



INSTITUTO SUPERIOR DE ENGENHARIA DE LISBOA
Departamento de Engenharia de Electrónica e Telecomunicações
e de Computadores



Integração de Informação Geográfica

Ester Fernandes Gonçalves
(Bacharel)

Dissertação de natureza científica realizada para obtenção do grau de
Mestre em Engenharia Informática e de Computadores

Orientadores:

Professor-adjunto Porfírio Pena Filipe, ISEL
Professor-adjunto Paulo Medeiros de Araújo, ISEL

Júri:

Presidente: Professor-coordenador Manuel Martins Barata, ISEL

Vogais:

Professor-adjunto Porfírio Pena Filipe, ISEL
Professor-adjunto Paulo Medeiros de Araújo, ISEL
Professor-adjunto João Carlos Amaro Ferreira, ISEL

Março de 2010

RESUMO

A integração de informação geográfica disponível em diferentes formatos levanta a necessidade de criar mecanismos de resolução de incompatibilidades, tendo em vista facilitar a partilha e reutilização de informação.

No âmbito desta dissertação, propõe-se uma arquitectura que efectue a integração de informação geográfica. A arquitectura processa diferentes fontes de informação geográfica com diferentes formatos. A informação geográfica apresentada nas entradas é transformada atendendo a um formato de representação interno e, após ser realizada a sua integração, é disponibilizada em múltiplos formatos de saída.

A arquitectura proposta é modular e inclui o módulo de leitura das entradas, o módulo que associa as fontes de informação aos respectivos formatos, o módulo de conversão dos dados de entrada no formato de representação interno e o módulo de integração que gera os dados nos diferentes formatos de saída.

Com o objectivo de avaliar experimentalmente a arquitectura, foi desenvolvido um protótipo. Foram processados exemplos de informação geográfica relativos a cenários de integração que cobrem as funcionalidades da arquitectura, nomeadamente a inclusão de novos formatos de informação geográfica e a integração de fontes de informação homogéneas ou heterogéneas.

Os resultados obtidos confirmam que a arquitectura proposta é adequada à integração de fontes de informação geográfica e que é uma contribuição válida para a resolução de problemas de interoperabilidade em sistemas de informação geográfica.

ABSTRACT

The integration of spatial information available in different formats raises the need to establish mechanisms for resolution of conflicts, in order to facilitate the sharing and reuse of information.

In this context, this thesis proposes an architecture that performs the integration of geographic information. The architecture handles many different sources of geographic information in different formats. The geographic information presented as input is transformed into a given format of internal representation, and after being integrated, is made available in multiple output formats.

The proposed architecture is modular, including the module of input reading, the module that links the sources of information to their formats, the conversion module of input data into the format of internal representation and the integration module that generates the data in different output formats.

With the aim to experimentally evaluate the architecture, a prototype was developed. Samples of spatial information were processed relating to integration scenarios that cover the architectural features, including the addition of new formats of spatial information, and the integration of homogeneous and heterogeneous information sources.

The results confirm that the proposed architecture is suitable for integration of sources of geographic information and is a valuable contribution to the resolution of interoperability issues in geographic information systems.

PALAVRAS-CHAVE

Sistemas de Informação Geográfica, Geoprocessamento, Georeferenciação, Característica geográfica

KEYWORDS

Geographic Information Systems, Geoprocessing, Georeferentiation, Feature

AGRADECIMENTOS

O trabalho desenvolvido no âmbito desta tese não teria sido possível sem o contributo de várias pessoas, aos quais não posso deixar de reiterar os meus agradecimentos.

Ao professor Porfírio Pena Filipe, pela orientação paciente, a clareza de ideias, a constante insistência nos pontos a melhorar e todo o apoio dado ao longo deste processo.

Ao professor Paulo Araújo, pela enorme disponibilidade e o contributo sempre presente em todo o processo de desenvolvimento do trabalho e nas constantes e intermináveis revisões.

Ao Eng. Pedro Machado, do Banco BPI, pelo apoio financeiro e logístico que facilitou o desenvolvimento deste trabalho.

À Eng. Sandra Castro, especialista em SIGs da ESRI, pelo enorme contributo na compreensão do funcionamento dos Sistemas de Informação Geográfica, e ainda pelo tempo que dispensou, sempre com uma enorme disponibilidade.

Ao Eng. Pedro Lopes, da Oracle, pela disponibilidade e apoio no esclarecimento de questões técnicas, e por contribuir com uma visão global da integração técnica da informação geográfica.

À minha família, sempre presente com uma palavra de apoio, e empurrando-me constantemente em direcção a esta meta.

Aos meus pais, Matilde e António, pelo amor e carinho incondicionais que sempre demonstraram, e pelo constante apoio sempre presente em todas as vertentes pessoais e profissionais.

Às minhas irmãs, por serem maravilhosas como só elas sabem ser.

Um agradecimento muito especial à minha irmã Rute, pela valiosa ajuda na revisão e formatação, pela discussão de aspectos relacionados com o tema desta dissertação, e sobretudo pela infinita paciência e apoio que sempre, e incondicionalmente, disponibiliza.

Aos meus amigos, pela força constante, sempre impulsionando para levar este trabalho a bom porto.

Para todos estes, e também para todos aqueles que não nomeio por nome, mas que de alguma forma contribuíram para este desfecho, o meu sincero obrigada.

22 de Março de 2010

ÍNDICE

1. INTRODUÇÃO.....	1
1.1. Motivação.....	2
1.2. Contribuição.....	3
1.3. Organização da Dissertação	5
2. SISTEMAS DE INFORMAÇÃO GEOGRÁFICA.....	7
2.1. A Interoperabilidade.....	8
2.2. Projecções, sistemas de coordenadas e registo.....	9
2.2.1. Geoide e Elipsoide.....	10
2.2.2. Sistemas de coordenadas	11
2.2.3. Datums.....	12
2.2.4. Projecções de mapas.....	12
2.3. Captura de informação geográfica	12
2.4. Metadados	13
2.5. Tipos de Formatos de Ficheiros Geográficos.....	14
2.5.1. Tipo de Formato Raster	15
2.5.2. Tipo de Formato Vectorial.....	16
2.5.3. Formato Raster e Formato Vectorial – Comparação.....	18
2.5.4. Tradução de Raster para Vectorial.....	18
2.6. Tipos de Informação num Mapa.....	19
2.6.1. Informação Geográfica	19
2.6.2. Informação sobre os Atributos	20
2.6.3. Informação de visualização	20
2.7. Os Sistemas de Informação Geográfica.....	20
3. REPRESENTAÇÃO DE INFORMAÇÃO GEOGRÁFICA	23
3.1. Integração de Informação Geográfica	23
3.2. Tipos de Formatos de Ficheiros	24
3.2.1. Formatos Vectoriais	24
3.2.2. Formatos Raster	26
3.3. Integração de Formatos de Ficheiros Geográficos	27
3.3.1. DPLAN – Sistema de Planeamento de Distribuição.....	27
3.3.1.1 DPLAN: O Problema.....	27

3.3.1.2	DPLAN: Desenho da solução	28
3.3.1.3	DPLAN: A Implementação	28
3.3.2.	SIG Empresarial – Orientação a Serviços	29
3.3.2.1	SIG Empresarial: O Problema	29
3.3.2.2	SIG Empresarial: Desenho da solução.....	29
3.3.2.3	SIG Empresarial: A Implementação	29
3.3.2.4	SIG Empresarial: A Arquitectura.....	30
3.3.3.	ATIS – Sistema de Controlo de Tráfego	32
3.3.3.1	ATIS: O Problema.....	32
3.3.3.2	ATIS: Desenho da solução	34
3.3.3.3	ATIS: A Implementação	34
3.3.4.	DISPRO – <i>A web based GIS</i>	36
3.3.4.1	DISPRO: O Problema	37
3.3.4.2	DISPRO: Desenho da solução	38
3.3.4.3	DISPRO: A Implementação	38
3.3.5.	Open Format Converter	40
3.3.5.1	OFC: O Problema	40
3.3.5.2	OFC: Desenho da solução.....	41
3.3.5.3	OFC: A Implementação	41
3.4.	Conclusões sobre as arquitecturas analisadas	44
3.5.	Identificação de trabalho em aberto	45
4.	OPENGISINTEGRATOR – PROPOSTA DE INTEGRAÇÃO DE INFORMAÇÃO GEOGRÁFICA....	47
4.1.	Descrição do OpenGisIntegrator	47
4.2.	Componentes Lógicos do OpenGisIntegrator	49
4.2.1.	Controller	50
4.2.2.	DataRW Engine.....	50
4.2.2.1	Leitura dos Ficheiros de Entrada.....	51
4.2.2.2	Escrita dos Ficheiros de Saída	51
4.2.3.	Converter Engine	52
4.2.3.1	Interface Recognizer	52
4.2.3.2	Converter Processor	54
4.2.4.	Integrator Engine.....	57
4.3.	Formato de Representação Interno	59
4.4.	Resolução de Conflitos	61
4.4.1.	Tipos de Conflitos.....	61
4.4.2.	Métodos de Resolução de Conflitos	61
4.4.2.1	Métodos de Verificação	62
4.4.2.2	Métodos de Adição	63

4.4.2.3	Métodos de Actualização.....	63
4.4.2.4	A Integração	63
4.5.	Fluxo de Integração de Ficheiros – <i>Putting It All Together</i>	64
4.5.1.	Dados de Entrada	65
4.5.2.	Conversão dos Dados de Entrada.....	65
4.5.3.	Integração dos Dados de Entrada	65
4.5.4.	Geração do Ficheiro de Saída	65
4.5.5.	Resumo do Processo.....	66
4.6.	OpenGisIntegrator - A Proposta de Arquitectura.....	66
5.	AVALIAÇÃO EXPERIMENTAL	67
5.1.	Interface Gráfica do Protótipo.....	67
5.2.	Parametrizações de Configurações	68
5.2.1.	Premissas	68
5.2.2.	Implementação do Processo de Integração	68
5.3.	Cenários de Avaliação	70
5.3.1.	Avaliação da Arquitectura	70
5.3.1.1	Avaliação da Escalabilidade da Arquitectura	70
5.3.1.2	Avaliação da Integração de Ficheiros de Formatos Diferentes.....	72
5.3.1.3	Avaliação dos Ficheiros de Saída em Diferentes Formatos	73
5.3.2.	Avaliação da Integração das Características Geográficas	75
5.3.2.1	Avaliação da Integração de Pontos.....	75
5.3.2.2	Avaliação da Integração de Linhas	78
5.3.2.3	Avaliação da Integração de Polígonos	81
5.4.	O Protótipo.....	83
6.	CONCLUSÕES E TRABALHO FUTURO	85
6.1.	Conclusões	85
6.2.	Trabalho Futuro	86

ÍNDICE DE TABELAS

Tabela 1 - Métodos de captura de informação geográfica	13
Tabela 2 - Comparação de formatos Raster e Vectorial	18
Tabela 3 - Tipos de Informação em Mapas Digitais	19
Tabela 4 - Exemplos de Formatos Vectoriais	25
Tabela 5 - Exemplos de Formatos Raster	26
Tabela 6 - Princípios INSPIRE.....	37
Tabela 7 - OpenGisIntegrator - Componentes Lógicos	49
Tabela 8 - OpenGisIntegrator - Responsabilidades do Controller.....	50
Tabela 9 - OpenGisIntegrator - Métodos das Interfaces	56
Tabela 10 - OpenGisIntegrator - Classes do formato de Representação Interno	60
Tabela 11 - Tipos de Conflitos na Integração Geográfica	61
Tabela 12 - Classe feature - Tipos de Métodos.....	62
Tabela 13 - Parametrizações do Protótipo	69

ÍNDICE DE FIGURAS

Figura 1 - Contexto de Integração de Informação	4
Figura 2 - Mapa de casos de cólera do Dr. John Snow.....	7
Figura 3 - Sistemas de coordenadas	11
Figura 4 - Imagem no formato Raster.....	15
Figura 5 - Imagem com informação Vectorial.....	16
Figura 6 - DPLAN – Processo de integração anterior	28
Figura 7 - DPLAN – Novo processo de integração.....	28
Figura 8 - SIG Empresarial – Arquitectura do sistema	30
Figura 9 - SIG Empresarial – Parte do ficheiro XML de registo.....	31
Figura 10 - SIG Empresarial – Pedido GetMap a partir de uma única fonte	32
Figura 11 - SIG Empresarial – Pedido GetMap a partir de várias fontes	32
Figura 12 - ATIS – Arquitectura proposta	35
Figura 13 - DISPRO – Arquitectura	39
Figura 14 - DISPRO – Implementação do WMS da OGC	39
Figura 15 - DISPRO – Integração do WMS.....	40
Figura 16 - Open Format Converter– Arquitectura	41
Figura 17 - OFC – Estrutura da herança dos objectos do Geo-Core	42
Figura 18 - OFC – Exemplo de ficheiro XML de configuração	43
Figura 19 - OpenGisIntegrator - Diagrama de Contexto.....	47
Figura 20 - OpenGisIntegrator - Arquitectura	48
Figura 21 - OpenGisIntegrator - O DataRW Engine	51
Figura 22 - OpenGisIntegrator - Geração dos Ficheiros de Saída.....	51
Figura 23 - OpenGisIntegrator - Arquitectura detalhada do Interface Recognizer	53
Figura 24 - OpenGisIntegrator - Identificação de Interfaces dos Ficheiros de Saída.....	53
Figura 25 - OpenGisIntegrator - Arquitectura detalhada do Converter Processor	54
Figura 26 - OpenGisIntegrator - Conversão dos Ficheiros de Saída	54
Figura 27 - OpenGisIntegrator - Diagrama de Classes do Converter Processor	55
Figura 28 - OpenGisIntegrator - Detalhe da Arquitectura do Integrator Engine	57
Figura 29 - OpenGisIntegrator - Instâncias da classe input_data	58
Figura 30 - OpenGisIntegrator - Detalhe do Integrator Engine.....	58
Figura 31 - OpenGisIntegrator - Diagrama de Classes da Representação Interna	59

Figura 32 - OpenGisIntegrator - Diagrama de Sequência – Ficheiros de Entrada.....	64
Figura 33 - OpenGisIntegrator - Diagrama de Sequência – Ficheiros de Saída	64
Figura 34 - Protótipo do OpenGisIntegrator.....	67
Figura 35 - Integração de Pontos - Ficheiro #1.....	76
Figura 36 - Integração de Pontos - Ficheiro #2.....	77
Figura 37 - Integração de Pontos - Ficheiro de Saída	77
Figura 38 - Integração de Linhas - Ficheiro #1	79
Figura 39 - Integração de Linhas - Ficheiro #2	79
Figura 40 - Integração de Linhas - Ficheiro de Saída.....	80
Figura 41 - Integração de Polígonos - Ficheiro #1	81
Figura 42 - Integração de Polígonos - Ficheiro #2.....	82
Figura 43 - Integração de Polígonos - Ficheiro de Saída	83

1. INTRODUÇÃO

In ancient maps of the world, expanses of unknown territory might hold a warning to would-be explorers: Here there be monsters.

For today's explorers seeking to navigate and understand the world of science, the monsters are the untamed collections of data that inhabit a largely uncharted landscape.

[Arlington, 2004]

Desde sempre, o ser humano sentiu necessidade de comunicar. Desde formas de comunicação ancestrais, como pinturas rupestres em paredes de cavernas, até aos dias de hoje, em que a comunicação se tornou possível a longas distâncias, permitidas pelos avanços tecnológicos, a troca de informação sempre desempenhou um papel central na evolução de civilizações.

Atualmente, os métodos de acesso à informação são diferentes dos métodos utilizados tradicionalmente durante centenas de anos. Até há bem pouco tempo atrás, a informação estava disponível em obras de referência, livros e outros suportes físicos disponíveis em bibliotecas, muitos deles autenticados por autoridades nas matérias em estudo. Hoje em dia, a informação está disponível através de computadores, que permitem o acesso rápido a um mundo de informação.

A difusão da informação conduziu à necessidade de a analisar e interpretar, de forma a conhecer o mundo que rodeia o homem e tentar compreendê-lo.

O acesso a informação geográfica não é exceção. Também neste campo se assistiu à mudança na forma como a informação está disponível. A tomada de decisões baseadas em geografia é uma tarefa básica do pensamento humano. Compreender a geografia e a relação das pessoas com os locais permite tomar decisões informadas sobre a forma como os seres humanos vivem.

É neste contexto que surgem os Sistemas de Informação Geográfica, vulgarmente abreviado por SIGs¹.

Um SIG é um sistema computacional de armazenamento, acesso, manipulação, análise, e apresentação de dados referenciados geograficamente [Hall et al, 2003].

¹ Em inglês, GIS, sigla para *Geographic Information System*, ou *Geographical Information System*

Hoje em dia, os SIGs são cada vez mais populares. Estão presentes em sistemas de navegação para automóveis, em aplicações a três dimensões na Internet ou como ferramentas de visualização de mapas interactivos [Bryntse, 2007].

Um SIG permite tratar informação geográfica computacionalmente. A diferença principal entre um sistema de informação convencional e um SIG é que este último tem a capacidade de armazenar atributos que descrevem os diferentes tipos de informação geográfica.

[Casanova et al., 2005] refere duas responsabilidades como sendo as principais de um SIG:

1. Permitir inserir e integrar informações geográficas numa única base de dados
2. Disponibilizar funcionalidades de consulta, manipulação e análise dos conteúdos da base de dados geográficos

Os SIGs são ferramentas tecnológicas que permitem compreender a geografia do terreno e tomar decisões com base nesse conhecimento. Permitem pesquisar e analisar informação espacial, e perceber, visualizar, compreender, questionar, interpretar e visualizar essa informação, de modo a revelar relações, padrões e tendências [ESRI, 2008].

1.1. Motivação

[Stoimenov, 2006] descreve um cenário de integração de informação geográfica relativa a um objecto geográfico existente numa comunidade local, alvo de interesse por parte de diferentes organizações.

Cada organização tem uma visão e um entendimento diferentes desse objecto. De acordo com isso, cada organização produz diferentes conjuntos de dados, ou atributos, que descrevem o mesmo objecto real. Alguns atributos são comuns a todas as organizações da comunidade local, alguns são comuns a algumas delas, e cada organização pode ter atributos específicos por conjunto de dados. Para além disso, as diferentes organizações podem usar termos diferentes para os mesmos atributos (sinónimos) ou o mesmo termo para diferentes atributos (homónimos). A forma como as diferentes organizações gerem os atributos que descrevem um mesmo objecto geográfico ilustra bem o modo como um mesmo objecto pode ser alvo de diferentes interpretações, descrito com diferentes designações. Integrar informação geográfica sobre o objecto em estudo, recorrendo a informação proveniente de diferentes modelos de dados das diferentes organizações pode levar a sérios conflitos na integração e interpretação da informação geográfica.

Em [Stoimenov, 2006] sugere-se como solução para a transferência de dados entre diferentes SIGs a existência de uma arquitectura e um conjunto de normas para dados

geoespaciais. Uma das importantes estratégias para a interoperabilidade é a conversão de diferentes formatos de dados em estruturas de dados comuns. Este tipo de estrutura de dados baseia-se usualmente num formato específico adoptado por um SIG existente. No entanto, é impossível conceber que a comunidade adopte uma arquitectura geoespacial ou uma norma de dados única. Isto significa que os esforços de normalização por si só não produzem a interoperabilidade.

Nos últimos anos, o OGC surgiu como uma força na tendência para a postura aberta. O OGC é um consórcio de vendedores de SIGs, agências e instituições académicas, que introduziu abordagens para a gestão efectiva de tecnologia de interoperabilidade em cooperação com a indústria e as universidades. Identificou ainda a necessidade de partilha aberta de dados geográficos e a troca de serviços SIG abertos.

Apesar das iniciativas de normalização, o uso de normas como o único esforço válido para atingir a interoperabilidade não é amplamente aceite. A heterogeneidade surge naturalmente a partir de um mercado livre de ideias e produtos, e por essa razão, a ideia de banir a heterogeneidade por decreto é inviável [Stoimenov, 2006]. Como uma consequência das duas principais características do sistema, fontes de dados distribuídos e a sua heterogeneidade, a interoperabilidade entre sistemas é um processo pesado.

1.2. Contribuição

A dificuldade na comunicação de informação geoespacial entre diferentes fontes reside não tanto na incompatibilidade dos formatos de dados, mas sobretudo na heterogeneidade dos modelos de dados. Estes modelos, categorizados de acordo com estruturas e esquemas próprios do sistema que os utiliza, são um obstáculo à interoperabilidade e consequentemente à integração. Este problema constitui uma limitação na capacidade dos sistemas trocarem informação entre si, dificultando a interoperabilidade entre diferentes sistemas.

A interoperabilidade levanta a necessidade de conceber mecanismos que facilitem a resolução de incompatibilidades ou heterogeneidades entre diferentes modelos de dados, de forma a permitir a partilha e reutilização da informação envolvida. Neste cenário, alcançar a interoperabilidade plena é um desafio de investigação relevante [Kavouras et al., 2002].

Tipicamente, são utilizadas abordagens empíricas para resolver as heterogeneidades de forma a facilitar a transmissão e reutilização de informação. No entanto, as abordagens empíricas apenas lidam com problemas específicos e não providenciam uma solução uniforme e abrangente. Assim, o desenvolvimento de metodologias específicas e formais para resolver sistematicamente as heterogeneidades e associar diferentes normas de categorização são um passo essencial para a integração de informação geográfica.

Perante este cenário, propõe-se nesta dissertação contribuir para a integração de informação geográfica representada em diferentes modelos de dados.

A proposta descrita nesta dissertação descreve como uma solução possível para o problema de integração de informação em diferentes formatos uma arquitectura que suporte a integração da informação.

A Figura 1 apresenta um diagrama representativo do contexto da transformação que concretiza a integração da informação. A figura mostra os dados provenientes das diferentes fontes $SIG_1, SIG_2, \dots, SIG_n$, que podem ser dados provenientes de outras aplicações ou sistemas, ou simplesmente fontes de informação autónomas, como ficheiros, ou ligações a bases de dados. A estes dados é aplicado um mecanismo de transformação. Este mecanismo visa adaptar para um formato pré-definido a informação recebida, e apresentar a informação integrada resultante a um visualizador, que os apresenta ao utilizador.

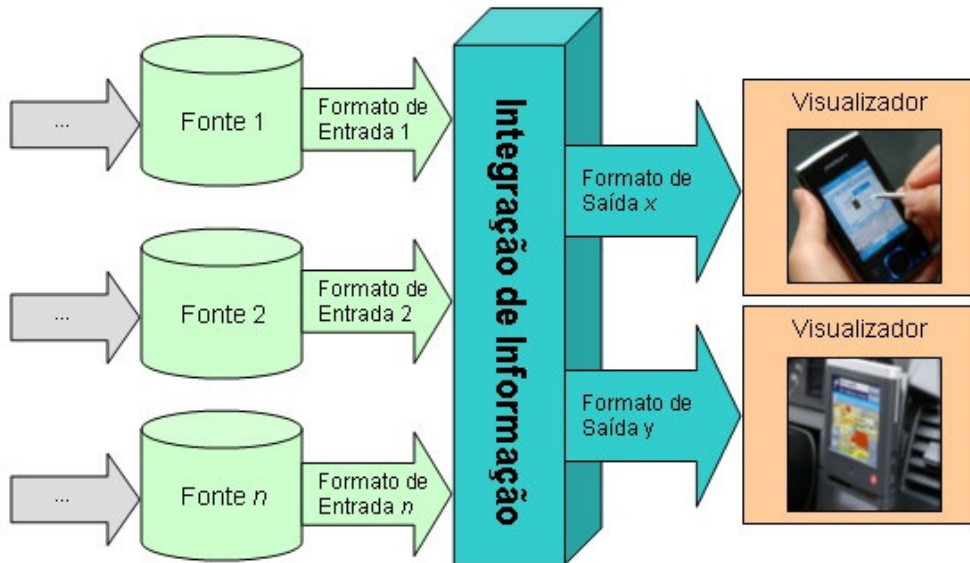


Figura 1 - Contexto de Integração de Informação

O objectivo da integração de informação é a integração de diferentes perspectivas existentes sobre um mesmo objecto real georreferenciado, localizado numa área geográfica específica de interesse para o utilizador. O foco desta dissertação é a definição de uma arquitectura que permita implementar processos de integração ou transformação desta informação.

1.3. Organização da Dissertação

A presente dissertação, sob o tema “Integração de Informação Geográfica”, encontra-se organizada da forma descrita em seguida.

A descrição de conceitos e metodologias de Sistemas de Informação Geográfica, bem como a análise de casos estudados é apresentada em dois capítulos, os capítulos 2 e 3. O capítulo 2 descreve os Sistemas de Informação Geográfica, explicando o que são e abordando alguns conceitos essenciais deste tipo de sistemas. O capítulo 3 descreve métodos de representação geográfica, onde se analisam tecnologias e metodologias desenvolvidas e adoptadas na integração de informação.

O capítulo 4 apresenta a proposta de integração de informação geográfica, onde é detalhada a arquitectura da solução desenhada.

O capítulo 5 descreve os resultados da avaliação experimental da arquitectura proposta, feita com recurso a um protótipo da solução, também descrito neste capítulo.

Por último, o capítulo 6 apresenta as conclusões obtidas e identifica trabalhos futuros a desenvolver.

2. SISTEMAS DE INFORMAÇÃO GEOGRÁFICA

Um Sistema de Informação Geográfica, vulgarmente designado por SIG, é um sistema informático com capacidade para armazenar, consultar, alterar, analisar e apresentar informação espacial, referenciada geograficamente [Hall et al, 2003].

Relacionar eventos com localizações geográficas não é uma abordagem recente. Em 1854, o Dr. John Snow, um médico britânico considerado um dos fundadores da epidemiologia, utilizou a referenciação geográfica para identificar a fonte de um surto de cólera em Londres [Rosenberg]. Snow desenhou um mapa que representava a localização dos locais de habitação das vítimas e onde se situavam as bombas de água que serviam essas zonas. A Figura 2 apresenta o mapa desenhado por Snow.



Figura 2 - Mapa de casos de cólera do Dr. John Snow

John Snow defendia que a cólera era disseminada pela água consumida, opinião contestada na época pela classe médica. Ao desenhar o mapa que relacionava os casos de cólera com as bombas que forneciam a água, Snow identificou um aglomerado de focos da doença junto de uma das bombas, situada em *Broad Street*. Quando a bomba foi fechada, os casos diminuíram.

Relacionando ocorrências de um evento com informação geográfica, neste caso, a localização dos casos de cólera e das bombas de água, Snow pôde chegar a conclusões sobre o surto de cólera que assolava Londres.

Esta história ilustra bem como um SIG permite relacionar informação referenciada geograficamente e chegar a conclusões com base na análise da informação referente a uma determinada zona geográfica em estudo.

O desenvolvimento de SIGs tal como os conhecemos hoje iniciou-se na década de 60, quando passou a ser possível fazer análise computacional de objectos geográficos. Actualmente, os SIGs tornam-se cada vez mais populares num vasto número de áreas de actuação. Ao sobrepor diferentes camadas de mapas, podem ser encontradas e analisadas novas relações.

Um exemplo de utilização de informação geográfica em diferentes camadas é um mapa que permita visualizar um conjunto de ruas, ao qual é sobreposta uma camada que mostre a densidade populacional naquela zona. Com base nas relações encontradas entre as diferentes camadas, são calculados os melhores pontos de colocação de paragens de autocarro [Ekhlund et al, 2003].

Como ferramenta de processamento e análise de informação geográfica, um SIG é composto por um conjunto de componentes que fazem com que o sistema funcione. Para isso, é essencial lidar com questões geográficas críticas, essenciais a ter em conta na análise geográfica da informação. São abordadas em seguida questões essenciais a ter em conta num SIG.

2.1. A Interoperabilidade

Interoperabilidade significa abertura na indústria de software, já que a disponibilização de estruturas de dados internas permitem aos utilizadores dos SIGs construir aplicações que integram componentes de software de diferentes programadores. Interoperabilidade também significa que deve haver troca livre de dados entre sistemas, porque cada sistema deve ter conhecimento dos formatos de dados de outros sistemas [Stoimenov, 2006].

Embora a distribuição de dados geográficos possa oferecer numerosas vantagens sobre bases de dados geográficas isoladas, há problemas técnicos e institucionais de partilha e interoperabilidade de informação geográfica.

O Open Geospatial Consortium, ou OGC, um consórcio de entidades públicas e privadas que trabalham no sentido de definir normas a respeitar nos SIGs, define a interoperabilidade no software como a capacidade de diferentes sistemas geridos localmente trocarem dados e instruções em tempo real que lhes permitam disponibilizar serviços computacionais (como serviços entre cliente e servidor ou serviços *web*). O desafio da interoperabilidade é balancear a necessidade de compatibilidade que os utilizadores têm com a autonomia e a heterogeneidade dos sistemas que interoperam entre si [OGC, 2009].

A interoperabilidade também se refere à interoperabilidade ao longo do tempo, tendo em conta as evoluções dos sistemas mantendo a compatibilidade, quer com versões anteriores, quer com versões futuras.

Em tempos, os protocolos criados para a Internet (IP – Internet Protocols) foram introduzidos como protocolos para comunicação entre redes, ou seja, para movimentar informação entre diferentes redes. A ligação entre redes abriu caminho para a utilização única dos protocolos IP. O resultado foi a Internet, que disponibiliza uma plataforma interoperável com suporte a uma enorme proliferação de serviços e aplicações. É este o modelo que se pretende para uma plataforma interoperável de geoprocessamento [OGC, 2009].

2.2. Projecções, sistemas de coordenadas e registo

Os mapas podem mostrar dados em diferentes escalas. A informação sobre mapas num SIG deve ser manipulada de modo a que se registe, ou se adapte, a informação obtida a partir de outros mapas. Antes que os dados digitais possam ser analisados, podem sofrer outras manipulações – projecções e conversão de coordenadas, por exemplo – que os integram no SIG.

O planeta Terra pode ser representado por diferentes modelos, cada um fornecendo um conjunto diferente de coordenadas (por exemplo, latitude, longitude, elevação) para qualquer ponto dado na superfície da Terra. O modelo mais simples é assumir que a Terra é uma esfera perfeita. À medida que mais medidas da Terra foram acumuladas, os modelos tornaram-se mais sofisticados e mais precisos. De facto, há modelos que se aplicam a diferentes áreas da Terra para providenciar maior precisão na informação. Por exemplo, o *datum* do Norte da América, de 1927, o NAD27, funciona bem no Norte da América, mas não na Europa.

A projecção é uma componente fundamental da construção de mapas. É um meio matemático utilizado para transferir informação a partir de um modelo da Terra, que representa uma superfície curva de três dimensões, para um meio de duas dimensões – papel ou o ecrã de um computador. Diferentes projecções são usadas para diferentes tipos de mapas porque cada projecção serve em particular determinadas utilizações. Por exemplo, uma projecção que represente com precisão as formas dos continentes irá distorcer os seus tamanhos relativos.

Hoje em dia, o utilizador comum está consciente da existência da utilização de GPS para localizações em termos de latitude, longitude e altitude, através das potencialidades de ferramentas como o *Google Earth* e até de SIGs. No entanto, a maioria das pessoas interpreta mal o que são latitude e longitude. [Sajeevan, G., 2008] explica a diferença.

As coordenadas autálicas são o que geralmente se designa por latitude e longitude, em que a Terra é assumida como se tivesse uma forma esférica. No dia a dia, as

coordenadas vistas nos mapas como os dos GPS são latitudes e longitudes geodésicas. É também imperativo o *datum* do mapa em uso. Se o *datum* for alterado, qualquer localização seleccionada pode ter diferentes coordenadas geodésicas.

Assume-se tipicamente que a longitude é o ângulo medido a partir do meridiano de Greenwich, e que a latitude é o ângulo medido a partir da linha do equador [Sajeevan, G., 2008]. Mas nem sempre assim é, como explica a secção seguinte.

2.2.1. Geoide e Elipsoide

Os parâmetros envolvidos na representação de características de superfícies topográficas irregulares (a 3 dimensões) em superfícies planas (a 2 dimensões) são do tipo geoide, esferóide e de projecção.

O geoide é a superfície equipotencial da gravidade da terra, em que a gravidade em cada ponto é igual à sua força ao nível do mar. É a forma que a terra assumiria ao nível do mar, se toda a sua superfície nesse nível fosse líquida. A utilização de geoides como base para a computação seria difícil dado que se trata de uma superfície irregular e não tem uma expressão matemática completa. Consequentemente, uma superfície mais suave, uma Elipsoide, que melhor se adequa a uma superfície geoide é considerada.

O termo elipsóide e esferóide são frequentemente utilizados. Uma elipsóide, também chamada de elipsóide de três eixos, tem eixos desiguais. Rodar uma elipse sobre os seus eixos menor e maior pode gerar um esferóide. Uma elipsóide de revolução e a elipsóide de dois eixos também descrevem um esferóide.

Uma elipsóide é representada por dois eixos, um semi-maior e um semi-menor (ou respectivamente excentricidade ou achatamento).

A altura ortométrica é medida a partir da superfície do geoide, e a altura geodésica é medida a partir da superfície da elipsóide. Enquanto a altura ortométrica de uma localização se mantém constante, a altura geodésica altera-se com o *datum*. Consequentemente, é importante calcular a nova altura ao longo da latitude e da longitude durante a transformação de projecção, no caso da altura geodésica. A altura do geoide é a distância da superfície da elipsóide até à superfície do geoide e pode ter valores positivos ou negativos.

As elipsóides podem ser categorizadas como locais, regionais e globais. Várias elipsóides são formadas para melhor se adaptar a determinadas regiões da terra. Uma elipsóide é seleccionada dependendo dos requisitos da aplicação e das exigências de precisão. Por exemplo, o elipsóide global WGS84 pode ser usado quando a aplicação abrange todo o planeta, dado que melhor se aproxima da forma da terra, embora possa não ser tão adequada para uma parte em particular da terra [Sajeevan, G., 2008].

2.2.2. Sistemas de coordenadas

Alguns sistemas de coordenadas importantes no contexto dos mapas são mostrados na Figura 3. Sistemas rectangulares normais, que utilizam coordenadas Cartesianas, são os sistemas de coordenadas convencionais XYZ normalmente utilizados. As coordenadas autálica e geodésica são duas categorias dos sistemas de coordenadas geográficas.

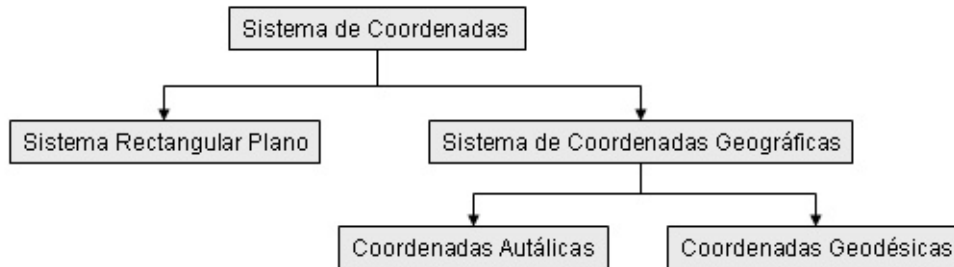


Figura 3 - Sistemas de coordenadas

As coordenadas autálicas baseiam-se na forma esférica da terra. Embora a forma da terra seja mais elipsóide do que esférica, os cartógrafos ainda usam a esfera autálica para mapeamentos e cálculos geodésicos. Uma esfera com a mesma área de superfície que uma elipsóide com o valor aproximado da área da superfície da terra é chamada de esfera autálica. Os conceitos vulgarmente conhecidos como latitude e longitude são na realidade latitudes e longitudes autálicas.

A latitude geodésica numa localização é o ângulo entre o plano equatorial e uma linha normal ao esferóide de referência. Linhas normais a superfícies esferoidais não passam pelo centro, excepto para as latitudes 0° e 90° . Mas a normal passará sempre pelo centro da esfera, no caso das latitudes autálicas. Assim, as latitudes autálicas e geodésicas em qualquer localização x são sempre diferentes.

A longitude geodésica numa localização é o ângulo entre o plano de referência e o plano que passa pela localização, sendo ambos os planos perpendiculares ao plano equatorial. A altura geodésica, ou altura elipsoidal, é a distância entre uma localização definida e a superfície de uma elipsóide de referência numa direcção normal à elipsóide.

A latitude autálica (α) baseia-se na esfera, e a latitude geodésica (Φ) baseia-se no esferóide. O sistema de coordenadas geodésicas, que se baseia no esferóide, é geralmente o seguido em SIGs e em mapeamentos relacionados. Mapas utilizados, por exemplo, no Google Earth são representados em latitude e longitude geodésicas [Sajeevan, G., 2008].

2.2.3. Datums

Um *datum* descreve a posição, orientação e as relações em escala de uma elipsóide de referência para a terra. Consequentemente, um número de *datums* pode ser gerado usando um único elipsóide. O WGS84 é o *datum* usado nos GPS e o seu centro é o centro da massa da terra.

À parte dos parâmetros da elipsóide, mais parâmetros são exigidos para definir um *datum* – três para definir a posição e três para definir o ângulo de rotação – com referência para um sistema de coordenadas Cartesiano a três dimensões com a sua origem a coincidir com o centro da massa da terra. No sistema Cartesiano espacial fixado para a terra (XYZ) o eixo Z coincide com o eixo rotacional da terra, o plano XY é o plano equatorial médio perpendicular a este eixo e o plano XZ é o plano meridional médio de Greenwich.

2.2.4. Projecções de mapas

Mapear envolve transferir localizações de superfícies topográficas para uma superfície geoidal e subsequentemente para uma superfície elipsóide. A informação numa superfície elipsóide pode ser transferida para uma superfície vulgar usando um tipo de projecção adequado.

Azimutal, cónica ou cilíndrica são três famílias principais de projecções. Nenhum mapa mantém uma escala, direcção, distância ou área uniforme. O Google Earth usa uma projecção cilíndrica simples com um *datum* WGS84 para a sua imagem de base.

A transformação das projecções numa aplicação de software é uma função importante exigida para alterar os parâmetros das projecções de mapas, já que parâmetros de projecções com dados disponíveis a partir de diferentes fontes podem ser diferentes daqueles que se destinam a um objectivo em particular [Sajeevan, G., 2008].

2.3. Captura de informação geográfica

Introduzir informação no sistema consome muito do tempo dos utilizadores de SIGs. Existe uma variedade de métodos utilizados para inserir dados num SIG, onde serão armazenados num formato digital.

A Tabela 1 descreve métodos de captura de informação geográfica.

Suporte	Descrição
Papel	Digitalização a partir de mapas existentes em papel. O resultado é guardado num formato Raster, que pode posteriormente ser processado para produzir informação Vectorial.
COGO – Coordinate Geometry	Método de inserção e análise de dados em que a inserção começa num ponto, num dado ângulo ou numa determinada direcção, durante uma distância definida, e continua nesse ângulo ou nessa direcção até que a característica geográfica seja completamente delineada.
Detecção remota	Detecção a partir de sensores anexados a uma plataforma. Os sensores são câmaras ou digitalizadores; as plataformas são tipicamente aparelhos de aviação (aeronaves) e satélites.
Fotografias aéreas	Digitalização de características a partir de pares de fotografias digitais. A fotografia digital de alta qualidade é já utilizada, descartando a necessidade de digitalização.
Sensores remotos via satélite	Utilização de pacotes de sensores para passivamente medir a relevância de partes do espectro electromagnético ou ondas de rádio que foram enviados a partir de um sensor activo, como um radar. Os dados detectados remotamente obtêm dados Raster que podem posteriormente ser processados utilizando diferentes bandas para identificar objectos e classes de interesse, como a cobertura do terreno.

Tabela 1 - Métodos de captura de informação geográfica

Além da informação geográfica capturada e inserida num SIG, é também necessário inserir informação sobre os atributos geográficos que incluem informação adicional sobre os objectos representados no sistema.

Após inserir dados num sistema, geralmente é necessário editá-los, para remover erros, ou para continuar o processamento dos dados. Por exemplo, numa rede de estradas as linhas devem ligar-se aos nós numa intersecção. Para mapas digitalizados, as manchas no mapa de origem podem precisar de ser removidas, pois um ponto de sujidade pode ser confundido com um ponto de intersecção [StateMaster, 2007].

2.4. Metadados

Os metadados são uma parte essencial dos SIGs. Metadados são “dados sobre dados”. Os metadados descrevem o conteúdo, o estado, a localização, ou outras características de um determinado dado. Podem ser armazenados em qualquer formato, como um ficheiro de texto, uma estrutura XML, ou um registo de uma base de dados. O objectivo dos Metadados é permitir identificar se um determinado dado existe, a qualidade da informação constante nesses dados, como está acessível e como pode ser utilizado [Casanova et al., 2005].

Os metadados são utilizados com diferentes objectivos. São listados em seguida os tipos de informação disponibilizada em SIGs com recurso a metadados:

- Informação detalhada sobre métodos de obtenção de dados, técnicas de integração e análise, aplicados aos dados de origem exigida como suporte à preparação de relatórios científicos;
- Informação sobre a precisão dos conjuntos de dados, história de processamento, e procedimentos de armazenamento exigidos para gerir e utilizar com eficácia os dados dentro de uma organização;
- Informação sobre especificações de projecções, escalas, formatos de troca, formatos de compressão e de ficheiros que devem acompanhar as transferências de dados para outras organizações;
- Descrições adequada dos conjuntos de dados ao nível do conteúdo, qualidade e o alcance geográfico disponíveis para disponibilizar nas consultas dos utilizadores;
- Descrições sumárias do conteúdo e da qualidade, bem como informações de contactos;
- Informação sobre software de acesso a conjuntos de dados, bem como parâmetros de software necessários para consultas e visualizações de dados.

Por causa do tamanho reduzido que têm em comparação com a informação que descrevem, os metadados são mais facilmente partilháveis. A criação e partilha de metadados tornam a informação sobre dados existentes imediatamente disponível para qualquer utilizador que a procure [ESRI, Metadata and GIS, 2002].

2.5. Tipos de Formatos de Ficheiros Geográficos

Os ficheiros são estruturas lógicas utilizadas para armazenar informação geográfica. Os formatos dos ficheiros são importantes porque em regra os SIGs não suportam todos os formatos. Ao pretender utilizar informação que não esteja num formato que o SIG em uso suporta, será necessário transformá-lo, encontrar outro conjunto de dados, ou até encontrar um outro SIG.

Quase todos os SIGs têm o seu próprio formato de representação interno [GeoComm.com, 2009]. Estes formatos são desenhados para uso optimizado dentro do sistema e frequentemente são proprietários. Não são desenhados para serem utilizados fora dos sistemas nativos.

Muitos sistemas suportam no entanto transferências de ficheiros em vários formatos. Alguns formatos são desenhados como formatos de transferência e são utilizados para importar ou exportar dados de, ou para, o SIG, por isso são geralmente normalizados ou bem documentados.

Se as necessidades de informação a utilizar forem simples, a preocupação principal será o formato interno suportado pelo SIG. Se as necessidades forem mais complexas, é importante conhecer um conjunto vasto de formatos de transferência, em especial quando se pretender misturar informação de diferentes fontes. Os formatos de transferência serão exigidos para importar alguns conjuntos de dados para o SIG [GeoComm.com, 2009].

Os dados nos SIGs representam objectos do mundo real (estradas, utilização de terrenos, elevações) com dados digitais. Os objectos reais podem ser divididos em duas abstracções: objectos discretos (por exemplo, uma casa), e informação contínua (por exemplo, a taxa de precipitação, ou elevações do terreno). Existem dois métodos alargados utilizados para armazenar dados num SIG para ambas as abstracções: Raster e Vectorial [StateMaster, 2007].

2.5.1. Tipo de Formato Raster

O formato Raster é essencialmente qualquer tipo de imagem digital representada em grelha. Por exemplo, em fotografia digital o pixel é a unidade mais pequena utilizada numa imagem. Uma combinação destes pixéis compõe uma imagem [StateMaster, 2007].

É geralmente utilizado para armazenar informação sobre imagens, como mapas em papel digitalizados ou fotografias aéreas. É também utilizado para armazenar dados capturados por satélite ou outros sistemas de captação aérea de imagens.

Um exemplo deste formato é mostrado na Figura 4.

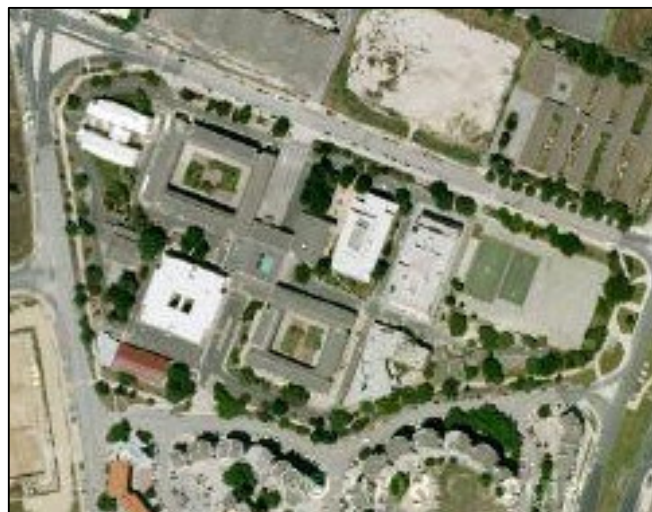


Figura 4 - Imagem no formato Raster

As imagens destes sistemas de captação aérea são frequentemente referidas como informação detectada remotamente. Ao contrário de outros ficheiros Raster, que definem a resolução em termos de tamanho de célula e *dots per inch* (dpi), a resolução em imagens detectadas remotamente expressa-se em metros, que indicam o tamanho da área do terreno coberto por cada célula [GeoComm.com, 2009].

A informação no formato Raster consiste em linhas e colunas de células, cada célula contendo um único valor. Podem ser imagens em que cada pixel ou célula contém um código de cor. Valores adicionais guardados para cada célula podem ser valores discretos, como a utilização do terreno, ou valores contínuos, como temperatura, ou um valor nulo, se não existir informação. Uma célula Raster armazena um único valor, mas pode incluir mais informação, utilizando bandas Raster para representar cores RGB (*red, green, blue*), mapas de cores (um mapeamento entre um código temático e um valor RGB), ou uma tabela de valores associada, com uma linha para cada valor único de célula. A resolução de um conjunto de dados Raster é a largura da célula em unidades de terreno [Smith, Goodchild, Longley, 2007].

A informação no formato Raster é armazenada em vários formatos: desde uma estrutura base, como por exemplo o formato TIF, ou o formato JPEG, ou até um BLOB (*Binary Large Object*) armazenado directamente numa base de dados relacional. O armazenamento em bases de dados, quando correctamente indexado, permite tipicamente a rápida devolução da informação Raster, mas exige um espaço considerável de armazenamento [StateMaster, 2007].

2.5.2. Tipo de Formato Vectorial

Muitos SIGs baseiam-se em tecnologia Vectorial, por isso os formatos Vectoriais são mais comuns. São também os mais complexos pois há diversas formas de armazenar coordenadas, atributos, ligações a atributos, estruturas de bases de dados, e informação a disponibilizar.

A Figura 5 mostra um exemplo de uma imagem com elementos Vectoriais que representam focos de interesse: pontos representam poços, linhas representam rios, e um polígono representa um lago.

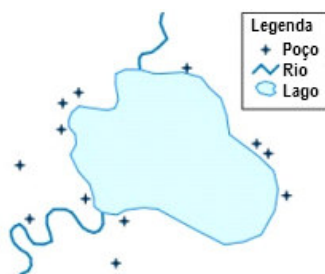


Figura 5 - Imagem com informação Vectorial

Num SIG, as características geográficas são frequentemente expressas em Vectores, sendo essas características consideradas como formas geográficas. Diferentes características geográficas são expressas em diferentes tipos geométricos [StateMaster, 2007]:

- Pontos: São utilizados como características geográficas e que são representadas por um simples ponto de referência, que indicam uma simples localização. Por exemplo, a localização de poços, picos de elevação de terreno ou trilhos são representados como pontos. Os pontos podem também ser utilizados para representar áreas quando visualizados em pequena escala. Por exemplo, cidades num mapa do mundo podem ser representadas como pontos em vez de polígonos. Não são possíveis quaisquer formas de medição com pontos;
- Linhas: São utilizadas como características lineares, como rios, estradas, caminhos-de-ferro, trilhos e linhas topográficas. Tal como com os pontos, características lineares visualizadas em pequena escala são representadas como linhas em vez de polígonos. As linhas permitem medir a distância;
- Polígonos: São utilizados para representar características geográficas que cobrem uma determinada área da superfície terrestre. Essas características podem incluir lagos, limites de parques, edifícios, limites de cidades, ou utilizações de terreno. Os polígonos transmitem o máximo de informação possível deste tipo de ficheiros. Os polígonos permitem medir perímetro e área.

Cada um destes tipos geométricos está descrito numa base de dados, que descreve também os seus atributos. Por exemplo, uma base de dados que descreve lagos pode conter a profundidade de um lago, a qualidade da água ou o nível de poluição. Esta informação pode ser utilizada para fazer um mapa que descreva um determinado atributo do conjunto de dados. Diferentes tipos geométricos podem também ser comparados. Ou pode ser usado um SIG para identificar todos os poços (ponto geométrico) que estejam num raio de 1 quilómetro de um lago (polígono geométrico) que tenha um nível alto de poluição.

Os formatos Vectoriais podem ser feitos de forma a respeitar a integridade espacial através da aplicação de regras topográficas, como por exemplo, “os polígonos não se podem sobrepor”. Dados Vectoriais podem também ser utilizados para representar continuamente fenómenos variáveis. Linhas de contorno ou redes triangulares irregulares (TIN – Triangule Irregular Networks) são utilizados para representar elevações ou outros valores variáveis contínuos. Estas redes gravam valores em pontos de localização, que são ligados por linhas para formar um conjunto irregular de triângulos. A face dos triângulos representa a superfície do terreno.

Dados não espaciais adicionais podem ser armazenados em conjunto com os dados espaciais representados pelas coordenadas de um Vector geométrico ou pela posição

de uma célula Raster. Em dados Vectoriais, os dados adicionais contém atributos da característica geográfica. Por exemplo, um polígono que representa uma floresta pode também ter um valor identificativo e informação sobre as espécies de árvores [StateMaster, 2007].

2.5.3. Formato Raster e Formato Vectorial – Comparação

Existem algumas vantagens em usar o formato Raster ou o Vectorial para representar a realidade, dependendo da utilização que se pretende.

A Tabela 2 apresenta algumas características a ter em conta ao optar por um ou outro formato.

Formato Raster	Formato Vectorial
Armazenamento de dados de toda a área coberta.	Armazenamento de dados apenas da área necessária.
Fácil implementação de operações em sobreposição.	Difícil implementação de operações em sobreposição
Visualização dos dados como imagem que, dependendo da resolução, pode ser de fraca qualidade.	Visualização dos dados como gráficos com qualidade Vectorial.
Dependendo da resolução, os dados armazenados podem ser 10 a 100 vezes maiores que os dados Vectoriais.	O tamanho de um ficheiro para dados Vectoriais é tipicamente mais pequeno.
A actualização é mais difícil.	Mais fácil de actualizar e manter.
Não possuem todos os atributos das características.	Suportam maior número de atributos.
-	Permite a combinação de diferentes camadas Vectoriais.
-	Compatíveis com bases de dados relacionais, beneficiando das operações disponíveis sobre bases de dados.

Tabela 2 - Comparação de formatos Raster e Vectorial

2.5.4. Tradução de Raster para Vectorial

A reestruturação da informação pode ser executada por um SIG com o objectivo de a converter para diferentes formatos. Por exemplo, um SIG pode ser usado para converter uma imagem de satélite de um mapa para uma estrutura Vectorial gerando linhas à volta de todas as células com a mesma classificação, enquanto determina a relação espacial da célula.

O processamento de dados mais avançados pode ocorrer com o tratamento de imagens, uma técnica desenvolvida nos finais dos anos 60 pela NASA e pelo sector privado para disponibilizar o realce de contrastes e uma variedade de outras técnicas, que incluem a utilização de transformadas de Fourier [StateMaster, 2007].

Visto que os dados digitais são capturados e armazenados de diversas formas, as duas fontes de dados podem não ser inteiramente compatíveis. Assim, um SIG deve ter a capacidade de converter informação geográfica de uma estrutura para a outra.

2.6. Tipos de Informação num Mapa

Qualquer mapa digital tem capacidade de armazenar mais informação que um mapa em papel que mostre a mesma área, mas geralmente não é claro à primeira vista que tipo de informação o mapa inclui. Por exemplo, há geralmente mais informação disponível num mapa digital do que aquela que se visualiza no ecrã. É necessário compreender que tipos de dados estão incluídos no mapa de modo a que possam ser utilizados apropriadamente.

A Tabela 3 descreve os tipos de informação tipicamente existentes em mapas digitais [GeoComm.com, 2009]:

Tipo de Informação	Descrição
Geográfica	Disponibiliza a posição e as formas de características geográficas específicas
Sobre os Atributos	Disponibiliza informação não geográfica adicional sobre cada característica
De visualização	Descreve como as características irão aparecer no ecrã

Tabela 3 - Tipos de Informação em Mapas Digitais

Alguns mapas digitais não incluem estes três tipos de informação. Por exemplo, geralmente os mapas Raster não incluem informação dos atributos e muitas fontes de dados Vectoriais não incluem informação de visualização.

Detalham-se em seguida os três tipos de informação:

2.6.1. Informação Geográfica

A informação geográfica num mapa digital disponibiliza a posição e a forma de cada característica do mapa. Por exemplo, a informação geográfica de um mapa de estradas é a localização de cada estrada no mapa.

Num mapa Vectorial, a posição da característica é normalmente expressa como um par XY ou uma tripla XYZ, utilizando o sistema de coordenadas definido para o mapa. Muitos sistemas de informação geográfica Vectoriais suportam os três objectos geométricos fundamentais:

- Pontos: Um par de coordenadas simples;
- Linha: Dois ou mais pontos numa sequência específica;
- Polígono: Uma área delimitada por uma linha.

Alguns sistemas também suportam entidades mais complexas, como regiões, círculos, elipses, arcos e curvas [GeoComm.com, 2009].

2.6.2. Informação sobre os Atributos

Os atributos descrevem características específicas do mapa, mas não são inerentemente gráficos. Por exemplo, um atributo associado a uma estrada pode ser o nome ou a data em que foi pavimentada a última vez. Os atributos são frequentemente armazenados em bases de dados separadas da porção gráfica do mapa. Os atributos pertencem apenas aos mapas Vectoriais; raramente são associados a imagens Raster [GeoComm.com, 2009].

Os SIGs mantêm hiperligações internas que unem a entidade gráfica do mapa à sua informação de atributos. A natureza dessas hiperligações varia conforme os sistemas. Em alguns sistemas os links estão implícitos e o utilizador não tem qualquer controle sobre isso. Noutros sistemas existem links explícitos que o utilizador pode modificar. Os links nestes sistemas tomam a forma de chaves da base de dados. Cada característica do mapa tem uma chave armazenada. Esta chave identifica o registo específico da base de dados que contém os atributos de informação da característica. Se surgirem problemas, é importante saber como o software estabelece e mantém as hiperligações de atributos [GeoComm.com, 2009].

2.6.3. Informação de visualização

A informação de visualização num mapa digital descreve como o mapa é visualizado ou parcelado. Informação de visualização comum inclui a cor das características, a grossura das linhas e os tipos de linhas (sólida, pontilhada, tracejada, simples ou dupla), como os nomes das estradas e outras características são mostradas no mapa, e se os lagos, parques ou outras áreas são codificadas com cores.

Muitos conjuntos de dados não incluem informação de visualização. Por exemplo, os ficheiros USGS Digital Line Graph não disponibilizam informação de visualização nenhuma. Cada característica contém um atributo que descreve a entidade mas não indica as características de visualização. Os utilizadores e o seu SIG tem de interpretar esses atributos e decidir como cada um irá ser visualizado.

2.7. Os Sistemas de Informação Geográfica

Um SIG é uma ferramenta de análise de informação geográfica. Nesta secção foram abordados conceitos essenciais que são tomados em consideração na integração dessa informação a ser utilizada em SIGs.

As questões relacionadas com a interoperabilidade entre sistemas geográficos com modelos de dados diferentes revestem-se de importância quando se fala de troca de informação entre sistemas que utilizam formatos diferentes. Um sistema que pretenda utilizar informação de outro sistema, guardada em outro formato diferente daquele que utiliza, tem que forçosamente garantir a interoperabilidade entre esses sistemas, sob pena de não poder reutilizar os dados.

A diversidade nos formatos de dados origina-se frequentemente na forma como são capturados. As diferentes formas de captura de dados geográficos dependem não só da forma como a captura é feita, mas também do formato escolhido pelo utilizador que efectua a captura e armazenamento no formato digital.

A disponibilização de diferentes formatos conduz por um lado à necessidade de conversão da informação para formatos diferentes daquele em que originalmente foram gravados, mas também a necessidades de edição da informação geográfica sempre que se pretenda acrescentar informação geográfica, como acontece no caso dos ficheiros Vectoriais.

Os tipos de formatos são também tomados em consideração, pois dependendo do formato em que está armazenada a informação, a utilização também variará. Não é possível obter atributos geográficos de um ficheiro Raster, por exemplo. No entanto, é possível converter a informação no formato Raster para o Vectorial. Assim, trabalhar com SIGs obriga necessariamente a conhecer ambos os tipos de formatos geográficos com que é necessário lidar, os formatos Raster e os formatos Vectoriais.

A troca de dados entre sistemas exige ainda que se conheçam as escalas utilizadas nos ficheiros geográficos. Se um SIG tiver disponíveis para trabalhar dois ficheiros com informação geográfica com *datums* diferentes, a utilização da informação dos ficheiros necessita ser adaptada, de modo a que o SIG possa trabalhar a informação geográfica recorrendo a dados geograficamente compatíveis.

Conhecer o funcionamento dos SIGs facilita a compreensão do processo de integração de informação geográfica armazenada em diferentes formatos. O próximo capítulo aborda um conjunto de arquitecturas analisadas, que integram informação geográfica utilizando diferentes abordagens.

3. REPRESENTAÇÃO DE INFORMAÇÃO GEOGRÁFICA

Existem diversas razões para o aumento do interesse na integração de informação proveniente de diferentes fontes. O número de SIGs tem vindo a aumentar e consequentemente as bases de dados de suporte a estes SIGs aumentam à mesma velocidade. Estas bases de dados são a componente mais dispendiosa do custo total dos SIGs. Assim, partilhar esta informação entre sistemas de diferentes entidades, agências e empresas é uma das formas de reduzir custos e aumentar a quantidade e qualidade da informação [Moyer, Niemann, 1993].

Numa primeira fase, logrou-se conhecer o mundo dos SIGs, os seus conceitos e as suas metodologias. A fase seguinte descreve a investigação de trabalhos relacionados no campo da integração de informação geográfica provenientes de diferentes fontes e armazenada em diferentes formatos, e na análise do trabalho de normalização dos formatos de transferência de dados entre SIGs diferentes.

3.1. Integração de Informação Geográfica

A integração de informação geográfica proveniente de diferentes fontes, a ser utilizada num mesmo suporte, implica necessariamente interoperabilidade entre os dados provenientes dessas diferentes fontes [Reichardt, 2004].

Como já referido, a interoperabilidade significa abertura ou a capacidade de trocar informações livremente entre sistemas. O OGC define a interoperabilidade como a capacidade de comunicar, executar programas, ou transferir informação entre várias unidades funcionais de uma forma que exija que um utilizador tenha pouco ou nenhum conhecimento das características únicas dessas unidades.

A inexistência de interoperabilidade impede a partilha de informação e de recursos informáticos, levando a que as organizações gastem mais do que o necessário em dados, software e hardware [Reichardt, 2004].

Quando falamos de informação geográfica, falamos de diferentes tipos de informação. Informação sobre a localização, sobre o formato dos objectos geográficos, sobre as relações existentes entre diferentes características e fenómenos geográficos. Isso torna a informação geográfica complexa, fundamentalmente porque existem também

diferentes sistemas de geoprocessamento, com a função de criar, armazenar, processar e mostrar informação geoespacial nos formatos que a suportam.

Na comunidade geoespacial, o significado de interoperabilidade mantém-se ambíguo, tal como os benefícios de ser interoperável. Assim, é sugerido o seguinte mandato:

“Para um sistema ser interoperável, a equipa que o gere deve estar activamente empenhada no processo contínuo de garantir que os sistemas, procedimentos e cultura de uma organização são geridos de forma a maximizar oportunidades para a troca e reutilização de informação e serviços, seja interna ou externamente” [OGC, 2005].

As especificações de interfaces abertas possibilitam aos fornecedores de conteúdos, programadores e integradores focar-se em entregar produtos e serviços mais potentes para consumidores em menos tempo, com menor custo, e mais flexíveis [OGC, 2005].

O OGC vem fornecer algumas ajudas ao se constituir como um organismo de gestão de normas que facilitam a interoperabilidade.

3.2. Tipos de Formatos de Ficheiros

A informação geográfica gerida pelos SIGs pode ser simplesmente descrita como informação geográfica à qual podem ser ligados atributos. Informação geográfica é toda a informação que se relaciona com espaço, como por exemplo, uma estrada ou um lago. As características geográficas da informação podem ser descritas como uma matriz (dados Raster), ou como pontos, linhas e polígonos (dados Vectoriais).

Os formatos suportados pelos SIGs descrevem a forma como as características geográficas foram gravadas. É muito importante considerar o formato da informação de um SIG porque os sistemas raramente suportam todos os tipos de formatos. Ao pretender utilizar informação geográfica gravada num formato particular não suportado pelo SIG em uso, será necessário ou transformar os dados ou simplesmente utilizar outro SIG.

Quase todos os SIGs têm os seus próprios formatos de armazenamento da informação. Estes formatos foram criados para otimizar a eficiência do próprio programa, e não foram desenhados para ser utilizados em outros programas. No entanto, muitos SIGs suportam outros formatos, disponibilizando funções de importação e exportação de conjuntos de dados. Estas funções estão geralmente bem documentadas e normalizadas [GISCentrum, 2009].

3.2.1. Formatos Vectoriais

Muitos SIGs baseiam-se em tecnologia Vectorial. Estes métodos são mais complexos pois há diversas formas de gravar coordenadas, atributos, estruturas de dados e informação de visualização.

A Tabela 4 apresenta alguns formatos de ficheiros Vectoriais de SIGs mais comuns [GISCentrum, 2009 e GeoComm.com, 2009].

Formato	Nome	Descrição
ARCEXport	ARCEXport	Formato de transferência da ESRI, em ASCII ou comprimido para um binário, utilizado para transferir ficheiros entre diferentes versões do ARCInfo.
DGN	MicroStation Design Files	Formato interno utilizado pela Bentley Systems Inc.'s MicroStation, um programa de CAD (Computer Aided Design). Contém informação de visualização detalhada.
DLG	Digital Line Graphs	Formato de transferência utilizado pelo USGS (United States Geological Survey), que representa informação Vectorial apresentada em mapas impressos em papel.
DWG	Autodesk Drawing Files	Formato interno proprietário do AutoCAD, um software de CAD (<i>Computer-Aided Design</i>). O AutoCAD consegue converter qualquer ficheiro DWG para um ficheiro DXF sem perda de informação gráfica.
DXF	Autodesk Drawing eXchange Format	Formato comum de transferência de dados para informação Vectorial. Inclui informação de visualização bastante completa, e quase todos os programas gráficos o conseguem ler.
E00	ARC/INFO interchange file	Formato de transferência da ESRI utilizado pelo ARCInfo. Permite combinar informação descritiva e espacial para ficheiros Vectoriais e Raster num único ficheiro ASCII.
GML	Geography Markup Language	Normalização XML para trocar e gravar informação geográfica Vectorial, com um Schema XML que descreve <i>schemas</i> aplicativos, bem como o transporte e armazenamento de informação geográfica.
KF85	Kommunförbundets transfereringsformat (ISOK)	Armazena tanto características geométricas como atributos. Consiste num ficheiro físico que pode ter qualquer extensão. A extensão .k85 é adicionada ao nome base do ficheiro KF85 quando escreve o ficheiro.
MIF/MID	MapInfo Interchange Format	Formato de transferência, utilizado pela MapInfo, possível de ser lido também por outros SIGs.
SDTS	Spatial Data Transfer System	Formato de transferência desenvolvido pelo governo norte-americano. Pode ser tanto binário como em ASCII, mas geralmente é binário. Pode codificar praticamente todos os conceitos geográficos.
SHP	ESRI Shapefile	Formato de transferência desenvolvido pela ESRI. Armazena informação geométrica não topológica e informação de atributos para as características espaciais de um conjunto de dados.
SVG	Scalable Vector Graphics	Normalização XML para apresentação de dados Vectoriais na internet. É aprovado pelo W3C (World Wide Web Consortium).
TIGER	Topologically Integrated Geographic Encoding and Referencing Files	Formato de transferência em ASCII utilizado pelo Gabinete de Censos Norte-Americano (US Census Bureau) para armazenar os mapas de estradas construídos para o censo de 1990. É baseado em linhas, não em polígonos.

Tabela 4 - Exemplos de Formatos Vectoriais

3.2.2. Formatos Raster

Os ficheiros Raster são geralmente utilizados para gravar imagens, como mapas em papel digitalizados, fotografias digitais ou imagens de satélite. São ainda utilizados para gravar variáveis que variam continuamente no espaço, como topografia e temperatura. Imagens de satélites ou outras aeronaves são conhecidas como informação detectada remotamente. A resolução deste tipo de informação e de mapas digitalizados refere-se à área do terreno coberta por um pixel, ao contrário de outras imagens, em que a resolução é dada em *dots per inch* (dpi).

A Tabela 5 apresenta alguns formatos de ficheiros Raster de SIGs mais comuns [GISCentrum, 2009].

Formato	Nome	Descrição
ADRG	Arc Digitized Raster Graphics	Formato criado pelos militares americanos para gravar mapas em papel no formato Raster.
BIL	Band Interleaved by Line	É um formato Computer Compatible Tape (CCT) que armazena todas as bandas de informação detectada remotamente num ficheiro de imagem.
BIP	Band Interleaved by Pixel	Cada linha de uma imagem é armazenada sequencialmente (na linha 1 todas as bandas, na linha 2 todas as bandas, etc.).
BSQ	Band Sequential	É um formato Computer Compatible Tape (CCT) que armazena cada banda de dados de satélite num ficheiro de imagem para todas as linhas digitalizadas num array de imagens. Os cabeçalhos CCT são gravados em cada banda.
DEM	Digital Elevation Model	Formato criado pelo USGS (US Geological Survey) para gravar informação de elevação. Os valores das células no DEM representam a elevação para essa posição na superfície da terra.
*.dem, *.hdr	DEM ArcInfo	Formato de dados de elevação do ArcINFO.
GTOPO30	Global 30 Arc Second Elevation Data Set	Modelo global digital de elevação com células horizontais de aproximadamente 1 km (30 segundos). Foi criado a partir de diferentes fontes Raster e Vectorial.
GeoTIFF	GeoTIFF	Formato do tipo TIFF (<i>Tag Image File Format</i>) para dados Raster georreferenciados.
GRIB	GRid In Binary	Normalização para dados meteorológicos baseados em grelhas, utilizado pelo WMO (<i>World Meteorological Organization</i>).
PCX	PC Paintbrush Exchange	Formato Raster comum encontrado em muitos programas gráficos e de digitalização.
SDTS	Spatial Data Transfer Standard	Formato para transferir informação geográfica. Uma variante do SDTS foi feita especificamente para transferir dados Raster.
TIFF	Tagged Image File Format	Formato produzido por programas de desenho e digitalizadores. Comprime os dados sem perda de informação.

Tabela 5 - Exemplos de Formatos Raster

3.3. Integração de Formatos de Ficheiros Geográficos

Ao longo dos anos foram desenvolvidas várias propostas de integração de ficheiros de diferentes formatos, umas para dar resposta a problemas de realidades específicas, outros para tentar responder de um modo geral a diferentes realidades, mas apresentando soluções adaptáveis a vários cenários.

Estudam-se em seguida algumas soluções desenhadas.

3.3.1. DPLAN – Sistema de Planeamento de Distribuição

Em 2007, a EDP apresentou uma proposta de arquitectura que visava integrar um SIG com um sistema de planeamento de distribuição, designado por DPLAN [Mira et al, 2007].

A EDP adoptou um Sistema de Informação Técnica, o SIT, e decidiu que este seria o repositório central de informação de rede, com recurso à tecnologia GE Smallworld GIS. O SIT visava integrar a informação proveniente de vários sistemas utilizados internamente.

Após integrar com sucesso os vários sistemas com o SIT, a EDP decidiu avançar com a integração do DPLAN, o que deu origem à proposta de arquitectura em análise.

3.3.1.1 DPLAN: O Problema

Após analisar a utilização do DPLAN, a EDP concluiu que a informação a integrar era de dois tipos:

- Estudos complexos: estudos mais demorados, que utilizam funções avançadas do DPLAN;
- Estudos simples: estudos de configurações de rede que culminam em desenho e propostas a clientes.

Nos estudos complexos, o resultado do estudo aprovado era carregado no SIT como um novo projecto manualmente, o que exigia que o estudo fosse carregado em dois sistemas diferentes.

Para os estudos mais simples, era necessário trabalhar com informação actualizada, o que exigia que fossem feitas frequentes actualizações de dados a partir do SIT.

3.3.1.2 DPLAN: Desenho da solução

Decidiu-se então fazer um desenvolvimento faseado:

- Fase 1: desenvolvimento de ferramentas de importação e exportação entre o SIT e o DPLAN;
- Fase 2: o DPLAN integrado seria executado em background no ambiente SIT.

3.3.1.3 DPLAN: A Implementação

Na primeira fase do desenvolvimento, foi desenvolvida uma solução que substituiu a utilização de uma ferramenta de exportação, o DINIS, que servia de intermediário entre o GIS utilizado e o ficheiro DPX, o formato interno do DPLAN, utilizado depois pelo DPLAN, como mostra a Figura 6.

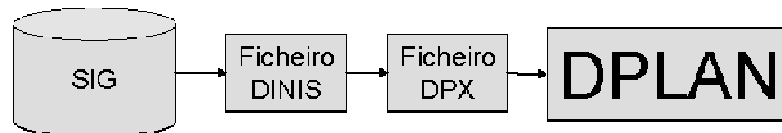


Figura 6 - DPLAN – Processo de integração anterior

Dado que o DPLAN apresentava alguns problemas no tratamento de alguns tipos de dados do SIT e na exportação unívoca de dados, eliminou-se a utilização do DINIS, passando a interface do DPLAN implementada no SIT a suportar o mapeamento exaustivo dos objectos e atributos SIT para os objectos e atributos do DPLAN, como representado na Figura 7.

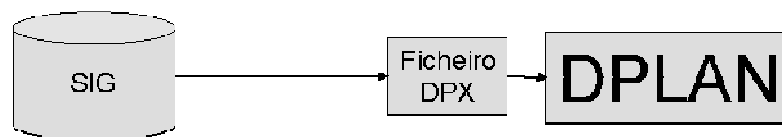


Figura 7 - DPLAN – Novo processo de integração

A interface do DPLAN implementada no SIT inclui também o mapeamento inverso, de objectos e atributos do DPLAN para os objectos e atributos SIT. Foi incluído um atribuído na interface, o atributo *status*, incluído em todos os objectos, com opções para instalar ou remover o objecto.

Na segunda fase de desenvolvimento, foi integrado uma versão do DPLAN no SIT, que visava essencialmente disponibilizar funcionalidades específicas na versão integrada que permitissem a comunicação entre o DPLAN e o Smallworld através da troca de comandos, em alguns casos recorrendo a uma janela nativa para os executar.

3.3.2. SIG Empresarial – Orientação a Serviços

Um SIG Empresarial é uma abordagem de implementação, operação e gestão de um SIG ao nível de uma organização. Disponibiliza acesso a informação geoespacial partilhada e recursos analíticos para um grande número de utilizadores concorrentes localizados em diversas partes de uma instituição [Ghosh et al, 2006].

Um SIG Empresarial pode ser descrito como um esforço para desenhar técnicas de gestão integrada de informação geoespacial para dar resposta a uma instituição complexa.

A arquitectura em análise propõe uma solução orientada a serviços utilizando serviços *web* na integração de diversos repositórios de dados espaciais.

A arquitectura utiliza as normas OGS, nomeadamente, o Web Map Service (WMS) e o Web Feature Service (WFS) para permitir o acesso centralizado a informação espacial em diferentes formatos.

3.3.2.1 SIG Empresarial: O Problema

Num cenário empresarial, em que diferentes tipos de informação são geridos com recurso a diversos sistemas, a tarefa de partilha de aplicações e de informação é difícil, senão impossível.

A gestão de informação geoespacial não é excepção, e a integração de dados existentes em diversos repositórios de informação espacial apresenta-se dificultada por este cenário.

3.3.2.2 SIG Empresarial: Desenho da solução

É proposta uma arquitectura orientada a serviços, utilizando serviços *web* para a integração dos dados espaciais provenientes de diferentes repositórios.

Foram contemplados dois serviços *web* propostos pelo OGC, o WMS e o WFS, como já referido. Estes serviços são utilizados para permitir o acesso centralizado a informação espacial em diferentes formatos.

O WFS permite que um cliente tenha acesso a dados geoespaciais a partir de múltiplos WFS. O WMS tem a capacidade de criar e mostrar mapas simultaneamente provenientes de várias fontes, em formatos normalizados, como .svg, .png, .gif ou .jpg.

3.3.2.3 SIG Empresarial: A Implementação

Para permitir o acesso centralizado à informação proveniente de várias fontes, cada instituição tem de registar os seus dados num registo central, o que permite manter um catálogo de dados. O catálogo deve conter informação sobre o tipo de dados, informação sobre características, etc. Ao ser feito um pedido, os dados são procurados

no catálogo, e o pedido é enviado para o módulo de gestão dos pedidos dessa entrada no catálogo. Independentemente do formato dos dados enviados para o sistema (dados de entrada), os dados que o sistema gera (dados de saída) são enviados no formato GML para assegurar a interoperabilidade.

A concretização desta arquitectura foi feita em Java, com integração dos serviços WFS e WMS do OGC. O objectivo é disponibilizar maior interoperabilidade geográfica recorrendo para isso às normas OGC e baixando as barreiras na integração de dados geográficos disponibilizados pelos fornecedores.

Os dados no pedido feito ao serviço WFS estão no formato GML. O pedido ao WMS mostra os dados gerados graficamente, ou seja, mostra o resultado num mapa.

O trabalho proposto nesta arquitectura integra dados num formato “flat file”, como o formato GML, e no formato relacional, como o formato do Oracle Spatial. Com esta abordagem, os dados podem ser publicados como mapas ou imagens, usando o WMS, e como dados reais, usando o WFS. Foca-se na facilidade de utilização do sistema e na utilização de formatos abertos, de modo a permitir a qualquer sistema partilhar rapidamente a sua informação espacial de forma interoperável.

3.3.2.4 SIG Empresarial: A Arquitectura

A arquitectura do sistema é mostrada na Figura 8. A integração dos repositórios de dados é feita com recurso às normas *web* WFS e WMS. O núcleo do sistema é um registo central que contem a informação de todos os repositórios de dados a partir dos quais é possível obter dados.

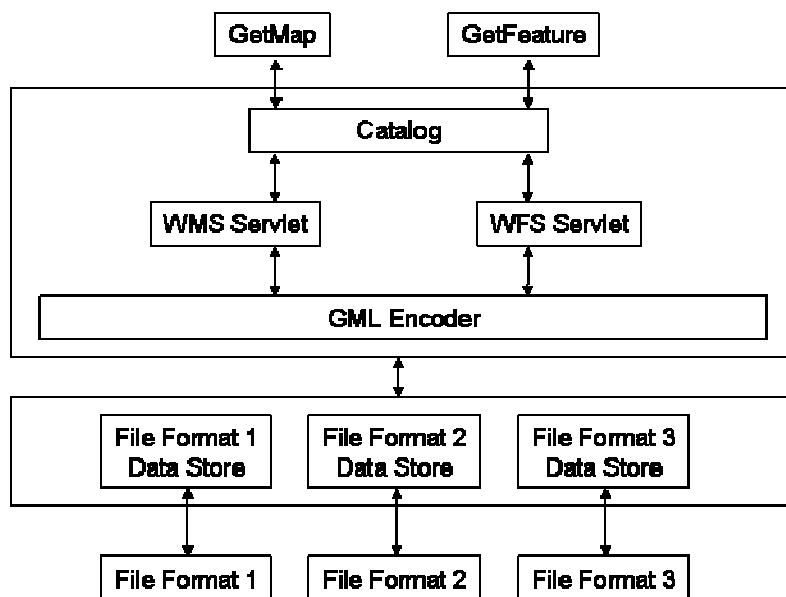


Figura 8 - SIG Empresarial – Arquitectura do sistema

Todos os fornecedores de dados precisam registrar os seus dados no registo central. Os dados são registados num catálogo, que é um ficheiro XML com a informação da estrutura dos dados, como o tipo de dados (shapefile, gml, etc.), o nome da característica, o *namespace* dos dados, etc. A Figura 9 representa um excerto do ficheiro de registo.

```
<datastore namespace="cite" enabled="true" id="ponds">
  <connectionParams>
    <parameter value="file:data/featureTypes/cite_Ponds/Ponds.sho" name="url" />
  </connectionParams>
</datastore>
<datastore namespace="soi" enabled="true" id="states_shapefile">
  <connectionParams>
    <parameter value="file:data/featureTypes/states/states.sho" name="url" />
  </connectionParams>
</datastore>
```

Figura 9 - SIG Empresarial – Parte do ficheiro XML de registo

Os dados também podem estar em sistemas de gestão de bases de dados relacionais. O Oracle é considerado como preferencial neste ponto dadas as suas capacidades de armazenamento de dados espaciais.

Por cada característica (feature), é armazenado também um ficheiro de estilo que é usado para formatar a visualização dos dados espaciais. O cliente pode aceder aos dados fazendo o pedido através da funcionalidade `getFeature`. O pedido é então enviado para o servidor como um GET ou um POST. O pedido inclui:

- O endereço do servidor;
- O tipo de pedido; neste caso, seria “getfeature”;
- O tipo de serviço; neste caso, seria “wfs”;
- A versão;
- O nome do tipo de dados a obter; por exemplo, “estradas”.

O pedido de mapas permite solicitar dados como parâmetros de formatação adicionais, como altura, largura, tipo de imagem, etc. O pedido inclui:

- O endereço do servidor;
- Dados do objecto a formatar; por exemplo, as dimensões de um rectângulo;
- O estilo a aplicar;
- O formato da imagem; por exemplo, image/gif;
- O tipo de pedido, neste caso, seria “GetMap”;
- As camadas;
- A altura e largura;
- O SRS.

A Figura 10 mostra o resultado de um pedido de dados provenientes de uma única fonte. Os dados também podem ser obtidos a partir de várias fontes. Dado que os dados provenientes de várias fontes estão disponíveis como características geográficas, podem ser especificadas várias.

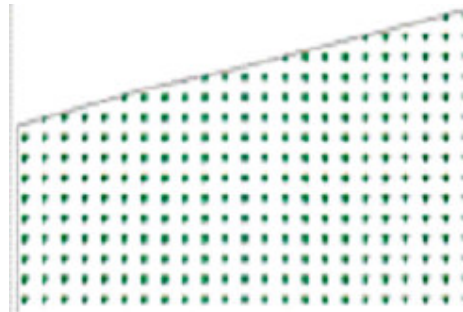


Figura 10 - SIG Empresarial – Pedido GetMap a partir de uma única fonte

A Figura 11 mostra os dados provenientes de diferentes fontes e o resultado correspondente.

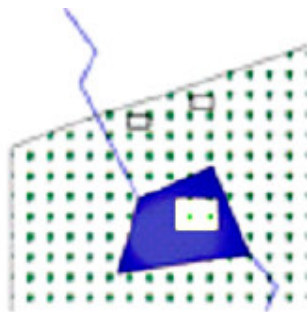


Figura 11 - SIG Empresarial – Pedido GetMap a partir de várias fontes

3.3.3. ATIS – Sistema de Controlo de Tráfego

O ATIS, Advanced Traveler's Information System, é um sistema de controlo de tráfego em tempo real, que usa a internet para disseminar esta informação. As diferentes partes do sistema são integradas através de um SIG disponível através da internet. A disseminação da informação de tráfego em tempo real é conseguida ligando o ATIS a sistemas de monitorização de tráfego, armazenando a informação em bases de dados e visualizando-a em cima da rede de estradas actual [Kotzinos, Pratacos, 2004].

3.3.3.1 ATIS: O Problema

Numa definição lata, o ATIS inclui qualquer tipo de sistema que forneça informação sobre tráfego e sobre viagens de uma forma organizada e consistente.

De forma a fornecer informação fiável e precisa, o ATIS tem primeiro de obter os dados a partir de diferentes fontes, organizá-los e apresentá-los aos utilizadores de forma

também consistente e intuitiva. Esta questão levanta o típico problema da interoperabilidade, já que os formatos dos dados em cada fonte de informação são geralmente diferentes. O ATIS tem de adaptar os dados de forma transparente para o utilizador. Os dados podem ser de natureza estática ou dinâmica e podem ser de vários tipos.

Os dados estáticos referem-se a dados que não mudam ao longo do tempo. Estes dados estão geralmente relacionados com a topologia ou a geografia da cidade ou da área em análise. São listados os dados que se enquadram nesta categoria:

- A rede de tráfego da cidade;
- A posição e tipo de sinais de trânsito;
- As luzes de trânsito;
- Informação das Páginas Amarelas;
- Os lugares de estacionamento.

Os SIGs permitem armazenar e manipular este tipo de dados, dado que conseguem lidar com bases de dados espaciais e todos estes tipos de dados são georreferenciados. Um SIG armazena os dados em diferentes camadas, permitindo assim a separação lógica dos diferentes tipos de informação. Uma topologia de rede que distingue entre cruzamentos, outros nós, direcções, etc., pode facilmente ser integrada num SIG.

Os dados dinâmicos são dados em tempo real relacionados com o tráfego. Ou são actualizados em intervalos regulares pré-definidos (*time-driven*), ou são actualizados após a ocorrência de um evento previamente especificado (*event-driven*). A lista deste tipo de ocorrências inclui:

- Dados de tráfego em tempo real, como o tempo, a velocidade, o nº de veículos para cada trecho de estrada da rede;
- Estado em tempo real das luzes de trânsito;
- Incidentes;
- Lugares de estacionamento disponíveis em parques de estacionamento;
- Relatórios de meteorologia;
- Informação de trânsito em tempo real como a localização de autocarros, hora de chegada aproximada às paragens de autocarros.

São necessários sistemas de informação em tempo real para gerir os dados dinâmicos.

3.3.3.2 ATIS: Desenho da solução

Pretendeu-se nesta arquitectura aumentar as capacidades de um SIG disponível na internet de forma a permitir a gestão em tempo real de dados tanto despoletados pelas actualizações regulares, como por ocorrências de eventos.

Definiu-se que o ATIS teria de ir ao encontro de três desafios de modo a ser considerado completo e de confiança. São eles:

- Gerir as diferentes fontes de dados de forma transparente para o utilizador final;
- Operar em tempo real;
- Operar num ambiente heterogéneo e cooperar com sistemas de gestão de tráfego diferentes e por vezes proprietários.

O acesso a sistemas de informação *web based* tem de ser conseguido independentemente da localização e dos meios, dentro de determinados limites.

3.3.3.3 ATIS: A Implementação

Para responder às especificações foi desenhada uma arquitectura modular e aberta.

A arquitectura possui as seguintes características:

- É aberta e pode ser facilmente expandida para permitir alterações de infraestrutura;
- É modular, o que significa que partes da arquitectura podem ser substituídas com relativa facilidade, se necessário;
- Disponibiliza um meio uniformizado para os utilizadores acederem à informação;
- Permite a personalização do nível de detalhe em que é visualizada;
- Disponibiliza informação em tempo real;
- Funciona com um mínimo de informação disponível e gere inclusivamente circunstâncias de ausência de informação;
- É compatível com qualquer infraestrutura de recolha de dados de tráfego.

O sistema *web based* é dividido em três módulos básicos, cada um composto por vários componentes. Os módulos estão representados na Figura 12, e representam as operações lógicas que tem de ser executadas para garantir a funcionalidade do sistema. São eles:

- Módulo de personalização e manipulação pelo utilizador;
- Módulo de visualização de informação;
- Modulo de interface com a infraestrutura existente.

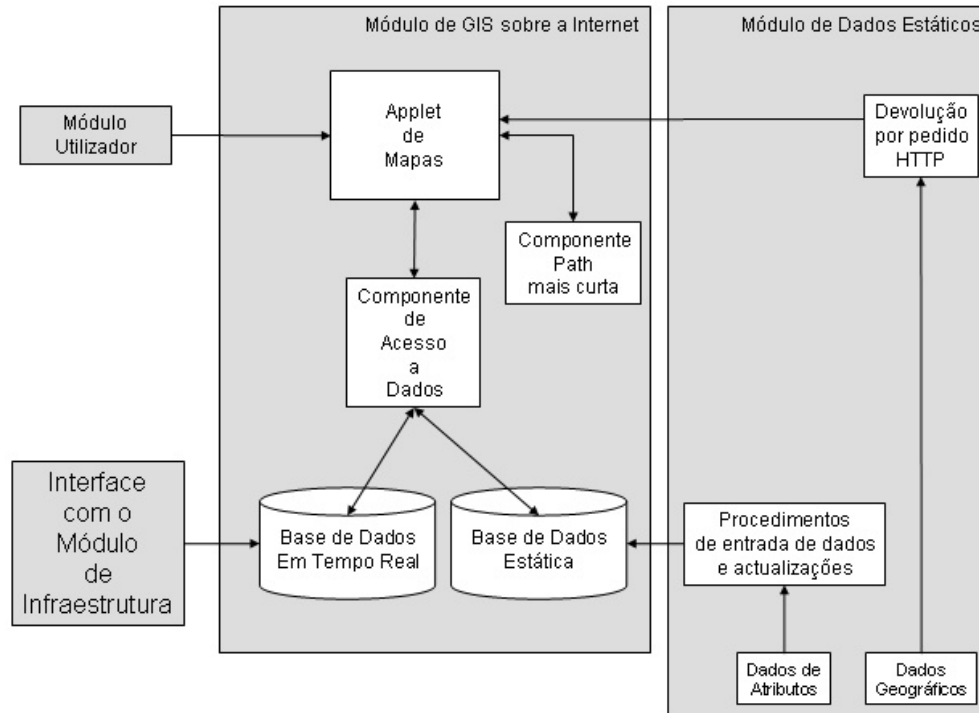


Figura 12 - ATIS – Arquitectura proposta

O módulo de personalização e manipulação pelo utilizador gere as operações relacionadas com o registo e personalização do site pela primeira vez. O utilizador pode escolher a quantidade e a qualidade da informação que pretende receber quando está ligado ao site. O sistema devolve para cada utilizador o perfil único correcto anteriormente armazenado. O perfil único contém informação sobre as suas preferências, as suas viagens, o tipo de informação em que esta interessado, etc., e é criado quando o utilizador acede ao site pela primeira vez.

O módulo de visualização de informação baseia-se na utilização de um SIG disponível na internet. É usado para obter a informação pedida a partir de varias bases de dados e apresentá-la como um mapa dentro de um *web browser*. O SIG age como a ligação entre as várias fontes de informação e os utilizadores, dado que tanto obtém como apresenta a informação disponível.

Este módulo tem cinco componentes:

- Visualizador: a parte do SIG que apresenta o mapa da cidade mostra informação geográfica, como a localização de pontos de interesse, e gere as operações sobre o mapa (zoom, panning, centrar e mover o mapa). Esta componente também funciona como interface com o utilizador que permite fazer pesquisas à base de dados e obter informação do sistema;
- Bases de dados com informação estática: bases de dados com informação não mutável em tempo real;
- Bases de dados com informação em tempo real: bases de dados com contagens de tráfego disponibilizados pelo sistema de gestão de tráfego ou outras infraestruturas, conjuntamente com outra informação disponível em tempo real sobre incidentes como acidentes de trabalhos de estrada, bloqueios de estrada, etc;
- Componente de acesso a dados: este componente gere a interacção entre o SIG e as bases de dados com informação estática e em tempo real;
- Componente de caminho mais curto: inclui o algoritmo que utiliza tanto dados das bases de dados estática e em tempo real, como os dados inseridos pelo utilizador relativos a pontos de origem e destino, para calcular a distância mais curta.

O módulo interface com a infraestrutura existente gere a interface entre o sistema de informação do viajante e a infraestrutura existente que disponibiliza os dados em tempo real. É separada do resto pelo sistema porque é específica ao site. A sua implementação é afectada pelo tipo de infraestrutura que um município tem para obter dados de tráfego. Este módulo recebe os volumes de tráfego e a ocupação das estradas a partir da infraestrutura existente para cada segmento de estrada ou para a parte da rede que está coberta. Passa então as medidas do tráfego para o sistema de informação do viajante. O objectivo deste módulo é concluído após passar as medidas da infraestrutura para as tabelas apropriadas nas bases de dados em tempo real.

3.3.4. DISPRO – *A web based GIS*

O DISMAR é um sistema de integração de dados criado no CMRC (Coastal & Marine Resources Centre, na Irlanda) para gerir problemas relacionados com a poluição das águas nas zonas costeiras europeias. É um acrónimo para Data Integration System for MARine Pollution and Water Quality [Tuama, 2004].

O desenho do sistema inclui um SIG como um componente base de um sistema distribuído, adoptando várias normas, como as do OpenGIS, do W3C ou da ISO, e trabalhando com software de código aberto sempre que possível. Os requisitos do

sistema reflectem as orientações da directiva EU INSPIRE, uma iniciativa que visa incentivar o desenvolvimento de uma infraestrutura de dados espaciais a nível europeu.²

O DISPRO, o primeiro protótipo do DISMAR, é um sistema distribuído *web based* que utiliza as normas WMS da OGC e adopta as directivas da norma ISO 19115, sobre catálogos de metadados.

3.3.4.1 DISPRO: O Problema

O termo SDI (Spatial Data Infrastructure), refere-se à colecção relevante de tecnologias, políticas e directrizes institucionais que facilitam o acesso a, e a disponibilidade de, dados espaciais. A directiva INSPIRE usa o termo análogo Infraestrutura de Informação Espacial, que se refere tanto a questões técnicas como não técnicas, e que englobam normas e protocolos técnicos, questões organizacionais, políticas de dados, como regras de acesso a dados, e a criação e manutenção de informação geográfica para um vasto conjunto de temas.

A directiva INSPIRE definiu alguns princípios em relação ao acesso e utilização de informação espacial no contexto de uma Infraestrutura de Dados Espaciais Europeia. A equipa do DISMAR procurou aderir o mais possível a estes princípios.

A Tabela 6 lista os princípios definidos pela directiva INSPIRE.

ID	Descrição
1	Os dados devem ser obtidos uma vez e mantidos no nível mais eficaz possível
2	Deve ser possível combinar dados espaciais de diferentes fontes da União Europeia e partilhá-los entre vários utilizadores e aplicações
3	Deve ser possível obter dados espaciais num determinado nível de governo e partilhá-los pelos outros níveis
4	Os dados espaciais necessários devem estar disponíveis em condições não restritivas
5	Deve ser fácil descobrir quais os dados espaciais disponíveis, para avaliar a sua adequabilidade a um determinado propósito e quais as condições para os utilizar

Tabela 6 - Princípios INSPIRE

O DISPRO procurou implementar estes princípios, sendo que os pontos 1, 2 e 5 foram conseguidos com recurso a normas abertas.

² A directiva INSPIRE baseia-se na infraestrutura de informação espacial estabelecida e operada pelos países membros da União Europeia. Exige que sejam adoptadas regras de implementação comuns em várias áreas especificadas (Metadados, Especificações de dados, Serviços de Redes, Partilha, Monitorização e Reporting de Dados e Serviços), que são adoptadas como Decisões ou Regulações de Comissão, e são vinculativas. Entrou em funcionamento em 15 de Maio de 2007. [<http://inspire.jrc.ec.europa.eu/>]

3.3.4.2 DISPRO: Desenho da solução

O objectivo do DISMAR seria disponibilizar um sistema que respeitasse os princípios delineados na directiva INSPIRE, como os listados na tabela, e outros relacionados com o tratamento de dados considerados essenciais, e que a directiva INSPIRE categoriza em dois grupos: dados básicos e dados temáticos frequentemente utilizados.

Os dados básicos são aqueles que requerem harmonização e partilha por várias áreas. A harmonização é exigida em vários níveis, incluindo no nível lógico (modelos e representações de dados), semântico (características e catalogação), topológico e de perfis de metadados.

Os dados temáticos frequentemente utilizados são conjuntos de dados usados em comum por várias áreas. São dados que apenas necessitam de mostrar consistência ao nível geométrico, ou seja, requerem normalização ao nível das características espaciais (linhas, pontos, polígonos).

O sistema deveria também recorrer à utilização de normas abertas na implementação da solução.

3.3.4.3 DISPRO: A Implementação

O sistema de gestão de dados DISMAR possui um SIG disponível numa estrutura *web* como um dos seus sistemas base. O desenho do sistema segue os princípios INSPIRE citados, e tirou partido dos esforços dos trabalhos do OGC e de projectos em código aberto. Inclui a utilização das normas da OGC, o WMS e o WFS. Utiliza XML para codificar e transportar os dados.

O DISPRO foi construído usando uma combinação de tecnologias de servidor *web* de código aberto (um servidor Apache HTTP, contentor de *servlets* Java Apache Tomcat, e uma aplicação de visualização de mapas certificada pelo OpenGIS, o UMN Server), e ferramentas de processamento Java/XML.

Implementa ainda um serviço *web* de mapas, recorrendo ao WMS da OGC.

A Figura 13 mostra a arquitectura do DISPRO.

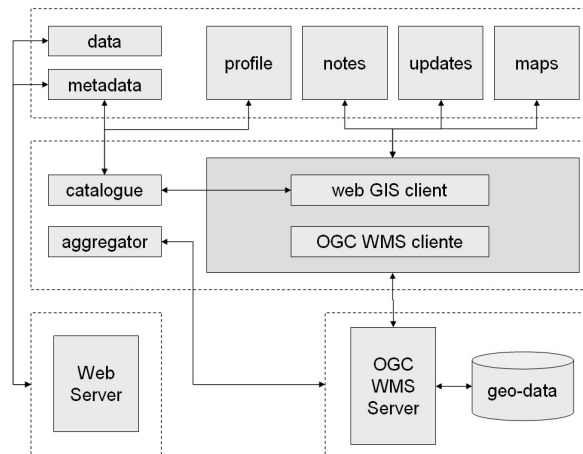


Figura 13 - DISPRO – Arquitectura

O DISPRO implementa o WMS da OGC com oito nós (NERSC, NIVA, Met.no, PML, GKSS, Ifremer, Ecoles des Mines, CMRC), que disponibilizam um vasto conjunto de dados a partir de vários sensores (por satélite, aparelhos aéreos, obtidos no local, etc.). Os nós são integrados através de um portal geográfico no CMRC, que constrói dinamicamente interfaces para os clientes do SIG, que lhes permitirá trabalhar com os dados disponíveis em cada nó.

A arquitectura do DISPRO contempla um sistema multi-camada, em que as várias aplicações-cliente comunicam com os servidores aplicativos, que estão por sua vez ligados aos repositórios de dados.

A Figura 14 mostra o sistema WMS implementado no DISPRO.

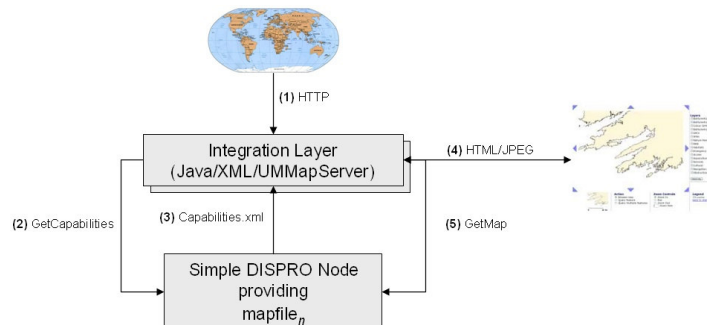


Figura 14 - DISPRO – Implementação do WMS da OGC

Quando um utilizador selecciona uma região (1), é feito um pedido GetCapabilities aos outros nós através da camada de software que contém o portal de integração (2). Os ficheiros Capabilities devolvidos (3) são processados com ferramentas Java/XML no portal de integração, e a interface do cliente SIG é gerada e mostrada (4). As camadas dos mapas estão agora visíveis através do cliente, com a camada de integração a enviar os pedidos GetMap necessários (5).

O sistema WMS implementado consiste em vários componentes:

- A interface *web* para o SIG;
- A aplicação *web* do cliente do SIG;
- O cliente WMS;
- Os servidores WMS;
- E o software que agrega todos estes componentes.

A Figura 15 mostra como o WMS está integrado na arquitectura do DISPRO.

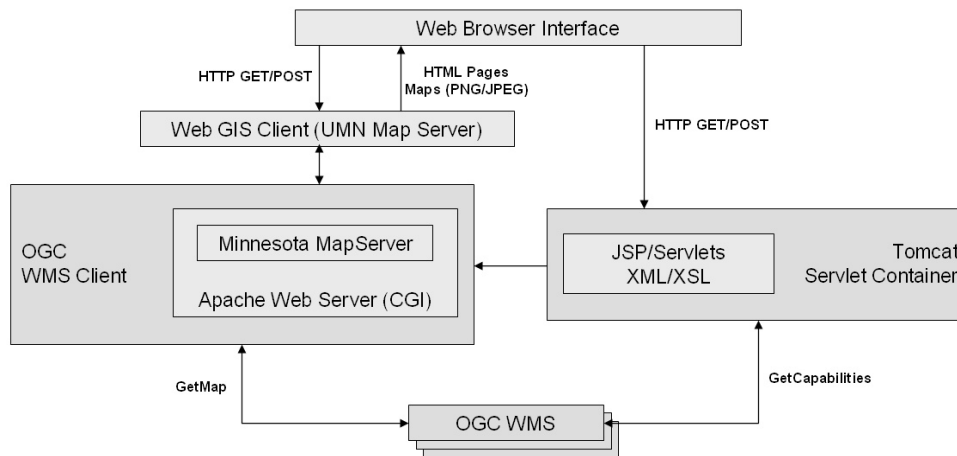


Figura 15 - DISPRO – Integração do WMS

3.3.5. Open Format Converter

O OFC (Open Format Converter) permite a conversão de dados Vectoriais geoespaciais para vários formatos. Define a estrutura para um conversor multifuncional que visa permitir a fácil implementação de formatos para além dos propostos [Mrozek, Fraczek, 2004].

3.3.5.1 OFC: O Problema

Muitas organizações e agências governamentais utilizam formatos especiais para obter dados relacionados com as suas áreas de interesse, como se dá com o formato Digital Line Graph (DLG), utilizado pelo USGS (US Geological Survey), ou o Topologically Integrated Geographic Encoding and Referencing Files (TIGER), utilizado pelo SDTS (US Census Bureau and Spatial Data Transfer System (SDTS)). Muitos formatos foram desenhados para facilitar a troca de informação entre sistemas, e muitos tem sido os esforços de normalização feitos por organizações internacionais no sentido de melhorar a interoperabilidade entre diferentes SIGs.

3.3.5.2 OFC: Desenho da solução

O OFC é uma proposta de arquitectura que define uma estrutura extensível para converter dados Vectoriais espaciais e não espaciais para formatos pretendidos.

A arquitectura do OFC utiliza tecnologia orientada a objectos para armazenar dados em contentores internos universais e um conjunto de interfaces de entrada e saída de dados, utilizadas para aceder à informação e escrevê-la em vários formatos. Fornece também funcionalidades de configuração que permitem controlar o comportamento do processo e definir transformações adicionais para atributos não espaciais.

3.3.5.3 OFC: A Implementação

A arquitectura do OFC inclui um módulo principal, designado de Translation Core, que contém dois módulos: o Geo-Core (Geospatial Core - GC) e a Attribute Matrix (AM).

O GC é o contentor universal para objectos espaciais Vectoriais de duas e três dimensões. A estrutura foi desenhada assumindo que haverá compatibilidade com as especificações do GML 2.0.

A Figura 16 apresenta a arquitectura genérica do conversor.

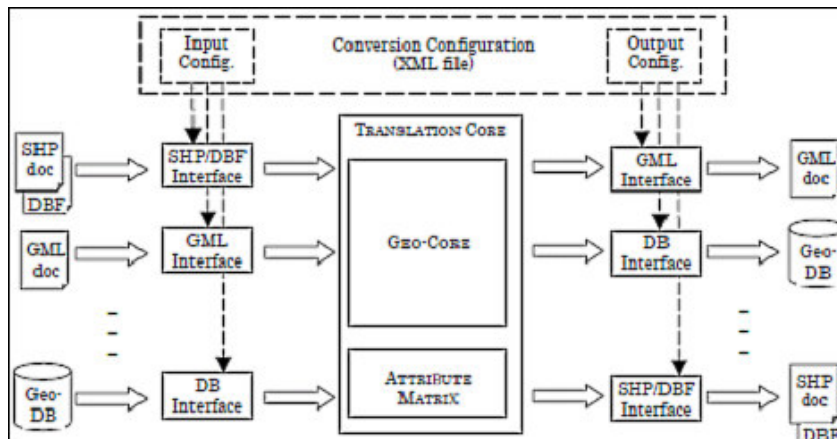


Figura 16 - Open Format Converter– Arquitectura

Apresenta o Translation Core, com os dois módulos que o compõem, o GC e o AM. As interfaces de entrada e saída de dados são representadas como dados de entrada do Translation Core e como dados de saída, representando os dados a converter como dados de input, e os dados já convertidos como dados de output.

A estrutura de objectos que representam os contentores internos universais são, como já referido, orientadas a objectos, e permitem o armazenamento dos dados espaciais, como os pontos, as linhas, os polígonos, e ainda conjuntos de dados espaciais, como multi-pontos, multi-linhas, multi-polígonos, que na prática definem hierarquias de objectos espaciais (por exemplo, multi-polígonos são vários polígonos, que são um conjunto de anéis lineares, que por sua vez são um conjunto de pontos ordenados).

A Figura 17 mostra o diagrama de objectos, representando as dependências de agregação entre as estruturas dos vários objectos armazenados no Translation Core.

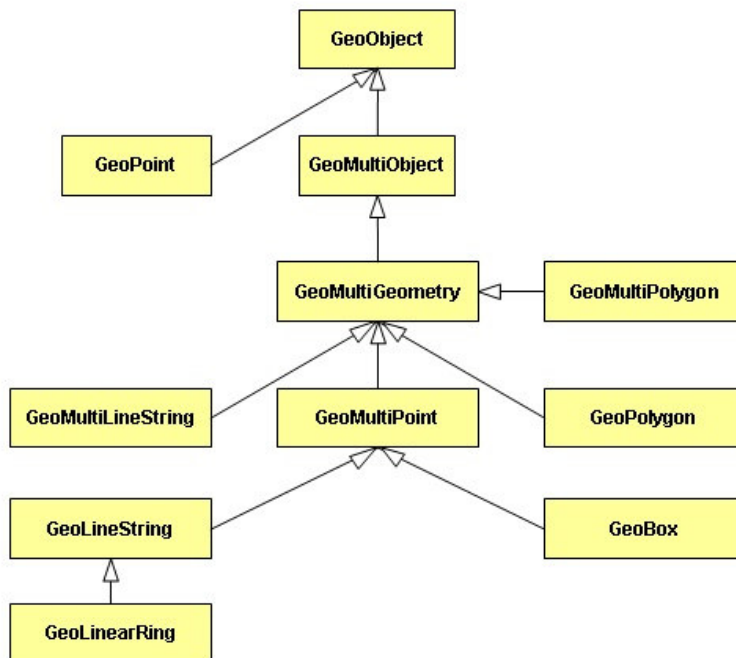


Figura 17 - OFC – Estrutura da herança dos objectos do Geo-Core

O AM é um contentor adicional para dados não espaciais. Este tipo de dados é armazenado na forma de pares atributo-valor. Os atributos representam informação adicional relacionada com um objecto espacial, como o nome de uma rua, a profundidade média de um reservatório de água, etc. O AM permite também armazenar outros tipos de dados, como as propriedades e características de um objecto, ou ainda informação de visualização do objecto, como cores, grossura e estilo das linhas, etc., que especificam como o objecto deve ser visualizado num mapa. Os dados armazenados no módulo AM podem vir directamente dos dados de entrada ou podem ser calculados no momento pelo processo de configuração.

Durante o processo de conversão, o GC armazena um único objecto processado. O AM é carregado com a lista de atributos do objecto em processamento. As interfaces de entrada lêem os dados a partir dos ficheiros ou de bases de dados, e armazenam-nos dentro dos módulos do Translation Core. As interfaces de saída podem também escrever dados para ficheiros ou para bases de dados.

Em muitos casos, o processo de conversão requer configurações adicionais. Isto pode ser feito por adicionar à aplicação a possibilidade de utilizar ficheiros de configuração externos. Para isso, o OFC utiliza ficheiros XML de configuração. É necessário haver um ficheiro de configuração para cada tipo de conversão.

A Figura 18 mostra um exemplo de um ficheiro de configuração.

```
<?xml version="1.0" encoding="UTF-8"?>
<Configuration>
  <General>
    <SourceFormat>Shape</SourceFormat>
    <DestinationFormat>OGC-GML</DestinationFormat>
    <TransferGeometry>true</TransferGeometry>
    <TransferAttributes>true</TransferAttributes>
    <GeoType>POINT</GeoType>
  </General>
  <Input>
    <Dictionaries>
      <Dictionary name="CITIES" type="DBF" file=".\\rivers.dbf"/>
    </Dictionaries>
    <Attribute dictionary="CITIES">
      <SrcAttName>Name</SrcAttName>
      <DstAttName>Name</DstAttName>
      <JoinCondition>Main.CityID = CITIES.ID</JoinCondition>
    </Attribute>
  </Input>
  <Output>
    <PartOfDoc parent="RefineryLayer">HydroMap.xml</PartOfDoc>
    <ExplicitCoords>true</ExplicitCoords>
    <SrsName>UTM zone 10</SrsName>
    <RecNumber>10000</RecNumber>
    <Element name="WaterRefinery">
      <Attribute>
        <SrcAttName>Bio</SrcAttName>
        <DstAttName>Biological</DstAttName>
      </Attribute>
      <Attribute>
        <SrcAttName>Chem</SrcAttName>
        <DstAttName>Chemical</DstAttName>
        <Value>True</Value>
      </Attribute>
    </Element>
  </Output>
</Configuration>
```

Figura 18 - OFC – Exemplo de ficheiro XML de configuração

O XML é composto por três secções principais, dentro do nó raiz, Configuration:

- **General:** contém informação comum sobre o processo, como o formato de origem, o formato de destino e tipos geométricos opcionais. É uma secção obrigatória no ficheiro de configuração;
- **Input:** esta secção é opcional e controla a interface de entrada. Pode ser definido um dicionário suplementar para obter valores de atributos adicionais que não estão disponíveis no conjunto de atributos principal;
- **Output:** esta secção é também opcional e controla a interface de saída. Pode conter instruções sobre como as coordenadas geométricas podem ser apresentadas no destino.

A principal vantagem da configuração avançada é a possibilidade de influenciar e alterar o comportamento do processo de conversão básico implementado. Configurações apropriadas podem alterar as actividades levadas a cabo pelas interfaces de entrada e saída.

3.4. Conclusões sobre as arquitecturas analisadas

A integração de informação geográfica significa tornar os dados espaciais acessíveis a partir de varias tecnologias e fornecedores, e tornar as funcionalidades espaciais e de dados disponíveis para outros sistemas, como CRM, logísticas, serviços baseados em localização para aparelhos sem fios, etc. A principal vantagem é disponibilizar aos utilizadores os meios para aceder, combinar e disseminar informação geoespacial a partir de várias fontes de informação distribuídas. No limite, a integração reduz custos de produção de informação, gestão e distribuição [McKee, OGC, 2005].

A integração é mais eficiente quando acompanhada com interfaces de normas abertas em vez de interfaces proprietárias. Sistemas que utilizam interfaces abertas podem tirar partido dos efeitos de rede que resultam de as mesmas interfaces serem utilizadas fora da empresa [McKee, OGC, 2005].

O objectivo do OGC é criar uma única infraestrutura neutra em termos de tecnologia proprietária para integração da informação, e que funcione em qualquer lado, em várias plataformas, tecnologias e tipos de aparelhos.

As arquitecturas analisadas neste documento permitem chegar a várias conclusões neste campo.

Em alguns casos, apresentam-se soluções para problemas específicos. É o caso da integração do SIG e do DPLAN, implementado pela EDP. Trata-se de um desenvolvimento feito à medida para responder a um problema específico, mas que em termos de desenho de solução contém algumas características de interesse para o estudo da integração da informação.

A parte da arquitectura que interessa realçar é o desenvolvimento de uma interface responsável por fazer o mapeamento dos objectos entre os dois sistemas. Como referido, é feito um mapeamento exaustivo dos objectos de e para cada um dos sistemas e são ainda incluídas no esquema de mapeamento formas de facilmente instalar ou remover o objecto.

Tratando-se de uma solução à medida, dificilmente poderia ser adaptada a um cenário como o que se discute nesta dissertação, de integração de diversos formatos. No entanto, a solução de desenvolvimento de uma interface com o mapeamento dos objectos entre os sistemas apresentada é uma estratégia interessante a ter em conta.

A proposta de [Ghosh et al, 2006], que propõe uma arquitectura de integração de dados de diferentes fontes numa abordagem orientada aos serviços para SIGs Empresariais, realça a importância de recorrer a normas. A arquitectura apresentada utiliza normas OGC amplamente aceites, as normas WMS e WFS, para integrar os dados. No entanto, o objectivo não é especificamente a integração de ficheiros de diferentes formatos, e sim a geração de mapas com informação de várias fontes a partir dos outputs do WMS e do

WFS. O output gerado é sempre no formato GML, o que constitui uma limitação numa arquitectura de integração de diferentes formatos.

O ATIS apresenta uma solução de implementação de uma arquitectura aberta e que permite escalabilidade pelo facto de ser modular.

O DISPRO, embora não foque especificamente a integração de dados Vectoriais, integra dados de diferentes fontes com o objectivo de gerar mapas, recorrendo para isso ao WMS. Não podendo ser utilizado como uma arquitectura genérica de integração, dado não se focar especificamente na integração de ficheiros de diferentes formatos, permite assimilar algumas ideias interessantes a utilizar, como o facto de utilizar normas de integração para processar os dados de input e gerar os mapas como output.

Outra das soluções estudadas é o Open Format Converter. Apresenta uma arquitectura aberta de integração de ficheiros, que permite a conversão de diferentes formatos, e inclui alguns aspectos interessantes na forma como são tratados os dados dentro da arquitectura. É o caso do Translation Core, que permite o tratamento de dados diferentes por módulos diferentes. Os dois módulos que apresenta, o GC, o contentor universal para objectos espaciais, e o AM, o contentor de dados não espaciais, têm responsabilidades distintas, apresentando também alguma modularidade, e sobretudo a distinção no tratamento dos dados de acordo com a natureza dos mesmos. No entanto, com o aparecimento de normas mais recentes, como o WFS e o WMS, alguns componentes desta solução passam a estar obsoletos.

3.5. Identificação de trabalho em aberto

A existência de informação geográfica que referencia os mesmos locais, mas que está dispersa por diferentes ficheiros em diferentes formatos é uma realidade. A informação geográfica utilizado em SIGs é heterogénea, e continua a estar disponíveis em vários formatos e a estar armazenada em diferentes suportes multimédia, conduzindo ao problema da interoperabilidade entre sistemas.

Têm sido desenvolvidos muitos trabalhos nos últimos anos no sentido de desenvolver SIGs mais robustos, e de obter normas que ajudem a ultrapassar o problema da interoperabilidade. O aparecimento das normas da OGC, o WFS e o WMS, vieram facilitar esta tarefa.

No entanto, de forma a ser possível utilizar a informação dispersa, é necessário efectuar um conjunto de tarefas, desde conversão para formatos reconhecidos pelo SIG que se pretende utilizar, até integração da informação de forma manual, ou através de funcionalidades de integração eventualmente disponibilizadas pelo SIG em utilização, e que pode ou não suportar os formatos com que tem de lidar.

Foram analisadas várias soluções de integração de informação, que disponibilizam já algumas abordagens facilitadoras da conversão e até mesmo da integração.

No entanto, falta uma solução de integração que facilite o acompanhamento de todo o processo, desde a análise da informação geográfica dispersa por vários ficheiros, até à disponibilização da informação integrada. São necessárias arquitecturas vocacionadas para sistemas de integração de informação geográfica, que integrem a informação proveniente de diferentes fontes, com diferentes formatos, e que permitam obter novas fontes em outros formatos, como resultado da integração da informação.

O capítulo seguinte apresenta uma proposta de solução para o problema da integração da informação geográfica dispersa por diferentes ficheiros, com diferentes formatos.

4. OPENGISINTEGRATOR – PROPOSTA DE INTEGRAÇÃO DE INFORMAÇÃO GEOGRÁFICA

A agregação de informação geográfica a partir de fontes de informação heterogéneas envolve a realização de um conjunto de operações de integração.

Neste capítulo é apresentada uma proposta de integração da informação geográfica, com recurso a diferentes módulos, responsáveis por efectuar diferentes operações sobre os dados, com o objectivo final de os agregar e apresentar em um ou mais formatos geográficos.

A arquitectura proposta recebeu a designação de OpenGisIntegrator, e a Figura 19 apresenta o diagrama de contexto onde se insere.

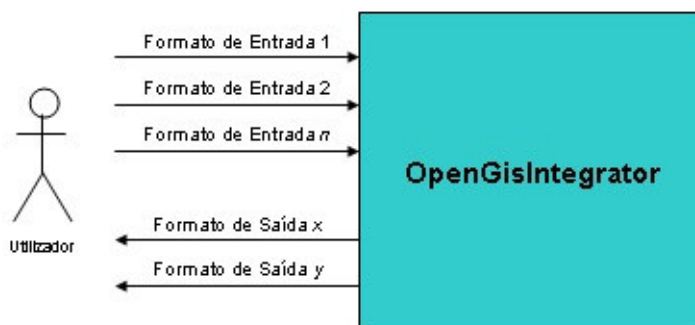


Figura 19 - OpenGisIntegrator - Diagrama de Contexto

O OpenGisIntegrator recebe n ficheiros de entrada, em diferentes formatos, e produz como saída os ficheiros de entrada agregados num único ficheiro, que é disponibilizado nos m formatos de saída definidos.

4.1. Descrição do OpenGisIntegrator

O OpenGisIntegrator é uma arquitectura que recebe como entrada dados Vectoriais disponibilizados em diferentes formatos. Estes dados são convertidos para um formato interno, sendo posteriormente integrados numa única representação interna de dados, que pode ser apresentada como saída em um ou vários formatos Vectoriais, escolhidos pelo utilizador.

A arquitectura genérica do OpenGisIntegrator é apresentada na Figura 20.

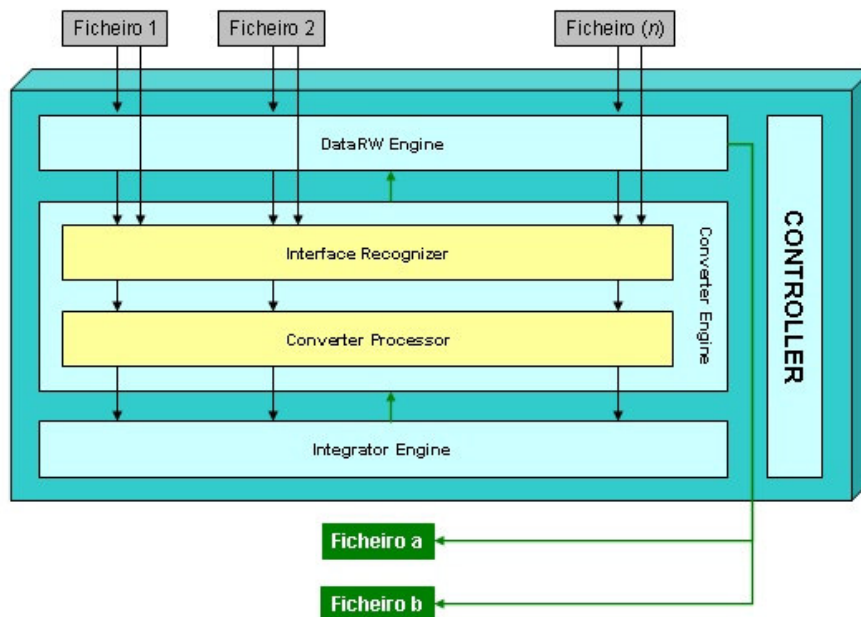


Figura 20 - OpenGisIntegrator - Arquitectura

O OpenGisIntegrator é uma arquitectura composta por módulos, ou componentes lógicos, com diferentes responsabilidades. Cada módulo é autónomo, desconhecendo a existência de outros módulos.

As próximas secções explicam em detalhe as responsabilidades de cada componente lógico.

4.2. Componentes Lógicos do OpenGisIntegrator

A arquitectura do OpenGisIntegrator é composta por um controlador e três componentes lógicos.

A Tabela 7 apresenta resumidamente os componentes lógicos e as responsabilidades de cada um.

Componente	Responsabilidades
Controller	Garantir a comunicação entre os diferentes componentes lógicos. É o único meio que os componentes lógicos possuem para comunicar entre si
DataRW Engine	Gerir os conteúdos de entrada e saída do OpenGisIntegrator
Converter Engine	<p>É constituído internamente por outros dois componentes:</p> <ul style="list-style-type: none">• O Interface Recognizer: responsável por identificar o formato a que cada ficheiro de entrada pertence, e associar a interface responsável por esse formato a cada ficheiro;• O Converter Processor: responsável por converter os dados existentes nos ficheiros de entrada para o formato de representação interno, feita por cada interface de cada ficheiro. <p>Após o processo de conversão e integração, o Converter Engine tem a responsabilidade de converter os dados integrados no formato de representação interno para os formatos de saída</p>
Integrator Engine	Integrar a informação dos diferentes ficheiros num único ficheiro de dados, representado no formato de representação interno.

Tabela 7 - OpenGisIntegrator - Componentes Lógicos

Os componentes do OpenGisIntegrator são explicados detalhadamente em seguida.

4.2.1. Controller

O Controller é responsável pelo controle da arquitetura. Todos os módulos comunicam com o Controller. É do Controller que recebem instruções e é a ele que enviam o resultado do processamento pelo qual são responsáveis

As suas responsabilidades são enumeradas na Tabela 8.

Responsabilidades do Controller
Identificar o conjunto de ficheiros de entrada.
Invocar o DataRW Engine, que obtém o conteúdo e a informação dos ficheiros de entrada. Obter o resultado do processamento do DataRW Engine.
Invocar o Converter Engine, que recebe como dados de entrada a informação retornada pelo DataRW Engine. Obter o resultado do processamento do Converter Engine.
Invocar o Integrator Engine, que recebe como dados de entrada a informação retornada pelo Converter Engine. Obter o resultado do processamento do Integrator Engine.
Invocar o Converter Engine, que recebe como dados de entrada a informação retornada pelo Integrator Engine. Obter o resultado do processamento do Converter Engine.
Retornar a informação retornada pelo Converter Engine, como resultado do processamento do OpenGisIntegrator.

Tabela 8 - OpenGisIntegrator - Responsabilidades do Controller

Desta forma, é garantida a modularidade dos vários componentes da arquitetura, sendo o controlo feito pelo Controller externamente a cada módulo.

4.2.2. DataRW Engine

O DataRW Engine tem como responsabilidade obter o conteúdo dos ficheiros de entrada e identificar o formato a que cada ficheiro pertence. É também responsável por gerar os ficheiros de saída, nos formatos previamente identificados.

4.2.2.1 Leitura dos Ficheiros de Entrada

A Figura 21 mostra a arquitectura do DataRW Engine.

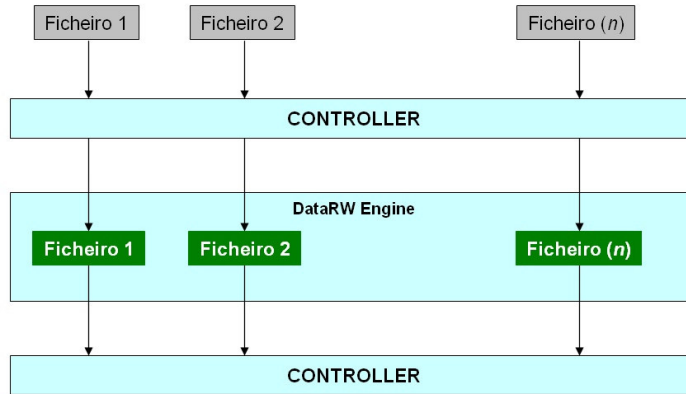


Figura 21 - OpenGisIntegrator - O DataRW Engine

O DataRW Engine lê os dados dos ficheiros de entrada e disponibiliza-os para o OpenGisIntegrator.

Recebe também os dados integrados e escreve-os nos ficheiros de saída, nos formatos pretendidos.

4.2.2.2 Escrita dos Ficheiros de Saída

Após percorrer os ficheiros a integrar, e concluir a integração de todas as características geográficas, o ficheiro eleito como principal (*mainFileX*) possui integrada a informação de todos os ficheiros de entrada. Resta então convertê-lo para os formatos de saída. Para isso o Integrator Engine disponibiliza o ficheiro integrado para o Controller.

A Figura 22 detalha a arquitectura do processo de integração dos dados.

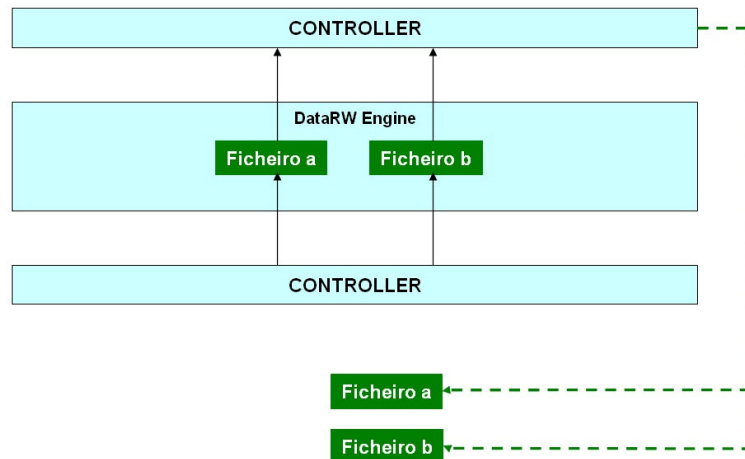


Figura 22 - OpenGisIntegrator - Geração dos Ficheiros de Saída

O Controller envia o ficheiro integrado para o Converter Engine, que identificou já os

formatos de saída. O Converter Engine efectua a conversão, e envia os ficheiros de saída nos formatos finais para o Controller. Este envia os ficheiros para o DataRW Engine, que se encarrega de disponibilizar os binários dos ficheiros gerados para o utilizador, nos vários formatos definidos por este.

O processo de conversão está assim concluído.

4.2.3. Converter Engine

A conversão dos ficheiros de entrada é da responsabilidade do Converter Engine. Este componente é composto por dois módulos: o Interface Recognizer e o Converter Processor.

O Interface Recognizer é o módulo responsável por receber os dados de entrada e colocá-los disponíveis para serem convertidos.

O Converter Processor é o módulo responsável pela conversão.

São explicados em seguida em detalhe os dois módulos que compõem este componente.

4.2.3.1 Interface Recognizer

O Interface Recognizer tem como função identificar o formato de cada ficheiro de entrada e associar a interface correspondente, responsável por conhecer esse formato.

É também responsável por identificar os formatos pretendidos para os ficheiros de saída, e associar a interface correspondente.

4.2.3.1.1 Reconhecimento dos Formatos dos Ficheiros de Entrada

A Figura 23 descreve o funcionamento interno do Interface Recognizer no tratamento dos ficheiros de entrada.

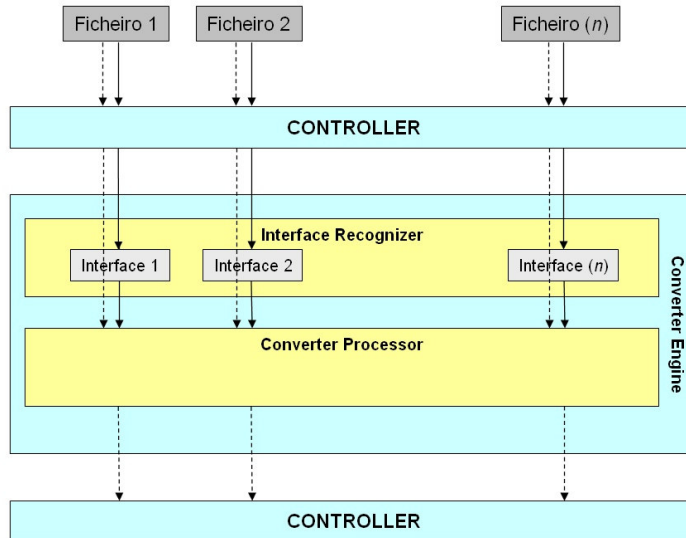


Figura 23 - OpenGisIntegrator - Arquitectura detalhada do Interface Recognizer

O Interface Recognizer recebe do Controller a lista de ficheiros de entrada, que percorre com vista a identificar o formato a que pertencem. Após identificar o formato, instancia a interface responsável por esse formato.

4.2.3.1.2 Reconhecimento dos Formatos dos Ficheiros de Saída

O Interface Recognizer é também responsável por identificar que formatos se pretendem nos ficheiros de saída, e instanciar as interfaces responsáveis por esses formatos.

A Figura 24 descreve o funcionamento interno do Interface Recognizer, agora na fase em que deverá tratar os ficheiros de saída.

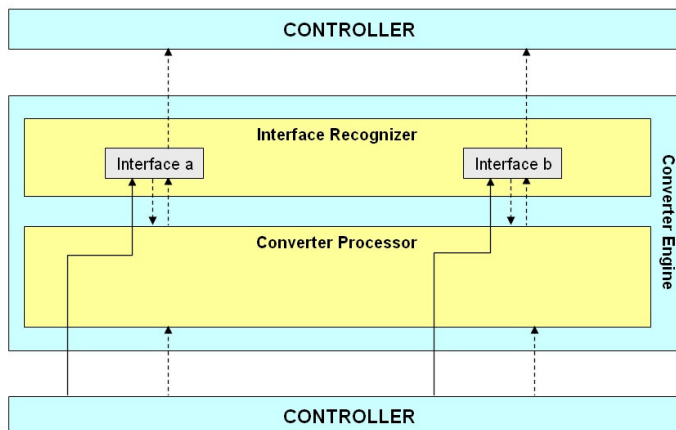


Figura 24 - OpenGisIntegrator - Identificação de Interfaces dos Ficheiros de Saída

As instâncias de todas as interfaces identificadas são devolvidas para o Controller do OpenGisIntegrator, que se encarregará de garantir a continuação do processo de integração.

4.2.3.2 Converter Processor

O Converter Processor é o conjunto de interfaces responsáveis por conhecer os detalhes dos vários formatos. A cada formato corresponde uma interface, responsável por conhecer e converter de e para esse formato.

A Figura 25 descreve o funcionamento do Converter Processor.

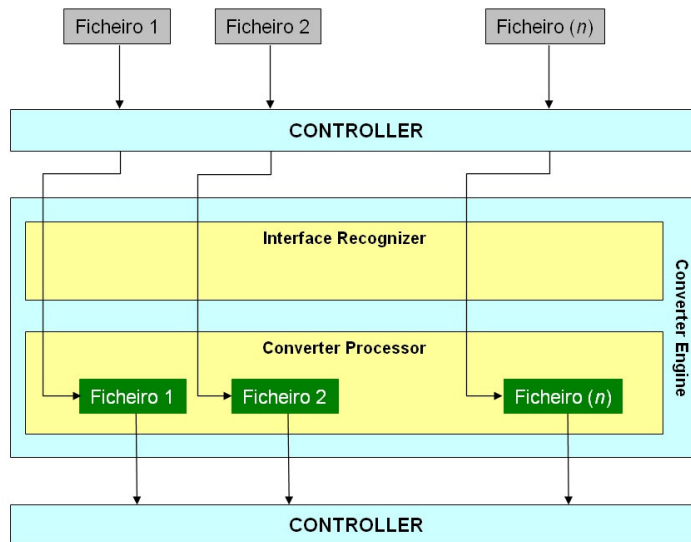


Figura 25 - OpenGisIntegrator - Arquitectura detalhada do Converter Processor

O Converter Processor é o conjunto de interfaces responsáveis por converter cada ficheiro de entrada, bem como os ficheiros de saída, de e para o formato de representação interna do OpenGisIntegrator.

A conversão dos ficheiros de saída está representada na Figura 26.

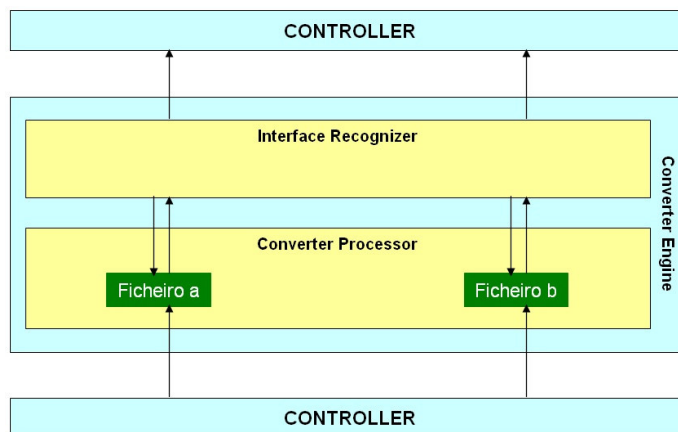


Figura 26 - OpenGisIntegrator - Conversão dos Ficheiros de Saída

Após identificação das interfaces que conhecem os formatos de saída, o Converter Processor encarrega-se de efectuar a conversão dos ficheiros de entrada, e notificar o Controller.

4.2.3.2.1 Modelo de Dados das Interfaces

As interfaces estão modeladas num modelo de dados implementado num conjunto de classes do OpenGisIntegrator.

A Figura 27 mostra o diagrama de classes das interfaces que integram o Converter Processor.

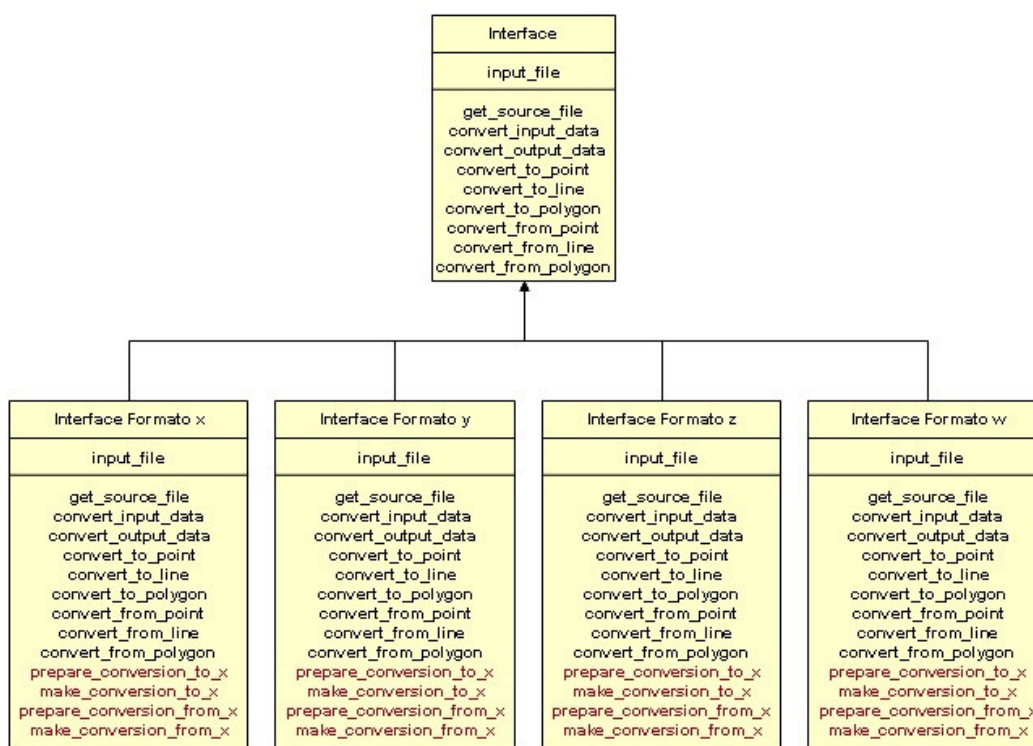


Figura 27 - OpenGisIntegrator - Diagrama de Classes do Converter Processor

Cada interface de cada formato herda da classe Interface. Cada interface tem como atributo o ficheiro de entrada, cujo conteúdo é guardado no atributo input_file.

A Tabela 9 descreve os métodos definidos para cada interface.

Método	Descrição
get_source_file	Método responsável por obter o conteúdo do ficheiro de entrada
convert_input_data	Método responsável por converter os dados de entrada para o formato de representação interno
convert_output_data	Método responsável por converter os dados de entrada no formato de representação interno para o formato da interface
convert_to_point	Método responsável por converter os dados de entrada sobre a característica geográfica "pontos" para o formato de representação interno
convert_to_line	Método responsável por converter os dados de entrada sobre a característica geográfica "linha" para o formato de representação interno
convert_to_polygon	Método responsável por converter os dados de entrada sobre a característica geográfica "polígono" para o formato de representação interno
convert_from_point	Método responsável por converter os dados de entrada sobre a característica geográfica "ponto" do formato de representação interno para o formato da interface
convert_from_line	Método responsável por converter os dados de entrada sobre a característica geográfica "linha" do formato de representação interno para o formato da interface
convert_from_polygon	Método responsável por converter os dados de entrada sobre a característica geográfica "polígono" do formato de representação interno para o formato da interface
prepare_conversion_to_x	Método responsável por validar se o conteúdo do ficheiro foi obtido com sucesso
make_conversion_to_x	Método responsável por fazer a conversão para o formato de representação interno; invoca os métodos convert_to_point, convert_to_line and convert_to_polygon
prepare_conversion_from_x	Método responsável por validar se o conteúdo do ficheiro foi gerado no formato da interface com sucesso
make_conversion_from_x	Método responsável por fazer a conversão do formato de representação interno para o formato da interface; invoca os métodos convert_from_point, convert_from_line and convert_from_polygon

Tabela 9 - OpenGisIntegrator - Métodos das Interfaces

Ao modelar as classes que representam as interfaces recorrendo à herança, as classes que modelam as várias interfaces partilham a mesma estrutura, generalizando assim a forma de acesso a cada interface.

O Converter Processor é assim constituído por uma estrutura modular e escalável. Caso se pretenda acrescentar uma nova interface a este componente (e por consequência, ao OpenGisIntegrator), bastará que a nova interface seja implementada herdando as características da classe interface.

4.2.4. Integrator Engine

Este é o componente principal do OpenGisIntegrator, já que é responsável por efectuar a integração dos ficheiros de diferentes formatos. É neste componente que são resolvidos conflitos de integração, são agrupadas as características geográficas dos vários ficheiros de entrada, e que é produzido o ficheiro de saída no formato interno, que será depois convertido para os diferentes formatos de saída pretendidos pelo utilizador. Explica-se em seguida o funcionamento deste componente.

Após receber os ficheiros de entrada já convertidos para o formato de representação interno, utilizado internamente pelo OpenGisIntegrator, estão reunidas as condições para integrar os dados provenientes de diferentes formatos num único ficheiro.

A Figura 28 detalha a arquitectura do processo de integração dos dados.

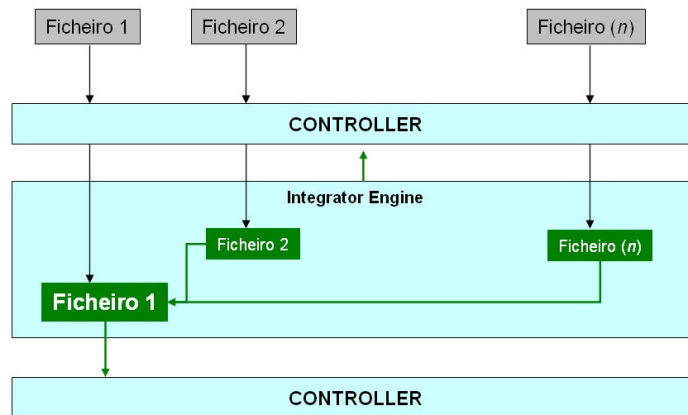


Figura 28 - OpenGisIntegrator - Detalhe da Arquitectura do Integrator Engine

A integração é feita elegendo um dos ficheiros como o ficheiro principal, e adicionando em seguida a este os dados existentes nos restantes ficheiros a integrar.

A Figura 29 e a Figura 30 detalham o processo de integração.

Os dados dos ficheiros de entrada, já disponíveis em instâncias da classe `input_data`, incluem os dados das características geográficas ponto, linha e polígono.

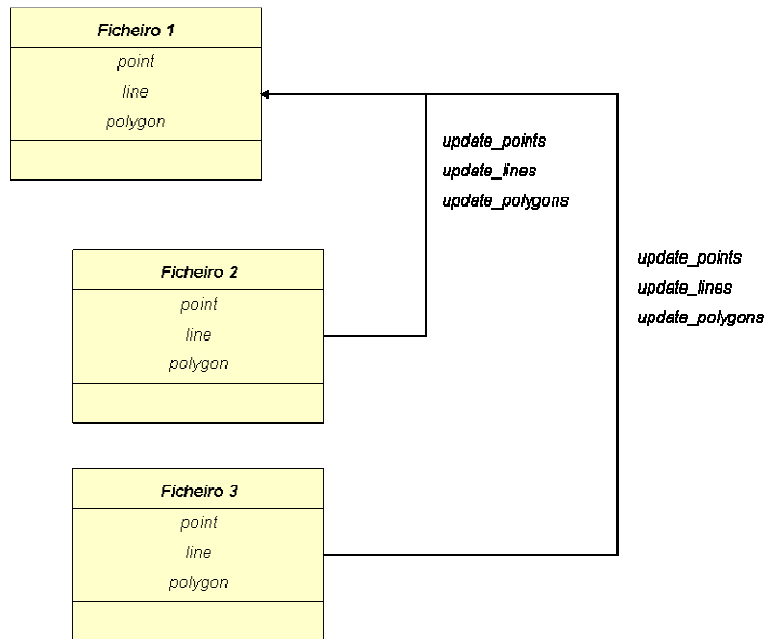


Figura 29 - OpenGisIntegrator - Instâncias da classe `input_data`

O Ficheiro 1 é eleito como o ficheiro principal.

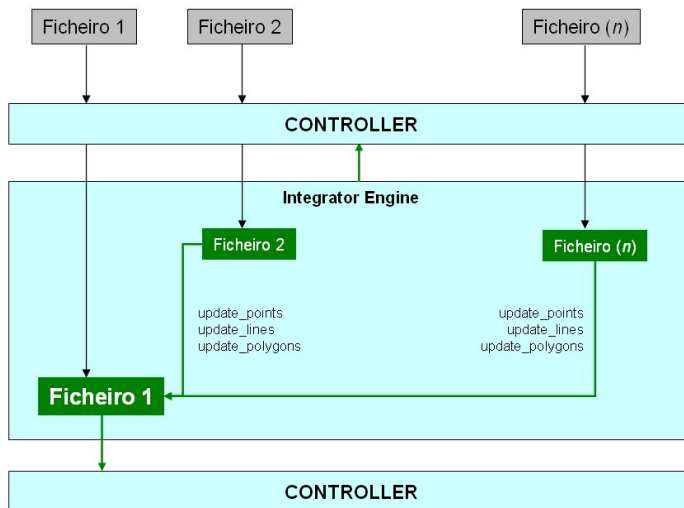


Figura 30 - OpenGisIntegrator - Detalhe do Integrator Engine

É ao Ficheiro 1 que serão adicionadas as características geográficas dos restantes ficheiros, recorrendo para isso aos métodos da classe `feature`, `update_points`, `update_lines` e `update_polygons`.

O resultado será o Ficheiro 1, que conterà os dados dos ficheiros Ficheiro 2 e Ficheiro 3.

4.3. Formato de Representação Interno

O formato de representação interno é um modelo de dados implementado através de um conjunto de classes que representam as características geográficas que os ficheiros representam. Cada interface conhece o formato de representação interno do OpenGisIntegrator, de modo a poder fazer conversões de e para este formato.

A Figura 31 representa o diagrama de classes que constituem o formato de representação interno.

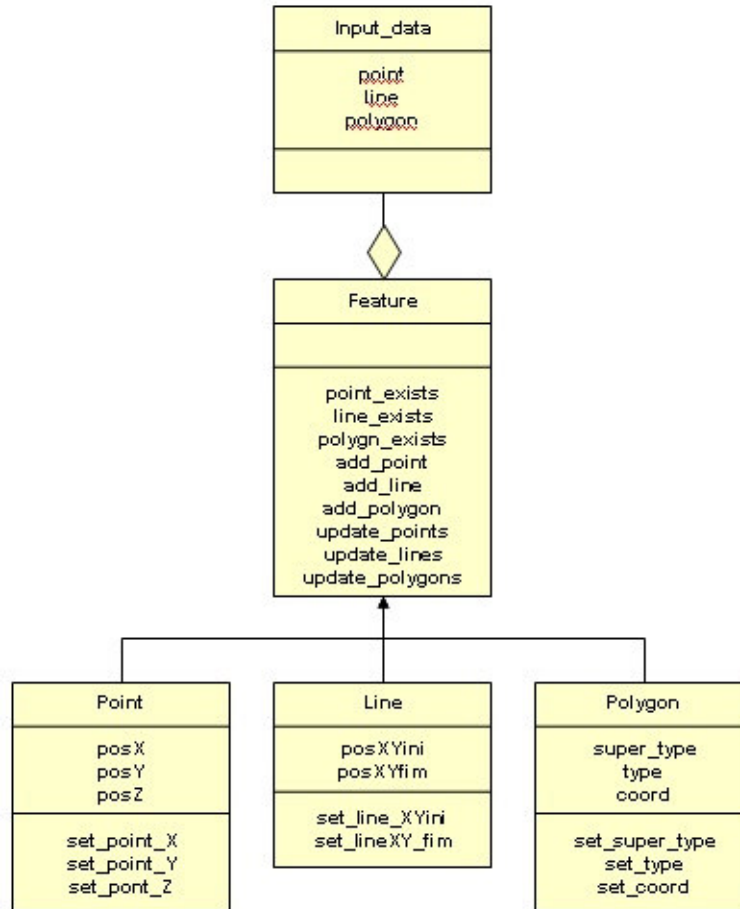


Figura 31 - OpenGisIntegrator - Diagrama de Classes da Representação Interna

A classe Input_data representa a instância de cada ficheiro de entrada, o que equivale a dizer que por cada ficheiro de entrada haverá uma instância da classe Input_data com os dados do ficheiro de entrada disponíveis no formato de representação interno do OpenGisIntegrator. Esta classe agrega várias características geográficas, disponibilizadas pela classe feature. As classes Point, Line e Polygon herdam da classe feature, e representam as características geográficas respectivas, existentes em cada ficheiro.

A Tabela 10 descreve as classes modeladas.

Classe	Descrição
input_data	Classe responsável por modelar o ficheiro de entrada; a cada ficheiro de entrada irá corresponder uma instância da classe input_data; esta classe possui como atributos os pontos, linhas e polígonos obtidos a partir dos ficheiros de entradas correspondentes
feature	Classe responsável por gerir as características geográficas existentes nos ficheiros; disponibiliza os seguintes métodos: . point_exists: verifica se determinado ponto já existe na classe . line_exists: verifica se determinada linha já existe na classe . polygon_exists: verifica se determinado polígono já existe na classe . add_point: adiciona um ponto à classe . add_line: adiciona uma linha à classe . add_polygon: adiciona um polígono à classe . update_points: atualiza um ponto da classe . update_lines: atualiza uma linha da classe . update_polygons: atualiza um polígono da classe
point	Classe responsável por gerir a característica geográfica "ponto"; disponibiliza os seguintes atributos: . posX: coordenada X do ponto . posY: coordenada Y do ponto . posZ: coordenada Z do ponto disponibiliza ainda os seguintes métodos: . set_point_X: atualiza o ponto X . set_point_Y: atualiza o ponto Y . set_point_Z: atualiza o ponto Z
line	Classe responsável por gerir a característica geográfica "linha"; disponibiliza os seguintes atributos: . posXYini: coordenadas XY do início da linha . posXYfim: coordenadas XY do fim da linha disponibiliza ainda os seguintes métodos: . set_line_XYini: atualiza as coordenadas do início da linha . set_line_XYfim: atualiza as coordenadas do fim da linha
polygon	Classe responsável por gerir a característica geográfica "polígono"; disponibiliza os seguintes atributos: . super_type: supertipo de polígono . type: tipo de polígono . coord: coordenadas do polígono disponibiliza ainda os seguintes métodos: . set_super_type: atualiza o supertipo de polígono . set_type: atualiza o tipo de polígono . set_coord: atualiza as coordenadas do polígono

Tabela 10 - OpenGisIntegrator - Classes do formato de Representação Interno

O Converter Engine, especificamente o Converter Processor, produz os ficheiros de saída a partir do formato de representação interno.

4.4. Resolução de Conflitos

O processo de integração de ficheiros com informação geográfica passa por agregar informação geográfica de diferentes ficheiros, muitas vezes referente à mesma área geográfica. Dado que um ficheiro pode referenciar as mesmas características geográficas que outro ficheiro a ser agregado, torna-se necessário validar se se trata da mesma característica ou de características semelhantes, mas que representem pontos diferentes no terreno.

Pela natureza do processo de integração, podem surgir conflitos na agregação, que a arquitectura do OpenGisIntegrator tem de identificar e permitir resolver, de forma a garantir que o ficheiro de saída integra realmente todas as características geográficas de todos os ficheiros de entrada.

4.4.1. Tipos de Conflitos

Como integrar dois pontos, linhas ou polígonos com coordenadas aproximadas são questões que se levantam na integração de características geográficas.

A Tabela 11 tipifica os conflitos possíveis na integração das características geográficas em análise nesta dissertação.

Tipo de Característica	Cenário	Ficheiro de Saída
point	n Pontos existentes em coordenadas que distam X graus entre si	Gerado um único ponto
		Gerados n pontos
line	n Linhas existentes em coordenadas que distam X graus entre si	Gerada uma única linha
		Geradas n linhas
polygon	n Polígonos existentes em coordenadas que distam X graus entre si	Gerado um único polígono
		Gerados n polígonos

Tabela 11 - Tipos de Conflitos na Integração Geográfica

4.4.2. Métodos de Resolução de Conflitos

O OpenGisIntegrator suporta a implementação de métodos responsáveis por verificar se determinada característica geográfica já existe.

O componente lógico responsável por garantir a resolução dos conflitos que eventualmente surjam na integração de características geográficas semelhantes é o Integrator Engine. Como já referido, o Integrator Engine percorre os ficheiros de entrada inseridos, com o objectivo de integrar a informação geográfica de todos os ficheiros em apenas um único ficheiro principal, eleito dentre os ficheiros de entrada, que incluirá toda a informação geográfica.

A proposta do OpenGisIntegrator para o processo de integração passa por implementar mecanismos no Integrator Engine que percorram todas as características geográficas de cada ficheiro com o objectivo de analisar cada uma e verificar se deve ser actualizada uma característica geográfica já existente, ou acrescentada a nova característica geográfica ao ficheiro principal.

A resolução de conflitos é feita recorrendo aos métodos da classe feature. Cada instância de cada ficheiro de entrada herda uma instância desta classe. Os métodos disponibilizados pela classe feature foram desenhados especificamente com o objectivo de gerir os conflitos na integração das características geográficas.

Os métodos da classe feature dividem-se em três tipos, de acordo com a Tabela 12.

Tipos de Métodos	Descrição	Métodos abrangidos
Verificação	Métodos de verificação da existência da característica geográfica em análise	point_exists
		line_exists
		polygon_exists
Adição	Métodos de adição da característica geográfica em análise	add_point
		add_line
		add_polygon
Actualização	Métodos de actualização de uma característica geográfica existente com informação da característica geográfica em análise	update_point
		update_line
		update_polygon

Tabela 12 - Classe feature - Tipos de Métodos

As próximas secções explicam o funcionamento dos tipos de métodos.

4.4.2.1 Métodos de Verificação

O resultado das verificações dita o que deverá acontecer com a característica geográfica em análise. Para efectuar essa verificação, deverão ser utilizados os métodos de verificação.

A implementação dos métodos de verificação deve incluir métricas de comparação de características geográficas adequadas às necessidades de cada realidade. Por exemplo, num determinado cenário dois pontos que distem 2 metros entre si podem ser considerados o mesmo ponto, descrito pela mesma característica geográfica. No entanto, noutro cenário esta distância já poderá ser pertinente, podendo a implementação optar por manter os dois pontos como características geográficas distintas.

4.4.2.2 Métodos de Adição

Caso o resultado da verificação da existência de uma característica geográfica indique que a característica não existe, então a implementação da arquitectura deve fornecer métodos para implementar os mecanismos de adição.

Os métodos de adição têm esse objectivo. Sempre que se verifique que se trata de características geográficas diferentes, então a característica geográfica em análise deverá ser adicionada ao ficheiro principal. A implementação dos métodos de adição deve garantir que as características geográficas novas são adicionadas.

4.4.2.3 Métodos de Actualização

Caso o resultado da verificação da existência de uma característica geográfica indique que se trata de uma característica geográfica já existente, deverá acontecer uma de duas situações:

- Verifica-se que se trata de uma característica geográfica já existente, cujos atributos descrevem também a nova característica geográfica em análise, e nesse caso não é executada qualquer operação;
- Verifica-se que se trata de uma característica geográfica já existente, mas cujo atributos necessitam de ser actualizados de modo a descreverem também a nova característica geográfica em análise. Nesse caso, os métodos de actualização deverão garantir a actualização dos atributos da característica geográfica em análise.

4.4.2.4 A Integração

O OpenGisIntegrator garante a integração recorrendo aos três tipos de métodos acima descritos. Desta forma é garantida a integração das características geográficas dos ficheiros de entrada, sendo integradas num único ficheiro, que descreverá as características geográficas de todos os ficheiros inseridos.

Como já referido, os critérios para verificar se determinada característica geográfica existe podem variar de implementação para implementação da arquitectura, de acordo com as necessidades da solução implementada. O OpenGisIntegrator propõe o Integrator Engine como o módulo responsável pela implementação de métodos que garantam a combinação de características geográficas semelhantes, disponibilizando um processo de integração que abrange todas as características geográficas armazenadas em todos os ficheiros de entrada.

4.5. Fluxo de Integração de Ficheiros – *Putting It All Together*

O OpenGisIntegrator permite a integração de ficheiros com informação geográfica em diferentes fontes, recorrendo aos componentes lógicos que o compõem.

O fluxo de integração de ficheiros é exemplificado na Figura 32, que apresenta um diagrama de sequência de integração de três ficheiros contendo informação geográfica.

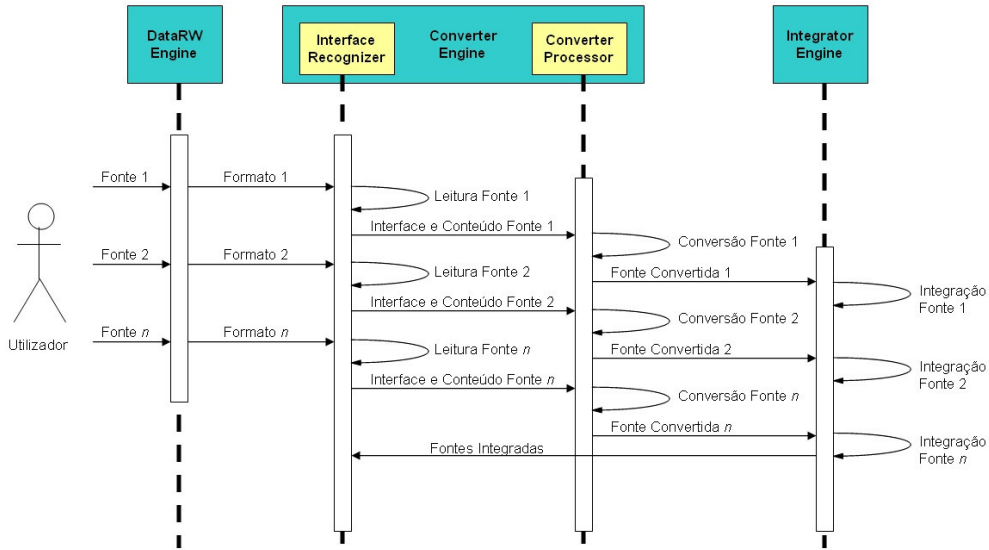


Figura 32 - OpenGisIntegrator - Diagrama de Sequência – Ficheiros de Entrada

A Figura 33 apresenta o diagrama que representa o ficheiro produzido com a informação geográfica destes três ficheiros integrada num só ficheiro, apresentado em dois ficheiros de diferentes formatos.

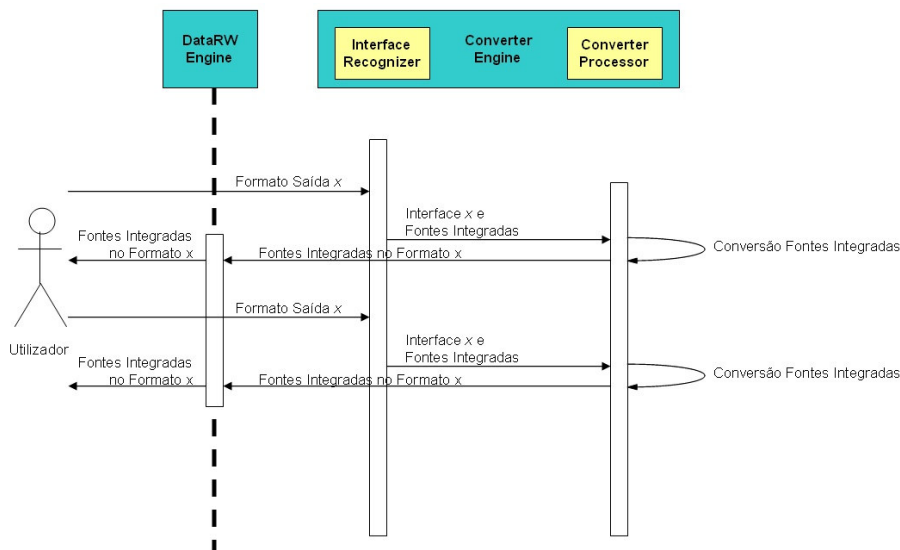


Figura 33 - OpenGisIntegrator - Diagrama de Sequência – Ficheiros de Saída

O diagrama de sequência apresentado é descrito em seguida.

4.5.1. Dados de Entrada

O utilizador insere os ficheiros “Ficheiro 1”, “Ficheiro 2” e “Ficheiro 3” no OpenGisIntegrator, e despoleta o processo de integração.

Os ficheiros são recebidos pelo DataRW Engine, responsável por gerir os ficheiros de entrada e saída do OpenGisIntegrator.

O DataRW Engine notifica o Controller

4.5.2. Conversão dos Dados de Entrada

O Controller notifica o Converter Engine que têm ficheiros. O Converter Engine associa a interface correcta a cada ficheiro, através do Interface Recognizer, e efectua a conversão para o formato interno através do Converter Processor.

Notifica então o Controller, que reconhece então cada ficheiro convertido no formato representação interno.

4.5.3. Integração dos Dados de Entrada

O Controller notifica agora o Integrator Engine, que obtém o conteúdo dos ficheiros, já no formato de representação interno, e integra todos os ficheiros, resolvendo todos os conflitos que surgem, e gerando um ficheiro final, no formato de representação interno, com a representação de toda a informação geográfica existente nos vários ficheiros de entrada.

Notifica então o Controller da integração dos ficheiros.

4.5.4. Geração do Ficheiro de Saída

O Controller notifica agora o Converter Engine que possui o ficheiro no formato de representação interno.

O Converter Engine efectua então a conversão do ficheiro integrado para os formatos dos ficheiros de saída definidos pelo utilizador. Utiliza novamente o Interface Recognizer para associar as interfaces aos formatos definidos pelo utilizador, e o Converter Processor para efectuar a conversão, agora do ficheiro no formato interno para os ficheiros de saída nos formatos pretendidos. Notifica então o Controller que possui os ficheiros prontos a entregar ao utilizador.

O Controller notifica então novamente o DataRW Engine, que se encarrega de gravar os ficheiros de saída numa localização física no disco, disponibilizando-os ao utilizador.

4.5.5. Resumo do Processo

O OpenGisIntegrator produz assim o ficheiro de saída, obtido a partir de ficheiros de entrada em diferentes formatos, convertendo cada um deles para o formato de representação interno, e integrando-os em seguida num único ficheiro. Após a integração, o ficheiro é convertido para os formatos de saída especificados, e estes são então devolvidos pela arquitectura.

4.6. OpenGisIntegrator - A Proposta de Arquitectura

Este capítulo descreveu a proposta de arquitectura de integração geográfica, designada por OpenGisIntegrator.

A integração de informação geográfica existente em diferentes formatos passa por diferentes processos. Perante a existência de diferentes ficheiros com diferentes formatos, e identificada a necessidade de utilizar a informação geográfica armazenada nesses ficheiros num SIG, é necessário efectuar um conjunto de passos até poder utilizar a informação no formato que se pretende. Exige-se que sejam executadas operações de conversão dos ficheiros para formatos aceites pelo SIG. É também necessário efectuar um trabalho de integração da informação num único ficheiro, ou em alternativa, inserir todos os ficheiros no SIG, caso este suporte a integração da informação. Na presença de um número considerável de ficheiros, o processo de conversão, integração e utilização da informação geográfica pode tornar-se complexo.

A proposta do OpenGisIntegrator é uma arquitectura que modela as diferentes fases do processo, desenhando uma solução de implementação que permita agilizar o processo de integração da informação geográfica.

O capítulo seguinte descreve a avaliação experimental feita, com recurso a um protótipo que demonstra uma implementação da arquitectura proposta.

5. AVALIAÇÃO EXPERIMENTAL

Para testar a implementação da arquitectura proposta, foi construído um protótipo da solução proposta, o OpenGisIntegrator, que implementa a arquitectura descrita no capítulo anterior. O protótipo tem como objectivo validar uma possível implementação da arquitectura descrita nesta dissertação.


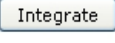
5.1. Interface Gráfica do Protótipo

O protótipo é composto por uma interface gráfica, conforme mostra a Figura 34.

Figura 34 - Protótipo do OpenGisIntegrator

O protótipo disponibiliza os seguintes campos:

- Localização dos ficheiros de entrada: caixas de inserção onde são inseridas as localizações dos ficheiros de entrada;
- Definição dos formatos dos ficheiros de entrada; caixas de selecção onde é possível seleccionar o formato do ficheiro de entrada, dentre os formatos disponíveis;
- Adição (+): funcionalidade que permite adicionar mais caixas de inserção para ficheiros de entrada;
- Definição dos formatos do ficheiro de saída: caixas de selecção onde é possível seleccionar o formato dos ficheiros de saída, dentre os formatos disponíveis;

- Adição (): funcionalidade que permite adicionar mais formatos para ficheiros de saída, dentre os formatos disponíveis;
- Botão  : funcionalidade que despoleta o processo de integração.

5.2. Parametrizações de Configurações

O protótipo implementado possui um conjunto de definições, utilizadas no processo de integração da informação geográfica inserida.

Estas definições dividem-se em premissas, sobre as quais o protótipo assenta, e em detalhes de implementação, incluídos na arquitectura e que definem o comportamento do protótipo em situações de integração específicas.

Estas parametrizações são detalhadas em seguida.

5.2.1. Premissas

O protótipo implementado assume um conjunto de premissas, listadas em seguida:

- Os ficheiros inseridos contemplam dados sobre as características geográficas ponto, linha e polígono, sendo omissos quaisquer dados adicionais, como atributos, metadados, ou outros;
- Assume-se que sempre que duas coordenadas distem entre si até 1 (um) metro, considera-se que se tratam das mesmas coordenadas, salvo se for necessário defini-las em separado;
- Assume-se que todos os ficheiros inseridos utilizam o mesmo sistema de coordenadas, não sendo contempladas conversões entre sistemas de coordenadas.

5.2.2. Implementação do Processo de Integração

O processo de integração implementado no protótipo compara todos os ficheiros de entrada. O processo de integração implementado assume, tal como definido na arquitectura proposta, o primeiro ficheiro de entrada lido como sendo o ficheiro principal. Os restantes ficheiros de entrada são integrados comparando as características que possuem com as características existentes neste ficheiro principal.

As linhas e os polígonos são definidos por conjuntos de pontos. O OpenGisIntegrator implementa os mecanismos de integração por comparação dos pontos que definem as características geográficas. Apenas são comparadas características geográficas do mesmo tipo. Pontos são comparados com pontos, linhas são comparadas com linhas e polígonos são comparados com polígonos. São comparados os pares de coordenadas

que compõem essa característica. Cada par de coordenadas, representado por um ponto definido por coordenadas (x,y), é comparado com um par de coordenadas já existente no ficheiro principal. Desta forma, em cada momento apenas são comparados dois pontos (x,y) de cada vez.

A implementação do processo de integração contempla a resolução de conflitos que surgem na integração de características geográficas do mesmo tipo. A Tabela 13 descreve a implementação dos métodos de resolução de conflitos que surgem na comparação de coordenadas (x,y), de acordo com a característica geográfica em análise em cada momento. Definem-se como pontos comuns os pontos que distam entre si até 1 (um) metro.

Característica	Resolução de Conflitos
Pontos	<ul style="list-style-type: none"> • O ponto a ser integrado é comum a um ponto já existente no ficheiro principal: o ponto em análise é ignorado; • O ponto a ser integrado não é comum a nenhum ponto já existente no ficheiro principal: o ponto em análise é acrescentado ao ficheiro principal.
Linhas	<p>São comparados todos os pontos que constituem a linha a ser integrada.</p> <ul style="list-style-type: none"> • O ponto a ser integrado é comum a um ponto já existente numa linha definida no ficheiro principal: <ul style="list-style-type: none"> . Se o ponto anterior analisado for um ponto comum, o ponto em análise é ignorado; . Se o ponto anterior analisado não for um ponto comum, a linha em análise é acrescentada ao ficheiro principal, adicionando o ponto em análise e o ponto anterior analisado; • O ponto a ser integrado não é comum a nenhum ponto já existente numa linha definida no ficheiro principal: a linha em análise é acrescentada ao ficheiro principal, adicionando o ponto em análise e o ponto anterior analisado.
Polígonos	<p>São comparados todos os pontos que constituem o polígono a ser integrada.</p> <ul style="list-style-type: none"> • O ponto a ser integrado é comum a um ponto já existente num polígono definido no ficheiro principal: <ul style="list-style-type: none"> . Se o ponto anterior analisado for um ponto comum, o ponto em análise é ignorado; . Se o ponto anterior analisado não for um ponto comum, a linha do polígono em análise é acrescentada ao ficheiro principal, adicionando o ponto em análise e o ponto anterior analisado; • O ponto a ser integrado não é comum a nenhum ponto já existente num polígono definido no ficheiro principal: a linha do polígono em análise é acrescentada ao ficheiro principal, adicionando o ponto em análise e o ponto anterior analisado.

Tabela 13 - Parametrizações do Protótipo

Foram implementados os seguintes formatos para a validação da arquitectura: KML, GML e geoRSS.

5.3. Cenários de Avaliação

O protótipo do OpenGisIntegrator foi avaliado em diferentes cenários de teste, de forma a validar a implementação da arquitectura proposta.

Foram efectuados dois tipos de avaliação:

- Avaliação da arquitectura: foram validados cenários que permitiram avaliar a escalabilidade da arquitectura, a produção de ficheiros de saída em mais que um formato, e a integração de ficheiros do mesmo formato e de diferentes formatos;
- Avaliação da integração das características geográficas: foram validados cenários de integração de pontos comuns ou distintos, linhas comuns ou distintas e polígonos comuns ou distintos. Foram avaliadas as formas de resolução de conflitos que surgiram decorrentes da existência de características geográficas cuja localização era aproximada, e foram testados os mecanismos que permitiram decidir se se tratava da mesma característica ou de uma característica diferente. Neste tipo de avaliação foram também testados cenários de integração de características geográficas diferentes.

A avaliação experimental, bem como os resultados obtidos, são descritos nas secções seguintes.

5.3.1. Avaliação da Arquitectura

A avaliação da arquitectura permitiu testar o comportamento do OpenGisIntegrator ao integrar informação geográfica existente em diferentes ficheiros, em diferentes formatos, e sempre que haja necessidade de acrescentar novas interfaces para suportar novos formatos até então não suportados.

Os cenários de avaliação são descritos em seguida.

5.3.1.1 Avaliação da Escalabilidade da Arquitectura

O OpenGisIntegrator permite acrescentar suporte a novos formatos, sem que este tipo de alterações à arquitectura tenha um impacto significativo na implementação.

Para validar este cenário, o OpenGisIntegrator foi inicialmente desenvolvido suportando dois formatos, o KML e o GML, tendo sido posteriormente acrescentado o formato geoRSS à arquitectura.

A implementação seguiu os seguintes passos:

- Foi acrescentada ao Converter Engine uma nova interface para o novo formato, que herda da classe Interface, actualizada com dados específicos do formato geoRSS:



- Foi alterado o Interface Recognizer, de modo a incluir o mapeamento para o novo formato:

```
get_interface : function (input_file) {
    var interface_recognizer;
    var file_format = input_file.file_format;

    switch (file_format) {
        case "gml" :
            interface_recognizer = new gml_interface(input_file);
            break;
        case "kml" :
            interface_recognizer = new kml_interface(input_file);
            break;
        case "geoRSS" :
            interface_recognizer = new georss_interface(input_file);
            break;

        default :
            interface_recognizer = null;
            break;
    }

    return interface_recognizer;
}
```

- Foi acrescentada ao DataRW Engine a indicação do suporte ao novo formato:

```
this.supported_formats = ["kml", "gml", "geoRSS"];
```

Esta alteração reflectiu-se na lista de formatos suportados visualizados no protótipo:



Para validar o cenário, foi feita uma conversão simples do formato geoRSS para o formato GML.

Foi inserido o ficheiro geoRSS com a seguinte estrutura:

```
<feed xmlns=" http://www.w3.org/2005/Atom" xmlns:georss="http://www.georss.org/georss">
  <entry>
    <georss:point>45.256 -71.92</georss:point>
  </entry>
</feed>
```

Foi dada instrução de integração, e foi gerado o ficheiro convertido para GML.

O ficheiro gerado possui a seguinte estrutura:

```
<gml:sample xmlns:gml="http://www.w3.org/TR/html4/">
  <gml:position xmlns:gml="http://www.w3.org/TR/html4/">
    <gml:Point xmlns:gml="http://www.w3.org/TR/html4/">
      <gml:coordinates>45.256,-71.92</gml:coordinates>
    </gml:Point>
  </gml:position>
</gml:sample>
```

A conversão simples foi efectuada com sucesso. As características geográficas existentes no ficheiro com o formato geoRSS foram convertidas para o formato GML, tal como se pretendia validar.

5.3.1.2 Avaliação da Integração de Ficheiros de Formatos Diferentes

O OpenGisIntegrator deverá permitir integrar informação geográfica existente em diferentes ficheiros de diferentes formatos, resultando num ficheiro de saída que deverá incluir a informação de entrada integrada.

Para verificar este cenário, foram inseridos três ficheiros de diferentes formatos, KML, GML e geoRSS, e foi seleccionado o formato KML como formato de saída. Após inserção da localização dos ficheiros foi dada a instrução de integração.

Foram inseridos três ficheiros com a seguinte estrutura:

- Ficheiro #1 – Formato geoRSS - inclui a característica geográfica Ponto:

```
<feed xmlns="http://www.w3.org/2005/Atom" xmlns:georss="http://www.georss.org/georss">
  <entry>
    <georss:point>45.256 -71.92</georss:point>
  </entry>
</feed>
```

- Ficheiro #2 – Formato GML - inclui a característica geográfica Linha:

```
<gml:LineString xmlns="http://www.opengis.net/examples" xmlns:gml="http://www.opengis.net/gml">
  <gml:coordinates>45.67,88.56 55.56,89.44</gml:coordinates>
</gml:LineString >
```

- Ficheiro #3 – Formato KML - inclui a característica geográfica Polígono:

```
<kml xmlns="http://www.opengis.net/kml/2.2">
  <Placemark>
    <Polygon>
      <outerBoundaryIs>
        <LinearRing>
          <coordinates>
            -77.05788457660967,38.87253259892824,100
            -77.05465973756702,38.87291016281703,100
            -77.05315536854791,38.87053267794386,100
            -77.05552622493516,38.868757801256,100
          </coordinates>
        </LinearRing>
      </outerBoundaryIs>
    </Polygon>
  </Placemark>
</kml>
```

Foi dada instrução de integração e foi gerado o ficheiro convertido para KML.

O ficheiro gerado possui a seguinte estrutura:

```
<kml xmlns="http://www.opengis.net/kml/2.2">
  <Placemark xmlns="">
    <Point>
      <coordinates>45.256,-71.92</coordinates>
    </Point>
  </Placemark>
  <Placemark xmlns="">
    <LineString>
      <coordinates>45.67,88.56 55.56,89.44</coordinates>
    </LineString>
  </Placemark>
  <Placemark xmlns="">
    <Polygon>
      <outerBoundaryIs>
        <LinearRing>
          <coordinates>
            -77.05788457660967,38.87253259892824,100
            -77.05465973756702,38.87291016281703,100
            -77.05315536854791,38.87053267794386,100
            -77.05552622493516,38.868757801256,100
          </coordinates>
        </LinearRing>
      </outerBoundaryIs>
    </Polygon>
  </Placemark>
</kml>
```

A integração de ficheiros de diferentes formatos foi efectuada com sucesso. O ficheiro resultante da integração inclui as características geográficas existentes nos três ficheiros de entrada, nomeadamente, as Linhas, os Pontos e os Polígonos, integrados num único ficheiro, no formato de saída definido, KML, tal como se pretendia.

5.3.1.3 Avaliação dos Ficheiros de Saída em Diferentes Formatos

O OpenGisIntegrator permite integrar os ficheiros de entrada num único ficheiro, que pode ser apresentado ao utilizador em diferentes formatos.

Para verificar este cenário, foi inserida a localização dos três ficheiros de diferentes formatos, KML, GML e geoRSS, utilizados no exemplo de avaliação anterior, e foram seleccionados como formatos de saída o formato KML e o formato GML.

O conteúdo dos ficheiros de entrada é o seguinte:

- Ficheiro #1 – Formato geoRSS - inclui a característica geográfica Ponto:

```
<feed xmlns=" http://www.w3.org/2005/Atom" xmlns:georss="http://www.georss.org/georss">
  <entry>
    <georss:point>45.256 -71.92</georss:point>
  </entry>
</feed>
```

- Ficheiro #2 – Formato GML - inclui a característica geográfica Linha:

```
<gml:LineString xmlns="http://www.opengis.net/examples" xmlns:gml="http://www.opengis.net/gml">
  <gml:coordinates>45.67,88.56 55.56,89.44</gml:coordinates>
</gml:LineString >
```

- Ficheiro #3 – Formato KML - inclui a característica geográfica Polígono:

```
<kml xmlns="http://www.opengis.net/kml/2.2">
  <Placemark>
    <Polygon>
      <outerBoundaryIs>
        <LinearRing>
          <coordinates>
            -77.05788457660967,38.87253259892824,100
            -77.05465973756702,38.87291016281703,100
            -77.05315536854791,38.87053267794386,100
            -77.05552622493516,38.868757801256,100
          </coordinates>
        </LinearRing>
      </outerBoundaryIs>
    </Polygon>
  </Placemark>
</kml>
```

Foi dada instrução de integração e foram gerados os ficheiros de saída nos formatos KML e GML.

O ficheiro gerado no formato KML possui a seguinte estrutura:

```
<kml xmlns="http://www.opengis.net/kml/2.2">
  <Placemark xmlns="">
    <Point>
      <coordinates>45.256,-71.92</coordinates>
    </Point>
  </Placemark>
  <Placemark xmlns="">
    <LineString>
      <coordinates>45.67,88.56 55.56,89.44</coordinates>
    </LineString>
  </Placemark>
  <Placemark xmlns="">
    <Polygon>
      <outerBoundaryIs>
        <LinearRing>
          <coordinates>
            -77.05788457660967,38.87253259892824,100
            -77.05465973756702,38.87291016281703,100
            -77.05315536854791,38.87053267794386,100
            -77.05552622493516,38.868757801256,100
          </coordinates>
        </LinearRing>
      </outerBoundaryIs>
    </Polygon>
  </Placemark>
</kml>
```

O ficheiro gerado no formato GML possui a seguinte estrutura:

```
<Model>
  <gml:Point>
    <gml:pos>45.256,-71.92</gml:pos>
  </gml:Point>
  <gml:LineString>
    <gml:coordinates>45.67,88.56 55.56,89.44</gml:coordinates>
  </gml:LineString >
  <gml:Polygon>
    <gml:outerBoundaryIs>
      <gml:LinearRing>
        <gml:coordinates>
          -77.05788457660967,38.87253259892824,100
          -77.05465973756702,38.87291016281703,100
          -77.05315536854791,38.87053267794386,100
          -77.05552622493516,38.868757801256,100
        </gml:coordinates>
      </gml:LinearRing>
    </gml:outerBoundaryIs>
  </gml:Polygon>
</kml>
```

A geração dos ficheiros de saída nos formatos seleccionados foi efectuada com sucesso. Os ficheiros resultantes da integração incluem as características geográficas existentes nos três ficheiros de entrada, nomeadamente, as Linhas, os Pontos e os Polígonos, integrados num único ficheiro, disponibilizado em dois formatos de saída, o formato KML e o formato GML, tal como se pretendia.

5.3.2. Avaliação da Integração das Características Geográficas

A avaliação da integração das características geográficas permitiu testar o comportamento do OpenGisIntegrator ao lidar com coordenadas geograficamente próximas. Para efeitos de simplificação e melhor compreensão dos cenários de teste, foram utilizados ficheiros de entrada no formato KML. Os ficheiros de entrada incluem coordenadas no *datum* WSG84, e os gráficos utilizados para representação das características geográficas são disponibilizados em metros.

Os cenários de avaliação são descritos em seguida.

5.3.2.1 Avaliação da Integração de Pontos

Neste cenário de utilização, foram avaliadas as funcionalidades de integração de características geográficas do tipo Ponto existentes em dois ficheiros distintos. Os ficheiros de entrada utilizados incluem um conjunto de pontos, alguns com coordenadas aproximadas geograficamente, devendo portanto ser detectada correspondência entre os pontos, e outros que não têm correspondência entre si.

Para efectuar a validação, foram inseridos dois ficheiros.

O Ficheiro #1 possui a seguinte estrutura:

```
<kml>
  <Placemark>
    <Point>
      <coordinates>58.51127403328638,0.000027058128099035313</coordinates>
    </Point>
    <Point>
      <coordinates>58.51128299228297,0.000036077504577906815</coordinates>
    </Point>
    <Point>
      <coordinates>58.51130091027761,0.000027058129102217802</coordinates>
    </Point>
    <Point>
      <coordinates>58.51129195128087,0.00000901937625594144</coordinates>
    </Point>
    <Point>
      <coordinates>58.51126507428996,0.000009019375921546616</coordinates>
    </Point>
  </Placemark>
</kml>
```

A Figura 35 representa os pontos do Ficheiro #1 em metros.

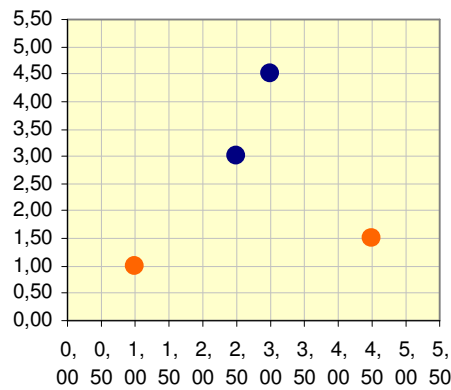


Figura 35 - Integração de Pontos - Ficheiro #1

O Ficheiro #2 possui a seguinte estrutura:

```
<kml>
  <Placemark>
    <Point>
      <coordinates>58.511274033286,0.000027058128</coordinates>
    </Point>
    <Point>
      <coordinates>58.511282992283,0.000036077505</coordinates>
    </Point>
    <Point>
      <coordinates>58.511291951281,0.000009019376</coordinates>
    </Point>
  </Placemark>
</kml>
```

A Figura 36 representa os pontos do Ficheiro #2 em metros.

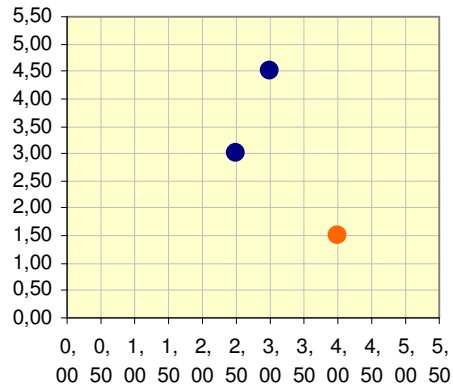


Figura 36 - Integração de Pontos - Ficheiro #2

O resultado da integração foi o ficheiro KML resultante, cuja estrutura é a seguinte:

```
<kml xmlns="http://www.opengis.net/kml/2.2">
  <Placemark>
    <Point>
      <coordinates>58.51127403328638,0.000027058128099035313</coordinates>
    </Point>
    <Point>
      <coordinates>58.51128299228297,0.000036077504577906815</coordinates>
    </Point>
    <Point>
      <coordinates>58.51130091027761,0.000027058129102217802</coordinates>
    </Point>
    <Point>
      <coordinates>58.51129195128087,0.00000901937625594144</coordinates>
    </Point>
    <Point>
      <coordinates>58.51126507428996,0.000009019375921546616</coordinates>
    </Point>
  </Placemark>
</kml>
```

A Figura 37 representa os pontos do Ficheiro de saída em metros.

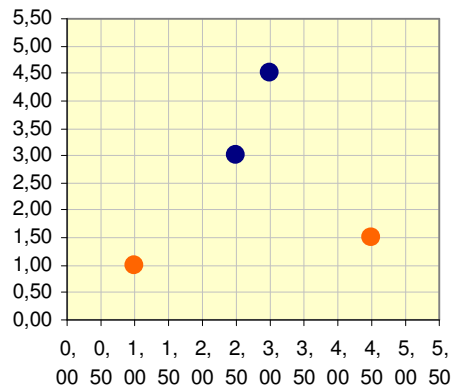


Figura 37 - Integração de Pontos - Ficheiro de Saída

O cenário descrito mostra três situações distintas de validação.

- Pontos em comum: os pontos assinalados com as coordenadas em metros (2.50, 3.00) e (3.00, 4,50) mantiveram-se pois eram comuns aos dois ficheiros;
- Pontos diferentes: o ponto assinalado com as coordenadas em metros (1.00, 1.00) foi incluído pois apenas existia em um dos ficheiros;
- Pontos aproximados: o ponto assinalado com as coordenadas em metros (4.50, 1.50) foi incluído pois existia no Ficheiro #1, e no Ficheiro #2, com uma distância inferior à definida, 1 (um) metro. Logo, o OpenGisIntegrator considerou-o o mesmo ponto, mantendo o ponto com as coordenadas já existentes no ficheiro principal.

Este cenário validou a integração da característica geográfica Ponto, testando a funcionalidade de resolução de conflitos, tendo sido obtidos os resultados de integração esperados.

5.3.2.2 Avaliação da Integração de Linhas

Neste cenário de utilização, foram avaliadas as funcionalidades de integração de características geográficas do tipo Linha existentes em dois ficheiros distintos. Os ficheiros de entrada utilizados incluem um conjunto de pontos que descrevem linhas, sendo algumas dessas linhas definidas por coordenadas aproximadas geograficamente, devendo portanto ser detectada correspondência entre elas, e sendo outras linhas definidas por coordenadas distantes geograficamente, que não tendo portanto correspondência entre si.

Para efectuar a validação, foram inseridos dois ficheiros.

O Ficheiro #1 possui a seguinte estrutura:

```
<kml>
  <Placemark>
    <LineString>
      <coordinates>
        58.511265074289,0.000036077504
        58.511274033287,0.000018038752
        58.511282992284,0.000018038752
        58.511291951281,0.00009019376
      </coordinates>
    </LineString>
  </Placemark>
</kml>
```

A Figura 38 representa os pontos do Ficheiro #1 em metros.

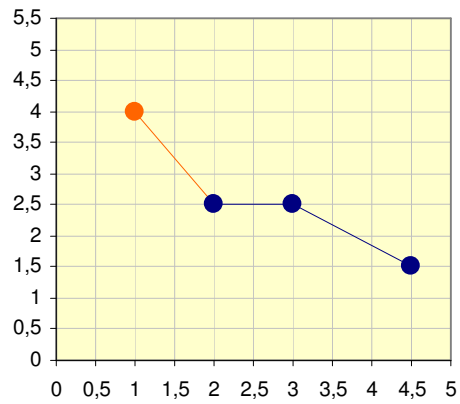


Figura 38 - Integração de Linhas - Ficheiro #1

O Ficheiro #2 possui a seguinte estrutura:

```
<kml>
  <Placemark>
    <LineString>
      <coordinates>
        58.511265074290,0.000009019376
        58.511274033287,0.000018038752
        58.511282992284,0.000018038752
        58.511291951281,0.000009019376
      </coordinates>
    </LineString>
  </Placemark>
</kml>
```

A Figura 39 representa os pontos do Ficheiro #2 em metros.

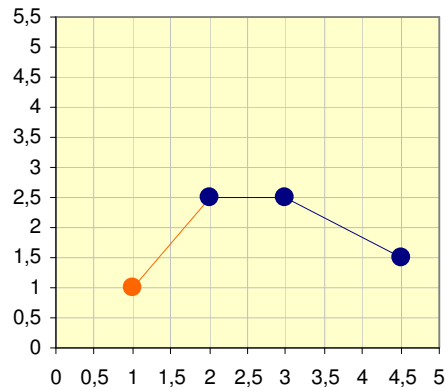


Figura 39 - Integração de Linhas - Ficheiro #2

O resultado da integração foi o ficheiro KML resultante, cuja estrutura é a seguinte:

```
<kml>
  <Placemark>
    <LineString>
      <coordinates>
        58.511265074289,0.000036077504
        58.511274033287,0.000018038752
        58.511282992284,0.000018038752
        58.511291951281,0.000009019376
      </coordinates>
    </LineString>
    <LineString>
      <coordinates>
        58.511265074290,0.000009019376
        58.511274033287,0.000018038752
      </coordinates>
    </LineString>
  </Placemark>
</kml>
```

A Figura 40 representa os pontos do Ficheiro de saída em metros.

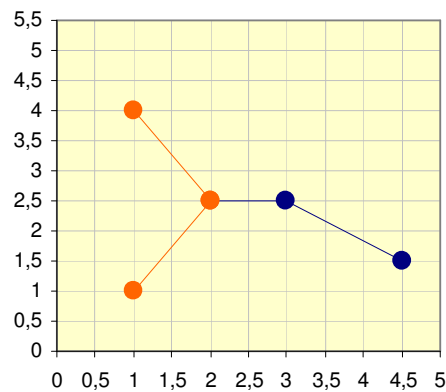


Figura 40 - Integração de Linhas - Ficheiro de Saída

O cenário descrito mostra três situações distintas de validação.

- Linhas em comum: as linhas assinaladas com as coordenadas em metros (3.00, 2.50) e (4.50, 1.50) mantiveram-se, pois eram comuns aos dois ficheiros;
- Linhas diferentes: a linha assinalada com as coordenadas em metros (1.00, 4.00) foi incluído pois apenas existia em um dos ficheiros;
- Linhas aproximadas: a linha assinalada com as coordenadas em metros (2.00, 2.50) foram incluídas em ambas as representações de linhas, pois apesar de o ponto ser comum a ambos os ficheiros, o ponto anterior não o era. Logo, o OpenGisIntegrator incluiu duas linhas distintas, que ao ser desenhadas sobrepõem o mesmo ponto, com as coordenadas em metros (2.00, 2.50), mas são representadas como duas linhas distintas.

Este cenário validou a integração da característica geográfica Linha, testando a funcionalidade de resolução de conflitos, tendo sido obtidos os resultados de integração esperados.

5.3.2.3 Avaliação da Integração de Polígonos

Neste cenário de utilização, foram avaliadas as funcionalidades de integração de características geográficas do tipo Polígono existentes em dois ficheiros distintos. Os ficheiros de entrada utilizados incluem um conjunto de pontos que descrevem linhas, que por sua vez desenham polígonos, sendo algumas dessas linhas definidas por coordenadas aproximadas geograficamente, devendo portanto ser detectada correspondência entre elas, e sendo outras linhas definidas por coordenadas distantes geograficamente, que não tendo portanto correspondência entre si.

Para efectuar a validação, foram inseridos dois ficheiros.

O Ficheiro #1 possui a seguinte estrutura:

```
<kml>
  <Placemark>
    <Polygon>
      <outerBoundaryIs>
        <LinearRing>
          <coordinates>
            58.511265074289,0.000036077504
            58.511291951280,0.000036077505
            58.511291951281,0.000018038753
            58.511265074290,0.000018038752
            58.511265074289,0.000036077504
          </coordinates>
        </LinearRing>
      </outerBoundaryIs>
    </Polygon>
  </Placemark>
</kml>
```

A Figura 41 representa os pontos do Ficheiro #1 em metros.

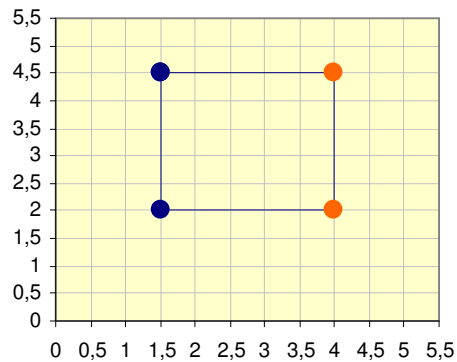


Figura 41 - Integração de Polígonos - Ficheiro #1

O Ficheiro #2 possui a seguinte estrutura:

```

<kml>
  <Placemark>
    <Polygon>
      <outerBoundaryIs>
        <LinearRing>
          <coordinates>
            58.511265074289,0.000036077504
            58.511282992283,0.000036077505
            58.511282992284,0.000018038752
            58.511265074290,0.000018038752
            58.511265074289,0.000036077504
          </coordinates>
        </LinearRing>
      </outerBoundaryIs>
    </Polygon>
  </Placemark>
</kml>

```

A Figura 42 representa os pontos do Ficheiro #2 em metros.

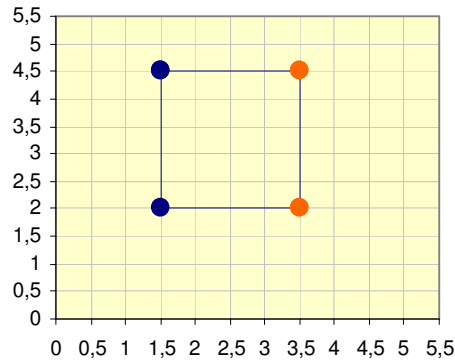


Figura 42 - Integração de Polígonos - Ficheiro #2

O resultado da integração foi o ficheiro KML resultante, cuja estrutura é a seguinte:

```

<kml>
  <Placemark>
    <Polygon>
      <outerBoundaryIs>
        <LinearRing>
          <coordinates>
            58.511265074289,0.000036077504
            58.511291951280,0.000036077505
            58.511291951281,0.000018038753
            58.511265074290,0.000018038752
            58.511265074289,0.000036077504
          </coordinates>
        </LinearRing>
      </outerBoundaryIs>
    </Polygon>
  </Placemark>
</kml>

```

A Figura 43 representa os pontos do Ficheiro de saída em metros.

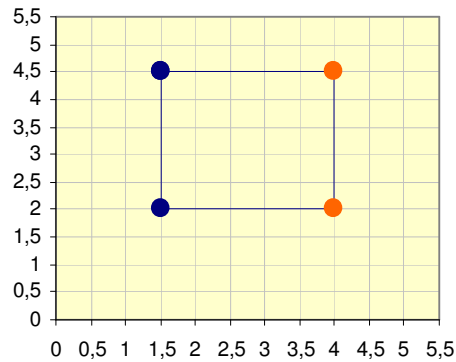


Figura 43 - Integração de Polígonos - Ficheiro de Saída

O cenário descrito mostra duas situações distintas de validação.

- Linhas em comum: as linhas assinaladas com as coordenadas em metros (1.50, 2.00) e (1.50, 4.50) mantiveram-se, pois eram comuns aos dois ficheiros;
- Linhas aproximadas: as linhas assinaladas com as coordenadas em metros (4.00, 2.00) e (4.00, 4.50) foram incluídas em ambas as representações de linhas, pois embora as coordenadas variassem entre os dois ficheiros, não variavam mais do que a distância máxima definida para a arquitectura, de 1 (um) metro. Assim, o OpenGisIntegrator descartou as linhas assinaladas no Ficheiro #2, mantendo as coordenadas definidas no Ficheiro #1.

Este cenário validou a integração da característica geográfica Polígono, testando a funcionalidade de resolução de conflitos, tendo sido obtidos os resultados de integração esperados.

5.4. O Protótipo

A implementação do protótipo da arquitectura do OpenGisIntegrator, descrita neste capítulo, tinha por objectivo validar a implementação da arquitectura proposta nesta dissertação.

Foram incluídos cenários de avaliação, onde se visava avaliar a arquitectura, quer do ponto de vista do desenho proposto, quer do ponto de vista da resolução de conflitos.

Os cenários demonstrados são exemplificativos do processo de integração proposto na arquitectura descrita, e validam o funcionamento da arquitectura na integração de informação geográfica existente em diferentes formatos, bem como a integração específica das características geográficas, permitindo, através de parametrizações definidas, tomar decisões na resolução de conflitos originados pela integração de características geográficas semelhantes.

6. CONCLUSÕES E TRABALHO FUTURO

A área que explora os Sistemas de Informação Geográfica é vasta. São muitas as questões em aberto no campo da informação geográfica, sendo uma área com vários caminhos a percorrer, explorar e contribuir.

Esta dissertação procurou dar um contributo na área da integração da informação geográfica, proveniente de diferentes fontes, em diferentes formatos.

6.1. Conclusões

A interoperabilidade entre sistemas que lidam com informação geográfica levanta problemas de reutilização da informação geográfica que, em casos de total incompatibilidade entre os formatos dos ficheiros e o SIG onde se pretende utilizá-los, podem levar à impossibilidade de utilização dessa informação.

Com vista a analisar a oferta de soluções que resolvam este problema, foram analisadas arquitecturas de integração de informação geográfica, que contribuíram para conhecer metodologias e estratégias de conversão e integração da informação geográfica, e ajudaram a perceber os problemas com que a integração de informação geográfica tem de lidar e quais as necessidades que ainda se fazem sentir neste campo.

Identificou-se a necessidade de implementação de arquitecturas vocacionadas especificamente para sistemas de integração de informação geográfica, e que permitam a integração da informação, com o objectivo de facilitar o acesso à informação geográfica dispersa por diferentes fontes e consequentemente em diferentes formatos.

Como resultado do trabalho de estudo realizado, esta dissertação propôs-se a apresentar o desenho de uma solução que permita agilizar o processo de integração de informação geográfica, através de uma arquitectura modular, que se propõe percorrer as várias fases do processo de análise e integração da informação geográfica dispersa, e facilitar o acesso à mesma.

O OpenGisIntegrator é a proposta de arquitectura apresentada nesta dissertação, e consiste numa proposta de uma arquitectura que recebe informação geográfica de diferentes fontes, em diferentes formatos, que a converte para um formato de representação interno. Após a conversão, a arquitectura integra a informação geográfica recebida nos vários ficheiros num único modelo de dados, que é então disponibilizado ao

utilizador nos formatos de saída por ele indicados, e que podem ser posteriormente utilizados em SIGs que os suportem.

Os cenários de utilização apresentados no capítulo da avaliação experimental permitiram avaliar a adequação da solução proposta ao problema da integração de informação geográfica dispersa por diferentes ficheiros com diferentes formatos, e demonstrar, através de um protótipo demonstrativo de uma implementação da arquitectura proposta, como a integração desta informação geográfica pode ser facilitada com uma solução desta natureza.

Foram demonstrados casos de utilização da integração de informação geográfica com diferentes cenários.

Os casos de utilização incluíram cenários em que se introduziram poucos ficheiros com a informação dispersa, e cenários em que se introduziu uma quantidade maior de ficheiros com informação geográfica a integrar. Os resultados obtidos permitiram avaliar o comportamento da arquitectura implementada, no que diz respeito à forma como a informação geográfica existente em cada ficheiro foi integrada num único modelo de dados, e apresentada em um ou vários formatos de dados.

Os casos de utilização permitiram também avaliar como a implementação da arquitectura no OpenGisIntegrator lidou com os conflitos que surgiram na integração. Foram parametrizadas configurações na implementação do protótipo, com o objectivo de simular casos reais de utilização que permitissem comparar dados geográficos semelhantes, e decidir, com base nos parâmetros definidos, como os integrar, optando por acrescentar os novos dados geográficos aos dados já integrados, ou por assumir que se tratam dos mesmo dados, e nesse caso não incluir informação repetida. Os resultados obtidos permitiram avaliar o comportamento do processo de integração na resolução de conflitos na integração, e garantir que esta é efectuada sem que da integração da informação geográfica dispersa resultem dados geográficos repetidos ou omissos.

6.2. Trabalho Futuro

O trabalho desenvolvido no decurso desta dissertação teve como objectivo dar uma contribuição na área de análise e integração de informação geográfica dispersa por diferentes ficheiros com diferentes formatos. No entanto, este trabalho não se encerra aqui, podendo ser explorando, numa perspectiva de continuidade, alguns temas relacionados tanto com esta área, como especificamente com o trabalho desenvolvido.

Algumas propostas de trabalhos em aberto identificadas e passíveis de ser continuadas são apresentadas em seguida:

- A extensão do modelo de dados apresentado no formato de representação interno da arquitectura proposta, que permita suportar a parametrização de informação relevante para a análise dos dados, como a configuração de *datums* diferentes, a conversão entre diferentes sistemas de projecções, entre outros tipos de configurações;
- A extensão do modelo de dados apresentado no formato de representação interno da arquitectura proposta, que permita suportar o tratamento de características geográficas complexas, baseadas nas características geográficas mais simples, o ponto, a linha ou o polígono;
- A inclusão de ficheiros de configuração de informação adicional, como metadados, como forma de personalizar processos específicos de integração de dados;
- A possibilidade de processar em paralelo os vários ficheiros de entrada, recorrendo à implementação de metodologias de *Multithreading*.

O trabalho desenvolvido nesta dissertação pretende dar uma contribuição na exploração de metodologias de integração de informação geográfica heterogénea, numa perspectiva de continuidade do processo de investigação e aprendizagem na área dos Sistemas de Informação Geográfica.

REFERÊNCIAS

- [Casanova et al., 2005] Casanova, M., Câmara, G., Davis C., Vinhas, L., Queiroz, G. R., “Bancos de Dados Geográficos”, MundoGEO, 2005
- [Kavouras et al., 2002] Kavouras, M., Kokla, M., “Developing Multi-scale, Multi-context databases through the Semantic Integration of Heterogeneous Datasets”, National Technical University of Athens, 2002
- [ESRI, 2008] “Essays on Geography and GIS”, ESRI, Sep 2008
- [Stoimenov, 2006] Stoimenov, L., “Geographic Information Systems Interoperability in Local Community Environment”, University of Nis, Sérvia, 2006
- [GeoComm.com, 2009],
<http://data.geocomm.com/helpdesk/formats.html>, acessido em Junho/2009
- [GIS.com, 2009], <http://www.gis.com/whatisgis/index.html>, acessido em Junho/2009
- [StateMaster, 2007],
<http://www.statemaster.com/encyclopedia/Geographic-information-system>, acessido em Junho/2009
- [Smith, Goodchild, Longley, 2007] Smith M. J., Goodchild M. F., Longley P. A., “Geospatial analysis: A comprehensive guide to principles, techniques and software tools”, 2nd edition, 2007, Troubador, UK, disponível *online* em
<http://www.spatialanalysisonline.com/output/>
- [OGC, 2009], <http://www.opengeospatial.org/ogc>, acessido em Junho/2009
- [Sajeevan, G., 2008] “Latitude and Longitude – A misunderstanding”, Março de 2008
- [Sajeevan, G., 2005] “Correct Usage of Projection Parameters for Spatial Accuracy”, 2005
- [ESRI, Metadata and GIS, 2002] “An ESRI ® White Paper”, Outubro 2002
- [GISCentrum, 2009],
<http://www.giscentrum.lu.se/english/whatisgisFileFormat.htm>, acessido em Junho/2009
- [Moyer, Niemann, 1993] Moyer D. D., Niemann, J. Jr., “The why, what and how of GIS Standards: Issues for discussion and resolution”, 1993

- [Arlington, 2004] Arlington, V., "Here There Be Data: Mapping the Landscape of Science", National Science Foundation, ScienceDaily: News & Articles in Science, Health, Environment & Technology, 2004
- [Hall et al, 2003] Hall, O., Alm, G., Ene, S. & Jansson, U. "Introduktion till kartografi och geografisk information. Denmark: Naryana Press", 2003
- [Ekhlund et al, 2003] Ekhlundh, L., Harrie, L., Hauska, H., Olsson, L. & Pilesjö, P. "Geografisk informationsbehandling: metoder och tillämpningar", 3ª edição, 2003
- [Bryntse, 2007] Bryntse, S., "A GIS tool for biological research - visualising flycatcher breeding data collected on Öland", 2007
- [FME – KF85, 2008] FME Readers and Writers, "Swedish KF85 Reader/Writer", 1998, modificado em 2008
- [ESRI – SHP, 2000] ESRI, "ESRI Shapefile Technical Description", 1998, modificado em 2000
- [Reichardt, 2004] Reichardt, M., "The Havoc of Non-Interoperability", 2004, OGC White Paper
- [OGC, 2005] "Interoperability & Open Architectures: An Analysis of Existing Standardization Processes & Procedures", OGC, 2005
- [McKee, OGC, 2005] McKee, L., "The Importance of Going 'Open'", OGC White Paper, 2005
- [Mira et al, 2007] Mira, F., Santos C., Jorge, L., Quaresma, E., El-Zein, A., Santos, N., Afonso, P., Ferreira, L., Carvalho, P., Silva, F., "An Architecture for Flexible GIS Integration of Distribution Planning and Analysis systems", EDP, 2007
- [Ghosh et al, 2006] Paul, M., Ghosh, S. K., De Sarkar, S. C., Acharya, P. S., "A Service Oriented Architecture for Enterprise-GIS", NRDMS, Department of Science & Technology, New Delhi, 2006, disponível *online* em http://www.gisdevelopment.net/proceedings/mapindia/2006/poster/mi06pos_89.htm
- [Hofman, 2004] Hofman, D., "MOIRA DSS – Architecture, Model Integration and User Interface Design", 2004
- [Kotzinos, Pratacos, 2004] Kotzinos, D., Pratacos, P., "Use of a Web-Based GIS for Real-Time Traffic Information Fusion and Presentation over the Internet", AGILE Conference on Geographic Information Science, 2004
- [Tuama, 2004] Tuama, E. O., "Implementing DISPRO – A Web Based GIS", Coastal & Marine Resources Centre, ERI, University College Cork, Ireland, 2004
- [INSPIRE, 2009] <http://inspire.jrc.ec.europa.eu/>, acedido em Setembro de 2009

[Hamre et al, 2005] Hamre, T., Sandven, S., Tuama, E. O. "Data Integration System for Marine Pollution and Water Quality (DISMAR)", MERIS-(A)ATSR User Workshop, Frascati, Italy, 26-30 September 2005

[Mrozek, Fraczek, 2004], Mrozek, D., Fraczek, J., "Architecture of the Open Format Converter", Proceedings of the FOSS/GRASS Users Conference - Bangkok, Thailand, 12-14 September 2004

[Rosenberg] Rosenberg, M., "Map Stops Cholera", About.com, artigo disponível em <http://geography.about.com/cs/medicalgeography/a/cholera.htm>, acedido em Março de 2010

[National Geographic]

<http://www.nationalgeographic.com/resources/ngo/education/ideas912/912choleraho3.html>, acedido em Março de 2010