

2 Dicionários e catálogos

A crescente disseminação de informações geográficas digitais como, por exemplo, cartas geográficas e imagens de satélite, aliada à crescente oferta de sistemas de informações geográficas, requerem métodos para melhor entendimento de dados geográficos disponíveis.

Dados geográficos digitais têm o propósito de representar o mundo real para análise por computador. São representações da realidade onde aproximações, simplificações e omissões podem ocorrer. Para garantir que um dado não será erroneamente utilizado as suposições e limitações envolvendo sua produção devem ser completamente documentadas. Metadados permitem que os produtores dos dados os descrevam apropriadamente, de modo que os usuários possam avaliar mais precisamente sua aplicabilidade e uso.

Dados geográficos são, normalmente, utilizados por terceiros, ou seja, pessoas, organizações, etc. Documentação apropriada permite, àqueles não familiarizados com os dados, utilizá-los corretamente.

Metadados, por definição, são dados sobre dados, isto é, dados que descrevem outros dados. Fornecem informações sobre o conteúdo de um dado, informações históricas, modo de obtenção do dado, sua qualidade, etc.

Iremos assumir que dados geográficos terão ao menos uma escala e uma descrição da área da superfície terrestre coberta por ele. Essa descrição é frequentemente um retângulo ou um paralelogramo cujos vértices são definidos por coordenadas em um dado sistema de referência (CRS). A escala, a descrição da área coberta e o CRS utilizado são tratados como metadados dos dados.

Catálogos de metadados geográficos contém metadados de dados geográficos [8]. São a estratégia mais simples de compartilhamento de dados geográficos entre aplicações. Tais catálogos oferecem serviços de consulta e

manipulação de metadados, bem como, serviços de intermediação de recuperação de dados geográficos. Catálogos não armazenam o dado geográfico em si, mas possuem uma referência para o seu local de armazenamento. Assim, na recuperação de um dado, o serviço é redirecionado para o sistema de armazenamento do dado que o localiza e o retorna ao cliente.

Para um compartilhamento efetivo de dados, não é suficiente a disponibilidade de descrições detalhadas. É necessário que o esquema de metadados e as interfaces de consulta e recuperação sejam internacionalmente padronizadas.

Duas autoridades propõem esquemas de metadados para dados geográficos: *US Federal Geographic Data Committee (FGDC)*¹ e o *International Organization of Standards (ISO)*².

O FGDC é um comitê dos EUA que coordena o compartilhamento de dados geográficos através do portal Geodata³. Utiliza dados da *National Spatial Data Infrastructure (NSDI) Clearing Network*. Essa rede é uma coleção de catálogos de metadados conhecida como *Clearinghouse Nodes*. Cada nó é hospedado por uma organização diferente e contém metadados que descrevem dados sob sua área de responsabilidade. O FDGC está, atualmente, conduzindo o desenvolvimento de uma adaptação para os EUA do padrão ISO 19115:2003 que veremos a seguir.

A ISO é uma organização internacional, formada por representantes de diversas empresas e governos, para estabelecer padrões internacionais. Seu trabalho é, normalmente, conduzido por comitês técnicos. Cada membro da organização interessado em algum assunto para o qual foi estabelecido um comitê tem o direito de se fazer representar nele. Organizações internacionais, governamentais e não governamentais, em colaboração com a ISO, também tomam parte no trabalho.

¹ <http://www.fgdc.gov>

² <http://www.iso.org>

³ <http://www.geodata.com>

O comitê técnico *ISO/TC211, Geographic Information/ Geomatic* foi constituído para especificar um esquema de metadados para dados geográficos. Esse esquema foi apresentado na norma ISO 19115:2003.

O *Open Geospatial Consortium (OGC)*⁴ é uma organização não lucrativa, internacional, que está conduzindo o desenvolvimento de padrões para serviços geo-espaciais e serviços baseados em localização. A OGC atua com governos, indústria e universidades para criar interfaces de programação abertas e extensíveis para Sistemas de Informações Geográficas (GIS).

O OGC desenvolveu uma especificação, atualmente na versão 2.0, para a interface de serviços de catálogo que suporta a publicação e consulta de metadados para dados e serviços geográficos, bem como a localização e recuperação de recipientes de dados geográficos selecionados.

2.1. Padrões de metadados

2.1.1. Padrão ISO 19115

A *ISO 19115:2003* propõe um padrão de metadados para dados e serviços geográficos digitais. A norma define elementos de metadados (identificação, extensão, qualidade, esquema temporal e espacial, referência espacial e distribuição do dado) e fornece um esquema e uma terminologia de metadados comum. Ela faz referência a outras normas da série 191xx que devem ser consultadas para esclarecimentos mais específicos.

Estruturas de dados básicas organizam os dados e seus relacionamentos com metadados. A Figura 2 apresenta o esquema dessas estruturas.

Dados geográficos (*DS_Dataset*) são coleções de dados identificáveis. Podem representar abstrações menores do mundo real que, limitadas por alguma restrição como extensão espacial ou tipo de objeto, estão fisicamente localizados dentro de outra representação maior do mundo real ou outro dado.

⁴ <http://www.opengeospatial.com>

Dados podem representar tanto um único objeto geográfico, quanto um mapa, imagem geográfica, ou qualquer outro dado geográfico.

Uma agregação de dados é uma abstração para agrupar conjuntos de dados geográficos relacionados. Pode ser especificada como um agrupamento de uso geral (*DS_OtherAggregate*), uma série de dados (*DS_Series*)⁵ ou como uma atividade especial (*DS_Initiative*). Por sua vez, uma série de dados pode ser do tipo plataforma de sensoriamento (*DS_Platform*), sensor (*DS_Sensor*) ou uma produção em série (*DS_ProductionSeries*). Uma agregação pode agrupar outras agregações e, assim formar uma hierarquia de dados.

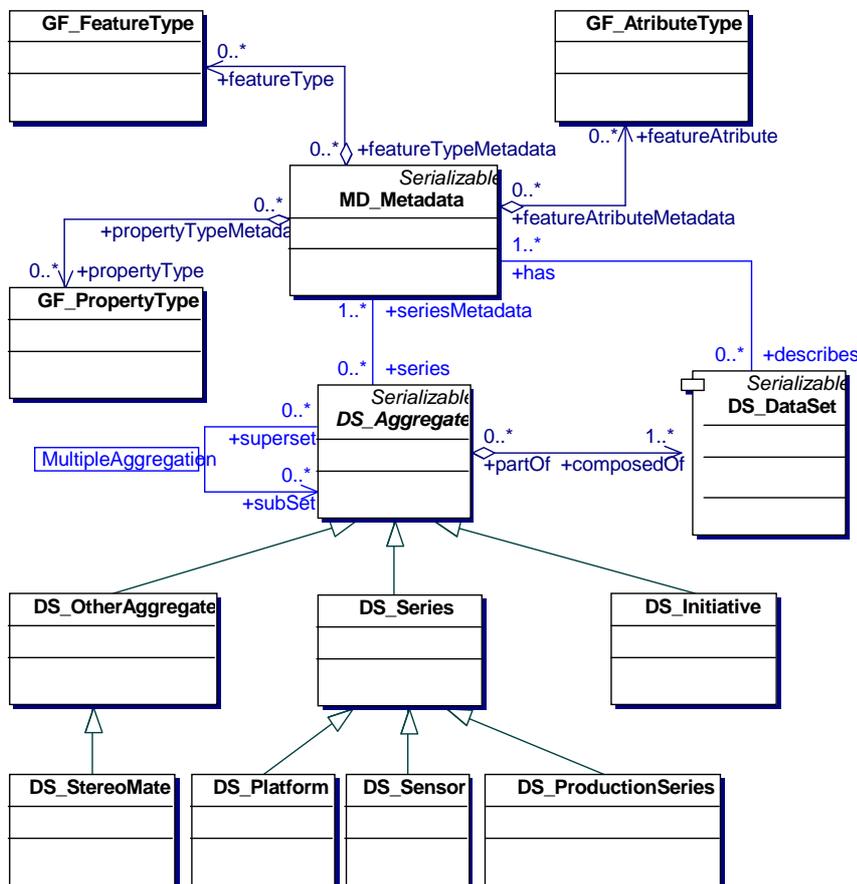


Figura 2 – Esquema de organização de dados geográficos definidos pela ISO 19115

Os metadados são normalmente associados ao conjunto de dados, que deve possuir pelo menos uma estrutura de metadados. Ao tratar-se de séries de

⁵ Um conjunto de dados do tipo série é uma coleção de dados que compartilham a mesma especificação de produto.

dados a norma propõe que a agregação dos dados possua pelo menos uma estrutura de metadados.

Os metadados estão distribuídos por várias estruturas de dados que organizam as informações segundo sua natureza. O esquema geral dessas estruturas pode ser visto na Figura 3. O conjunto principal de metadados (*MD_Metadata*) pode ser complementado com metadados sobre a representação espacial (*MD_SpatialRepresentation*) do dado geográfico, sistemas de geo-referenciamento (*MD_ReferenceSystemInformation*), qualidade do dado (*MD_DataQuality*), identificação do dado (*MD_Identification*), etc.

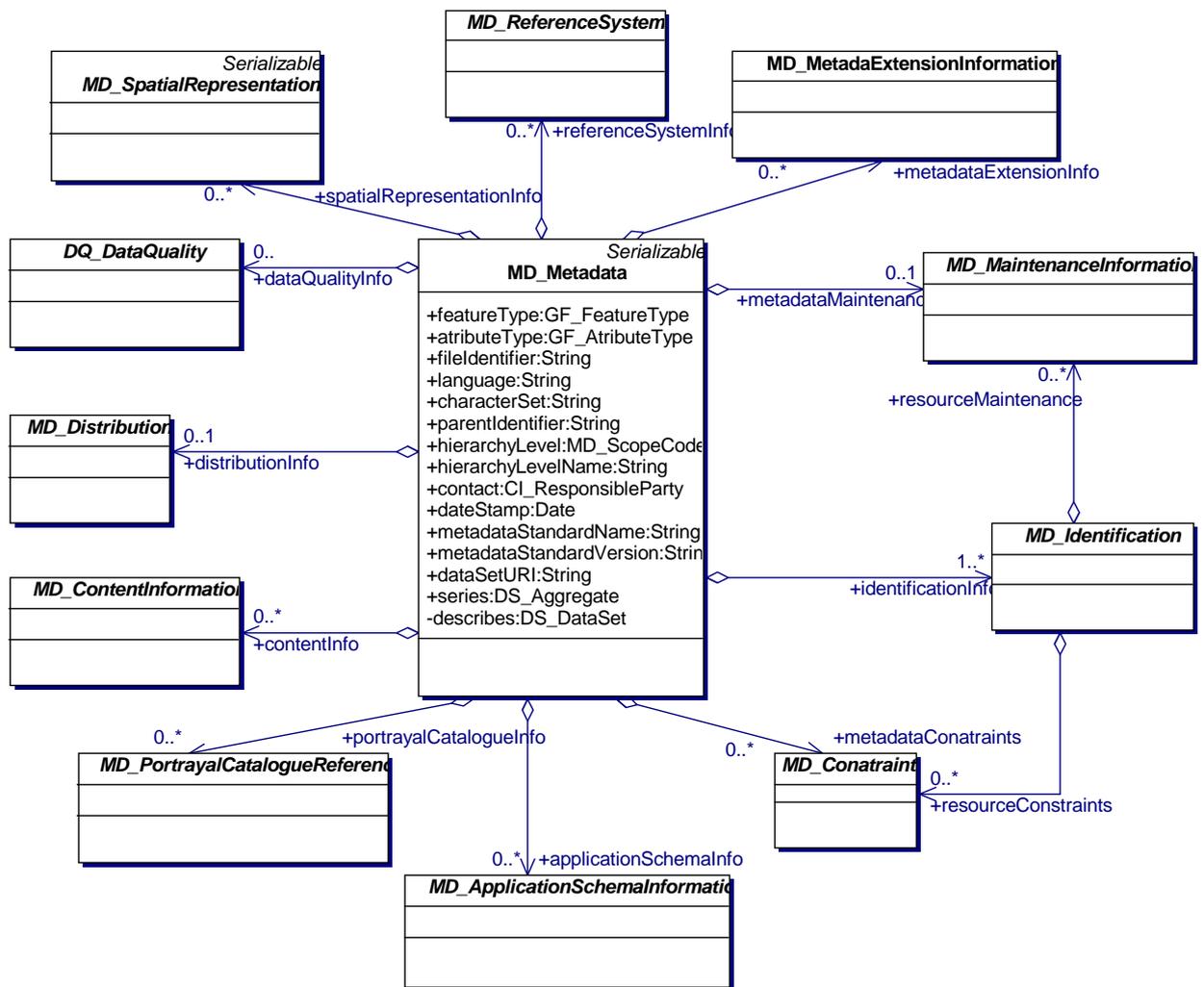


Figura 3 – Esquema geral de metadados geográficos definidos pela ISO19115

Existe um conjunto extenso de elementos de metadados, do qual somente um subconjunto é, normalmente, utilizado. Entretanto, é essencial que um conjunto básico, mínimo, de elementos de metadados seja utilizado para

descrever um conjunto de dados. A Tabela 2 lista os elementos de metadados principais para identificar dados geográficos, tipicamente para propósitos de catalogação. Os elementos com o símbolo (M) são definidos como obrigatórios, aqueles com (O) são opcionais e aqueles com (C) são obrigatórios em algumas circunstâncias. Essa lista contém elementos para responder às seguintes questões: “Existe dado sobre um tópico específico (‘O que’)?”, “Sobre algum lugar (‘Onde’)?”, “Em alguma data ou período (‘Quando’)?” e “Qual o ponto de contato para obter mais informações sobre o dado ou obter cópia (‘Quem’)?”.

Elemento de metadado	Descrição
MD_Metadata>MD_DataIdentification.citation>CI_Citation.title	Dataset title (M)
MD_Metadata>MD_DataIdentification.spatialRepresentationType	Spatial representation type (O)
MD_Metadata>MD_DataIdentification.citation>CI_Citation.date	Dataset reference date (M)
MD_Metadata>MD_ReferenceSystem	Reference system (O)
MD_Metadata>MD_DataIdentification.pointOfContact>CI_ResponsibleParty	Dataset responsible party (O)
MD_Metadata>DQ_DataQuality.lineage>LI_Lineage	Lineage (O)
MD_Metadata>MD_DataIdentification.extent>EX_Extent>EX_GeographicExtent>EX_GeographicBoundingBox or EX_GeographicDescription	Geographic location of the dataset (by four coordinates or by geographic identifier) (C)
MD_Metadata>MD_Distribution>MD_DigitalTransferOption.onLine>CI_OnlineResource	On-line resource (O)
MD_Metadata>MD_DataIdentification.language	Dataset language (M)
MD_Metadata.fileIdentifier	Metadata file identifier (O)
MD_Metadata>MD_DataIdentification.characterSet	Dataset character set (C)

Tabela 1 – Conjunto de metadados principal definidos pela ISO 19115

Elemento de metadado	Descrição
MD_Metadata.metadataStandardName	Metadata standard name (O)
MD_Metadata>MD_DataIdentification.topicCategory	Dataset topic category (M)
MD_Metadata.metadataStandardVersion	Metadata standard version (O)
MD_Metadata>MD_DataIdentification.spatialResolution>MD_Resolution.equivalentScale or MD_Resolution.distance	Spatial resolution of the dataset (O)
MD_Metadata.language	Metadata language (C)
MD_Metadata>MD_DataIdentification.abstract	Abstract describing the dataset (M)
MD_Metadata.characterSet	Metadata character set (C)
MD_Metadata>MD_Distribution>MD_Format.nameandMD_Format.version	Distribution format (O)
MD_Metadata.contact>CI_ResponsibleParty	Metadata point of contact (M)
MD_Metadata>MD_DataIdentification.extent>EX_Extent>EX_TemporalExtent or EX_VerticalExtent	Additional extent information for the dataset (vertical and temporal) (O)
MD_Metadata.dateStamp	Metadata date stamp (M)

Tabela 2 – Conjunto de metadados principal definidos pela ISO 19115 (cont.)

A utilização de itens opcionais, além dos itens obrigatórios aumentará a interoperabilidade entre aplicações, permitindo aos usuários compreender os dados geográficos e os metadados relacionados livres de ambigüidades. Versões customizadas do conjunto de metadados devem incluir os últimos.

2.1.2. Padrão FGDC

O *Content Standard for Digital Metadata (CSDGM)* (FGDC 1998) estabelece um conjunto de terminologias e definições para a documentação de dados geográficos digitais, incluindo elementos de metadados para os seguintes tópicos:

- Identificação: nome, desenvolvedor, área geográfica coberta, temas de informação incluídos, restrições de acesso.
- Qualidade de dado: precisão de posição e de atributos, integridade, consistência, procedência.

- Organização espacial do dado: modelo de dados espacial utilizado para codificar o dado geográfico, número de objetos geográficos, métodos além de coordenadas, utilizados para codificar localizações, tais como ruas e endereços.
- Referências espaciais: codificação das localizações das coordenadas, projeção de mapa ou sistema grade utilizado, parâmetros para converter o dado para outro sistema de coordenadas.
- Informação de entidade e atributos: informação geográfica (estradas, casas, elevação, temperatura, etc.)
- Distribuição: agência de distribuição, formatos e mídias disponíveis, preço, distribuição on-line.
- Referência dos metadados: *timestamp* e agência responsável pela compilação dos metadados.

A **Erro! Fonte de referência não encontrada.** e a Tabela 5 mostram os elementos de metadados obrigatórios no padrão FGDC.

XML Tag Elemento de Metadado	Descrição
Data Originator /metadata/idinfo/citation/citeinfo/origin/	O nome de uma organização ou de um indivíduo que gerou o dado
Data Title /metadata/idinfo/citation/citeinfo/title/	O nome pelo qual o dado é conhecido
Abstract /metadata/idinfo/descript/abstract/	Um a breve resumo sobre o dado
Progress /metadata/idinfo/status/progress	O estado do dado
West Bounding Coordinate /metadata/idinfo/spdom/bounding/westbc/	Coordenada mais a oeste da área coberta pelo dado expressa em longitude
East Bounding Coordinate /metadata/idinfo/spdom/bounding/eastbc/	Coordenada mais a leste da área coberta pelo dado expressa em longitude
North Bounding Coordinate /metadata/idinfo/spdom/bounding/northbc/	Coordenada mais ao norte da área coberta pelo dado expressa em latitude

Tabela 3 – Elementos de metadado obrigatórios no padrão FGDC

XML Tag Elemento de Metadado	Descrição
South Bounding Coordinate /metadata/idinfo/spdom/bounding/southbc/	Coordenada mais ao sul da área coberta pelo dado expressa em latitude
Theme Keyword /metadata/idinfo/keywords/theme/themekey/	Palavra ou frase mais frequentemente utilizada para descrever o conteúdo do dado
Metadata Contact Organization /metadata/metainfo/metac/ctinfo/ctorgp/ctorg/	O nome da organização para a qual o tipo de contato se aplica
Metadata Contact Person /metadata/metainfo/metac/ctinfo/ctperp/ctper/	O nome do indivíduo para o qual o tipo de contato se aplica
Metadata Contact Address City	A cidade do endereço
Metadata Contact Address State or Province /metadata/metainfo/metac/ctinfo/ctaddr/state/	O estado ou provincial do endereço
Metadata Contact Address Postal Code /metadata/metainfo/metac/ctinfo/ctaddr/postal/	Código postal do endereço
Publication Date /metadata/idinfo/citation/citeinfo/pubdate/	A data de liberação para publicação ou data de publicação
Purpose /metadata/idinfo/descript/purpose/	Um resumo do propósito de geração do dado
Time Period of Content: Single Date /metadata/idinfo/timeperd/timeinfo/sngdate/caldate	O período de tempo pelo qual o dado corresponde à referência
Time Period of Content: Range of Dates, Beginning Date /metadata/idinfo/timeperd/timeinfo/rngdates/begdate/	
Time Period of Content: Range of Dates, Ending Date /metadata/idinfo/timeperd/timeinfo/rngdates/enddate/	
Currentness Reference /metadata/idinfo/timeperd/current/	A base de tempo sob a qual o período de tempo é determinado
Maintenance and Update Frequency /metadata/idinfo/status/update	A frequência com a qual alterações e adições são realizadas sobre o dado após sua primeira produção

Tabela 4 – Elementos de metadado obrigatórios no padrão FGDC (cont.)

XML Tag Elemento de Metadado	Descrição
Theme Keyword Thesaurus /metadata/idinfo/keywords/theme/themekt/	Referência a um tesouro formalmente registrado ou a uma autoridade similar
Access Constraints /metadata/idinfo/acconst/	Restrições e pré-requisitos para acesso ao dado
Use Constraints /metadata/idinfo/useconst/	Restrições e pré-requisitos legais para utilizar o dado
Metadata Contact Address Type /metadata/metainfo/metc/cntinfo/cntaddr/addrtype/	Tipo de endereço
Metadata Contact Phone number /metadata/metainfo/metc/cntinfo/cntvoice/	O número de telefone do contato
Metadata Date	Data de criação ou última atualização do dado

Tabela 5 – Elementos de metadado obrigatórios no padrão FGDC (cont.)

2.2. Interface de serviço de catálogos

2.2.1. OpenGIS Services Framework

As especificações OGC seguem uma arquitetura de referência, chamada *OpenGIS Services Framework (OSF)* [19], que identifica serviços, interface e protocolos de troca de dados que as aplicações podem utilizar. Resumidamente o OSF consiste no seguinte (Figura 4):

- Padrões de codificação: padrões que especificam intercâmbio e armazenamento de dados geográficos, incluindo sistemas de referenciamento espacial e topologia e geometria de objetos geográficos. Incluem a *Geography Markup Language (GML)*.
- Interface de serviços: componentes que, no lado cliente, interagem com usuários e no lado servidor, interagem com a aplicação de catálogo e servidores de dados.

- Serviços de registro: componentes que oferecem mecanismos para classificar, registrar, descrever, procurar, manter e acessar dados sobre recursos disponíveis. Incluem *Web Registry Service (WRS)*.
- Serviços de visualização: componentes que oferecem suporte específico para visualização de dados geográficos na forma de mapas, perspectivas tridimensionais de terreno, imagens anotadas, etc. Incluem o *Web Map Service (WMS)*, o *Coverage Portrayal Service (CPS)* e o *Style Management Service (SMS)*.
- Serviços de dados: componentes que fornecem serviços básicos de acesso a dados geográficos. Incluem o *Web Object Service (WOS)*, o *Web Feature Service (WFS)*, o *Sensor Collection Service (SCS)*, o *Image Archive Service (IAS)* e o *Web Coverage Service (WCS)*.
- Serviços de processamento: componentes que operam sobre dados geográficos e fornecem serviços especiais às aplicações. Incluem serviços de dicionários geográficos.

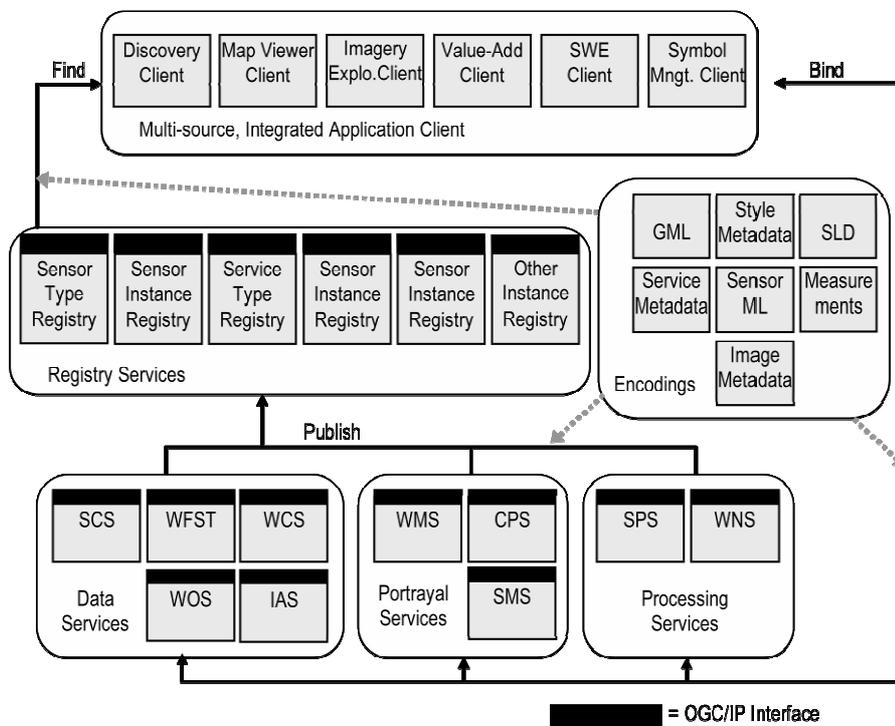


Figura 4 – Componentes do *OpenGIS Services Framework*

2.2.2. Geospatial Web Services Architecture

O OSF foi desenvolvido independentemente da arquitetura de protocolo para *Web services* do W3C. Somente recentemente, uma série de propostas de especificações coletivamente chamada de *OpenGIS Web Service 2 Initiative*, definiram interfaces que seguiam os padrões W3C. A OGC também iniciou uma experiência para testar o conceito de Web Geo-espacial Semântica.

A arquitetura *OGC Web Services (OWS)* é uma arquitetura orientada a serviço, com todos os componentes fornecendo um ou mais serviços para outro serviço ou clientes. Essa arquitetura é baseada nos conceitos fundamentais de provedor e consumidor de serviço dentro de um sistema computacional distribuído. Trata, também, de interações entre componentes definindo requisições de serviço (*service requests*), respostas de serviços (*service responses*) e exceções ocorridas nas requisições (*service exceptions*).

De acordo com Whiteside [20], definimos:

- um serviço é uma parte distinta de funcionalidade que é fornecida por uma entidade através de interfaces;
- uma interface é um conjunto nomeado de operações que caracterizam o comportamento de uma entidade;
- uma operação é uma especificação de uma transformação ou consulta que um objeto pode ser solicitado a executar. A operação tem um nome e listas de dados de entrada e saída.

As *OGC Web service interfaces* utilizam padrões abertos e fornecem algumas poucas operações estáticas por serviço. Formatos de codificação de dados e linguagens padrões baseados em XML são utilizados em muitas transferências de dados entre clientes e servidores. Serviços fracamente organizados em quatro camadas, como mostra a Figura 5.

A camada de serviços de gerenciamento de informação contém adaptadores que encapsulam fontes de dados. Os serviços nessa camada normalmente incluem algum processamento sobre os dados recuperados. Por exemplo, WMS, WCS, e WFS podem executar transformação de coordenadas e

conversão de formatos. A camada de processamento de serviço contém serviços projetados para o processamento dos dados. A camada de serviços de aplicação contém serviços projetados para suportar clientes, especialmente “clientes magros” como Web browsers.

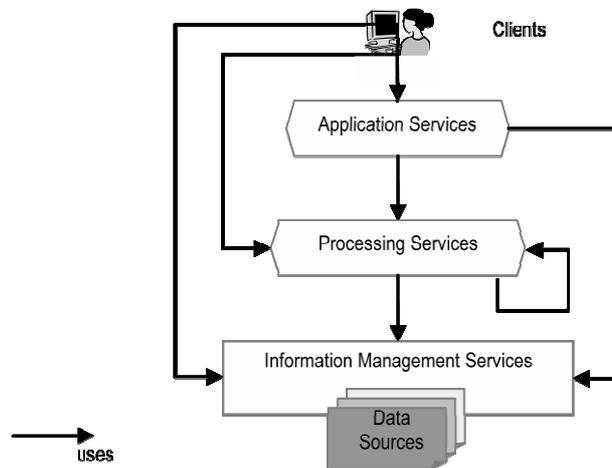


Figura 5 – Camadas de serviços na arquitetura OWS

2.2.3. OGC Catalogue Service 2.0

A especificação *OGC Catalogue Service (CS) 2.0* define interfaces que devem estar disponíveis em aplicações de catálogo de metadados sobre dados e serviços geográficos. Define operações para consulta e manutenção de metadados, localização e recuperação de recursos⁶ de dados e serviços e controle de sessão. Para auxiliar no processo de consulta dos metadados o CS traz, ainda, a especificação de uma linguagem de consulta de metadados baseada na sintaxe SQL. Toda a especificação é extensível de modo a acomodar heterogeneidade de ambientes e variedade de áreas de conhecimento.

A especificação OGC CS apresenta um conjunto completo de operações de catálogo. Um serviço de catálogo para ser compatível com a OGC CS 2.0 deve implementar ao menos um subconjunto das operações de interface. Assim

⁶ Um recurso de um dado ou serviço geográfico é o local onde está armazenado o dado geográfico ou o ponto de acesso ao serviço. No caso de dados geográficos, o recurso é o recipiente do dado.

um serviço de catálogo poderá implementar operações de consulta, outro poderá permitir consulta e atualização de metadados, outro ainda poderá implementar consultas em federações de catálogos e assim por diante.

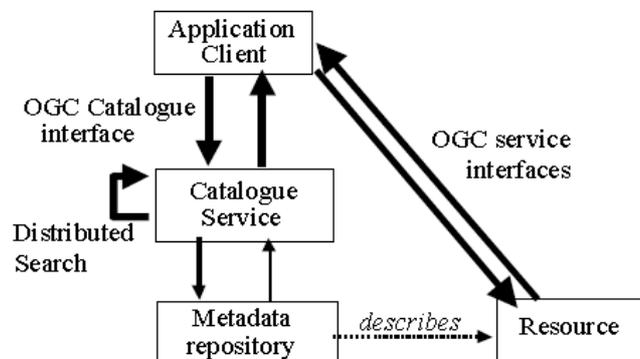


Figura 6 – Arquitetura interna de um serviço de catálogo

A Figura 6 ilustra a arquitetura interna do serviço de catálogo proposto. Além da interface, o esquema apresenta os componentes internos do serviço. A consulta ou a manutenção de metadados se faz por meio das interfaces de catálogo OGC. Nesse caso o serviço de catálogo irá utilizar seu próprio repositório de metadados ou desviará a operação para outro serviço de catálogo, como em uma federação de serviços. A recuperação de recursos de informação ocorre quando o recurso não está disponível ao cliente via rede e pode ser feito por meio de interfaces OGC ou proprietárias.

O conjunto completo das operações de um serviço de catálogo é:

Interface de serviço de catálogo

- *getCapabilities* → retorna metadados sobre o próprio serviço de catálogo.

Interface de consulta

- *query* → procura no conjunto de metadados disponível (local ou em outros serviços) aqueles que atendam à consulta definida pelo usuário e retorna ao cliente um conjunto contendo referências aos dados e serviços descritos pelos metadados selecionados. Opcionalmente, o conjunto resultado poderá conter também metadados selecionados na consulta.

- *present* → retorna metadados selecionados referentes a dados previamente selecionados por uma consulta.
- *describeRecordType* → retorna a definição de tipo para metadados.
- *getDomain* → retorna valores válidos para algum metadado.

Interface de transação

- *initialize* → inicia uma transação.
- *close* → encerra uma transação.
- *status* → retorna o estado atual da transação.
- *cancel* → termina uma transação previamente iniciada.

Interface de gerenciamento

- *transaction* → executa um conjunto especificado de operações de inclusão, alteração e atualização de metadados.
- *harvestRecords* → carrega no repositório de metadados um conjunto de metadados de uma origem especificada.

Interface de intermediação.

- *order* → retorna um dado ou endereço de serviço identificado registrado no catálogo mas não acessível ao usuário.

A especificação define, ainda, um conjunto de metadados mínimo como retorno de consultas para favorecer a interoperabilidade entre aplicações de diferentes características e ambientes. A Tabela 7 mostra a lista desses elementos utilizando a sintaxe Dublin Core (ISSO 15836).

Elemento	Descrição
dc:title	Um nome dado ao recurso
dc:creator	Uma entidade responsável por produzir o conteúdo do recurso
dc:subject	Um tópico do conteúdo do recurso É o lugar onde uma taxonomia de classificação poderia ser utilizada
dct:abstract	Um resumo do conteúdo do recurso
dc:publisher	Uma entidade responsável por tornar o recurso disponível

Tabela 6 – Conjunto de metadados mínimo

Elemento	Descrição
dc:contributor	Uma entidade responsável por fazer contribuições ao conteúdo do recurso
dc:date	Data de criação ou atualização
dc:type	Natureza do conteúdo do recurso
dc:format	Manifestação física ou digital do recurso
dc:identifier	Identificador não ambíguo em um contexto
dc:source	Referência a um recurso de origem
dc:language	Língua do conteúdo do recurso
dc:relation	Referência a recursos relacionados
dct:spatial	Extensão espacial do conteúdo do recurso
dc:rights	Direitos sobre o conteúdo do recurso

Tabela 7 – Conjunto de metadados mínimo (cont.)

2.3.

Dicionários geográficos

2.3.1.

Exemplos de dicionários geográficos

O sistema mais comum para identificação de objetos geográficos é a atribuição de um nome para cada objeto. Essa forma de referência geográfica é normalmente utilizada por muitos dicionários geográficos. Eles contêm listas de nomes de objetos geográficos, suas respectivas localizações geográficas e outras informações descritivas.

Muitos Atlas possuem uma seção de dicionário geográfico para auxiliar na localização das páginas que fazem referência a um nome específico. Dicionários geográficos disponíveis na Web têm sido desenvolvidos para atender às necessidade de sistemas de informação. Eles armazenam nome, tipo, coordenadas que representam o objeto, e uma área administrativa que contém o objeto para cada objeto. Eles também oferecem serviços para manipulação dos dados armazenados.

A seguir, são apresentadas descrições breves de alguns dicionários geográficos disponíveis na Web.

- *Getty Thesaurus of Geospatial Names (TGN)*⁷ → É um vocabulário estruturado de abrangência mundial com cerca de 1,3 milhões de nomes e 892.000 objetos geográfico desenvolvido pelo *Getty Research Institute*. Tem o objetivo de auxiliar a criação de catálogos de arte e literatura de arquitetura. Para cada objeto, o TGN possui um identificador único, nome preferido, um conjunto de classificações obtido do *Art Architecture Thesaurus (AAT)*, nomes alternativos, localização geográfica, entre outros.
- *U. S. Geographic Names Information System (GNIS)*⁸ → Desenvolvido pelo *U. S. Geological Survey (USGS)* em cooperação com o *U.S. Board on Geographic Names (BGN)*, contém informações sobre cerca de 2 milhões de objetos geográficos culturais e físicos, atuais e históricos dos Estados Unidos e de seus territórios. Alguns dos atributos mantidos são a localização geográfica, nome oficial, classificação segundo tipos definidos pelo BGN, entre outros.
- *GEOnet Names Server (GNS)*⁹ → Desenvolvido pela *National Geospatial-Intelligence Agency (NGA)*, possui cerca de 4 milhões de objetos geográficos com 5,5 milhões de nomes provenientes da NGA e BGN com abrangência mundial. Alguns atributos mantidos para cada objeto são identificador único, nome oficial, região, coordenadas geográficas classificação segundo o BGN, entre outros.
- *Alexandria Digital Library Gazetteer (ADL Gazetteer)*¹⁰ → Desenvolvido no contexto do projeto *Alexandria Digital Libray* da *Universidade da Califórnia em Santa Bárbara*, possui cerca de 5.9 milhões de nomes provenientes do GNIS, GNS e outros dicionários, classificados segundo o *ADL Feature Type Thesaurus (ADL FTT)*. Mantém para cada entrada um identificador único, nome oficial, nomes alternativos, classificação segundo o ADL FTT, coordenadas geográficas, descrição, entre outros.

⁷ http://www.getty.edu/research/conducting_research/vocabularies/tgn

⁸ <http://nhd.usgs.gov/gnis.html>

⁹ <http://164.214.2.53/gns/html>

¹⁰ <http://www.alexandria.ucsb.edu/gazetteer/>

2.3.2. Classificação de objetos geográficos

Um objeto é uma abstração de um fenômeno do mundo real e um objeto geográfico é um objeto associado a uma localização sobre a superfície terrestre [18]. Um objeto geográfico representa um lugar, tal como a cidade do Rio de Janeiro. No jargão familiar da Ciência da Computação possui um atributo especial que descreve sua posição na superfície da terra, usando um dado sistema de geo-referenciamento (*Coordinate Reference System CRS*). São normalmente classificados por tipos como, por exemplo, baías, rios, cidades, etc. que por sua vez são organizados de acordo com alguma característica como, por exemplo, rios e riachos são objetos hidrográficos e cidades e estados são unidades administrativas. Dados geográficos como, por exemplo, mapas e imagens de sensoriamento remoto, normalmente se referem a um conjunto de objetos geográficos.

Atributo	Abrev.	Descrição
Scope Note	SN	Descrição do significado do conceito
Use	USE	o termo que se segue ao símbolo é o termo preferido quando é utilizado como sinônimo ou quase-sinônimo ¹¹
Use For	UF	O termo que se segue ao símbolo não é o termo preferido quando é utilizado como sinônimo ou quase-sinônimo
Top Term	TT	O termo que se segue ao símbolo é o termo mais geral ao qual o termo anterior ao símbolo pertence
Broader Term	BT	o termo que se segue ao símbolo é um termo de significado mais amplo do que o termo anterior ao símbolo
Narrower Term	NT	o termo que se segue ao símbolo é um termo de significado mais específico do que o termo anterior ao símbolo
Related Term	RT	o termo que se segue ao símbolo é um termo associado ao termo anterior ao símbolo mas não é um sinônimo ou quase-sinônimo

Tabela 8 – Relacionamentos entre conceitos de um tesouro

¹¹ Um quase-sinônimo é um termo com significado semelhante porém mais abrangente, por exemplo, rio pode ser um quase-sinônimo de riacho.

Um tesouro [9] é o vocabulário de uma linguagem controlada de indexação, formalmente organizada, de tal forma que os relacionamentos entre termos são feitos explícitos. São utilizados por dicionários geográficos para classificar suas entradas.

Um tesouro usualmente fornece seis tipos de propriedades para cada termo (Tabela 8), cinco das quais expressam relacionamentos entre termos e uma é a descrição do significado do termo.

O *ADL Gazetteer Feature Type Thesaurus (FTT)* é o resultado da mistura de vários vocabulários de classificação de vários dicionários geográficos. O FTT versão 2.0 de julho de 2002 contém perto de 200 termos e 3000 relacionamentos. A fig mostra um fragmento desse tesouro ilustrando a hierarquia entre os termos.

regions	regions (cont.)
. agricultural	. climatic regions
regions	. coastal zones
. biogeographic	. economic regions
regions	. land regions
. . barren lands	. . continents
. . deserts	. . islands
. . forests	. . . archipelagos
. . . petrified	. . subcontinents
forests	. linguistic
. . . rain	regions
forests	. map regions
. . . woods	. . chart regions
. . grasslands	. . map quadrangle
. . habitats	regions
. . jungles	. . UTM zones
. . oases	
. . shrublands	
. . snow regions	
. . tundras	
. . wetlands	

Figura 7 – Fragmento da hierarquia de termos do ADL Gazetteer Feature Type Thesaurus

2.3.3. Padrões para dicionários geográficos

Da mesma forma que catálogos utilizam especificações padrões para permitir interoperabilidade de aplicações e compartilhamento de dados,

dicionários também se beneficiam de padrões internacionalmente aceitos. O comitê técnico *ISO/TC211, Geographic Information/ Geomatic* publicou um padrão para dicionários geográficos, o *ISO 19112:2003 Spatial Referencing by Geospatial Identifiers*, que define um esquema conceitual de dados (Tabela 9) e linhas gerais para a descrição de sistemas de referenciamento indireto.

O ADL Gazetteer Content Standard suporta descrições de dados geográficos bastante ricas que vão muito além de alguns dicionários geográficos tradicionais. A Tabela 10 contém as principais seções do ADL Standard, que pode ser subdividido em subseções para codificar informações mais detalhadamente. Por exemplo, *streetAddressSection* é dividido em *streetAddress*, *city*, *stateProvince*, *postalCode* e *country*.

Attribute	Comment
geospatial identifier	a unique identifier for the feature taken from a given identifier space
alternative identifiers	one or more alternative identifiers
name	the preferred name for the feature
type	the type of the feature, typically taken from a feature type thesaurus
description	a description of the feature
position	coordinates of a representative point of the feature
temporal extent	a description of the temporal extension of the feature
geospatial extent	a description of the geospatial extension of the feature
administrator	organization responsible for maintaining the feature's name
parent	identifier of the parent feature of which the entry is a subdivision
child	identifiers of the features that the entry is a parent

Tabela 9 – Alguns atributos selecionados do esquema conceitual da ISO 19112

Attribute	Comment
featureID	unique identifier for the feature
featureName	the names for the feature
classSection	the primary type of the feature
codeSectionType	the code associated with the feature
spatialLocation	the map location of the feature
streetAddressSection	the street address of the feature
relatedFeature	relationships of the feature with other features
description	a short narrative description of the feature
featureData	data about the feature, such as population or elevation
featureLink	website which provides information on the features
supplementalNote	note explaining unusual circumstance of the feature
entryMetadata	documents about entry and modification dates

Tabela 10 – Alguns atributos selecionados do ADL Gazetteer Content Standard

O *Open GIS Consortium* desenvolveu um padrão para serviços de dicionários geográficos [22] para acesso distribuído. A interface de serviço especifica quatro operações que podem ser solicitadas por um cliente e executadas pelo serviço de dicionário.

- **GetCapabilities:** permite que um cliente recupere metadados sobre o service de dicionário sendo oferecido.
- **DescribeFeatureType:** permite que um cliente obtenha um esquema XML descrevendo um objeto geográfico.
- **GetFeature:** permite que um cliente obtenha instâncias de objetos geográficos. Além disso, o cliente pode escolher os atributos a recuperar e condições de seleção de objetos.
- **GetGMLObject:** permite ao cliente obter instâncias através de referências aos identificadores de objetos geográficos.

2.3.4. Alexandria Digital Library Gazetteer

A Alexandria Digital Library (ADL) é uma biblioteca digital distribuída contendo coleções de materiais geo-referenciados. O projeto visa modelar, prototipar e avaliar arquiteturas de bibliotecas digitais aplicações de dicionários geográficos, aplicações de educação e componentes de software.

O ADL Gazetteer é um dicionário geográfico, disponível via Web, que utiliza a ADL como infra-estrutura. Oferece uma interface HTTP/XML para ser utilizada em aplicações geográficas.

As entradas desse dicionário são classificadas segundo três tesouros independentes:

- ADL Feature Type Thesaurus (ADL FTT) – É uma coleção de termos que classifica o tipo do objeto descrito por uma entrada do ADL Gazetteer. Todas as entradas são classificadas segundo o FTT. Nele existem termos hierarquicamente definidos como, por exemplo, “canais” e “estruturas hidrográficas”. O último é um subtipo do primeiro. Todas as entradas classificadas como “canais” são, também, uma “estrutura hidrográfica”. Diz-se que “estrutura hidrográfica” é o *Broader term* de “canal”. Entretanto, pode haver mais de uma hierarquia de conceitos admitida por um tipo. Tomemos o exemplo anterior do canal, segundo a visão apresentada um “canal” é um tipo de “estrutura hidrográfica”, mas poderíamos ainda pensar em uma outra classificação de objeto como “dispositivos de transporte” para a qual um “canal” estaria hierarquicamente subordinado, como uma espécie de dispositivo. Para acomodar esse conceito, uma vez que só é possível definir uma hierarquia no modelo do FTT, existem os *Related terms*. São todos os tipos que por alguma semântica se relacionam com o primeiro. Diz-se, então, que “dispositivos de transporte” é um “*Related term*” de “canal”.
- *Gazetteer type terms of U.S. Geological Survey* – apenas algumas entradas são classificadas segundo esse dicionário
- *U.S. Geospatial-Intelligence* – apenas algumas entradas são classificadas segundo esse dicionário

Cada entrada do ADL Gazetteer descreve único lugar geográfico conceitual. Seus atributos e cardinalidades, são:

- um *identificador* → *String* que identifica univocamente a entrada. É único dentro do dicionário, mas não universalmente único.
- Zero ou mais *códigos* → Identifica o lugar em um esquema de código específico, *namespace* ou sistema.
- um *status* de lugar → Status da existência do lugar, pode ser: *former* (o lugar não existe mais), *current* (o lugar existe) ou *proposed* (o lugar não existe, mas sua criação foi antecipada).
- um *nome de exibição*
- um ou mais *nomes*
- *coordenadas* da diagonal de um retângulo de contorno
- um ou mais *footprints* → É uma aproximação, expressa em coordenadas de latitude e longitude, de uma região da superfície terrestre, coberta pelo lugar. Não precisa ser contínuo.
- zero ou mais *classes* → Classifica o lugar conforme o dicionário utilizado.
- Zero ou mais *relacionamentos*

Os atributos *nome*, *footprint* e *classes* são qualificados por quatro descritores:

- *primary* → indica que o valor do atributo é oficial ou preferido
- *former* → indica que o valor do atributo não existe mais
- *current* → indica que o valor do atributo existe
- *proposed* → indica que o valor do atributo ainda não existe, mas foi criado antecipadamente.

As seguintes condições devem existir para cada entrada do dicionário:

- Exatamente um *nome* precisa ser marcado como *primary*.
- Exatamente um *footprint* precisa ser marcado como *primary*.
- Se houver *classificação*, ao menos uma delas precisa ser marcada como *primary*.

Os nomes geográficos do dicionário podem ser consultados através de um serviço Web utilizando-se o protocolo *ADL Gazetteer Service Protocol*, versão 1.2¹². Trata-se de um protocolo leve, sem estado, baseado em XML sobre HTTP. O protocolo suporta consulta ao dicionário (por nome de lugar, por área geográfica, por classe de objeto, por relacionamento entre nomes de lugares), recuperação de relatório padrão de nomes de lugares e *download* das entradas do dicionário.

Os serviços oferecidos são:

- *get-capabilities* → retorna um resumo de todos os recursos do dicionário: os tipos de serviços e consultas, o dicionário que o dicionário usa, etc.
- *query* → retorna um relatório (padrão ou estendido) das entradas do dicionário selecionadas por uma consulta expressa em GQL (*Gazetteer Query Language*).
- *download* → semelhante à *query* na seleção do tipo de relatório mas retorna todas as entradas do dicionário.

Esses serviços são invocados enviando-se requisições HTTP POST, contento a descrição do serviço desejado codificada em XML, para uma URL que representa o ponto de acesso comum a todos os serviços. O Apêndice A apresenta o esquema XML utilizado na comunicação com dicionário pela Web.

Uma consulta freqüente ao dicionário seria: “Quais os objetos contidos em uma região geográfica específica?”. O fragmento de código a seguir apresenta uma consulta ao dicionário, de acordo com o protocolo, buscando todos os objetos contidos na região geográfica de forma retangular cujas coordenadas da diagonal do retângulo são (latitude = -22,892, longitude = -43,26) e (latitude = -22,932, longitude = -43,203).

¹² <http://www.alexandria.ucsb.edu/gazetteer/protocol>

```
<?xml version="1.0" encoding="UTF-8"?>
<schema xmlns="http://www.w3.org/2001/XMLSchema" />
<gazetteer-service
xmlns="http://www.alexandria.ucsb.edu/gazetteer"
  xmlns:gml="http://www.opengis.net/gml" version="1.2">
  <query-request>
    <gazetteer-query>
      <footprint-query operator="within">
        <gml:Box>
          <gml:coordinates>
            -43.26,-22.891,-43.203,-22.932
          </gml:coordinates>
        </gml:Box>
      </footprint-query>
    </gazetteer-query>
    <report-format>standard</report-format>
  </query-request>
</gazetteer-service>
```