.org/class/yago/Method105660268</a>
	event	<a href="http://dbpedia.org/class/yago/Event100029378">http://dbpedia.org/class/yago/Event100029378</a>
	activity	<a href="http://dbpedia.org/class/yago/Activity100407535">http://dbpedia.org/class/yago/Activity100407535</a>
	analysis	<a href="http://dbpedia.org/class/yago/Analysis106014043">http://dbpedia.org/class/yago/Analysis106014043</a>
	basin	<a href="http://dbpedia.org/class/yago/Basin102801525">http://dbpedia.org/class/yago/Basin102801525</a>
	lake	<a href="http://dbpedia.org/class/yago/Lake109328904">http://dbpedia.org/class/yago/Lake109328904</a>
	river	<a href="http://dbpedia.org/class/yago/River109411430">http://dbpedia.org/class/yago/River109411430</a>
	resource	<a href="http://dbpedia.org/class/yago/Resource113331778">http://dbpedia.org/class/yago/Resource113331778</a>
	analysis	<a href="http://dbpedia.org/class/yago/Analysis105772667">http://dbpedia.org/class/yago/Analysis105772667</a>
Water Analyses (Winter)	analysis	<a href="http://dbpedia.org/class/yago/Analysis100634276">http://dbpedia.org/class/yago/Analysis100634276</a>
	analysis	<a href="http://dbpedia.org/class/yago/Analysis107067876">http://dbpedia.org/class/yago/Analysis107067876</a>
	analysis	<a href="http://dbpedia.org/class/yago/Analysis106376014">http://dbpedia.org/class/yago/Analysis106376014</a>
	water	<a href="http://dbpedia.org/class/yago/Water114845743">http://dbpedia.org/class/yago/Water114845743</a>
	method	<a href="http://dbpedia.org/class/yago/Method105660268">http://dbpedia.org/class/yago/Method105660268</a>
	event	<a href="http://dbpedia.org/class/yago/Event100029378">http://dbpedia.org/class/yago/Event100029378</a>
Water Analyses (Winter)	activity	<a href="http://dbpedia.org/class/yago/Activity100407535">http://dbpedia.org/class/yago/Activity100407535</a>
	analysis	<a href="http://dbpedia.org/class/yago/Analysis106014043">http://dbpedia.org/class/yago/Analysis106014043</a>
	basin	<a href="http://dbpedia.org/class/yago/Basin102801525">http://dbpedia.org/class/yago/Basin102801525</a>
	lake	<a href="http://dbpedia.org/class/yago/Lake109328904">http://dbpedia.org/class/yago/Lake109328904</a>
	river	<a href="http://dbpedia.org/class/yago/River109411430">http://dbpedia.org/class/yago/River109411430</a>
Water Analyses (Winter)	resource	<a href="http://dbpedia.org/class/yago/Resource113331778">http://dbpedia.org/class/yago/Resource113331778</a>
	analysis	<a href="http://dbpedia.org/class/yago/Analysis105772667">http://dbpedia.org/class/yago/Analysis105772667</a>
	analysis	<a href="http://dbpedia.org/class/yago/Analysis100634276">http://dbpedia.org/class/yago/Analysis100634276</a>
	analysis	<a href="http://dbpedia.org/class/yago/Analysis107067876">http://dbpedia.org/class/yago/Analysis107067876</a>
	analysis	<a href="http://dbpedia.org/class/yago/Analysis106376014">http://dbpedia.org/class/yago/Analysis106376014</a>

Continua na próxima página

Titulo	Nome do Conceito	URI do Conceito
	analysis	<a href="http://dbpedia.org/class/yago/Analysis106376014">http://dbpedia.org/class/yago/Analysis106376014</a>
Water Areas	water	<a href="http://dbpedia.org/class/yago/Water114845743">http://dbpedia.org/class/yago/Water114845743</a>
Water Management Units	water management management unit	<a href="http://dbpedia.org/class/yago/Water114845743">http://dbpedia.org/class/yago/Water114845743</a> <a href="http://dbpedia.org/class/yago/Management101133281">http://dbpedia.org/class/yago/Management101133281</a> <a href="http://dbpedia.org/class/yago/Management108381165">http://dbpedia.org/class/yago/Management108381165</a> <a href="http://dbpedia.org/class/yago/UnitOfMeasurement113583724">http://dbpedia.org/class/yago/UnitOfMeasurement113583724</a>
Zone of Subsidence	zone subsidence subsidence subsidence	<a href="http://dbpedia.org/class/yago/Zone108688247">http://dbpedia.org/class/yago/Zone108688247</a> <a href="http://dbpedia.org/class/yago/Remission107419960">http://dbpedia.org/class/yago/Remission107419960</a> <a href="http://dbpedia.org/class/yago/Settling107362218">http://dbpedia.org/class/yago/Settling107362218</a> <a href="http://dbpedia.org/class/yago/CaveIn107361416">http://dbpedia.org/class/yago/CaveIn107361416</a>
Napier City Council Election Wards	ward city council election physical entity municipality urban area administrative division	<a href="http://dbpedia.org/class/yago/Ward110767020">http://dbpedia.org/class/yago/Ward110767020</a> <a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a> <a href="http://dbpedia.org/class/yago/Election100181781">http://dbpedia.org/class/yago/Election100181781</a> <a href="http://dbpedia.org/class/yago/PhysicalEntity100001930">http://dbpedia.org/class/yago/PhysicalEntity100001930</a> <a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a> <a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a> <a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
NapierCityCouncilImContours	city council city council contour municipality urban area administrative division	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a> <a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a> <a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a> <a href="http://dbpedia.org/class/yago/Contour105851517">http://dbpedia.org/class/yago/Contour105851517</a> <a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a> <a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a> <a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
NapierCityCouncilAddressPoints	city council city council address municipality urban area administrative division	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a> <a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a> <a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a> <a href="http://dbpedia.org/class/yago/Address106356515">http://dbpedia.org/class/yago/Address106356515</a> <a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a> <a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a> <a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
NapierCityCouncilCadastralParcels	city council city council parcel municipality urban area administrative division	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a> <a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a> <a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a> <a href="http://dbpedia.org/class/yago/Parcel1101085098">http://dbpedia.org/class/yago/Parcel1101085098</a> <a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a> <a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a> <a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
NapierCityCouncilCarParks	city council city council car park municipality urban area administrative division	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a> <a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a> <a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a> <a href="http://dbpedia.org/class/yago/Car102958343">http://dbpedia.org/class/yago/Car102958343</a> <a href="http://dbpedia.org/class/yago/Park108615149">http://dbpedia.org/class/yago/Park108615149</a> <a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a> <a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a> <a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
NapierCityCouncilDogExerciseAreas	city council city council dog exercise neighbourhood neighbourhood neighbourhood municipality urban area administrative division	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a> <a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a> <a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a> <a href="http://dbpedia.org/class/yago/Dog102084071">http://dbpedia.org/class/yago/Dog102084071</a> <a href="http://dbpedia.org/class/yago/Exercise107454452">http://dbpedia.org/class/yago/Exercise107454452</a> <a href="http://dbpedia.org/class/yago/OttawaNeighbourhoods">http://dbpedia.org/class/yago/OttawaNeighbourhoods</a> <a href="http://dbpedia.org/class/yago/AgartalaNeighbourhoods">http://dbpedia.org/class/yago/AgartalaNeighbourhoods</a> <a href="http://dbpedia.org/class/yago/EdmontonNeighbourhoods">http://dbpedia.org/class/yago/EdmontonNeighbourhoods</a> <a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a> <a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a> <a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
NapierCityCouncilHeritageBuildingDatabase	city council city council heritage	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a> <a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a> <a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a> <a href="http://dbpedia.org/class/yago/Heritage105669797">http://dbpedia.org/class/yago/Heritage105669797</a>

*Continua na próxima página*

Titulo	Nome do Conceito	URI do Conceito
	heritage	<a href="http://dbpedia.org/class/yago/Inheritance104921011">http://dbpedia.org/class/yago/Inheritance104921011</a>
NapierCityCouncilHeritageBuildingDatabase	heritage	<a href="http://dbpedia.org/class/yago/Inheritance113262913">http://dbpedia.org/class/yago/Inheritance113262913</a>
	heritage	<a href="http://dbpedia.org/class/yago/Inheritance100083585">http://dbpedia.org/class/yago/Inheritance100083585</a>
	edifice	<a href="http://dbpedia.org/class/yago/Building102913152">http://dbpedia.org/class/yago/Building102913152</a>
	database	<a href="http://dbpedia.org/class/yago/Database106637824">http://dbpedia.org/class/yago/Database106637824</a>
	municipality	<a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a>
	urban area	<a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a>
NapierCityCouncilLeasedParking	administrative division	<a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
	city council	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a>
	city	<a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a>
	council	<a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a>
	parking	<a href="http://dbpedia.org/class/yago/Parking113778671">http://dbpedia.org/class/yago/Parking113778671</a>
	parking	<a href="http://dbpedia.org/class/yago/Parking100168505">http://dbpedia.org/class/yago/Parking100168505</a>
	municipality	<a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a>
NapierCityCouncilLiquorBan	urban area	<a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a>
	administrative division	<a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
	park	<a href="http://dbpedia.org/class/yago/Park108615149">http://dbpedia.org/class/yago/Park108615149</a>
	city council	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a>
	city	<a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a>
	council	<a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a>
	liquor	<a href="http://dbpedia.org/class/yago/Liquor114941230">http://dbpedia.org/class/yago/Liquor114941230</a>
NapierCityCouncilMobilityParks	ban	<a href="http://dbpedia.org/class/yago/Bans">http://dbpedia.org/class/yago/Bans</a>
	ban	<a href="http://dbpedia.org/class/yago/Ban113702315">http://dbpedia.org/class/yago/Ban113702315</a>
	ban	<a href="http://dbpedia.org/class/yago/Ban113685578">http://dbpedia.org/class/yago/Ban113685578</a>
	municipality	<a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a>
	urban area	<a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a>
	administrative division	<a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
NapierCityCouncilMotorcycleParks	city council	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a>
	city	<a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a>
	council	<a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a>
	mobility	<a href="http://dbpedia.org/class/yago/Mobility104773351">http://dbpedia.org/class/yago/Mobility104773351</a>
	park	<a href="http://dbpedia.org/class/yago/Park108615149">http://dbpedia.org/class/yago/Park108615149</a>
	municipality	<a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a>
	urban area	<a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a>
NapierCityCouncilOnStreetParks	administrative division	<a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
	city council	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a>
	city	<a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a>
	council	<a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a>
	street	<a href="http://dbpedia.org/class/yago/Street104334599">http://dbpedia.org/class/yago/Street104334599</a>
	park	<a href="http://dbpedia.org/class/yago/Park108615149">http://dbpedia.org/class/yago/Park108615149</a>
	municipality	<a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a>
NapierCityCouncilOperativePlanZones	urban area	<a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a>
	administrative division	<a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
	city council	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a>
	city	<a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a>
	council	<a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a>
	operative	<a href="http://dbpedia.org/class/yago/BritishIntelligenceOperatives">http://dbpedia.org/class/yago/BritishIntelligenceOperatives</a>
NapierCityCouncilPayandDisplayCarparks	plan	<a href="http://dbpedia.org/class/yago/SeatingPlans">http://dbpedia.org/class/yago/SeatingPlans</a>
	zone	<a href="http://dbpedia.org/class/yago/Zone108688247">http://dbpedia.org/class/yago/Zone108688247</a>
	municipality	<a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a>
	urban area	<a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a>
	administrative division	<a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
	city council	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a>

Continua na próxima página

<b>Título</b>	<b>Nome do Conceito</b>	<b>URI do Conceito</b>
NapierCityCouncilPayandDisplayCarparks	city	<a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a>
NapierCityCouncilPayandDisplayCarparks	council display display display display display municipality urban area administrative division	<a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a> <a href="http://dbpedia.org/class/yago/LiquidCrystalDisplays">http://dbpedia.org/class/yago/LiquidCrystalDisplays</a> <a href="http://dbpedia.org/class/yago/Display106879521">http://dbpedia.org/class/yago/Display106879521</a> <a href="http://dbpedia.org/class/yago/Display103210940">http://dbpedia.org/class/yago/Display103210940</a> <a href="http://dbpedia.org/class/yago/Display107213717">http://dbpedia.org/class/yago/Display107213717</a> <a href="http://dbpedia.org/class/yago/Display106887726">http://dbpedia.org/class/yago/Display106887726</a> <a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a> <a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a> <a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
NapierCityCouncilSubdivisionApplications	city council city council subdivision application program municipality urban area administrative division	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a> <a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a> <a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a> <a href="http://dbpedia.org/class/yago/Subdivision108674251">http://dbpedia.org/class/yago/Subdivision108674251</a> <a href="http://dbpedia.org/class/yago/Application106570110">http://dbpedia.org/class/yago/Application106570110</a> <a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a> <a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a> <a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
NapierCityCouncilTitleParcels	city council city council parcel municipality urban area administrative division title title title title title title	<a href="http://dbpedia.org/class/yago/CityCouncil108311282">http://dbpedia.org/class/yago/CityCouncil108311282</a> <a href="http://dbpedia.org/class/yago/City108524735">http://dbpedia.org/class/yago/City108524735</a> <a href="http://dbpedia.org/class/yago/Council108310949">http://dbpedia.org/class/yago/Council108310949</a> <a href="http://dbpedia.org/class/yago/Parcel1101085098">http://dbpedia.org/class/yago/Parcel1101085098</a> <a href="http://dbpedia.org/class/yago/Municipality108626283">http://dbpedia.org/class/yago/Municipality108626283</a> <a href="http://dbpedia.org/class/yago/UrbanArea108675967">http://dbpedia.org/class/yago/UrbanArea108675967</a> <a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a> <a href="http://dbpedia.org/class/yago/Title106343117">http://dbpedia.org/class/yago/Title106343117</a> <a href="http://dbpedia.org/class/yago/Title106346461">http://dbpedia.org/class/yago/Title106346461</a> <a href="http://dbpedia.org/class/yago/Title106343520">http://dbpedia.org/class/yago/Title106343520</a> <a href="http://dbpedia.org/class/yago/Title106345993">http://dbpedia.org/class/yago/Title106345993</a> <a href="http://dbpedia.org/class/yago/Title106346220">http://dbpedia.org/class/yago/Title106346220</a> <a href="http://dbpedia.org/class/yago/Title105181199">http://dbpedia.org/class/yago/Title105181199</a>
ne_50m_boundary_lines_land	boundary land	<a href="http://dbpedia.org/class/yago/WritingSystemsWithoutWordBoundaries">http://dbpedia.org/class/yago/WritingSystemsWithoutWordBoundaries</a> <a href="http://dbpedia.org/class/yago/Land113250048">http://dbpedia.org/class/yago/Land113250048</a>
ne_50m_coastline	coastline Ne NE	<a href="http://dbpedia.org/class/yago/Coastline108613000">http://dbpedia.org/class/yago/Coastline108613000</a> <a href="http://dbpedia.org/class/yago/Neon114646152">http://dbpedia.org/class/yago/Neon114646152</a> <a href="http://dbpedia.org/class/yago/Northeast113831843">http://dbpedia.org/class/yago/Northeast113831843</a>

# Apêndice B

## Linha de Base para a Anotação no Nível de Atributos

A linha de base para a anotação no nível de atributos é composta de cem nomes de atributos de *feature types* e os respectivos *ids* dos conceitos relevantes ligados semanticamente a eles. Para melhor ilustrar a linha de base, serão listados na Tabela B.1 os títulos dos atributos com os nomes e as URI dos conceitos relevantes.

Tabela B.1: Linha de base para a anotação no nível de atributos.

Nome do Atributo	Nome do Conceito	URI do Conceito
ACTION	action	<a href="http://dbpedia.org/class/yago/Action100037396">http://dbpedia.org/class/yago/Action100037396</a>
ADDRESS	address	<a href="http://dbpedia.org/class/yago/Address106356515">http://dbpedia.org/class/yago/Address106356515</a>
ADVICE	advice	<a href="http://dbpedia.org/class/yago/Advice106671484">http://dbpedia.org/class/yago/Advice106671484</a>
AGENT	agent	<a href="http://dbpedia.org/class/yago/Agent109190918">http://dbpedia.org/class/yago/Agent109190918</a>
AIRZONE	air air air air air zone	<a href="http://dbpedia.org/class/yago/Air114841267">http://dbpedia.org/class/yago/Air114841267</a> <a href="http://dbpedia.org/class/yago/Air108653314">http://dbpedia.org/class/yago/Air108653314</a> <a href="http://dbpedia.org/class/yago/Air104727214">http://dbpedia.org/class/yago/Air104727214</a> <a href="http://dbpedia.org/class/yago/Atmosphere108499057">http://dbpedia.org/class/yago/Atmosphere108499057</a> <a href="http://dbpedia.org/class/yago/Air114842703">http://dbpedia.org/class/yago/Air114842703</a> <a href="http://dbpedia.org/class/yago/Zone108688247">http://dbpedia.org/class/yago/Zone108688247</a>
APPLICANT	applicant	<a href="http://dbpedia.org/class/yago/Applicant109607280">http://dbpedia.org/class/yago/Applicant109607280</a>
APPLICATION_NUMBER	application program figure	<a href="http://dbpedia.org/class/yago/Application106570110">http://dbpedia.org/class/yago/Application106570110</a> <a href="http://dbpedia.org/class/yago/Number105121418">http://dbpedia.org/class/yago/Number105121418</a>
AQUATIC_TERRESTRIAL_WETLAND	aquatic wetland wetland wetland	<a href="http://dbpedia.org/class/yago/Aquatic111536673">http://dbpedia.org/class/yago/Aquatic111536673</a> <a href="http://dbpedia.org/class/yago/Wetlands">http://dbpedia.org/class/yago/Wetlands</a> <a href="http://dbpedia.org/class/yago/FictionalWetlands">http://dbpedia.org/class/yago/FictionalWetlands</a> <a href="http://dbpedia.org/class/yago/Wetland109477890">http://dbpedia.org/class/yago/Wetland109477890</a>
AREAHECTARE	area hectare	<a href="http://dbpedia.org/class/yago/Area108497294">http://dbpedia.org/class/yago/Area108497294</a> <a href="http://dbpedia.org/class/yago/Hectare113613985">http://dbpedia.org/class/yago/Hectare113613985</a>
AUDIT_YEAR	audit year year year year	<a href="http://dbpedia.org/class/yago/Audit100141176">http://dbpedia.org/class/yago/Audit100141176</a> <a href="http://dbpedia.org/class/yago/EducationalYears">http://dbpedia.org/class/yago/EducationalYears</a> <a href="http://dbpedia.org/class/yago/Years">http://dbpedia.org/class/yago/Years</a> <a href="http://dbpedia.org/class/yago/Year115204297">http://dbpedia.org/class/yago/Year115204297</a> <a href="http://dbpedia.org/class/yago/Year115201505">http://dbpedia.org/class/yago/Year115201505</a>
BEACHPHOTO	beach	<a href="http://dbpedia.org/class/yago/Beach109217230">http://dbpedia.org/class/yago/Beach109217230</a>

Continua na próxima página

Nome do Atributo	Nome do Conceito	URI do Conceito
BOUNDARYSOURCE	boundary source	<a href="http://dbpedia.org/class/yago/WritingSystemsWithoutWordBoundaries">http://dbpedia.org/class/yago/WritingSystemsWithoutWordBoundaries</a> <a href="http://dbpedia.org/class/yago/Beginning108507558">http://dbpedia.org/class/yago/Beginning108507558</a>
BUILD_COUNT	count	<a href="http://dbpedia.org/class/yago/Count109969218">http://dbpedia.org/class/yago/Count109969218</a>
BYLAW	bylaw	<a href="http://dbpedia.org/class/yago/Bylaw106537951">http://dbpedia.org/class/yago/Bylaw106537951</a>
CAPACITY	capacity capacity capacity capacity capacity	<a href="http://dbpedia.org/class/yago/Capacity105203397">http://dbpedia.org/class/yago/Capacity105203397</a> <a href="http://dbpedia.org/class/yago/Capacity100914795">http://dbpedia.org/class/yago/Capacity100914795</a> <a href="http://dbpedia.org/class/yago/Capacity100720951">http://dbpedia.org/class/yago/Capacity100720951</a> <a href="http://dbpedia.org/class/yago/Capacity113755053">http://dbpedia.org/class/yago/Capacity113755053</a> <a href="http://dbpedia.org/class/yago/Capacity105033569">http://dbpedia.org/class/yago/Capacity105033569</a>
CASE_OFFICER	case officer	<a href="http://dbpedia.org/class/yago/CaseOfficer109899134">http://dbpedia.org/class/yago/CaseOfficer109899134</a>
CATCHMENT	catchment	<a href="http://dbpedia.org/class/yago/Catchment102983357">http://dbpedia.org/class/yago/Catchment102983357</a>
CATEGORY	category	<a href="http://dbpedia.org/class/yago/Class107997703">http://dbpedia.org/class/yago/Class107997703</a>
CHANGE	change	<a href="http://dbpedia.org/class/yago/PhaseChanges">http://dbpedia.org/class/yago/PhaseChanges</a>
CLASSIFICATION	compartmentalization	<a href="http://dbpedia.org/class/yago/Categorization101012712">http://dbpedia.org/class/yago/Categorization101012712</a>
CODE	code	<a href="http://dbpedia.org/class/yago/Code106667317">http://dbpedia.org/class/yago/Code106667317</a>
COMMENT	comment comment comment	<a href="http://dbpedia.org/class/yago/Remark106765044">http://dbpedia.org/class/yago/Remark106765044</a> <a href="http://dbpedia.org/class/yago/Comment106762711">http://dbpedia.org/class/yago/Comment106762711</a> <a href="http://dbpedia.org/class/yago/Gossip107223170">http://dbpedia.org/class/yago/Gossip107223170</a>
COMMODITY_PRODUCED	commodity	<a href="http://dbpedia.org/class/yago/Commodity103076708">http://dbpedia.org/class/yago/Commodity103076708</a>
CONFIRMATION_DATE	confirmation confirmation confirmation date	<a href="http://dbpedia.org/class/yago/Confirmation106650070">http://dbpedia.org/class/yago/Confirmation106650070</a> <a href="http://dbpedia.org/class/yago/Confirmation101038895">http://dbpedia.org/class/yago/Confirmation101038895</a> <a href="http://dbpedia.org/class/yago/Confirmation101038761">http://dbpedia.org/class/yago/Confirmation101038761</a> <a href="http://dbpedia.org/class/yago/Date115159583">http://dbpedia.org/class/yago/Date115159583</a>
COST	cost cost cost cost	<a href="http://dbpedia.org/class/yago/Costs">http://dbpedia.org/class/yago/Costs</a> <a href="http://dbpedia.org/class/yago/Cost113275847">http://dbpedia.org/class/yago/Cost113275847</a> <a href="http://dbpedia.org/class/yago/MonetaryValue105145118">http://dbpedia.org/class/yago/MonetaryValue105145118</a> <a href="http://dbpedia.org/class/yago/Price105163807">http://dbpedia.org/class/yago/Price105163807</a>
COUNTY	county	<a href="http://dbpedia.org/class/yago/County108546183">http://dbpedia.org/class/yago/County108546183</a>
COVERAGE	coverage	<a href="http://dbpedia.org/class/yago/Coverage105123760">http://dbpedia.org/class/yago/Coverage105123760</a>
DATE_DATA_REFRESHED	date data	<a href="http://dbpedia.org/class/yago/Date115159583">http://dbpedia.org/class/yago/Date115159583</a> <a href="http://dbpedia.org/class/yago/Data108462320">http://dbpedia.org/class/yago/Data108462320</a>
DEPTH	depth	<a href="http://dbpedia.org/class/yago/Depth105134547">http://dbpedia.org/class/yago/Depth105134547</a>
DESCRIPTION	description description description description	<a href="http://dbpedia.org/class/yago/LogoDescriptions">http://dbpedia.org/class/yago/LogoDescriptions</a> <a href="http://dbpedia.org/class/yago/Description106724763">http://dbpedia.org/class/yago/Description106724763</a> <a href="http://dbpedia.org/class/yago/Description107201365">http://dbpedia.org/class/yago/Description107201365</a> <a href="http://dbpedia.org/class/yago/Description105840076">http://dbpedia.org/class/yago/Description105840076</a>
DISPLAY	display display display display	<a href="http://dbpedia.org/class/yago/Display106879521">http://dbpedia.org/class/yago/Display106879521</a> <a href="http://dbpedia.org/class/yago/Display103210940">http://dbpedia.org/class/yago/Display103210940</a> <a href="http://dbpedia.org/class/yago/Display107213717">http://dbpedia.org/class/yago/Display107213717</a> <a href="http://dbpedia.org/class/yago/Display106887726">http://dbpedia.org/class/yago/Display106887726</a>
EDENMONENTITY	Mon Mon entity	<a href="http://dbpedia.org/class/yago/Mon110327696">http://dbpedia.org/class/yago/Mon110327696</a> <a href="http://dbpedia.org/class/yago/Mon106937441">http://dbpedia.org/class/yago/Mon106937441</a> <a href="http://dbpedia.org/class/yago/Entity100001740">http://dbpedia.org/class/yago/Entity100001740</a>
FEATURE	characteristic	<a href="http://dbpedia.org/class/yago/Feature105849789">http://dbpedia.org/class/yago/Feature105849789</a>
FINISH_DAT	finish finish finish DAT	<a href="http://dbpedia.org/class/yago/Finish107353376">http://dbpedia.org/class/yago/Finish107353376</a> <a href="http://dbpedia.org/class/yago/Finish107333162">http://dbpedia.org/class/yago/Finish107333162</a> <a href="http://dbpedia.org/class/yago/Finish105717747">http://dbpedia.org/class/yago/Finish105717747</a> <a href="http://dbpedia.org/class/yago/DigitalAudiotape103195959">http://dbpedia.org/class/yago/DigitalAudiotape103195959</a>
FURTHER_INFO_PUBLIC	info public public	<a href="http://dbpedia.org/class/yago/Information106634376">http://dbpedia.org/class/yago/Information106634376</a> <a href="http://dbpedia.org/class/yago/MinistreFrançaisDesTravauxPublics">http://dbpedia.org/class/yago/MinistreFrançaisDesTravauxPublics</a> <a href="http://dbpedia.org/class/yago/Public107965817">http://dbpedia.org/class/yago/Public107965817</a>
GROUNDWATER_DEPENDENT	solid ground water dependent	<a href="http://dbpedia.org/class/yago/Land109334396">http://dbpedia.org/class/yago/Land109334396</a> <a href="http://dbpedia.org/class/yago/Water114845743">http://dbpedia.org/class/yago/Water114845743</a> <a href="http://dbpedia.org/class/yago/Dependant110004804">http://dbpedia.org/class/yago/Dependant110004804</a>
HABITAT_SPECIES	habitat species species	<a href="http://dbpedia.org/class/yago/Habitat108580583">http://dbpedia.org/class/yago/Habitat108580583</a> <a href="http://dbpedia.org/class/yago/Species108110373">http://dbpedia.org/class/yago/Species108110373</a> <a href="http://dbpedia.org/class/yago/Species105844881">http://dbpedia.org/class/yago/Species105844881</a>
HIGH_SCHOOL_CLUSTER	high high school cluster	<a href="http://dbpedia.org/class/yago/High114405621">http://dbpedia.org/class/yago/High114405621</a> <a href="http://dbpedia.org/class/yago/High114405452">http://dbpedia.org/class/yago/High114405452</a> <a href="http://dbpedia.org/class/yago/School108276720">http://dbpedia.org/class/yago/School108276720</a> <a href="http://dbpedia.org/class/yago/Bunch107959943">http://dbpedia.org/class/yago/Bunch107959943</a>

*Continua na próxima página*

Nome do Atributo	Nome do Conceito	URI do Conceito
HIGH_SCHOOL_CLUSTER	high	<a href="http://dbpedia.org/class/yago/High105097536">http://dbpedia.org/class/yago/High105097536</a>
	high	<a href="http://dbpedia.org/class/yago/High114520670">http://dbpedia.org/class/yago/High114520670</a>
HOUSING_MARKET_AREA	housing	<a href="http://dbpedia.org/class/yago/Housing103546766">http://dbpedia.org/class/yago/Housing103546766</a>
	market	<a href="http://dbpedia.org/class/yago/Market101097292">http://dbpedia.org/class/yago/Market101097292</a>
	market	<a href="http://dbpedia.org/class/yago/GroceryStore103461385">http://dbpedia.org/class/yago/GroceryStore103461385</a>
	market	<a href="http://dbpedia.org/class/yago/Marketplace103722288">http://dbpedia.org/class/yago/Marketplace103722288</a>
HYPERLINK	area	<a href="http://dbpedia.org/class/yago/Area108497294">http://dbpedia.org/class/yago/Area108497294</a>
	hyperlink	<a href="http://dbpedia.org/class/yago/Hyperlink106586672">http://dbpedia.org/class/yago/Hyperlink106586672</a>
INSPECTION_MONTH	month	<a href="http://dbpedia.org/class/yago/Months">http://dbpedia.org/class/yago/Months</a>
	month	<a href="http://dbpedia.org/class/yago/BlackMonths">http://dbpedia.org/class/yago/BlackMonths</a>
	month	<a href="http://dbpedia.org/class/yago/Month115206296">http://dbpedia.org/class/yago/Month115206296</a>
INTERIOR	interior	<a href="http://dbpedia.org/class/yago/Inside108588294">http://dbpedia.org/class/yago/Inside108588294</a>
	interior	<a href="http://dbpedia.org/class/yago/Inside108588152">http://dbpedia.org/class/yago/Inside108588152</a>
LABEL	label	<a href="http://dbpedia.org/class/yago/Label107202579">http://dbpedia.org/class/yago/Label107202579</a>
LATITUDE	latitude	<a href="http://dbpedia.org/class/yago/Latitude108595531">http://dbpedia.org/class/yago/Latitude108595531</a>
	latitude	<a href="http://dbpedia.org/class/yago/Latitude113995148">http://dbpedia.org/class/yago/Latitude113995148</a>
	latitude	<a href="http://dbpedia.org/class/yago/Latitude105127640">http://dbpedia.org/class/yago/Latitude105127640</a>
LICENCESTATUSDESCRIPTION	status	<a href="http://dbpedia.org/class/yago/Condition113920835">http://dbpedia.org/class/yago/Condition113920835</a>
	description	<a href="http://dbpedia.org/class/yago/LogoDescriptions">http://dbpedia.org/class/yago/LogoDescriptions</a>
	description	<a href="http://dbpedia.org/class/yago/Description106724763">http://dbpedia.org/class/yago/Description106724763</a>
	description	<a href="http://dbpedia.org/class/yago/Description107201365">http://dbpedia.org/class/yago/Description107201365</a>
LOCAL_PLAN_REFERENCE	description	<a href="http://dbpedia.org/class/yago/Description105840076">http://dbpedia.org/class/yago/Description105840076</a>
	reference	<a href="http://dbpedia.org/class/yago/MediaWikiReferences">http://dbpedia.org/class/yago/MediaWikiReferences</a>
LOCATION	reference	<a href="http://dbpedia.org/class/yago/Reference105923314">http://dbpedia.org/class/yago/Reference105923314</a>
	location	<a href="http://dbpedia.org/class/yago/Location100027167">http://dbpedia.org/class/yago/Location100027167</a>
MARINEWATER_DEPENDENT	marine	<a href="http://dbpedia.org/class/yago/Marines">http://dbpedia.org/class/yago/Marines</a>
	water	<a href="http://dbpedia.org/class/yago/Water114845743">http://dbpedia.org/class/yago/Water114845743</a>
	dependent	<a href="http://dbpedia.org/class/yago/Dependant110004804">http://dbpedia.org/class/yago/Dependant110004804</a>
NATIONAL_PARK	national park	<a href="http://dbpedia.org/class/yago/NationalPark108600992">http://dbpedia.org/class/yago/NationalPark108600992</a>
NUMBER_OF_TURBINES	figure	<a href="http://dbpedia.org/class/yago/Number105121418">http://dbpedia.org/class/yago/Number105121418</a>
OFFICIAL_GRID_REFERENCE	official	<a href="http://dbpedia.org/class/yago/Official110372373">http://dbpedia.org/class/yago/Official110372373</a>
	grid	<a href="http://dbpedia.org/class/yago/Grid105931512">http://dbpedia.org/class/yago/Grid105931512</a>
	grid	<a href="http://dbpedia.org/class/yago/Grid103458753">http://dbpedia.org/class/yago/Grid103458753</a>
	reference	<a href="http://dbpedia.org/class/yago/MediaWikiReferences">http://dbpedia.org/class/yago/MediaWikiReferences</a>
	reference	<a href="http://dbpedia.org/class/yago/Reference105923314">http://dbpedia.org/class/yago/Reference105923314</a>
OPERATOR	operator	<a href="http://dbpedia.org/class/yago/Operator113786413">http://dbpedia.org/class/yago/Operator113786413</a>
PERIMETER	perimeter	<a href="http://dbpedia.org/class/yago/Perimeter113871452">http://dbpedia.org/class/yago/Perimeter113871452</a>
	perimeter	<a href="http://dbpedia.org/class/yago/Circumference105101261">http://dbpedia.org/class/yago/Circumference105101261</a>
POLICY_TYPE	policy	<a href="http://dbpedia.org/class/yago/Policy105901508">http://dbpedia.org/class/yago/Policy105901508</a>
	type	<a href="http://dbpedia.org/class/yago/Type105840188">http://dbpedia.org/class/yago/Type105840188</a>
PREVIOUS_MINERAL_WORKINGS	workings	<a href="http://dbpedia.org/class/yago/Working104601473">http://dbpedia.org/class/yago/Working104601473</a>
	mineral	<a href="http://dbpedia.org/class/yago/Mineral114662574">http://dbpedia.org/class/yago/Mineral114662574</a>
PROCESSINGSTATUS	processing	<a href="http://dbpedia.org/class/yago/Processing113541167">http://dbpedia.org/class/yago/Processing113541167</a>
	status	<a href="http://dbpedia.org/class/yago/Condition113920835">http://dbpedia.org/class/yago/Condition113920835</a>
PROPOSAL	proposal	<a href="http://dbpedia.org/class/yago/Proposal107162194">http://dbpedia.org/class/yago/Proposal107162194</a>
PUBLIC_FURTHER_INFO	public	<a href="http://dbpedia.org/class/yago/MinistreFrancaisDesTravauxPublics">http://dbpedia.org/class/yago/MinistreFrancaisDesTravauxPublics</a>
	public	<a href="http://dbpedia.org/class/yago/Public107965817">http://dbpedia.org/class/yago/Public107965817</a>
	info	<a href="http://dbpedia.org/class/yago/Information106634376">http://dbpedia.org/class/yago/Information106634376</a>
RAIL_ROAD_VISIBILITY	rail	<a href="http://dbpedia.org/class/yago/Rail104046810">http://dbpedia.org/class/yago/Rail104046810</a>
	rail	<a href="http://dbpedia.org/class/yago/Rail104046590">http://dbpedia.org/class/yago/Rail104046590</a>
	rail	<a href="http://dbpedia.org/class/yago/Rail102014941">http://dbpedia.org/class/yago/Rail102014941</a>
	route	<a href="http://dbpedia.org/class/yago/Road104096066">http://dbpedia.org/class/yago/Road104096066</a>
	visibility	<a href="http://dbpedia.org/class/yago/Visibility114434329">http://dbpedia.org/class/yago/Visibility114434329</a>
RECOMMENDED_USAGE	visibility	<a href="http://dbpedia.org/class/yago/Visibility104702545">http://dbpedia.org/class/yago/Visibility104702545</a>
	usage	<a href="http://dbpedia.org/class/yago/Custom100413239">http://dbpedia.org/class/yago/Custom100413239</a>
RECORD_STATUS	usage	<a href="http://dbpedia.org/class/yago/Usage106283459">http://dbpedia.org/class/yago/Usage106283459</a>
	record	<a href="http://dbpedia.org/class/yago/Record106647206">http://dbpedia.org/class/yago/Record106647206</a>
RECORD_STATUS	status	<a href="http://dbpedia.org/class/yago/Condition113920835">http://dbpedia.org/class/yago/Condition113920835</a>
REFERENCE	reference	<a href="http://dbpedia.org/class/yago/MediaWikiReferences">http://dbpedia.org/class/yago/MediaWikiReferences</a>
	reference	<a href="http://dbpedia.org/class/yago/Reference105923314">http://dbpedia.org/class/yago/Reference105923314</a>
REGION	region	<a href="http://dbpedia.org/class/yago/Region108630039">http://dbpedia.org/class/yago/Region108630039</a>

*Continua na próxima página*

Nome do Atributo	Nome do Conceito	URI do Conceito
REMARK	remark	<a href="http://dbpedia.org/class/yago/Remark105710573">http://dbpedia.org/class/yago/Remark105710573</a>
REVIEW_DATE	review date reevaluation	<a href="http://dbpedia.org/class/yago/Reviews">http://dbpedia.org/class/yago/Reviews</a> <a href="http://dbpedia.org/class/yago/Date115159583">http://dbpedia.org/class/yago/Date115159583</a> <a href="http://dbpedia.org/class/yago/Reappraisal105747582">http://dbpedia.org/class/yago/Reappraisal105747582</a>
RINGS_NOK	rings	<a href="http://dbpedia.org/class/yago/Rings104093223">http://dbpedia.org/class/yago/Rings104093223</a>
SALINITY	salinity salinity	<a href="http://dbpedia.org/class/yago/Salt105717342">http://dbpedia.org/class/yago/Salt105717342</a> <a href="http://dbpedia.org/class/yago/Brininess104993604">http://dbpedia.org/class/yago/Brininess104993604</a>
SCALE	scale of measurement	<a href="http://dbpedia.org/class/yago/Scale113850304">http://dbpedia.org/class/yago/Scale113850304</a>
SCHOOL_NAME	school name name name name name name	<a href="http://dbpedia.org/class/yago/School108276720">http://dbpedia.org/class/yago/School108276720</a> <a href="http://dbpedia.org/class/yago/Name106333653">http://dbpedia.org/class/yago/Name106333653</a> <a href="http://dbpedia.org/class/yago/Name114438788">http://dbpedia.org/class/yago/Name114438788</a> <a href="http://dbpedia.org/class/yago/Name107972279">http://dbpedia.org/class/yago/Name107972279</a> <a href="http://dbpedia.org/class/yago/Name110344443">http://dbpedia.org/class/yago/Name110344443</a> <a href="http://dbpedia.org/class/yago/Name101139636">http://dbpedia.org/class/yago/Name101139636</a> <a href="http://dbpedia.org/class/yago/Name106720964">http://dbpedia.org/class/yago/Name106720964</a>
SETTLEMENT	administrative division	<a href="http://dbpedia.org/class/yago/AdministrativeDistrict108491826">http://dbpedia.org/class/yago/AdministrativeDistrict108491826</a>
SHAPE	shape	<a href="http://dbpedia.org/class/yago/Shape105064037">http://dbpedia.org/class/yago/Shape105064037</a>
SITE_CAPACITY	land site capacity capacity capacity capacity capacity	<a href="http://dbpedia.org/class/yago/Site108651247">http://dbpedia.org/class/yago/Site108651247</a> <a href="http://dbpedia.org/class/yago/Capacity105203397">http://dbpedia.org/class/yago/Capacity105203397</a> <a href="http://dbpedia.org/class/yago/Capacity100914795">http://dbpedia.org/class/yago/Capacity100914795</a> <a href="http://dbpedia.org/class/yago/Capacity100720951">http://dbpedia.org/class/yago/Capacity100720951</a> <a href="http://dbpedia.org/class/yago/Capacity113755053">http://dbpedia.org/class/yago/Capacity113755053</a> <a href="http://dbpedia.org/class/yago/Capacity105033569">http://dbpedia.org/class/yago/Capacity105033569</a>
SITE_CHARACTERISTICS	land site	<a href="http://dbpedia.org/class/yago/Site108651247">http://dbpedia.org/class/yago/Site108651247</a>
SIZE	size size size size	<a href="http://dbpedia.org/class/yago/Size105098942">http://dbpedia.org/class/yago/Size105098942</a> <a href="http://dbpedia.org/class/yago/Size105095691">http://dbpedia.org/class/yago/Size105095691</a> <a href="http://dbpedia.org/class/yago/Size113937727">http://dbpedia.org/class/yago/Size113937727</a> <a href="http://dbpedia.org/class/yago/Size105099231">http://dbpedia.org/class/yago/Size105099231</a>
SOIL_GROUP	group	<a href="http://dbpedia.org/class/yago/Group100031264">http://dbpedia.org/class/yago/Group100031264</a>
SOURCE	source	<a href="http://dbpedia.org/class/yago/Beginning108507558">http://dbpedia.org/class/yago/Beginning108507558</a>
SPECIES	species species	<a href="http://dbpedia.org/class/yago/Species108110373">http://dbpedia.org/class/yago/Species108110373</a> <a href="http://dbpedia.org/class/yago/Species105844881">http://dbpedia.org/class/yago/Species105844881</a>
START_DATE	start start start date	<a href="http://dbpedia.org/class/yago/Start107325190">http://dbpedia.org/class/yago/Start107325190</a> <a href="http://dbpedia.org/class/yago/Start100241699">http://dbpedia.org/class/yago/Start100241699</a> <a href="http://dbpedia.org/class/yago/StartingSignal106792188">http://dbpedia.org/class/yago/StartingSignal106792188</a> <a href="http://dbpedia.org/class/yago/Date115159583">http://dbpedia.org/class/yago/Date115159583</a>
STATIONPOSITION	station position	<a href="http://dbpedia.org/class/yago/Station104306080">http://dbpedia.org/class/yago/Station104306080</a> <a href="http://dbpedia.org/class/yago/Position108621598">http://dbpedia.org/class/yago/Position108621598</a>
STATUS	status	<a href="http://dbpedia.org/class/yago/Condition113920835">http://dbpedia.org/class/yago/Condition113920835</a>
STRUCTURE	structure	<a href="http://dbpedia.org/class/yago/Structure104341686">http://dbpedia.org/class/yago/Structure104341686</a>
STYLE	mode	<a href="http://dbpedia.org/class/yago/Manner104928903">http://dbpedia.org/class/yago/Manner104928903</a>
SUBCLASS	subclass	<a href="http://dbpedia.org/class/yago/Subclass108106661">http://dbpedia.org/class/yago/Subclass108106661</a>
SUBSITEPREFIX	prefix prefix prefix	<a href="http://dbpedia.org/class/yago/Prefixes">http://dbpedia.org/class/yago/Prefixes</a> <a href="http://dbpedia.org/class/yago/SIPrefixes">http://dbpedia.org/class/yago/SIPrefixes</a> <a href="http://dbpedia.org/class/yago/Prefix106308304">http://dbpedia.org/class/yago/Prefix106308304</a>
SURFACEWATER_DEPENDENT	surface water dependent	<a href="http://dbpedia.org/class/yago/Surface104362025">http://dbpedia.org/class/yago/Surface104362025</a> <a href="http://dbpedia.org/class/yago/Water114845743">http://dbpedia.org/class/yago/Water114845743</a> <a href="http://dbpedia.org/class/yago/Dependant110004804">http://dbpedia.org/class/yago/Dependant110004804</a>
SURROUNDING_USE_1	use use	<a href="http://dbpedia.org/class/yago/Use105149978">http://dbpedia.org/class/yago/Use105149978</a> <a href="http://dbpedia.org/class/yago/Habit100414179">http://dbpedia.org/class/yago/Habit100414179</a>
SURVEY	survey survey survey	<a href="http://dbpedia.org/class/yago/AstronomicalSurveys">http://dbpedia.org/class/yago/AstronomicalSurveys</a> <a href="http://dbpedia.org/class/yago/Survey100644503">http://dbpedia.org/class/yago/Survey100644503</a> <a href="http://dbpedia.org/class/yago/Sketch106469694">http://dbpedia.org/class/yago/Sketch106469694</a>
SYSTEM	system	<a href="http://dbpedia.org/class/yago/System104377057">http://dbpedia.org/class/yago/System104377057</a>
TIME_LIMIT	time limit	<a href="http://dbpedia.org/class/yago/TimeLimit115224156">http://dbpedia.org/class/yago/TimeLimit115224156</a>
TREE_NUMBER	tree figure	<a href="http://dbpedia.org/class/yago/Tree113104059">http://dbpedia.org/class/yago/Tree113104059</a> <a href="http://dbpedia.org/class/yago/Number105121418">http://dbpedia.org/class/yago/Number105121418</a>
TYPE	type	<a href="http://dbpedia.org/class/yago/Type105840188">http://dbpedia.org/class/yago/Type105840188</a>
URL	URL	<a href="http://dbpedia.org/class/yago/URL106358747">http://dbpedia.org/class/yago/URL106358747</a>
VERSION	version version	<a href="http://dbpedia.org/class/yago/Version107173585">http://dbpedia.org/class/yago/Version107173585</a> <a href="http://dbpedia.org/class/yago/Version101267901">http://dbpedia.org/class/yago/Version101267901</a>

*Continua na próxima página*

Nome do Atributo	Nome do Conceito	URI do Conceito
VOLUME	volume	<a href="http://dbpedia.org/class/yago/Volume113779032">http://dbpedia.org/class/yago/Volume113779032</a>
	volume	<a href="http://dbpedia.org/class/yago/Volume106413666">http://dbpedia.org/class/yago/Volume106413666</a>
	volume	<a href="http://dbpedia.org/class/yago/Volume113779244">http://dbpedia.org/class/yago/Volume113779244</a>
WARD	ward	<a href="http://dbpedia.org/class/yago/Ward110767020">http://dbpedia.org/class/yago/Ward110767020</a>
WATER_DEPENDANCY	water	<a href="http://dbpedia.org/class/yago/Water114845743">http://dbpedia.org/class/yago/Water114845743</a>
WATERTYPE	water	<a href="http://dbpedia.org/class/yago/Water114845743">http://dbpedia.org/class/yago/Water114845743</a>
	type	<a href="http://dbpedia.org/class/yago/Type105840188">http://dbpedia.org/class/yago/Type105840188</a>
WISEREFERENCEDATA	wise	<a href="http://dbpedia.org/class/yago/Wise104931733">http://dbpedia.org/class/yago/Wise104931733</a>
	reference	<a href="http://dbpedia.org/class/yago/MediaWikiReferences">http://dbpedia.org/class/yago/MediaWikiReferences</a>
	reference	<a href="http://dbpedia.org/class/yago/Reference105923314">http://dbpedia.org/class/yago/Reference105923314</a>
YEAR_REVIEWED	data	<a href="http://dbpedia.org/class/yago/Data108462320">http://dbpedia.org/class/yago/Data108462320</a>
	year	<a href="http://dbpedia.org/class/yago/EducationalYears">http://dbpedia.org/class/yago/EducationalYears</a>
	year	<a href="http://dbpedia.org/class/yago/Years">http://dbpedia.org/class/yago/Years</a>
	year	<a href="http://dbpedia.org/class/yago/Year115204297">http://dbpedia.org/class/yago/Year115204297</a>
ZONE_DOCS	year	<a href="http://dbpedia.org/class/yago/Year115201505">http://dbpedia.org/class/yago/Year115201505</a>
	zone	<a href="http://dbpedia.org/class/yago/Zone108688247">http://dbpedia.org/class/yago/Zone108688247</a>

# Apêndice C

## Resultado da Anotação no Nível de Camadas

Os resultados do desempenho da abordagem proposta para **anotação semântica** no nível de camadas estão listados na Tabela C.1. Cada linha desta tabela contempla, respectivamente, o título da camada, a sua métrica de precisão, a sua métrica de cobertura e o tipo de anotação realizado (via LOD ou via casamento de *strings*).

Tabela C.1: Resultado da anotação semântica da linha de base das camadas

<b>Título da Camada</b>	<b>Precisão</b>	<b>Cobertura</b>	<b>Tipo de Anotação</b>
Allotment Locations	1.0	0.33	Casamento de strings
Community Council Boundaries	1.0	0.66	Casamento de strings
Conservation Areas	1.0	1.0	Casamento de strings
Core Path Network	1.0	1.0	Casamento de strings
Defined Policy Areas	1.0	0.33	Casamento de strings
Development Proposals Area	1.0	0.5	Casamento de strings
Development Proposals Line	1.0	0.5	Casamento de strings
Employment Land Audit	1.0	0.75	Casamento de strings
Fife Common Good Land	0.66	0.66	Casamento de strings
Fife Primary School Catchments	1.0	0.5	Casamento de strings
Fife RC Primary School Catchments	1.0	0.4	Casamento de strings
Fife RC Secondary School Catchments	0.25	0.4	Casamento de strings

*Continua na próxima página*

<b>Título da Camada</b>	<b>Precisão</b>	<b>Cobertura</b>	<b>Tipo de Anotação</b>
Fife Secondary School Catchments	0.33	0.5	Casamento de strings
Housing Land Audit 2014	1.0	0.5	Casamento de strings
Listed Buildings	1.0	1.0	Casamento de strings
Local Landscape Areas	1.0	0.5	Casamento de strings
Natural Heritage Sites	0.6	0.375	LOD
Polling Districts	1.0	0.125	Casamento de strings
Polling Places	1.0	0.16	Casamento de strings
Tree Preservation Order Areas	1.0	1.0	Casamento de strings
Tree Preservation Order Points	1.0	1.0	Casamento de strings
Vacant And Derelict Land	1.0	0.33	Casamento de strings
Wind Turbines	0.66	1.0	LOD
Windfarm Broad Area Of Search	1.0	1.0	Casamento de strings
AIR EMEP Monitoring Sites	1.0	0.375	Casamento de strings
Air Monitoring Sites	1.0	0.375	Casamento de strings
Air Zones	1.0	0.33	Casamento de strings
Bathing Water Quality	1.0	0.4	Casamento de strings
Coal Restricted Areas	1.0	1.0	Casamento de strings
Coastal waterbodies	0.0	0.0	Não anotado
EPA Offices	1.0	0.66	Casamento de strings
Geological Linear Features	1.0	1.0	Casamento de strings
Ground waterbodies	1.0	0.25	Casamento de strings
Historical Groundwater Quality	1.0	1.0	Casamento de strings
Large Towns	1.0	0.5	Casamento de strings
Licenced IPPC Facilities	0.0	0.0	Não anotado
Licenced Waste Facilities	1.0	0.75	Casamento de strings
Linear Mine Features	1.0	1.0	Casamento de strings
Linear Waste Features	1.0	0.25	Casamento de strings
Medium Towns	1.0	1.0	Casamento de strings
Mine District	1.0	1.0	Casamento de strings
Mine Site Boundaries	1.0	0.66	Casamento de strings

*Continua na próxima página*

<b>Título da Camada</b>	<b>Precisão</b>	<b>Cobertura</b>	<b>Tipo de Anotação</b>
Mine Site Location	1.0	1.0	Casamento de strings
Other Waste Areas	1.0	1.0	Casamento de strings
Point Mine Features	1.0	1.0	Casamento de strings
Point Waste Features	1.0	0.25	Casamento de strings
Radiation Monitoring	1.0	0.2	Casamento de strings
Radon Map	1.0	1.0	Casamento de strings
River Basin District	1.0	1.0	Casamento de strings
River Catchments	1.0	0.5	Casamento de strings
River Sub Basins	1.0	1.0	Casamento de strings
SAC CO Habitats	1.0	0.16	Casamento de strings
SAC CO Species	1.0	1.0	Casamento de strings
SPA CO Habitats	1.0	0.2	Casamento de strings
Site Features	1.0	1.0	Casamento de strings
Small Towns	1.0	0.33	Casamento de strings
Solid Waste Analyses	0.72	0.66	LOD
Solid Waste Heaps	1.0	0.25	Casamento de strings
Stream Sediment Analyses	0.31	0.66	LOD
Subsoils	1.0	1.0	Casamento de strings
Transitional waterbodies	1.0	0.25	Casamento de strings
Water Analyses (Summer)	0.46	0.5	Casamento de strings
Water Analyses (Winter)	0.78	0.85	Casamento de strings
Water Areas	1.0	1.0	Casamento de strings
Water Management Units	1.0	0.5	Casamento de strings
Zone of Subsidence	1.0	0.25	Casamento de strings
Napier City Council Election Wards	0.25	0.28	LOD
NapierCityCouncil1mContours	0.0	0.0	Não anotado
NapierCityCouncilAddressPoints	1.0	0.43	Casamento de strings
NapierCityCouncilCadastralParcels	1.0	0.43	Casamento de strings
NapierCityCouncilCarParks	1.0	0.5	Casamento de strings
NapierCityCouncilDogExerciseAreas	1.0	0.27	Casamento de strings

*Continua na próxima página*

<b>Título da Camada</b>	<b>Precisão</b>	<b>Cobertura</b>	<b>Tipo de Anotação</b>
NapierCityCouncilHeritageBuildingDatabase	0.14	0.33	LOD
NapierCityCouncilLeasedParking	1.0	0.33	Casamento de strings
NapierCityCouncilLiquorBan	1.0	0.2	Casamento de strings
NapierCityCouncilMobilityParks	1.0	0.375	Casamento de strings
NapierCityCouncilMotorcycleParks	1.0	0.5	Casamento de strings
NapierCityCouncilOnStreetParks	1.0	0.5	Casamento de strings
NapierCityCouncilOperativePlanZones	1.0	0.33	Casamento de strings
NapierCityCouncilPayandDisplayCarparks	1.0	0.54	Casamento de strings
NapierCityCouncilSubdivisionApplications	1.0	0.5	Casamento de strings
NapierCityCouncilTitleParcels	1.0	0.69	Casamento de strings
ne_50m_boundary_lines <sub>land</sub>	1.0	0.5	Casamento de strings
ne_50m_coastline	1.0	1.0	Casamento de strings

# Apêndice D

## Resultado da Anotação no Nível de Atributos

Os resultados do desempenho da abordagem proposta para **anotação semântica** no nível de atributos estão listados na Tabela D.1. Cada linha desta tabela contempla, respectivamente, o nome do atributo, a sua métrica de precisão e a sua métrica de cobertura.

Tabela D.1: Resultado da anotação semântica da linha de base dos atributos

Nome do Atributo	Precisão	Cobertura
ACTION	1.0	1.0
ADDRESS	1.0	1.0
ADVICE	1.0	1.0
AGENT	1.0	1.0
AIRZONE	1.0	0.16
APPLICANT	1.0	1.0
APPLICATION_NUMBER	1.0	0.5
AQUATIC_TERRESTRIAL_WETLAND	1.0	1.0
AREAHECTARE	1.0	0.5
AUDIT_YEAR	1.0	1.0
BEACHPHOTO	1.0	1.0
BOUNDARYSOURCE	1.0	0.5
BUILDCOUNT	1.0	1.0

*Continua na próxima página*

<b>Nome do Atributo</b>	<b>Precisão</b>	<b>Cobertura</b>
BYLAW	1.0	1.0
CAPACITY	1.0	1.0
CASE_OFFICER	1.0	1.0
CATCHMENT	1.0	1.0
CATEGORY	1.0	1.0
CHANGE	0.0	0.0
CLASSIFICATION	1.0	1.0
CODE	1.0	1.0
COMMENT	1.0	1.0
COMMODITY_PRODUCED	1.0	1.0
CONFIRMATION_DATE	1.0	0.25
COST	1.0	1.0
COUNTY	1.0	1.0
COVERAGE	1.0	1.0
DATE_DATA_REFRESHED	1.0	0.5
DEPTH	1.0	1.0
DESCRIPTION	1.0	1.0
DISPLAY	1.0	1.0
EDENMONENTITY	1.0	0.33
FEATURE	1.0	1.0
FINISH_DAT	1.0	1.0
FURTHER_INFO_PUBLIC	1.0	1.0
GROUNDWATER_DEPENDENT	1.0	0.33
HABITAT_SPECIES	1.0	0.33
HIGH_SCHOOL_CLUSTER	1.0	0.16
HOUSING_MARKET_AREA	1.0	0.6
HYPERLINK	1.0	1.0
INSPECTION_MONTH	1.0	1.0
INTERIOR	1.0	1.0
LABEL	1.0	1.0

*Continua na próxima página*

<b>Nome do Atributo</b>	<b>Precisão</b>	<b>Cobertura</b>
LATITUDE	1.0	1.0
LICENCESTATUSDESCRIPTION	1.0	1.0
LOCALPLAN_REFERENCE	1.0	1.0
LOCATION	1.0	1.0
MARINEWATER_DEPENDENT	1.0	0.33
NATIONAL_PARK	1.0	1.0
NUMBER_OF_TURBINES	1.0	1.0
OFFICIAL_GRID_REFERENCE	1.0	0.2
OPERATOR	1.0	1.0
PERIMETER	1.0	1.0
POLICY_TYPE	1.0	0.5
PREVIOUS_MINERAL_WORKINGS	1.0	0.5
PROCESSINGSTATUS	1.0	1.0
PROPOSAL	1.0	1.0
PUBLIC_FURTHER_INFO	1.0	1.0
RAIL_ROAD_VISIBILITY	1.0	0.16
RECOMMENDED_USAGE	1.0	1.0
RECORD_STATUS	1.0	0.5
REFERENCE	1.0	1.0
REGION	1.0	1.0
REMARK	1.0	1.0
REVIEW_DATE	1.0	0.33
RINGS_NOK	1.0	1.0
SALINITY	1.0	1.0
SCALE	1.0	1.0
SCHOOL_NAME	1.0	0.86
SETTLEMENT	0.0	0.0
SHAPE	1.0	1.0
SITE_CAPACITY	1.0	0.83
SITE_CHARACTERISTICS	1.0	1.0

*Continua na próxima página*

<b>Nome do Atributo</b>	<b>Precisão</b>	<b>Cobertura</b>
SIZE	1.0	1.0
SOIL_GROUP	1.0	1.0
SOURCE	1.0	1.0
SPECIES	1.0	1.0
START_DATE	1.0	0.25
STATIONPOSITION	1.0	0.5
STATUS	1.0	1.0
STRUCTURE	1.0	1.0
STYLE	1.0	1.0
SUBCLASS	1.0	1.0
SUBSITEPREFIX	1.0	1.0
SURFACEWATER_DEPENDENT	1.0	0.33
SURROUNDING_USE_1	1.0	1.0
SURVEY	1.0	1.0
SYSTEM	1.0	1.0
TIME_LIMIT	1.0	1.0
TREE_NUMBER	1.0	0.5
TYPE	1.0	1.0
URL	1.0	1.0
VERSION	1.0	1.0
VOLUME	1.0	1.0
WARD	1.0	1.0
WATER_DEPENDANCY	1.0	1.0
WATERTYPE	1.0	0.5
WISEREFERENCEDATA	1.0	1.0
YEAR_REVIEWED	1.0	1.0
ZONE_DOCS	1.0	1.0

# Apêndice E

## URL dos Serviços da Amostra

As URL dos trinta e um serviços que compuseram a amostra utilizada na avaliação experimental estão listadas a seguir:

Tabela E.1: URL dos serviços da amostra

<a href="http://ogc.bgs.ac.uk/digmap625k_gsm131_gp/wfs">http://ogc.bgs.ac.uk/digmap625k_gsm131_gp/wfs</a>
<a href="http://arcgisweb.fife.gov.uk/geoserver/fife/wfs?version=1.1.0&amp;request=Getcapabilities">http://arcgisweb.fife.gov.uk/geoserver/fife/wfs?version=1.1.0&amp;request=Getcapabilities</a>
<a href="http://sedsh127.sedsh.gov.uk/arcgis/services/HS/World_Heritage_Sites/MapServer/WFSServer?request=GetCapabilities">http://sedsh127.sedsh.gov.uk/arcgis/services/HS/World_Heritage_Sites/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://sedsh127.sedsh.gov.uk/arcgis/services/HS/Conservation_Areas/MapServer/WFSServer?request=GetCapabilities">http://sedsh127.sedsh.gov.uk/arcgis/services/HS/Conservation_Areas/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://sedsh127.sedsh.gov.uk/arcgis/services/HS/Gardens_and_Designed_Landscapes/MapServer/WFSServer?request=GetCapabilities">http://sedsh127.sedsh.gov.uk/arcgis/services/HS/Gardens_and_Designed_Landscapes/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://sedsh127.sedsh.gov.uk/arcgis/services/HS/Scheduled_Monuments/MapServer/WFSServer?request=GetCapabilities">http://sedsh127.sedsh.gov.uk/arcgis/services/HS/Scheduled_Monuments/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://sedsh127.sedsh.gov.uk/arcgis/services/HS/Historic_Marine_Protected_Areas/MapServer/WFSServer?request=GetCapabilities">http://sedsh127.sedsh.gov.uk/arcgis/services/HS/Historic_Marine_Protected_Areas/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://sedsh127.sedsh.gov.uk/arcgis/services/HS/Historic_Scotland/MapServer/WFSServer?request=GetCapabilities">http://sedsh127.sedsh.gov.uk/arcgis/services/HS/Historic_Scotland/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://sedsh127.sedsh.gov.uk/arcgis/services/HS/Listed_Buildings/MapServer/WFSServer?request=GetCapabilities">http://sedsh127.sedsh.gov.uk/arcgis/services/HS/Listed_Buildings/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://gis.epa.ie/geoserver/wfs">http://gis.epa.ie/geoserver/wfs</a>
<a href="http://www.gis.napier.govt.nz/geoserver/wfs/kml_reflect?layers=gn:NapierArtDeco_point">http://www.gis.napier.govt.nz/geoserver/wfs/kml_reflect?layers=gn:NapierArtDeco_point</a>
<a href="http://sims-wfs.niwa.co.nz/cgi-bin/sims-wfs">http://sims-wfs.niwa.co.nz/cgi-bin/sims-wfs</a>
<a href="http://www.sitges.cat/absAtlasWEB2/absopenGIS/wfs.ashx">http://www.sitges.cat/absAtlasWEB2/absopenGIS/wfs.ashx</a>
<a href="http://ndgishub.nd.gov/arcgis/services/Features/Locations/MapServer/WFSServer">http://ndgishub.nd.gov/arcgis/services/Features/Locations/MapServer/WFSServer</a>
<a href="http://mgswb2.mngs.umn.edu/ArcGIS/services/MNBoreholeLithIntervals/MapServer/WFSServer?request=GetCapabilities">http://mgswb2.mngs.umn.edu/ArcGIS/services/MNBoreholeLithIntervals/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://services.azgs.az.gov/ArcGIS/services/aasggeothermal/ORTermalSprings/MapServer/WFSServer?request=GetCapabilities">http://services.azgs.az.gov/ArcGIS/services/aasggeothermal/ORTermalSprings/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://kgs.uky.edu/arcgis/services/aasggeothermal/ALBoreholeLithIntervals/MapServer/WFSServer?request=GetCapabilities">http://kgs.uky.edu/arcgis/services/aasggeothermal/ALBoreholeLithIntervals/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://kgs.uky.edu/arcgis/services/aasggeothermal/ARWellLogs/MapServer/WFSServer?request=GetCapabilities">http://kgs.uky.edu/arcgis/services/aasggeothermal/ARWellLogs/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://kgs.uky.edu/arcgis/services/aasggeothermal/KYWellHeaders/MapServer/WFSServer?request=GetCapabilities">http://kgs.uky.edu/arcgis/services/aasggeothermal/KYWellHeaders/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://kgs.uky.edu/arcgis/services/aasggeothermal/ALSeismicHypocenters/MapServer/WFSServer?request=GetCapabilities">http://kgs.uky.edu/arcgis/services/aasggeothermal/ALSeismicHypocenters/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://kgs.uky.edu/arcgis/services/aasggeothermal/ALaqWellChemistry/MapServer/WFSServer?request=GetCapabilities">http://kgs.uky.edu/arcgis/services/aasggeothermal/ALaqWellChemistry/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://geothermal.isgs.illinois.edu/ArcGIS/services/aasggeothermal/MAThermalConductivity/MapServer/WFSServer?request=GetCapabilities">http://geothermal.isgs.illinois.edu/ArcGIS/services/aasggeothermal/MAThermalConductivity/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://kgs.uky.edu/arcgis/services/aasggeothermal/ARBoreholeTemperatures/MapServer/WFSServer?request=GetCapabilities">http://kgs.uky.edu/arcgis/services/aasggeothermal/ARBoreholeTemperatures/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://kgs.uky.edu/arcgis/services/aasggeothermal/GABoreholeTemperatures/MapServer/WFSServer?request=GetCapabilities">http://kgs.uky.edu/arcgis/services/aasggeothermal/GABoreholeTemperatures/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://services.azgs.az.gov/ArcGIS/services/aasggeothermal/AZWellHeaders/MapServer/WFSServer?request=GetCapabilities">http://services.azgs.az.gov/ArcGIS/services/aasggeothermal/AZWellHeaders/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://mgswb2.mngs.umn.edu/ArcGIS/services/MNWellheaders/MapServer/WFSServer?request=GetCapabilities">http://mgswb2.mngs.umn.edu/ArcGIS/services/MNWellheaders/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://kgs.uky.edu/arcgis/services/aasggeothermal/MDWellHeaders/MapServer/WFSServer?request=GetCapabilities">http://kgs.uky.edu/arcgis/services/aasggeothermal/MDWellHeaders/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://geothermal.isgs.illinois.edu/ArcGIS/services/aasggeothermal/MABoreholeTemperatures/MapServer/WFSServer?request=GetCapabilities">http://geothermal.isgs.illinois.edu/ArcGIS/services/aasggeothermal/MABoreholeTemperatures/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://kgs.uky.edu/arcgis/services/aasggeothermal/VAHeatFlow/MapServer/WFSServer?request=GetCapabilities">http://kgs.uky.edu/arcgis/services/aasggeothermal/VAHeatFlow/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://kgs.uky.edu/arcgis/services/aasggeothermal/VAThermalConductivity/MapServer/WFSServer?request=GetCapabilities">http://kgs.uky.edu/arcgis/services/aasggeothermal/VAThermalConductivity/MapServer/WFSServer?request=GetCapabilities</a>
<a href="http://kgs.uky.edu/arcgis/services/aasggeothermal/FLWellLogs/MapServer/WFSServer?request=GetCapabilities">http://kgs.uky.edu/arcgis/services/aasggeothermal/FLWellLogs/MapServer/WFSServer?request=GetCapabilities</a>