

C&S SIG

***BIODIVERSITY OPEN MAP (BIODOM)
SIGWeb Colaborativos e Software Geoespacial
Livre e Aberto Numa Abordagem da Neogeografia
Aplicada ao Conhecimento da Biodiversidade***

Paulo André Ribeiro

Trabalho de Projecto apresentado como requisito parcial
para obtenção do grau de Mestre em Ciência e Sistemas
de Informação Geográfica

Instituto Superior de Estatística e Gestão de Informação
da Universidade Nova de Lisboa

BIODIVERSITY OPEN MAP (BIODOM)

SIGWeb colaborativos e *Software* Geoespacial Livre e Aberto Numa Abordagem da Neogeografia Aplicada ao Conhecimento da Biodiversidade

Trabalho de Projeto orientado por
Professor Doutor Rui Pedro Julião

Novembro de 2012

AGRADECIMENTOS

Gostaria de prestar os meus sinceros agradecimentos a todos os que, direta ou indiretamente, tornaram a realização deste sonho realidade.

Estou profundamente grato ao Prof. Dr. Rui Pedro Julião, por ter aceite ser o orientador científico deste Projeto, por me ter inspirado e apoiado, sempre, na concretização do mesmo e na redação desta dissertação.

Ao Ricardo Silva e restantes intervenientes da equipa ProData, pela abertura demonstrada para a realização deste Projeto e pelos vários e enriquecedores debates de ideias e sugestões.

Alargo o meu agradecimento ao André Barriguinha e ao Miguel Fernandes, Engenheiros, Mestres em C&SIG e Mestres na amizade e no apoio, ao longo dos últimos anos desde que a grande aventura começou.

À Teresa, a quem este Projeto “roubou” horas infinitas, mas que sempre me acompanhou com amor, carinho e compreensão.

Aos mestres espirituais e ao maior guru de todos: a vida.

Finalmente gostaria de exprimir o meu amor e a minha gratidão à minha mãe e ao meu pai, por todo o apoio e compreensão, sem os quais esta dissertação não teria vindo a lume. Este trabalho é dedicado a eles.

Notas prévias: Este texto foi escrito ao abrigo do Acordo Ortográfico da Língua Portuguesa, de 16 de Novembro de 1990, que entrou em vigor em Janeiro de 2009 (em período transitório até 2015). Optou-se por incluir alguns termos em *itálico* que representam designações anglo-saxónicas de uso comum do domínio dos Sistemas de Informação Geográfica e das Tecnologias da Informação. O seu desenvolvimento foi realizado, integralmente, com recurso a ferramentas informáticas de código aberto.

BIODIVERSITY OPEN MAP (BIODOM)

SIGWeb colaborativos e Software Geoespacial Livre e Aberto Numa Abordagem da Neogeografia Aplicada ao Conhecimento da Biodiversidade

RESUMO

A Biodiversidade é o bem mais valioso da Humanidade. Sendo um domínio muito abrangente tem vindo a ser alvo de estudo, de modo indireto, desde há alguns séculos ao nível de abordagens da Taxonomia e Sistemática. Mais recentemente, também, ao nível da determinação da riqueza e abundância de espécies. Pese embora este facto, apesar de atualmente surgirem cada vez mais iniciativas dedicadas a este domínio, é um facto universalmente aceite que apenas se conhece uma pequena fração da Biodiversidade do planeta. Menos ainda, ao nível da sua distribuição espacial. O presente Projeto pretende constituir-se como uma nova abordagem ao nível da caracterização geoespacial da distribuição de espécies e da sua riqueza. Deve ser visto como complemento às iniciativas e tendências de estudos atuais e não como uma abordagem substituta ou por oposição. Agregam-se os conceitos dos Sistemas de Informação Geográfica (SIG) Colaborativos e da Web Geoespacial, para se construir um sistema aplicacional de SIGWeb. Esta conceptualização permite uma abordagem Neogeográfica em que os colaboradores, de modo voluntário, se tornam simultaneamente produtores e consumidores de Informação Geográfica (Voluntária), relativa à distribuição de espécies, com base nas suas observações territoriais e nos seus próprios conhecimentos. A utilização de *software* geoespacial livre e aberto no desenvolvimento do sistema aplicacional é um elemento chave do Projeto. Não só devido à inexistência de custos e, conseqüentemente, a possibilidade de as ferramentas chegarem a um largo espectro de utilizadores, mas também porque a natureza colaborativa envolvida na sua tipologia de desenvolvimento se funde parcialmente com a natureza colaborativa do Projeto. Se se concluir afirmando que o sistema desenvolvido visa a integração numa das maiores iniciativas de caracterização e quantificação da Biodiversidade em Portugal: o portal Naturdata, fica assim caracterizado o *Biodiversity Open Map (BiodOM)*.

BIODIVERSITY OPEN MAP (BIODOM)

Collaborative WebGIS and Geospatial Free and Open Source Software in a Neogeographic Approache Applied to Biodiversity Knowledge

ABSTRACT

To talk about Biodiversity is to talk about the most valuable asset of mankind. Being a vast domain it has been a subject of indirect study since some centuries ago, through the approaches of Taxonomy and Systematic and, more recently, through the determination of species richness and abundance. Although in the modern times there is an ever growing number of initiatives dedicated to this domain, it is a universally accepted fact that it is only known a small fraction of planet Earth Biodiversity. Even less regarding to it's spatial distribution. This Project aims to be a new approach at the level of geospatial characterization of species distribution and it's richness. It should be seen as complementary to the initiatives and present studies tendencies and not as a substitute or by opposition to them. It aggregates the concepts of Collaborative Geographic Information Systems (GIS) and Geospatial Web to develop a WebGIS application system. This conceptualization allows space for a Neogeographic approach in which the contributors, in a volunteer way, become simultaneously producers and consumers of (Volunteered) Geographic Information. This kind of Geographic Information has origin in the contributors field perceptions and observations, regarding to species distribution. The use of geospatial free and open source software in the deployment of the application system is a key concept of the Project. Not only it reduces implementation costs which allows to reach a larger spectrum of volunteer users, as well as it's collaborative development nature partially merges with the collaborative nature of the Project. If we conclude saying that the developed system aims to be integrated in one of the largest initiatives of characterization and quantification of Biodiversity in Portugal: Naturdata Portal, the Project Biodiversity Open Map (BiodOM) is resumed.

PALAVRAS-CHAVE

Biodiversidade

Sistemas de Informação Geográfica Colaborativos

Neogeografia

Software Geoespacial Livre e Aberto

Informação Geográfica Voluntária

Mapa Aberto

KEYWORDS

Biodiversity

Collaborative Geographic Information Systems

Neogeography

Free and Open Source Software for Geospatial

Volunteered Geographic Information

Open Map

ACRÓNIMOS

- ABCD** - Access to Biological Collection Data
- API** - Application Programming Interfaces
- BD** - Base de Dados
- BDG** - Base de Dado Geoespacial
- BiodOM** - Biodiversity Open Map
- C&SIG** - Ciência e Sistemas de Informação Geográfica
- CAOP** - Carta Administrativa Oficial de Portugal
- CE** - Comissão Europeia
- CIG** - Ciência de Informação Geográfica
- CRUD** - Create, Read, Update e Delete
- DwC** - Darwin Core
- EDIT** - European Distributed Institute of Taxonomy
- FOSS4G** - Free and Open Source Software for Geospatial
- GBIF** - Global Biodiversity Information Facility
- GML** - Geography Markup Language
- GPS** - Global Positioning Systems
- HTML** - Hyper Text Markup Language
- HTTP** - HyperText Transfer Protocol
- ICNB** - Instituto de Conservação da Natureza e Biodiversidade
- ICV** - Identificadores da Ciência da Vida
- ICZN** - International Code of Zoological Nomenclature
- IDE** - Infraestrutura de Dados Espaciais
- IG** - Informação Geográfica
- IGT** - Instrumentos de Gestão Territorial
- IGU** - Interface Gráfica do Utilizador
- IGV** - Informação Geográfica Voluntária
- IUCN** - International Union for Conservation of Nature
- MSc** - Mestrado
- OGC** - Open Geospatial Consortium
- OSM** - OpenStreetMap
- PdC** - Prova de Conceito
- PHP** - PHP Hypertext Preprocessor
- PSRN2000** - Plano Sectorial da Rede Natura 2000
- SGBD** - Software Gestor de Bases de Dados
- SGBDGR** - Software Gestor de Bases de Dados Geoespaciais Relacionais
- SGLA** - Software Geoespacial Livre e Aberto
- SIGP** - Sistema de Informação Geográfica Participativo
- SIGPP** - Sistema de Informação Geográfica de Participação Pública
- SIGWeb** - Sistema de Informação Geográfica na Web
- SSH** - Secure Shell
- SO** - Sistema Operativo

SQL - Structured Query Language
TI - Tecnologias da Informação
TWDG - Taxonomic Databases Working Group
UE - União Europeia
WCS - Web Coverage Service
WFS - Web Feature Service
WFS-T - Web Feature Service Transaccional
WMS - Web Map Service
ZEC - Zona Especial de Conservação
ZPE - Zona de Proteção Especial

ÍNDICE DO TEXTO

AGRADECIMENTOS.....	iii
RESUMO.....	v
ABSTRACT.....	vi
PALAVRAS-CHAVE.....	vii
KEYWORDS.....	vii
ACRÓNIMOS.....	viii
ÍNDICE DE FIGURAS.....	xiii
ÍNDICE DE TABELAS.....	xiv
1 INTRODUÇÃO.....	1
1.1 Enquadramento.....	1
1.2 Objetivos.....	3
1.3 Premissas.....	4
1.4 Metodologia e faseamento.....	4
1.5 Estrutura do Trabalho de Projeto.....	5
2 BIODIVERSIDADE.....	6
2.1 Introdução.....	6
2.2 Conceito.....	6
2.3 O valor (escondido) da Biodiversidade.....	7
2.4 Tempo e espaço em Biodiversidade.....	9
2.5 Investigação em Biodiversidade.....	10
2.5.1 Taxonomia e espécie.....	10
2.5.2 Aferir Biodiversidade.....	12
2.5.3 Identificadores únicos globais para as nomenclaturas.....	15
2.5.4 Protocolos e padrões de interoperabilidade em Biodiversidade.....	16
2.6 Modelação e SIG em Biodiversidade.....	18
2.7 Conservação da Biodiversidade em Portugal e a Rede Natura 2000.....	19
2.8 O Projeto Naturdata.....	21
2.9 Conclusões do capítulo.....	23
3 INTERNET, SIG COLABORATIVOS E A PARTILHA DE CONHECIMENTO.....	25
3.1 Introdução.....	25
3.2 SIGWeb e Web 2.0: a Web Geoespacial.....	25
3.2.1 Web Standards, OGC e Interoperabilidade.....	27
3.2.1.1 WMS.....	27
3.2.1.2 WFS.....	28
3.3 SIG de Participação Pública (SIGPP) e SIG Colaborativos (SIGC).....	29
3.4 O conceito de Neogeografia.....	33

3.4.1 Desigualdade de informação geográfica em Neogeografia.....	34
3.5 Informação Geográfica Voluntária (IGV).....	35
3.5.1 Cartografia colaborativa.....	37
3.5.2 A visualização de informação cartográfica na sociedade da informação atual...38	
3.5.3 A qualidade da informação e o conceito de vago em IGV.....	38
3.5.4 IGV como input de dados geoespaciais para investigação.....	40
3.6 Integração de IGV em Infraestruturas de Dados Espaciais (IDE).....	41
3.7 SIG Colaborativos e Informação Geográfica Voluntária (IGV) aplicados à distribuição da Biodiversidade.....	42
3.8 Conclusões do capítulo.....	43
4 SOFTWARE GEOESPACIAL LIVRE E ABERTO E DADOS GEOESPACIAS ABERTOS. .45	
4.1 Introdução.....	45
4.2 Software Geoespacial Livre e Aberto.....	45
4.3 A fundação OSGeo.....	46
4.4 Convergência entre SGLA e padrões abertos de informação.....	47
4.5 Dados geoespaciais abertos.....	48
4.6 Comunidades como fio-condutor de requisitos tecnológicos de SGLA.....	49
4.7 SGLA e Dados Geoespaciais Abertos como Valor para o Utilizador Final.....	50
4.8 Componentes de um sistema aplicacional para um SIGWeb Colaborativo.....	51
4.8.1 Debian 6.0.5 “Squeeze”: um sistema operativo.....	51
4.8.2 Geoserver 2.1.3.....	52
4.8.3 Base de Dados Geográfica: PostgreSQL/PostGIS.....	53
4.8.4 Um Cliente: Openlayers 2.12.....	55
4.9 Conclusões do capítulo.....	55
5 A ESTRUTURA E O MODELO DE DADOS DO BIODOM.....	57
5.1 Introdução.....	57
5.2 A modelação de dados geográficos	57
5.3 Modelação de Dados de IGV em SIG Colaborativos.....	58
5.4 IGV vaga e incorporação de incerteza num modelo de dados de IGV.....	58
5.5 Modelo de dados BiodOM.....	60
5.5.1 Premissas de base e requisitos.....	60
5.5.2 Fontes de entrada de dados de IGV.....	61
5.5.3 Desenho do modelo de dados.....	61
5.6 Cartografia temática de base do BiodOM.....	65
5.7 OpenStreetMap.....	66
5.8 A segurança dos dados do BiodOM.....	68
5.9 Os formatos dos dados do BiodOM.....	69
5.10 Conclusões do capítulo.....	69

6 DESENVOLVIMENTO E IMPLEMENTAÇÃO DO SISTEMA APLICACIONAL: BIODOM...	71
6.1 Introdução.....	71
6.2 Requisitos operacionais.....	71
6.3 Conceptualização da arquitetura do sistema.....	72
6.3.1 Camada de recursos: PostgreSQL/PostGIS.....	74
6.3.2 Camada de processamento: Geoserver.....	74
6.3.3 Camada de Interface Gráfica do Utilizador: OpenLayers.....	74
6.3.3.1 Mapa da ficha de espécie (modo visualização).....	74
6.3.3.2 Mapa da ficha de espécie: modo de edição (módulo browser editor).....	75
6.3.3.3 Mapa global BiodOM.....	77
6.3.4 API Google Maps.....	78
6.4 Implementação.....	79
6.4.1 O sítio www.biodom.pt.....	79
6.5 Especificações particulares: interface Naturdata/BiodOM.....	80
6.5.1 Criação de utilizadores na BDG BiodOM após o registo no portal Naturdata....	80
6.5.2 Importação das tabelas MySQL Naturdata e integração no BiodOM.....	80
6.6 Especificações funcionais: ferramentas do sistema e produção de IGV.....	81
6.6.1 Browser editor.....	81
6.6.2 Exif georeader.....	81
6.6.3 GPX georeader.....	82
6.6.4 QGIS editor.....	82
6.7 Especificações de operações de pós-processamento: a IGV aplicada ao conhecimento geoespacial da Biodiversidade	82
6.7.1 ST_Intersects.....	84
6.7.2 ST_ConvexHull.....	84
6.8 Aplicação a uma prova de conceito (PdC).....	85
6.9 Conclusões do capítulo.....	85
7 CONSIDERAÇÕES FINAIS.....	88
REFERÊNCIAS BIBLIOGRÁFICAS.....	91
Anexo A.....	104
Anexo B.....	107
Anexo C.....	117
Anexo D.....	119
Anexo E.....	123
Anexo F.....	125

ÍNDICE DE FIGURAS

Figura 1: Exemplo de ficha de espécie do portal Naturdata	23
Figura 2: Cubo de SIG Colaborativos	31
Figura 3: Estrutura conceptual de abordagem aos SIG colaborativos	32
Figura 4: O modelo “gema de ovo” na representação de ponto, linha e polígono, vagos	60
Figura 5: Modelo de dados implementado no BiodOM	62
Figura 6: Ilustração do código SQL para o esquema “current” da BDG BiodOM	64
Figura 7: Passo 1, criação do esquema “cartbase” novo na biodomdb	66
Figura 8: Passo 2, conversão de shapefile para o formato .sql da BDG PostGIS com shp2pgsql	66
Figura 9: Passo 3, importação de um ficheiro .sql para a biodomdb	66
Figura 10: Importação de um ficheiro .osm para a biodomdb	67
Figura 11: Script do crontab programado para a importação da camada de edificado do OSM para o território de Portugal	67
Figura 12: Script do cronjob programado para a realização de dumps das BDG do BiodOM	68
Figura 13: Arquitetura do sistema aplicacional	73
Figura 14: Exemplo do mapa da ficha de espécie	75
Figura 15: Exemplo do mapa da ficha de espécie em modo de edição	76
Figura 16: Aspeto geral do mapa global BiodOM	77
Figura 17: Código utilizado de implementação da camada aerofotográfica da API Google Maps	78
Figura 18: Exemplo da sintaxe SQL de ficheiro .sql externo exemplo, invocado pelo cronjob de interseção das nuvens de pontos com o tema de localidades	83
Figura 19: Produto do cruzamento da ocorrência de uma espécie com a grelha de Portugal	84

ÍNDICE DE TABELAS

Tabela 1: Estimativa de espécies descobertas e por descobrir	7
Tabela 2: Cartografia temática de base do sistema aplicativo BiodOM	65

1 INTRODUÇÃO

“Todos os pequenos passos que se dão, são maiores que as grandes obras que se planeiam.” -
(Anon)

1.1 Enquadramento

Em Dezembro de 2010 (no Ano Internacional da Biodiversidade), a assembleia das Nações Unidas, declarou 2011-2020 como a década das Nações Unidas, da Biodiversidade (UN 2011). Com a definição da estratégia *EU 2020 Biodiversity Strategy*, também a União Europeia (UE) se comprometeu com a proteção da Biodiversidade. Esta estratégia, adotada oficialmente em Maio de 2011, assume 2 compromissos fundamentais: parar a perda da Biodiversidade e dos consequentes *serviços prestados pelos ecossistemas* até 2020 e, simultaneamente, restaurar esses ecossistemas tão fielmente quanto possível.

Pese embora se esteja na era da história de maior riqueza quantitativa de Biodiversidade, paradoxalmente, é reconhecido por grande parte da comunidade científica que o ritmo da extinção de espécies tem sido maior agora do que durante qualquer outro período na história humana. Durante o último século tem-se vindo a observar um demarcado decréscimo significativo de riqueza biológica. A habilidade de determinar e prever alterações contínuas à diversidade e distribuição dos organismos na Terra e a sua relação com os fatores ambientais, são um fator chave para definir estratégias e mecanismos de conservação da Biodiversidade.

Em termos de Sistemas de Informação Geográfica (SIG) têm-se observado nos últimos anos diferentes iniciativas associadas à caracterização geoespacial e de produção de mapas da Biodiversidade. Estas iniciativas, normalmente, desenvolvem-se a nível local e passam pela modelação em função de variáveis bio-climáticas desse local e pela aplicação de modelos preditivos, o que manifestamente se tem vindo a mostrar insuficiente face às crescentes necessidades em termos de conhecimento de Biodiversidade.

Em 2008, concretizando as ideias e o trabalho de base inicial que lhe deram origem, foi lançado na Internet o projeto Naturdata. É um projeto colaborativo, independente e nascido de uma iniciativa particular, cujo objetivo principal é a recolha, produção e divulgação de informação sobre a Biodiversidade de Portugal. A oportunidade de desenvolvimento do presente Projeto surgiu do contacto do autor deste relatório com o projeto Naturdata, aquando da frequência do curso de Mestrado (MSc) em Ciência e Sistemas de Informação Geográfica (C&SIG). Dado o seu carácter inovador, a sua tipologia, o conhecimento envolvido e a sua filosofia, desde cedo despertou ao autor interesse e curiosidade. Não foi muito longo o passo que levou ao início da colaboração com a equipa do Naturdata, com a consequente sinergia de esforços e conhecimentos. Com efeito, a principal lacuna do

Naturdata identificada desde logo, foi a inexistência de informação de carácter espacial associado à caracterização das espécies e Biodiversidade.

Na Ciência de Informação Geográfica (CIG) é universalmente aceite que os SIG são um conjunto de ferramentas computadorizadas (*hardware* e *software*) que permitem a recolha, armazenamento, pesquisa, transformação e apresentação de dados e informação referenciada espacialmente. Permitem a integração de bases de dados (BD) alfanuméricos com uma componente gráfica (base cartográfica e mapas temáticos) (Longley *et al.* 2005). Com a vulgarização do acesso à *Web*, tem-se registado um crescimento da importância dos SIG em diversas vertentes, sendo aberta uma nova linha de perspetivas para pessoas que necessitam de utilizar esses dados (Sousa *et al.* 2005). A necessidade da criação de sistemas independentes de um equipamento padrão ou sistema é assim cada vez mais premente (Miranda 2004).

Deste modo, um dos principais desenvolvimentos ao nível de ferramentas SIG, atualmente, tem sido a disponibilização de dados e informação geográfica (IG) via Internet. O objetivo é que o utilizador tenha acesso a grandes quantidades de informação, de forma simples e a custos reduzidos (Machado *et al.* 2002). Os *SIGWeb* permitem, assim, combinar duas poderosas tecnologias: os SIG, analisando e integrando IG e a *Internet*, fornecendo conectividade a um nível global. O resultado desta sinergia resulta numa maior facilidade em encontrar e disponibilizar dados, partilhar ferramentas analíticas e no facto de ambos poderem chegar a um muito maior número de utilizadores. Utilizando todos os dados espaciais existentes de forma eficiente e efetiva, os SIG desempenham um papel crítico. Não apenas na disseminação de dados em bruto, mas também pela disseminação informação geoespacial útil, constituindo uma mais valia para os potenciais utilizadores.

Pairava no horizonte a possibilidade de conferir uma nova dimensão ao conhecimento, a dimensão geográfica. Os SIG, não sendo tecnologias recentes, são tecnologias em permanente evolução e atualização. É uma área de aplicação transversal a diversas vertentes do conhecimento, que procura descrever, explicar e prever padrões e processos, a partir de objetos georreferenciados, isto é, objetos geográficos representados num determinado sistema informático que liga com exatidão esses objetos a uma determinada posição na superfície terrestre, pela atribuição de um determinado posicionamento e sistema de coordenadas. Ao georreferenciar objetos e associar-lhe um vasto conjunto de informação sob a forma de base de dados geográfica (BDG), estas tecnologias afiguravam-se para permitir, em termos de Biodiversidade, um vasto conjunto de possibilidades e operações com a informação disponível, tais como: prever distribuições de espécies, obter nova informação a partir de informação existente, realizar consultas à informação disponível, visualizar graficamente informação georreferenciada, realizar operações relativas a BD para construir nova informação, modelar sistemas, entre outras.

Por outro lado, o conceito da *Web 2.0* introduziu um novo paradigma relacionado com o modo como a informação é criada, ordenada, descoberta e atualizada. Veio trazer a oportunidade de diferentes utilizadores se juntarem em comunidades que podem desenvolver uma consciencialização coletiva e providenciar o acesso à informação a cada indivíduo. Aplicando este conceito aos SIG, Goodchild (2007) refere que esta é cada vez mais um canal que permite aos utilizadores de forma simples e amigável a criação de redes de *sensores humanos* que criam séries de dados originais de informação geográfica voluntária (IGV) com um vasto potencial de aplicação. Chega mesmo a descrever a IGV como *um caso particular dos conteúdos gerados por utilizadores para Web (Web 2.0) em que os utilizadores produzem informação referenciada espacialmente*.

Pode-se dizer que se deu o primeiro passo na construção de uma ferramenta de potenciação do conhecimento da Biodiversidade. Integrando estes conceitos, pretendeu-se desenvolver o: ***Biodiversity Open Map (BiodOM)***, para posterior integração no Naturdata.

1.2 Objetivos

O principal objetivo do Projeto, foi o desenvolvimento e a implementação do sistema aplicacional de *SIGWeb do BiodOM*, para posterior integração no portal Naturdata. Recorreu-se exclusivamente a tecnologias de *software* geoespacial livre e aberto (SGLA). Pretende-se que os colaboradores e utilizadores do portal partilhem de um sistema comum de disponibilização e integração de informação geoespacial de Biodiversidade, de modo a que, numa perspetiva de SIG colaborativos para produção de IGV, se dê outra dimensão ao conhecimento fornecido pelo portal, bem como um novo conjunto de funcionalidades.

No que respeita ao SIG de suporte foi necessário: definir o modelo de dados da informação geoespacial (e BDG); recolher, processar e inserir informação espacial sub-dividida, nomeadamente: a) pelas principais áreas temáticas do portal; b) pela tipologia de informação associada a cada situação particular em função do modelo de dados definido; e, desenvolver análise espacial específica que sirva os objetivos do portal Naturdata e globalmente da distribuição de Biodiversidade. No que respeita ao *SIGWeb* foi necessário: identificar as funcionalidades a disponibilizar pela *Internet*, nomeadamente atendendo à tipologia colaborativa do portal; e, identificar e escolher as diferentes tecnologias para construir o sistema aplicacional adequado.

Deste modo, o sistema desenvolvido pretende servir como um repositório da IG de Biodiversidade produzida, bem como gerir e partilhar essa informação de forma devidamente articulada por todos os colaboradores do portal e comunidade científica, servindo assim de suporte a uma nova dimensão do conhecimento da Biodiversidade. É pretendido, também, avaliar a fidedignidade e o potencial do SGLA, em contexto de trabalho colaborativo, afigurando-se o sistema aplicacional simultaneamente como uma ferramenta

de SIG colaborativos. Será necessário acautelar as devidas questões relativas à interoperabilidade dos dados do sistema.

Em termos de área de estudo, pretende-se que o sistema aplicativo seja o tão abrangente quanto possível, atendendo à área de atuação do portal Naturdata: o território Português. Pelo que o modelo de dados contemplará este aspeto, nomeadamente no que respeita aos sistemas de projeção adequados.

1.3 Premissas

Consideraram-se como premissas iniciais para o desenvolvimento do trabalho, as seguintes:

- O portal Naturdata está já em funcionamento pleno à data de desenvolvimento do Projeto, pelo que o sistema aplicativo a desenvolver deverá adaptar-se às especificidades existentes;
- A inexistência de um SIG associado ao portal, traduz-se em algumas limitações daquilo que é o objeto do mesmo, o conhecimento da Biodiversidade;
- Com a vulgarização do acesso à *Web*, tem-se registado um crescimento da importância dos SIG em diversas vertentes, sendo aberta uma nova linha de perspectivas para pessoas que necessitam de utilizar dados geográficos;
- O sistema aplicativo tem que ser baseado em SGLA, dado que para além de possibilitar a implementação de uma solução a custos reduzidos, constitui simultaneamente uma forma de todos os intervenientes no portal Naturdata intervirem mais ativamente na produção de conhecimento de Biodiversidade;
- O sistema aplicativo deverá ser amigável para os utilizadores, dada a possibilidade de utilização por parte de colaboradores com pouca experiência na utilização de ferramentas e tecnologias da área de conhecimento dos SIG;
- O sistema aplicativo deverá atender à natureza colaborativa do portal Naturdata na produção de conhecimento.

1.4 Metodologia e faseamento

Em termos gerais a metodologia adotada neste Projeto desenvolveu-se segundo as seguintes etapas: 1ª - Descrição e enquadramento da temática da Biodiversidade e do portal Naturdata; 2ª - Avaliação das tipologias de troca de dados e desenvolvimento do modelo de dados a implementar; 3ª - Pesquisa e revisão de literatura sobre SIG Colaborativos e diferentes SGLA disponíveis; 4ª - Desenho da arquitetura do sistema aplicativo; 5ª - Instalação e configuração do *software*, das suas dependências e dos seus requisitos de modo a garantir a sua operacionalidade; 6ª - Construção de repositório de

dados do projeto; 7ª - Realização de testes de operacionalidade e integridade ao sistema aplicacional; 8ª - Aplicação a uma "prova de conceito"; 9ª - Memorando sobre a avaliação do sistema, limitações reveladas e principais conclusões; 10ª -Redação da dissertação final.

1.5 Estrutura do Trabalho de Projeto

Esta dissertação está dividida em sete capítulos. O presente capítulo inicia-se destacando os principais objetivos e assunções do Projeto. Apresenta o panorama geral e uma contextualização do BiodOM. De seguida, o segundo capítulo evidencia o conceito e tendências atuais no domínio da Biodiversidade e a aplicação e uso potencial dos SIG neste domínio. Passa-se em revista o portal Naturdata antes da implementação do sistema aplicacional desenvolvido, enquadrando-o no contexto da Biodiversidade em Portugal.

No terceiro capítulo aborda-se a publicação de IG pela *Internet*. Explora-se o conceito de trabalho colaborativo e SIG colaborativo, fazendo-se a devida distinção entre os conceitos colaborativo e cooperativo. São discutidas as principais tendências atuais em termos de SGLA, bem como as vantagens custo-benefício inerentes à utilização ao projeto em questão.

No quarto capítulo aborda-se as principais tendências em termos de software SIGWeb aberto, fazendo-se a ligação à sua pertinência num contexto de SIG colaborativos. Analisam-se as principais tendências atuais deste segmento de *software* abordando diferentes alternativas que serviram de base à escolha para a construção do sistema aplicacional desenvolvido neste Projeto.

A abordagem ao modelo de dados adotado, o seu desenvolvimento e a discussão das opções tomadas no que se refere a este aspeto, é realizada no capítulo cinco. São ainda abordadas técnicas de modelação espacial que poderão ser utilizadas para produzir informação útil ao desenvolvimento do conhecimento biodiversificado.

O sexto capítulo é dedicado ao desenvolvimento e implementação do protótipo do sistema aplicacional BiodOM. Ilustram-se as principais características e os passos envolvidos no desenvolvimento da arquitetura do sistema, englobando: o SGLA adotado, cartografia temática, funcionalidades do sistema e interface do SIGWeb.

O último capítulo encerra a dissertação, discutindo as vantagens e limitações do sistema aplicacional proposto. Aborda diferentes perspetivas de desenvolvimento futuro realçando, simultaneamente, as oportunidades e a potenciação que o sistema poderá transmitir ao conhecimento no domínio da Biodiversidade.

2 BIODIVERSIDADE

“Para existir como nação, prosperar como estado e viver como pessoas, temos que ter árvores.” - Theodore Roosevelt

2.1 Introdução

Segundo o sítio *Web da Comissão Europeia (CE): Nature & Biodiversity (2012)* nos últimos 25 anos a UE construiu uma vasta rede de 26.000 áreas protegidas distribuída por todos os estados membros numa área total de 750.000Km², 18% da área total da UE. Esta rede, conhecida como *Rede Natura 2000*, é a mais larga rede de áreas protegidas do mundo e o testemunho à importância que os cidadãos da UE atribuem à Biodiversidade (tradução livre do original). Falar de Biodiversidade é falar do resultado de 3,5 mil milhões de anos de evolução. Apesar de a origem da vida não ter sido ainda definida com rigor pela ciência, diferentes evidências sugerem que “apenas” algumas centenas de milhões de anos depois da formação da Terra, esta já estava bem estabelecida no nosso planeta. Pese embora estarmos na era da história de maior riqueza quantitativa de Biodiversidade de sempre, paradoxalmente, é reconhecido por grande parte da comunidade científica que o ritmo da extinção de espécies tem sido maior agora do que durante qualquer outro período na história humana. Durante o último século tem-se vindo a observar um demarcado decréscimo significativo de riqueza biológica.

Alguns autores consideram que o planeta está mergulhado numa crise biótica quase certamente causada pela atividade humana. Esta crise é de uma amplitude e taxa sem precedentes e poderá levar à extinção, até ao final do século, de metade das espécies que se estima existirem (Wilson 1991, Sala *et al.* 2000, Loreau *et al.* 2006). Em 2007, o ministro Federal do Ambiente Alemão citou estimativas de extinção de 30% de todas as espécies até 2050 (Sigmar 2007). A habilidade de determinar e prever alterações contínuas à diversidade e distribuição dos organismos na Terra e a sua relação com os fatores ambientais é um fator chave para definir estratégias e mecanismos de conservação da Biodiversidade existente.

2.2 Conceito

O termo diversidade biológica ou Biodiversidade tem tido interpretações diferentes ao longo dos últimos anos. A Cimeira da Terra das Nações Unidas em 1992, definiu diversidade biológica como a variabilidade entre organismos vivos de todas as origens, incluindo os terrestres, marinhos e outros ecossistemas aquáticos, e ainda os complexos ecológicos dos quais fazem parte: incluindo diversidade dentro de espécies, diversidade entre espécies e diversidade de ecossistemas. Esta definição continua a ser utilizada na Convenção de Diversidade Biológica das Nações Unidas (Hawksworth 1996).

Atualmente, o termo define o grau de variação de formas de vida dentro de uma dada

espécie, um ecossistema, um bioma ou todo o planeta. Em parte é função do clima: os habitats terrestres tropicais são ricos em Biodiversidade, ao passo que as regiões polares suportam uma menor variedade de espécies. É comum encontrarmos cientistas que definem o termo como a totalidade de genes, espécies e ecossistemas de uma região (Larsson 2001, Davis e Cornwell 2012). Uma vantagem desta abordagem é que descreve a maioria das circunstâncias e apresenta uma visão unificada dos diferentes níveis de diversidade.

Um aspeto que ressalta da definição é que medir a Biodiversidade em determinado nível num grupo de organismos pode não corresponder exatamente à mesma Biodiversidade noutro nível de análise. No entanto, a título de exemplo, nos vertebrados terrestres a diversidade taxonómica e a diversidade ecológica apresenta uma correlação muito aproximada. Segundo a *International Union for Conservation of Nature* (IUCN) (2012a) até ao ano de 2010, estimava-se que cientistas de todo o mundo tivessem descoberto e descrito aproximadamente 1,7 milhões de espécies de animais, plantas e algas (Tabela 1).

Grupo	Estimativa de espécies descobertas		Estimativa de espécies por descobrir	
	(nº)	% (relativa)	(nº)	% (relativa)
Insetos	950.000	9,60%	8.950.000	90,40%
Plantas	270.000	41,54%	380.000	58,46%
Aracnídeos	75.000	9,20%	740.000	90,80%
Cogumelos	72.000	13,28%	470.000	86,72%
Moluscos	80.000	24,24%	250.000	75,76%
Vertebrados	56.000	47,86%	61.000	52,14%
Algas	40.000	9,09%	400.000	90,91%
Protozoários	30.000	12,50%	210.000	87,50%
Crustáceos	75.000	29,41%	180.000	70,59%
Outros invertebrados	120.000	23,08%	400.000	76,92%

(Fonte: IUCN - Red List of Threatened Species 2012b)

Tabela 1: Estimativa de espécies descobertas e por descobrir.

Sendo um número impressionante, não deixa de encerrar algumas nuances. Se por um lado os números apontam para que haja um maior número de espécies de mamíferos, aves e plantas coníferas, por outro lado os grupos dos insetos, aracnídeos e algas, rondando os 90% de novas espécies estimadas por descobrir, refletem que há ainda muito a fazer no domínio da Biodiversidade. Neste particular, destaca-se o grupo dos insetos, com um impressionante número aproximado de 9 milhões de espécies por identificar.

2.3 O valor (escondido) da Biodiversidade

Uma estimativa do valor da Biodiversidade é uma pré-condição necessária para qualquer

discussão sobre a distribuição da riqueza da Biodiversidade. Numa primeira análise o valor da Biodiversidade relaciona-se com o suporte aos “serviços” dos ecossistemas, que incluem: clima e sequestro de CO₂, polinização, prevenção da erosão dos solos, manutenção da qualidade do ar, purificação da água, entre outros, que indiretamente se relacionam com a atividade humana (Costanza *et al.* 1997). Uma equipa de investigadores liderada por Costanza (1997), levou a abordagem do valor da Biodiversidade a um extremo, tentando calcular o valor monetário dos “serviços ecológicos” prestados pela natureza. A ideia base era contabilizar o valor anual de mão de obra de diferentes atividades realizadas pelo homem que tentariam reproduzir funções de diferentes espécies na natureza: p.e., a mão de obra envolvida na polinização de plantas. O trabalho envolveu a análise de vários destes “serviços ecológicos” e chegou a uma impressionante estimativa média de 33 biliões¹ de dólares anuais, equivalente a duas vezes o produto interno bruto mundial.

Em termos de valor da Biodiversidade, este pode-se dividir em 4 categorias distintas: valor funcional - em que cada espécie tem um papel funcional num dado ecossistema; valor de uso direto - relativo ao uso direto de diversas espécies pela sociedade humana, como alimentos ou como matérias primas para produção de bens; valor de uso indireto - em que outras espécies são indiretamente utilizadas pela sociedade, p.e., criar abelhas em laranjas favorece a polinização das flores de laranja, resultando numa melhor produção de frutos; e, valor potencial - relativo ao valor que muitas espécies podem vir a ter futuramente em termo de utilização direta, p.e., espécies de plantas que possuem princípios ativos a partir dos quais podem ser desenvolvidos medicamentos. Numa perspetiva do valor direto da Biodiversidade para a atividade humana, esta pode observar-se a vários e diferentes sectores, nomeadamente: agricultura, saúde humana, atividade industrial e valores culturais.

Ao nível da agricultura, para se compreender a importância da Biodiversidade e, por conseguinte, o seu valor para o homem, ilustra-se um caso. Quando uma estirpe de um vírus específica de plantações de arroz (*rice grassy stunt virus*) atacou vastas plantações de arroz desde a Indonésia até à Índia, nos anos 70, foram testadas 6.273 variedades de espécies de arroz em termos de resistência ao vírus. Apenas foi encontrada uma espécie resistente, à data conhecida pela comunidade científica, apenas desde 1966. Esta espécie permitiu formar híbridos resistentes ao vírus e tem agora uma vasta distribuição. A Biodiversidade (diferentes espécies de plantas) constituiu um importante fator diferenciador quando uma espécie dominante foi atacada por uma doença (Lumrix 2009).

Muitos materiais industriais derivam diretamente de fontes biológicas, entre os quais se encontram materiais de construção, fibras, corantes, borracha e óleos. Complementarmente, a Biodiversidade é também importante para a segurança de recursos

¹ Na notação portuguesa de bilião

como: água, madeira, papel, fibras e alimentos. Deste modo, em função dessa importância relativa, a sua perda é um risco no desenvolvimento industrial e, consequentemente, uma ameaça para a sustentabilidade económica a longo prazo (WRI *et al.* 2006).

De igual modo, a observação de evidências científicas no impacto na saúde global decorrente da perda de Biodiversidade releva, por analogia, o potencial impacto na saúde humana decorrente desse fenómeno (Wilson 2002). Este aspeto está muito relacionado com o fenómeno das alterações climáticas, no sentido de que muitos dos riscos para saúde decorrentes de alterações climáticas associam-se, normalmente, a flutuações de Biodiversidade (Krishna 2010).

Outro aspeto importante que expressa o valor resiliente da Biodiversidade, é o facto de 80% da população mundial depender de uma forma ou de outra, de medicamentos com origem natural em termos de cuidados de saúde (Chivian e Bernstein 2008). A Biodiversidade providencia um papel de suporte importante na descoberta de novas drogas e princípios ativos e, consequentemente, na existência e disponibilidade de recursos médicos.

2.4 Tempo e espaço em Biodiversidade

Decorre da definição de Biodiversidade que esta não se distribui de modo uniforme mas, ao invés, tem uma variação significativa pela superfície terrestre em geral e uma variação dentro de diferentes regiões, em particular. Sabe-se que, entre outros fatores, a diversidade de todos os seres vivos depende de, nomeadamente: precipitação, temperatura, características do solo, aspetos fisiográficos, altitude e da presença de outras espécies.

A Biogeografia é o ramo da ciência que estuda a distribuição de espécies, organismos e ecossistemas no espaço e ao longo dos tempos geológicos. A pesquisa moderna em biogeografia, combina informação e ideias de muitos campos, desde as restrições fisiológicas e ecológicas da dispersão de organismos, até aos fenómenos geológicos e climatológicos que operam à escala global, em diferentes episódios de tempo evolutivo.

A biogeografia moderna também recorre aos SIG, nomeadamente para tentar compreender os fatores que podem afetar a distribuição espacial de organismos e para tentar prever as suas tendências de distribuição futura (Cavalcanti 2007). Frequentemente diferentes modelos matemáticos em combinação com técnicas de SIG são utilizados para resolver problemas ecológicos com componente espacial (Whittaker e Fernandez-Palacios 2007).

Um aspeto de realce refere-se ao facto de a Biodiversidade não ser estática. É um sistema em constante evolução, quer do ponto de vista das espécies, quer do ponto de vista de um único organismo. Um dos padrões bem conhecidos em ecologia é o fenómeno do enriquecimento da diversidade de espécies que ocorre desde os polos até aos trópicos no planeta. É designado frequentemente como gradiente latitudinal de diversidade. De um modo geral, verifica-se que regiões a baixas latitudes apresentam um maior número de

espécies do que as regiões a altas latitudes (IUCN 2012b). A justificação deste gradiente de Biodiversidade tem sido um dos grandes desafios contemporâneos da Biogeografia (Cardillo *et al.* 2005). Mais do que se querer atingir objetivos científicos, o conhecimento deste fenómeno pode-se revelar fulcral em variadas questões que afetam a humanidade, tais como: a disseminação de espécies invasoras, o controlo dos vetores de determinadas doenças e os efeitos das alterações climáticas na manutenção da Biodiversidade. O que, por outro lado, revela o muito que há ainda a desenvolver neste domínio do conhecimento.

É também comum neste domínio o termo *hotspot* ou ponto crítico de Biodiversidade. Caracteriza-se, genericamente, como sendo uma dada região com um número significativo de espécies endémicas. Podemos encontrar *hotspots* distribuídos por todo o planeta, sendo mais provável que a sua distribuição ocorra nos trópicos. A sua maioria são sobretudo áreas de floresta. Um exemplo de um dos *hotspots* mais conhecidos mundialmente é a floresta da Amazónia. Segundo dados bibliográficos, estima-se que habitem neste *hotspot* aproximadamente 20.000 espécies de plantas, 1.350 vertebrados e milhões de espécies de insetos, não ocorrendo metade delas em outro lugar do planeta (Myers 1990, Normile 2010).

É aceite na generalidade que o planeamento ao nível da conservação se deve focar em áreas ao invés de espécies (Margules e Pressey 2000). No entanto, de um modo geral, os investimentos em conservação continuam aquém daquilo que é necessário para manter mesmo que seja uma rede insuficiente de reservas e áreas protegidas (James *et al.* 2001, Davies e Cadotte 2011). A definição de prioridades é, assim, crítica.

A identificação de *hotspots* permite que se foquem os poucos recursos que a generalidade da humanidade coloca à disposição dos objetivos de conservação, de modo a maximizar o retorno desse investimento em conservação (Myers *et al.* 2000). Pese embora os *hotspots* se referirem a áreas de riqueza de espécies e de endemismo, na literatura é frequentemente utilizar a expressão de modo a considerar a evolução histórica e a diversidade funcional. Neste contexto os *hotspots* representam áreas que capturam um conjunto desproporcional de Biodiversidade para a área em questão. A identificação de *hotspots* para definir prioridades de conservação requer assim: a atribuição de um valor à unidade de conservação - a espécie; o somatório de valores de áreas em competição; e, atribuir uma classificação em conformidade (Davies e Cadotte 2011).

2.5 Investigação em Biodiversidade

2.5.1 Taxonomia e espécie

A espécie é um conceito fundamental da Biologia que designa a unidade básica do sistema taxonómico utilizado na classificação científica dos seres vivos. Apesar de terem sido propostas múltiplas definições com maior ou menor grau de precisão, a dificuldade em encontrar uma definição universal para o conceito levou, ao aparecimento do chamado

problema da espécie. Consequentemente adotaram-se definições mais flexíveis, em função das especificidades do grupo biológico a que se aplicado (Dobzhansky 1955, Queiroz 2005).

O conceito de espécie mais comum é o conceito biológico de espécie proposto por Dobzhansky (1955) e Mayr (citado por Queiroz 2005): *espécies são grupos de populações naturais que se inter cruzam ou têm esse potencial, e que estão reprodutivamente isolados de outros grupos*. Deste conceito resulta que a espécie será o *conjunto de indivíduos que partilham o mesmo fundo génico, sendo morfologicamente semelhantes e capazes de se cruzarem entre si em condições naturais, estando isoladas reprodutivamente de outros grupos semelhantes com os quais, quando se cruzam, não originam indivíduos férteis* (tradução livre do original) (Dobzhansky 1955, Margulis e Sagan 2003).

Do ponto de vista estritamente sistemático e taxonomia, é a hierarquia compreendida entre o género (ou o subgénero, se existir) e a variedade. Corresponde a cada um dos grupos em que se dividem os géneros e forma agrupamentos compostos por indivíduos que, para além dos caracteres genéticos, têm em comum outros caracteres pelos quais se assemelham e se distinguem das demais espécies. Aplicando este conceito, indivíduos de espécies diferentes não se cruzam por falta de condições anatómicas ou por desinteresse sexual.

Mayr (1963) definiu a classificação biológica como *o agrupamento de indivíduos numa série hierárquica de classes na qual, as classes semelhantes ou relacionadas num dado nível hierárquico, se combinam de um modo abrangente em classes mais inclusivas no nível hierárquico superior seguinte*, definindo-se a classe como *um agrupamento de entidades semelhantes* (tradução livre do original) (Mayr 1963, Judd et al. 2008).

A taxonomia, que faz parte da Biologia Sistemática, compreende os métodos e técnicas para agrupar e categorizar os organismos em grupos, tais como o género ou a espécie. Esses grupos são conhecidos como *taxa* (ou no singular *taxon*). A classificação moderna tem origem nos trabalhos de Carolus Linnaeus, que agrupou as espécies de acordo com as suas semelhanças físicas. Ao longo do tempo essas classificações foram revistas para integrar o princípio Darwinista da descendência comum. Mais tarde, no século XX, com a introdução de métodos cladísticos, surge a taxonomia filogenética na qual os organismos são agrupados com base na sua relação evolutiva, ignorando as semelhanças morfológicas. Mais recentemente a filogenética molecular, que utiliza a sequenciação de DNA, tem contribuído para várias revisões de classificação.

A classificação definida deste modo é hierárquica. O *International Code of Zoological Nomenclature* (ICZN), ou em Português, Código Internacional de Nomenclatura Zoológica, define categoria taxonómica como: *o nível, para fins de nomenclatura, de um taxon numa hierarquia taxonómica (p.e., todas as famílias são para fins de nomenclatura do mesmo filo, que fica entre super-família e subfamília)* (tradução livre do original) (Kirk et al. 2008). O

ICZN define 7 agrupamentos principais: reino, filo, classe, ordem, família, género e espécie. Os agrupamentos situados entre os 7 principais podem ser produzidos adicionando determinados prefixos tais como “super-”, “sub-” ou “infra-”. Assim, uma subclasse tem uma classificação entre classe e ordem, uma super-família entre ordem e família. De um modo geral, na prática, existem agrupamentos ligeiramente diferentes entre zoologia e botânica, incluindo subdivisões como a tribo (Knapp 2010).

2.5.2 Aferir Biodiversidade

Existem diversas formas de quantificar a Biodiversidade. Uma forma é a abundância relativa, que se caracteriza pela uniformidade com que os indivíduos se distribuem pela espécie numa dada comunidade. Porém é a denominada riqueza de espécies, que continua a ser a métrica mais comumente utilizada (Bonn e Gaston 2005, Fleishman *et al.* 2006). É o método mais simples de descrever comunidades e a diversidade regional (Magurran 1988 citado por Gotelli e Colwell 2001). Esta variável, conjuntamente com o número de espécies, forma a base de muitos modelos ecológicos da estrutura de comunidades (Gotelli e Colwell 2001).

Não sendo o objetivo desta dissertação a discussão da existência de métodos mais ou menos corretos que outros, ou que apresentem maiores ou menores vantagens em relação a outros, podemos afirmar que o método baseado na contagem de espécies do ponto de vista de uma abordagem mais pragmática, apresenta as seguintes características: a riqueza de espécies pode ser facilmente comparada entre diferentes locais e esquemas de conservação; variadas regiões apresentam já informação detalhada sobre a riqueza de espécies; a espécie representa uma entidade identificável e palpável que podem ser útil para angariar apoio público e alavancar fundos de apoio à conservação; e, as espécies fornecem um foco para implementação de políticas e legislação (Davies e Cadotte 2011).

A quantificação da riqueza de espécies é importante não apenas em termos relativos, de comparação básica entre diferentes locais, mas também para endereçar questões ao nível da saturação de comunidades locais colonizadas a partir de focos regionais (Cornell 1999 citado por Gotelli e Colwell 2001). A maximização da riqueza de espécies é, em última análise, um objetivo dos estudos relativos à conservação. As taxas de extinção de espécies são também, muitas vezes, calibradas em função de padrões de riqueza de espécies (Gotelli e Colwell 2001).

Uma parte significativa de estudos ecológicos de comunidades preconizam uma outra abordagem com enfoque na área e na amostragem. Nesta abordagem, grande parte das comparações ecológicas de Biodiversidade são comparações de densidade de espécies: o número de espécies por unidade de área (Simpson 1964 citado por Gotelli e Colwell 2001). Estudos deste género assumem o pressuposto que as amostras são retiradas de

populações de indivíduos que estão a níveis comparáveis de densidade. Porém, a densidade das espécies depende quer da riqueza de espécies, quer da densidade média de indivíduos (não contemplando espécies). Consequentemente a ordenação de comunidades pode ser distinta quando seriadas por riqueza de espécies por oposição à densidade de espécies (McCabe e Gotelli 2000).

É então legítimo questionar qual a métrica mais adequada: riqueza de espécies ou densidade de espécies. Para finalidades de conservação e problemáticas que se focam em áreas consideráveis, a densidade de espécies será, provavelmente, mais interessante, dado que mede o número de espécies numa dada área (Gotelli e Colwell 2001). Por outro lado, para modelos de teste e avaliações teóricas de predição em ecologia, a riqueza de espécies pode ser mais apropriada. Parte significativa dos modelos teóricos em ecologia das comunidades não contém termos explícitos relativos a áreas e densidade. Ao invés, as unidades de aferição utilizadas são a abundância (N) e as taxas de crescimento populacionais (dN/dt), as quais são alteradas por coeficientes *per capita* que descrevem interações com outras espécies (Gotelli 2001).

Importa no entanto salientar, que nenhuma das métricas é “o modo correto” de medição em Biodiversidade, apesar de os padrões de Biodiversidade serem demarcadamente sensíveis relativamente à medida utilizada. Algumas decisões relativas a questões de conservação podem ser complicadas, quando algumas áreas de reserva contenham densidades de espécies superiores e outras contenham uma maior riqueza de espécies. As perturbações (internas e externas) e os modelos de gestão que afetem a abundância de espécies devem ser considerados na escolha de entre essas áreas (Gotelli e Colwell 2001).

Tal como ilustrado anteriormente, a Biodiversidade prende-se com a variação espacial das espécies e indivíduos, bem como dos complexos ecológicos dos quais fazem parte. Inclui-se diversidade dentro das espécies, entre as espécies e de ecossistemas (Parviainen e Paivinen 1998 citados por Nangendo *et al.* 2002). O uso de índices de diversidade de espécies que inerentemente combinam riqueza de espécies com abundância relativa, é uma referência constante em diferentes revisões bibliográfica (Gotelli e Colwell 2001). O índice de Shannon ou índice de Shannon-Wiener (H) é uma medida da abundância e riqueza de espécies que permite aferir a Biodiversidade. Este índice atende quer à abundância de espécies como a riqueza de espécies e é sensível à importância da mudança nas classes mais raras (Spellerberg e Fedor 2003). É também um dos índices mais utilizado. Calcula-se pela expressão (Shannon 1948):

$$H' = - \sum_{i=1}^S p_i \ln p_i$$

(em que s corresponde ao número de espécies e p_i à razão de indivíduos de espécies i ,

dividido por todos os indivíduos N de todas as espécies). Adicionalmente pode-se também considerar o índice de Simpson (1949), como métrica de cálculo da dominância de espécies, uma medida do equilíbrio do espalhamento. O índice de Simpson, diminui com o aumento da Biodiversidade e é dado por:

$$D = \frac{\sum_{i=1}^S n_i(n_i - 1)}{N(N - 1)}$$

(em que n_i é o nº de indivíduos na i-ésima espécie e N equivale ao número total de indivíduos). Tuomisto (2010) define *verdadeira diversidade*, ou o número efetivo de indivíduos igualmente abundantes, referindo-se ao número de indivíduos igualmente abundantes necessários para que a abundância média proporcional dos indivíduos iguale a observada no conjunto de dados de interesse (na qual todos os indivíduos podem ser igualmente abundantes). É dada pela expressão:

$${}^qD = \frac{1}{\sqrt[q]{\sum_{i=1}^R p_i p_i^{q-1}}}$$

(em que o denominador é igual à abundância proporcional média dos indivíduos no conjunto de dados, calculado com a média ponderada generalizada com expoente $q - 1$, R é riqueza (o número total de indivíduos no conjunto de dados), e a abundância do tipo proporcional é o i-ésimo indivíduo de p_i). Em aferições de Biodiversidade é também frequente o recurso a métodos estatísticos e geoestatísticos para descrever e comparar quantitativos. Os métodos estatísticos permitem aferir da significância de interpretações visuais, ao passo que os métodos geoestatísticos mostram o modo como a quantidade de variação espacial é afetada (Nangendo *et al.* 2002).

Contudo, contagens de riqueza de espécies também apresentam limitações. Um *hotspot* de riqueza específica para um dado taxon pode não corresponder, ou corresponder de modo desajustado, a um *hotspot* de riqueza para outra espécie. Mais acresce que é frequente à escala biogeográfica a riqueza de espécies entre diferentes taxa mais altos na hierarquia frequentemente co-variarem de modo mais acentuado, sobretudo no gradiente latitudinal de diversidade (Hillebrand 2004). Pode ser explicado por séries similares de variáveis ambientais e climáticas (Currie *et al.* 2004). Os *hotspots* de Biodiversidade não nos fornecem a solução “milagrosa” para a crise atual no panorama da Biodiversidade, dado que não existe uma métrica que consiga agregar todos os aspetos da Biodiversidade que possamos valorizar. Contudo, é urgente o desenvolvimento de uma solução universal e integrada, que permita endereçar as questões da crise global da Biodiversidade e que permita a duplicação de esforços (Mace *et al.* 2000).

Na generalidade dos estudos de variabilidade espacial de espécies, um objetivo comum costuma ser a obtenção de mapas a partir de pontos de observação. Estas operações

(ICV) (tradução do Inglês *Life Science Identifiers*) para todos os nomes biológicos. Isto permitiria a todos os autores citar nomes de forma desambigua pelos variados meios eletrônicos, reduzindo de forma significativa os erros associados à enunciação das nomenclaturas de nomes e abreviações (COL 2012). Este processo foi iniciado por 3 grandes BD de nomenclaturas: a *Index Fungorum*, a *International Plant Names Index* e a *ZooBank*. Outras BD, que publicam dados de taxonomia, têm também vindo a utilizar os ICV na identificação de diferentes taxa, destacando-se o *Catalogue of Life* (COL 2012).

2.5.4 Protocolos e padrões de interoperabilidade em Biodiversidade

A informação sobre taxonomia abrange mais de 250 anos de pesquisas concretizadas em publicações científicas. Apesar da digitalização destes conteúdos ser um processo em curso, esta informação, para ser útil, precisa de ser convertida em BD estruturadas. Nas últimas três décadas, tem vindo a haver um crescente reconhecimento que a Biodiversidade é um bem de grande valor para as gerações presentes e futuras (CBD 2012). Estes 2 fatores combinados têm levado a um crescente número de iniciativas e projetos que visam a utilização de dados no domínio da Biodiversidade. A Informática para a Biodiversidade utiliza dados de coleções preservadas (museus de história natural e herbários), coleções vivas (jardins botânicos e zoológicos e culturas de espécies), bem como dados recolhidos em pesquisas (monitorização, mapas florísticos e faunísticos), ou ainda de iniciativas e eventos particular (p.e., observação de aves). Outra fonte de informação importante é a já referida literatura, que resulta de 250 anos de investigação e que continua a ser relevante nas pesquisas atuais. Os resultados de pesquisas no âmbito da sistemática e taxonomia, fornecem um vasto e crescente conjunto de dados que se estende a novos tipos de dados. São uma crescente evidência sobre os importantes atributos funcionais das espécies.

Um aspeto patente é que os esforços desenvolvidos para partilhar estes dados de diferentes fontes levaram, desde cedo, à compreensão que a recolha e armazenamento de dados não é, por si só, suficiente. Apesar de muitos dos atributos poderem ser partilhados, as séries de dados são de difícil integração ou ligação entre elas. Na base desta questão está a falta de vocabulários comuns e a diversidade de estruturas de dados, que impedem a partilha destes dados. Esta partilha é fundamental para facilitar a colaboração entre diferentes entidades e análises de larga-escala, fundamentais para a resposta adequada às questões emergentes relacionadas com a Biodiversidade. O desenvolvimento de padrões para a informação fornece, assim, uma representação consistente dos dados a serem representados, permitindo que se combinem dados de diferentes fontes, ao mesmo tempo que se minimizam perdas e duplicação de dados (Berendsohn *et al.* 2011).

A *Biodiversity Information Standards* – TWDG, isto é, a organização para os Padrões de Informação em Biodiversidade, opera para definir esses padrões no domínio da Informática

para a Biodiversidade². Foi originalmente estabelecida como a *Taxonomic Databases Working Group* (TWDG) (grupo de trabalho de BD taxonómicas) por diferentes instituições botânicas e projetos de todo o mundo (Anon 2007 citado por Berendsohn *et al.* 2011). O seu âmbito de atuação foi sendo alargado de modo a incluir diferentes grupos, o que levou mais tarde à reformulação do seu nome (TDWG 2007). Dois padrões desenvolvidos com foco no suporte para a partilha de dados de coleções e dados relativos a observações, são o ABCD (*Access to Biological Collection Data*, isto é, Acesso a Coleções de Dados Biológicos) e o DwC (*Darwin Core*, isto é, Núcleo de Darwin). Ambos foram ratificados pelo TDWG como esquemas XML³ padrão. O padrão ADBC foi lançado para agregar todos os atributos utilizados nos espécimes e coleções de dados resultantes de observação que possam ser fornecidos por sistemas de informação. É compreendido por aproximadamente 1200 elementos e atributos (incluindo algumas centenas que são descritores específicos de elementos, como p.e.: linguagens). Pese embora as coleções de dados utilizarem uma série variada desses elementos, não existem coleções que utilizem mais do que uma fração da sua globalidade. Este protocolo é utilizado diretamente pela *Global Biodiversity Information Facility* (GBIF) e pela *Biological Collection Access Service* (Berendsohn *et al.* 2011). O padrão DwC de um modo geral descreve a ocorrência de espécies e a existência de espécimes em coleções. É uma série de dados de definição de elementos mais pequena, também desenhado para servir de suporte à integração e partilha de dados primários de Biodiversidade, nos quais se incluem: nomes de organismos, taxonomias, informações de espécies, dados factuais, distribuição e literatura (Wieczorek *et al.* 2009).

Mas todos os esforços de modelação e padrões resultantes teriam aplicação pouco significativa, se as aplicações utilizadas por investigadores não pudessem importar e exportar dados padronizados. Houve uma necessidade de uma abordagem de fluxos de trabalho para converter e integrar dados e, simultaneamente, proteger os utilizadores das complexidades de dados e da sua estruturação. Com base neste pressuposto a *European Distributed Institute of Taxonomy* (EDIT) criou a plataforma EDIT para ciber-taxonomia. Esta plataforma serve de suporte a todo o fluxo de trabalho desde a importação e exportação de dados até à sua padronização (EDIT 2012).

Um dado notável é que, até há poucos anos, não existiam plataformas ou ferramentas para editar e partilhar dados sobre Biodiversidade, em geral, ou dados de base taxonómicos, em particular. A maioria das aplicações eram produtos personalizados, desenhados para lidar explicitamente com nomenclaturas taxonómicas. Não serviam para importar e exportar dados em formatos padrão, à exceção das BD da rede GBIF (GBIF 2012).

² Apesar de as questões relativas aos protocolos e padrões de interoperabilidade em Biodiversidade serem fora do âmbito prático do Projeto, como se verá no Capítulo 5, a adesão futura por parte do portal Naturdata a um destes padrões, influenciou algumas das decisões adotadas no modelo de dados e arquitetura do sistema

³ eXtensible Markup Language

2.6 Modelação e SIG em Biodiversidade

A teoria da hierarquia da Biodiversidade sugere que o que acontece a níveis mais elevados de organização ecológica, como é o caso da paisagem ou ao nível do ecossistema, irá restringir os níveis inferiores: p.e, o nível de espécie ou nível genético (Noss 1990). Esta hierarquia é composta pelos níveis: genético, espécies-população, ecossistema-comunidade e paisagem (ou nível regional). Adicionalmente estes níveis de Biodiversidade podem ser afetados tanto pela escala temporal, como pela escala espacial. Numa aproximação tradicional à quantificação em Biodiversidade, a primeira fase do processo para uma dada região envolve a estimativa da Biodiversidade nesse local, num dado momento. A segunda fase, a monitorização, normalmente consiste em estimar a diversidade biológica no mesmo local, mas em mais do que um período temporal, de modo a poder-se inferir sobre eventuais alterações. Pese embora este facto, há que referir que a escala continuará a ser um fator preponderante. A variação da Biodiversidade a uma escala grande de análise será menos propícia a induções de alteração temporal que a uma menor.

A pesquisa em Biodiversidade requer a medição e inferência sobre a localização e abundância de espécies. Dado que numa abordagem tradicional a elaboração de inventários exaustivos e compreensivos da realidade em estudo de determinada área, são difíceis de concretizar, a ciência tem recorrido, normalmente, a diferentes modelos indicadores da diversidade das espécies. Estes modelos baseiam-se na combinação de dados *in situ* com variáveis ambientais para preverem a distribuição de espécies numa dada área geográfica (Fook *et al.* 2009). Num estudo realizado por Wilson *et al.* (1996) identificou-se um conjunto de atributos de Biodiversidade que podem ser avaliados em cada nível de organização ecológica. Ao nível da paisagem os atributos que podem ser monitorizados incluem a identidade, distribuição e proporções de cada *habitat*, bem como a distribuição das espécies nesse *habitat*. Ao nível do ecossistema, a riqueza, igualdade e diversidade das espécies, assim como conjuntos e comunidades, são importantes atributos. Ao nível da espécie, a abundância, densidade e biomassa de cada população, são os atributos de maior interesse. No que respeita ao nível genético, a variabilidade genética de organismos individuais da população é o fator mais importante. Uma abordagem holística de quantificação e monitorização de Biodiversidade será mais abrangente se interpretar a Biodiversidade em todos os níveis organizacionais utilizando aproximações diferenciadas em termos de escala espacial e temporal (Wilson *et al.* 1996, Gaines *et al.* 1999).

Porém, os cientistas que trabalham com modelos preditivos de distribuição de espécies encontram uma dificuldade inerente. Trata-se da necessidade de acesso a largos conjuntos de informação geoespacial. Isto cria um desafio adicional de representar, gerir, armazenar e aceder, a dados geoespaciais relacionados com Biodiversidade. Este cenário aponta para a necessidade de uma infraestrutura computacional que suporte estudos colaborativos de

Biodiversidade, mas também que permita a partilha de dados, modelos e resultados (Ramamurthy 2006). A partilha de informação necessita da informação associada a repositórios, formatos de arquivos e informação sobre semântica. De igual modo a partilha de modelos de modelação de Biodiversidade, requer o reconhecimento da aplicabilidade de cada algoritmo à espécie modelanda. A colaboração entre investigadores é não só sobre a partilha e troca de dados, mas também sobre a comparação entre diferentes modelos científicos e resultados experimentais (Fook *et al.* 2009).

É discutido frequentemente na comunidade científica que não existe uma solução simples para o problema da integração de informação de múltiplas disciplinas científicas, em termos de Biodiversidade. Consequentemente, esta situação reflete-se na adoção de SIG para produção de mapas de Biodiversidade. Por outro lado Bowker (2000) afirma que no campo dos SIG é necessário ter em linha de conta e representar a diversidade ontológica das várias BD de Biodiversidade, de diferentes projetos em elaboração. Este autor vai ainda mais longe ao discutir as questões que emergem quando diferentes cientistas de diferentes disciplinas se juntam na tentativa de construir um panorama da Biodiversidade global (com o objetivo de pesquisa científica). Afirma que *de modo a construir tal mapa deverá existir uma base de acordo comum acerca dos padrões de dados, sistemas de classificação e infra-estruturas*. No entanto, demonstra que tal pretensão e acordo, raramente são alcançados e envolvem as necessárias mas difíceis trocas entre as necessidades dos campos científicos e agências governamentais (tradução livre do original) (Bowker 2000)

2.7 Conservação da Biodiversidade em Portugal e a Rede Natura 2000

Na UE e particularmente em Portugal, a Rede Natura 2000 é uma rede ecológica resultante da aplicação das Diretivas n.º 79/409/CEE (Diretiva Aves) e da Diretiva n.º 92/43/CEE (Diretiva Habitats). Tem como fim a conservação da diversidade biológica e ecológica dos estados-membros da UE, atendendo às exigências económicas, sociais e culturais das diferentes regiões que a constituem. O objetivo fulcral é contribuir para parar a perda de Biodiversidade e constitui o principal instrumento para a conservação da natureza na UE (ICNB 2005, EC 2012b). É composta por Zonas de Proteção Especial (ZPE), estabelecidas ao abrigo da Diretiva Aves, que se destinam essencialmente a garantir a conservação das espécies de aves e seus habitats, bem como pelas Zonas Especiais de Conservação (ZEC), criadas ao abrigo da Diretiva Habitats, com o objetivo expresso de *contribuir para assegurar a Biodiversidade, através da conservação dos habitats naturais (Anexo I) e dos habitats de espécies da flora e da fauna selvagens (Anexo II), considerados ameaçados no espaço da UE* (ICNB 2005).

O Plano Sectorial da Rede Natura 2000 (PSRN2000) surge neste contexto. Aprovado com a publicação da Resolução do Conselho de Ministros n.º 115-A/2008, de 21 de Julho, trata-se de um Instrumento de Gestão Territorial (IGT), de concretização da política nacional de

conservação da diversidade biológica, visando a salvaguarda e valorização dos sítios e das ZPE do território continental, bem como a manutenção das espécies e habitats num estado de conservação favorável nestas áreas. Vincula as entidades públicas, definindo orientações estratégicas e normas programáticas para a atuação da administração central e local assumindo-de, sobretudo, como um instrumento para a gestão da Biodiversidade (Resolução do Conselho de Ministros n.º 115-A/2008).

O PSRN2000 *foi desenvolvido a uma macro-escala para o território continental, que caracteriza os habitats naturais e semi-naturais e as espécies da flora e da fauna presentes nos Sítios e ZPE, e define as orientações estratégicas para a gestão do território abrangido por aquelas áreas, considerando os valores naturais que nelas ocorrem* (Resolução do Conselho de Ministros n.º 115-A/2008). Da listagem dos objetivos principais constantes no PSRN2000 e enumerados no relatório do Plano, dá-se particular destaque às seguintes: estabelecer o regime de salvaguarda dos recursos e valores naturais das ZPE e Sítios, orientando a uma macro-escala a fixação dos usos e o regime de gestão compatíveis com a utilização sustentável do território a efetuar, posteriormente, através da inserção das normas e orientações nos IGT que vinculam diretamente os particulares (planos municipais e planos especiais de ordenamento do território); e, representar cartograficamente, em função dos dados disponíveis, a distribuição dos habitats naturais e semi-naturais e das espécies da flora e da fauna, presentes no Sítios e ZPE (ICN 2006).

Numa primeira fase, no Volume II do PSRN2000, foi compilada e sistematizada a informação base relativa aos valores naturais das áreas (ZPE e Sítios) definidas. Procurou-se sistematizar e uniformizar a informação em fichas descritivas e cartograficamente. A cartografia apresentada resulta de um trabalho de tratamento e aferição da informação disponível, em vários formatos e escalas de levantamento, sendo a escala de apresentação: 1:100.000. Verificou-se que em algumas das áreas classificadas se registaram lacunas de informação, devidamente assinaladas no Relatório do PSRN2000, sobretudo respeitantes a alguns habitats e grupos de espécies (ICN 2006).

No Volume III encontra-se sistematizada a informação relativa aos Sítios (da Lista Nacional de Sítios e aos Sítios de Importância Comunitária) e às ZPE. Do total das 118 orientações de gestão identificadas para os valores naturais, foram selecionadas as 55 orientações passíveis de ser cartografadas. Realizou-se deste modo uma abordagem à leitura da incidência no território de um conjunto considerável de orientações de gestão, que coexistem numa mesma área em função dos valores naturais que nela ocorrem. De acordo com o Relatório do PSRN2000 *esta cartografia deverá ser utilizada apenas a título indicativo, como orientação e enquadramento genérico, não substituindo o confronto do uso e ocupação actual do solo com as fichas de caracterização de valores naturais, sempre que as opções concretas a adoptar exijam maior rigor que o possível para uma escala de*

referência de 1:100 000 (ICN 2006). Mais acresce, que as limitações ao uso desta cartografia têm ainda em conta o dinamismo inerente aos sistemas naturais, descrito anteriormente. Este fenómeno amplifica a inerente desatualização da informação de base, contribuindo para, ao longo do tempo, aumentar as lacunas de informação associadas aos levantamentos cartográficos produzidos.

2.8 O Projeto Naturdata

O projeto Naturdata foi lançado na Internet em 2008. Trata-se de projeto colaborativo, independente e nascido de uma iniciativa particular, cujo objetivo principal era a recolha, produção e divulgação de informação sobre a Biodiversidade de Portugal. Apenas alguns meses volvidos desde o seu lançamento na *Internet*, este foi ganhando uma rápida e notória popularidade. Este ganho de projeção e dimensão, tornou-o no mais significativo projeto do género conhecido em Portugal. A título indicativo o projeto Naturdata conta com 35.146 espécies seriadas e um total de 13.789 fotografias de espécies catalogadas. Atualmente, é composto por uma BD e pelo portal Naturdata sendo este último de acesso livre e gratuito (pode ser consultado em www.naturdata.com). Ambos são desenvolvidos através de colaboradores, em regime voluntário, que de alguma forma se interessam pelo tema da Biodiversidade e que queiram colaborar com o mesmo.

O núcleo principal, é a equipa de desenvolvimento do projeto. Encontra-se subdividida por áreas temáticas, p.e.: Entomologia, Botânica e Ornitologia, entre outras. Para cada uma das áreas temáticas, existe um grupo de trabalho que é formado por um ou mais coordenadores, normalmente com ligação ao meio académico, bem como um número indefinido de colaboradores. Os coordenadores têm a função de, como o nome indica, coordenar a informação que é produzida e disponibilizada para cada grupo e são assistidos pelos colaboradores de cada área. Presentemente conta com mais de 70 colaboradores, no seu todo (Naturdata 2012).

Em termos de funcionalidades do portal Naturdata, este apresenta um design moderno, elegante e simples, com elevada ênfase na usabilidade e interoperabilidade *Web 2.0*. Foi otimizado quer para computadores *Desktop*, quer para *tablets*, sendo composto por vários módulos que acrescentam funcionalidades úteis ao objeto de retrata. Imediatamente abaixo do logótipo, encontra-se o "menu de topo". É neste menu que se encontram os botões para as secções mais estruturantes do Naturdata: [entrada](#) - a página inicial com as novidades; [espécies](#) - acesso à consulta da BD e diretório de espécies; [equipa](#) - listagem dos colaboradores Naturdata; e, [contacto](#) - os contactos principais associados ao projeto. No lado esquerdo, encontra-se o "menu principal". Trata-se de um menu vertical que garante acesso a todas as restantes secções e páginas do portal. Os botões podem variar ao longo do tempo, de acordo com a organização das páginas e requisitos necessários. De entre alguns dos diversos módulos instalados, destacam-se: o [módulo Estatísticas](#) onde se pode

acompanhar o número de fotografias inseridas no Naturdata e o número de espécies válidas atualizado em cada instante, em que este número representa o número de espécies conhecidas para todo o território Português; o módulo "Diretório de Espécies" – que contém as “Últimas Fichas Atualizadas” onde é apresentada uma miniatura de cada uma das espécies cuja ficha foi alterada recentemente pelos colaboradores. No que respeita ao módulo "diretório de espécies", quando se entra nesta secção, são apresentadas várias opções de pesquisa e navegação. A primeira e mais rápida é a "pesquisa de palavras-chave", que permite localizar fichas individuais de espécies a partir do nome científico, nome comum ou outros termos chave. Nesta pesquisa pode-se também obter a lista de qualquer grupo taxonómico, seja género, família, ordem, classe, filo ou reino, bastando para tal digitar o nome correto desse grupo. Esta pesquisa pode ser usada em conjunto com o módulo de "restringir pesquisa" para navegar apenas dentro de determinado grupo (Naturdata 2012).

Associado à pesquisa no diretório, está a funcionalidade "restringir a grupos taxonómicos". Esta pode servir por si só como elemento de pesquisa, pois se restringirmos até ao nível de género, ele apresenta a lista das espécies desse género com respetivos *links* para as fichas. Pode ser também utilizado como uma pesquisa rápida à lista de grupos taxonómicos. A título de exemplo, se desejarmos saber quais os géneros que compõem a família das Rosáceas, se restringirmos até esta família, na caixa de seleção seguinte, irão aparecer todos os géneros da família e apenas esses (Naturdata 2012). Ao escolher um dos reinos, é aberta uma lista automática com todos os filos listados no Naturdata e uma miniatura representativa. Quando selecionamos em determinado filo, o portal abre as respetivas classes e assim sucessivamente, na hierarquia taxonómica. Com esta forma de navegar, em cada passo tem-se acesso apenas às categorias que se encontram dentro da categoria superior que se selecionou. Fica-se, também, com uma noção clara de que grupos se encontram ilustrados (os que não têm nenhuma imagem apresentam uma câmara com um sinal de proibição indicando que faltam imagens nesse grupo). Quando se chega ao nível de espécie, ao clicar num nome, abre a ficha correspondente. A Figura 1 ilustra um exemplo de ficha de espécie:

Tyto alba (Scopoli, 1769)

Reino	Animalia
Filo	Chordata
Classe	Aves
Ordem	Strigiformes
Família	Tytonidae
Género	Tyto
Nome Comum	Coruja-das-torres

Colaboradores
João Guilherme
Tomás Martins
Ricardo Lima
Carlos Rio
Roberto Araújo
João Tavares
Nuno Barros
Faisca







© Faisca
Zoom

Ecologia de Tyto alba

Origem
Distribuição
Habitat
Local
Situação
Guilde Trófica
Estatuto
Protecção

Morfologia de Tyto alba

Ainda não existem caracteres morfológicos associados à taxonomia de Tyto alba.

(Fonte: Naturdata 2012)

Figura 1: Exemplo de ficha de espécie do portal Naturdata.

2.9 Conclusões do capítulo

Este capítulo inicia com uma abordagem à temática da Biodiversidade, explorando-se as suas definições mais universalmente aceites, que culminam numa ilustração do seu valor e importância, aspetos cada vez mais reconhecidos atualmente. Neste contexto explica-se o desenvolvimento e implementação do portal Naturdata, que surge numa iniciativa particular mas com raízes no meio académico, com o objetivo de recolher, produzir e divulgar informação sobre a Biodiversidade de Portugal. Pretende ultrapassar algumas das lacunas de informação existentes atualmente neste domínio. Quando se faz a descrição de algumas das características e funcionalidades do portal, já em funcionamento, fica patente a inexistência de um sistema de informação associado à distribuição geoespacial da Biodiversidade o que, como se constatou no capítulo inicial, foi um dos fatores que abriu espaço à oportunidade de desenvolvimento do Projeto.

Outro aspeto chave abordado, foi o contributo da caracterização da informação ao nível da espécie em termos de investigação em Biodiversidade. É aqui que se inclui a Taxonomia, que fornece inúmeros parâmetros palpáveis, que importa ter em linha de conta e analisar. Mostrou-se também que a investigação moderna em Biodiversidade, começa a integrar a dimensão computacional, da qual se incluem novas iniciativas de abordagem à sua produção de mapas e distribuição geoespacial. Porém as abordagens que existem neste

sub-domínio da distribuição geoespacial da Biodiversidade, são muito incipientes e não existem ainda verdadeiras metodologias de produção de mapas, nem modelos definidos. Verifica-se ainda que o desenvolvimento de projetos de distribuição geoespacial de Biodiversidade, tendem a exigir um esforço humano considerável.

Por outro lado, surgem diferentes iniciativas de standards de classificação e taxonomia relativamente as espécies conhecidas e identificadas que, previsivelmente, irão assumir um papel futuro importante, na distribuição e partilha de conhecimento científico neste domínio. Pese embora estas iniciativas ainda não serem oficiais no mundo científico, será, pois, vantajoso o Projeto absorver alguns dos conceitos destas iniciativas, dado que, tal como enunciado, ao uniformizar a informação científica, são os *standards* que permitem que possa haver o seu intercâmbio entre diferentes organismos e entidades e deste modo contribuir para a evolução do conhecimento.

Explorou-se que em diferentes níveis de organização ecológica, isto é, em diferentes escalas de análise, existem diferentes atributos de Biodiversidade que podem ser aferidos. Daqui se depreende que o modelo de dados a definir terá que ter este aspeto da escala e a sua relação com o nível de organização biológica em linha de conta, para a finalidade pretendida com o Projeto. Em termos de SIG têm-se observado nos últimos anos diferentes iniciativas associadas à caracterização geoespacial e produção de mapas da Biodiversidade. Estas iniciativas, normalmente, desenvolvem-se a nível local e passam pela modelação de variáveis bio-climáticas desse local e pela aplicação de modelos preditivos.

Por último, destaca-se que em Portugal as iniciativas de conservação da Biodiversidade resultaram na criação de redes de áreas protegidas com legislação específica. Pese embora a existência destes zonamentos, previstos em termos de IGT, sabe-se que a escala aquisição da informação de base e, conseqüentemente a IG disponibilizada, pode ser melhor ajustada de modo a ter-se um melhor conhecimento da distribuição da Biodiversidade. Este facto torna-se mais evidente dado que é reconhecida a existência de lacunas de informação das áreas classificadas. Também os próprios documentos orientadores reconhecem que a cartografia disponível deverá ser utilizada apenas a título indicativo como orientação genérica. Como se viu ao longo do capítulo, a componente geoespacial é não só um fator chave em termos de conhecimento da Biodiversidade, como também um fator que através da aplicação de SIG, pode permitir realizar novas inferências futuras pela relação ecológica entre espécies associada à sua distribuição e, assim, permite produzir novo conhecimento. Deste modo, o desenvolvimento do presente Projeto assume uma pertinência considerável, porque pode vir a permitir ocupar uma lacuna atual de um domínio do conhecimento tão importante, mas ainda simultaneamente com uma vasta margem de progressão em termos de conhecimento.

3 INTERNET, SIG COLABORATIVOS E A PARTILHA DE CONHECIMENTO

“Os problemas mais significativos que enfrentamos, não podem ser resolvidos ao mesmo nível de pensamento em que estávamos quando eles foram criados” - Albert Einstein

3.1 Introdução

Alguns sítios *Web* de mapas mais populares como o *Google Maps*, o *Virtual Earth* e o *Yahoo! Maps*, proporcionam aos seus utilizadores a resposta à questão genérica “Onde é que está?” com a apresentação da resposta num moderno, interativo e detalhado mapa. Na sociedade atual, estamos cada vez menos restringidos à descrição textual de uma localização: “passando a bomba de gasolina é a 3ª casa à esquerda”. Analisando esta abordagem de um modo mais aprofundado, diferentes organizações descobriram que o fenómeno de mapear, nomeadamente online, pode ser um grande recurso para analisar padrões de dados. Paralelamente a denominada *Web 2.0* tem vindo a modelar a forma como produzimos conhecimento: utilizadores com interesses comuns organizados em comunidades que se auto-regulam e auto-democratizam, produzem informação de modo colaborativo, tendo a *Internet* como suporte. Assume-se assim como tendência natural a crescente pertinência dos SIG colaborativos e a produção de IGV, mas também a perspetiva de uso destas tecnologias para potenciação do conhecimento científico e não só.

3.2 SIGWeb e Web 2.0: a Web Geoespacial

O termo *SIGWeb* define um conjunto de serviços de IG baseados numa rede que utiliza diferentes formas de acesso à *Internet*⁴ para aceder a IG, ferramentas analíticas e diferentes serviços de SIG. A disponibilização e divulgação dos temas de forma dinâmica pela *Internet* através do *SIGWeb*, permite integrar, disseminar e comunicar IG visualmente (Barriguinha e Ribeiro 2008). O *SIGWeb* permite combinar duas poderosas tecnologias: os SIG, analisando e integrando IG e a *Internet*, fornecendo conectividade a um nível global. O resultado desta sinergia resulta numa maior facilidade em encontrar e disponibilizar dados, partilhar ferramentas analíticas e no facto de ambos poderem chegar a um muito maior número de utilizadores. Utilizando todos os dados espaciais existentes de forma eficiente e efetiva, os SIG desempenham um papel crítico, não apenas na disseminação de dados em bruto, mas também pela disseminação informação geoespacial útil, constituindo uma mais valia para os potenciais utilizadores. Contudo, para permitir aceder e utilizar dados geoespaciais, esses ambientes SIG têm que estar disponíveis ao público. Assim, a *Internet*, e as comunicações sem fios, fornecem uma plataforma ideal para dotar diferentes utilizadores de tecnologias

⁴ Apesar de serem confundidos frequentemente, os termos *Internet* e *World Wide Web* ou mais vulgarmente: *Web*, diferem. A *Internet* refere-se à rede massiva de redes e infraestruturas de redes que liga milhões de computadores globalmente formando uma rede em que cada computador pode comunicar com outro computador. A *Web* refere-se a um modelo de partilha e acesso da informação, utilizando o protocolo de comunicação HTTP, desenvolvido sobre a *Internet*. Os SIG, mais do que a *Web*, podem utilizar outros protocolos de comunicação.

SIG através de sistemas *SIGWeb* (Barriguinha e Ribeiro 2008).

O termo *Web 2.0* não tem uma definição genérica concreta adotada (De Longueville 2010). Por esse motivo torna-se importante ilustrar as suas características, de modo a que se possa compreender a sua natureza inovadora. Passam por parâmetros como: conteúdos gerados pelo utilizador em que a barreira que delimitava o consumidor e o produtor é muito ténue, surgindo muitas vezes o termo *prosumidor* e havendo lugar a uma comunicação bidirecional; modularidade, em que qualquer sistema pode agregar informação de fontes diferentes graças à implementação de *standards* de interoperabilidade; e, uma dimensão social em que os utilizadores não surgem isolados, mas começam a criar ligações entre eles de modo a poderem partilhar informação mais facilmente (De Longueville 2010). A *Web 2.0* veio trazer a oportunidade de diferentes utilizadores se juntarem em comunidades que podem desenvolver uma consciencialização coletiva e providenciar o acesso à informação, a cada indivíduo. Assim, introduziu um novo paradigma relacionado com o modo como a informação é criada, ordenada, descoberta e atualizada. O ganho de notoriedade dos últimos anos dos mapas com base na *Web*, está intimamente associado ao recente advento e implementação de tecnologias de *software* deste conceito (*Web 2.0*) (Hall et al. 2010).

Em contraste com o sucesso de iniciativas de mapas na *Web 2.0* tais como o *Google Maps*, o *Google Earth*, o *Yahoo! Maps* ou o *Bing Maps*, entre outras, a indústria principal dos SIG continuou a manter-se focada na produção institucional de mapas *Web*, organizacionais e orientados à investigação. No entanto, muitos *softwares* proprietários facilitaram a disponibilização das suas ferramentas com a sua integração em ambientes *Web 2.0* (Dangermond 2008). Não obstante, algumas alternativas de *software* de mapas para a *Web* (Hall e Leahy 2008, Steiniger e Bocher 2009) têm vindo a mostrar-se como alternativas populares para determinados grupos e organizações, com possibilidade de aceder a dados e com alguns conhecimentos de programação, mas aos quais falta a motivação para pagar licenças de *software* proprietário (Hall et al. 2010).

Neste contexto surge a possibilidade e a crescente motivação de incorporar dados gerados por utilizadores, que tem vindo a transformar a *Internet* numa vasta plataforma de rede, com potencial de desempenhar um papel fulcral na denominada: *Web* geoespacial (Erle e Gibson 2006). A definição de *Web* geoespacial caracteriza-se como o conjunto integrado de uma coleção de serviços *Web* relacionados geograficamente e o conjunto dados que abrangem diferentes regiões geográficas (Rouse et al. 2007). Num sentido lato, a *Web* geoespacial refere-se ao conjunto global de serviços e dados que servem de suporte ao uso de dados geográficos numa gama de aplicações de dado domínio, que utilizam a *Internet* como meio de disseminação e partilha desses serviços e desses dados.

3.2.1 Web Standards, OGC e Interoperabilidade

Os recentes desenvolvimentos em termos de tecnologias da Internet, associados à adoção em larga escala do paradigma dos serviços *Web*, bem como dos padrões de interoperabilidade, transformaram a *Web* num veículo de transmissão de informação geoespacial e de geoprocessamento. Um dos aspetos em foco da pesquisa académica em termos de *Web* geoespacial, tem sido os aspetos técnicos da abertura geoespacial da *Internet*, para que seja possível a informação ser pesquisada e acedida na *Web* utilizando a localização como parâmetro (Egenhofer 2002). O recente advento da *Internet* e de todo o *software* de SIG e SIGWeb foram os percursos da consequente necessidade crescente de armazenamento de IG. Surgem diferentes questões de incompatibilidade entre sistemas, relacionados com os formatos de dados específicos de diferentes produtores de *software*. Isto constituiu o principal desafio recente da interoperabilidade. Assim, uma das principais questões inerentes ao desenvolvimento da *Web* geoespacial foi a construção de *Web standards* (padrões) comuns e de protocolos que permitissem que a localização seja um fator de unificação na interrogação e apresentação de resultados de pesquisa da *Web*.

A *Open Geospatial Consortium* (OGC), a principal organização responsável pelo desenvolvimento e implementação de padrões e protocolos abertos de interoperabilidade geoespacial, na *Internet*, veio dar resposta a estas questões e sistematizar as questões de interoperabilidade. Trata-se de uma organização não lucrativa que visa a definição de *standards* para os dados geoespaciais e para serviços baseados em localização (OGC 2012). Considera-se que a interoperabilidade é a propriedade que se reflete na capacidade para diferentes sistemas trocarem informação entre si e integrar informação de sistemas diferentes, com significado. Decorre desta definição que a interoperabilidade implica a existência de conceitos comuns sobre os quais se apoia a troca da informação. A missão do consórcio incide diretamente sobre questões de interoperabilidade na *Web* geoespacial, isto é, segundo a OGC (2012) este tem vindo a desenvolver normas para a transmissão de dados e de serviços geoespaciais, independentes da plataforma física ou lógica em que os próprios dados se encontram armazenados. No âmbito dos objetivos do Projeto, de entre os padrões abertos da OGC, consideraram-se: *Web Map Service* (WMS), *Web Feature Service* (WFS) e *Web Feature Service Transaccional* (WFS-T). A sua consideração, constitui um passo importante em termos de interoperabilidade do BiodOM.

3.2.1.1 WMS

A implementação da especificação WMS fornece uma interface HTTP⁵ simples para requisitar imagens de mapas de uma ou mais BD espaciais distribuídas (OGC 2012). Um pedido WMS define uma ou várias camadas geográficas e áreas de interesse a processar. A

⁵ O HyperText Transfer Protocol (HTTP) ou Protocolo de Transferência de Hipertexto, é um protocolo utilizado por diferentes aplicações, nomeadamente o *browser*, responsável pelo tratamento de pedidos e respostas entre cliente e servidor na *Web*, que surgiu da necessidade de distribuir informação pela *Internet*

resposta a esse pedido é uma ou várias imagens georreferenciadas (devolvidas no formato .jpeg, .png ou .svg, entre outras) que pode ser apresentado na aplicação cliente, frequentemente o *browser*⁶. Este serviço especificação padrão da OGC, disponibiliza assim ao *software* cliente, uma imagem produzida a partir dos dados de base, de acordo com o nível de zoom e sistema de projeção solicitado, atendendo às especificações disponibilizadas pelo servidor. Dado tratarem-se de dados matriciais, a sua simbologia ao chegar ao *software* cliente encontra-se já pré-definida pelo servidor. Na eventualidade de se necessitar modificar a extensão geográfica da informação visualizada, será necessário realizar um novo pedido de dados ao servidor, que enviará nova imagem. É o serviço que melhor protege a propriedade dos dados, relativamente à informação de base residente no lado do servidor. No entanto, em virtude de o *software* cliente apenas receber uma imagem, as operações de manipulação dos dados são mais difíceis, quer em termos de simbologia, quer termos de uso, para realizar operações de geoprocessamento e análise espacial.

3.2.1.2 WFS

Neste serviço o servidor envia ao cliente dados em formato vetorial, na extensão geográfica solicitada previamente. Posteriormente, é função da aplicação cliente (quer se trate de um *browser* ou um software cliente SIG “tradicional”) a definição da sua simbologia e arranjo gráfico. Geralmente os dados são recebidos no formato *Geography Markup Language* (GML) (um formato padrão da OGC), podendo o *software* cliente manipular a informação recebida, gravar essa mesma informação ou exportar para outros formatos. Atendendo ao formato dos dados transferidos, também é possível efetuar operações de geoprocessamento e análise espacial sobre estes. Os dados geográficos de origem são de natureza vetorial, que podem ter associados atributos alfanuméricos. Apesar de ser um serviço que não protege de forma tão eficiente a propriedade dos dados, aumenta consideravelmente a sua versatilidade em termos de operações e manipulação.

Neste padrão da OGC, as operações que deverão poder processar-se entre cliente e servidor são: *GetCapabilities* - possibilidade de descrever as suas capacidades, que especificamente deverá ser capaz de indicar os tipos de dados que disponibiliza e quais as operações que suporta sobre esses dados; *DescribeFeatureType* - possibilidade, depois de solicitado, de descrever a estrutura dos tipos de dados que disponibiliza; *GetFeature* - capacidade de devolver os dados solicitados num pedido, o cliente deverá ter a possibilidade de especificar quais os atributos pretendidos, e efetuar restrições espaciais e não espaciais aos dados que pretende obter; *GetGmlObject* - dar resposta a um pedido de elementos “ligados” por XLinks aos seus próprios dados e, adicionalmente, quando um cliente solicitar dados, deverá poder especificar se deseja obter também os dados ligados por XLinks aos dados que solicitou; *Transaction* - Possibilidade de resposta a pedidos de

⁶ *Web browser*, *Internet browser* ou mais simplesmente *browser*, refere-se a um programa de computador que permite aceder, descarregar e visualizar recursos de informação da *World Wide Web*

transações, em que um pedido de transação é composto por operações que modificam os dados, isto é, criar, modificar ou eliminar dados geográficos; *LockFeature* - capacidade de processar um pedido de bloquear o acesso a uma ou mais instâncias de um elemento durante a duração de uma transação.

3.3 SIG de Participação Pública (SIGPP) e SIG Colaborativos (SIGC)

Um dos campos de pesquisa desde os anos 90 no domínio da SIG, tem sido a temática da permissão do acesso livre à informação geoespacial e as suas implicações. Diferentes publicações sobre as implicações dos SIG neste aspeto, levantaram questões acerca da potenciação de comunidades, acesso a dados, participação pública e a incorporação de conhecimento local em sistemas especializados. Estas publicações levaram a determinadas iniciativas no seio dos SIG de Participação Pública (SIGPP⁷). Este domínio tem-se vindo a desenvolver ao longo do tempo sendo nos dias de hoje mais frequentemente denominado (e talvez mais corretamente) SIG Participativo (SIGP). A premissa base dos SIGP é a potenciação das comunidades pela facilitação do acesso dessas comunidades a dados e tecnologias geoespaciais, mapas comunitários e operações de análise espacial, como suporte à tomada de decisão no âmbito de determinados projetos (Rouse *et al.* 2007).

Atualmente qualquer indivíduo dotado de um computador com acesso à *Internet* tem a possibilidade de visualizar e interpretar dados geoespaciais e acrescentar valor a essa informação, mesmo que não seja um especialista em tecnologias de SIG. Em 2011, ocorreu um sismo na costa do Pacífico: o *Grande Sismo do Este do Japão*. Atingiu a magnitude de 8.9 na escala de *Richter* e despoletou um grande tsunami que causou mediáticos acidentes nucleares no reator de *Fukushima*. A rede social *Twitter*, através *tweets geoetiquetados*⁸ de habitantes da região, forneceu informação vital em tempo real a outras pessoas e autoridades: comunicou ao mundo o acontecimento e o local preciso do acontecimento, apenas instantes depois, quando os meios tradicionais de comunicação social podem levar algumas horas até difundir informação adequada; permitiu rapidamente às autoridades identificar a magnitude (normalmente baseada em interpretação visual dos efeitos sentidos); e, foi determinante a fornecer localização automática de pessoas em zonas mais afetadas, permitindo uma mais eficiente atuação dos meios de socorro (Vrba e Wotawa 2011).

No que respeita a uma das primeiras abordagens de SIGPP, Harris e Weiner (1998 citados por Rouse *et al.*, 2007) demonstraram um sistema que integrando uma dada comunidade participativa num projeto de SIG, permitiu endereçar questões como a distorção do conhecimento local (relativamente a um SIG “normal”) e questões relacionadas com acesso

⁷ Termo derivado da sigla em Inglês: PPGIS – Public Participation GIS, muito universalizada no mundo dos SIG

⁸ Um tweet é uma mensagem de até 160 caracteres que se pode transmitir via *Web* pelo serviço *Twitter* que, em Agosto de 2009, passou a permitir anexar às mensagens enviadas pelos utilizadores a informação da sua localização geográfica, com base nos GPS de dispositivos móveis e serviços de localização dos *browsers*

diferencial, através da incorporação de conhecimento espacial local (fornecido pela comunidade) num SIG multimédia (Rouse *et al.* 2007). Neste exemplo, foi solicitado aos membros de uma comunidade local para criarem mapas mentais das áreas circundantes às respetivas aldeias, para posteriormente se integrar essa informação bem como texto, narrativas orais e fotografias, num SIG, o que veio a contribuir para um processo de reforma do planeamento de espaços (Rouse *et al.* 2007). Porém, convém ressaltar que neste exemplo apesar de ser possível integrar essa informação da comunidade num SIG, foi necessário continuar a contar com especialistas e fontes externas de informação, não só para desenhar o sistema mas também para a sua implementação.

Paralelamente à evolução dos SIGPP, a *Internet* (e consequentemente a *Web*) sofreram uma rápida evolução. Desde cedo, cientistas e investigadores descobriram o potencial de soluções baseadas na *Internet*, associada à massificação da rede (Kingston 2002, Rouse *et al.* 2007). A crescente disponibilidade de computadores com ligação à *Internet*, ao remover as limitações de alocar o SIG num sistema local central, providenciou a oportunidade de alargar o acesso a projetos SIGPP e, consequentemente, permitiu à comunidade interagir remotamente com o SIG. Por outro lado, os SIGWeb facilitam a incorporação de formatos multimédia, tais como: fotografias, vídeo e áudio (Rouse *et al.*, 2007). Kingston (2002) num projeto em que a participação *online* era utilizada pelos utilizadores para explorar individualmente o SIGPP e abordar questões do acesso a dados, demonstrou que apesar de alguns utilizadores não se sentirem confortáveis com a utilização da tecnologia, de um modo geral o SIGPP promoveu a participação pública. Apesar dos sistemas SIGPP assumirem o potencial de alargar o acesso de comunidades a ferramentas de SIG e aos SIG enquanto sistema global, estes sistemas continuam a requerer elevados conhecimentos técnicos e domínio dos mesmos para se poderem implementar.

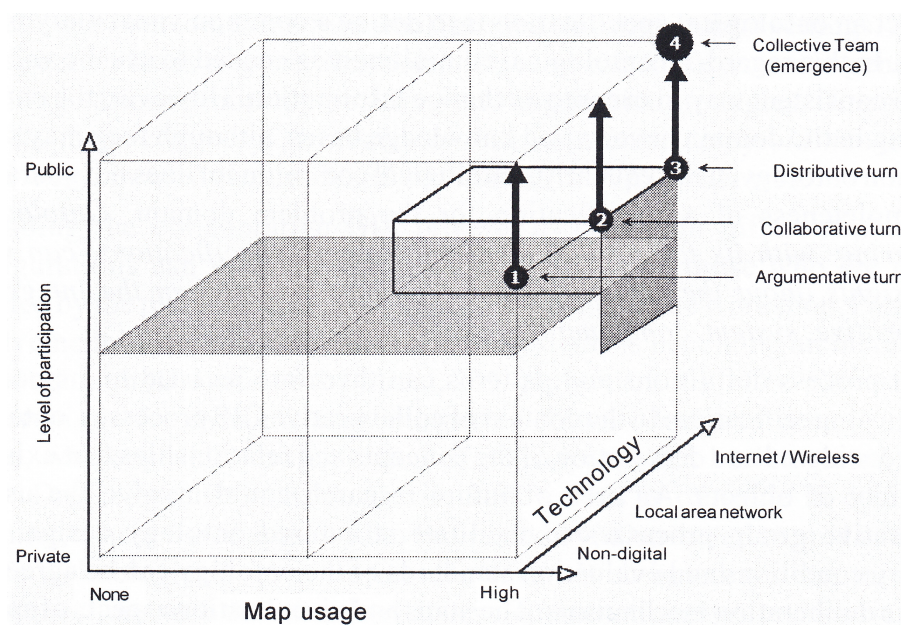
Segundo Mason e Dragičević (2006) os SIGWeb são frequentemente construídos e implementados com recurso a *software* proprietário⁹. As soluções de *software* SIG aberto para mapas *Web* requerem servidores *Web* e ligação à *Internet* para armazenar e servir projetos da tipologia SIGPP. Acresce que muitas aplicações de SIG com base em *software* aberto, tendo em vista a sua correta implementação, requerem conhecimentos técnicos avançados e experiência em desenvolvimento de *software* (Kishor e Ventura 2006). Consequentemente, um crescente número de projetos de SIGPP com base na *Web*, conduzidos por especialistas, têm tido sucesso, alargando o acesso à informação de diferentes SIG e aumentando a colaboração de comunidades no planeamento e tomada de decisão. Porém, esta tipologia de projetos tem vindo a ser limitada devido às especificidades técnicas e grau de conhecimentos SIG necessários à sua implementação.

Complementarmente, surgem os SIG colaborativos. Definem-se como a integração de

⁹ No capítulo 4 é abordada a noção de *software* proprietário

teorias, ferramentas e tecnologias, cujo foco principal é a estruturação da participação humana em processos de decisão espacial de grupo, não sendo em si limitadas por este processo. Os intervenientes são um misto entre utilizadores públicos e peritos no domínio, sendo as ferramentas computadores em rede (Balram e Dragičević 2006a, Coftas e Diosteanu 2010). Os seus resultados caracterizam-se não pela implementação de uma abordagem orientada à tarefa, mas antes por emergirem de contextos de exploração de problemas espaciais de forma estruturada, de modo a beneficiar o planeamento e promover eficientes processos de tomada de decisão (Balram e Dragičević 2006a).

Neste particular será pertinente distinguir entre “colaborativo” e “cooperativo”, uma diferença subtil, mas importante no contexto do Projeto. Diferentes autores (Jankowski e Neyerges 2001, Balram e Dragičević 2006a) caracterizam nos SIG colaborativos, diferentes termos, conceptualizados numa hierarquia de 4 níveis: comunicação, cooperação, coordenação e colaboração. Ao passo que a cooperação utiliza as ideias geradas no nível da comunicação para desenvolver um acordo global, apesar de os indivíduos poderem não interagir entre si, o nível colaborativo lida com a sensação de um significado e de um propósito ao nível do processo do grupo. Na participação colaborativa, o grupo interveniente, a tecnologia e os dados, operam com um único sistema. A integração de dados e conhecimento resultante de representação geoespacial, bem como o desenvolvimento de cenários, raciocínio crítico e aprendizagem adaptativa contínua, são aspetos essenciais dos SIG colaborativos (Jankowski & Nyerges 2001). A Figura 2 ilustra um *cube de SIG colaborativo*:



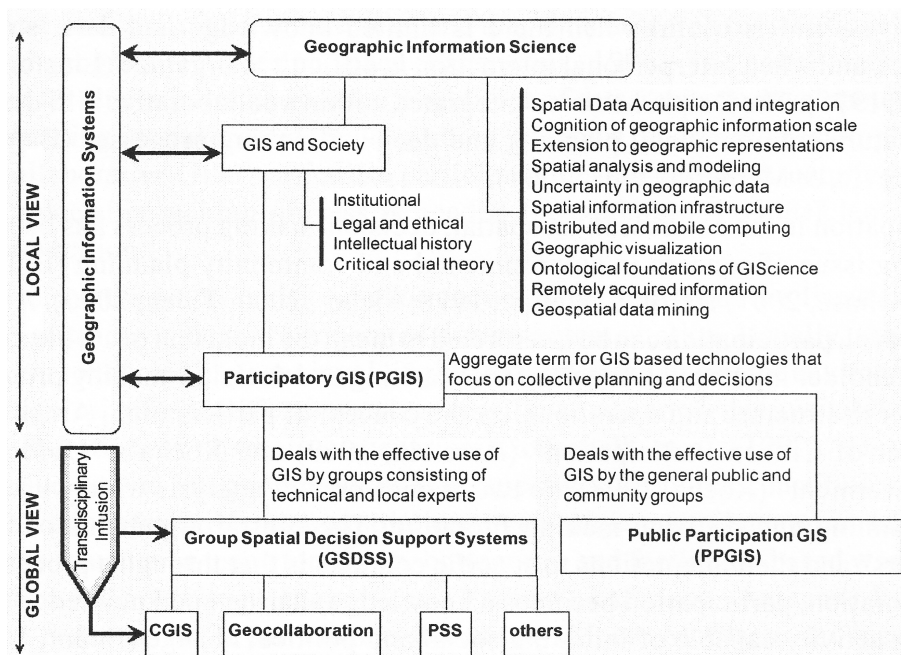
(Fonte: Balram e Dragičević 2006b)

Figura 2: Cubo de SIG Colaborativos.

Justapõe o nível de participação, utilização de mapas, e variáveis tecnológicas, bem como a

localização no cubo das possíveis tendências ao nível de planeamento e tomada de decisão: passo argumentativo, passo colaborativo, passo distributivo e o passo coletivo (de equipa). Os SIG colaborativos centram-se na secção do cubo no plano vertical sombreado. Pode-se observar as 2 tendências mais extremas a partir da análise do cubo. Por um lado o uso de mapas tradicionais em papel ou outros formatos não digitais, numa orientação mais focada ao nível do grupo e, num outro extremo, em que a tecnologia é maioritariamente digital, com tecnologias de rede utilizadas na partilha de informação no processo colaborativo. Neste último caso a utilização de mapas é elevada e inclui a exploração, análise e construção de cenários. Balram e Dragičević (2006b) afirmam que a recente e crescente necessidade de expandir a participação para incluir um espectro participativo mais diverso leva gradualmente a um nível distributivo no planeamento. Esta transição ocorre, primeiro, por adotar uma tecnologia apropriada (SIGWeb e Web 2.0) e depois a tecnologia adotada facilita a expansão para permitir um público mais vasto. Prevê-se que nesta fase possa ocorrer um passo coletivo, em que um grupo largo de intervenientes se funde com a tecnologia como um sistema único de apoio à decisão.

Em ciência, uma abordagem transdisciplinar em geral, lida com a integração e transformação de múltiplas perspectivas de múltiplas disciplinas, para criar novo conhecimento de modo a resolver problemas. Segundo Balram & Dragičević (2006b), pode ocorrer uma *infusão transdisciplinar* ao nível da CIG, do papel dos SIG na sociedade, dos SIGPP e dos sistemas de apoio à decisão, ao nível de grupos. Acresce ainda que os SIG intervêm em todos os níveis da hierarquia da CIG, tal como ilustrado na Figura 3:



(Fonte: Balram & Dragičević 2006a)

Figura 3: Estrutura conceitual de abordagem aos SIG colaborativos.

Nesta perspetiva, pode imprimir um novo dinamismo ao conhecimento que será mais rápido do que aquele que seria expectável com a difusão de novos métodos pela hierarquia depois de integrados na CIG. Um aspeto de realce em SIG colaborativos prende-se com a relação entre indivíduos e impacte dos resultados. Se se mantiver o problema-situação como constante para diferentes grupos estruturados de forma semelhante, variando sistematicamente a composição do grupo, pode revelar como os resultados das decisões espaciais representadas em mapas, são influenciadas por grupos de interesses específicos. Deste modo fica implícito, que o reconhecimento do efeito da composição do grupo pode justificar a robustez dos resultados esperados para determinado modelo. Desenhar um SIG colaborativo para lidar com diferentes combinações tempo-espço é um modo de aumentar o processo participativo para além dos intervenientes envolvidos diretamente no sistema. De acordo com Balram e Dragičević (2006a) o envolvimento de intervenientes em determinado SIG em sistemas colaborativos suportados pela *Web* tem sido uma solução viável em crescendo. A *Web*, eliminando barreiras espaço-tempo, permite criar oportunidades de integrar uma mais vasta participação de intervenientes em determinado SIG em torno de uma temática comum. Sobretudo em situações de planeamento comunitário, em que os resultados necessitam de ser criados, distribuídos e acessíveis a utilizadores menos tecnologicamente conscientes.

3.4 O conceito de Neogeografia

O termo Neogeografia foi inicialmente caracterizado por Scott (2006 citado por Turner 2006). O autor caracterizou um âmbito mais alargado de intervenção em SIG, para incluir artistas, *psicogeografia* e outros. Turner (2006) agregou ao termo os aspetos de origem mais técnica, traduzindo uma terminologia mais bem definida. O uso contemporâneo do termo deve muito da sua inspiração ao movimento dos meios multimédia de localização, que procuram expandir o uso de tecnologias baseadas na localização geográfica, para incluir a expressão pessoal e da sociedade. Descreve conteúdos gerados pelo utilizador que são adicionados a uma plataforma central de mapas *Web* (Turner 2006, Jackson 2006).

A Neogeografia é, assim, um termo ou domínio que se reporta a técnicas, ferramentas e práticas de geografia que têm sido usadas tipicamente, para além do âmbito de geógrafos profissionais e SIG profissionais (Turner 2006). De acordo com Szott (2006 citado por Graham 2009) caracteriza-se como um conjunto diversificado de práticas que operam à margem das práticas de geógrafos profissionais. Em vez de assumir notações sob a forma de padrões científicos, as metodologias de Neogeografia tendem para o intuitivo, expressivo e pessoal. No entanto, podem ser aplicações idiossincráticas de técnicas geográficas “reais” e, por isso, não quer dizer que estas práticas não sejam de utilidade para as ciências cartográficas ou geográficas. Sobretudo se houver mecanismos de controle da integridade da informação neogeográfica. Apesar disso, geralmente, não são metodologias que estejam

em conformidade com os diferentes protocolos e normas da prática profissional.

Significando numa interpretação mais literal "nova geografia", é frequentemente aplicado ao uso de técnicas geográficas e ferramentas utilizadas para atividades pessoais e da comunidade ou para utilização por um grupo de utilizadores não-especialistas. Os domínios de aplicação da Neogeografia são tipicamente não formal ou não analíticos (Turner 2006). Os SIG tradicionais, historicamente, têm desenvolvido ferramentas e técnicas orientadas a aplicações formais que exigem precisão e exatidão. Em contraste, a Neogeografia tende a ser aplicada a mais acessíveis áreas de aplicações de comunidades. Pese embora este fato, ambos os domínios podem ter alguma sobreposição, dado que os mesmos problemas geográficos são muitas vezes apresentados para diferentes conjuntos de utilizadores: especialistas e não especialistas (Graham 2009). Uma característica importante da Neogeografia que se tem vindo a observar na *Internet*, é que os utilizadores voluntários não estão apenas a investir o seu trabalho em criação de conteúdo. Verificou-se também que eles têm vindo progressivamente a desempenhar um papel de realce na edição, ordenação e categorização de conteúdos (Graham 2009).

Os *global positioning systems* (GPS), isto é, sistemas de posicionamento global, em particular, têm permitido que variados tipos de dados espaciais possam ser criado por utilizadores não profissionais (Brunn *et al.* 2004, Dykes 2006). As tecnologias de GPS têm incentivado a produção de mapas personalizados e transformaram os movimentos quotidianos em expressões criativas que podem ser carregados e compartilhados com outros utilizadores. Além disso, tecnologias espacialmente conscientes não se limitem apenas a dispositivos dedicados GPS conectados à Internet através de computadores.

A comunicação através de mapas tradicional era realizada num sentido unidirecional: dos produtores dessa cartografia para os utilizadores de mapas. Porém a crescente utilização *online* de mapas de base e o recurso a API¹⁰ abriram o caminho à emergente comunidade de "Neogeógrafos" que vieram democratizar a produção de mapas de tal modo, que a distinção entre produtores e consumidores se apresenta muitas vezes ofuscada (Turner 2006, Hall *et al.* 2010). Atendendo a que muitos destes novos utilizadores e produtores não têm formação sobre princípios cartográficos ou geográficos, fica implícito que vários séculos de experiência no que concerne aos domínios cartográficos de generalização, simplificação e conteúdos temáticos de cartografia, entre outros, ou serão irrelevantes para o desenvolvimento de uma comunicação eficiente com mapas ou, ao invés, essas competências estarão embutidas e codificadas no próprio *software* (Goodchild 2008).

3.4.1 Desigualdade de informação geográfica em Neogeografia

Inerente à Neogeografia surge o facto de apesar existir uma multiplicidade de utilizadores que estão a criar enormes quantidades de conteúdos na *Internet*, sobre uma multiplicidade

¹⁰ A sigla API designa *application programming interfaces* e aborda-se mais em detalhe no capítulo 4.

de lugares físicos, isto não significa que não haja uma geografia distinta para a produção deste conhecimento. Fontes variadas de bibliografia (Warf 2001, Zook *et al.* 2004, Recabarren *et al.* 2008) demonstram que as redes físicas de utilizadores dedicados à produção de conteúdo na *Internet* são caracterizadas por geografias altamente desiguais. Não será de estranhar que um olhar mais superficial sobre diferentes projetos de Neogeografia, revelem por vezes análises que não sejam representadas com um grau de detalhe adequado. Em alguns desses projetos, as caracterizações de diferentes lugares são provavelmente sub-representadas devido a barreiras tecnológicas, económicas e educacionais, enfrentadas por pessoas com profundo conhecimento sobre esses lugares. Um exemplo, pode ser observado na Coreia do Norte (um dos países com os níveis mais baixos de acesso à Internet) em que se pode observar que apenas alguns detalhes foram adicionados, em termos de pequena escala, ao projeto OpenStreetMap (OSM).

3.5 Informação Geográfica Voluntária (IGV)

Devido aos elevados custos relativos à produção de IG, durante muitos anos esta foi considerada uma área reservada a agências oficiais e grandes organizações. No entanto, a *Web 2.0* veio revolucionar os conteúdos gerados pelos utilizadores. Nos últimos anos, apoiada pelas tecnologias *Web 2.0*, tem vindo a surgir uma nova fonte de IG sob a forma de conteúdos *Web* gerados por utilizadores (Vossen e Hagemann 2007). Dois casos particulares e mediáticos são a Wikipedia e o OSM. Constituem dois dos muitos esforços de construir mapas globais com recurso a contribuições voluntárias dos seus utilizadores. Outro caso, menos mediático, é o *Cornell Lab of Ornithology* que constitui um esforço de *ciência cidadã*, sendo parte deste projeto a produção de IG (COL 2008, Elwood *et al.* 2011).

A evolução dos últimos anos, favorável a uma rede colaborativa e social com base no modelo *Web 2.0*, tem tido um impacto na natureza e forma dos SIG (colaborativos) e nas questões mais alargadas da produção de IG, a sua utilização e as suas características (Elwood 2008). Esta inovação tecnológica veio permitir que os utilizadores contribuam para adquirir e partilhar informação que dificilmente poderia ser recolhida e organizada no passado. Fritz *et al.* (2009) revelam uma variedade de terminologias para descrever a criação de conteúdos geoespaciais criados por utilizadores. Termos como *crowdsourcing*, IG contribuída colaborativamente, SIGPP via *Web*, mapas *Web 2.0*, Neogeografia e IGV, têm sido utilizados. Diferentes autores (Goodchild 2007, Coftas e Diosteanu 2010) referem-se a esta produção IG por parte de utilizadores apoiados pela *Web 2.0* como IGV.

Goodchild (2007) refere que a IGV incide sobre os dados geoespaciais, que são voluntariamente criados por utilizadores, na maioria das situações, sem formação nas disciplinas de cartografia, geografia e áreas afins. Esta informação é construída através de uma “*ciência coletiva*”, que consiste em redes de observadores amadores com apetência para serem qualificados e treinados. A produção de IGV, em que a distinção clássica entre

profissionais de SIG e outros utilizadores se apresenta cada vez mais ofuscada, tem sido muitas vezes identificada como uma forma de Neogeografia (Turner 2006, Graham 2009, Rana e Joliveau 2009). Os Neogeógrafos têm sido potenciados pela disponibilização em larga escala de equipamentos de posicionamento, imagens aéreas de alta resolução (sem custos) e diferentes *softwares* cartográficos que permitem produzir mapas às necessidades individuais (por oposição a mapas mais generalistas da cartografia tradicional) (Goodchild 2009, Hudson-Smith *et al.* 2009). Ilustra-se a possibilidade do importante papel que a IGV tem em pesquisas e investigações de profissionais de SIG, bem como as diferentes linhas de investigação científica em curso no que concerne aos domínios dos benefícios deste fenómeno e da compreensão dos seus impactes sociais (Elwood *et al.* 2011).

O interesse na interação com mapas e dados de mapas na *Internet* tem vindo a produzir um crescente inventário de IGV disponível (Goodchild 2007, Hall *et al.* 2009). Esta convergência entre novas tecnologias *Web* interativas e crescentes práticas de conteúdos gerados por utilizadores e disseminados pela *Internet*, está a gerar esta nova forma de IG. Na sua base encontram-se os utilizadores que, em função de diferentes motivações, recorrem a recetores GPS de mão, como telemóveis e outros, para recolher dados e contribuir para vastas séries de IG. Utilizando depois interfaces de mapas baseadas na *Web* para marcar e anotar figuras geográficas ou adicionar localizações geográficas a fotografias, texto e outros conteúdos média partilhados *online*. A IGV representa assim uma mudança de paradigma do modo como a IG é criada e partilhada e por quem é criada e partilhada, bem como no seu conteúdo e características (Elwood *et al.* 2011).

Neste âmbito, é pertinente considerar a existência de uma nova orientação generalizada na criação de dados espaciais. Passa da criação tradicional de dados geoespaciais por profissionais num laboratório ou local de trabalho, para utilizadores não especialistas, de variados âmbitos e valências, em redes de informação globais. Segundo Sieber (2006) e Hall *et al.* (2009) as formas de interação alteraram-se com a disseminação do fenómeno da IGV. Este compreende o uso dos SIG para alargar o envolvimento do público em geral nos processos de tomada de decisão política, na promoção dos objetivos de organizações não governamentais (ONG) e organizações baseadas em comunidades. Acresce ainda que a forma relativamente facilitada de interação com aplicações de mapas na *Web* tem servido para reinventar o modo como as pessoas exploram espaços geográficos (Elwood 2006).

Comparando a IG produzida de modos mais convencionais com a IGV, esta última apresenta diferenças a diferentes níveis, que importa considerar: o conteúdo e teor da informação, as tecnologias de base envolvida na sua aquisição, questões em torno da sua qualidade, os métodos e técnicas para trabalhar esta informação e os aspetos sociais que medeiam entre a sua criação e os impactes da sociedade da mesma (Elwood *et al.* 2011). O modo de como o rápido crescimento da IGV pode afetar a utilização de informação

geoespacial e respetivas tecnologias nos processos de participação pública é uma matéria complexa que só muito recentemente os investigadores começaram a dedicar atenção (Hall *et al.* 2009). Conceptualmente, numa perspetiva de orientação à pesquisa, o fenómeno da IGV relaciona-se com o fenómeno dos SIG colaborativos e, sem dúvida, integra a área mais abrangente da CIG. Porém as pesquisas bibliográficas atuais ainda não esclarecem o grau de extensão desta interligação.

Facto é, que este fenómeno tem contribuído para disponibilizar um vasto conjunto de dados, que constituem em si uma fonte rica e imediata de informação para uma variedade de fins. A IGV é de interesse não apenas para investigadores no domínio dos SIG, mas também apresenta potencial para contribuir de modo significativo para vários ramos da ciência: Biodiversidade, Geografia Física (Lawrence 2006), Geografia Humana (Zook e Graham 2009) e Geografia Educacional (Moulder 2009). *A integração destas múltiplas fontes de IG pode fornecer aos geógrafos uma oportunidade sem precedentes de conduzirem investigação numa variedade de tópicos, a várias escalas* (Elwood *et al.* 2011) (tradução livre do original). Não obstante, esta abordagem *bottom-up* de criação de IG pela contribuição de utilizadores apresenta também grande potencial de integração com as abordagens *top-down* mais clássicas de produção de informação (Elwood *et al.* 2011).

3.5.1 Cartografia colaborativa

Um dos aspetos mais proeminentes dos SIG colaborativos que tem vindo a ganhar forma através da *Web* geoespacial, é a potenciação das comunidades através da elaboração coletiva de mapas. Grupos e indivíduos começam a ter em mãos a possibilidade de recolher e adicionar dados a uma plataforma central de *Web mapping*. Ao trabalhar em conjunto, é possível produzir e recolher mais informação que um único indivíduo ou grupo isolado, com custos muito reduzidos para o global da comunidade envolvida. Um exemplo, já enunciado anteriormente, é o OSM. Trata-se de um projeto iniciado em 2004 no Reino Unido cujo objetivo primordial é construir um mapa global que não seja restringido por restrições proprietárias ou de direitos de autor (Rouse *et al.* 2007).

As possibilidades criadas de permitir a utilizadores comuns carregarem trilhos pedestres ou rotas de automóveis, adquiridos com dispositivos pessoais, bem como georreferenciar as suas fotos online, ou até partilharem experiências e opiniões relativas a locais que frequentaram, têm vindo a enriquecer a caracterização das especificidades de localizações, que grandes companhias não dispõem de recursos ou motivação para recolher e publicar (Goodchild 2007). Estas novas ferramentas de mapas na *Web* têm tido o importante papel de dar a conhecer ao público, em particular aos utilizadores da *Internet*, sobre geografia, informação geoespacial e geovisualização. Diferentes projetos têm dado enfoque a locais com significância para diferentes indivíduos e estas várias perspetivas informam todos nós.

3.5.2 A visualização de informação cartográfica na sociedade da informação atual

Têm-se observado em diferentes projetos inovações em termos de programação *Web*, para criar plataformas de mapas facilmente personalizáveis. Estas fornecem-nos as camadas de informação base, tais como redes de estradas e ortofotografias. A partir desta informação os utilizadores podem gerar aplicações próprias de mapas que combinam os seus dados com a cartografia de base fornecida. Este método conhecido como *mashup* tem vindo a crescer significativamente e é um elemento caracterizador da *Web 2.0*. (Rouse *et al.* 2007)

Pressupõe-se que talvez um dos impactes da disponibilização ao público de plataformas de mapas pela *Web* tenha sido o aparecimento de um interesse em tecnologias geoespaciais mais para além dos meios académicos e indústria SIG. O fenómeno do *mashup* tem vindo a demonstrar, que existindo as ferramentas adequadas, utilizadores das mais variadas áreas têm vindo a contribuir na criação de aplicações de mapas para a *Web* que ligam a propriedade da localização a uma variada série de dados de informação. (Rouse *et al.* 2007). As API podem ser usadas livremente por qualquer pessoa com os conhecimentos suficientes para criar um *sítio Web* e mapas personalizáveis através de um *mashup* entre os seus dados e cartografia de sites de mapas *Web*. Tal como foi sugerido recentemente na CNET *os mapas estão a evoluir para um nexa histórico de tecnologias dispare e comunidades que, juntos, estão a alterar o uso principal da Internet, bem como a redefinir o conceito de mapas na nossa cultura. Os mashups de mapas vão fornecendo uma ideia clara, ainda que de aplicações comerciais, para a geração da chamada tecnologia social que eles representam* (tradução livre do original) (Mills 2005).

Permanecem por resolver algumas questões relacionadas com o nível de conhecimento técnico necessário para implementar este tipo de aplicações que permitem a interação com dados geoespaciais, bem como a origem da informação cartográfica de base ter condições de uso restritas. Acresce-se a grande maioria das aplicações envolvidas ser proprietária e por vezes acarreta custos elevados. No entanto, é inequívoco que estas soluções (incluindo as API) dão um passo para uma rede de conhecimento geoespacial acessível a todos. Rouse *et al.* (2007) afirmam que com o surgimento da *Web* geoespacial os projetos SIG começam a pouco e pouco a sair da sala de aulas para uma mais vasta circulação.

3.5.3 A qualidade da informação e o conceito de vago em IGV

Pode-se supor que se um dado aplicativo foi desenhado de modo semelhante à *Wikipedia* (2012) e que sendo as suas entradas de informação, até determinado ponto, monitorizadas por outros voluntários, sendo assim aberto a reter informações adicionais por qualquer um utilizador que discorde desses mesmos dados, a aplicação pode-se tornar-se bem sucedida do ponto de vista de produção de informação voluntária. Em 2005, numa investigação realizada pela revista *Nature*, sobre o uso da revisão por parceiros, que pretendia comparar a enciclopédia *Wikipedia* produzida de modo voluntário, e a enciclopédia *Britannica*,

mostrou que a diferença de precisão entre os dois não era grande (Giles 2005). Apesar da aplicação de projetos de IGV para a aquisição de IG produzir, normalmente, dados menos precisos, é considerada com uma opção viável em casos em que os métodos clássicos são caros. Um projeto de IGV típico não implica custos e permite não apenas juntar grande quantidade de informação num período de tempo relativamente curto, mas também manter essa informação atualizada. É também frequente um utilizador que habite nas proximidades de uma determinada figura mapeada, poder facilmente descobrir e corrigir informação incorreta dentro do projeto de SIG de IGV (Coftas e Diosteanu 2010).

Alguns sistemas populares de IGV, já anteriormente referidos, tais como o OSM ou o *Wikimapia*, herdaram a visão da orientação ao objeto em termos de IG, o que implica uma certa precisão espacial. Dependendo do tipo de IGV em questão a precisão espacial desta informação pode ser alta. Isto aplicar-se-á a situações particulares, como é o caso de se a informação referenciar uma morada ou for criada utilizando um aparelho de GPS (Goodchild 2007). Contudo, as percepções das pessoas relativamente aos fenómenos ambientais e de Biodiversidade podem não corresponder a uma morada conhecida. Pode também acontecer situações em que as pessoas não tenham uma noção precisa da extensão do fenómeno ambiental que estejam a observar dado que, muitas vezes, essa percepção pode ocorrer de memória ou a sua percepção desse fenómeno atingirá apenas uma parte do todo. As suas percepções ou observações de determinado fenómeno podem ser vagas. Nos sistemas de IGV existentes têm-se vindo a observar lacunas no que concerne a este aspeto da vaguidade espacial, quer quando a informação é introduzida por um parceiro, quer quando essa informação é pesquisada e disponibilizada por outros (De Longueville *et al.* 2009).

Importa ter em conta a distinção entre percepções e medidas, dado que as percepções dos intervenientes em IGV são muitas vezes vagas por definição. Os intervenientes na aquisição de dados de IGV *não são sensores digitais que enviam valores numéricos através de uma rede de sensores* (tradução livre do original). São seres humanos que interagem com uma máquina, computador ou dispositivo móvel, para partilhar as suas percepções (Goodchild 2007). O trabalho de Fisher (2000) sobre dados geográficos à base das percepções dos utilizadores, permite aferir que *os intervenientes vivem num mundo mergulhado de vaguidade em que realizam funções de modo eficiente e em que pensam sobre a geografia e espaço como conceitos vagos* (tradução livre do original). Deste modo, o conceito de vaguidade de informação, é um conceito que faz parte da experiência humana no domínio geográfico. É importante que uma BDG incorpore essa vaguidade na interação com os utilizadores (De Longueville *et al.* 2009). Porém os SIG são, tradicionalmente, baseados em modelos orientados ao objeto não compatíveis com a integração do conceito de vago na sua modelação (Cross e Firat 2000 citados por De Longueville *et al.* 2009). Estes modelos orientados ao objeto permitem integrar qualquer objeto da vida real, como uma unidade

espacial com informação sob a forma de atributos alfa-numéricos, simultaneamente com métodos que permitem realizar operações com esses objetos. Assim, esta informação não é considerada vaga. As componentes espaciais (p.e.: pontos, linhas e polígonos) e respetivos atributos contém um valor preciso. Em alternativa o modelo matricial é utilizado frequentemente para descrever variáveis espaciais contínuas, em que a cada pixel é atribuído um valor (De Longueville *et al.* 2009). Porém os sistemas que utilizam IGV, normalmente, utilizam o primeiro modelo para codificar, armazenar e disponibilizar dados.

Para endereçar esta questão, diferentes autores (Fisher 2000, Dilo *et al.* 2007, De Longueville *et al.* 2009) referem-se ao conceito de grau de verdade (*degree of truth*), para conciliar o modelo orientado ao objeto com a vaguidade de informação. Na base deste conceito está uma lógica de multi-avaliação que visa substituir a visão da lógica booleana das características de objetos. É possível expressar com o grau de verdade que um dado objeto pode ter a tendência de apresentar determinadas características numa certa extensão (expressas numa escala contínua de 0 a 1), ao invés de afirmar que esse mesmo objeto terá esta característica (1 = verdadeiro) ou não terá esta característica (0 = falso), patentes na lógica booleana. Segundo Fisher (2000) a aplicação deste conceito permite-nos dizer que se “A está perto de B” tem um grau de verdade superior quando a distância entre A e B é 10Km, do que se essa mesma distância for 1000Km. Este conceito tem utilização frequente em diferentes técnicas de análise espacial: um determinado conjunto matricial mostra-nos a distribuição espacial dos graus de verdade de um determinado parâmetro (Dilo *et al.* 2007). Em suma, para lidar com as questões de informação vaga no domínio da IGV, apresentam-se duas técnicas de modelação, que integram o conceito: objetos com atributos específicos, que expressam um grau de relação para expressar a relação de vaguidade sobre as suas características espaciais e não-espaciais, que podem ser utilizados nesses sistemas; e, camadas matriciais que contêm informação relativa à vaguidade da informação nas fronteiras geográficas de um dado objeto (De Longueville *et al.*, 2009).

3.5.4 IGV como *input* de dados geoespaciais para investigação

Sabe-se que a IG é uma componente importante das infraestruturas de variadas ciências que variam da geofísica à antropologia. É, também, essencial na administração pública e na vida quotidiana de cidadãos. É por esse motivo que estas novas e importantes abordagens à criação, compilação e partilha de vastos dados com informação descritiva, tem o potencial de causar impacto em vários domínios do conhecimento científico (Elwood *et al.* 2011). O domínio da Neogeografia, aponta no sentido destes novos desenvolvimentos de aquisição de IG. No entanto, segundo Elwood *et al.* (2011) é legítimo argumentar se as competências, ferramentas e redes teóricas de geógrafos profissionais são ou não essenciais para endereçar muitas das questões de fundo associadas à IGV. Incluem-se as questões do potencial enquanto fonte de dados para pesquisa e investigação, questões da qualidade dos

dados e o papel da IGV em diferentes métodos de pesquisa científica (Elwood *et al.* 2011).

A experiência em estudos piloto mostra que se se endereçar suficiente atenção à escolha e formação dos participantes, voluntários e utilizadores que integrem projetos de IGV, estes podem fornecer contribuições úteis para a aquisição de dados científicos de base. Observa-se também que em diferentes projetos de IGV que operam sem atender a este controlo, a informação produzida não reúne o design rigoroso e as garantias de qualidade que a comunidade científica exige, o que reforça a necessidade desta abordagem (Elwood *et al.* 2011). Porém, nem sempre são necessários instrumentos de questionários de larga-escala ou amostragem altamente estruturada e protocolos de observação. Diferentes estágios no processo de pesquisa e investigação científica requerem diferentes tipos de dados e, consequentemente, de informação, assim como diferentes modos de aquisição dessa informação. Frequentemente são utilizadas observações preliminares para formular hipóteses conceptuais, seleccionar locais de estudo e estratificar populações. Em muitas situações a informação proveniente de IGV é adequada a esses fins (Elwood *et al.* 2011).

Em alguns desses projetos a IGV constitui uma fonte apropriada de informação. Tem também demonstrado ser útil a endereçar questões do domínio da pesquisa e investigação que envolvem a percepção humana, conceitos de lugar e espaço e outras conceptualizações difíceis de endereçar através de outras técnicas e fontes de informação, mais tradicionais. A título de exemplo, Girardin *et al.* (2009 citado por Elwood *et al.* 2011) utilizaram fotografias geotiquetadas colocadas no sítio Web Flickr.com como fonte primária de informação para reconstruir o movimento turístico em diferentes locais. Noutro exemplo, Graham (2009) analisou IGV para localizar comunidades pela pesquisa de palavras chave partilhadas.

3.6 Integração de IGV em Infraestruturas de Dados Espaciais (IDE)

Pierce *et al.* (2009) evidenciaram recentemente como é que a integração dos conceitos da Web 2.0 em ciber-infraestruturas irá contribuir para a criação de comunidades científicas. Na sua pesquisa, aponta que tanto as comunidades de pesquisa de Infraestruturas de Dados Espaciais (IDE) e ciber-infraestruturas, estão a caminhar no mesmo sentido, de uma perspectiva distribuída mais técnica para uma perspectiva mais social, através do desenvolvimento de comunidades virtuais que aperfeiçoam a colaboração. Os geoportais constituem-se como um elemento que desempenha um papel fulcral nesta visão colaborativa das IDE. O conceito de *portal Web* está intimamente relacionado com o conceito de recursos. Maguire e Longley (2005) caracterizam os geoportais como *sítios Web que funcionam como portas de entrada para uma determinada coleção de recursos, incluindo séries de dados, serviços, notícias, tutoriais, ferramentas e uma coleção organizada de ligações a outros sítios Web através de catálogos*. Em termos de recursos geoespaciais Yang *et al.* (2007) definem a aplicação do termo geoportal, quando estes envolvem quer dados geoespaciais quer serviços geoespaciais, sendo uma categoria

particular de portais *Web* que lida, sobretudo mas não exclusivamente, com recursos geoespaciais. Consequentemente, são a *front end* de uma longa cadeia espacial de informação suportada por uma IDE (Bernard *et al.* 2005). De facto, de toda a cadeia de componentes que constituem uma IDE moderna, os geoportais são a componentes mais “exposta” ao mundo exterior, atuando como a componente que expõe dados, metadados e serviços *Web* que permitem potenciar os utilizadores.

É notório que estes exemplos demonstram o alcance da informação que integra a IGV. Por um lado, observam-se infraestruturas de dados que descrevem a localização de objetos sobre a superfície terrestre, tradicionalmente produzidos por organizações governamentais e empresas particulares, mas agora complementados e ampliados pela IGV. Numa outra perspetiva, as agregações de dados, tais como observações por utilizadores de determinadas condições ambientais, eventos ou atividades, nas suas respetivas localizações. Normalmente, estes dados são adquiridos pela georreferenciação *online* de conteúdos multimédia de natureza não cartográfica. Estas duas vertentes de IGV constituem um complemento ou alternativa produtivos, às formas tradicionais de IG, estendendo as formas tradicionais de produção de informação (Elwood *et al.* 2011).

Segundo Elwood *et al.* (2011) a IGV tem o potencial de resolver diferentes constrangimentos e omissões que são recorrentes em várias IDE. Numa abordagem tradicional, isto é, numa abordagem prévia à proliferação da IGV, as estruturas de dados eram criadas e mantidas por organizações governamentais ou empresariais, que poderiam não disponibilizar esses dados ao público. Muitas vezes, quando disponibilizados eram de custos elevados ou implicavam várias restrições ao seu uso. Diferentes destas convencionais estruturas de dados, são também atualmente criadas e atualizadas com recurso a deteção remota. No entanto, os dados provenientes de deteção remota apresentam restrições de aquisição, não só devido às órbitas dos satélites como também devido à presença em certas situações de nuvens e fumo que obscurecem os objetos. Deste modo, muitas vezes, não constituem uma fonte adequada de informação para limites administrativos, moradas e outros pontos de interesse que são tão importantes para diferentes serviços. Noutras situações ainda, verifica-se que dados não integrantes de estruturas de dados, isto é, desenvolvidos através de observação direta (tais como certos levantamentos realizados por governos locais para determinar o uso do solo ou as condições de determinadas propriedades), são limitados por constrangimentos temporais e orçamentais, que leva a séries de dados incompletos ou imprecisos, sobretudo em situações de alterações frequentes (Elwood *et al.* 2011).

3.7 SIG Colaborativos e Informação Geográfica Voluntária (IGV) aplicados à distribuição da Biodiversidade

As questões ambientais atuais são afetadas por vários fatores de dimensão física,

ecológica, sócio-cultural e política. Esta complexidade surge cada vez mais como um desafio para a gestão ambiental, planeamento e processos de tomada de decisão. Mason e Dragičević (2006) sugerem que as abordagens que envolvam um processo colaborativo e gestão de conhecimento podem fornecer uma via eficaz para gerir a complexidade dos sistemas ambientais. De igual modo, a complexidade dos sistemas naturais e do problemas ambientais, bem como as suas soluções, requerem que uma mais vasta secção de participantes seja integrada no processo de tomada de decisão (Balram e Dragičević 2006b). Esta participação através de diferentes configurações espaço-tempo e o modo de as possibilitar de modo eficiente é um dos desafios dos SIGWeb modernos. Para configurações síncronas, os processos de SIG colaborativos embebidos são estruturados com mecanismos de enfoque no grupo e exploração de ferramentas digitais para partilha, anotação, análise e visualização de dados e conhecimento num espaço de solução conjunto. Em configurações assíncronas para se estruturar o processo de discussão on-line são frequentemente utilizadas abordagens que envolvem mapas e SIGWeb colaborativos.

Das fundações aqui apresentadas, relativamente às várias linhas de análise da IGV, é legítimo dizer que esta constitui uma profunda transformação sobre o modo como conhecemos o mundo, produzimos e validamos conhecimento e, ao fazê-lo, reproduzimos e desafiamos predicados como a precisão ou a fiabilidade que guia a produção de conhecimento, na investigação (Elwood *et al.* 2011).

3.8 Conclusões do capítulo

Este capítulo inicia-se com a procura de uma caracterização das novas tendências emergentes de SIG distribuídos pela *Internet*, que constituem a *Web* geoespacial. Ilustra algumas das suas vantagens, nomeadamente uma maior facilidade em encontrar e disponibilizar dados, partilhar ferramentas analíticas e no facto da combinação das tecnologias de SIG e da *Internet* permitirem chegar a um número maior de utilizadores. Com o advento da *Web* 2.0 esse efeito é ainda mais potenciado, dado que surge a oportunidade de diferentes utilizadores se juntarem em comunidades, que podem desenvolver uma consciencialização coletiva e providenciar o acesso à informação a cada indivíduo.

Demonstrou-se que a adoção de padrões de interoperabilidade abertos está intimamente relacionada com o desenvolvimento da interoperabilidade da *Web* geoespacial que, consequentemente, potencia a transferência de dados entre diferentes sistemas. Por outro lado, explora-se o conceito de SIG colaborativos, um conceito mais integrador do que o conceito de SIGPP, enquanto sistemas que permitem a potenciação de comunidades em torno de uma temática comum, pela facilitação do acesso a dados e tecnologias geoespaciais, consulta de mapas interativos e realização de operações de análise espacial. Essas próprias comunidades intervêm simultaneamente como produtores e consumidores de informação geoespacial, estruturando a participação humana em processos de apoio à

tomada de decisão. Demonstra-se que a adoção destes sistemas é muito potenciada pela introdução dos conceitos e práticas da *Web* geoespacial e que, normalmente, os participantes nestes sistemas são um misto entre utilizadores públicos e peritos na matéria. Demonstrou-se ainda que é possível aproveitar as oportunidades que as ferramentas disponibilizadas pela *Web 2.0* fornecem, para integrar utilizadores locais na recolha de IG e participarem em debates sobre o conhecimento territorial das áreas em que habitam.

Esta abordagem leva a um novo conceito de um novo domínio: a Neogeografia, que se reporta a técnicas, ferramentas e práticas de geografia que têm sido usadas tipicamente, para além do âmbito de geógrafos profissionais e SIG profissionais. Esta inovação tecnológica, apoiada na *Web* geoespacial, leva também a uma nova forma de produção de dados. Utilizadores comuns passam a contribuir para adquirir e partilhar IGV, que dificilmente poderia ser recolhida e organizada no passado, ou que, devido aos seus elevados custos durante muitos anos foi considerada uma área reservada a agências oficiais e grandes organizações. Abrem-se novas perspetivas de produção de informação para novas áreas do conhecimento que levam, inevitavelmente, a novas formas de visualização de informação cartográfica, na *Web*. Combinam-se diferentes camadas de informação de diferentes fontes, para construir aplicações interativas, que ampliam a visão dos objetos georreferenciados do mundo e potenciam o conhecimento. Este fenómeno tem contribuído para disponibilizar um vasto conjunto de dados não apenas para investigadores no domínio dos SIG, mas também apresentando potencial para contribuir de modo significativo para vários ramos da ciência.

O capítulo conclui-se apresentando algumas especificidades da integração de dados geoespaciais provenientes de IGV em IDE, nomeadamente que se começa a caminhar de uma perspetiva distribuída mais técnica para uma perspetiva mais social, através do desenvolvimento de comunidades virtuais que aperfeiçoam a colaboração e em que os geoportais se apresentam como um elemento que desempenha um papel preponderante nesta visão colaborativa de uma IDE. Um último aspeto que é possível concluir, prende-se com a demonstração de que uma visão integradora destes conceitos não só se torna viável, como constitui um imperativo no caminho do vasto domínio que é a abordagem ao conhecimento da distribuição da Biodiversidade, enquanto necessidade emergente de valor incalculável para a humanidade, que nas últimas décadas tem surgido como ameaçado.

4 SOFTWARE GEOESPACIAL LIVRE E ABERTO E DADOS GEOESPACIAIS ABERTOS

“If you want to accomplish something in the world, idealism is not enough, you need to choose a method that works to achieve the goal, you need to be pragmatic.” - Richard Stalman

4.1 Introdução

Existem diversos fatores que devem ser tidos em linha de conta quando optamos pela utilização de determinado pacote de *software*, mais especificamente, *software* aberto. Para além dos nossos próprios requisitos operacionais, destacam-se: a maturidade do *software*, o nível de envolvimento da comunidade, as funcionalidades específicas disponíveis e a facilidade de implementação. Alguns destes fatores são analisados neste capítulo para um conjunto de pacotes de *software* que constituem um sistema aplicacional de SIGWeb. Procura-se também descrever o conceito de dados geoespaciais abertos, dos quais se descrevem as principais tendências e a sua importância atual, relacionando-os com o SGLA. Irá procurar demonstrar-se o papel fulcral que esta inter-relação pode permitir desempenhar numa abordagem de SIG colaborativos para produção de IGV, como forma de potenciação de conhecimento. Neste capítulo efetua-se ainda uma breve descrição de diferentes tipos de SGLA com diferentes funcionalidades ao nível de servidores, BD, clientes e manipulação de IG.

4.2 Software Geoespacial Livre e Aberto

Software Livre, *software* de código aberto, *software* livre e aberto ou, apenas, *software* aberto, são expressões que designam um programa de computador cujo código-fonte é disponibilizado de modo a permitir o uso, a cópia, o estudo e sua redistribuição. O conceito de livre e aberto opõe-se ao conceito de *software* proprietário, mas não ao *software* que é vendido com intenção de lucro: *software* comercial. Ao distribuir o *software* livre, o detentor dos direitos escolhe uma das tipologias de licença de *software* livre que normalmente é anexada ao código fonte. Esta licença informa quais os direitos que o autor está a garantir e quais as condições aplicadas (OSI 2012).

A *Free Software Foundation* considera um *software* como livre quando este respeita os quatro tipos de liberdade para os utilizadores: 1ª Liberdade: executar o programa, para qualquer propósito; 2ª Liberdade: estudar como o programa funciona, e adaptá-lo às necessidades individuais de cada utilizador; 3ª Liberdade: redistribuir cópias do programa de modo a que cada utilizador possa ajudar utilizadores com necessidades semelhantes; 4ª Liberdade: modificar o programa e distribuir estas modificações, de modo a que todos os utilizadores interessados beneficiem dessas alterações. Para que as quatro liberdades sejam satisfeitas é necessário que o programa seja distribuído juntamente com o seu código fonte e que não sejam colocadas restrições para que os utilizadores alterem e redistribuam

esse código (FSF 2012). Já a *Open Source Initiative* devido à ambiguidade da palavra “free” em inglês, prefere a expressão *Open Source*, que em Português se traduz para as expressões indicadas anteriormente (OSI 2012). Por sua vez, o movimento FOSS4G, sigla que deriva da terminologia inglesa: *Free and Open Source Software for Geospatial*, ou em Português, *Software Geoespacial Livre e de Código Aberto*, consiste no evento global anual organizada pela *Open Source Geospatial Foundation* (OSGeo) desde a sua criação em 2006, com ênfase nesta tipologia de *software*. Começa a ganhar forma e integridade um novo nicho, especificamente aplicado a estas tecnologias (OSGeo 2012).

Numa definição holística de SIG apresenta-os como *uma disciplina especial de informação que mantém o registo não apenas de eventos, atividades e “coisas”, mas também o onde esses eventos, atividades e “coisas” acontecem ou existem* (Longley et al. 2005). Esta definição complementa-se com a abordagem integradora que nos é fornecida por Wade e Sommer (2006), segundo a qual os SIG são *uma coleção integrada de hardware, software computacional e dados, utilizados para ver e gerir informação sobre locais geográficos. Analisar relações espaciais e modelar processos espaciais. Um SIG fornece um quadro de trabalho para recolher e organizar dados espaciais e informação relacionada, para ser disponibilizada e analisada* (tradução livre do original). Assim, o SGLA é o ecossistema de *software* livre e de código aberto que é vocacionado para aplicações de SIG. Genericamente, podemos dizer que o ecossistema de *software* geoespacial aberto é mais pequeno que o fenómeno do *software* aberto em geral, tal como SIG é mais pequeno que as tecnologias da informação (TI). Nesta dissertação, quando se fala em SGLA refere-se aos SIG, sistemas de GPS¹¹, gestão de dados espaciais, ferramentas de desenvolvimento e aplicações destinadas a utilizadores, abrangidos por uma licença de *software* aberto.

4.3 A fundação OSGeo

Não é possível falar em SGLA sem se falar no importante papel desempenhado pela fundação OSGeo. Segundo o sítio *Web OSGeo* (2012) trata-se de uma organização sem fins lucrativos cuja missão é apoiar o desenvolvimento colaborativo de SGLA, e promover a sua utilização generalizada. Para tal proporciona uma infraestrutura partilhada para melhorar a colaboração entre projetos. A fundação fornece também apoio financeiro, organizacional e legal a este nicho de *software*. Funciona ainda como uma entidade jurídica independente, em que os membros da comunidade podem contribuir para o financiamento de código e outros recursos, tendo a garantia de que suas contribuições serão mantidos para benefício público.

Destacam-se além destes, outros objetivos de relevo da OSGeo: promover a disponibilização gratuita de dados geográficos; incentivar a implementação de padrões abertos e interoperabilidade baseada em padrões abertos, nos projetos da fundação;

¹¹ *Global Positioning Systems* ou, em Português, *Sistemas de Posicionamento Global*

garantir um alto grau de qualidade nos projetos da fundação, para construir e preservar a fundação enquanto marca; e, tornar a fundação e *software* relacionado com esta mais acessíveis aos utilizadores (p.e.: pacotes integrados de *software*, documentação cruzada entre pacotes de *software*, entre outros). Atualmente a OSGeo patrocina 22 projetos, entre *software* de SIG *desktop*, aplicações de mapas para a *Web*, bibliotecas geoespaciais e projetos de dados abertos. Entre estes, destacam-se: Mapserver, Geoserver, Mapfish, OpenLayers, GRASS GIS, Quantum GIS, GDAL, PostGIS, entre outros (OSGeo 2012).

4.4 Convergência entre SGLA e padrões abertos de informação

As especificações OGC (descritas no capítulo 3) são suportadas por vasto conjunto de *software* de código aberto e proprietário (quer de ambiente *desktop*, quer de ambiente *Web*). Torna-se assim evidente, que os padrões abertos gradualmente promovem o seu impacto na comunidade geoespacial. Destacam-se os padrões geoespaciais abertos como o formato ESRI Shapefile ou GeoRSS, que têm ganho muito enfoque nos últimos anos, tendo sido adotados por várias comunidades de TI (Kralidis 2008).

Os processos colaborativos e parcerias que envolvem a utilização de padrões abertos, fornecem às organizações a oportunidade da criação de interfaces abertos e mecanismos de comunicação de computação distribuída (Kralidis 2008). Na ausência dos padrões geoespaciais abertos, os mecanismos de interoperabilidade de *softwares* clientes são restringidos às interfaces e operações prescritas pela organização ou fornecedor do serviço. Daí resulta que sempre que há a modificação de alguma parametrização ou da lógica de algum processo ao nível de serviço, os clientes terão que se alinhar em consonância com essas mudanças de modo a garantirem o mesmo nível de informação e serviços.

Assim, o uso de padrões abertos diminui as barreiras à integração. Padrões bem conhecidos podem promover o desenvolvimento e uso de tecnologias e ferramentas comuns, que podem surgir como peças de construção para os técnicos que implementem determinado sistema (Kralidis 2008). É exemplo num hipotético sistema aplicacional, a implementação da biblioteca de *software* aberto: GeoExt, combina o *software* de sistema de mapas OpenLayers com a aplicação *Web* ExtJS (uma API), permitindo criar poderosas aplicações ao estilo SIG *desktop* na *Web* com recurso a JavaScript. Ao invés de criar algo de raiz, permite que se disponibilizem mais recursos nos domínios necessários para a organização. Atualmente existem variadas ferramentas e tecnologias de código aberto que implementam padrões abertos. A existência de padrões abertos beneficia ainda o desenvolvimento e manutenção de diversos pacotes de *softwares* das diferentes camadas da cadeia de valor (p.e.: servidores, clientes e BDs) (OSGeo 2012).

Por outro lado, a Internet tem vindo a possibilitar o intercâmbio entre redes de dados e informação pelo planeta, como nunca antes tinha sido possível. Alarga o horizonte de

comunicação, ao permitir que comunidades de informação, peritos em determinadas áreas, cientistas, investigadores e o público em geral, troquem informação e comuniquem de modo mais eficiente independentemente da sua localização física (Kralidis 2008). Os padrões abertos fornecem uma infraestrutura básica de interoperabilidade que permite que a informação seja integrada de modo consistente dentro dessas redes. Cabe a diferentes organizações como a W3C ou a OGC fornecer formatos padrões de base que diminuam a complexidade do desenvolvimento de *software* e permitir que o intercâmbio de dados e informação entre ferramentas dispare e independentes, se faça de um modo mais transparente que nunca. (W3C 2012). Adicionalmente às vantagens, em geral, do *software* aberto, este permite satisfazer os crescentes requisitos de interoperabilidade ao influenciar positivamente os padrões abertos e, deste modo, contribuir para que as tecnologias de IDE e a troca de dados espaciais, sejam ferramentas de informação efetivas e de suporte à tomada de decisão (Kralidis 2008).

4.5 Dados geoespaciais abertos

Uma parte integrante da OSGeo é o comité dos dados geoespaciais públicos. Este comité, dedicado exclusivamente à questão dos dados geoespaciais, opera com foco nos objetivos seguintes: promover o uso de padrões abertos de formatos de informação; promover o acesso público aos dados recolhidos por entidades públicas; gerir um repositório de dados abertos; e, difundir as licenças relativos aos dados geoespaciais abertos (OSGeo 2012). Segundo Gray (2010) apesar de *evidentes avanços tecnológicos na partilha de dados geoespaciais, muitos destes dados produzidos por cientistas raramente ou apenas uma pequena parte deles, fica disponível para os seus colegas. Esta impossibilidade de acesso livre a dados geoespaciais apresenta-se mais trágica do que a falta de acesso livre a outros tipos de dados científicos. Não só porque a humanidade enfrenta desafios críticos ambientais, mas também porque todos os dados geoespaciais referem-se à mesma Terra e, assim, cada nova camada de dados é rica em possibilidades para a exploração de relações com outras camadas de dados* (tradução livre do original).

Este mesmo autor considera haver ainda muitos obstáculos técnicos e institucionais para superar antes que a ciência se torne tão habilitada pela tecnologia da informação como outros domínios. Os obstáculos à interoperabilidade técnica têm vindo a ser superados pela OGC através da constituição de grupos de trabalho em díspares domínios (Gray 2010). Um possível contributo para o aumento da disponibilidade de dados geoespaciais, tem vindo a ser o crescimento exponencial dos dados geoespaciais abertos e existem tecnologias chave que têm vindo a permitir que isso aconteça: nomeadamente a *Web 2.0* e a massificação dos recetores de GPS (Zeiss 2008).

Numa análise ao panorama dos dados geoespaciais abertos, Gray (2010) apresenta 17 razões e motivações para a publicação de dados geoespaciais online, utilizando padrões

abertos da OGC. De entre todos os fatores referidos, destacam-se: a verificabilidade dos dados: a ciência exige verificabilidade da informação, isto é, qualquer pessoa competente deve ser capaz de examinar os dados de um pesquisador para ver se esses dados apoiam as conclusões do pesquisador; unificação de observações: ser capaz de caracterizar os parâmetros de sensores, sistemas de sensores e redes de sensores integrados de processamento (incluindo intervenções humanas), permite a unificação útil de muitos tipos de observações; partilha de dados e estudos interdisciplinares transversais: diversos conjuntos de dados com modelos de dados bem documentados podem ser partilhados entre diferentes comunidades¹², com esta partilha de dados a proporcionar melhores oportunidades de estudos interdisciplinares; maximização de valor: o valor dos dados aumenta com o aumento potencial do número de utilizadores o que beneficia a ciência e o conhecimento de um modo geral; o ritmo da ciência e do conhecimento: os dados abertos permitem um ritmo acelerado de descobertas científicas, com a automação e melhores arranjos institucionais é dado aos investigadores mais tempo para estudos de campo, trabalho e comunicação. Ainda neste ponto Gray (2010) refere que as *alterações na superfície terrestre ao nível de fenómenos biofísicos que costumavam levar 10.000 anos, podem agora levar três, e este é uma das razões porque precisamos de ciência em tempo real* (tradução livre do original).

4.6 Comunidades como fio-condutor de requisitos tecnológicos de SGLA

Segundo McIlhagga (2008) o tradicional modelo do *software* proprietário de SIG tem vindo a manifestar-se problemático sendo um dos fatores no cerne dessa questão os diferentes requisitos colocados na *Web* geoespacial por diferentes e variadas indústrias. O resultado foi que a tecnologia produzida apenas endereçou as necessidades básicas de um segmento de mercado reduzido. O sucesso do SGLA neste domínio particular advém da sua abordagem de desenvolvimento que *evita as armadilhas dos produtores de software proprietário, ao tornarem disponível o núcleo do software livremente para uma pessoa, grupo ou organização que o deseje e, mais importante, a uma comunidade que o ajuda a construir* (tradução livre do original) (McIlhagga 2008).

Verifica-se assim uma repentina mudança na relação produtor-utilizador no nicho de SGLA de mapas na *Web*. A relação deixa de ser passiva, por parte do utilizador que se limitava a consumir um produto de *software*, mas passa a ser mais ativa e dinâmica. Surge uma nova tendência emergente de produtores e consumidores simultaneamente: os *prosumidores*¹³, um termo cunhado por Toffler (1989) e que tem estado mais em evidência nos tempos atuais. A *Web* 2.0 tem vindo a potenciar novas tendências emergentes em que há uma fusão parcial ou total desses dois papéis e o *software* aberto e trabalho colaborativo utilizam

¹² A OGC define uma comunidade de informação como um grupo de pessoas ou utilizadores que têm como característica comum a partilha de um dicionário de dados geoespaciais, incluindo a definição de relações de objetos e um esquema de metadados comum.

¹³ Termo que designa simultaneamente produtores e consumidores.

estas tecnologias, tal como focado recentemente por Tapscott e Williams (2008).

Com a implementação de soluções de SGLA é criada uma perspetiva em que as taxas de licenciamento e restrições de utilização do *software* proprietário deixam de ser impedimentos ao desenvolvimento, nomeadamente de soluções de SIG colaborativos para produção de IGV. Ao invés, a implementação de soluções de SGLA, usufruindo da 3ª liberdade do *software* livre e, portanto, na prática sem custos para os utilizadores, permite que se alcance uma mais vasta gama e número de utilizadores, do que com soluções implementadas de acordo com os princípios do *software* proprietário. Verifica-se também que, deste modo, diferentes colaboradores (prosumidores) trabalham em conjunto para desenvolver soluções apropriadas que depois ficam disponíveis a outros produtores e utilizadores, continuando-se assim o ciclo de desenvolvimento (McIlhagga 2008).

O sucesso de um fornecedor de *software* aberto é medido não pelo número de licenças vendidas mas pelo facto de a solução ir de encontro às necessidades específicas dos utilizadores. Esta abordagem de desenvolvimento tecnológico assume particular relevância no caso de entidades que desenvolvem tecnologias de *software* aberto para uso em determinadas comunidades que recorrem a estes recursos. As comunidades envolvem-se com as aplicações de mapas na *Web* para a resolução de problemas dos seus domínios, de modo a apresentar a informação de modo diferente de outras tecnologias (McIlhagga 2008).

4.7 SGLA e Dados Geoespaciais Abertos como Valor para o Utilizador Final

As tecnologias de SIGWeb têm vindo a ser desenvolvidas e expandidas com base nas necessidades de aplicativos específicas e de comunidades de utilizadores, quer se trate de *software* proprietário quer se trate de *software* aberto (McIlhagga 2008). Pressupõe-se que este fenómeno tenha vindo a ocorrer, sobretudo, porque o núcleo das tecnologias de SIGWeb não são mais do que ferramentas que servem os requisitos e interesses das continuamente crescentes comunidades, que procuram incorporar conteúdos geoespaciais nos requisitos das suas aplicações de SIGWeb.

O efeito de rede, do valor do *software* aberto residir no favorecimento de soluções por contraponto à tecnologia em si, tem vindo a alterar a dinâmica do desenvolvimento de *software* de mapas *Web*. O valor passa a residir na ida do encontro desta tecnologia aos requisitos específicos dos utilizadores, esses sim de valor acrescido. Com a crescente disseminação deste tipo de *software* observada nos últimos anos, o resultado passa a ser o de existir uma tendência emergente de um novo nicho. Este permite que quer a atividade comercial quer a colaborativa respondam a determinados requisitos específicos das comunidades e utilizadores, através do avanço das tecnologias de mapas *Web*, dentro do ecossistema mais abrangente que é o do SGLA.

4.8 Componentes de um sistema aplicativo para um SIGWeb Colaborativo

4.8.1 Debian 6.0.5 “Squeeze”: um sistema operativo

Os SIG tendem a ser vistos como importantes ferramentas de apoio à resolução de problemas geográficos (Longley *et al.* 2005). Não sendo o objetivo deste capítulo ou desta dissertação realizar uma abordagem à definição de SIG, importa enquadrar a abordagem realizada ao nível do Sistema Operativo (SO) como parte integrante e fundamental de um SIG. Tal como foi referido anteriormente um SIG é todo o conjunto formado por *software* e *hardware* que permite operar com informação geoespacial. Apesar de diferentes revisões bibliográficas omitirem frequentemente esta componente, torna-se pertinente iniciar este sub-capítulo englobando os SO como parte fundamental do SIG. Tanto mais que SO de naturezas diferentes limitam a possibilidade de utilização de *software* SIG, em si, de naturezas diferentes para um dado sistema aplicativo. Um servidor com um SO baseado na arquitetura *amd64* (de 64 bits) implementado num sistema de *hardware* compatível com essa arquitetura, como é o caso da opção deste Projeto, permite a alocação de mais do que 4 *gigabytes* de memória do servidor quando comparado com uma arquitetura de SO de 32 bits (p.e.: arquitetura *i386*). Este aspeto é fulcral em termos da performance de uma moderna BDG que manipule uma vasta série de dados geoespaciais com vasto volume de informação. Além disso uma arquitetura de 64 bits traduz um ganho de performance em termos de execução de todos os componentes de *software* de um sistema aplicativo. No entanto, será necessário a disponibilidade de pacotes de *software* desenvolvidos e compilados para serem utilizados nessa arquitetura.

O termo Debian designa, simultaneamente, o nome de uma distribuição de um SO Linux¹⁴ livre e de código aberto e de um grupo organizado de voluntários de todo o mundo que mantêm essa distribuição. Segundo o seu sítio Web o Debian (2012) *é um SO livre para o seu computador, um SO é a série de programas e utilitários base que permitem fazer funcionar um computador. O Debian utiliza o kernel Linux (sendo o Kernel o núcleo do SO que faz a ponte de ligação entre o hardware e o software), mas a grande maioria das ferramentas têm origem no projeto GNU, daí o nome GNU/Linux* (tradução livre do original). Atualmente a versão estável para utilização em ambiente de produção é a versão 6.0.4 (nome de código *Squeeze*). O SO Debian é especialmente conhecido pelo seu sistema de gestão de pacotes, chamado APT, que permite atualizações de pacotes de aplicações de *software* de modo relativamente fácil a partir de um vasto repositório de pacotes de aplicações e a remoções automática dos pacotes obsoletos. Outra característica desta distribuição face a outros SO Linux, é o modo de disponibilização de novas versões, em que o Debian procura sempre manter os pacotes de *software* mais estáveis após um período de testes considerável, garantindo assim uma grande estabilidade e fiabilidade, o que o torna

¹⁴ Linux é o termo geralmente usado para designar um SO que utiliza a arquitetura de sistema assente no núcleo Linux. Foi desenvolvido pelo finlandês Linus Torvalds e o seu código fonte está disponível sob licença GPL.

num SO vantajoso para ambientes de produção (Debian 2012).

Em termos de *software*, para além do SO, na base de um sistema aplicacional de SIG distribuídos pela *Internet*, encontramos sempre um servidor HTTP ou servidor *Web*, isto é, o *software* cuja função é disponibilizar conteúdos que podem ser acedidos pela *Internet*. Segundo a *Apache Software Foundation* o projeto *Apache HTTP Server* é um projeto colaborativo de desenvolvimento de *software* que visa a criação de um código fonte de implementação de um servidor *Web HTTP* robusto, de nível comercial, rico em funcionalidades e livre (tradução livre do original) (Apache 2012). Desde o seu lançamento em 1996 é o *software* de servidores *Web* mais popular, estimando-se que seja utilizado por 57% de todos os servidores *Web* mundiais (Apache 2012). De entre as suas principais características, destacam-se: a potência e flexibilidade como servidor *Web* compatível com o protocolo HTTP/1.1; a flexibilidade de configuração e extensão com módulos de *software* de terceiros; a possibilidade de correr na maioria dos SO atuais; e, a implementação de funcionalidades relacionadas com BD relacionais e não relacionais.

4.8.2 Geoserver 2.1.3

De acordo com o sítio *Web* Geoserver (2012), este é um *software* livre e aberto servidor de mapas, escrito em linguagem Java, que permite aos utilizadores partilhar e editar informação geoespacial. A sua arquitetura foi desenvolvida dando particular atenção a questões de interoperabilidade, pelo que a publicação de dados a partir de fontes de dados geoespaciais é realizada utilizando normas abertas (definidas pela OGC). Como é apologia dos *softwares* livres e abertos, foi desenvolvido e é mantido por uma vasta comunidade que envolve pessoas e organizações de todo o mundo, que colaboram ativamente na sua programação, teste e suporte. Ainda segundo a Geoserver (2012) o *software* consiste na implementação de referência das normas *WFS* e *Web Coverage Service* (WCS) do OGC e é, simultaneamente, um servidor certificado e de alta performance do serviço *WMS* (tradução livre do original).

Em termos de características principais deste *software* destacam-se: a implementação das especificações do OGC, nomeadamente: *WMS* 1.1.1 e 1.3.0; *WFS* 1.0.0 e 1.1.0; *WFS-T* 1.0.0; e, *WCS* 1.0.0 e 1.1.1; o suporte a vários formatos de *back-end*, nomeadamente: *ArcSDE*, *Oracle Spatial*, *SQL Server*, *ESRI shapefile*, *GeoTIFF*, *MrSID* e *JPEG2000*; os vários formatos de output de informação, nomeadamente: *GML*, *shapefile*, *KML*, *GeoJSON*, *PNG*, *JPEG*, *TIFF*, *SVG*, *PDF* e *GeoRSS*; e, a interface de administração gráfica, via *Web*. O Geoserver tem um largo suporte como servidor geoespacial a uma variedade de *softwares* clientes e servidores, nos quais se incluem: as DB de SGLA: *PostGIS* e *MySQL*; clientes de SGLA: *OpenLayers*, *uDig*, *Quantum GIS* e *gvSIG*; BD proprietárias: *ArcSDE*, *Oracle Locator*, *SQL Server 2008*; e, *softwares* clientes proprietários: *ArcGIS*, *Google Earth* e *Virtual Earth*.

4.8.3 Base de Dados Geográfica: PostgreSQL/PostGIS

De acordo com o seu sítio *Web*, o PostgreSQL (2012) é um poderoso *software* gestor de bases de dados (SGBD) orientadas ao objeto. Trata-se de SGLA com mais de 15 anos de desenvolvimento e uma arquitetura que tem ganho ao longo do tempo forte reputação em termos de fiabilidade, integridade e precisão (PostgreSQL 2012). É altamente amplificável quer ao nível da quantidade de dados que pode gerir, quer ao nível do número de utilizadores simultâneos que pode acomodar e com que pode interagir. Existem sistemas PostgreSQL ativos em ambientes de produção que gerem acima de 4 *terabytes* de dados, o que é de realçar ao nível dos sistemas de informação (PostgreSQL 2012).

Em termos de funcionalidades adicionais e compatibilidade com padrões de BD, a sua implementação SQL (isto é, em Inglês, *Structured Query Language*) é totalmente consonante com o padrão ANSI-SQL:2008. O GiST (*Generalized Search Tree*) constitui um sistema de indexação avançado deste SGBD. Apresenta algoritmos de ordenação e procura, bem como uma interface que permitem tanto a criação de tipos de dados personalizados como métodos de consulta. É o GiST que permite ainda que diferentes projetos utilizem o PostgreSQL: tal como o PostGIS (PostgreSQL 2012). Outra característica de realce é o facto de ser suportado por um vasto e detalhado sistema de documentação. Este é mantido não só pelo seu grupo de desenvolvimento global, como também integra variados contributos da sua vasta comunidade de utilizadores, o que torna o suporte a este SGBD extremamente completo. O PostgreSQL reclama a si ser o SGBD aberto mais avançado. Apresenta a performance e funcionalidades que lhe permitem não só competir com outros *softwares* proprietários bastante populares, como também chega a apresentar funcionalidades únicas que os permitem transcender (Obe e Hsu 2011). Existem disponíveis quatro versões estáveis correspondentes aos ramos ativos do PostgreSQL: 8.3.18, 8.4.11, 9.0.7 e 9.1.3. Da análise efetuada sobre o *software* registou-se que as diferentes versões não trariam melhorias significativas de umas relativamente às outras para efeitos do sistema aplicacional a desenhar e das funcionalidades pretendidas. A escolha da versão do PostgreSQL a utilizar no Projeto atendeu assim a dois aspetos fundamentais: a estabilidade da versão; e, a existência e disponibilidade de documentação de referência de suporte, para além da documentação oficial.

Uma BDG é um tipo particular de BD que define tipos de dados para objetos geométricos e permite que se armazenem dados geográficos em tabelas normais da BD. Para além disso, dispõe de funcionalidades especiais e índices relativos à interrogação e manipulação desses dados, frequentemente utilizando a linguagem *Structured Query Language* (SQL). Uma BDG fornece-nos simultaneamente uma ferramenta de armazenamento e uma ferramenta de análise (Obe e Hsu 2011). O PostGIS é um pacote de SGLA liderado pela organização *Refraction Research*, sob a forma de extensão para o SGBD PostgreSQL.

Acrescenta o suporte a objetos geográficos ao PostgreSQL, permitindo que este SGBD seja utilizado como o suporte de BDG para SIG (à semelhança da ESRI SDE). A versão mais atual do PostGIS providencia mais de 700 funções espaciais, operadores, tipos de dados e capacidade de indexação dos dados. Este *software* segue as especificações da OGC: *Simple Features Specification for SQL*. (Obe e Hsu 2011, PostGIS 2012). Fornece assim ao PostgreSQL: representações espaciais de diferentes tipos de geometrias; suporte operações de análise espacial e operações de criação de geometrias e conversão da dados; apoio geodésico para medições na superfície terrestre; e, linha de comandos e ferramentas gráficas para uma gestão flexível das BDG. Em muitas situações, trata-se de um *software* utilizado como um espaço de *backend* de BDG quer para *softwares* clientes, quer para *softwares* servidores, de entre os quais se destacam: GeoServer, Mapserver, Mapnik, GRASS, QGIS, uDig, gvSIG ArcServer, ArcGIS, MapInfo Professional, entre outros. Fornece-nos uma panóplia de tipos de geometrias para implementar na modelação do mundo real. Genericamente, os tipos geométricos implementados pelo PostGIS são: pontos, linhas, polígonos, multi-pontos, multi-linhas, multi-polígonos e coleções geométricas.

Segundo Shamal (2012), num estudo comparativo entre as BDG Oracle Spatial e PostgreSQL/PostGIS, estas BDG constituem as implementações mais maduras destes sistemas geoespaciais e são conhecidos pela crescente relevância que este tipo de BD tem vindo a ganhar. A investigação realizada sugere que o PostGIS é mais vantajoso em relação ao Oracle Spatial pelas razões seguintes: poupança de custos (dado tratar-se de *software* aberto); facilidade de integração, dado que uma vasta gama de produtos suportam o PostGIS (alguns dos quais enunciados anteriormente); facilidade de utilização; performance superior, uma vez que o PostGIS imprime uma implementação leve utilizando uma quantidade mínima de bytes para produzir geometrias e índices, o que reduz o número de operações de escrita no disco com a consequente melhoria de desempenho; variedade de funcionalidades, muitas das quais não estão disponíveis na Oracle Spatial.

Noutra perspetiva pode-se questionar a vantagem do PostGIS relativamente ao formato *shapefile* (e outros) que durante vários anos terão sido a forma padrão de armazenar e interagir com dados espaciais. No entanto, estes formatos apresentam diferentes desvantagens: o formato em arquivos requer *software* especial para ler e escrever, ao passo que a SQL é uma abstração para acesso a dados e análise aleatória e em essa abstração seria necessário escrever muito mais código para o acesso e análise de dados geoespaciais o que se repercute na velocidade de processamento da informação; e, vários utilizadores em simultâneo, pode causar corrupção de dados o que dificilmente acontecerá com um BD geoespacial (um fator vital num projeto colaborativo que envolva vários utilizadores a aceder à mesma BDG simultaneamente) (OpenGeo 2012).

4.8.4 Um Cliente: Openlayers 2.12

O OpenLayers é uma biblioteca de código aberto desenvolvida em Javascript, que permite construir mapas *Web* “ricos” e interativos, visíveis na generalidade dos *browsers*. Uma vez que é uma biblioteca do lado do cliente e sem dependências do lado do servidor, não é necessário descarregar qualquer *software* ou configurações especiais do lado do servidor, pois implementa uma API¹⁵ Javascript. Pode-se assim utilizar sem a necessidade de descarregar qualquer aplicação para o PC (Hazzard 2011, OpenLayers 2012).

Em termos de relação e semelhanças com o Google Maps, o Yahoo! Maps, o Bing Maps e outras API de aplicações de mapas na *Web*, o OpenLayers também permitem que se ligue com o *backend* de servidor de mapas. Estas API, normalmente, também providenciam uma interface do lado do cliente. No entanto, existem algumas diferenças substanciais: não se tem qualquer controlo ou opção sobre o *backend*; não se pode personalizar o *backend* do servidor de mapas, e este pode mudar a qualquer momento, sofrendo alterações significativas; existem diversas restrições de uso em termos de licenciamento ou, em diferentes situações, alguns custos envolvidos; e, essas API não fornecem a flexibilidade de personalização que o OpenLayers fornece. O OpenLayers permite utilizar múltiplas camadas de diferentes *backends* de diferentes servidores para produzir um dado mapa. Normalmente, para aceder a um servidor de mapas na *Web*, cria-se uma camada de objeto e adiciona-se ao mapa com o OpenLayers. A título de exemplo, se se quisesse ter um Yahoo! Maps e um serviço WMS exibido em determinado mapa, utilizaríamos o OpenLayers para criar uma camada de objeto com o Yahoo! Maps e uma camada objeto WMS. De seguida é possível adicioná-las ao mapa pretendido com o OpenLayers (Hazzard 2011).

4.9 Conclusões do capítulo

Este capítulo começa por introduzir a noção de *software* geoespacial aberto caracterizando as suas quatro liberdades. Neste domínio, dá-se enfoque às duas liberdades que são premissas chave para o desenvolvimento do presente trabalho de projeto: a liberdade de estudar como o *software* funciona e adaptá-lo a necessidades particulares e específicas; e, a liberdade de redistribuir cópias gratuitamente. Num projeto de produção de IGV as vantagens deste tipo de *software* apresentam-se como um dado fulcral. Os custos comparativos de distribuição de *software* proprietário a toda uma comunidade, no contexto de um projeto desta natureza, poderiam por si só ser limitantes ao seu desenvolvimento. Acresce que a própria filosofia do *software* aberto, neste caso *software* geoespacial aberto, apresenta um largo grau de compatibilidade com a filosofia de um projeto de SIG colaborativo para a produção de IGV, o que, por sua vez, se reflete numa maior facilidade de integração de recursos humanos adicionais e ferramentas de apoio, bem como em

¹⁵ API ou Application Programming Interface (Interface de Programação de Aplicativos) é um conjunto de rotinas e padrões estabelecidos por um dado software para a utilização das suas funcionalidades por aplicativos que não pretendem envolver-se em detalhes da implementação desse software, mas somente usar seus serviços.

repercussões ao nível da interação homem-máquina analisadas do ponto de vista da comunidade enquanto produtora de IGV.

Demonstra-se que existe uma convergência entre os padrões abertos de interoperabilidade e o SGLA, tal como se poderia supor à partida. É tendencialmente mais fácil a adaptação de padrões abertos em SGLA do que em *software* proprietário e este aspeto relaciona-se com a 2ª e 4ª liberdades do *software* aberto. Por sua vez, no domínio do SGLA para SIGWeb, a adoção de padrões abertos de interoperabilidade diminui as barreiras de intercâmbio e integração de nova informação, passando a atuar o *software* como fator potenciador da agregação de dados e, em última, análise do conhecimento. Um aspeto de importância no âmbito do Projeto e de relevância enquanto vantagem de utilização deste tipo de *software*.

Realiza-se uma abordagem aos dados geoespaciais abertos para demonstrar não só a sua importância no domínio científico, porque uma disponibilização de dados a um maior espectro de utilizadores permite uma melhor análise e diferentes abordagens à informação nomeadamente ao nível da sua verificabilidade, mas também porque no domínio das ciências da terra e da vida e, em particular, da Biodiversidade, a humanidade enfrenta desafios críticos. A disponibilização de dados abertos amplifica o seu conhecimento através das relações cruzadas com outros dados externos, de diferentes camadas de informação.

Ao longo do capítulo realiza-se uma abordagem, não exaustiva, a diversos pacotes de *software* geoespacial aberto, ilustrando-se os seus pontos fortes, que permitem a constituição de um sistema aplicacional SIGWeb. De entre eles destacam-se os pacotes chave do sistema: PostgreSQL com suporte PostGIS, Geoserver e OpenLayers, suportados por um SO Debian. Explana-se a importância e algumas das vantagens da utilização do SO, que muitas vezes surge com um papel subestimado. Uma arquitetura amd64 permite uma utilização mais completa da memória de uma máquina e consequentemente uma maior eficiência. A segurança e fiabilidade de uma arquitetura Linux constituem também uma vantagem, associada ao seu baixo custo de implementação.

O *software* central do sistema ilustrado é, talvez, o PostGIS. Este é na sua essência uma BDG moderna e muito robusta, de alto desempenho. Traduz várias vantagens relativamente a outros *software* de SIG tradicionais e ao formato de trabalho com ficheiros de dados (p.e., o shapefile), nomeadamente a possibilidade de vários utilizadores em simultâneo poderem utilizar a mesma BDG para produzir informação, sem causar corrupção de dados. O Geoserver, enquanto servidor de mapas, prima pelo potencial de interoperabilidade e ligação a vários sistemas. Destaca também a implementação da especificação OGC: WFS-T, que num projeto de SIG colaborativo, em última análise, pode constituir mais um método de aquisição de informação por parte de utilizadores, de modo simples (através do *browser*).

5 A ESTRUTURA E O MODELO DE DADOS DO BIODOM

“All models are, by definition, wrong. But some are useful.” - George Box

5.1 Introdução

Os modelos de dados convencionais em SIG, enfatizam representações estáticas de uma dada realidade. A IG para uma determinada área é normalmente decomposta num conjunto de camadas temáticas matriciais ou vetoriais, como se se tratasse de um mosaico (Mark e Frank 1996). Estas camadas limitam as funcionalidades dos SIG para representar informações dinâmicas, como transições ou movimento. Apesar das células matriciais codificarem valores de diferentes atributos num dado local, não têm em conta as características geoespaciais dinâmicas do tema que representam. Por outro lado, os objetos vetoriais forçam uma segmentação das entidades em camadas separadas, sempre que interagem em função do tempo ou espaço. É a adoção deste método de representação de entidades e atributos que compromete grande parte da modelação do domínio ambiental (Raper e Livingstone 1995). Os SIG precisam de uma rede complexa e rigorosa para a modelação de dados geográficos, de modo a superarem a dificuldade em lidar com complexidade geográfica, diferenças de escala, generalização e precisão (Goodchild 1992).

5.2 A modelação de dados geográficos

Segundo Longley *et al.* (2005) uma análise geográfica será apenas tão boa quanto a qualidade da BDG em que assenta e, por sua vez, essa BDG será tão boa quanto o modelo de dados do qual deriva. Para estes autores, a modelação de dados geográficos inicia-se com: uma definição clara dos objetivos e respetivo faseamento das necessidades específicas de determinado projeto; uma definição dos objetos e as suas relações; a formulação de um modelo lógico; e, posteriormente, a criação de um modelo físico. Na modelação de dados um aspeto importante é a compreensão do propósito do exercício de modelação em questão, o que pode ser compreendido pela recolha dos requisitos de determinado sistema e grupo de utilizadores. A formulação de um modelo lógico requer que se identifiquem os objetos e as relações entres estes, não sendo essencial que todos os objetos e relações fiquem determinados numa primeira abordagem, porque os modelos lógicos podem ser refinados ao longo do tempo. Quando um modelo lógico, independente da implementação, é desenhado, pode ser virtualmente transformado num modelo físico (dependente do sistema). Um modelo físico irá resultar num esquema de BD vazio, isto é, uma coleção de tabelas e de relações entre elas. Por vezes, por questões de otimização de performance ou alterações de requisitos, é necessário alterar o modelo físico.

Longley *et al.* (2005) enfatizam que é importante ter noção de que não existirá um modelo de dados geográficos correto. Cada problema geográfico pode ser representado de várias formas e vários modelos de dados podem representar o mesmo problema geográfico. Cada

modelo de dados pode ser desenhado com um fim específico e poderá ser adequado para outros fins. A modelação de dados geográficos é, simultaneamente, uma arte e uma ciência. Requer o conhecimento científico das características geográficas chave de sistemas do mundo real, incluindo o estado e o comportamento dos objetos e as relações entre eles. Os modelos de dados geográficos são de importância crítica porque têm uma importância particular ao nível do controle do tipo de dados que pode ser representado e das operações que podem ser realizadas. Modelos orientados ao objeto são o melhor tipo de modelos de dados para representar os *objetos do tipo rico* e as relações num sistema de infraestruturas, ao passo que modelos de representação simples de objetos serão suficientes para aplicações mais elementares (Longley *et al.* 2005).

5.3 Modelação de Dados de IGV em SIG Colaborativos

A modelação e desenho de um SIG colaborativo permite uma maior reutilização do conhecimento processual. Uma abordagem possível será associar à ontologia do sistema a utilização do desenho orientado ao objeto de modo a que fique especificada a hierarquia de informação genérica do mundo real. O primeiro passo para construir uma especificação ontológica é definir o vocabulário comum do domínio (Balram e Dragičević 2006a). Neste particular é importante o compromisso ontológico de modo a que o vocabulário seja utilizado de uma forma consistente para que o intercâmbio de informação seja possível. Normalmente, as partes envolvidas podem ter variadas bases de conhecimento, apesar de poderem partilhar uma ontologia comum. Como consequência o compromisso ontológico nem sempre garante que se complete totalmente as especificações do problema.

No contexto dos processos de SIG colaborativos as ontologias podem ser consideradas como uma descrição dos conceitos e relações que existem para a sociedade de objetos (Balram e Dragičević 2006a). Assim, os mapas e os aspetos de visualização da colaboração geoespacial do grupo podem permitir consolidar várias perspetivas da ontologia, ao passo que a *Geographic Markup Language* (GML) permite documentar essa mesma ontologia. A especificação ontológica criada deste modo, poderá então ser melhorada de modo iterativo para melhor modelar o domínio. Neste contexto, a escala é um tema fundamental em qualquer análise de IG. Segundo Sheppard e McMaster (2004) apesar de existirem inúmeras definições, as categorizações físicas e sociais de escala fornecem-nos uma determinada agregação de conceitos no que respeita à distinção dos elementos fulcrais das diferentes definições. *A escala física lida com o grau de detalhe e o alcance das medidas tomadas, ao passo que a escala social lida com a construção de espaços e locais sobre os quais os sistemas sociais operam* (tradução livre do original).

5.4 IGV vaga e incorporação de incerteza num modelo de dados de IGV

Em SIG é do conhecimento comum que se pode atribuir, a título de exemplo, uma dada localização a uma casa e saber se um jardim está ou não contido dentro dessa casa,

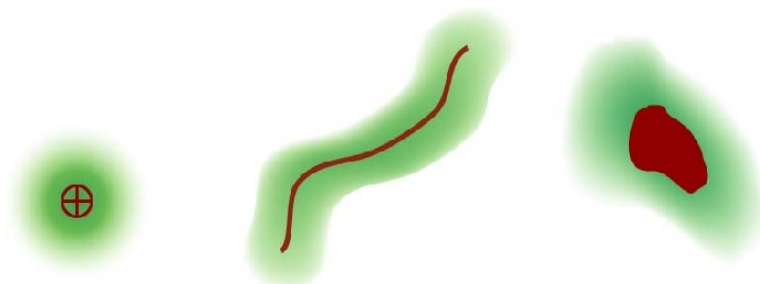
porque esta informação se baseia em geometrias e localizações precisas. No entanto, nem todos os objetos geoespaciais têm uma localização precisa. De acordo com Hart e Dolbear (2006) pode haver objetos que não correspondem a esse padrão. Se considerarmos um lago, este pode não ter fronteiras muito bem definidas apesar de podermos saber, com certeza, que determinado objeto está contido dentro desse lago: p.e., uma ilha. No entanto as fronteiras não são bem definidas o que se pode traduzir numa imprecisão. Esta imprecisão em registar dados geoespaciais, também se pode traduzir em incerteza.

Para dar resposta a esta questão De Longueville *et al.* (2009) desenvolveram uma IDE em que na camada de apresentação se mostrava uma variável que continha o testemunho do próprio utilizador em termos da vaguidade da informação. Integrava sob a forma de metadados as perceções dos utilizadores no que concerne aos aspetos relacionados com a vaguidade da informação de *input*, a partir dos seus testemunhos e atribuindo-lhe uma valoração. Foram incluídos dois tipos de metadados de vaguidade de informação. O primeiro tipo, refere-se à vaguidade da informação aferida pelo próprio utilizador, de modo qualitativo, que serve de auto-avaliação da IGV. Ao introduzir determinados dados o utilizador escolhe de entre as opções de uma lista de menu, que grau de precisão considera adequado ao seu testemunho. Foi proposta a seguinte lista de valores: 5 = “está exatamente ali”; 4 = “está ali”; 3 = “está mais ou menos ali”; 2 = “deve estar algures ali”; 1 = “não tenho a certeza se está algures por ali”; 0 = “não sei onde está”. O segundo tipo era gerado pelo próprio sistema, constituindo uma medida mais objetiva da vaguidade de IGV. Consistia num mecanismo em que o cliente da IDE guarda automaticamente a escala em que os elementos de IGV eram esboçados (De Longueville *et al.* 2009).

Segundo Zhang e Goodchild (2002) a escala à qual se codificam elementos geográficos constitui um bom indicador de precisão geográfica. Por ser expectável que um utilizador que tenha um bom conhecimento da localização do elemento geográfico em questão, ao desenhar esse elemento num SIG ele deverá utilizar uma escala (utilizando as funcionalidades de zoom) em que possa visualizar lagos, rios, estradas ou outros elementos, que o apoiem na sua localização no mapa. Por oposição, um utilizador com vagos conhecimentos da localização de um dado elemento geográfico terá a tendência, na maioria das vezes, de não utilizar o *zoom* até um grau de detalhe tão preciso. Os metadados introduzidos pelo utilizador e gerados pelo sistema aplicativo são muito complementares e dão-nos uma boa estimativa da vaguidade que caracteriza todas as partes de um todo de IGV (De Longueville *et al.* 2009).

No modelo da “gema de ovo” de Cohn e Gotts (1996) cada objeto geoespacial com um dado grau de vaguidade é representado por dois elementos: a “gema” é um objeto geográfico que representa a parte mais correta do objeto vago, ao passo que a “clara” é a fronteira alargada que define os limites do alcance da vaguidade desse objeto. O “grau de verdade” decresce

de 1 na fronteira entre a “gema” e a “clara” até 0, que corresponde ao espaço fora da “clara” (Dilo *et al.* 2007). Na Figura 4 é ilustrado um exemplo que representa este modelo.



(Fonte: De Longueville *et al.* 2009)

Figura 4: O modelo “gema de ovo” na representação de ponto, linha e polígono, vagos.

Na IDE desenvolvida por De Longueville *et al.* (2009) completa-se a informação relativa à incerteza espacial com a utilização de um conjunto matricial para representar essa mesma incerteza (Bastin *et al.* 1999 citados por De Longueville *et al.* 2009). Deste modo a informação vetorial de base com os metadados, terá que ser processada, de modo a que se possa calcular o valor do pixel dessa série de dados matriciais. Este valor do pixel expressa um “grau de verdade” para cada elemento vetorial, isto é, indica o quanto um objeto codificado por um dado utilizador pode ser extrapolado para esse pixel. Pese embora estas inferências decorrentes do seu trabalho, De Longueville *et al.* (2009) descrevem que fica por conceptualizar uma função matemática que descreva a distribuição espacial da vaguidade. Estes autores chegam mesmo a avançar com três parâmetros para essa função: a distância ao elemento geográfico codificado pelo utilizador (que será inversamente proporcional ao “grau de verdade” à medida que este decresce com a distância); a escala à qual o elemento geográfico é codificado (proporcional ao “grau de verdade”); e, a avaliação da vaguidade efetuada pelo próprio utilizador (em que um valor mais alto significaria maior precisão). Essa eventual função seria, deste modo, inversamente proporcional ao “grau de verdade”, dado que informação mais precisa terá um menor “grau de verdade” fora do elemento vetorial.

5.5 Modelo de dados BiodOM

5.5.1 Premissas de base e requisitos

Na base do desenho do modelo de dados do BiodOM, estiveram diferentes sub-premissas específicas e requisitos do sistema aplicacional, relativamente aos seus dados: 1 – a definição de esquemas e tabelas diferentes, dentro da BD, que deve ser relacionada com a hierarquização dos utilizadores para a manipulação de dados, bem como da necessidade de validação de toda a IGV produzida, pelos coordenadores dos grupos taxonómicos; 2 – a definição de diferentes tabelas de entrada de dados em função das fontes de entrada dados de IGV a conceptualizar; 3 - dado tratar-se de um sistema cuja IG é produzida, alterada e modificada pelos utilizadores, será pertinente a implementação de um sistema que permita

manter um registo das operações de *Create, Read, Update e Delete* (CRUD), isto é: Criar, Ler, Atualizar e Apagar, realizadas por estes na BDG; 4 - a necessidade de interligação com as BD do portal Naturdata e garantia de que na migração futura desses dados para padrões de dados Biodiversidade, por parte do portal Naturdata, estes mantenham a consistência; 5 - os dados deverão ser estruturados de modo a permitir as operações de pós-processamento conceptualizadas; 6 - o modelo de dados deverá estar ajustado à arquitetura do sistema aplicacional (aspeto que se verá em pormenor no próximo capítulo); 7 - o modelo de dados deverá dar resposta não só às questões de armazenamento de dados de IGV, como também de visualização desses dados, nomeadamente na ficha de espécie e no mapa global BiodOM; e, 8 - deverão ser integradas variáveis como: definição de unidades de área, escala de aquisição, “grau de verdade” (como modo de input por parte dos utilizadores) e incorporação de áreas de influência, parâmetros essenciais em IGV.

Na BDG do BiodOM não constam as questões de interoperabilidade e padrões de dados de Biodiversidade (aspetos abordados no 2º capítulo). Esta questão está relacionada pela previsível futura adoção e migração das tabelas de informação relativas às espécies, por parte do portal Naturdata, para tabelas que sejam em conformidade com as normas de padrões de dados Biodiversidade. Porém a BDG utiliza um esquema que permite a importação da informação das BD do portal Naturdata (presentemente no SGBD MySQL), e que permite relacionar a informação taxonómica das espécies com a informação geoespacial providenciada pelo BiodOM.

5.5.2 Fontes de entrada de dados de IGV

Como se verá em detalhe no capítulo 6.6, o sistema aplicacional na sua componente de IGV, apresenta como fontes de entrada de dados por parte dos seus utilizadores quatro módulos: **browser editor**: edição direta no *browser* com a implementação do protocolo WFS-T, para marcação de pontos e registo de ocorrências de espécies; **exif georeader**: carregamento de fotografias geoetiquetadas e importação da informação georreferenciada a partir dos campos de coordenadas do “exif” da foto; **GPX georeader**: importação de *tracks* e nuvens de pontos de recetores GPS; e, **QGIS editor**: ligação direta do software QGIS ao esquema “qgis_igv” da BD BiodOM e carregamento de IG relativa à distribuição de espécies padronizada, em função dos campos da tabela. A cada fonte de entrada de dados pelos utilizadores, correspondem tabelas específicas na BDG, por uma questão de lógica, organização dos dados e consistência da informação. Pretende-se manter uma certa integridade, dado que os *inputs* para estas quatro tipologias são de fontes diferentes. Naturalmente, tratando-se de uma BDG relacional, é possível efetuar análises e interrogar a BDG, relacionando a informação das diferentes tabelas.

5.5.3 Desenho do modelo de dados

A figura seguinte ilustra o modelo de dados desenvolvido e implementado para o BiodOM:

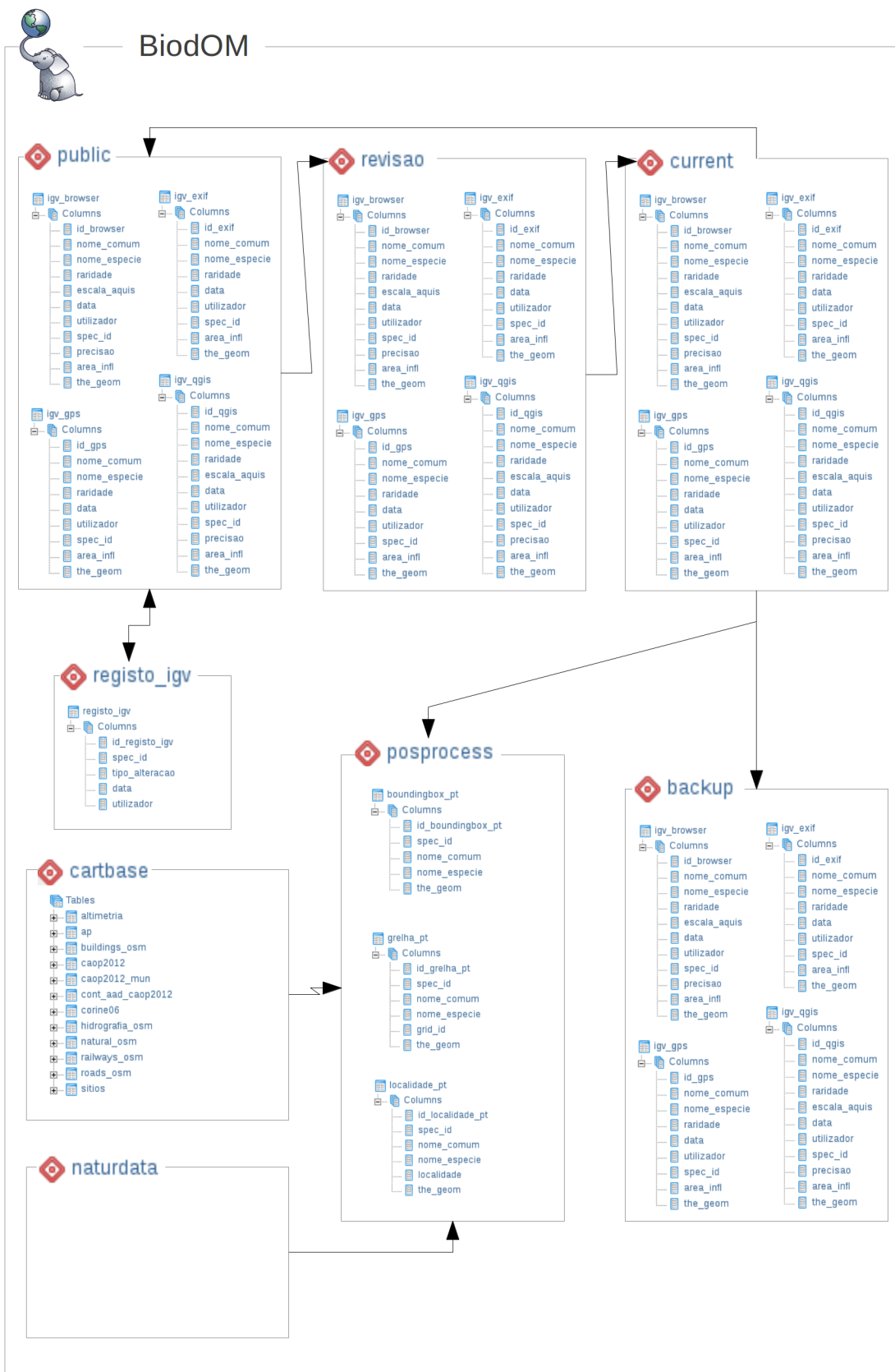


Figura 5: Modelo de dados implementado no BiodOM.

Uma BDG PostgreSQL/PostGIS contém, normalmente, um ou mais esquemas com nome próprio que, por sua vez, contém tabelas. Cada esquema contém ainda outros objetos, nos quais se incluem: tipos de dados, funções e operadores. Um objeto com o mesmo nome pode ser utilizado em esquemas diferentes sem haver conflito na BDG (PostgreSQL 2012, PostGIS 2012). Na BDG “biodomdb” o esquema “public” e o esquema “current”, entre outros, apresentam tabelas e funções com os mesmos nomes. Um utilizador pode aceder a objetos que estejam em diferentes esquemas da BDG à qual ele está ligado (caso tenha privilégios de acesso). A utilização de esquemas apresenta assim diferentes vantagens: permitem que vários utilizadores utilizem uma mesma BDG sem interferir entre eles; permitem a organização dos objetos das BDG em grupos lógicos, de modo a torná-los mais facilmente manipuláveis; e, as aplicações de terceiros podem ser compartimentadas num esquema, de modo a os nomes não colidirem com outros objetos (PostgreSQL 2012).

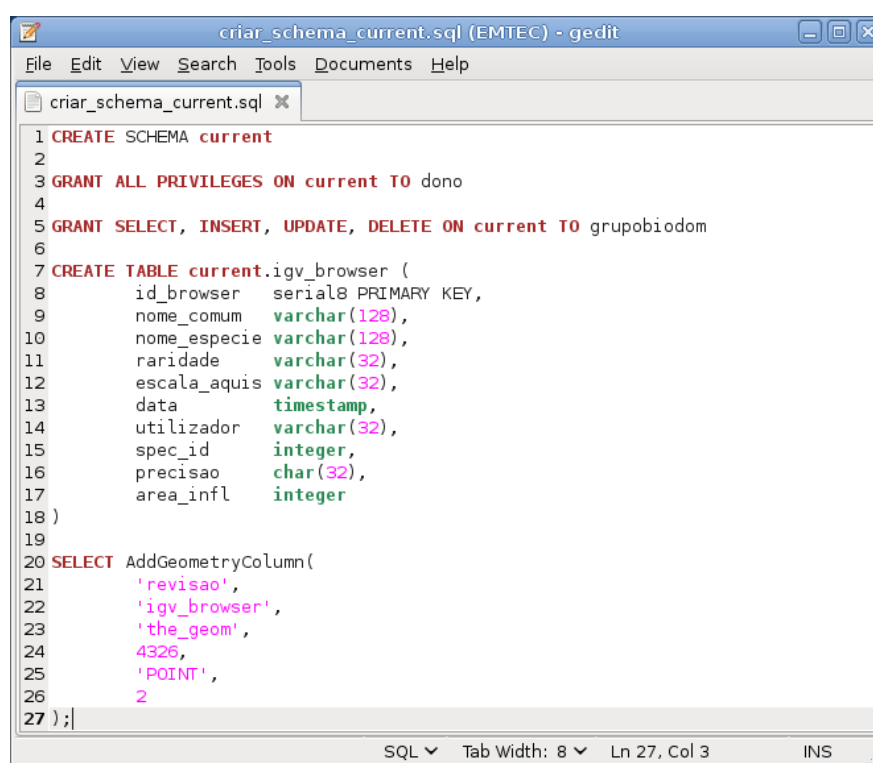
A BDG central organizada em diferentes esquemas de informação é, assim, um componente chave do BiodOM. Tal como o nome indica, este conjunto, é o espaço lógico onde reside a globalidade dos dados de IGV. Cada esquema da BDG contém diferentes tabelas em função da tipologia de informação e atendendo às diferentes fontes de *input* de dados de IGV: exif georeader, QGIS editor, *browser* editor e GPX georeader. Tal como se pode observar na Figura 5, o variado conjunto de esquemas da “biodomdb” contém um conjunto variado de tabelas. Importa referir também que as tabelas, para além dos dados das espécies e geometrias relacionadas com a sua distribuição, registam também o seu historial das operações de CRUD no esquema: “registro_igv”. Paralelamente, integram na fase de recolha, de parâmetros relacionados com a escala de introdução (nos módulos QGIS editor e *browser* editor) de informação (diretamente relacionada com a precisão de recolha de informação por parte de um colaborador de IGV), bem como pela integração num campo de metadados de uma avaliação qualitativa da própria informação que o colaborador está a integrar (no módulo *browser* editor). Posteriormente essa informação será cruzada com os próprios dados pelo sistema aplicacional, de modo a produzir a IGV com integração de incerteza em que, num dado matricial, cada pixel expressa um determinado “grau de verdade” relacionado com determinado objeto referenciado.

Existem quatro versões das tabelas principais, em três diferentes esquemas, para cada conjunto de dados primários: o esquema com as tabelas mais recentes (“public”), o esquema com as tabelas mais atualizadas que providenciam a visualização da IG depois de validada a sua integração (“current”); o esquema de revisão (“revisao”) (para efeitos de revisão por parte dos coordenadores de cada grupo taxonómico); e, o esquema “backup” onde se irão criar as cópias de segurança de modo automatizado, de toda a informação.

O esquema com as tabelas mais recentes: “public”, corresponde à versão mais recente utilizada para a introdução dos dados pelos diferentes utilizadores e colaboradores. Estes

dados são depois validados pelos coordenadores que trabalham essa informação sobre o esquema de revisão. Do esquema de revisão os dados passam para o esquema: “current”, que passa a constituir a informação validada e consolidada. A validação da informação introduzida pelos colaboradores e utilizadores, por parte dos coordenadores, é realizada com recurso às ferramentas Phpmyadmin e Quantum GIS. Sendo editada diretamente na BDG e após a sua validação, é importada para o esquema “public”, atualizando a informação. Paralelamente o esquema “posprocess” contém as tabelas relativas à informação resultante do pós-processamento dos dados, tal como descrita no ponto 6.7.

Estes esquemas, em função da sua estruturação, permitem que diferentes aplicações utilizem diferentes esquemas, em termos da realização de operações de consulta, gestão e de realização de análise espacial. Na figura seguinte, pode-se observar um exemplo detalhado do código SQL envolvido na criação do esquema “current” e respetivas tabelas:



```
1 CREATE SCHEMA current
2
3 GRANT ALL PRIVILEGES ON current TO dono
4
5 GRANT SELECT, INSERT, UPDATE, DELETE ON current TO grupobiodom
6
7 CREATE TABLE current.igv_browser (
8     id_browser      serial8 PRIMARY KEY,
9     nome_comum      varchar(128),
10    nome_especie     varchar(128),
11    raridade         varchar(32),
12    escala_aquis     varchar(32),
13    data             timestamp,
14    utilizador       varchar(32),
15    spec_id          integer,
16    precisao         char(32),
17    area_infl        integer
18 )
19
20 SELECT AddGeometryColumn(
21     'revisao',
22     'igv_browser',
23     'the_geom',
24     4326,
25     'POINT',
26     2
27 );
```

Figura 6: Ilustração do código SQL para o esquema “current” da BDG BiodOM.

Ter o conhecimento do modelo de dados BiodOM permite, de futuro, interagir com os seus dados no formato nativo, por parte de diferentes utilizadores e partes interessadas, para manipular e utilizar esses mesmos dados noutros formatos que potenciem diferentes utilizações. O conhecimento do modelo de dados será assim pertinente, se se pretender elaborar novos mapas ou converter os dados do BiodOM num outro formato para utilização em diferentes aplicações, entre outras possibilidades.

5.6 Cartografia temática de base do BiodOM

A cartografia temática de base do sistema aplicacional é a cartografia que representa fenómenos localizáveis de diferente natureza e fenómenos temáticos não localizáveis, quantitativos ou qualitativos, sobre uma base topográfica mais ou menos simplificada. Representa, entre outros, os acidentes naturais e artificiais, de acordo com a escala de representação ou a delimitação atual da rede de áreas protegidas em Portugal. Constitui-se, portanto, como toda a cartografia de apoio ao BiodOM, assente em diferentes bases cartográficas de diferentes fontes que, no seu conjunto, tem como objetivo constituir a base geográfica do sistema. Esta cartografia respeita, naturalmente, todos os direitos de propriedade e são utilizadas, na medida do possível, fontes cartográficas com dados abertos. Nos casos de fontes de dados não abertos, mas cujas licenças de utilização se compatibilizam com os objetivos do BiodOM, esses dados foram à mesma integrados.

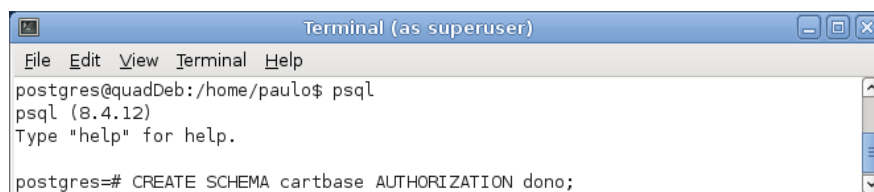
Foi integrada no sistema aplicacional e é disponibilizada, de dois modos diferentes, atendendo não apenas às questões de propriedade dos dados mas também de operacionalidade. Assim, parte dessa informação foi integrada diretamente na BDG em esquemas próprios sendo processada e servida pelo Geoserver. Esta informação é disponibilizada sob a forma de diferentes camadas, na camada do interface gráfico do utilizador (IGU) (ver capítulo 6.3.3), num mapa único e individualizado, onde posteriormente é sobreposta a IGV do BiodOM. Complementarmente, outra parte da informação é disponibilizada através de *mashups* e com recurso a API, que integram camadas de dados cartográficos de base, diretamente da sua fonte. Na tabela seguinte, apresenta-se a cartografia temática de base do BiodOM e as suas principais características:

Nome	Descrição	Fonte	Formato	Data de aquisição	Escala de ref. ^a
Altimetria (30m)	Curvas de nível de Portugal continental com espaçamento de 30 metros. MDT base: ASTER GDEM	Faunália	.shp	2009	-
Áreas Protegidas	Manchas das áreas protegidas de Portugal, disponibilizadas pelo ICNB	ICNB	.shp	2004	-
CAOP 2012	Cartografia Administrativa Oficial de Portugal	IGP	.shp	2012	1:25.000
CLC 2006	Corine Land Cover 2006	IGP	.shp	2006	1:100.000
Fotografia Aérea Google Maps	API de Ortofotografia Aérea disponibilizada pela Google	Google	API	2010	-
Hidrografia	Rede Hidrográfica classificada, disponibilizada pelo INAG	INAG	.shp	2008	1:50.000
Núcleos Urbanos	Principais núcleos urbanos do OSM	OSM	.osm	2012	-
Mapa geral	Mapa geral do OpenStreetMap	OSM	API	2012	-
Rede Viária	Rede viária do OSM	OSM	.osm	2012	-

Tabela 2: Cartografia temática de base do sistema aplicacional BiodOM.

No que respeita à parte da cartografia importada para o BiodOM, o seu processo de

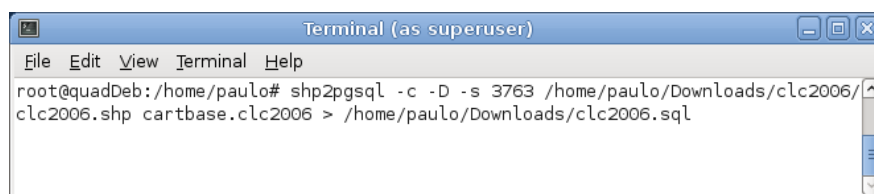
importação realiza-se, sobretudo, de 2 modos, atendendo ao formato de origem. Para os casos cuja origem foram ficheiros shapefile, utiliza-se a aplicação shp2pgsql. Noutras situações, relativamente à importação de ficheiros (.osm) cuja fonte foi o OSM, utiliza-se a ferramenta osm2pgsql, seguindo os procedimentos definidos no ponto 5.7.1. A título de exemplo, para o primeiro caso, a importação da Corine Land Cover 2006 (CLC2006) segue os passos representados da Figura 7 à Figura 9:



```
Terminal (as superuser)
File Edit View Terminal Help
postgres@quadDeb:/home/paulo$ psql
psql (8.4.12)
Type "help" for help.

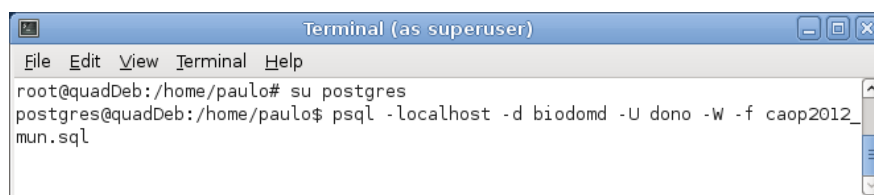
postgres=# CREATE SCHEMA cartbase AUTHORIZATION dono;
```

Figura 7: Passo 1, criação do esquema “cartbase” novo na biodomdb.



```
Terminal (as superuser)
File Edit View Terminal Help
root@quadDeb:/home/paulo# shp2pgsql -c -D -s 3763 /home/paulo/Downloads/clc2006/
clc2006.shp cartbase.clc2006 > /home/paulo/Downloads/clc2006.sql
```

Figura 8: Passo 2, conversão de shapefile para o formato .sql da BDG PostGIS com shp2pgsql.



```
Terminal (as superuser)
File Edit View Terminal Help
root@quadDeb:/home/paulo# su postgres
postgres@quadDeb:/home/paulo$ psql -localhost -d biodomdb -U dono -W -f caop2012_
mun.sql
```

Figura 9: Passo 3, importação de um ficheiro .sql para a biodomdb.

5.7 OpenStreetMap

O OSM, referido anteriormente, apresenta-se como um projeto colaborativo que tem como principal objetivo criar um mapa livre e editável do planeta. Na sua base estão duas grandes forças motrizes: as constantes restrições ao uso ou disponibilidade de informações de mapas e o advento do aparecimento de aparelhos GPS de baixo custo acessíveis a vários utilizadores. Os mapas do OSM são criados usando dados de dispositivos GPS, fotografia aérea, de outras fontes gratuitas ou simplesmente a partir do conhecimento local dos seus utilizadores. Quer as imagens *renderizadas*¹⁶ e o conjunto de dados vetoriais que compõem o OSM estão disponíveis para *download* sob uma licença Creative Commons: *Attribution-ShareAlike 2.0 licence*, uma licença de dados abertos. O mapa padrão é processado por um pacote de *software* denominado Mapnik e os dados armazenados em PostGIS. O conjunto é alimentado por um módulo do Apache (*mod_tile*) (OpenStreetMap 2012).

Sendo constituído exclusivamente por dados abertos o OSM oferece a possibilidade de os

¹⁶ Em computação (gráfica) o processo de *rendering* consiste na geração de uma imagem a partir de um modelo com recurso a um programa de computador

utilizar como uma camada de IG com recurso à API OpenstreetMap e, também, de descarregar a informação e integra-la na BDG, p.e.: com a ferramenta *osm2pgsql* (exclusiva para o ambiente Linux). O bem conhecido Planet.osm é um ficheiro que agrupa todos os dados do OSM: nós, formas e relações que compõem o mapa global. A cada semana é lançada uma nova versão deste ficheiro. Existem também disponíveis arquivos chamados *extracts* que contêm dados do OSM para continentes individuais, países e regiões metropolitanas. Atualmente o ficheiro Planet.osm tem o tamanho, quando comprimido, de 22,6 Gigabytes, sendo os ficheiros nativos do OSM são no formato (.osm). Neste caso, utiliza-se o comando *wget* para obter os extracts em formato shapefile para Portugal, a partir do sítio *Web Geofabrik* (2012), tal como se ilustra na Figura 10:

```

Terminal (as superuser)
File Edit View Terminal Help
root@quadDeb:/home/paulo/Downloads# wget http://download.geofabrik.de/osm/europe
/portugal.osm.bz2
--2012-08-04 12:55:18-- http://download.geofabrik.de/osm/europe/portugal.osm.bz
2
Resolving download.geofabrik.de... 46.4.99.10, 176.9.53.72
Connecting to download.geofabrik.de|46.4.99.10|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 70790181 (68M) [application/x-bzip2]
Saving to: "portugal.osm.bz2"

100%[=====] 70,790,181 5.49M/s in 14s

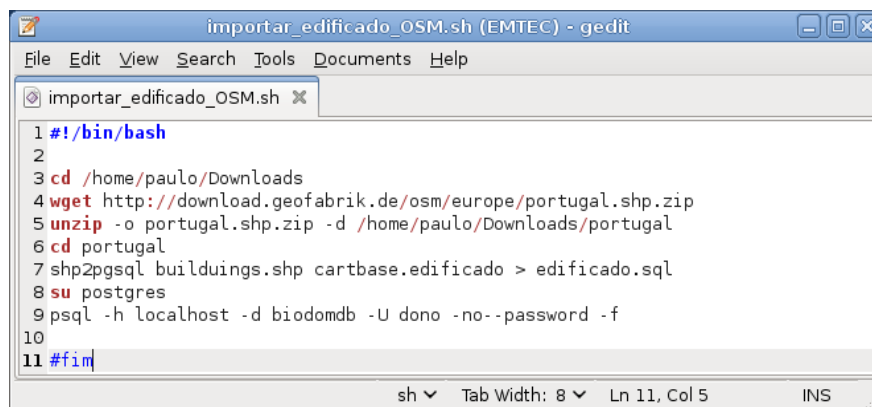
2012-08-04 12:55:32 (4.88 MB/s) - "portugal.osm.bz2" saved [70790181/70790181]

root@quadDeb:/home/paulo/Downloads# osm2pgsql -c -d biodomdb -m -p biodom_osm_
-U dono -W portugal.osm.bz2

```

Figura 10: Importação de um ficheiro .osm para a biodomdb.

A ferramenta *osm2pgsql* converte dados OSM (.osm) num formato que pode depois ser carregado numa BDG PostgreSQL/PostGIS, para um esquema adequado. Assim, de seguida, procedeu-se à importação de dados de uma camada de IG, com recurso a um *cronjob* programado do SO. Um *cron* ou *cronjob* é uma tarefa que se pode adicionar sob a forma de *script* ao *crontab* de um SO de base Unix. O *crontab* é um recurso de agendamento de tarefas baseadas numa data, numa hora ou num intervalo de tempo. Deste modo, a tarefa *cron* é definida por um ficheiro - *crontab* (*cron table*) - localizado normalmente na pasta /etc do SO, onde está descrito o momento e os comandos a executar para cada tarefa, que um *daemon* do SO lê e executa (Debian 2012). Permite aos utilizadores agendar tarefas (através de comandos ou *scripts* utilizados na linha de comandos) para serem executados periodicamente em determinados momentos ou determinadas datas. É frequentemente utilizado em servidores para automatizar a manutenção do próprio sistema a sua administração, pese embora, poder ser utilizado para outros fins dada a sua natureza geral e versatilidade (Debian 2012). No exemplo seguinte importa-se o edificado do OSM para o esquema da cartografia temática de base da BGD do BiodOM, recorrendo ao seguinte *script* implementado no *crontab* do SO (Figura 11):

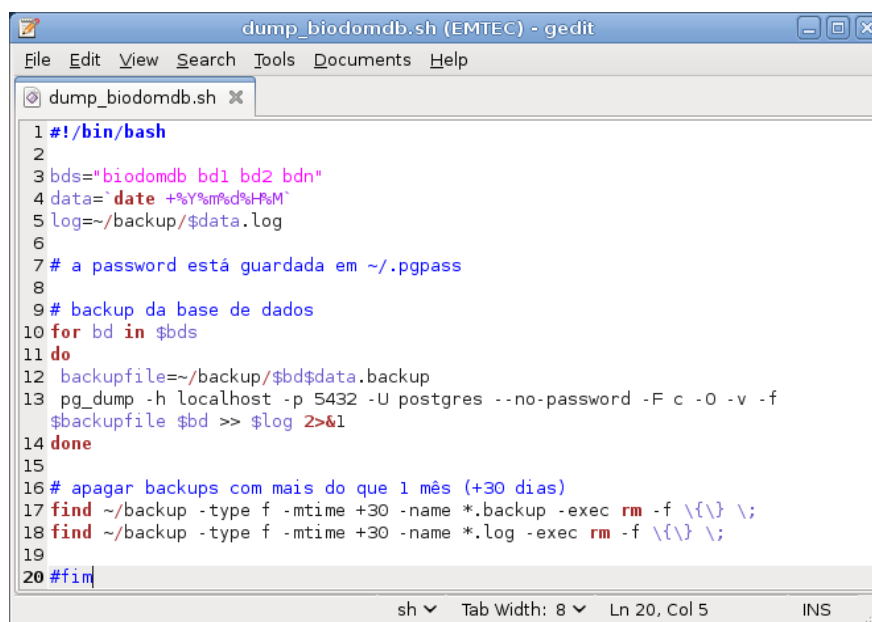


```
1 #!/bin/bash
2
3 cd /home/paulo/Downloads
4 wget http://download.geofabrik.de/osm/europe/portugal.shp.zip
5 unzip -o portugal.shp.zip -d /home/paulo/Downloads/portugal
6 cd portugal
7 shp2pgsql builduings.shp cartbase.edificado > edificado.sql
8 su postgres
9 psql -h localhost -d biodomdb -U dono -no--password -f
10
11 #fim
```

Figura 11: Script do *crontab* programado para a importação da camada de edificado do OSM para o território de Portugal.

5.8 A segurança dos dados do BiodOM

As cópias de segurança da IG do BiodOM são realizadas regularmente, de modo automatizado, recorrendo-se também à implementação de um *cronjob* no *crontab* do Debian. Assim em termos de dados do BiodOM, implementaram-se 2 tipos de *cronjobs*: para **realização automática de dumps** (isto é, cópias de segurança de tabelas e dados das tabelas) frequentes, com periodicidade diária, das BD do BiodOM; e, para a **importação regular (semanal) das camadas de dados do OSM**, que fazem parte da cartografia temática de base (como descrito no ponto anterior). Na Figura 12 ilustra-se um exemplo do *cronjob* programado (*script*) para a realização de *dumps* da BDG:



```
1 #!/bin/bash
2
3 bds="biodomdb bd1 bd2 bdn"
4 data=`date +%Y%m%d%H%M`
5 log=~/.backup/$data.log
6
7 # a password está guardada em ~/.pgpass
8
9 # backup da base de dados
10 for bd in $bds
11 do
12 backupfile=~/.backup/$bd$data.backup
13 pg_dump -h localhost -p 5432 -U postgres --no-password -F c -O -v -f
14 $backupfile $bd >> $log 2>&1
15 done
16 # apagar backups com mais do que 1 mês (+30 dias)
17 find ~/.backup -type f -mtime +30 -name *.backup -exec rm -f {} \;
18 find ~/.backup -type f -mtime +30 -name *.log -exec rm -f {} \;
19
20 #fim
```

(Fonte: Rocha 2012)

Figura 12: Script do *crontab* programado para a realização de dumps das BDG do BiodOM.

Este recorre à função: **pg_dump**, para realizar um *backup* consistente, mesmo que a BDG esteja a ser utilizada no momento. Não bloqueia o acesso de leitura ou escrita à BDG

enquanto o *backup* se realiza. A utilização genérica da função `pg_dump` é: [`pg_dump` [option...] [dbname]]. Os *dumps* podem ser sob o formato de *script*: ficheiros de texto simples contendo os comandos SQL necessários para reconstruir a BDG para o estado em que estava, no momento em que foi realizado o *backup*. Estes ficheiros podem ser utilizados para reconstruir a BDG mesmo noutras máquinas com outras arquiteturas (com algumas modificações). Os formatos de arquivo alternativos podem, se necessário, ser utilizados com o comando `pg_restore` para repor a BD (PostgreSQL 2012).

5.9 Os formatos dos dados do BiodOM

Tal como visto anteriormente, os dados nativos de base do BiodOM estão centralizados numa BDG, neste caso o PostgreSQL/PostGIS. Esta BDG centralizada permite consultar e inquirir dados sob uma multiplicidade de meios, nomeadamente: através de aplicações de extração e importação de dados (tais como a `shp2pgsql`, o `ogr2ogr` ou o `pgsql2shp`); através de aplicações de mapas baseadas na *Web* (tais como o Mapserver ou o Geoserver) que implementam os protocolos WMS, WFS, WCS e WFS-T, para consultar e operar sobre os dados nos formatos GeoRSS, KML, GeoTiff, GIF, PNG, JPEG, CSV, GML, GeoJSON, ESRI Shapefile, entre outros; através do acesso direto de *software* de SIG *Desktop* que ligam diretamente à BDG, com possibilidades de exportação e visualização dos dados.

Para além da possibilidade de visualização dos dados no portal Naturdata, que agrega o sistema aplicacional BiodOM, com a implementação efetiva deste sistema, prevê-se que futuramente a distribuição dos dados de distribuição da Biodiversidade a todos os utilizadores e colaboradores, produto da IGV, seja feita com recurso ao formato XML. Este formato permite representar todos os objetos geográficos dos tipos de dados identificados anteriormente. O formato XML, porém, pode tornar-se extremamente volumoso, ocupando várias centenas de megabytes facilmente. Por esse motivo será distribuído utilizando eficientes algoritmos de compressão de dados tais como o `gzip` e o `bz2`.

5.10 Conclusões do capítulo

Neste quinto capítulo discute-se a estrutura dos dados e, mais particularmente, o modelo de dados do BiodOM. No início do capítulo realiza-se uma abordagem à modelação de dados geográficos em geral. Discute-se a importância da definição dos requisitos de um determinado sistema para melhor se poder proceder a um modelo de dados consistente, bem como o facto de diferentes autores enfatizarem que não haverá um dado modelo de dados correto. Antes, um modelo de dados torna-se mais representativo proporcionalmente ao conhecimento das características geográficas chave, do sistema que representa.

É amplamente reconhecido em vários domínios da ciência, nomeadamente em SIG, que todos os exercícios de modelação consistem numa representação da realidade e nunca reproduzem completamente e em rigor essa realidade observada. Não obstante, como

consequência deste aspeto, um modelo desenvolvido em determinado contexto é sempre suscetível de ser melhorado pelo contínuo desenvolvimento e pesquisa de novas tecnologias respeitantes à realidade a que se refere.

No domínio da modelação de dados de IGV, em particular, focam-se dois importantes aspetos relacionados com a própria natureza da IGV. O primeiro, trata-se da incorporação de um vetor de vaguidade (relacionado diretamente com a incerteza da IG) num modelo de dados com contribuições voluntárias de IG, com recurso a observação (sem utilização de equipamentos de posicionamento), por parte de diferentes utilizadores, apresentando-se uma metodologia que tem vindo a ser adotada por diferentes autores para a integração de IG com determinado grau de vaguidade, em diferentes modelos de dados. O segundo aspeto, prende-se com a própria natureza da IGV e a necessidade do registo das operações de CRUD por parte dos seus utilizadores. Ambos estes aspetos são tidos em linha de conta.

Em termos de modelo de dados BiodOM em si, apresentam-se ao longo do capítulo as premissas de base que levaram ao desenho e implementação do modelo apresentado. Destaca-se a necessidade de criação de diferentes esquemas na BDG em função das fontes de entrada de dados no sistema, bem como da implementação de uma hierarquia de utilizadores, para a implementação de questões relacionadas com a validação dos dados por parte dos coordenadores científicos de cada taxa.

Mais adiante no capítulo foca-se a diferente cartografia temática de base do BiodOM. Esta constituindo-se como toda a cartografia de apoio ao BiodOM, quer em termos de visualização de informação quer em termos de edição de dados de IGV, assenta em diferentes bases cartográficas de diferentes fontes. Integram um conjunto de temas de informação estruturada, pertinente e de algum modo relacionada com o âmbito do BiodOM, ou que, não estando relacionada diretamente com o seu âmbito, no seu conjunto tem como objetivo constituir a base geográfica de referencia do sistema. Dá-se particular importância ao aspeto dos direitos de propriedade desses dados, utilizando-se sobretudo dados abertos. Implementa-se no sistema esta cartografia sob forma de camadas de informação com recurso a API ou integrando-a diretamente na BDG, num esquema próprio, que é atualizada com periodicidade com recurso a um *cronjob*.

6 DESENVOLVIMENTO E IMPLEMENTAÇÃO DO SISTEMA APLICACIONAL: BIODOM

“A stone is simple. But you can build cathedrals of stones.” - George Whitesides

6.1 Introdução

O BiodOM pretende constituir-se como uma ferramenta de produção de IGV no domínio da Biodiversidade. Trata-se de um SIG colaborativo que pretende agregar informação georreferenciada massiva, relativa à distribuição de espécies de seres vivos. Dado que o domínio da Biodiversidade é muito vasto e são necessárias vastas quantidades de informação para que se possam realizar inferências neste domínio, tal como foi considerado nos capítulos precedentes, o sistema aplicativo desenvolvido pretende envolver a contribuição de uma comunidade de utilizadores e interessados na matéria que possam, de algum modo, contribuir para esses dados, sob o ponto de vista de IGV. Neste capítulo descreve-se o processo de implementação da solução conceptualizada que foi de encontro às ideias enunciadas anteriormente. Numa primeira etapa, para suportar o conjunto de funcionalidades e objetivos específicos do Projeto, foram identificados os requisitos operacionais do sistema. Estes permitiram conceptualizar uma arquitetura adequada, numa segunda etapa. É feita a descrição das suas soluções tecnológicas e funcionalidades principais sendo, de igual modo, apresentadas as atuais aplicações práticas. No que respeita a este último ponto evidencia-se também a descrição detalhada da solução implementada ao nível da interface: Portal Naturdata/BiodOM, dado que foi necessário considerar no sistema a integração de aspetos relativos a uma plataforma já existente.

6.2 Requisitos operacionais

Antes de se discutir a conceção heurística, o desenvolvimento e a estratégia de implementação do BiodOM, foi necessário levar a cabo uma avaliação dos requisitos e das ferramentas necessárias. Deste modo, considerando tudo o que foi exposto nos capítulos precedentes, em termos de requisitos operacionais e necessidades do sistema, era esperado que este apresentasse:

- uma solução de baixo custo devido a restrições económicas (em que se partiu do princípio que o SGLA seria o caminho para atingir este objetivo);
- uma BD geoespacial robusta ;
- capacidade de edição simultânea multi-utilizador ;
- possibilidade de contribuição por uma vasta gama de utilizadores (de interessados sem conhecimentos prévios de SIG até profissionais no domínio da Biodiversidade);
- hierarquia de permissões para validação e credibilização do trabalho colaborativo

de produção de IGV ;

- um modelo de dados adequado ;
- questões relativas à adaptação com a interface Naturdata já existente ;
- requisitos específicos no que respeita à produção de IGV, nomeadamente o *tracking*¹⁷ de alterações por parte dos utilizadores/colaboradores;
- uma gestão de utilizadores simplificada;
- interface intuitiva para edição de informação pelos utilizadores;
- diferentes esquemas dentro da BDG com possibilidade de interligação entre eles (por motivos de organização lógica, segurança e integridade da BDG) .

6.3 Conceptualização da arquitetura do sistema

É frequente verificar em diferentes plataformas SIGWeb uma arquitetura multi-camada. Isto é, a presença de uma arquitetura cliente-servidor em que a apresentação, o processamento e a gestão de IG, são processos distintos (Martins 2010). Yang *et al.* (2007) descrevem uma arquitetura centrada no recurso para o desenvolvimento deste tipo de sistemas aplicacionais. Caracterizam o recurso como sendo uma peça estruturada de informação criada para ser pesquisada, descoberta e visualizada *online*. A arquitetura, em si, apresentada por estes autores, inclui três camadas principais: a **camada dos recursos**, correspondente ao armazenamento físico da informação estruturada em BD ou ficheiros; a **camada de acesso**, que inclui todo o código e *software* desenhados para providenciar acesso aos recursos no formato apropriado; e, a **camada da IGU**: que, na sua essência, é a componente da arquitetura do lado cliente.

A opção por uma arquitetura em camadas centrada no recurso dotou o sistema de grande modularidade. Permite a escolha de diferentes e variados *softwares*, independentemente de questões relacionadas com dependências e interoperabilidade. Em muitas situações os *softwares* escolhidos e implementados, podem tratar-se de SGLA que, desenvolvidos e adaptados atendendo a necessidades específicas do sistema sempre numa lógica de seguimento da filosofia do *software* livre, contribuem para o próprio desenvolvimento do *software*. Esta abordagem permite simultaneamente uma melhoria na manutenção do sistema, dado que permite ao técnico que implementa o sistema aplicacional adicionar, modificar ou até reutilizar componentes específicos enquanto os restantes se mantêm em funcionamento. Na figura seguinte, esquematiza-se a arquitetura projetada e implementada, para o sistema aplicacional BiodOM:

¹⁷ Registo informático automatizado de operações e alterações realizadas relativamente a um estado prévio

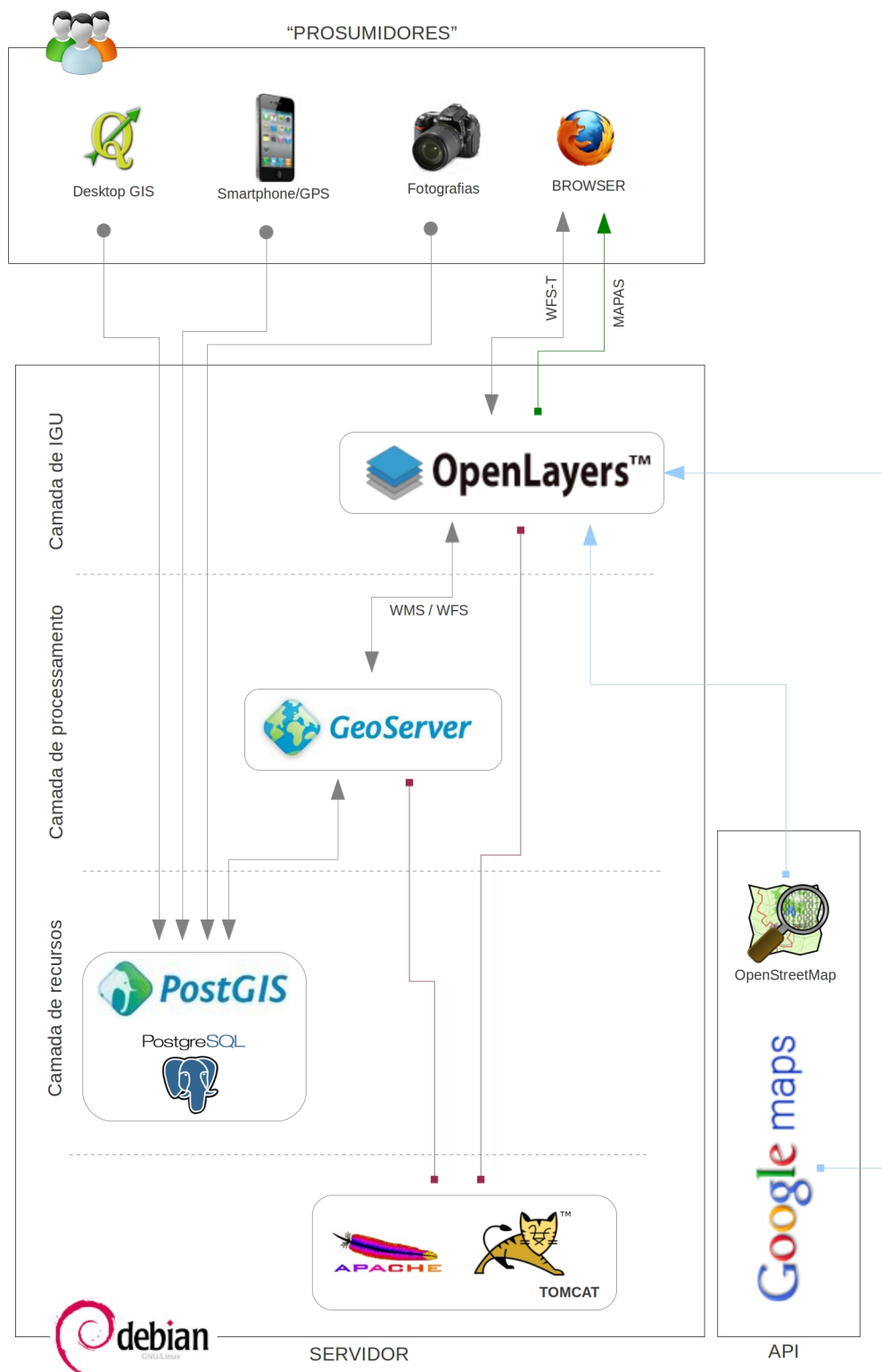


Figura 13: Arquitetura do sistema aplicacional.

6.3.1 Camada de recursos: PostgreSQL/PostGIS

Ao nível da camada de recursos, implementou-se uma BDG PostgreSQL ativada espacialmente pela extensão PostGIS. A camada de acesso, garante o acesso à camada de recursos em 2 níveis, *input* de informação e *output* de informação (do ponto de vista do servidor). Os *inputs* de informação, tal como foi considerado no capítulo anterior, são constituídos por quatro modos de entrada de IGV por parte dos colaboradores do BiodOM: **browser editor**; **exif georeader**; **GPX georeader**; e, **QGIS editor** (descritos em detalhe no ponto 6.6). De realçar ainda que na implementação do sistema aplicacional, se teve particular cuidado com a administração do BiodOM ao nível da camada de recursos. Esta é garantida com recurso a configurações na linha de comandos através de ligação com recurso ao protocolo *Secure Shell* (SSH) ao servidor e, ainda, diretamente no *browser*. As operações de gestão e manutenção por parte dos coordenadores dos diferentes grupos taxonómicos do portal Naturdata, recorrem à interface de administração *Web* do Geoserver e ao Phpmyadmin para a BDG (PostgreSQL/PostGIS).

6.3.2 Camada de processamento: Geoserver

Na camada de processamento, foram implementados diferentes servidores. O Apache é um dos servidores HTTP mais populares, de código aberto e de instalação e integração fácil, tendo a escolha recaído sobre o mesmo. A operar sobre o Apache (com o módulo Tomcat) implementou-se o Geoserver como servidor cartográfico. Este *software*, tal como foi detalhado anteriormente, tem capacidade de servir dados espaciais em conformidade com as especificações OGC: WMS e WFS. Apresenta ainda uma vantagem face a outros servidores cartográficos mais populares (p.e. o Mapserver), de suportar o padrão WFS-T, uma condição fundamental para processar e editar dados geoespaciais vetoriais via *Web*.

6.3.3 Camada de Interface Gráfica do Utilizador: OpenLayers

Ao nível da camada de IGU considerou-se a biblioteca de código aberto OpenLayers, para preencher as restantes funcionalidades necessárias do sistema. Neste particular definiram-se como requisitos práticos específicos (do ponto de vista Naturdata), três tipologias de visualização e/ou acesso à IG que se prendem com a configuração de três tipologias de mapas diferentes com diferentes funcionalidades específicas: **ficha de espécie**; **ficha de espécie (em modo edição)**; e, **mapa global BiodOM**. O desenvolvimento do sistema operacional BiodOM, contemplou a configuração e estilização com recurso a esta biblioteca do código HTML¹⁸ das páginas *Web*, aplicado a estas tipologias de mapas.

¹⁸ HTML é a sigla inglesa *HyperText Markup Language*, isto é, Linguagem de Marcação de Hipertexto, utilizada para produzir páginas na Web que podem ser interpretadas por um *browser*

6.3.3.1 Mapa da ficha de espécie (modo visualização)

O mapa da ficha de espécie é o mapa que se integra na ficha detalhada de cada espécie, do portal Naturdata. Apresenta, por isso, como IG principal a distribuição da espécie no território de Portugal, tal como se pode observar no exemplo da figura seguinte:

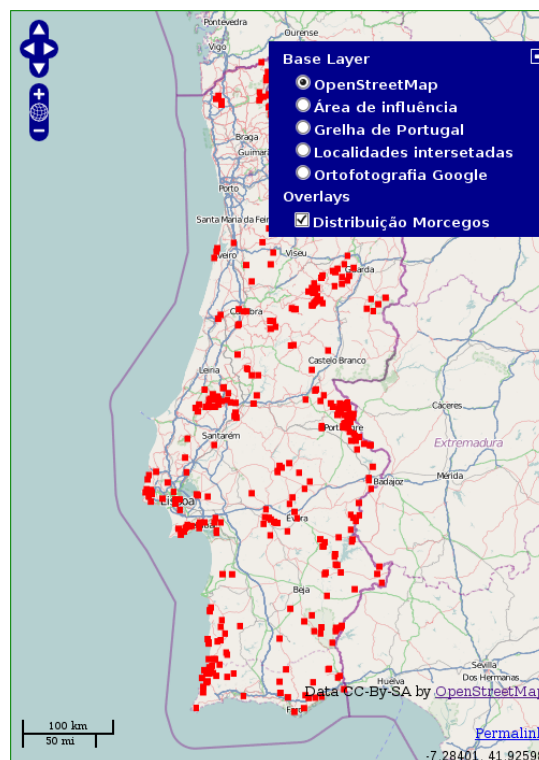


Figura 14: Exemplo do mapa da ficha de espécie.

A renderização dos mapas de cada ficha de espécie tem na sua base uma instrução em PHP¹⁹, no código da página Web, em que para essa dada espécie adiciona o código ID da espécie à estilização do OpenLayers, nomeadamente nos parâmetros e classes, de modo a “chamar” o mapa respetivo com os parâmetros corretos. Para esta funcionalidade, foi necessário padronizar a informação no Geoserver, em termos de nomenclatura das *stores* e *layers*, adicionando-lhe o ID da espécie. Em termos de seleção de camadas do mapa, são apresentadas: camada de cobertura aerofotográfica Google Maps, camada OSM, IGV relativa à distribuição dessa espécie, tema área de influência, tema grelha_pt e tema localidades intersetadas. Os elementos de função de mapa da ficha de espécie estilizados são: botões de *pan*, botão de zoom, elemento de escala numérica e elemento de seleção de camadas.

6.3.3.2 Mapa da ficha de espécie: modo de edição (módulo *browser editor*)

O mapa da ficha de espécie em modo de edição, corresponde ao módulo *browser editor*, de

¹⁹ PHP é um acrónimo recursivo para *PHP Hypertext Preprocessor*, uma linguagem de programação utilizada para o desenvolvimento de aplicações que operam do lado do servidor e capazes de gerar conteúdo dinâmico na Web

input de IG no BiodOM por parte dos seus colaboradores. Este módulo consiste numa extensão das funcionalidades da estilização do mapa da ficha de espécie, para criação de um modo de edição diretamente no *browser*, tendo por base uma camada de ortofotografia aérea do Google Maps. Em termos de funcionamento, o módulo ativa-se mediante o *login* no portal Naturdata por parte do colaborador, para poder ter acesso às possibilidades de edição do mapa da ficha de espécie em questão, que são disponibilizadas após esse *login*. Em termos de sistema aplicacional, posteriormente e à semelhança do mapa anterior, terá uma instrução PHP em que para cada espécie é adicionado o código ID da espécie à estilização do OpenLayers, nomeadamente nos parâmetros e classes, para “chamar” o mapa respetivo com os parâmetros pretendidos corretos, nomeadamente a IG geográfica relativa à espécie em questão a partir da BDG (Figura 15).



Figura 15: Exemplo do mapa da ficha de espécie em modo de edição.

Para esta instrução padronizou-se a informação no Geoserver, em termos de nomenclatura das *stores* e *layers*, adicionando-lhe o código ID da espécie, visto que a requisição WMS/WFS é também neste caso efetuada de modo automático. Os elementos de função de mapa são: botões de *pan*, botão de zoom, elemento de escala numérica, elemento de seleção de camadas e indicação das coordenadas do ponteiro. Naturalmente, quando este mapa é apresentado, é acrescido da barra de ferramentas de edição (no *browser*). Em termos de seleção de camadas, são apresentadas na caixa de seleção as *camadas*: imagem aérea Google Maps e pontos de distribuição da espécie (aos quais se vai adicionar nova IGV). O input e introdução de IGV relativa à distribuição de espécies com recurso a

este módulo realiza-se assim, por definição, no próprio *browser* sobre ortofotografia aérea, com recurso às ferramentas de edição disponibilizadas. O sistema aplicacional comunica depois com a BDG através da especificação WFS-T, com recurso ao Geoserver. No Anexo A ilustra-se o código de estilização da requisição WFS-T com recurso ao OpenLayers.

6.3.3.3 Mapa global BiodOM

O mapa global BiodOM é o elemento da IGU mais visível e, talvez, de maior destaque. Constitui a porta de entrada para o BiodOM, apresentando-se como um mapa global onde se pode consultar seletivamente toda a informação que o BiodOM disponibiliza, nomeadamente a produzida por IGV (Figura 16).



Figura 16: Aspeto geral do mapa global BiodOM.

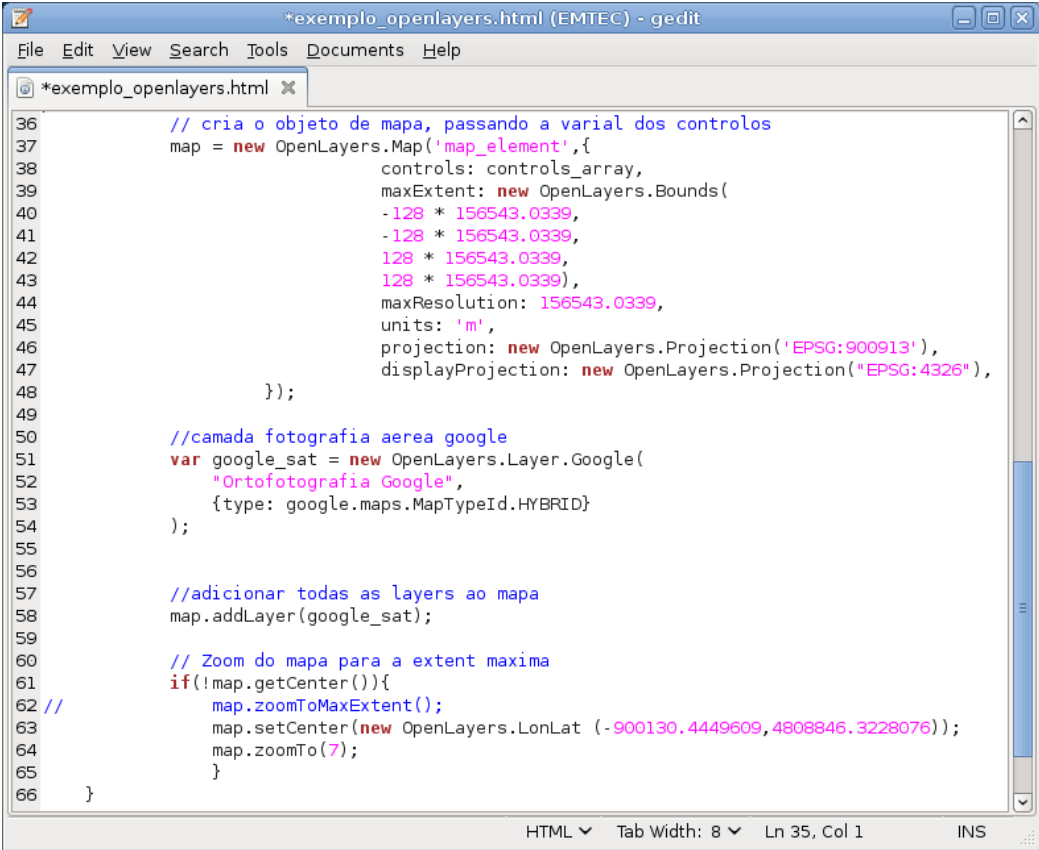
Os elementos de função de mapa do mapa global BiodOM são constituídos por: botões de *pan*, barra de zoom, elemento de escala gráfica e elemento de seleção de camadas. Dois elementos adicionais diferenciais neste mapa são: 1 – a caixa de pesquisa (por nome, id, taxa ou outros); 2 – a camada da cartografia temática de base do BiodOM. No que respeita à caixa de pesquisa, é necessário efetuar uma pesquisa nesta caixa para o mapa apresentar informação relevante, além da apresentada por definição. Isto porque, dada a complexidade do sistema e a quantidade de informação envolvida, não é possível apresentar toda a informação sob a forma de camadas de seleção. Este aspeto irá ser referido em detalhe num futuro guia de utilização do BiodOM.

Deste modo, em termos de seleção de camadas, são apresentadas no seletor de camadas de informação, as seguintes camadas por defeito: camada aerofotográfica Google Maps,

camada OSM, camada da cartografia temática de base BiodOM e camada da espécie em destaque (uma camada temática que será atualizada manualmente com determinada frequência, que se prende com questões de marketing e promoção do Projeto).

6.3.4 API Google Maps

A camada de cobertura aerofotográfica é garantida no sistema com recurso à implementação da API Google Maps. A escolha desta, face a outras populares fontes de cobertura aerofotográfica tal como o Yahoo Maps ou o Bing Maps, entre outras, recaiu sobretudo devido a questões relacionadas com a licença de utilização e com a qualidade da ortofotografia. A API Google Maps permite criar um *mashup* embutindo uma camada de IG em diferentes aplicações *Web* de mapas, nomeadamente o OpenLayers. Nos últimos anos tem sido grande a tendência de variados sítios *Web* disponibilizarem conteúdos sobre mapas fornecidos por diferentes entidades com recursos à API. Provavelmente estará no topo da lista das mais populares. A figura seguinte ilustra um extrato de parte do código da sua implementação:



```
36      // cria o objeto de mapa, passando a varial dos controlos
37      map = new OpenLayers.Map('map_element',{
38          controls: controls_array,
39          maxExtent: new OpenLayers.Bounds(
40              -128 * 156543.0339,
41              -128 * 156543.0339,
42              128 * 156543.0339,
43              128 * 156543.0339),
44          maxResolution: 156543.0339,
45          units: 'm',
46          projection: new OpenLayers.Projection('EPSG:900913'),
47          displayProjection: new OpenLayers.Projection("EPSG:4326"),
48      });
49
50      //camada fotografia aerea google
51      var google_sat = new OpenLayers.Layer.Google(
52          "Ortofotografia Google",
53          {type: google.maps.MapTypeId.HYBRID}
54      );
55
56
57      //adicionar todas as layers ao mapa
58      map.addLayer(google_sat);
59
60      // Zoom do mapa para a extent maxima
61      if(!map.getCenter()){
62          map.zoomToMaxExtent();
63          map.setCenter(new OpenLayers.LonLat (-900130.4449609,4808846.3228076));
64          map.zoomTo(7);
65      }
66  }
```

Figura 17: Código utilizado de implementação da camada aerofotográfica da API Google Maps.

O Google Maps, no formato ortofotográfico, apresenta informação detalhada disponível para os Estados Unidos, Canadá, Porto Rico e a maioria dos países europeus, com um grau de detalhe considerável. Possui ainda imagens de satélite de resolução moderada para todo o

mundo, estando disponíveis resoluções de 1-2 metros para os Estados Unidos e algumas cidades selecionadas em todo o mundo (Google 2012).

6.4 Implementação

O desenvolvimento de todo o Projeto decorreu numa máquina local. Pretendeu-se a criação de um ambiente protótipo para estudos, testes, configurações e implementação do sistema aplicacional, que posteriormente possibilitasse a sua transferência para um servidor dedicado de produção de um modo relativamente fácil. Em termos de *hardware*, como principais características físicas do servidor, destacam-se: máquina Intel QuadCore 2.6Ghz, 8 Gigabytes de memória RAM, RAID de discos rígidos (2x500Gb); e, ligação fibra 100/10Mbit. No que respeita ao SO, a escolha recaiu sobre o Debian, na sua versão 6.0.5 “Squeeze” (amd64), um SO Linux. Não sendo o objetivo deste trabalho prático debater as opções tomadas em termos das escolhas dos diferentes pacotes de *software* de entre a multiplicidade de softwares existentes importa, no entanto, referir que a opção pelo Debian se deve: à facilidade de atualização dos pacotes de software, através do seu sistema de gestão de pacotes - APT, que permite atualizações de pacotes de aplicações de software de modo relativamente fácil a partir de um vasto repositório de pacotes de aplicações; ao modo de disponibilização de novas versões, em que o Debian procura sempre manter os pacotes de software mais estáveis após um período de testes considerável, garantindo assim uma grande estabilidade e fiabilidade, o que o torna num SO vantajoso para utilização quer em ambientes de produção; à vasta comunidade, uma das mais antigas em termos de SO Linux; e, aos recursos e informação disponíveis sobre este SO.

Outro aspeto, refere-se ao facto de longo do trabalho prático se ter mantido o cuidado de, sempre que possível, recorrer à linha de comandos para realizar as diferentes instalações e configurações do sistema aplicacional, tal como apresentado no Anexo B. Este aspeto, para além de encetar algumas vantagens, nomeadamente de rapidez de execução de instalações, configurações e testes, uma vez ambientado com a mesma, traduz-se noutro aspeto muito importante. Trata-se da previsível relativa facilidade de implementar e testar o sistema aplicacional remotamente, quando este for transferido para um servidor dedicado de produção, com recurso à própria linha de comandos via SSH.

6.4.1 O sítio www.biodom.pt

O sítio *Web* www.biodom.pt trata-se de um espaço protótipo onde se implementou a interface do Projeto. Pretende servir o duplo objetivo de implementar/demonstrar a parte visível do Projeto, ao nível da camada de IGU; e, servir de espaço protótipo do BiodOM, para realização de testes operacionais na componente interface Naturdata/BiodOM do sistema aplicacional (previamente à sua integração no portal Naturdata). Neste espaço constaram ainda outras secções, tais como a descrição e objetivos do projeto BiodOM, bem como uma pequena wiki relativa ao BiodOM e às operações de produção de IGV,

destinadas aos colaboradores que desejem intervir, p.e.: “Como produzir dados com o Quantum GIS no BiodOM?”, num contexto de Projeto (SIG) colaborativo.

Esta tipologia de configuração e funcionamento do espaço protótipo, perdurou até à configuração final do Projeto. Num futuro breve, após concluídas todos os testes e conceptualizações, irá migrar-se o sistema aplicacional para um servidor dedicado, atualizando-se os *links* de ligação ao servidor e se realizará a integração do BiodOM no Portal Naturdata. Porém, para efeitos da conceptualização, testes e otimização de configurações, a tipologia de funcionamento foi a descrita.

6.5 Especificações particulares: interface Naturdata/BiodOM

As especificações particulares da interface Naturdata/BiodOM referem-se aos aspetos que foram necessários configurar, para permitir a correta integração com o portal Naturdata, para além dos módulos de entrada de aquisição de IGV e da IGU. Não constituem mais do que requisitos do sistema da parte do portal (Naturdata). São eles: 1 - **a agilização da gestão entre os utilizadores (colaboradores) Naturdata existentes e a sua replicação no BiodOM (PostGIS)**, do ponto de vista do registo e permissões dos mesmos; 2 – **a criação de utilizadores na BDG do BiodOM após o registo no portal Naturdata**; 3 – **a importação das tabelas MySQL do Naturdata e a integração da sua informação no BiodOM**. Dado que para o ponto 1 se prevê a criação manual de utilizadores, detalham-se de seguida os pontos 2 e 3 e as configurações envolvidas na sua implementação.

6.5.1 Criação de utilizadores na BDG BiodOM após o registo no portal Naturdata

O registo de novos colaboradores no BiodOM é realizado a partir do seu registo, em primeira instância, no portal Naturdata. Quando se assiste ao processo de registo de utilizadores no portal Naturdata, uma instrução em PHP irá automaticamente criar o colaborador no BiodOM. Este registo, em última análise, não corresponde mais do que a criação de um novo utilizador na BDG PostgreSQL/PostGIS. Genericamente, utilizam-se duas instruções em PHP embutidas no código HTML da página de criação de utilizador do Naturdata, que são executadas de modo automático logo após a criação do utilizador. A primeira instrução é a `pg_connect`, cuja utilização genérica é: `[resource pg_connect ( string $connection_string )]` e abre uma ligação com um SGBD PostgreSQL, cujos parâmetros são especificados na `$connection_string`. Retorna um recurso de ligação em caso de sucesso, ou retorna o valor *FALSE* se a ligação não puder ser estabelecida. De seguida, a instrução `pg_query` cuja utilização genérica é: `[resource pg_query ([ resource $connection ], string $query )]` vai executar a *query* na BD à qual se ligou anteriormente, neste caso uma query de criação de utilizador com o mesmo nome do utilizador criado no Naturdata (PHP 2012). No Anexo C apresenta-se o código envolvido na criação de um novo utilizador.

6.5.2 Importação das tabelas MySQL Naturdata e integração no BiodOM

A importação dos dados da BD de informação de Biodiversidade do portal Naturdata construída em MySQL, que constitui a informação existente relativa às espécies (no entanto sem dados sobre a distribuição de espécies) para a BDG do BiodOM desenvolvida em PostgreSQL/PostGIS é realizada com recurso a um processo manual e com a periodicidade que se julgar necessária. Esta importação torna-se pertinente para possibilitar a conjugação de determinados aspetos de informação sobre Biodiversidade do Naturdata, com a componente de distribuição espacial das espécies conferida pelo BiodOM. Porém, prevê-se que estes dois sistemas permaneçam individualizados, no que respeita às suas BD. Daí que seja necessário proceder regularmente a essa importação.

Assim, a importação das tabelas MySQL do Naturdata compreende a importação para um esquema próprio da BDG do BiodOM (ver capítulo 5). Apesar do processo de *dump* das tabelas de uma BD em MySQL providenciar um modo de compatibilidade para o PostgreSQL, o processo de importação contudo não é linear. Existem limitações ao nível da compatibilidade dos tipos de dados entre ambos os SGBD. Genericamente, a cada importação é necessário: 1 - realizar um *dump* da BD em MySQL: [# mysqldump --compatible=postgresql Naturdata > naturdata2biodom.sql]; 2 - alterar as propriedades de *quote escaping* do ficheiro (.sql): [# sed -i "s/\\'\\'/'/' naturdata2biodom.sql]; 3 - modificar o tipo de dados do ficheiro (.sql), com recurso a um *script* de conversão fornecido por <http://www.xach.com/aolserver/mysql2psql.pl>; e, 4 - importar, finalmente, o ficheiro (.sql) trabalhado com o comando *psql* do PostgreSQL (Wikibooks 2012).

6.6 Especificações funcionais: ferramentas do sistema e produção de IGV

No desenvolvimento inicial do sistema aplicacional BiodOM conceptualizaram-se quatro modos diferentes de aquisição e integração de IGV a produzir pelos seus colaboradores. Estas configurações, apresentam-se sob a forma de módulos ou ferramentas, que correspondem a requisitos específicos do sistema em termos de *input* de IGV. São deste modo fontes de entrada de dados por parte dos colaboradores BiodOM os módulos descritos a seguir.

6.6.1 Browser editor

O funcionamento deste módulo abordou-se em detalhe no ponto 6.3.3.2.

6.6.2 Exif georeader

Relativamente ao módulo Exif georeader, genericamente, este corresponde ao carregamento de fotografias geoetiquetadas e importação da IG a partir dos campos de coordenadas do “exif” da fotografia. De um modo geral, na área de utilizador do portal Naturdata existe uma opção para o carregamento de fotografias que retratem espécies. Estas fotografias são carregadas para o repositório de fotografias do portal Naturdata,

contendo no repositório de fotografias o nome da foto a identificação do utilizador, a data de carregamento da informação e um id sequencial. Depois de validadas e identificadas as espécies pelos coordenadores do respetivo grupo taxonómico, será acrescentado à informação um código “spec_id” semelhante à referência na BDG de cada espécie. Por uma questão de direitos de utilização das fotografias, no momento de desenvolvimento do trabalho prático não se prevê a sua cópia para os repositórios de dados de IGV do próprio BiodOM (o que não invalidará a possibilidade de vir a ocorrer num futuro próximo).

Uma instrução programada corre posteriormente do lado do portal Naturdata e lê o “exif” das fotografias adicionadas. Envia uma instrução via *Web* para o BiodOM, que carrega essa informação geoespacial relativa à fotografia na BDG. O módulo Exif georeader, dada a extensão do código envolvido na sua programação, apresenta-se no Anexo D.

6.6.3 GPX georeader

Este módulo consiste na importação de *tracks* GPS (ficheiros .gpx) e nuvens de pontos de levantamentos de ocorrências de espécies, de recetores GPS (tais como smartphones e aparelhos GPS dedicados) para a BDG. Do lado do portal Naturdata, na área de utilizador, existe uma opção para carregar ficheiros para o repositório de base de ficheiros de IGV do servidor BiodOM, nomeadamente os ficheiros (.gpx). Posteriormente um *script* programado no SO, encarrega-se de incumbir à aplicação ogr2ogr a sua integração direta na BDG. Esse *script* invoca num primeiro passo a conversão do (.gpx) para o formato (.shp), com recurso à aplicação gpx2shp. Posteriormente realiza a conversão do ficheiro (.shp) em (.sql) com o comando shp2pgsql e, finalmente, conclui a sua integração na BDG: biodomdb.

6.6.4 QGIS editor

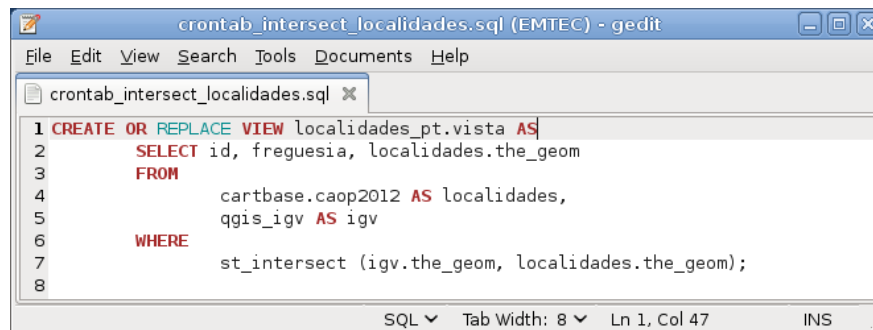
Este módulo consiste na ligação direta do *software* QGIS ao esquema “qgis_igv” da BD BiodOM para carregamento de dados, padronizados pelos campos da tabela. Permite edições diretas e de modo manual na BDG com recurso ao *software* QGIS, mediante um conjunto de regras e orientações específicas predefinidas. A edição de IGV com recurso ao *software* QGIS, por parte dos colaboradores do BiodOM, é bastante vertical. Quando se regista no portal Naturdata é criado automaticamente o utilizador na BDG, que permite simultaneamente a sua ligação à BDG biodomdb com recurso ao *software* QGIS. Mediante as instruções de como se ligar à BDG facultadas ao colaborador, utilizando esse mesmo nome de utilizador, este fica com a possibilidade de editar a IG de relevância deste modo.

6.7 Especificações de operações de pós-processamento: a IGV aplicada ao conhecimento geoespacial da Biodiversidade

No que respeita às tarefas de pós-processamento dos dados, conceptualizaram-se diferentes *cronjobs* no Debian e *triggers* no PostGIS. Um *trigger* é uma funcionalidade existente no SGBDGR que mediante o reconhecimento da execução de determinados comandos SQL numa vista ou numa tabela, permite que decorra algum tipo de

processamento adicional e condicional sobre esses comandos (Obe e Hsu 2011). No caso particular do sistema BiodOM irão utilizar-se *triggers* que executam funções adicionais aos comandos *INSERT*, *UPDATE* e *DELETE*. Os *triggers* conceptualizados, em termos de pós-processamento, foram: **posprocess#1: trigger para tratamento dos dados**, de modo a incorporar o vetor de vaguidade e incerteza de IGV, em esquema próprio: “vagueness_igv”; e, **posprocess#2: trigger de tracking (registo) das operações de CRUD**, fruto dos inputs de IGV por parte dos colaboradores do BiodOM.

Os *cronjobs* conceptualizados permitem realizar o seguinte tipo de operações: **posprocess#3: intersecção dos pontos e nuvens de pontos de ocorrências de espécies com o esquema “grelha_pt”**, para delimitar as quadrículas de ocorrências de cada espécie no território de Portugal; **posprocess#4: delimitação de bounding boxes** no esquema “bounding_box_pt” (tema poligonal), relativamente aos pontos e nuvens de pontos de ocorrências de espécies, atendendo também à informação do campo “area_inf” (relativamente à área de influência de cada espécie), para cada espécie; **posprocess#5: intersecção dos pontos e nuvens de pontos de ocorrências de espécies com o esquema “localidade_pt”**, obtido a partir dos lugares e freguesias da carta administrativa oficial de Portugal (CAOP), para determinar as localidades de ocorrência de cada espécie. A Figura 18 representa a sintaxe SQL para a intersecção das nuvens de pontos de ocorrências de espécies com os lugares/freguesias e respetiva produção do tema na BDG:



```

1 CREATE OR REPLACE VIEW localidades_pt.vista AS
2   SELECT id, freguesia, localidades.the_geom
3   FROM
4     cartbase.caop2012 AS localidades,
5     qgis_igv AS igv
6   WHERE
7     st_intersect (igv.the_geom, localidades.the_geom);
8

```

Figura 18: Exemplo da sintaxe SQL de ficheiro .sql externo exemplo, invocado pelo cronjob de intersecção das nuvens de pontos com o tema de localidades.

As operações de pós-processamento de informação constituem um aspeto de relevo do BiodOM. Em primeira instância porque conferem consistência à IGV nas suas dimensões de vaguidade e *tracking* das alterações efetuadas pelos colaboradores. Sendo em termos de IGV o registo do histórico das alterações efetuadas pelos utilizadores do BiodOM um aspeto importante, sempre que ocorre uma operação CRUD, cria-se um registo dessa operação numa tabela de histórico, que regista não apenas a operação realizada, como também o nome do próprio utilizador que a realizou. Este *trigger* produz um novo tema na BDG, tal como apresentado no modelo de dados. A sintaxe do *trigger* de tracking das operações de CRUD implementado no BiodOM apresenta-se no Anexo E.

Para além disso, as operações de pós-processamento, permitem realizar uma primeira inferência de análise espacial sobre a IGV produzida, de modo a estruturarem essa informação para a tornarem em IG mais importante do ponto de vista do conhecimento da Biodiversidade, ao nível da distribuição de espécies pelo território. Assim, a interseção das ocorrências de espécies com o tema das localidades confere uma dimensão visual prática dessa distribuição. Já a delimitação poligonal de *bounding boxes* para cada espécie, tem em consideração o campo área de influência dessa espécie, um dado que apesar de ser mais empírico do que científico garante uma maior aproximação da realidade, para melhor se estimarem probabilidades de distribuição de espécies. Já o cruzamento do registo de ocorrências de espécies com a grelha 10x10km de Portugal, pretende construir uma malha das áreas de maior riqueza de Biodiversidade de Portugal, tal como se ilustra na Figura 19:

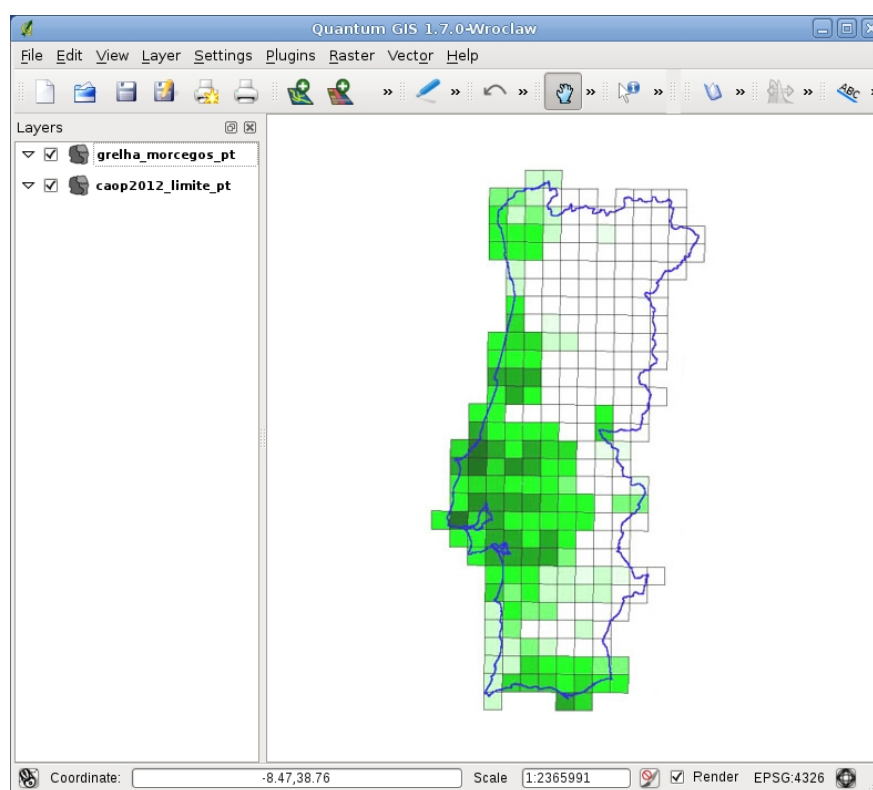


Figura 19: Produto do cruzamento da ocorrência de uma espécie com a grelha de Portugal.

6.7.1 ST_Intersects

Na base de cada *cronjob* existem funções específicas do SGBDG PostgreSQL/PostGIS. Os *posprocess#1* e *posprocess#3* têm na sua base a função **ST_Intersects**. Utiliza as colunas de geometria das tabelas, para devolver um valor “verdadeiro” se duas geometrias se intersectam (partilham uma porção de espaço) ou “falso” em caso contrário (se forem disjuntas). Utilização genérica: [ST_Intersects(geometry geomA, geometry geomB)]

6.7.2 ST_ConvexHull

O *posprocess#2* consiste na delimitação de *bounding boxes* no tema *bounding_box_PT*

(tema poligonal), relativamente aos pontos e nuvens de pontos de ocorrências de espécies. Atende à informação do campo “area_inf” (relativamente à área de influência de cada espécie) para cada espécie identificada, para produzir um polígono envolvente às geometrias de origem. Esse polígono engloba também uma área limítrofe de influência na periferia do menor polígono envolvente da união das geometrias em apreço, dada pelo valor do campo “area_inf”. Neste caso, utiliza-se a função do PostGIS: **ST_ConvexHull**. Representa a geometria mínima convexa que inclui todas as geometrias dentro do conjunto. Sendo utilizada com coleções geométricas, determina uma área afetada com base num conjunto de observações de pontos. A utilização genérica é: [ST_ConvexHull(geometry geomA)].

6.8 Aplicação a uma prova de conceito (PdC)

Uma prova de conceito (PdC) é um modelo prático que permite comprovar uma proposta conceptual do modelo (teórico) estabelecido por uma determinada pesquisa ou investigação. Pode também ser considerada como uma implementação, geralmente resumida, de uma metodologia ou conceptualização, com o intuito de verificar que a referida proposta conceptual é suscetível de ser explorada de uma maneira útil (Pinheiro 2012).

A PdC constitui-se, assim, como um passo importante no processo de criação de um protótipo efetivamente operacional. Em sistemas de informação, e em particular nos SIG, a PdC é uma demonstração de que um sistema está, em princípio, operacional sem a necessidade de que a sua completa implementação esteja concluída. Ainda segundo Pinheiro (2012), *a PdC serve para validar um projeto de rede de computadores antes que este seja executado na prática. Todos os resultados dos testes devem ser plenamente avaliados na PdC, uma vez que nas fases posteriores, que envolvem a instalação e operação da rede, será maior a relutância às mudanças ou à re-interpretação dos requisitos.*

O sítio Web www.biodom.pt constitui a PdC do sistema aplicacional conceptualizado, bem como a implementação prototípica, previamente à sua integração completa no portal Naturdata. Este espaço permitiu testar a operacionalidade do sistema globalmente. Para tal foi criado na BDG um tema da espécie de morcegos cavernícolas: *Pipistrellus pygmaeus* (neste caso obtido a partir do sítio Web do ICNB). Foram também criados diferentes utilizadores para testar todas as diferentes funcionalidades do sistema (inclusive em modo de edição simultânea). De seguida, foram realizados os testes de operacionalidade do sistema aplicacional e anotadas os resultados, tal como constantes no Anexo F.

6.9 Conclusões do capítulo

O sexto e último capítulo culmina na caracterização física do Projeto realizado, cuja fundamentação se foi ilustrando ao longo desta dissertação. Representa o culminar de uma sucessão de etapas que, de um modo geral, se concretizou em: pesquisa e revisão de

literatura sobre diferentes SGLA disponíveis; conceptualização da arquitetura do sistema aplicativo; definição de um modelo de dados adequado; instalação e configuração do *software*, das suas dependências e dos seus requisitos de modo a garantir a sua operacionalidade; configurações várias do sistema; construção de um pequeno repositório de dados amostra; e, realização de pequenos testes de funcionalidade.

Realça-se que se considera que a opção por um SO Linux, constitui um dado grau de inovação, em termos de Projeto de SIG. Atualmente, é comum observar-se o desenvolvimento de projetos de SGLA assentes em SO proprietários, ou com recurso a máquinas virtuais. Se por um lado esta abordagem é muito encorajadora para estudar diferentes sistemas ou para quem pretende passar por uma fase transição no que respeita a passar a utilizar SGLA ao invés de *software* proprietário, por outro lado, em projetos de grande amplitude, pode-se tornar limitante.

Por outro lado, a escolha de um SGBDG como repositório de dados para o GeoServer com recurso a SGLA, tal como efetuado no Projeto, é possível graças ao poderoso binómio PostgreSQL/PostGIS. Considerando que o SIG arquitetado revela algum grau de complexidade, é de notar que as capacidades deste SGBDG são mais vastas do que as aqui experimentadas. Sobretudo graças às funções avançadas de gestão de vistas permitidas pelo PostgreSQL, bem como às capacidades de análise espacial do PostGIS. Ainda ao nível da arquitetura do sistema, a utilização do OpenLayers apresenta-se como promissora, quer pela simplicidade com que se conseguem os *mashups*, quer pela relativa facilidade de integração no sistema aplicativo, desde que haja suficientes conhecimentos de programação em JavaScript. Combinado com as restantes tecnologias *Web* implementadas, permite a disponibilização de uma autêntica *janela geográfica* de Biodiversidade no *browser* (de comunicação bidirecional de dados). Contudo, apesar de a edição de dados via *browser* ser possível, é necessário desenvolver mais estas potencialidades, eventualmente aguardando por uma maior maturação desta tecnologia.

Evidencia-se que, com a utilização de SGLA e com recurso a algumas linguagens de programação (Javascript, Linux shell *scripting*, PHP e SQL), é possível construir um sistema aplicativo de SIGWeb de suporte a um mapa aberto aplicado à Biodiversidade, estável, de baixo custo e ainda com alguma margem de desenvolvimento e melhorias. Permitiu dar resposta aos requisitos de produção de IGV, nomeadamente: diferentes fontes de *input* de IG, capacidades de edição multi-utilizador simultânea e *tracking* de registo das operações, entre outros. Mais relevo ganha este aspeto, se adicionarmos as potencialidades de pós-processamento da IGV implementadas. Estas melhoram significativamente a utilidade da informação das contribuições dos diferentes colaboradores. Dado que estas ferramentas de SIGWeb têm vindo a alcançar cada vez mais uma vasta gama de utilizadores, os seus resultados permitem também ser úteis e alcançar uma mais vasta comunidade.

A utilização de diferentes SGLA num sistema aplicacional funcional, ilustram o seu grau de maturação e sofisticação, bem como o potencial que eventualmente podem representar quando aplicados a outros sistemas aplicacionais de SIGWeb. O facto de estarem acessíveis na *Internet* (muitas vezes em repositórios de aplicações) e o empenho das suas comunidades no suporte aos utilizadores e na partilha de conhecimento, transcende largamente eventuais lacunas na documentação de suporte, tal como foi possível sentir no desenvolvimento do sistema aplicacional. Muitas vezes essas comunidades não só fornecem suporte à resolução de problemas, como também fornecem diferentes abordagens e perspectivas de análise o que, em última análise, pode potenciar a utilização de SGLA.

As quatro ferramentas implementadas para produção de IGV por parte dos colaboradores (*Browser editor*, *Exif georeader*, *GPX georeader* e *QGIS editor*) garantem uma cobertura de parte significativa das possibilidades de edição/*input* de informação num SIG colaborativo. Em muitos casos, dadas as interfaces amigáveis, considera-se que esta inferência é válida para colaboradores com pouca experiência ao nível de SIG, porém será sempre necessário algum grau de à vontade com sistemas informáticos (nomeadamente *browser* e plataformas *Web*). Também a camada de IGU com as três soluções implementadas (mapa de ficha de espécie, mapa de ficha de espécie: modo de edição e mapa global BiodOM), se afigura como adequada, dada a possibilidade de integração completa com o portal Naturdata.

Por último, importa realçar que a breve aplicação da prova de conceito, com os devidos resultados obtidos considerou-se, globalmente, bastante satisfatória. Porém, julga-se necessário uma maior janela temporal de testes com um mais vasto grupo de colaboradores, para poder realizar uma inferência mais assertiva.

7 CONSIDERAÇÕES FINAIS

“Death is the only pure, beautiful conclusion of a great passion. Therefore until then, no great passion is really ever finished” - David Lawrence

Nos anos recentes a *Internet* tem-se vindo a tornar num vasto meio de disponibilização e transação de dados espaciais (Zhang e Tsou 2009). Os SIG e as tecnologias *Web* têm vindo a fornecer gradualmente a oportunidade de disponibilizar dados e efetuar a sua análise, apresentando-os de um modo que se torna acessível a um vasto público. O BiodOM é um sistema aplicacional *Web* que integra IG sobre a distribuição das espécies produzida colaborativamente e voluntariamente, num esforço para compreender o conhecimento global sobre a Biodiversidade, providenciando o acesso à distribuição de espécies e a mapas dessa mesma distribuição. Cria-se assim uma plataforma de base de conhecimento relativa à distribuição de espécies e respetivos mapas, juntamente com um conjunto de ferramentas para interrogar, aceder, descarregar e visualizar esses dados geoespaciais.

Em termos de SIG e Biodiversidade têm-se vindo a observar diferentes iniciativas associadas à sua caracterização geoespacial e produção de mapas. Maioritariamente de carácter local, consistem normalmente na modelação de variáveis bio-climáticas desse local e na aplicação de modelos preditivos. A abordagem BiodOM como mapa aberto, contrasta com estas abordagens clássicas. Deve parte da sua inspiração a outras iniciativas como o OSM e a Wikipedia e pretende ser uma nova abordagem à temática da distribuição da Biodiversidade. Apresenta funcionalidades de edição de IG facilitadas, de modo colaborativo, mas neste caso desenvolvidas especificamente para o sistema. Os colaboradores registados, podem carregar pontos GPS de distribuição e editar a informação vetorial numa interface *Web* ou com recurso à ferramenta QGIS. Existe igualmente um registo do histórico de alterações efetuadas por parte dos colaboradores, a incorporação de um vetor de vaguidade (relacionado diretamente com a incerteza em IGV) e a implementação do *software* PostGIS, uma BDG robusta e de alto desempenho que confere ao sistema a possibilidade de ligação de vários utilizadores em simultâneo. Providencia-se ao BiodOM as características próprias de um SIG colaborativo. O sistema aplicacional, pelo facto de agregar informação e dados sobre distribuição de espécies, bem como servir de plataforma à produção de IGV também relativa a distribuição de espécies, apresenta-se como um modelo de base de integração, constituindo-se desse modo como um sistema que permite aos seus utilizadores a construção e potenciação do conhecimento.

Entra-se assim num novo domínio, a Neogeografia. Utilizadores comuns passam a contribuir para adquirir e partilhar IGV, que dificilmente poderia ser recolhida e organizada

no passado, ou que, devido aos seus elevados custos durante muitos anos foi considerada uma área reservada a agências oficiais e grandes organizações. Abrem-se novas perspectivas de produção de informação para novas áreas do conhecimento, em que comunidades intervêm simultaneamente como produtores e consumidores de informação geoespacial, estruturando a participação humana em processos de apoio à tomada de decisão. No entanto, é legítimo questionar a aplicabilidade dessa informação, para além a utilização por parte das comunidades envolvidas. Diferentes autores demonstram que pelo menos nas etapas iniciais do método científico, tem potencial de utilização. Essa utilização é potenciada e amplificada quando se trata de dados geoespaciais abertos. Acresce que, uma eventual tentativa de caracterizar a distribuição de Biodiversidade, requer não só dados massivos mas também, numa abordagem clássica, tende esta a envolver um esforço humano considerável ao nível dos recursos exigidos. Os SIG colaborativos e a IGV alinham-se para dar resposta a esses requisitos.

A visão integrada e transversal entre Biodiversidade, CIG e SGLA, durante a fase de conceptualização revelou-se como um fator preponderante para definir os requisitos e a arquitetura do BiodOM. Uma visão integradora destes conceitos não só se torna viável, como constitui um imperativo no caminho do vasto domínio que é a abordagem ao conhecimento da distribuição da Biodiversidade. Assim, para o BiodOM, pretende-se que a sua missão seja permitir à comunidade a amplificação do conhecimento sobre Biodiversidade, tendo em vista a sua preservação.

Em termos de melhorias futuras, um sistema de informação, tradicionalmente, pode sempre incorporar novas funcionalidades de modo a ir amplificando a sua experiência de utilização gradualmente. Apesar da conceptualização do BiodOM ser recente há já linhas de desenvolvimento futuro que se perspectivam, nomeadamente: a atualização para as versões mais recentes dos pacotes de *software*; ferramentas de exportação de dados mais úteis e intuitivas; melhorias ao nível da cartografia temática de base; novas funcionalidades ao nível do pós-processamento apoiadas em revisão bibliográfica mais ampla de modo a extrair o máximo de conhecimento da IGV produzida; desenvolvimento de uma aplicação para smartphones iOS e/ou Android; entre outras.

Por outro lado à semelhança da implementação de outros modelos de dados por diferentes autores, aceita-se que o modelo aqui apresentado também pode ser suscetível de melhorias futuras. Nomeadamente atendendo às seguintes aproximações hipotéticas: sendo a IGV um domínio relativamente recente, é possível ao longo do tempo, o desenvolvimento de metodologias comprovadamente mais adequadas para a incorporação de incerteza no modelo de dados; ao longo do tempo, esperando-se uma crescente colaboração por parte de diferentes colaboradores e voluntários de IGV, a tipologia de dados poderá ser melhorada em função da quantidade de informação e da própria disponibilidade de

introdução de diferente informação por parte dos colaboradores que à partida do projeto não se esperaria; a cartografia temática de base poderá evoluir, podendo-se adotar novas fontes, sobretudo atendendo a que dos dados abertos começam agora a sentir uma grande massificação.

Em termos de disseminação dos resultados da pesquisa, após uma fase de maturação e utilização do sistema, pretende-se participar no mínimo em duas conferências de especialidade. Não só para efeitos de apresentar e divulgar o produto mas, sobretudo para promover a discussão e recolher *feedback* para melhorias futuras. É também objetivo a publicação de artigo científico para submissão a uma revista da especialidade. Apesar disso, dois principais desafios continuarão a permanecer. O primeiro desafio é continuar a atrair uma ampla gama de voluntários, como p.e.: estudantes universitários ou cidadãos comuns que gostam de se envolver em atividades de Biodiversidade. Através da divulgação de material educativo e ferramentas fora da comunidade científica, pode ser demonstrado como uma comunidade pode se envolver em exercícios de validação e distribuição de espécies. Em termos de eventuais novas instalações de baixo custo de divulgação, uma opção seria a utilização de redes sociais e com os grupos existentes que foram criados. Em particular aqueles que incluem as pessoas que têm algum tipo de experiência em geografia e cartografia. O segundo desafio é ser capaz de garantir a qualidade e integridade dos dados, para além de se garantir que a ferramenta aplicacional é bem utilizada. Como discutido por vários autores, a questão da credibilidade dessas contribuições voluntárias públicas é crucial.

Um último e importante aspeto é o facto de o sistema aplicacional poder ser utilizado de modo diferente para diferentes áreas de conhecimento, além daquelas propostas neste trabalho. Com as devidas alterações e adaptações perspectiva-se também para poder ser replicado para outros projetos semelhantes noutras regiões do globo. Aqui a disseminação dos resultados irá desempenhar um papel vital. Quem sabe, o sistema aplicacional possa vir a ser parte integrante de uma rede global.

REFERÊNCIAS BIBLIOGRÁFICAS

- Apache, 2012. *The apache software foundation* [online]. Apache Software Foundation. Disponível em: <http://www.apache.org/> [Acesso em: 31-03-2012].
- Balram, S. e Dragičević, S., 2006. Collaborative geographic information systems: Origins, boundaries, and structures. *Collaborative geographic information systems*. IGI Global, 1-23.
- Barriguinha, A. e Ribeiro, P., 2008. WebGIS opensource no apoio à consultadoria e gestão agro-florestal: O "Eco@gro Digital". *Actas do X Encontro de Utilizadores de Informação Geográfica – ESIG2008*, USIG (Ed.) (TagusPark, Oeiras) 139-150.
- Berendsohn, W., Güntsch, A., Hoffmann, N., Kohlbecker, A., Luther, K. e Müller, A., 2011. Biodiversity information platforms: From standards to interoperability. *ZooKeys*, 150 (0), 71-87.
- Bernard, L., Kanellopoulos, I., Annoni, A. e Smits, P., 2005. The european geoportal: One step towards the establishment of a european spatial data infrastructure. *Computers, Environment and Urban Systems*, 29 (1), 15-31.
- Bonn, A. e Gaston, K.J., 2005. Capturing biodiversity: Selecting priority areas for conservation using different criteria. *Biodiversity and Conservation*, 14 (5), 1083-1100.
- Bowker, G.C., 2000. Mapping biodiversity. *International Journal of Geographical Information Science*, 14 (8), 739-754.
- Brunn, S.D., Cutter, S.L. e Harrington, J.W., 2004. *Geography and technology*: Springer.
- Canhos, V.P., Souza, S., Giovanni, R. e Canhos, D.A.L., 2004. Global biodiversity informatics: Setting the scene for a "new world" of ecological modeling. *Biodiversity Informatics*, 1 (1), 1-13.
- Cardillo, M., Orme, C.D.L. e Owens, I.P.F., 2005. Testing for latitudinal bias in diversification rates: An example using new world birds. *Ecology*, 86 (9), 2278-2287.
- Cavalcanti, M., 2007. *Biogeography in a changing world* [online]. Biogeografia. Disponível em: <https://sites.google.com/site/maurobio/publications/articles> [Acesso em: 31-04-2012].
- CBD, 2011. *Climate change and biodiversity* [online]. Convention on Biological Diversity. Disponível em: <http://www.cbd.int/climate/> [Acesso em: 31-05-2012].
- CBD, 2012. *Convention on biological diversity: Technology transfer and cooperation information database* [online]. United Nations Environment Programme. Disponível

- em: <http://www.cbd.int/programmes/cross-cutting/technology/search.aspx> [Acesso em: 31-04-2012].
- EC, 2012a. *European Comission nature e biodiversity* [online]. European Comission. Disponível em: http://ec.europa.eu/environment/nature/index_en.htm [Acesso em: 31-04-2012].
- EC, 2012b. *Natura 2000 network* [online]. European Comission. Disponível em: http://ec.europa.eu/environment/nature/natura2000/index_en.htm [Acesso em: 31-04-2012].
- Chivian, E. e Bernstein, A., 2008. *Sustaining life: How human health depends on biodiversity*. Oxford University Press, USA.
- CLO, 2008. *About citizen science* [online]. The Cornell Lab of Ornithology. Disponível em: <http://www.birds.cornell.edu/citsci/about> [Acesso em: 31-05-2012].
- Coftas, L. e Diosteanu, A., 2010. Evaluating accessibility in crowdsourcing GIS. *The Journal of Applied Collaborative Systems* 2(1), 45-49.
- Cohn, A.G. e Gotts, N.M., 1996. Representing spatial vagueness: A mereological approach. In Aiello, L.C., Doyle, J. e Shapiro, S.C. eds. *KR*. Morgan Kaufmann, 230-241.
- COL, 2012. *Catalogue of life* [online]. Catalogue Of Life. Disponível em: <http://www.catalogueoflife.org/content/about#usage> [Acesso em: 31-06-2012].
- Costanza, R., D'Arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R.V., Paruelo, J., Raskin, R.G., Sutton, P. e Van Den Belt, M., 1997. The value of the world's ecosystem services and natural capital. *Ecological Economics*, 25 (1), 3-15.
- Council Directive 79/409/EEC (on the conservation of wild birds), de 2 Abril 1979.
- Council directive 92/43/EEC (on the conservation of natural habitats and of wild fauna and flora), de 21 de Maio de 1992.
- Currie, D.J., Mittelbach, G.G., Cornell, H.V., Field, R., Guégan, J.F., Hawkins, B.A., Kaufman, D.M., Kerr, J.T., Oberdorff, T., O'Brien, E. e Turner, J.R.G., 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters*, 7 (12), 1121-1134.
- Dangermond, J., 2008. *GIS and the geoweb* [online]. ESRI. Disponível em: <http://www.esri.com/news/arcnews/summer08articles/gis-and-geoweb.html> [Acesso em: 31-08-2012].
- Davies, T.J. e Cadotte, M., 2011. Quantifying biodiversity: Does it matter what we measure?

- In Zachos, F.E. e Habel, J.C. eds. *Biodiversity hotspots*. Springer Berlin Heidelberg, 43-60.
- Davis, M. e Cornwell, D., 2012. *Introduction to environmental engineering*: McGraw-Hill Companies, Incorporated.
- De Longueville, B., 2010. Community-based geoportals: The next generation? Concepts and methods for the geospatial web 2.0. *Computers, Environment and Urban Systems*, 34 (4), 299-308.
- De Longueville, B., Ostländer, N. e Keskitalo, C., 2009. Addressing vagueness in volunteered geographic information (VGI). *International Journal of Spatial Data Infrastructures*.
- De Queiroz, K., 2005. Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences of the United States of America*, 102 (Supl 1), 6600-6607.
- Debian, 2012. *About Debian* [online]. Debian. Disponível em: <http://www.debian.org/intro/about> [Acesso em: 31-03-2012].
- Dilo, A., De By, R.A. e Stein, A., 2007. A system of types and operators for handling vague spatial objects. *International Journal of Geographical Information Science*, 21 (4), 397-426.
- Dobzhansky, T.G., 1955. *Evolution, genetics, and man*: Wiley.
- Dragičević, S. e Balam, S., 2006. Collaborative geographic information systems and science: A transdisciplinary evolution. *Collaborative geographic information systems*. IGI Global, 341-350.
- Dykes, J., 2006. Progress in our representation of geographic phenomena and our evaluation, use and analysis of geographic information. *Computers, Environment and Urban Systems*, 30 (2), 125-129.
- E-Biosphere, 2009. *International conference on biodiversity informatics* [online]. e-Biosphere. Disponível em: <http://www.e-biosphere09.org/> [Acesso em: 31-07-2012].
- Edit, 2012. *European distributed institute of taxonomy (edit)* [online]. European Distributed Institute of Taxonomy (EDIT). Disponível em: <http://www.e-taxonomy.eu/> [Acesso em: 31-05-2012].
- Egenhofer, M.J., 2002. Toward the semantic geospatial web. *Proceedings of the 10th ACM international symposium on Advances in geographic information systems*. McLean, Virginia, USA: ACM, 1-4.

- Elwood, S., 2006. Critical issues in participatory GIS: Deconstructions, reconstructions, and new research directions. *Transactions in GIS*, 10 (5), 693-708.
- Elwood, S., 2008. Volunteered geographic information: Future research directions motivated by critical, participatory, and feminist GIS. *GeoJournal*, 72 (3), 173-183.
- Elwood, S., Goodchild, M.F. e Sui, D.Z., 2011. Researching volunteered geographic information: Spatial data, geographic research, and new social practice. *Annals of the Association of American Geographers*, 102 (3), 571-590.
- Erle, S. e Gibson, R., 2006. *Google Maps hacks*. Sebastopol, USA: O'Reilly Media, Inc.
- Fisher, P., 2000. Sorites paradox and vague geographies. *Fuzzy Sets and Systems*, 113 (1), 7-18.
- Fleishman, E., Noss, R.F. e Noon, B.R., 2006. Utility and limitations of species richness metrics for conservation planning. *Ecological Indicators*, 6 (3), 543-553.
- Fook, K.D., Vieira Monteiro, A.M., Camara, G., Casanova, M.A. e Amaral, S., 2009. Geoweb services for sharing modelling results in biodiversity networks. *Transactions in GIS*, 13 (4), 379-399.
- Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, R., Achard, F., Kraxner, F. e Obersteiner, M., 2009. Geo-wiki.Org: The use of crowdsourcing to improve global land cover. *Remote Sensing*, 1 (3), 345-354.
- FSF, 2004. *Free Software Foundation: About* [online]. Free Software Foundation. Disponível em: <http://www.fsf.org/about/> [Acesso em: 31-06-2012].
- Gabriel, S., 2007. *Biodiversity 'fundamental' to economics* [online]. BBC News. Disponível em: <http://news.bbc.co.uk/2/hi/science/nature/6432217.stm> [Acesso em: 31-04-2012].
- Gaines, W.L., Harrod, R.J., Lehmkuhl, J.F. e Station, P.N.R., 1999. *Monitoring biodiversity: Quantification and interpretation*. U.S. Dept. of Agriculture, Forest Service, Pacific Northwest Research Station.
- GBIF, 2012. *Global biodiversity information facility data portal* [online]. Global Biodiversity Information Facility. Disponível em: <http://data.gbif.org/welcome.htm> [Acesso em: 31-05-2012].
- Geofabrik, 2012. *Geofabrik downloads* [online]. Geofabrik. Disponível em: <http://www.geofabrik.de/data/download.html> [Acesso em: 31-09-2012].
- Geoserver, 2012. *Geoserver user manual* [online]. Geoserver. Disponível em: <http://docs.geoserver.org/stable/en/user/> [Acesso em: 31-03-2012].

- Giles, J., 2005. Internet encyclopaedias go head to head. *Nature*, 438 (7070), 900-901.
- Goodchild, M., 2007. Citizens as sensors: The world of volunteered geography. *GeoJournal*, 69 (4), 211-221.
- Goodchild, M., 2008. Commentary: Whither VGI? *GeoJournal*, 72 (3), 239-244.
- Goodchild, M., 2009. Neogeography and the nature of geographic expertise. *Journal of Location Based Services*, 3 (2), 82-96.
- Goodchild, M.F., 1992. Geographical information science. *International Journal of Geographical Information Systems*, 6 (1), 31-45.
- Google, 2012. *Google's complete map content specifications* [online]. Google. Disponível em: <http://support.google.com/mapcontentpartners/bin/answer.py?hl=en&answer=144284> [Acesso em: 31-09-2012].
- Gotelli, N.J., 2008. *A primer of ecology*: Sinauer Associates.
- Gotelli, N.J. e Colwell, R.K., 2001. Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, 4 (4), 379-391.
- Graham, M., 2009. Neogeography and the palimpsests of place: Web 2.0 and the construction of a virtual earth. *Tijdschrift voor economische en sociale geografie*, 101 (4), 422-436.
- Gray, J., 2010. *Open geoprocessing standards and open geospatial data* [online]. Open Knowledge Foundation. Disponível em: <http://blog.okfn.org/2010/06/21/open-geoprocessing-standards-and-open-geospatial-data/> [Acesso em: 31-05-2012].
- Guralnick, R.P., Hill, A.W. e Lane, M., 2007. Towards a collaborative, global infrastructure for biodiversity assessment. *Ecology Letters*, 10 (8), 663-672.
- Hall, G.B., Chipeniuk, R., Feick, R.D., Leahy, M.G. e Deparday, V., 2010. Community-based production of geographic information using open source software and web 2.0. *International Journal of Geographical Information Science*, 24 (5), 761-781.
- Hall, G.B. e Leahy, M.G., 2008. Design and implementation of a map-centred synchronous collaboration tool using open source components: The mapchat project. In Hall, G.B. e Leahy, M.G. eds. *Open source approaches in spatial data handling*. Springer Berlin Heidelberg, 221-245.
- Hanson, C., Ranganathan, J., Iceland, C. e Finisdore, J., 2012. *The corporate ecosystem services review: Guidelines for identifying business risks e opportunities arising from ecosystem change* [online]. World Resources Institute. Disponível em:

- <http://www.wri.org/publication/corporate-ecosystem-services-review> [Acesso em: 30-05-2012].
- Hart, G. e Dolbear, C., 2007. What's so special about spatial? In Scharl, A. e Tochtermann, K. eds. *The geospatial web: How geobrowsers, social software and the web 2.0 are shaping the network society*. Springer London, 39-44.
- Hawksworth, D.L., 1996. *Biodiversity: Measurement and estimation*. London: Chapman and Hall. The Royal society.
- Hazzard, E., 2011. *Openlayers 2.10 beginner's guide*: Packt Publishing, Limited.
- Hillebrand, H., 2004. On the generality of the latitudinal diversity gradient. *American Naturalist*, 163, 192-211.
- Hudson-Smith, A., Crooks, A., Gibin, M., Milton, R. e Batty, M., 2009. Neogeography and Web 2.0: Concepts, tools and applications. *Journal of Location Based Services*, 3 (2), 118-145.
- ICN, 2006. *Plano sectorial da Rede Natura 2000 - relatório* [online]. Instituto de Conservação da Natureza. Disponível em: http://www.icn.pt/psrn2000/conteudo_plano.htm [Acesso em: 31-04-2012].
- ICNB, 2005. *Instituto de conservação da natureza e biodiversidade: Rede Natura 2000* [online]. Instituto de Conservação da Natureza e Biodiversidade. Disponível em: <http://portal.icnb.pt/ICNPortal/vPT2007/O+ICNB/Rede+Natura+2000+2010/> [Acesso em: 31-06-2012].
- IUCN, 2012a. *The IUCN red list of threatened species (summary statistics)* [online]. International Union for Conservation of Nature. Disponível em: <http://www.iucnredlist.org/about/summary-statistics> [Acesso em: 31-05-2012].
- IUCN, 2012b. *Red list of threatened species* [online]. International Union for Conservation of Nature. Disponível em: <http://www.iucnredlist.org/> [Acesso em: 31-04-2012].
- Jackson, J., 2006. *"Neogeography" blends blogs with online maps* [online]. National Geographic News. Disponível em: http://news.nationalgeographic.com/news/2006/04/0425_060425_map_blogs.html [Acesso em: 31-07-2012].
- James, A., Gaston, K.J. e Balmford, A., 2001. Can we afford to conserve biodiversity? *BioScience*, 51 (1), 43-52.
- Jankowski, P. e Nyerges, T., 2001. GIS-supported collaborative decision making: Results of an experiment. *Annals of the Association of American Geographers*, 91 (1), 48-70.

- Judd, W.S., Campbell, C.S., Kellog, E.A., Stevens, P.F. e Donogue, M.J., 2008. *Plant systematics: A phylogenetic approach*. Sunderland: Sinauer Associates, Inc.
- Kirk, P.M., Cannon, P.F., Minter, D.W. e Stalpers, J.A., 2008. *Dictionary of the fungi*. Netherlands: CABI.
- Kishor, P. e Ventura, S., Year. What can GIS learn from floss? *9th International Conference of the Global Spatial Data Infrastructure (GSDI-9-2006)*, Santiago, Chile.
- Knapp, S., 2010. *What's in a name? A history of taxonomy* [online]. Natural History Museum. Disponível em: <http://www.nhm.ac.uk/nature-online/science-of-natural-history/taxonomy-systematics/history-taxonomy/index.html> [Acesso em: 31-06-2012].
- Kralidis, A.T., 2008. Geospatial open source and open standards convergences. In Hall, G.B. e Leahy, M.G. eds. *Open source approaches in spatial data handling*. Springer Berlin Heidelberg, 1-20.
- Larsson, T.B., 2001. *Ecological bulletins, biodiversity evaluation tools for european forests*: Wiley.
- Lawrence, A., 2006. 'No personal motive?' volunteers, biodiversity, and the false dichotomies of participation. *Ethics, Place e Environment*, 9 (3), 279-298.
- Longley, P.A., Goodchild, M.F., Maguire, D.J. e Rhind, D.W., 2005. *Geographic information systems and science*: John Wiley and Sons.
- Loreau, M., Oteng-Yeboah, A., Arroyo, M.T.K., Babin, D., Barbault, R., Donoghue, M., Gadgil, M., Häuser, C., Heip, C., Larigauderie, A., Ma, K., Mace, G., Mooney, H.A., Perrings, C., Raven, P., Sarukhan, J., Schei, P., Scholes, R.J. e Watson, R.T., 2006. Diversity without representation. *Nature*, 442 (7100), 245-246.
- Lumrix, 2009. *Rice grassy stunt virus* [online]. Lumrix.net. Disponível em: http://www.lumrix.net/health/Rice_grassy_stunt_virus.html [Acesso em: 31-04-2012].
- Mace, G.M., Balmford, A., Boitani, L., Cowlshaw, G., Dobson, A.P., Faith, D.P., Gaston, K.J., Humphries, C.J., Vane-Wright, R.I., Williams, P.H., Lawton, J.H., Margules, C.R., May, R.M., Nicholls, A.O., Possingham, H.P., Rahbek, C. e Van Jaarsveld, A.S., 2000. It's time to work together and stop duplicating conservation efforts. *Nature*, 405 (6785), 393-393.
- Machado, J., Cabral, P. e Paíño, M., 2002. Aplicações de SIG na web – o atlas do ambiente dinâmico. *VII Encontro de Utilizadores de Informação Geográfica, ESIG 2002* (USIG – TagusPark).

- Maguire, D.J. e Longley, P.A., 2005. The emergence of geoportals and their role in spatial data infrastructures. *Computers, Environment and Urban Systems*, 29 (1), 3-14.
- Margules, C.R. e Pressey, R.L., 2000. Systematic conservation planning. *Nature*, 405 (6783), 243-253.
- Margulis, L. e Sagan, D., 2003. *Acquiring genomes: A theory of the origin of species*: Basic Books.
- Mark, D.M. e Frank, A.U., 1996. Experiential and formal models of geographic space. *Environment and Planning B: Planning and Design*, 23 (1), 3-24.
- Martins, H., 2010. *Sigla WebGIS portal: An open source web-based GIS application for the portuguese bluetongue entomological surveillance program*. Dissertação para obtenção do grau de Mestre em C&SIG. Instituto Superior de Estatística e Gestão da Informação da Universidade Nova de Lisboa.
- Mason, B.C. e Dragicevic, S., 2006. WebGIS and knowledge management systems: An integrated design for collaborative community planning. *Collaborative geographic information systems*. IGI Global, 263-284.
- Mayr, E., 1963. *Animal species and evolution*: Belknap Press of Harvard University Press.
- Mccabe, D.J. e Gotelli, N.J., 2000. Effects of disturbance frequency, intensity, and area on assemblages of stream macroinvertebrates. *Oecologia*, 124 (2), 270-279.
- McIhagga, D., 2008. Communities of practice and the business of open source web mapping. In Hall, G.B. e Leahy, M. eds. *Open source approaches in spatial data handling*. Springer Berlin Heidelberg, 49-64.
- Mills, E., 2005. *Mapping a revolution with 'mashups'* [online]. CNET. Disponível em: http://news.cnet.com/Mapping-a-revolution-with-mashups/2009-1025_3-5944608.html [Acesso em: 31-04-2012].
- Miranda, J.I., 2004. Springweb: Um aplicativo para publicar mapas na web. *Embrapa Informática Agropecuária*, (41).
- Moulder, C., 2009. Google Earth meets higher ed: Reflections on neogeography. *Bulletin [Association of Canadian Map Libraries and Archives]*, 134 (134), 10-15.
- Myers, N., 1990. The biodiversity challenge: Expanded hot-spots analysis. *The Environmentalist*, 10 (4), 243-256.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., Da Fonseca, G.a.B. e Kent, J., 2000. Biodiversity hotspots for conservation priorities. *Nature*, 403 (6772), 853-858.

- Nangendo, G., Stein, A., Gelens, M., De Gier, A. e Albricht, R., 2002. Quantifying differences in biodiversity between a tropical forest area and a grassland area subject to traditional burning. *Forest Ecology and Management*, 164 (1–3), 109-120.
- Naturdata, 2012. *Naturdata - biodiversidade online* [online]. Prodata. Disponível em: <http://naturdata.com/o-que-e-o-naturdata> [Acesso em: 31-03-2012].
- Normile, D., 2010. Saving forests to save biodiversity. *Science*, 329 (5997), 1278-1280.
- Noss, R.F., 1990. Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology*, 4 (4), 355-364.
- Obe, R.O. e Hsu, L.S., 2011. *PostGIS in action*: Manning Publications.
- OGC, 2008. *Open geospatial consortium* [online]. Open Geospatial Consortium. Disponível em: <http://www.opengeospatial.org/> [Acesso em: 31-05-2012].
- OpenGeo, 2012. *Opengeo case studies* [online]. OpenGeo. Disponível em: <http://opengeo.org/publications/casestudies/> [Acesso em: 31-05-2012].
- OpenLayers, 2012. *Openlayers: Free maps for the web* [online]. OpenLayers. Disponível em: <http://www.openlayers.org/> [Acesso em: 31-04-2012].
- OpenStreetMap, 2012. *Openstreetmap documentation* [online]. OpenStreetMap. Disponível em: http://wiki.openstreetmap.org/wiki/Main_Page [Acesso em: 31-05-2012].
- OSGeo, 2012a. *About the open source geospatial foundation* [online]. Open Source Geospatial Foundation. Disponível em: <http://www.osgeo.org/content/foundation/about.html> [Acesso em: 31-06-2012].
- OSGeo, 2012b. *FOSS4G* [online]. OSGeo. Disponível em: <http://wiki.osgeo.org/wiki/FOSS4G> [Acesso em: 31-09-2012].
- OSI, 2012. *The open source definition (annotated)* [online]. Open Source Initiative. Disponível em: <http://opensource.org/osd.html> [Acesso em: 31-07-2012].
- PHP, 2012. *PHP manual* [online]. PHP Documentation Group. Disponível em: <http://www.php.net/manual/en/> [Acesso em: 31-09-2012].
- Pierce, M.E., Fox, G.C., Choi, J.Y., Guo, Z., Gao, X. e Ma, Y., 2009. Using web 2.0 for scientific applications and scientific communities. *Concurrency and Computation: Practice and Experience*, 21 (5), 583-603.
- Pinheiro, J., 2010. *Prova de conceito (PdC) no projeto de redes de computadores* [online]. Desmonta e CIA. Disponível em: <http://desmontacia.wordpress.com/2010/12/21/prova-de-conceito-poc-no-projeto->

- de-redes-de-computadores/ [Acesso em: 31-09-2012].
- PostGIS, 2012. *PostGIS 1.5.5 manual* [online]. PostGIS. Disponível em:
<http://postgis.refractory.net/documentation/manual-1.5/>
- PostgreSQL, 2012. *Postgresql 8.4.14 documentation* [online]. PostgreSQL Global Development Group. Disponível em:
<http://www.postgresql.org/docs/8.4/interactive/index.html> [Acesso em: 28-04-2012].
- Ramamurthy, M.K., 2006. A new generation of cyberinfrastructure and data services for earth system science education and research. *Adv. Geosci.*, 8, 69-78.
- Ramanujan, K., 2010. *Study: Loss of species is bad for your health* [online]. Cornell University. Disponível em:
<http://www.news.cornell.edu/stories/Dec10/BiodiversityHealth.html> [Acesso em: 31-03-2012].
- Rana, S. e Joliveau, T., 2009. Neogeography: An extension of mainstream geography for everyone made by everyone? *Journal of Location Based Services*, 3 (2), 75-81.
- Raper, J. e Livingstone, D., 1995. Development of a geomorphological spatial model using object-oriented design. *International Journal of Geographical Information Systems*, 9 (4), 359-383.
- Recabarren, M., Nussbaum, M. e Leiva, C., 2008. Cultural divide and the internet. *Computers in Human Behavior*, 24 (6), 2917-2926.
- Resolução do Conselho de Ministros n.º 115-a/2008, de 21 de Julho.
- Richard, K., 2002. Web-based PPGIS in the united kingdom. *Community participation and geographical information systems*. CRC Press.
- Rocha, J.G., 2012, Apontamentos da Unidade Curricular de Geospatial Free and Open Source Software. (11ª Edição do Mestrado em Ciência e Sistemas de Informação Geográfica: ISEGI- UNL).
- Rouse, L.J., Bergeron, S.J. e Harris, T.M., 2007. Participating in the geospatial Web: Collaborative mapping, social networks and participatory GIS. In Scharl, A. e Tochtermann, K. eds. *The geospatial web: How geobrowsers, social software and the Web 2.0 are shaping the network society*. Springer London, 153-158.
- Sala, O.E., Stuart Chapin, F., Iii, Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Hueneke, L.F., Jackson, R.B., Kinzig, A., Leemans, R., Lodge, D.M., Mooney, H.A., Oesterheld, M.N., Poff, N.L., Sykes, M.T., Walker, B.H., Walker, M. e Wall, D.H., 2000. Global biodiversity scenarios for the year 2100. *Science*, 287 (5459), 1770-1774.

- Shamal, M., 2012. *Comparative study of oracle spatial and postgres spatial*. Dissertação para obtenção do grau de Mestre em SIG. San Diego State University.
- Shannon, C.E., 1948. A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379-423, 623-656.
- Sheppard, E. e McMaster, R.B., 2004. *Scale and geographic inquiry: Nature, society, and method*: John Wiley e Sons.
- Sieber, R., 2006. Public participation geographic information systems: A literature review and framework. *Annals of the Association of American Geographers*, 96 (3), 491-507.
- Simpson, E.H., 1949. Measurement of diversity. *Nature*, 163, 688.
- Spellerberg, I.F. e Fedor, P.J., 2003. A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the ‘Shannon–Wiener’ index. *Global Ecology and Biogeography*, 12 (3), 177-179.
- Steiniger, S. e Bocher, E., 2009. An overview on current free and open source desktop GIS developments. *International Journal of Geographical Information Science*, 23 (10), 1345-1370.
- Tapscott, D. e Williams, A.D., 2006. *Wikinomics: How mass collaboration changes everything*: Portfolio.
- TDWG, 2007. *TDWG standards* [online]. Biodiversity Information Standards (TDWG). Disponível em: <http://www.tdwg.org/standards/> [Acesso em: 31-05-2012].
- Toffler, A., 1989. *The third wave*: Bantam Books.
- Tuomisto, H., 2010. A consistent terminology for quantifying species diversity? Yes, it does exist. *Oecologia*, 164 (4), 853-860.
- Turner, A., 2006. *Introduction to neogeography*: O'Reilly Media.
- UN, 2011. *UN launches decade on biodiversity to stem loss of ecosystems* [online]. UN News Centre. Disponível em: <http://www.un.org/apps/news/story.asp?NewsID=40766#.UHX2NRXIX8A> [Acesso em: 31-03-2012].
- V., S., M., O., T., V. e H., A., 2005. Portal vertical geosolos – spring web na divulgação dos resultados de pesquisa em regiões cafeeiras de minas gerais. *XII Simpósio Brasileiro de Sensoriamento Remoto Goiânia*, (Brasil: INPE).
- Vossen, G. e Hagemann, S., 2007. Unleashing Web 2.0: From concepts to creativity. *Ubiquity*, 2007 (December), 1-1.
- Vrba, V. e Wotawa, G., 2011. *How Fukushima's radioactive cloud influenced social networks*

- [online]. Social Intensity. Disponível em: <http://www.socialintensity.org/#fukushima> [Acesso em: 31-05-2012].
- W3C, 2012. *W3C standards* [online]. World Wide Web Consortium (W3C). Disponível em: <http://www.w3.org/standards/> [Acesso em: 31-07-2012].
- Wade, T. e Sommer, S., 2006. *A to z GIS: An illustrated dictionary of geographic information systems*: ESRI Press.
- Warf, B., 2001. Segueways into cyberspace: Multiple geographies of the digital divide. *Environment and Planning B: Planning and Design*, 28, 3-19.
- Whittaker, R.J. e Fernandez-Palacios, J.M., 2007. *Island biogeography: Ecology, evolution, and conservation*: Oxford University Press, USA.
- Wieczorek, J., Döring, M., Giovanni, R.D., Robertson, T. e Vieglaes, D., 2009. *Darwin core* [online]. Biodiversity Information Standards (TDWG). Disponível em: <http://rs.tdwg.org/dwc/> [Acesso em: 31-09-2012].
- Wikibooks, 2012. *Converting mysql to postgresql* [online]. Wikibooks. Disponível em: http://en.wikibooks.org/wiki/Converting_MySQL_to_PostgreSQL [Acesso em: 31-09-2012].
- Wikipedia, 2012. *Wikipedia: About* [online]. Wikimedia. Disponível em: <http://en.wikipedia.org/wiki/Wikipedia:About> [Acesso em: 31-06-2012].
- Wilson, D.E., Cole, R.F., Nichols, J.D., Rudran, R. e Foster, M.S., 1996. *Measuring and monitoring biological diversity: Standard methods for mammals*: Smithsonian Institution Press.
- Wilson, E.O., 1991. The current state of biological diversity. In Willers, W.B. ed. *Learning to listen to the land*. Island Press, 295.
- Wilson, E.O., 2002. *The future of life*: Vintage Books.
- WRI, IUCN, Earthwatch e WBCSD, 2006. *Business and ecosystems issue brief: Ecosystem challenges and business implications* [online]. World Resources Institute. Disponível em: <http://www.wri.org/publication/business-and-ecosystems-issue-brief-ecosystem-challenges-and-business-implications> [Acesso em: 31-03-2012].
- Yang, P., Evans, J., Cole, M., Marley, S., Alameh, N. e Bambacus, M., 2007. The emerging concepts and applications of the spatial web portal. *American Society for Photogrammetry and Remote Sensing* 73 (6), 691-698.
- Zeiss, G., 2008. *Open geospatial data* [online]. Between The Poles Blog. Disponível em: <http://geospatial.blogs.com/geospatial/2008/11/open-geospatial-data.html> [Acesso

em: 31-05-2012].

Zhang, J. e Goodchild, M.F., 2002. *Uncertainty in geographical information*: Taylor e Francis.

Zhang, T. e Tsou, M.H., 2009. Developing a grid-enabled spatial web portal for internet giservices and geospatial cyberinfrastructure. *International Journal of Geographical Information Science*, 23 (5), 605-630.

Zook, M., Dodge, M., Aoyama, Y. e Townsend, A., 2004. New digital geographies: Information, communication, and place. *Geography and technology*, 155-176.

Zook, M., Graham, M., Shelton, T. e Gorman, S., 2010. Volunteered geographic information and crowdsourcing disaster relief: A case study of the haitian earthquake. *World Medical e Health Policy*, 2 (2), 7-33.

Anexo A

Código envolvido na estilização WFS-T com recurso ao OpenLayers

```

var map, wfs;
OpenLayers.ProxyHost = "proxy.cgi?url=";

var DeleteFeature = OpenLayers.Class(OpenLayers.Control, {
    initialize: function(layer, options) {
        OpenLayers.Control.prototype.initialize.apply(this,
[options]);
        this.layer = layer;
        this.handler = new OpenLayers.Handler.Feature(
            this, layer, {click: this.clickFeature}
        );
    },
    clickFeature: function(feature) {
        // se o elemento nao tiver uma fid, destrui-lo
        if(feature.fid == undefined) {
            this.layer.destroyFeatures([feature]);
        } else {
            feature.state = OpenLayers.State.DELETE;
            this.layer.events.triggerEvent("afterfeaturemodified",
                {feature: feature});
            feature.renderIntent = "select";
            this.layer.drawFeature(feature);
        }
    },
    setMap: function(map) {
        this.handler.setMap(map);
        OpenLayers.Control.prototype.setMap.apply(this, arguments);
    },
    CLASS_NAME: "OpenLayers.Control.DeleteFeature"
});

function init() {
    var extent = new OpenLayers.Bounds(
        -128*156543.0339, -128*156543.0339, 128*156543.0339,
128*156543.0339
    );

    map = new OpenLayers.Map('map', {
        projection: new OpenLayers.Projection("EPSG:900913"),
        displayProjection: new OpenLayers.Projection("EPSG:4326"),
        restrictedExtent: extent,
        controls: [
            new OpenLayers.Control.Navigation({}),
            new OpenLayers.Control.ScaleLine({}),
            new OpenLayers.Control.ScaleLine({}),
            new OpenLayers.Control.PanZoom({}),
            new OpenLayers.Control.Attribution({}),
            new OpenLayers.Control.MousePosition({})
        ]
    });

    var gorto = new OpenLayers.Layer.Google(
        "Ortofotografia Google",
        {type: google.maps.MapTypeId.HYBRID, sphericalMercator:
true}
    );

```

```

var saveStrategy = new OpenLayers.Strategy.Save();

wfs = new OpenLayers.Layer.Vector("Editable Features", {
  strategies: [new OpenLayers.Strategy.BBOX(), saveStrategy],
  projection: new OpenLayers.Projection("EPSG:4326"),
  protocol: new OpenLayers.Protocol.WFS({
    version: "1.1.0",
    srsName: "EPSG:4326",
    url: "http://biodom.dynip.sapo.pt/geoserver/wfs",
    featureNS : "http://biodom.dynip.sapo.pt",
    featureType: "restricted",
    geometryName: "the_geom",
    schema:
"http://biodom.dynip.sapo.pt/geoserver/wfs/DescribeFeatureType?
version=1.1.0&typename=og:restricted"
  })
});

map.addLayers([gorto, wfs]);

var panel = new OpenLayers.Control.Panel({
  displayClass: 'customEditingToolbar',
  allowDepress: true
});

var draw = new OpenLayers.Control.DrawFeature(
  wfs, OpenLayers.Handler.Polygon,
  {
    title: "Desenhar elemento",
    displayClass: "olControlDrawFeaturePolygon",
    multi: true
  }
);

var edit = new OpenLayers.Control.ModifyFeature(wfs, {
  title: "Modificar elemento",
  displayClass: "olControlModifyFeature"
});

var del = new DeleteFeature(wfs, {title: "Delete Feature"});

var save = new OpenLayers.Control.Button({
  title: "Salvar",
  trigger: function() {
    if(edit.feature) {
      edit.selectControl.unselectAll();
    }
    saveStrategy.save();
  },
  displayClass: "olControlSaveFeatures"
});

panel.addControls([save, del, edit, draw]);
map.addControl(panel);
map.zoomToExtent(extent, true);
}

```

Anexo B

**Listagem de comandos utilizados nas diferentes instalações e configurações
do sistema aplicativo**

Instalação do sistema aplicativo BiodOM (notas pessoais retiradas ao longo da parte prática):

```
=====
| BiodOM |
=====

Notas prévia:
$ -> comando executado como utilizador
# -> comando executado como root do sistema

foram adicionados ao /etc/apt/sources.list, os seguintes
repositórios:
qgis master - versao mais recente qgis,
deb http://qgis.org/debian-nightly squeeze main
deb-src http://qgis.org/debian-nightly squeeze main
backports para instalar postgres 9.1,
deb http://backports.debian.org/debian-backports squeeze-backports
main

=====
| sistema aplicativo, instalação e config |
=====

sistema operativo:
Debian 6.0.6 "squeeze" amd64

aplicações:
apache 2.2.16
openJDK java runtime 6b18-1.8.7
apache tomcat 6.0.35
geoserver 2.1.2
postgres 8.4
postgis 1.5.2
php 5.3.3
pgAdminIII 1.14.0
phpPgAdmin 4.2.3
qgis 1.8.0
osm2pgsql (?)

frameworks:
openlayers 2.11

bibliotecas:
proj4 4.8
geos 3.3.3
gdal 1.9
python 2.6

=====
| .: apache2 |
=====

O apache2 faz parte do repositório normal do Debian. Versão
corrente v2.2.16
```

```

comando instalação:
# apt-get install apache2

configuração:
O ficheiro de configuração principal está localizado em:
/etc/apache2/apache2.conf

ficheiro de configuração:

pasta default:
Default document root for apache2 is
/var/www

comandos comuns:
#/etc/init.d/apache2 restart

=====
| .: java (OpenJDK) |
=====
pre-requisito para o tomcat
OpenJDK Java runtime, using Hotspot JIT 6b18-1.8.7

O OpenJDK faz parte do repositório normal do Debian. Versão
corrente 6b18-1.8.7

comando:
# apt-get install openjdk-6-jre

=====
| .: tomcat 6 |
=====
Package: tomcat6 (6.0.35-1)

NOTA:
O pacote Tomcat7 é um requisito para correr o Geoserver, segundo o
site do Geoserver
Porém, depois das pesquisas efectuadas, verificou-se que o pacote
tomcat7 não esta nos repositórios squeeze-backports, apenas no
testing.
Por esse motivo recorreu-se ao tomcat6

comandos:
# apt-get install tomcat6 tomcat6-admin tomcat6-docs

administração e testar:
http://localhost:8080/
http://localhost:8080/docs

Nomes de utilizadores, passwords e papéis (groups) podem ser
definidos de forma centralizada em um container Servlet. No Tomcat
6.0 isto é realizado através do arquivo

gedit
/etc/tomcat6/tomcat-users.xml

```

```

adicionar
<role rolename="manager"/>
<role rolename="admin"/>
<user username="tomcat" password="tomcat" roles="manager,admin"/>

# /etc/init.d/tomcat6 restart

gestao Tomcat
/etc/init.d/tomcat6 start
/etc/init.d/tomcat6 stop
/etc/init.d/tomcat6 restart

a pasta de trabalho, para colocar os containers java, é:
/var/lib/tomcat6/webapps/

utilização do apache tomcat:
http://localhost:8080/

=====
| .: geoserver |
=====
GeoServer 2.1.3

$ wget http://downloads.sourceforge.net/geoserver/geoserver-2.1.3-war.zip
$ unzip geoserver-2.1.3-war.zip

copiei como root a web application archive do geoserver para o
tomcat6 webapp directory
# cp geoserver.war /var/lib/tomcat6/webapps

# /etc/init.d/tomcat6 restart

Observar que foi criada a pasta /var/lib/tomcat6/webapps/geoserver/
contendo os arquivos Java executáveis do Geoserver.

testar o geoserver:
num browser,
http://localhost:8085/geoserver

sites guia utilizados:
http://concani3.wordpress.com/2012/02/22/instalar-geoserver/
http://docs.geoserver.org/stable/en/user/installation/linux/debian.html
http://docs.geoserver.org/stable/en/user/index.html#

(ver livro PostGIS in Action, para ver a configuração de adicionar
a BD PostGIS ao geoserver, para publicação)

=====
| Postgres 8.4.11 / PostGIS 1.5.1 |
=====

notas para os comandos:
a) Do shell, iniciam-se com:

```

```
postgres$
```

```
b) Do SQL, iniciam-se com:  
postgres=#
```

```
USERS e PASS:
```

```
postgres  
postgres
```

```
dono => template_postgis  
dono
```

```
tipo  
tipo
```

```
BD:  
biodom
```

```
=====  
| PostgreSQL/PostGIS |  
=====
```

```
INSTALAÇÃO:
```

```
1 - verificar se está instalado e que versão  
$ apt-cache policy postgresql
```

```
2 - para instalar o PostgreSQL 8.4:  
# apt-get install postgresql postgresql-doc  
# apt-get install pgadmin3
```

```
3 - criar password para o utilizador "postgres":  
# passwd postgres
```

```
<!--  
4 - libertar o acesso local à base de dados, no arquivo  
/etc/phppgadmin/config.inc.php confirmar as seguintes linhas:  
$conf['servers']['0']['host'] = 'localhost';  
$conf['servers']['0']['defaultdb'] = 'postgres';  
$conf['extra_login_security'] = true;  
-->
```

```
5 - confirmar se a instalação do PostgreSQL foi bem sucedida:  
$ apt-cache policy postgresql
```

```
6 - instalar o PostGIS referente a versão 8.4 do PostgreSQL  
# apt-get install postgis postgresql-8.4-postgis
```

```
7 - confirmar se a instalação do PostGIS foi bem sucedida:  
$ apt-cache policy postgis
```

```
=====
```

| CONFIGURAÇÃO PostgreSQL: | =====

No Debian os parâmetros de configuração encontram-se em:
/etc/postgresql/8.4/main/

pg_hba.conf => este ficheiro determina as condições de acesso
(host-based access)
PostgreSQL.conf =>

Os binários de stop/start:
sudo /etc/init.d/postgresql restart
sudo /etc/init.d/postgresql stop
sudo /etc/init.d/postgresql start

gedit /etc/postgresql/8.4/main/postgresql.conf

change the line listen_addresses = 'localhost' to
listen_addresses = '*'
This opens the server to connections from outside the local
loopback

Optimização de PostgreSQL:
shared_buffers = 256MB
work_mem = 16MB
maintenance_work_mem = 16MB
wal_buffers = 1MB
checkpoint_segments = 6
random_page_cost = 2.0

Criação de utilizadores:

1 - por defeito existe um utilizador "postgres" com privilégios de
superuser

2 - podemos tomar "posse" do utilizador "postgres"
sudo passwd postgres
exit

3 - para poder gerir os utilizadores PostgreSQL é necessário entrar
no ambiente PSQL, depois de ter efectuado o login com o utilizador
"postgres":
su postgres

4 - digitar:
\$ psql
postgres=# \q (para sair do ambiente psql)

5 - Criação de um utilizador:
postgres=# CREATE ROLE utilizador LOGIN PASSWORD '****';
é equivalente a:
postgres=# CREATE USER utilizador PASSWORD '****';
Para removê-lo:
postgres=# DROP ROLE utilizador;
Criação de um superuser:
postgres=# CREATE ROLE utilizador LOGIN SUPERUSER PASSWORD '****';
é equivalente a:

```
postgres=# CREATE USER utilizador SUPERUSER PASSWORD '****';
Em alternativa pode-se fazer :
CREATE ROLE utilizador LOGIN PASSWORD '****' CREATEDB CREATEROLE;
(isto cria um utilizador, que pode por sua vez criar utilizadores e
BD, mas sem todos os privilégios do superuser)
podem-se visualizar os utilizadores com o comando:
\du
```

6 - Criação de um grupo de utilizadores:
 cria-se um novo role sem login:
 CREATE ROLE grupo;
 para adicionar utilizadores a um grupo:
 GRANT grupo TO utilizador_1, utilizador_2, utilizador_n;

Criação de uma BD (normal):

```
1 - criar a BD
CREATE DATABASE database OWNER utilizador TEMPLATE template1;
2 - atribuir/retirar privilégios às tabelas:
GRANT SELECT, INSERT, UPDATE, DELETE ON tabela TO role_name;
REVOKE SELECT, INSERT, UPDATE, DELETE ON tabela FROM role_name;
```

```
=====
| CONFIGURAÇÃO PostGIS: |
=====
```

Criação de novas BD espaciais

O suporte para o PostGIS deve ser habilitado para cada uma das BD que requer a sua utilização. Isto faz-se através do ficheiro postgres.sql (script de activação). O script de activação requer o uso da linguagem PL/pgSQL de forma a operar correctamente, podendo-se usar para tal o programa createlang, que é disponibilizado com a instalação do PostgreSQL.

1 - efectuar o “login” com o utilizador “postgres”:
 sudo su postgres

```
2 - criar um utilizador (com permissão de login e com password):
createuser utilizador l W
# Shall the new role be a superuser? (y/n) n
# Shall the new role be allowed to create databases? (y/n) n
# Shall the new role be allowed to create more new roles? (y/n) n
# Password: <type the user password>
```

CRIAR UMA BD TEMPLATE POSTGIS, para usar na criação de novas BD:

3 - criar uma BD propriedade desse utilizador:
 createdb template_postgis o utilizador

4 - habilitar a linguagem plpgsql para a BD, antes de carregar as funções:
 createlang plpgsql template_postgis
 (ou em ambiente PSQL:
 CREATE LANGUAGE plpgsql;

```

)

5 - carregar as funções e as tabelas de referência espacial:
psql f /usr/share/postgresql/8.4/contrib/postgis-1.5/postgis.sql d
template_postgis
psql f /usr/share/postgresql/8.4/contrib/postgis-
1.5/spatial_ref_sys.sql d template_postgis

6 - Se desejar-mos acrescentar comentários às funções do PostGIS, o
último passo é carregar o script postgis_comments.sql para a nova
BDG:
psql d template_postgis f
/usr/share/postgresql/8.4/contrib/postgis_comments.sql

(Os comentários podem ser visualizados digitando simplesmente:
\dd [function_name]
)

7 - podemos testar o funcionamento da BD template_postgis que
acabamos de criar:
psql -d template_postgis -c "SELECT postgis_full_version();"

CRIAR NOVA BD a partir da template_postgis:

8 - a) Do shell:
postgres$ createdb -T template_postgis -O utilizador biodomdb

b) Do SQL:
postgres=# CREATE DATABASE biodomdb TEMPLATE=template_postgis
OWNER=utilizador;

9 - aceder a BD criada (ainda como superuser "postgres"):
psql biodomdb

10 - atribuir os privilégios standard ogc (relativamente às
tabelas) aos utilizadores:
GRANT ALL
ON geometry_columns
TO utilizador;

GRANT SELECT
ON spatial_ref_sys
TO utilizador;

(sair do PSQL:
\q
)

fazer logout do "postgres":
exit

11 - e como utilizador "normal" entrar na nova BD:
psql h servidor U utilizador d biodomdb

confirmar as tabelas:
\dt

```

```

confirmar as funções:
\df

12 - testar a BD
$ psql -d postgistemplate -c "SELECT postgis_full_version();"

=====
| Carregar BD geoespacial de testes: |
=====

a) Usando a forma SQL,
$ psql -d [database] -f roads.sql

b) Utilizando a aplicação shp2pgsql
$ shp2pgsql shaperoads myschema.roadstable > roads.sql
$ psql -d [database] -f roads.sql

=====
| .: php5 (/apache) v5.3.3 |
=====

# apt-get install php5

=====
| .: phppgadmin v4.2.3 |
=====
.: phppgadmin v4.2.3
# apt-get install phppgadmin

=====
| Open Layers |
=====

1 - descarregar do site openlayers
# wget http://openlayers.org/download/OpenLayers-2.11.zip
# unzip OpenLayers-2.11.zip
# mv OpenLayers-2.11.zip openlayers

2 - colocar, como root, na pasta /var/www
# mv openlayers /var/www

3 - editar como root, o ficheiro /var/www/map.html
ver map01.html, como exemplo

4 - adicionar o caminho /openlayers-2.11/lib/OpenLayers.js ao
ficheiro map.html
(lembrar do workshop da OpenGeo, para o exemplo de utilização do
OpenLayers)

```

```
=====
| proj4js |
=====
```

```
1 - para permitir a conversão de coordenadas entre sistemas
diferentes do Mercator - WGS84 (EPSG:4326) e o Google Mercator
(EPSG:900913) é necessário instalar a biblioteca proj4js:
# wget http://download.osgeo.org/proj4js/proj4js-1.1.0.zip
# unzip proj4js-1.1.0.zip
# mv proj4js /var/www/
```

```
=====
| .: QGIS |
=====
```

```
# apt-get -t squeeze-backports install qgis
```

Anexo C

**Código envolvido na criação de cada novo utilizador a partir do portal
Naturdata**

criar_novo_user.php

```
<?php

// efectua a ligação
$dbh = pg_connect("host=localhost dbname=biodomdb user=postgres");
if (!$dbh) {
    die("Error in connection: " . pg_last_error());
}

// executa a query
$sql = "CREATE USER $user PASSWORD
'md5b5f5ba1a423792b526f799ae4eb3d59e' NOSUPERUSER NOCREATEDB
NOCREATEROLE INHERIT LOGIN;";
$result = pg_query($dbh, $sql);
if (!$result) {
    die("Erro na execução do código SQL: " . pg_last_error());
}

// liberta a memória
pg_free_result($result);

// fecha a ligação aberta
pg_close($dbh);

?>
```

Anexo D

Código desenvolvido na construção do módulo Exif georeader

exif_georeader.php

```
<?php
function toDecimal($deg, $min, $sec, $hem) {
    $d = $deg + $min/60 + $sec/3600;
    return ($hem=='S' || $hem=='W') ? $d*=-1 : $d;
}
function divide($a) {
    $e = explode('/', $a);
    if (!$e[0] || !$e[1]) {
        return 0;
    } else {
        return $e[0] / $e[1];
    }
}
function getGPS() {
    global $exif, $lat_deg, $lat_min, $lat_sec, $lat_hem, $log_deg,
    $log_min, $log_sec, $log_hem;
    if ($exif) {
        $lat = $exif['GPS']['GPSLatitude'];
        $log = $exif['GPS']['GPSLongitude'];
        if (!$lat || !$log) return null;
        $lat_deg = divide($lat[0]);
        $lat_min = divide($lat[1]);
        $lat_sec = divide($lat[2]);
        $lat_hem = $exif['GPS']['GPSLatitudeRef'];
        $log_deg = divide($log[0]);
        $log_min = divide($log[1]);
        $log_sec = divide($log[2]);
        $log_hem = $exif['GPS']['GPSLongitudeRef'];
        $ltd_dec = toDecimal($lat_deg, $lat_min, $lat_sec, $lat_hem);
        $lgd_dec = toDecimal($log_deg, $log_min, $log_sec, $log_hem);
        return array(round($ltd_dec, 7), round($lgd_dec, 7));
    } else {
        return null;
    }
}
function hasSection($str) { global $exif; return
    strstr($exif['FILE']['SectionsFound'], $str); }
function strTrunc($str, $strunc) { return substr($str, 0,
    strlen($str)-$strunc); };
function normalize() {
    global $lat_min, $lat_sec, $log_min, $log_sec;
    $x = (int)$lat_min;
    $y = (int)$log_min;
    $lat_sec = round(($lat_min - $x) * 60, 2);
    $log_sec = round(($log_min - $y) * 60, 2);
    $lat_min = $x;
    $log_min = $y; }
function getEXIF($dir) {
    global $exif;
    if (is_dir($dir)) {
        if ($dh = opendir($dir)) {
            $count = 1;
            while (($file = readdir($dh)) !== false) {
                if (strstr($file, '.jpg')) {
                    $exif = exif_read_data($file, 0, true);
                    $str = " <tr>\n <td><div><a href=\"gps_exif.php?img=\" .
```

```

strTrunc($file, 4) . "\"><img src=\"";
    $str .= (hasSection('THUMBNAIL')) ? "thumbnail.php?file=" .
$file . "\" : $file . "\" width=\"195\"";
    $str .= "></a></div></td>\n <td><div>";
    $str .= "File: <b>" . $exif['FILE']['FileName'] .
"</b><br>\n";
    $str .= "Timestamp : " . date("m/d/Y h:i:s A T", $exif['FILE']
['FileDateTime']) . "<br>\n";
    $str .= "Date taken: " . $exif['EXIF']['DateTimeOriginal'] .
"<br>\n";
    $str .= "Dimensions: " . $exif['COMPUTED']['Width'] . " x " .
$exif['COMPUTED']['Height'] . " <br>\n";
    $gps = (hasSection('GPS')) ? getGPS() : null;
    if ($gps != null) {
        $str .= "Latitude : " . $gps[0] . "&deg;<br>\n";
        $str .= "Longitude : " . $gps[1] . "&deg;<br>\n";
        $str .= "<a href=\"http://www.wikimapia.org/#lat=" .
$gps[0] . "&lon=" . $gps[1] . "&z=17\">Map Reference " .
$count++ . "</a>\n";
    }
    $str .= " </div></td>\n";
    $str .= " </tr>\n";
    echo $str;
}
}
closedir($dh);
}
}
}
function getImg($img) {
    global $exif;
    if (file_exists($img)) {
        $str = "<img src=\"$img\" ";
        $chtml = $exif['COMPUTED']['html'];
        if (!$chtml) {
            $fil_wid = $exif['COMPUTED']['Width'];
            $fil_hgt = $exif['COMPUTED']['Height'];
            $str .= "width=\"\" . (($fil_wid > 0) ? $fil_wid : "100%") . "\"
height=\"\" . (($fil_hgt > 0) ? $fil_hgt : "100%") . "\"";
        } else {
            $str .= $chtml;
        }
        $str .= " alt=\"\" . strTrunc($img, 4) . "\">";
        echo $str;
    } else {
        echo "Image not found.";
    }
}
function readExif($img) {
    global $exif, $gps_alt, $lat_dec, $lon_dec, $use_com, $errStr;
    $errStr = "";
    if (file_exists($img)) {
        if (exif_read_data($img)) {
            $exif = exif_read_data($img, 0, true);
            echo "<div id=\"exifdump\">\n";
            print_r($exif);
            echo "</div>\n";
        }
    }
}

```

```

$use_com = (hasSection('COMMENT')) ? $exif['COMMENT'][0] : null;
if (hasSection('GPS')) {
    $gps = $exif['GPS']['GPSAltitudeRef'];
    $alt_ref = ($gps !== null) ? $gps : null;
    $gps = $exif['GPS']['GPSAltitude'];
    $gps_alt = ($gps !== null) ? round(divide($gps), 4) : "N/A";
    $gps = getGPS();
    if ($gps !== null) {
        $lat_dec = $gps[0];
        $lon_dec = $gps[1];
    }
    normalize();
} else {
    $errStr .= "    <tr>\n<td colspan=\"3\"><p>Não foram
encontradas etiquetas de GPS.<p></td>\n";
}
} else {
    $errStr .= "    <tr>\n<td colspan=\"3\"><p>Não foram encontradas
etiquetas no EXIF.<p></td>\n";
}
}
}
function print_data() {
    global $errStr, $exif, $gps_alt, $lat_deg, $lat_min, $lat_sec,
    $lat_hem, $lat_dec, $log_deg, $log_min, $log_sec, $log_hem,
    $lon_dec, $use_com, $h;
    $str = "";
    if ($use_com) { $str .= "    <tr>\n        <td colspan=\"3\">\n
<p class=\"comment\">$use_com</p>\n        </td>\n        </tr>\n"; }
    if (!$errStr) {
        $str .= "    <tr><th scope=\"col\">LATITUDE</th><th
scope=\"col\">LONGITUDE</th><th scope=\"col\">ALTITUDE</th></tr>\n
<tr>\n";
        $str .= "        <td>\n            <p>$lat_deg&deg; $lat_min&rsquo;
$lat_sec&rdquo; $lat_hem</p>\n            <p>$lat_dec&deg;</p>\n
</td>\n";
        $str .= "        <td>\n            <p>$log_deg&deg; $log_min&rsquo;
$log_sec&rdquo; $log_hem</p>\n            <p>$lon_dec&deg;</p>\n
</td>\n";
        $z = isset($h) ? ((intval($h) <= 4 && intval($h) >= -4) ?
"&z=" . (string)(17 + intval($h)) : "&z=17") : "&z=17";
        $str .= "        <td>\n            <p>" . (($gps_alt == "N/A") ? $gps_alt :
"$gps_alt m" . (($alt_ref) ? "Abaixo" : "Acima") . " Nível do
Mar.") . "</p>\n            <p>Ir para esta <a
href=\"http://www.biodom.pt/#lat=" . $lat_dec . "&lon=" .
$lon_dec . $z . "\">localização no mapa</a></p>\n        </td>\n";
    } else {
        $str .= $errStr;
    }
    $str .= "    </tr>";
    echo $str;
}
?>

```

(código adaptado de: <http://www.quietless.com/kitchen/extract-exif-data-using-php-to-display-gps-tagged-images-in-google-maps/> e <http://www.webproworld.com/webmaster-forum/threads/119120-Extracting-and-displaying-EXIF-data-with-PHP/page2>)

Anexo E

**Código desenvolvido na construção do trigger de tracking das operações de
CRUD**

CRUD_trigger.sql

```
CREATE TABLE emp (  
    empname      text NOT NULL,  
    valor        integer  
);  
  
CREATE TABLE registo (  
    operacao     char(1) NOT NULL,  
    stamp        timestamp NOT NULL,  
    userid       text NOT NULL,  
    empname      text NOT NULL,  
    valor        integer  
);  
  
CREATE OR REPLACE FUNCTION process_emp_audit() RETURNS TRIGGER AS  
$emp_audit$  
    BEGIN  
        --  
        -- Cria uma linha em registo para reflectir a operação  
realizada,  
        -- fazendo uso da variável especial TG_OP para trabalhar a  
operação.  
        --  
        IF (TG_OP = 'DELETE') THEN  
            INSERT INTO registo SELECT 'D', now(), user, OLD.*;  
            RETURN OLD;  
        ELSIF (TG_OP = 'UPDATE') THEN  
            INSERT INTO registo SELECT 'U', now(), user, NEW.*;  
            RETURN NEW;  
        ELSIF (TG_OP = 'INSERT') THEN  
            INSERT INTO registo SELECT 'I', now(), user, NEW.*;  
            RETURN NEW;  
        END IF;  
        RETURN NULL; -- o resultado é ignorado dado que este é um  
trigger à posteriori  
    END;  
$registo$ LANGUAGE plpgsql;  
  
CREATE TRIGGER registo  
AFTER INSERT OR UPDATE OR DELETE ON emp  
FOR EACH ROW EXECUTE PROCEDURE process_emp_audit();
```

Anexo F

Resultado dos testes de operacionalidade do sistema aplicativo

Componente a testar	Teste	Notas	Resultado
Camada de acesso: Geoserver	Teste de publicação de GeoTIFF	-	Operacional
Camada de acesso: Geoserver	Getcapabilities e respostas WMS e WFS	-	Operacional
Camada de recursos: BDG	Utilização geral: importação e exportação de dados	-	Operacional
Camada de recursos: BDG	Acesso de multi-utilizadores simultâneos à BDG em modo de edição.	-	Operacional
Camada de recursos: processo	Criação automática de utilizadores	-	Operacional
Camada de IGU: ficha de espécie	Testes gerais de usabilidade	-	Operacional
Camada de IGU: mapa biodom	Testes gerais de usabilidade e operacionalidade	-	Operacional
Módulo Browser Editor	Digitalização de um tema de 50 pontos via browser e	-	Operacional
Módulo Exif Georeader	Carregamento de um conjunto de 20 fotos e acesso à informação na BDG.	-	Operacional
Módulo GPX Georeader	Carregamento de 3 tracks GPS e acesso à informação na BDG.	-	Operacional
QGIS Editor	Ligação à BDG com QGIS e user "teste1". Edição de informação. Descarga de informação.	-	Operacional