



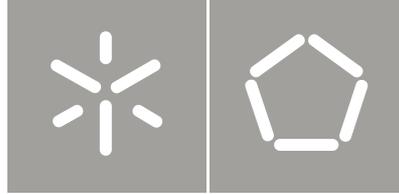
Universidade do Minho
Escola de Engenharia

Sara Alexandra da Costa Fernandes Coelho | Plataforma Online para Casos de Estudo em Cidades Inteligentes

Sara Alexandra da Costa Fernandes Coelho
**Plataforma Online para Casos de Estudo em
Cidades Inteligentes**

UMinho | 2016

Outubro de 2016



Universidade do Minho
Escola de Engenharia

Sara Alexandra da Costa Fernandes Coelho
**Plataforma Online para Casos de Estudo em
Cidades Inteligentes**

Tese de Mestrado
Ciclo de Estudos Integrados Conducentes ao Grau de
Mestre em Engenharia de Telecomunicações e
Informática

Trabalho efetuado sob a orientação de
Professor Doutor Nuno Vasco Lopes
Professora Doutora Elsa Estevez

Agradecimentos

Ao finalizar esta etapa tão desejada e bastante enriquecedora, a qual exigiu muita dedicação, persistência e empenho, pretendo agradecer a um conjunto de pessoas que me ajudaram a tornar este objetivo numa realidade.

Por forma da concretização deste projeto foi fundamental a contribuição e o apoio de diversas pessoas essenciais e de elevada importância, às quais devo o meu sincero agradecimento.

Primeiramente, o meu agradecimento é dirigido ao meu orientador, Professor Nuno Vasco Lopes, pela confiança depositada em mim, pela total disponibilidade, pela orientação, e pelas palavras avultadas de incentivo e motivação durante o desenvolvimento da dissertação. Agradeço, também, à minha coorientadora Professora Elsa Estevez pela partilha do conhecimento acerca da tecnologia OLAP e da área de *Business Intelligence*.

Aos membros da Unidade Operacional da Universidade das Nações Unidas (UNU), situada em Guimarães, agradeço o carinho e simpatia com que sempre me receberam, neste local, para as reuniões de esclarecimento com os orientadores.

Seguidamente, agradeço a todos os meus colegas e amigos do curso, pelo companheirismo, pela amizade e pela entreaajuda nos diversos momentos ao longo do percurso académico. Também, agradeço à Universidade do Minho, bem como a todos os professores, pelos ensinamentos e partilha de experiências durante esta formação académica.

Em especial, aos meus pais, José e Rosa, o meu eterno e profundo agradecimento, por todo o esforço e dedicação em proporcionar-me a melhor educação e pelos sublimes valores que me inculcaram. Agradeço-lhes toda a ternura e carinho com que sempre me confortaram. Manifesto, ainda, o meu agradecimento a eles, por toda a força e motivação nas diversas ocasiões, pelo incansável apoio e ajuda no alcance dos meus objetivos, acreditando sempre em mim. Gratifico também, ao meu irmão

Ricardo e à minha madrinha Ana Margarida pelo constante ânimo e incentivo concedidos durante este percurso. Aos meus primos João e Filipa Silva, e restante família, agradeço os seus contributos de apoio prestados.

Por último, mas não menos importante, ao meu namorado, Diogo Lima, agradeço profundamente por toda a compreensão, carinho, pela perseverante ajuda nos momentos menos bons, por todo o encorajamento e motivação incondicionais transmitidos, particularmente, no decorrer deste projeto.

A todos o meu muito obrigada!

Resumo

Nos dias de hoje, com o crescimento populacional nas cidades, há cada vez mais a necessidade de adaptar as áreas urbanas de forma inteligente, face ao aumento da procura de eletricidade, água, alimentos, transportes e diversos serviços, visando para isso o desenvolvimento e produção de soluções inovadoras. Para tal, é necessário que empresas, organizações, parceiros financeiros, entidades públicas e governos se afirmem na medida de alcançar, produzir e experimentar produtos e serviços que promovam cidades mais habitáveis e sustentáveis, denominadas por Cidades Inteligentes (*Smart Cities*). As Cidades Inteligentes estão a tornarem-se numa prioridade nas políticas dos países, na medida de serem uma estratégia de reindustrialização e de inovação inteligente, impulsionando a reinvenção das cidades a vários níveis como sustentabilidade, desenvolvimento económico, inclusão social e qualidade de vida. O conceito de Cidade Inteligente exige um conjunto de competências e uma capacidade em integrar soluções e sistemas em diversas áreas como em redes de energia, mobilidade, edifícios, serviços públicos e em Tecnologias de Informação e Comunicação (TIC). Os projetos em Cidades Inteligentes têm como objetivo solucionar problemas urbanos através de soluções inovadoras e inteligentes.

Muitas dessas soluções são obtidas através da utilização de sistemas *Business Intelligence*, que oferecem suporte à tomada de decisão e gestão de negócios, possibilitando uma análise de um grande volume de dados mais eficiente e rápida. E por essas vantagens o *Business Intelligence* é cada vez mais adotado pelo setor empresarial.

Esta dissertação é desenvolvida em colaboração com a Unidade Operacional *Operating Unit on Policy-Driven Electronic Governance*, UNU-EGOV, sediada na Universidade do Minho, em Guimarães. O objetivo desta dissertação é o desenvolvimento de uma plataforma *online* capaz de analisar e armazenar dados sobre iniciativas em Cidades Inteligentes. Esta plataforma permite reunir informação sobre as diversas iniciativas em Cidades Inteligentes através de um formulário *online*. Os dados recolhidos são processados e analisados, com recurso à tecnologia OLAP (*Online Analytical Processing*) *open source*, para a elaboração de estatísticas e mapas geográficos referentes às Cidades Inteligentes. Deste modo, a plataforma proposta pretende constituir uma ferramenta útil para os decisores políticos e gestores, que tencionam implementar novas iniciativas em Cidades Inteligentes e aperfeiçoar as existentes, necessitando, para tal, de alguma orientação.

Palavras-Chave: Cidades Inteligentes, *Business Intelligence*, OLAP, *open source*.

Abstract

Currently, the fast population growth in cities demands intelligent and innovative solutions capable of answer to the increasing given the increasing demand for electricity, water, food, transport and other services. In order to achieve this goal companies, organizations, financial partners, public authorities and governments, i.e. all stakeholders, must work together to produce new smart products and services that contribute affectively to livable and sustainable cities, called Smart Cities. The smart cities are becoming a priority in the strategies of several countries around the world. They are seen as the new reindustrialization era , which will allow the reinvention of cities at all levels: economic, environmental, mobility, governance, social and quality of live. The Smart City concept requires a set of skills and an ability to integrate solutions and systems in different areas, such as energy networks, mobility, buildings, public services and Information and Communications Technologies (ICT). The projects in Smart Cities aim to solve urban problems through innovative and intelligent solutions. Many of these solutions are obtained through the use of Business Intelligence systems, that give support to the decision-making process and to the business management, enabling analysis of a large amounts of data more efficiently and quickly. The advantages offered by Business Intelligence systems are being harnessed by public and private sectors.

This dissertation was developed in collaboration with Operating Unit on Policy-Driven Electronic Governance, UNU-EGOV, hosted in University of Minho and located in Guimarães. The goal of this dissertation was to develop an online plataform able to analyse and store the smart cities initiatives. The platform allows to collect data about the smart cities initiatives through out an online form. The data collected is processed and treated using open-source OLAP (Online Analytical Processing) technologies for producing aggregated statistics and maps regarding smart cities. The proposed platform intends to be an useful tool for policy-makers and managers, which want to implement new smart cities initiatives and need some guidance to run it.

Keywords: Smart Cities, Business Intelligence, OLAP, open source.

Conteúdo

Agradecimentos	v
Resumo	vii
Abstract	ix
Lista de Figuras	xv
Lista de Tabelas	xix
Lista de Acrónimos	xxi
1 Introdução	1
1.1 Contextualização	1
1.2 Motivação	4
1.3 Finalidade e Objetivos	5
1.4 Estrutura da Dissertação	7
2 Estado da Arte	9
2.1 Cidades Inteligentes	9
2.1.1 A importância das TIC em Cidades Inteligentes	13
2.1.2 A importância de Iniciativas Sustentáveis em Cidades Inteligentes	15
2.1.3 Projeto “ <i>Smart Cities for Sustainable Development</i> ”	16
2.2 <i>Business Intelligence</i>	21
2.2.1 Processo de ETL (<i>Extract, Transform and Load</i>)	25
2.2.2 <i>Data Warehouse</i>	26
2.2.3 Modelação de dados	33
2.2.4 Cubos multidimensionais	39
2.2.5 OLAP (<i>Online Analytical Processing</i>)	41

2.2.6	<i>Data Mining</i>	45
2.2.7	Aplicações <i>Front-End</i>	45
2.3	Conclusões	46
3	Estudo de ferramentas de BI <i>open source</i>	49
3.1	Critérios para avaliação de ferramentas OLAP	49
3.2	Descrição das ferramentas de BI <i>open source</i> OLAP	53
3.2.1	<i>SpagoBI Suite</i>	54
3.2.2	<i>JasperSoft</i>	56
3.2.3	<i>Vanilla</i>	58
3.2.4	<i>Palo/Jedox</i>	59
3.2.5	<i>Pentaho Business Intelligence</i>	60
3.3	Tabela Comparativa das Ferramentas do tipo OLAP	65
3.4	Revisão literária de ferramentas de BI <i>open source</i>	69
3.5	Conclusão do estudo e análise das ferramentas de BI	70
4	Especificação do sistema	73
4.1	Levantamento e especificação de requisitos	73
4.1.1	Requisitos Não Funcionais	74
4.1.2	Requisitos Funcionais	75
4.2	Descrição dos casos de uso	77
4.2.1	Utilizador sem <i>login</i> efetuado no sistema	78
4.2.2	Utilizador com autenticação no sistema	78
4.2.3	Administrador	79
4.3	Esboço da Plataforma <i>Online</i>	81
4.4	Arquitetura do sistema	83
4.5	Conclusões	86
5	Implementação do sistema	87
5.1	Ferramentas e tecnologias utilizadas	87
5.1.1	Ferramentas para análises estatísticas	88
5.1.2	Ferramentas e tecnologias/linguagens para o desenvolvimento <i>web</i>	95
5.2	Sistema de <i>Data Warehousing</i>	99
5.2.1	Modelo de dados do <i>Data Warehouse</i>	100
5.2.2	Cubo OLAP	136

5.2.3	Criação de Gráficos com a ferramenta analítica <i>Pentaho Community Edition</i>	138
5.3	Plataforma <i>Online</i>	141
5.3.1	Formulário <i>online</i> : “ <i>Smart Cities for Sustainable Development</i> ”	144
5.3.2	Análises Estatísticas	148
5.4	Aplicação <i>web</i> de Administração	148
5.5	Conclusões	150
6	Resultados do sistema	153
6.1	Plataforma <i>Online</i>	153
6.2	Aplicação <i>web</i> de Administração	166
6.3	Conclusões	171
7	Conclusões	173
7.1	Síntese	173
7.2	Contribuições	177
7.3	Trabalho Futuro	179
	Referências	181

Lista de Figuras

2.1	Principais características dos domínios em Cidades Inteligentes	12
2.2	Modelo conceptual para o Desenvolvimento Sustentável em Cidades Inteligentes, retirado de [1]	19
2.3	Secção do documento “ <i>Smart Cities for Development - State of Practice Survey</i> ”, retirado de [1]	20
2.4	Arquitetura geral de um sistema de <i>Business Intelligence</i> , baseado em [2]	23
2.5	Arquitetura geral de um sistema de <i>Data Warehouse</i> , baseado em [3, 4, 5]	27
2.6	Abordagem <i>top-down</i> da arquitetura de um sistema de <i>Data Warehouse</i>	32
2.7	Abordagem <i>bottom-up</i> da arquitetura de um sistema de <i>Data Warehouse</i>	33
2.8	Esquema em Estrela	36
2.9	Esquema em Floco de Neve	37
2.10	Esquema em Constelação de Factos	38
2.11	Exemplo de um cubo com três dimensões: Produtos, Tempo e Localização	40
3.1	Exemplo de gráficos interativos na <i>suite SpagoBI</i> , retirado de [6] . . .	56
3.2	Exemplo de análise interativa na <i>suite JasperSoft</i> , retirado de [7] . . .	57
3.3	Exemplo de <i>dashboards</i> e gráficos interativos na <i>suite Vanilla</i> , retirado de [8]	59
3.4	Exemplo de relatório com gráficos interativos na <i>suite Palo/Jedox</i> , retirado de [9]	60
3.5	Arquitetura da ferramenta <i>suite Pentaho BI</i> , retirado de [10]	61
3.6	Exemplo de gráficos e <i>dashboards</i> interativos na <i>suite Pentaho BI</i> , retirado de [11]	65
4.1	Diagrama de caso de uso - Utilizador sem <i>login</i> efetuado no sistema .	78

4.2	Diagrama de caso de uso - Utilizador autenticado no sistema	79
4.3	Diagrama de caso de uso - Administrador	80
4.4	Esboço da plataforma <i>online</i> - Inicial	81
4.5	Esboço da plataforma <i>online</i> - <i>Login</i> efetuado: Formulário	82
4.6	Esboço da plataforma <i>online</i> - <i>Login</i> efetuado: Análises (Gráficos) . .	83
4.7	Arquitetura da plataforma <i>online</i>	84
5.1	Ferramentas e tecnologias/linguagens utilizadas no desenvolvimento da plataforma <i>online</i>	88
5.2	Ambiente de trabalho da ferramenta <i>Pentaho Schema Workbench</i> . .	89
5.3	Ambiente de trabalho da aplicação <i>front-end</i> PUC	90
5.4	Ambiente de trabalho do <i>plugin Saiku Analytics</i>	91
5.5	Ambiente de trabalho do <i>plugin</i> CDE	92
5.6	Ambiente de trabalho do Editor <i>Sublime Text 3</i>	95
5.7	Ambiente da ferramenta <i>XAMPP Control Panel</i>	96
5.8	Ambiente de trabalho da ferramenta <i>MySQL Workbench 6.3 CE</i> . . .	97
5.9	Modelo multidimensional do sistema de <i>Data Warehouse</i>	101
5.10	Tabela de factos <i>Fact_Use_Case</i>	103
5.11	Tabela de dimensão <i>Practices</i>	105
5.12	Tabelas de dimensão <i>Surveys</i> e <i>Sources</i>	106
5.13	Tabelas de dimensão <i>Dim_Surveyors</i> e <i>Dim_Who_Surveyors</i> e tabela de ponte <i>Surveyors_has_Surveyors</i>	107
5.14	Tabelas de dimensão <i>Dim_Responsibles_Institutions</i> , <i>Dim_RI_Name</i> , <i>Dim_Types_RI</i> e <i>Dim_Resp_Inst_Role</i>	109
5.15	Tabelas de dimensão <i>Partners_BTTF</i> , <i>Dim_Partners</i> , <i>Dim_Types_Partners</i> , <i>Dim_Partners_Role</i>	112
5.16	Tabela de dimensão <i>Dim_Date</i>	115
5.17	Tabela de dimensão <i>Dim_Cities</i>	116
5.18	Tabelas de dimensão <i>Dim_What</i> e <i>Dim_Concepts</i> e tabela de ponte <i>Concepts_BDWhat</i>	118
5.19	Tabelas de dimensão <i>Dim_What</i> e <i>Dim_Innovations</i> e tabela de ponte <i>Innovations_BDWhat</i>	119
5.20	Tabelas de dimensão <i>Dim_What</i> e <i>Dim_Lessons_Learnt</i> e tabela de ponte <i>Lessons_BDWhat</i>	121

5.21	Tabelas de dimensão <i>Dim_Why</i> e <i>Dim_Drivers</i> e tabela de ponte <i>Drivers_BDWhy</i>	123
5.22	Tabelas de dimensão <i>Dim_Why</i> e <i>Dim_Benefits</i> e tabela de ponte <i>Benefits_BDWhy</i>	125
5.23	Tabelas de dimensão <i>Dim_Why</i> e <i>Dim_Values</i> e tabela de ponte <i>Values_BDWhy</i>	126
5.24	Tabela de ponte <i>Governances_BTF</i> e tabela de dimensão <i>Dim_Governances</i>	128
5.25	Tabela de ponte <i>Risks_BTF</i> e tabela de dimensão <i>Dim_Risks</i>	129
5.26	Tabela de ponte <i>Technologies_BTF</i> e tabela de dimensão <i>Dim_Technologies</i>	131
5.27	Tabela de ponte <i>Maturities_BTF</i> e tabela de dimensão <i>Dim_Maturities</i>	132
5.28	Tabela de ponte <i>Tools_BTF</i> e tabela de dimensão <i>Dim_Tools</i>	134
5.29	Tabela de ponte <i>Challenges_BTF</i> e tabela de dimensão <i>Dim_Challenges</i>	135
5.30	Esquema do cubo OLAP desenvolvido na ferramenta PSW	138
5.31	Ambiente de trabalho do <i>plugin</i> CDE - Menu de configuração <i>Layout</i>	139
5.32	Ambiente de trabalho do <i>plugin</i> CDE - Menu de configuração <i>Components</i>	140
5.33	Ambiente de trabalho do <i>plugin</i> CDE - Menu de configuração <i>Data-sources</i>	141
5.34	Mapa do <i>website</i> relativo à plataforma <i>online</i>	142
5.35	Esquema do procedimento de uma solicitação HTTP de um cliente e de uma resposta HTTP do servidor	143
5.36	Exemplo da utilização do elemento de entrada de texto	145
5.37	Exemplo da utilização do elemento de entrada do tipo rádio	145
5.38	Exemplo da utilização do elemento de entrada do tipo caixa de seleção	146
5.39	Exemplo da utilização do elemento <i>select</i>	146
5.40	Elementos para inserção da localização da iniciativa	147
5.41	Mapa do <i>website</i> relativo à aplicação <i>web</i> de administração	149
6.1	Demonstração da página <i>web</i> inicial - <i>Home</i>	154
6.2	Demonstração da página <i>web</i> <i>UNU Project</i>	155
6.3	Demonstração da página <i>web</i> <i>Form</i> - Parte inicial	156

6.4	Demonstração da página <i>web Form</i> - Questão “ <i>What</i> ”	157
6.5	Demonstração da página <i>web Form</i> - Validação	157
6.6	Demonstração da página <i>web About form</i>	158
6.7	Demonstração da página <i>web List of Initiatives</i>	159
6.8	Demonstração da página <i>web Statistical Analysis 3 - Why</i> - Análises estatísticas	160
6.9	Demonstração da página <i>web Map of Initiatives</i> - Mapa	161
6.10	Demonstração da página <i>web Contact us</i>	162
6.11	Demonstração da página <i>web Login/Register</i> - Secção <i>Login</i>	163
6.12	Demonstração da página <i>web Login/Register</i> - Secção <i>Register</i>	164
6.13	Demonstração do <i>e-mail</i> enviado ao utilizador após o registo	165
6.14	Mensagem de confirmação via <i>web browser</i>	165
6.15	Demonstração da página <i>web Login</i> - Administração	166
6.16	Demonstração da página <i>web Home</i> - Administração	167
6.17	Demonstração da página <i>web Insert Admin's</i> - Administração	168
6.18	Demonstração da página <i>web List Requests - Users</i> - Administração	169
6.19	Demonstração da página <i>web</i> para gestão da permissão dos utilizadores - Administração	169
6.20	Demonstração do <i>e-mail</i> de autorização de <i>login</i> ao utilizador	170
6.21	Demonstração da página <i>web List/Edit/Remove Forms</i> - Administração	171

Lista de Tabelas

2.1	Comparação entre Cidades Tradicionais e Cidades Inteligentes, baseado em [12]	14
2.2	Comparação entre Sistemas OLTP e Sistemas OLAP, baseado em [13, 14, 15, 16]	29
3.1	Doze regras de Codd para a avaliação de produtos OLAP, baseado em [17, 18]	50
3.2	Funcionalidades e características de ferramentas <i>open source</i> , baseado em [19, 20, 21, 22, 23, 24, 25]	66
5.1	Descrição de requisitos de <i>Hardware</i> e de <i>Software</i> do Servidor da ferramenta <i>Pentaho</i> , adaptado de [26]	93
5.2	Descrição de requisitos de <i>Hardware</i> e de <i>Software</i> da ferramenta <i>Pentaho Schema Workbench</i> , adaptado de [26]	94
5.3	Descrição da tabela de factos <i>Fact_Use_Case</i>	103
5.4	Descrição da tabela de dimensão <i>Dim_Practices</i>	105
5.5	Descrição da tabela de dimensão <i>Dim_Surveys</i>	106
5.6	Descrição da tabela de dimensão <i>Dim_Sources</i>	107
5.7	Descrição da tabela de dimensão <i>Dim_Surveyors</i>	108
5.8	Descrição da tabela de dimensão <i>Dim_Who_Surveyors</i>	108
5.9	Descrição da tabela de ponte <i>Surveyors_has_Surveyors</i>	108
5.10	Descrição da tabela de dimensão <i>Dim_Responsibles_Institutions</i>	110
5.11	Descrição da tabela de dimensão <i>Dim_RI_Name</i>	110
5.12	Descrição da tabela de ponte <i>Dim_Types_RI</i>	111
5.13	Descrição da tabela de dimensão <i>Dim_Resp_Inst_Role</i>	111
5.14	Descrição da tabela de dimensão <i>Partners_BTF</i>	113
5.15	Descrição da tabela de dimensão <i>Dim_Partners</i>	113

5.16	Descrição da tabela de dimensão <i>Dim_Types_Partners</i>	114
5.17	Descrição da tabela de dimensão <i>Dim_Partners_Role</i>	114
5.18	Descrição da tabela de dimensão <i>Dim_Date</i>	115
5.19	Descrição da tabela de dimensão <i>Dim_Cities</i>	116
5.20	Descrição da tabela de dimensão <i>Dim_What</i>	117
5.21	Descrição da tabela de ponte <i>Concepts_BDWhat</i>	118
5.22	Descrição da tabela de dimensão <i>Dim_Concepts</i>	119
5.23	Descrição da tabela de ponte <i>Innovations_BDWhat</i>	120
5.24	Descrição da tabela de dimensão <i>Dim_Innovations</i>	120
5.25	Descrição da tabela de ponte <i>Lessons_BDWhat</i>	121
5.26	Descrição da tabela de dimensão <i>Dim_Lessons_Learnt</i>	122
5.27	Descrição da tabela de dimensão <i>Dim_Why</i>	122
5.28	Descrição da tabela de ponte <i>Drivers_BDWhy</i>	124
5.29	Descrição da tabela de dimensão <i>Dim_Drivers</i>	124
5.30	Descrição da tabela de ponte <i>Benefits_BDWhy</i>	125
5.31	Descrição da tabela de dimensão <i>Dim_Benefits</i>	126
5.32	Descrição da tabela de ponte <i>Values_BDWhy</i>	126
5.33	Descrição da tabela de dimensão <i>Dim_Values</i>	127
5.34	Descrição da tabela de ponte <i>Governances_BTF</i>	128
5.35	Descrição da tabela de dimensão <i>Dim_Governances</i>	129
5.36	Descrição da tabela de ponte <i>Risks_BTF</i>	130
5.37	Descrição da tabela de dimensão <i>Dim_Risks</i>	130
5.38	Descrição da tabela de ponte <i>Technologies_BTF</i>	131
5.39	Descrição da tabela de dimensão <i>Dim_Technologies</i>	132
5.40	Descrição da tabela de ponte <i>Maturities_BTF</i>	133
5.41	Descrição da tabela de dimensão <i>Dim_Maturities</i>	133
5.42	Descrição da tabela de ponte <i>Tools_BTF</i>	134
5.43	Descrição da tabela de dimensão <i>Dim_Tools</i>	135
5.44	Descrição da tabela de ponte <i>Challenges_BTF</i>	136
5.45	Descrição da tabela de dimensão <i>Dim_Challenges</i>	136

Lista de Acrónimos

AJAX	<i>Asynchronous Javascript And XML</i>
BI	<i>Business Intelligence</i>
CDE	<i>Community Dashboard Edition</i>
CE	<i>Community Edition</i>
CRM	<i>Customer Relationships Management</i>
CSS	<i>Cascading Style Sheets</i>
DM	<i>Data Mart</i>
DW	<i>Data Warehouse</i>
EJB	<i>Enterprise Java Beans</i>
EPL	<i>Eclipse Public License</i>
ERP	<i>Enterprise Resources Planning</i>
ETL	<i>Extract, Transform, Load</i>
FASMI	<i>Fast Analysis of Shared Multidimensional Information</i>
FK	<i>Foreign Key</i>
GNU	<i>GNU's Not Unix</i>
GPL	<i>General Public License</i>
LGPL	<i>Lesser General Public License</i>
MPL	<i>Mozilla Public License</i>
HOLAP	<i>Hybrid OnLine Analytical Processing</i>
HSQLDB	<i>Hyperthreaded Structured Query Language DataBase</i>
HTML	<i>HyperText Markup Language</i>
HTTP	<i>HyperText Transfer Protocol</i>
IDRC	<i>International Development Research Center</i>
IP	<i>Internet Protocol</i>
JDBC	<i>Java DataBase Connectivity</i>
JNDI	<i>Java Naming and Directory Interface</i>
JS	<i>JavaScript</i>

JSON	<i>JavaScript Object Notation</i>
KPI	<i>Key Performance Indicator</i>
Mac OS	<i>Macintosh Operating System</i>
MD5	<i>Message-Digest algorithm 5</i>
MDX	<i>MultiDimensional eXpressions</i>
MOLAP	<i>Multidimensional OnLine Analytical Processing</i>
MPL	<i>Mozilla Public License</i>
OCI	<i>Oracle Call Interface</i>
ODBC	<i>Open Database Connectivity</i>
OLAP	<i>OnLine Analytical Processing</i>
OLTP	<i>OnLine Transaction Processing</i>
PAS	<i>Pentaho Analysis Services</i>
PDI	<i>Pentaho Data Integration</i>
PHP	<i>PHP: Hypertext Preprocessor</i>
PK	<i>Primary Key</i>
POJO	<i>Plain Old Java Object</i>
PSW	<i>Pentaho Schema Workbench</i>
PUC	<i>Pentaho User Console</i>
RAM	<i>Random Access Memory</i>
RDBMS	<i>Relational DataBase Management Systems</i>
ROLAP	<i>Relational OnLine Analytical Processing</i>
SAD	<i>Sistemas de Apoio à Decisão</i>
SGBD	<i>Sistema de Gestão de Base de Dados</i>
SQL	<i>Structured Query Language</i>
TCP	<i>Transmission Control Protocol</i>
TIC	<i>Tecnologias de Informação e Comunicação</i>
UML	<i>Unified Modeling Language</i>
UNU	<i>United Nations University</i>
EGOV	<i>Electronic Governance</i>
URL	<i>Uniform Resource Locator</i>
WSN	<i>Wireless Sensor Network</i>
XML	<i>eXtensible Markup Language</i>
XMLA	<i>eXtensible Markup Language for Analysis</i>

Capítulo 1

Introdução

Este capítulo inicia-se com uma contextualização do âmbito da dissertação (secção 1.1), bem como a sua motivação (secção 1.2), a finalidade e os principais objetivos (secção 1.3). Por último, este capítulo termina com a apresentação da estrutura deste documento (secção 1.4).

1.1 Contextualização

A cidade, enquanto uma unidade de governo, tem vindo a crescer e a tornar-se mais complexa e importante, face ao crescimento elevado da população nas zonas urbanas. Devido a um conjunto de fatores, como o rápido aumento da população das cidades, as alterações climáticas e o esgotamento de recursos naturais, têm conduzido as cidades a enfrentarem vários riscos, preocupações e problemas, como a degradação das condições do ar, da terra e da água, riscos económicos e sociais. Todos esses fatores, riscos e problemas levam a uma necessidade urgente das cidades encontrarem soluções inteligentes em prol dos constantes desafios que vão surgindo [27]. Muitas dessas soluções e estratégias inteligentes, para tornarem as cidades mais habitáveis e sustentáveis, passam por diversos parâmetros, tais como proporcionar uma urbanização sustentável, obter uma economia atrativa, fomentar a inovação, melhorar a eficiência, promover a resiliência e estimular a igualdade social. Estas cidades são intituladas por Cidades Inteligentes, em inglês *Smart Cities*.

Giffinger *et al.* destacam, como características de uma Cidade Inteligente, a economia, pessoas, governação, mobilidade, ambiente e qualidade de vida, dando uma importância à utilização das TIC como elemento facilitador da melhoria dessas mesmas características [28].

Embora o conceito de Cidade Inteligente seja bastante discutido, o facto é que ainda não existe um consenso de uma definição genérica. Muitas das vezes o termo é utilizado em circunstâncias diferentes gerando confusões entre as várias perspectivas existentes de Cidade Inteligente [29].

No entanto, o que se tem verificado, progressivamente, é que muitas das práticas associadas ao termo Cidade Inteligente têm sido possíveis e realizáveis devido à utilização de Tecnologias de Informação e Comunicação (TIC). A utilização das TIC tem permitido facultar melhores condições de vida e até favorecer a sustentabilidade.

Neste contexto, várias TIC podem ser aplicadas, tais como a Internet das Coisas (*Internet of Things*); sensores (RFID e ZigBee, por exemplo) e atuadores; sistemas de informação; computação em nuvem; utilização de dispositivos móveis, como *smartphones* e *tablets*, permitindo mobilidade; *Business Intelligence*, como forma de extração, monitorização e análise de dados para apoio à tomada de decisão, entre outros [30].

Relativamente aos setores público e empresarial, atualmente, as empresas procuram soluções revolucionárias de modo a corresponderem às constantes evoluções, como a intensificação da concorrência, o rápido aumento do volume da informação e o desenvolvimento tecnológico, por forma a tornarem-se mais competitivas e inovadoras perante outras.

A importância de tomar decisões fulcrais aliada à demanda de obtenção rápida e confiável de grandes quantidades de informações, em espaços de tempo cada vez mais curtos, num mercado bastante competitivo, leva à necessidade de as empresas implementarem sistemas de informação cada vez mais complexos.

Com esta necessidade, e no âmbito das Tecnologias de Informação, as empresas gradualmente têm adotado e utilizado o modelo de *Business Intelligence* (Negócio Inteligente). Este conceito possibilita às empresas uma gestão de negócios mais eficiente, no qual são realizados um conjunto de processos que facilitam a tomada de decisões estratégicas. Para tal, são utilizadas ferramentas essenciais para reunir, armazenar, manipular, fornecer o acesso e análise a dados de grande volume. Destas ferramentas têm-se destacado as ferramentas *open source* e *software* livre, sobretudo nas pequenas e médias empresas, devido aos seus benefícios, tais como, custo reduzido ou inexistente, atualizações frequentes e possibilidade de integração com outras ferramentas.

De entre muitas práticas de *Business Intelligence*, evidencia-se *Data Warehousing*, que são sistemas que permitem reunir dados, bem como a integração e ma-

nipulação de dados de uma ou mais fontes, criando, assim, repositórios de dados denominados por *Data Warehouses* (DW). O *Data Warehouse* permite armazenar toda a informação referente às atividades de uma empresa ou organização, possibilitando a análise de grandes quantidades de dados. Num *Data Warehouse*, existem unidades lógicas denominadas por *Data Mart*, que são bases de dados que servem para armazenar informações de gestão de setores ou áreas específicas de negócio de uma empresa, como por exemplo, os setores das finanças e o da contabilidade.

Os sistemas utilizados antes do surgimento deste tipo de base de dados, denominados por sistemas OLTP (*Online Transaction Processing*), não permitem efetuar análises complexas de modo a apoiar a tomada de decisões estratégicas. Estes sistemas são do tipo operacional e permitem registrar as transações de negócios que vão ocorrendo no dia-a-dia numa empresa, armazenando os dados de forma detalhada das mesmas, tais como atualizar uma conta bancária e levantamentos de dinheiro. Embora os sistemas OLTP primam pela sua simplicidade e eficiência, apresentam alguns cuidados a ter em conta, como as questões de segurança e os custos, e além disso, o esquema da base de dados não está delineado para o apoio à tomada de decisão.

Uma vez que os sistemas OLTP estão projetados e otimizados para corresponder de forma eficiente e rápida ao processamento de numerosas transações diárias de uma empresa, acabam por desvalorizar outros aspetos, como a capacidade de análise e gestão estratégica. Esta desvalorização conduz a que estes sistemas tornem-se lentos ou mesmo inaptos para o desempenho dessas tarefas de análise e gestão de dados. Assim, neste contexto de *Business Intelligence*, surgiram os sistemas OLAP (*Online Analytical Processing*), que com base na recolha de informação produzida por sistemas OLTP, reorganizam-na permitindo a análise mais eficiente e rápida de um grande volume de informação e através de várias perspetivas distintas [5].

Como o âmbito desta dissertação se depara com a análise de vários dados sobre projetos de Cidades Inteligentes, optou-se por uma implementação de um sistema com base nos sistemas OLAP, no qual se destacam conceitos como: (i) Multidimensionalidade, para a modelação multidimensional da base de dados e (ii) Cubo OLAP, que é uma estrutura para armazenar esses dados multidimensionais, permitindo consultas e análises rápidas da informação nas diferentes perspetivas. Além disso, a realização destas análises tem por base a utilização de ferramentas OLAP *open source* do tipo *suite*, pois conjugam o servidor e cliente OLAP num só pacote, e não apresentam qualquer custo, sendo livre e de código aberto.

A criação da plataforma *online* permite reunir a informação acerca de cada

projeto de Cidade Inteligente. Esta plataforma disponibiliza métodos de análise a essa informação, como gráficos analíticos e um mapa, que possibilitam tirar conclusões acerca dos dados reunidos. A plataforma *online* foi desenvolvida com recurso à tecnologia OLAP, programação *web* e Base de Dados.

1.2 Motivação

A escolha do tema da dissertação teve em consideração dois aspetos importantes: interesse pessoal e enquadramento do tema num contexto atual. Dessa forma, permitir consolidar competências e aptidões úteis num futuro relativamente próximo.

Assim, surgiu a oportunidade de criar uma plataforma *online* universal, no sentido das várias organizações e instituições poderem contribuir com os seus projetos em Cidades Inteligentes, facilitando a procura e análise de iniciativas existentes.

O desenvolvimento da plataforma *online* permite reunir informações sobre práticas em Cidades Inteligentes implementadas por todo o mundo, através de um formulário. Na plataforma são disponibilizados gráficos analíticos e um mapa ilustrativo das localizações das mesmas, elaborados com base nos dados extraídos do formulário *online* através do recurso da tecnologia OLAP. Com isto, é pretendido que a plataforma contribua para tomadas de decisão de gestores políticos e de todas as organizações relacionados com as Cidades Inteligentes.

A recolha das informações de um conjunto de atributos relativos aos vários domínios relacionados com Cidades Inteligentes, bem como a criação de gráficos analíticos permitem analisar os aspetos tidos em conta e os resultados obtidos das diversas iniciativas. Neste sentido, a plataforma permite contribuir para que todos os envolvidos nas iniciativas, possam melhorar determinados aspetos nas práticas existentes, levando, assim, ao aumento da eficiência do desenvolvimento sustentável e socioeconómico.

Além disso, a plataforma *online* pode também auxiliar novas iniciativas, na medida que viabiliza, aos decisores políticos responsáveis e a organizações interessadas, possuírem uma base, um guia, para o estudo, análise e projeção de iniciativas futuras através da obtenção de informações pertinentes. Deste modo, através da exploração das análises e da recolha de informações sobre as práticas existentes, estes poderão se focar em aspetos ainda pouco desenvolvidos. Posto isto, a implementação da plataforma *online* viabilizará aos decisores políticos e responsáveis pelas iniciativas o aperfeiçoamento de determinados aspetos que possam aumentar a eficiência do de-

envolvimento sustentável e socioeconómico, podendo também servir de apoio para a projeção de iniciativas futuras.

1.3 Finalidade e Objetivos

De forma a tornar este projeto exequível foi necessário proceder a dois pontos fundamentais: (i) uma fase teórica, com recurso a uma revisão literária sobre os aspetos que englobaram o tema em questão e (ii) uma fase prática, a qual consistiu na análise, esboço e implementação do sistema para casos de estudo em Cidades Inteligentes.

A implementação da plataforma *online* permite reunir a informação relevante sobre cada iniciativa em Cidade Inteligente através de um formulário baseado num modelo conceptual disponibilizado no projeto “*Smart Cities for Sustainable Development*” [1]. O formulário *online* permite automatizar a inserção de iniciativas na base de dados criada. Além disso, esta plataforma disponibiliza gráficos analíticos e um mapa com base nessas informações reunidas. As análises apresentadas permitem aos gestores e a todas as partes interessadas, relacionados com as iniciativas, analisarem a evolução das diversas características nos principais domínios referentes a uma Cidade Inteligente. Posto isto, a plataforma pretende contribuir para o melhoramento de práticas existentes e/ou servir de suporte para o planeamento e desenvolvimento de novas iniciativas.

Esta plataforma serve de apoio para a identificação dos vários objetivos e da natureza de cada iniciativa para que viabilizem o desenvolvimento sustentável. Ao reunir as principais informações de um conjunto de atributos relativos a uma Cidade Inteligente é esperado que permita analisar quais os aspetos tidos em conta e os resultados obtidos com uma iniciativa. Deste modo, pretende-se que possa contribuir para que os gestores e responsáveis pelas iniciativas possam melhorar determinados pontos, aumentando, assim, a eficiência do desenvolvimento sustentável e socioeconómico. Além disso, esta plataforma *online* pode auxiliar novas iniciativas, na medida que possibilita às organizações e interessados a adquirirem informações essenciais para o estudo e projeção de iniciativas futuras, com a exploração das análises e do modelo conceptual, podendo, assim, dedicarem-se a aspetos ainda pouco desenvolvidos.

Como principais objetivos desta dissertação pretendeu-se a automatização da inserção de dados relativos a práticas em Cidades Inteligentes através de uma plataforma *online*. Inicialmente a base de dados armazena a informação sobre 21 casos

de estudo em Cidades Inteligentes referenciados e analisados no projeto “*Smart Cities for Sustainable Development*” [1]. Também foi proposto a realização de algumas análises estatísticas desses dados, a partir de uma ferramenta OLAP *open source*.

Portanto, o projeto desta dissertação teve como finalidade uma análise, o estudo, a concepção e a implementação de um sistema para casos de estudo em Cidades Inteligentes, com recurso a uma ferramenta OLAP *open source*.

Neste âmbito, de forma a alcançar o intuito deste projeto, foram traçados um conjunto de objetivos, tanto a nível pessoal como a nível do tema abordado, em especial:

- Obtenção de conhecimentos em sistemas *Business Intelligence*, nomeadamente os sistemas OLAP;
- Estudo das metodologias/tecnologias essenciais para o sistema;
- Estudo e esboço geral da arquitetura do sistema a desenvolver e do seu funcionamento, especificando os componentes, ferramentas e tecnologias a utilizar;
- Aprofundamento e desenvolvimento de conhecimentos em programação *web* e em bases de dados;
- Estudo das principais ferramentas OLAP *open source* existentes, bem como a exploração da ferramenta selecionada;
- Elaboração de uma base de dados de acordo com a semântica de modelação multidimensional;
- Implementação de um formulário *online* para reunir e armazenar informação das iniciativas em Cidades Inteligentes na base de dados;
- Elaboração de gráficos analíticos e de um mapa, baseados na informação contida na base de dados, com recurso à ferramenta OLAP *open source* selecionada.
- Disponibilização desses gráficos analíticos e do mapa, na plataforma *online*.

Com a conclusão deste projeto é esperado que o sistema permita reunir, explorar e analisar informações dos mais variados casos de estudo em Cidades Inteligentes, e com isso possa incentivar a concretização de novas práticas e a melhorar as existentes.

1.4 Estrutura da Dissertação

Este documento, que descreve todo o trabalho desenvolvido durante o projeto de dissertação, é composto pelos seguintes 7 capítulos, apresentados sucintamente:

Capítulo 1 - Introdução: este capítulo tem como propósito apresentar a contextualização do tema de dissertação, descrevendo a motivação, a finalidade e os principais objetivos propostos para a sua elaboração.

Capítulo 2 - Estado da Arte: o presente capítulo menciona a revisão literária essencial para a compreensão dos conceitos relacionados com o projeto e necessários para o seu desenvolvimento.

Capítulo 3 - Estudo de ferramentas de BI *open source*: retrata os principais critérios para avaliação de ferramentas OLAP. Em seguida, descreve o conjunto de 5 ferramentas de *Business Intelligence open source* analisadas com o intuito de proceder à seleção e exploração de uma para a elaboração das análises OLAP. Posteriormente, este capítulo termina com uma tabela comparativa dessas ferramentas e de uma revisão literária sobre esse estudo.

Capítulo 4 - Especificação do sistema: este capítulo apresenta os requisitos não funcionais e requisitos funcionais levantados para o sistema. Além disso, são apresentados os casos de uso relativos aos intervenientes do sistema, os principais esboços da plataforma *online* e a arquitetura geral dessa plataforma.

Capítulo 5 - Implementação do sistema: o presente capítulo retrata o processo de desenvolvimento do sistema. Inicialmente, são descritas as ferramentas, tecnologias e linguagens utilizadas para a elaboração do projeto. Seguidamente, são demonstradas a modelação multidimensional dos dados, a criação do Cubo OLAP e a exploração das análises aos dados. Por último, são descritos os principais processos do desenvolvimento da plataforma *online* e da aplicação *web* de administração.

Capítulo 6 - Resultados do sistema: este capítulo apresenta os principais resultados obtidos com a elaboração da plataforma *online*, bem como com o desenvolvimento da aplicação *web* de administração.

Capítulo 7 - Conclusões: neste último capítulo são expostas a síntese e as conclusões de todo o trabalho elaborado. Além disso, apresenta as principais contribuições e algumas propostas de trabalho futuro.

Capítulo 2

Estado da Arte

Neste capítulo é apresentada toda a revisão da literatura que permitiu compreender os principais conceitos, fundamentos e metodologias que serviram de apoio no trabalho desta dissertação, dos quais se destacam o conceito de Cidade Inteligente, de *Business Intelligence* e de OLAP.

Primeiramente, na secção 2.1 é apresentada a noção de Cidades Inteligentes, as principais dimensões interligadas ao conceito e a aplicação das Tecnologias de Informação e de Comunicação (TIC) em Cidades Inteligentes. Também, é mencionada, nesta secção, a importância das iniciativas em Cidades Inteligentes nos dias de hoje para promover a sustentabilidade e a idealização do projeto “*Smart Cities for Sustainable Development*”.

Em seguida, na secção 2.2 é exposto o estado da arte relativamente ao termo de *Business Intelligence*, destacando o seu conceito, a história, arquitetura, técnicas e aplicações. Ainda, nesta secção, é dado a conhecer a definição de *Data Warehousing*, as suas principais técnicas e tecnologias associadas, como ETL, *Data Warehouse* e OLAP.

2.1 Cidades Inteligentes

O crescimento urbano, que se tem verificado ao longo dos anos, conduz a uma urgência em encontrar soluções inteligentes face aos desafios e problemas urbanos que surgem constantemente nas cidades. Existem práticas recentes que tornam as cidades melhores para viver, mais eficientes e sustentáveis, tornando-se casos de sucesso de novas estratégias de desenvolvimento da cidade, sendo esta intitulada por Cidade Inteligente, em inglês, *Smart City* [27]. Além disso, muitas dessas práticas aplicadas

têm sido concretizáveis devido à utilização eficiente das TIC como ferramenta para a construção e integração de infraestruturas e de serviços, que conseqüentemente proporcionam melhor qualidade de vida [30].

A reabilitação urbana inteligente é uma necessidade e uma oportunidade para os países promoverem a sustentabilidade. Porém, é necessária uma estratégia global para dinamizar essa reabilitação urbana no mercado das Cidades Inteligentes, de modo que haja atração de investimento, tanto privado como público.

De acordo com o relatório de pesquisa [31], Navigant Research define Cidade Inteligente como sendo uma cidade que integra tecnologia numa abordagem estratégica para promover a sustentabilidade, o bem-estar dos cidadãos e o desenvolvimento económico.

Hall e os seus colaboradores em [32] referem uma Cidade Inteligente como sendo uma cidade que monitoriza e integra condições das suas infraestruturas críticas, tais como pontes, túneis, metros, aeroportos, portos marítimos, comunicações, água, energia e até mesmo grandes edifícios. Além disso, os autores mencionam que uma Cidade Inteligente pode otimizar melhor os seus recursos, planear as suas atividades de manutenção preventiva e monitorizar os aspetos de segurança, maximizando os serviços dos cidadãos.

Schuurman, Baccarne, Marez e Mechante [33] referem que o conceito de Cidade Inteligente pode ser visto como um reconhecimento da importância crescente das tecnologias digitais, essenciais para uma posição competitiva e um futuro sustentável das cidades.

Em [34] os autores indicam que as Cidades Inteligentes proporcionam, como benefícios, serviços mais adequados para os cidadãos; melhor governação; qualidade de ambientes mais favoráveis; uma indústria mais avançada. E desta forma, mais amiga do ambiente e das pessoas, com uma economia mais dinâmica e inovadora.

Embora o conceito de Cidade Inteligente seja bastante discutido nos últimos anos, o facto é que ainda não existe um consenso de uma definição genérica, sendo este termo, muitas vezes, utilizado em circunstâncias diferentes, gerando significados diferentes. Em consequência disso, tem-se suscitado confusões entre as partes interessadas com entre as várias perspetivas de Cidade Inteligente [29].

Uma vez que não há um acordo relativamente ao conceito, também se verifica que não existe unanimidade sobre as principais áreas ou dimensões que compõem a Cidade Inteligente.

Segundo um modelo europeu, Giffinger *et al.* identificam seis domínios que en-

volvem as Cidades Inteligentes: Economia, Pessoas, Governação, Mobilidade, Meio-Ambiente e Vida. Neste mesmo contexto, os autores referem Cidade Inteligente como sendo uma cidade que apresenta um desempenho adequado, promissor e eficiente desse conjunto de seis domínios, baseado numa combinação inteligente de investimentos e atividades de cidadãos autocríticos, independentes e conscientes [28].

Em [35] a IBM menciona que o crescimento urbano, oferece às cidades maior poder económico, político e tecnológico. Ainda, é referenciado que as cidades são baseadas em seis sistemas centrais interligados funcionando como um só, destacando como domínios as Pessoas, Negócios, Transporte, Comunicação, Água e Energia.

T. Nam e T. A. Pardo em [27] expõem princípios estratégicos em três dimensões: Tecnologia, Pessoas e Instituições.

Embora se destaquem estes domínios, não se pode enumerar uma lista que determine os domínios alvo de uma Cidade Inteligente, pois devem variar conforme o contexto e prioridades de cada cidade [30].

No contexto deste trabalho, o impacto da iniciativa de uma Cidade Inteligente foca-se na sua contribuição perante as dimensões Economia, Pessoas, Governação, Mobilidade, Meio-Ambiente e Vida. A figura 2.1 retrata as características das Cidades Inteligentes mais vezes citadas durante a revisão da literatura.

Segundo Hedlund [12], a mudança de uma cidade em Cidade Inteligente melhora a eficiência, gera um impacto ambiental positivo, aumenta a segurança, melhora a saúde e simplifica as licenças de construção. Esta mudança, conseqüentemente, torna a cidade mais habitável, com uma economia mais viável, atraindo, assim, novos cidadãos e empresas. Deste modo, o autor salienta que uma Cidade Inteligente tem capacidades para satisfazer as mais altas expectativas a nível da sustentabilidade ambiental, social e económica.

Para A. Caragliu, C. Del Bo, e P. Nijkamp, uma cidade é inteligente quando o investimento em capital humano e social, em transportes e em infraestruturas modernas para TIC, incentivam um crescimento económico sustentável e aumentam a qualidade de vida. Sendo isto baseado numa boa gestão dos seus recursos naturais e riquezas, através de um governo participativo [36]. Os autores destacam como características próprias de uma Cidade Inteligente a utilização de infraestruturas de rede para melhorar a eficiência económica, política e social e promover o desenvolvimento urbano, social e cultural. Além disso, os autores evidenciam o foco de uma Cidade Inteligente perante o alcance da inclusão social de vários cidadãos urbanos nos serviços públicos, e a importância do papel da indústria de alta tecnologia e do capital social

para o desenvolvimento urbano. Também, é destacada a capacidade da comunidade de uma Cidade Inteligente em inovar e adaptar-se face à evolução da tecnologia. O desenvolvimento da sustentabilidade social e ambiental constituem, portanto, para os autores, dois importantes componentes estratégicos, requerendo especial atenção por parte de empresas, organizações e governos face às constantes mudanças das cidades e das dinâmicas complexas que apresentam. Desse modo, é essencial as cidades modernas adotarem soluções inteligentes que permitam prosperarem-se.



Figura 2.1: Principais características dos domínios em Cidades Inteligentes

Uma cidade, enquanto Cidade Inteligente, deve incentivar a sustentabilidade no domínio económico, social e ambiental, facultando ao cidadão urbano serviços mais eficazes e inovadores, atendendo a uma boa gestão urbana. Neste contexto, salienta-se o papel essencial das TIC para as iniciativas em Cidades Inteligentes. Contudo,

as iniciativas em Cidades Inteligentes enfrentam diversos desafios, na sua maioria de vertente político e económico.

As iniciativas de tornar uma cidade em Cidade Inteligente surgiram como um modelo para mitigar e retificar os problemas urbanos, fazendo com que as cidades se tornem mais habitáveis e sustentáveis. Por isso, é necessário construir mais Cidades Inteligentes no futuro, de modo a contribuírem para um desenvolvimento mais sustentável, eficiente e considerável da economia, do ambiente e da qualidade de vida dos habitantes urbanos.

2.1.1 A importância das TIC em Cidades Inteligentes

Muitos são os autores que destacam a importância da utilização das Tecnologias de Informação e Comunicação no contexto de Cidade Inteligente, quer em programas, projetos ou iniciativas, como modo de melhorar as infraestruturas e os serviços de uma cidade. E consequentemente, permitirem maximizar a qualidade de vida dos cidadãos, do ambiente e a competitividade económica, promovendo a sustentabilidade entre os vários domínios.

A criação de uma Cidade Inteligente seria bastante dificultada sem a utilização das TIC. No entanto, as TIC apenas constituem um forma de facilitar o desenvolvimento dos vários domínios associados a Cidades Inteligentes [37].

Em 2008, Hollands [29], destaca a necessidade de uma verdadeira mudança no equilíbrio de poder entre a utilização das TIC pelas empresas, governo, comunidades e os cidadãos urbanos. Além disso, o autor menciona a importância de equilibrar o crescimento económico com a sustentabilidade.

Neste mesmo contexto, existem vários exemplos de integração das TIC em Cidades Inteligentes, tais como computação em nuvem, redes de sensores sem fios (WSNs), redes elétricas inteligentes, sistemas geográficos e dispositivos móveis. A utilização das TIC abrange desde a recolha de dados nas ruas para serem analisados, de forma a extrair informação, até níveis mais altos. Dessa forma, a exploração e o emprego das TIC de forma eficiente e otimizada fornecem mecanismos para o auxílio na monitorização da cidade. Por conseguinte, as TIC promovem inovações, e proporcionam suporte em tomadas de decisões e na resolução das adversidades encontradas em iniciativas para Cidades Inteligentes [30].

A Cidade Inteligente difere em vários pontos de uma cidade tradicional, particularmente na centralização no cidadão, dado que este gradualmente exige, aos governos, melhores serviços públicos, mais rápidos e eficientes. Além disso, a Cidade

Inteligente distingue-se no investimento e na gestão equilibrada das TIC, que de alguma forma, permitem menores custos de operações e maiores aproveitamentos de recursos. A tabela 2.1, que se segue, apresenta algumas características comparativas entre uma Cidade Inteligente e uma cidade tradicional, relativamente à vertente tecnológica e humana.

Tabela 2.1: Comparação entre Cidades Tradicionais e Cidades Inteligentes, baseado em [12]

Cidade Tradicional	<i>Smart City</i>
- Compra ou desenvolve aplicações específicas.	- Investe ou reutiliza capacidades comuns das TIC e dos negócios; - Adapta em vez de comprar ou desenvolver aplicações específicas.
- Mudança de serviços lenta e cara; - Serviços pouco flexíveis e raramente atualizados.	- Serviços adaptáveis à mudança, associados a menores custos; - Serviços ágeis, rapidamente atualizados.
- Serviços fáceis de serem utilizados por especialistas; - Serviços pouco amigáveis para o cidadão.	- O cidadão é o principal utilizador; - Serviços simples de serem utilizados, sendo que o utilizador não tem de ser um especialista; - Serviços eletrónicos retiram sobrecarga aos recursos humanos da cidade.
- A integração e reutilização dos dados é dispendiosa.	- O uso e a integração dos dados e das capacidades de negócio são facilitados por arquiteturas de referência e outras <i>guidelines</i> ; - Maior qualidade dos dados.
- Custos de infraestruturas baixos; - Custos das Tecnologias de Informação elevados.	- Custos de infraestruturas elevados; - Custos das Tecnologias de Informação baixos.
- Exclusão dos cidadãos nos processos da cidade.	- Transparência, participação e poder de decisão facultados aos cidadãos.
- Não está preparada para serviços em nuvem.	- Acessível para serviços em nuvem.
- Desperdício de recursos humanos; - Automatização reduzida.	- Bom aproveitamento dos recursos humanos; - Serviços centralizados de apoio ao cidadão, reduzindo os custos das empresas.
- Especialistas efetuam trabalho não especializado.	- Especialistas apenas realizam trabalho especializado, encaminham tarefas não especializadas

A integração das TIC na evolução de uma cidade, além de trazer melhorias

significativas na prestação de serviços, cada vez mais exigidas pelos cidadãos, permite obter um caminho sustentável, mais rápido e inteligente, para o desenvolvimento social, económico e ambiental [38].

A evolução tecnológica, nos dias de hoje, permite aperfeiçoar os vários aspetos nos domínios que envolvem as Cidades Inteligentes, viabilizando, progressos eficientes e benéficos nesses domínios nas cidades. E em consequência, essa evolução proporciona o aumento das potencialidades e dos recursos disponibilizados por essas cidades.

2.1.2 A importância de Iniciativas Sustentáveis em Cidades Inteligentes

As iniciativas em Cidades Inteligentes estão a tornar-se fundamentais para a resolução das adversidades que vão surgindo nas áreas de maior urbanização.

O desenvolvimento sustentável tem sido definido de várias formas. Contudo, tem-se destacado a seguinte definição, muitas vezes citada, do relatório “*Our Common Future* de *World Commission on Environment and Development*”¹: “o desenvolvimento sustentável é o desenvolvimento que satisfaz as necessidades do presente sem comprometer a capacidade das gerações futuras satisfazerem as suas próprias necessidades” [12].

As iniciativas em Cidades Inteligentes pretendem obter um desenvolvimento sustentável através de soluções que integram os aspetos mais essenciais, tais como a energia, a mobilidade, o ambiente, a economia, o transporte, as comunicações, as TIC, a água e os cidadãos.

As iniciativas em Cidades Inteligentes pretendem, assim, melhorar o desempenho sustentável urbano nos vários domínios. Para tal, de entre várias soluções viáveis, procuram servir-se de dados, de informação e das TIC para diversas finalidades. Dessas finalidades, aponta-se como exemplos, para o fornecimento de serviços, aos cidadãos, mais eficientes; para uma monitorização e otimização, mais eficaz, das infraestruturas existentes; para o aumento da colaboração dos agentes económicos; e para o incentivo à inovação de novos modelos de negócio nos setores privados e públicos [39].

É importante que as iniciativas em Cidades Inteligentes promovam a inovação, porém, esta apresenta não só oportunidades, como também acarreta riscos. Deste

¹Comissão que procura unir os países para alcançar o desenvolvimento sustentável em conjunto.

modo, é fundamental estabelecer um esforço para gerir esses riscos associados à inovação [40].

H. Chourabi *et al.* em [41], identificam oito fatores críticos em iniciativas de Cidades Inteligentes: gestão e organização, tecnologia, governação, contexto político, pessoas e comunidades, economia, infraestrutura construída e o meio-ambiente. Estes fatores constituem a base de uma estrutura integrada, que pode ser utilizada para analisar a forma como os governos locais estão a antever as iniciativas.

A sustentabilidade em Cidades Inteligentes, enquanto ambiente, inclui a eficiência na utilização dos recursos, a proteção do ambiente e o equilíbrio dos ecossistemas. Além disso, é abrangido, nessa sustentabilidade, uma gestão eficiente da água, dos resíduos e da energia, o uso de energias renováveis, a construção sustentável e energeticamente eficiente e a mobilidade [42]. No entanto, a sustentabilidade, além de compreender um conjunto de preocupações ambientais, também engloba preocupações sociais e económicas de forma interligada. Na qualidade de sustentabilidade social, é englobado por exemplo, a integração de imigrantes, tornando-os cidadãos produtivos, a educação, a saúde, o bem-estar e a segurança de todos cidadãos [12]. Em sustentabilidade económica, está integrado conceitos como dados abertos, na medida em que os cidadãos e desenvolvedores de *software* possuem acesso a dados públicos. Isto permite, sobretudo aos desenvolvedores de *software*, criarem aplicações úteis reduzindo, assim, custos e recursos das organizações. Esta sustentabilidade económica também inclui a utilização de sistemas *Business Intelligence* para suporte em tomadas de decisões, uma vez que permitem integrar e reunir dados de diversas fontes e obter análises eficientes e rápidas [12].

Em [43], destaca, também, que as TIC apresentam um grande potencial para sustentar a transição das cidades para cidades mais sustentáveis, tanto a nível da gestão dos sistemas urbanos, como no aumento de apoio aos estilos de vida urbanos sustentáveis.

2.1.3 Projeto “*Smart Cities for Sustainable Development*”

O projeto “*Smart Cities for Sustainable Development*” foi desenvolvido pelos investigadores Elsa Estevez, Nuno Vasco Lopes e Tomasz Janowski, com a parceria de *United Nations University* (UNU-EGOV) e *International Development Research Center* (IDRC). Este projeto teve o objetivo de explorar a tese de que Cidades Inteligentes permitem alcançar o desenvolvimento sustentável [1].

As Cidades Inteligentes surgiram como resposta face aos desafios e oportuni-

des criados pela rápida urbanização. Este estudo analisou 876 publicações científicas. Essas publicações analisadas surgiram de recomendações de 51 organizações *think tanks*², de pesquisa independente, de 119 iniciativas concretas em Cidades Inteligentes, e ainda, através de 7 entrevistas a gestores municipais, projetistas e investigadores responsáveis por iniciativas em Cidades Inteligentes.

As Cidades Inteligentes, segundo o relatório deste projeto [1], surgiram como resposta em virtude dos desafios e das oportunidades resultantes da rápida urbanização. Além disso, estas cidades implementam sistemas urbanos inteligentes para servir o desenvolvimento socioeconómico e ecológico, para melhorar a qualidade de vida e abordar as origens da instabilidade social nas zonas citadinas.

Uma Cidade Inteligente Sustentável (*Smart Sustainable City*) é descrita no relatório como um processo de transformação contínua. Sendo este processo baseado no envolvimento e colaboração de partes interessadas e na construção de diferentes tipos de capacidades, como a nível de infraestrutura, humano, institucional e técnico. Dessa forma, permite à Cidade Inteligente fortificar as suas capacidades no aperfeiçoamento da qualidade de vida dos seus cidadãos, proporcionar um ambiente saudável e uma economia competitiva, obter o desenvolvimento social, criar uma infraestrutura eficaz e sustentável e proteger os recursos naturais, entre outros.

Os governos estão a tentar corresponder a esses desafios da rápida urbanização, através de modelos de urbanização, também, assentes em tecnologia digital.

Este estudo permite reconhecer que as iniciativas em Cidades Inteligentes podem ajudar a superar os limites de desenvolvimento urbano tradicional. Porém a transformação em Cidade Inteligentes é, simultaneamente, um processo altamente complexo e desafiador.

Além disso, este estudo permite descobrir uma grande variação de iniciativas em Cidades Inteligentes localizadas por todo o mundo, englobando tanto países desenvolvidos como países em desenvolvimento.

A maioria das iniciativas são implementadas pelos governos, seguindo-se, a indústria e as organizações não governamentais. Este relatório também teve o objetivo de explorar os benefícios, desafios e possíveis caminhos para inovações em Cidades Inteligentes nas diferentes dimensões: governação, economia, vida, mobilidade, pessoas, ambiente, incluindo o ambiente institucional, socioeconómico, político e cultural.

Neste relatório são apresentadas, as conclusões de investigação, a identificação

² *Think tanks* são organizações ou instituições que produzem e expandem conhecimento e estratégias de determinados assuntos, atuando em campos de grupos de interesse.

de instrumentos de política e de ferramentas para iniciativas, bem como a referência de diversas iniciativas e a descrição de alguns casos de estudo. Ainda, é exposta uma proposta de um modelo conceptual a ter em conta para promover os objetivos de desenvolvimento sustentável em Cidades Inteligentes.

Do subconjunto das 119 iniciativas analisadas, nesse projeto, foram selecionados 21 casos de estudo. Essa seleção foi efetuada de acordo com os seguintes critérios: foco, sobretudo, em países em desenvolvimento; riqueza da informação encontrada durante a pesquisa; e distribuição geográfica equilibrada, sendo abrangidas iniciativas de várias regiões do mundo. Os 21 casos de estudo englobam países desenvolvidos e países em desenvolvimento, nos quais 3 são da África, 8 das Américas, 6 da Ásia e 4 da Europa. Os casos de estudo exploram vários aspetos, tais como os benefícios, os desafios, as abordagens e as mudanças, ocasionados pelas Cidades Inteligentes sustentáveis. A informação sobre estes 21 casos de estudo estará armazenada na base de dados da plataforma *online*. Além disso, as análises OLAP a serem apresentadas nesta plataforma, inicialmente, irão refletir-se sobre os dados extraídos dessa informação.

Neste estudo foi desenvolvido um modelo conceptual composto pelos atributos *Technologies* (Tecnologias) e *Tools* (Ferramentas) como *Inputs* (Entradas), que constituem elementos específicos que o processo de transformação de uma iniciativa pode utilizar. Além disso, este modelo é constituído pelos atributos *Approaches* (Abordagens), *Stakeholders* (Partes interessadas), *Governance* (Governança), *Maturity Models* (Modelos de maturidade) como *Transformation* (Transformação). Esta *Transformation* é o processo de construção de diferentes tipos de capacidades para tornar a cidade mais inteligente e com melhor qualidade de vida. Também, é composto pelos atributos *Values* (Valores), *Drivers* (Impulsionadores), *Challenges* (Desafios), *Risks* (Riscos) e *Region* (Região) como *Context* (Contexto), que representam as características específicas da cidade que devem ser consideradas para o desenvolvimento da Cidade Inteligente. Por último, este modelo abrange como *Outcomes* (Resultados), os atributos *Innovations* (Inovações) e *Benefits* (Benefícios), que apresentam os resultados obtidos pelos processos de transformação em Cidade Inteligente.

A figura 2.2, que se segue, retrata o modelo conceptual desenvolvido neste projeto, apresentando esses atributos relativos a cada fase no processo de transformação de uma cidade em Cidade Inteligente Sustentável.

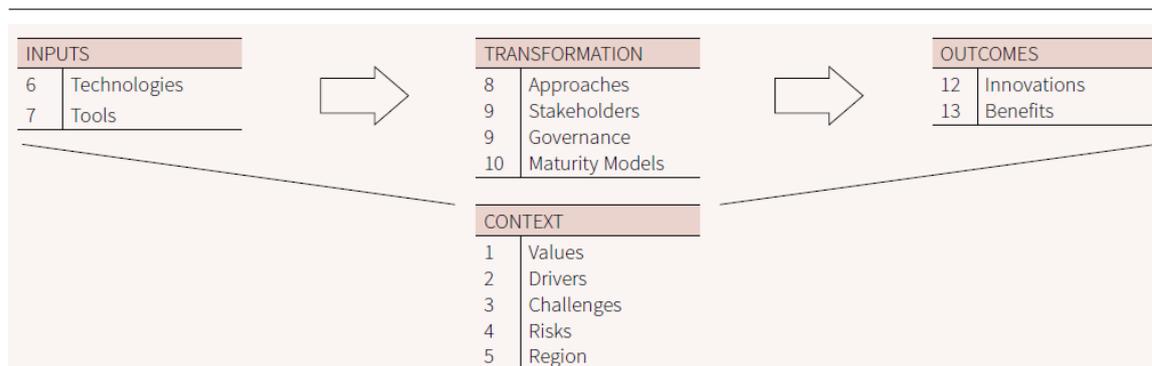


Figura 2.2: Modelo conceitual para o Desenvolvimento Sustentável em Cidades Inteligentes, retirado de [1]

Ao longo desse projeto foi desenvolvido um documento, denominado por “*Smart Cities for Sustainable Development - State of Practice Survey*”, baseado no modelo conceitual. Este documento tem a finalidade de obter descrições uniformes das diferentes iniciativas para o desenvolvimento sustentável em Cidades Inteligentes. É com base neste modelo que será desenvolvido o formulário *online*. A figura 2.3 seguinte apresenta uma parte desse documento que servirá de base para a elaboração do formulário *online*. Nesta imagem é possível observar uma breve explicação da estrutura desse documento, bem como alguns atributos para obtenção de informação relativa aos casos de estudo.

Smart Cities for Sustainable Development – State of Practice Survey		  UNITED NATIONS UNIVERSITY UNU-EGOV <small>Operating Unit on Policy-Driven Electronic Governance</small>	
CASE			
INITIATIVE			
COUNTRY			
INTRODUCTION			
<p>This document is the result of a survey, which aims at identifying best practices in the areas related to Smart Cities for Sustainable Development. Of interest are Smart City initiatives – at national, local or regional level, contributing to sustainable development.</p> <p>In order to obtain uniform descriptions of different programs, the survey relies on a conceptual framework for Smart Cities for Sustainable Development (SCV4SD) with five questions applied to each practice:</p> <ol style="list-style-type: none"> 1. Who developed the practice? The question determines the institution developing the practice as well as its partners. 2. Where the practice has been developed? The question determines the country, state/province, city or region where the practice has been developed as well as the time frame of its development. 3. What is the practice about? The question determines the aim of the practice and summarizes its main features and lessons learnt. 4. Why the practice is relevant? The question determines the impact of the practice in terms of its contribution to the social, economic or environment dimensions. 5. How the practice is delivered? The question determines the mechanism used for delivering the practice, such as policy instruments, government initiatives, tools, governance mechanisms and others. <p>This document is structured into the corresponding five sections – who, where, what, why and how; preceded by information: 1) About the practice – name and case; 2) About the survey – major sources used to carry out the survey with URLs and descriptions, the most representative highlights used to describe the practice and any auxiliary comments; and 3) About the surveyor – who carried out the survey and when.</p>			
PRACTICE			
Name			
Case			
SURVEY			
Sources	No	Description	Reference
	1		
	2		
Highlights			
Comments			
SURVEYOR			
Who			
When			

Figura 2.3: Secção do documento “*Smart Cities for Development - State of Practice Survey*”, retirado de [1]

2.2 *Business Intelligence*

Atualmente, no mundo empresarial, a competitividade crescente juntamente com as constantes evoluções, conduzem a grandes pressões nas empresas. Desta forma, têm exigido às empresas a aplicação de soluções rápidas e eficientes, em curtos períodos de tempo, fazendo com que as tornem mais competitivas e inovadoras.

O aumento da necessidade de transformar dados, provenientes de várias fontes e sistemas, em informação útil de forma lépida e confiável, para auxílio nas tomadas de decisão, conduz à implementação de sistemas mais complexos, capazes de suportar essas transformações. Os sistemas que permitem às empresas uma gestão de negócios mais eficiente, facultando análises rápidas e eficazes de grandes volumes de informação, são apontados como essenciais para o sucesso, inovação e competitividade das mesmas.

Para tal, é necessário utilizar sistemas e tecnologias de informação e da comunicação de forma a permitirem abordagens rápidas e eficientes dos dados fundamentais para tomar decisões relevantes. É neste contexto que progressivamente o modelo de *Business Intelligence* tem sido amplamente utilizado pelas empresas, com o objetivo de fornecer suporte na gestão de negócios.

A utilização das TIC, também, tem vindo a desempenhar um papel cada vez mais importante numa era tão tecnológica. Estas têm gerado um enorme impacto na produtividade e inovação, favorecendo a criatividade, a gestão e a modernização de serviços públicos, como saúde, educação e transporte. Além disso, as TIC têm permitido avanços significativos na ciência e na tecnologia, assegurando a cooperação e o acesso à informação [24].

O termo *Business Intelligence* foi marcado pela primeira vez, por Hans Peter Luhn, investigador da IBM, em 1958. Nesta altura o investigador propôs um sistema automático, baseado em máquinas de processamento de dados, que resume e codifica automaticamente os documentos das organizações [44, 45]. Luhn em [45] mostra a sua preocupação na comunicação, mais especificamente, na dispersão da informação nas empresas. O autor salienta, ainda, a necessidade de métodos eficientes para quantidades de informação cada vez maiores e de tomadas de decisões mais rápidas. Para Luhn [45] o termo *Business* é um conjunto de atividades que levam a um objetivo a nível, de por exemplo, da ciência, da tecnologia, da indústria e do comércio. O investigador define, ainda, *Intelligence*, referindo que é a capacidade em aprender as relações entre os factos, de tal forma a orientar ações para atingir o objetivo desejado.

No entanto, o prosseguimento destes sistemas era inviável e distantes da realidade, face ao desenvolvimento dos computadores e documentos naquela época.

Entretanto, com a evolução da tecnologia, nos anos 90, o desenvolvimento de sistemas de *Business Intelligence* voltam a ser evidenciados, pelo Gartner Group³. Esta empresa definiu *Business Intelligence* como sendo um termo genérico que inclui aplicações, arquiteturas e ferramentas que permitem o acesso e análise de informações para melhorar e otimizar decisões e desempenho das organizações. [46, 44].

Business Intelligence, na teoria e na prática, refere-se a um conjunto de conceitos, métodos e recursos tecnológicos para a recolha, processamento e análise de informações. Assim, BI permite às empresas fornecer informações exatas, para que possam avaliar os resultados e interpretá-los, com o propósito de apoiar e melhorar a capacidade de tomada de decisão [47, 48]. Os sistemas de BI são fundamentais para se obter um nível de decisão bastante eficiente, aumentando a produtividade das empresas.

M. Golfarelli, S. Rizzi e I. Cella definem BI como um processo que transforma dados em informação e, em seguida, em conhecimento [49].

Um sistema de BI converte dados em informações úteis, que por meio da análise humana, permitem obter fundamentos baseados no conhecimento existente numa organização, viabilizando melhores resultados. De entre muitas funções que os sistemas de *Business Intelligence* permitem, destacam-se as seguintes [50]:

- Analisar grandes quantidades de dados;
- Criar previsões e estimativas direcionadas para o futuro, assentes em dados históricos e até mesmo atuais;
- Analisar os impactos de mudanças e cenários alternativos;
- Permitir o acesso *ad-hoc* dos dados para responder a questões específicas e não rotineiras;
- Analisar as atividades de uma organização, permitindo obter um conhecimento melhor e mais profundo da mesma.

Em torno deste conceito, é necessário que cada empresa adote uma estratégia para a escolha de um sistema de BI, de acordo com as suas necessidades [47]. Um

³Empresa de consultoria e de pesquisa em Tecnologia da Informação.

sistema de BI difere consoante a finalidade e os objetivos que a organização pretende alcançar. Na definição de soluções de BI numa organização, o foco deve constar numa análise funcional, num desenho (*design*) da solução e nos resultados ou informações [47]. O objetivo geral dos mais bem sucedidos sistemas de BI é ser destinado a qualquer utilizador, independentemente das suas competências tecnológicas. E desse modo, permitir que os utilizadores tenham à sua disposição a informação essencial ocultada em grandes quantidades de dados, por forma a tomarem decisões benéficas e importantes para a organização [47].

Para conseguir um sistema eficiente é necessário ter em conta uma arquitetura que possa suportar eficazmente todos os sistemas que englobam *Business Intelligence*. Durante a revisão da literatura verificou-se diversas arquiteturas de sistemas de *Business Intelligence*, que variavam consoante o autor e o modo como se enquadravam nos respetivos argumentos/temas [2, 50, 51, 52].

A figura 2.4, que se segue, expõe um esquema de uma arquitetura geral de um sistema de *Business Intelligence*. Esta arquitetura é fundamentada nas várias arquiteturas e conceitos analisados durante a revisão literária, nomeadamente na arquitetura demonstrada em [2].

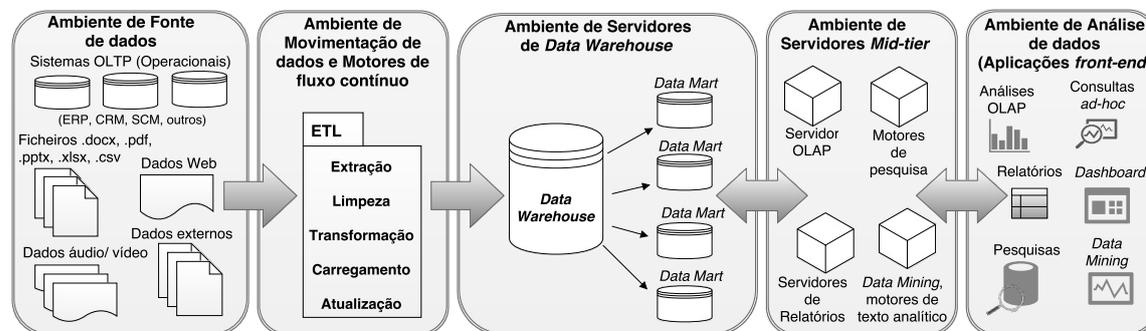


Figura 2.4: Arquitetura geral de um sistema de *Business Intelligence*, baseado em [2]

A arquitetura encontra-se dividida em cinco camadas, tal como se pode observar na imagem anterior, sendo possível, também, visualizar cada ambiente constituinte de um sistema de BI, bem como as tecnologias associadas a cada um. Essas cinco camadas são as seguintes:

- Na primeira camada situa-se o **Ambiente de fontes dos dados**. Estas fontes de dados podem ser internas e/ou externas à organização, e através das quais

são extraídos todos os dados que irão sustentar o sistema de *Business Intelligence*. Os dados, na sua gênese, podem estar ou não estruturados, ou seja, podem ser originados a partir de sistemas de informação com dados organizados e normalizados e podem, também, serem obtidos através de várias fontes como texto puro, vídeo e outros formatos. Exemplos de fontes de dados são os seguintes: folhas de cálculo, *websites*, dados históricos, bases de dados operacionais, bases de dados relacionais, sistemas de transação como ERP (*Enterprise Resources Planning*) e CRM (*Customer Relationships Management*);

- A camada seguinte é constituída pelo **Ambiente de movimentação de dados e motores de fluxo contínuo**, no qual é realizado o processo de ETL (*Extract Transform Load*). O ETL é o processo de extração, transformação e carregamento dos dados, oriundos das diversas fontes, nos repositórios de armazenamento adequados;
- A próxima camada é denominada por **Ambiente de servidores de *Data Warehouse***, o qual inclui os repositórios de armazenamento de dados organizacionais, *Data Warehouse* e *Data Marts*. A escolha mais usual de motores para armazenar e consultar dados em *Data Warehouses* são os Sistemas de Gestão de Base de Dados Relacionais (SGBD) (em inglês, *Relational Database Management Systems, RDBMS*) [5];
- A camada que se segue é formada pelo **Ambiente de servidores *mid-tier***. Esta camada é composta por servidores que permitem a exploração e análise multidimensional dos dados armazenados no DW e DM, através de cubos multidimensionais. Estes cubos, também denominados por cubos OLAP, são estruturas de dados que fornecem suporte a análises eficientes, através de factos e dimensões associadas. Esta exploração e análise de grandes volumes de dados é realizada através da utilização de várias técnicas como OLAP e *Data Mining*;
- A última camada é constituída pelo **Ambiente de Análise de dados**, que abrange as aplicações de *front-end*. Estas aplicações analíticas permitem o acesso, a manipulação e a análise das informações obtidas. Estas análises podem permitir descobrir e prever padrões ou tendências. As informações são disponibilizadas em diversos formatos, tais como relatórios, gráficos, consultas *ad-hoc* e *dashboards*.

De acordo com o contexto desta dissertação as camadas exploradas são os ambientes de fonte de dados, o ambiente de servidores de *Data Warehouse*, o ambiente de *mid-tier*, mais precisamente os servidores OLAP, e o ambiente que engloba as aplicações de *front-end*.

As ferramentas de BI são vistas como tecnologia que permitem a eficiência de operações de negócios, fornecendo mais valias às informações que cada vez são maiores, e à forma como os dados são utilizados. O objetivo final de sistemas de *Business Intelligence* é tornar a informação disponível com melhor qualidade e mais oportuna, podendo levar as organizações a tomarem os melhores rumos e decisões. As pessoas envolvidas em processos de *Business Intelligence* podem usar a aplicação e outras tecnologias para recolher, armazenar, aceder e analisar os dados, de forma simples e útil [53].

O desenvolvimento das tecnologias de informação e comunicação tem aumentado a importância de sistemas de *Business Intelligence* nos processos de tomada de decisão [54].

Vários são os conceitos que envolvem sistemas de BI, tal como: (1) ferramentas ETL para extração, transformação e carregamento de dados, (2) *Data Warehouse*, (3) modelação multidimensional, (4) *Data Warehouse*, (5) Análise OLAP e (6) *Data Mining* [53].

2.2.1 Processo de ETL (*Extract, Transform and Load*)

O processo de Extração, Transformação e Carregamento, denominado por ETL engloba um conjunto de procedimentos. Inicialmente, é realizada uma importante recolha de dados das bases de dados transacionais, operacionais e de fontes externas, sendo transformados num padrão, através de processos de limpeza, tratamento e classificação. Em seguida, esses dados transformados são carregados em repositórios que armazenam os dados (*Data Warehouse* ou *Data Marts*) [44]. Este processo é fundamental quando existem, nas organizações, várias fontes de dados, em diversos formatos. A implementação deste processo consiste, assim, em definir as fontes de dados para se proceder à extração dos dados considerados mais importantes. Por conseguinte, é efetuada a transformação e limpeza desses dados, os quais são devidamente verificados, modelados e reorganizados. Estes dados após serem carregados nos *Data Warehouses* e *Data Marts* estão disponibilizados para serem acedidos rapidamente e facilmente [55]. O ETL é responsável pela integração, agregação e qualidade das informações provenientes de diversos tipos de fontes até ao armazenamento em repo-

sitórios para esse efeito. Além disso, este processo ocorre na fase inicial de um sistema de BI/*Data Warehouse*, sendo as atualizações dos dados realizadas periodicamente, de acordo com o volume de dados e o processamento envolvido.

2.2.2 *Data Warehouse*

O *Data Warehouse* é um conjunto de tecnologias de suporte à decisão, destinado a permitir aos utilizadores das organizações, como gestores e analistas, fazerem melhores e mais rápidas decisões [5]

Um *Data Warehouse*, segundo Inmon, considerado o mentor deste conceito, é “uma coleção de dados organizados por assunto, integrados, variantes no tempo e não voláteis, que suporta os processos de tomada de decisão dos gestores” [56]. Seguindo esta descrição de Inmon, as características de um *Data Warehouse* são [3]:

- **Organizado por assunto:** O *Data Warehouse* pode ser utilizado para a análise de dados de um determinado assunto, uma vez que os mesmos estão armazenados e organizados de acordo assuntos e temas específicos importantes para o negócio da organização;
- **Integrado:** O *Data Warehouse* seleciona, integra e armazena os vários dados provenientes das diversas fontes de dados, tais como SGBD relacionais, ficheiros e registos OLTP;
- **Variante no tempo:** O *Data Warehouse* preserva os dados históricos, bem como informações atuais, de modo a fornecerem informações a partir de uma perspetiva histórica;
- **Não volátil:** Uma vez que os dados estão armazenados no *Data Warehouse*, estes não são alterados.

De acordo com Ralph Kimball, um *Data Warehouse* é uma cópia de dados de transação especificamente estruturados para pesquisas e análises [57]. Esta definição está direcionado para uma visão funcional, um vez que Kimball centra-se na funcionalidade de um *Data Warehouse* ao contrário de Inmon que aborda a forma como é elaborado [58].

R. Kimball, L. Reeves, M. Ross e W. Thornthwaite em [59], referem que um dos recursos mais importantes de uma organização é a informação. Esta informação é mantida por meio de dois sistemas com necessidades diferentes: sistemas operacionais,

em que os dados são colocados, e o *Data Warehouse*, local que contém os dados de saída. Para os autores os objetivos fundamentais, em termos funcionais, de um *Data Warehouse* são:

- Permitir que a informação de uma organização seja facilmente acessível, com bom desempenho e num curto espaço de tempo;
- Tornar a informação de uma organização consistente, ou seja, informação com alta qualidade de modo que a mesma seja contabilizada e completa. As informações de uma parte da organização podem ser combinadas com informações de outra parte da organização, desde que seja garantida a integridade semântica;
- Consistir numa fonte de informações adaptativa e resiliente à mudança. O *Data Warehouse* deve ser projetado para suportar mudanças contínuas, no qual os dados e tecnologias existentes não deverão ser alterados ou interrompidos;
- Constituir um suporte seguro que protege as informações ativas. O *Data Warehouse* deve controlar o acesso aos dados de forma eficaz, como também permitir grande visibilidade sobre a utilização dos mesmos;
- Servir de apoio para uma melhor tomada de decisões. O *Data Warehouse* apresenta dados corretos que facilitam a tomada de decisão.

A arquitetura de um sistema de *Data Warehouse*, geralmente, é uma arquitetura, denominada por *three-tier architecture* [3, 4], ou seja, uma arquitetura cliente-servidor de três camadas. Estas camadas são desenvolvidas e mantidas como módulos independentes em plataformas separadas, tal como demonstra a figura 2.5.

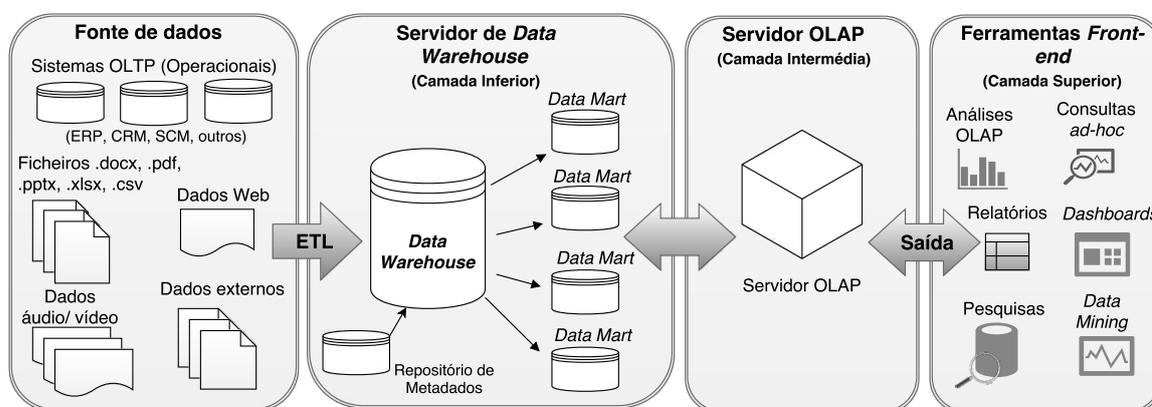


Figura 2.5: Arquitetura geral de um sistema de *Data Warehouse*, baseado em [3, 4, 5]

A arquitetura do *Data Warehouse* tem a finalidade de consolidar, organizar e acessar a grandes quantidades de dados para análises complexas, as quais podem ser utilizadas para auxílio de tomadas de decisão.

As camadas constituintes de um sistema de *Data Warehouse* vão de encontro com o modelo de *Business Intelligence*. Essas camadas são as seguintes:

- A primeira camada, imediatamente a seguir à camada da Fonte de dados, é o servidor de *Data Warehouse*, que normalmente utiliza um sistema de base de dados relacional. Os dados, que têm origem em bases de dados operacionais e em outras fontes de dados, são, sempre que necessário, extraídos, transformados e carregados no servidor de base de dados do *Data Warehouse*, através do processo de ETL. Esta camada pode conter, também, um repositório de metadados, para armazenar informação sobre o *Data Warehouse* e o seu conteúdo.
- A camada seguinte é composta por um servidor OLAP. Este servidor permite explorar os dados através de cubos multidimensionais. As principais arquiteturas de um servidor OLAP são o ROLAP (*Relational OLAP*), MOLAP (*Multi-dimensional OLAP*) e HOLAP (*Hybrid OLAP*).
- A camada superior é um cliente *front-end* que contém as ferramentas necessárias para consulta, pesquisa, exploração, mineração e análise dos dados.

O conceito de *Data Warehouse*, tal como se pode visualizar na figura anterior, engloba dois componentes principais, o repositório de dados e o repositório de metadados [60]. Os repositórios de dados, geralmente, referidos como *Data Warehouses*, são um conjunto lógico de informações integradas, que foram recolhidas a partir de várias fontes de dados numa organização [60]. Metadados, na sua definição simples, são dados sobre dados. Neste contexto, metadados são um conjunto de regras e orientações que orientam a extração, transformação, limpeza e carregamento de dados no *Data Warehouse* [60].

Um *Data Mart* (DM) é um pequeno repositório de dados, que contém um subconjunto de dados de uma organização, ou seja, dados relativos a um determinado departamento ou área de negócio. Este pequeno repositório de dados armazena dados que tendem a ser resumidos. Os *Data Marts* podem ser classificados como independentes ou dependentes, dependendo da fonte de dados que os sustenta. Assim, *Data Marts* independentes são sistemas completos que integram dados das diversas fontes operacionais e/ou externas, dentro num departamento específico numa organização.

Por sua vez, os *Data Marts* dependentes carregam e armazenam dados provenientes do *Data Warehouse* central da organização [15].

Por norma, o *Data Warehouse* é mantido separadamente das bases de dados operacionais da organização, por diversas razões. O sistema de *Data Warehouse* é normalmente apresentado como analítico, pois suporta OLAP. Os requisitos funcionais e de desempenho de sistemas OLAP são muito diferentes dos do sistema OLTP tradicionalmente apoiados por bases de dados operacionais [5]. Além disso, estes sistemas OLAP, focados em análises de dados e relatórios, são direcionados aos executivos, gestores e analistas de negócio para apoiar as tomadas de decisão, nos quais, a informação engloba dados históricos e atuais, resumidos e consolidados.

Os sistemas OLTP normalmente automatizam as tarefas de processamento de dados administrativos, ou seja, registam as transações que ocorrem no dia-a-dia de uma organização, como transações bancárias e pedidos de encomendas. As transações são estruturadas, repetitivas e consistem em transações curtas, atômicas e isoladas. Estas transações incluem operações de inserção, atualização, modificação, eliminação, leitura e detalhe de dados [5]. A estrutura de sistemas OLTP não são adequados para responder a consultas complexas, de grandes volumes de dados [61].

A tabela 2.2 mostra a comparação entre as características de sistemas OLTP relativamente a sistemas OLAP, para uma melhor compreensão das suas diferenças.

Tabela 2.2: Comparação entre Sistemas OLTP e Sistemas OLAP, baseado em [13, 14, 15, 16]

Caraterísticas	Sistemas OLTP	Sistemas OLAP
Fonte de dados	<ul style="list-style-type: none"> - Dados atuais em formatos muito detalhados; - Dados operacionais; - São as fontes originais dos dados. 	<ul style="list-style-type: none"> - Dados históricos e arquivados; - Consolidação (agregação) dos dados; - Os dados surgem de vários sistemas OLTP.
Orientado a	<ul style="list-style-type: none"> - Transações. 	<ul style="list-style-type: none"> - Análises.
Finalidade dos dados	<ul style="list-style-type: none"> - Controlar e executar tarefas importantes dos negócios. 	<ul style="list-style-type: none"> - Auxiliar no planeamento, resolução de problemas e suporte de apoio à tomada de decisão.
Foco	<ul style="list-style-type: none"> - Entrada de dados; - Atualização de dados existentes. 	<ul style="list-style-type: none"> - Saída de informações; - Relatórios e recuperação de informações.

Aplicações	- Serviços <i>web</i> , de gestão, operacionais e cliente-servidor.	- Indústrias de produção, de retalho, serviços financeiros, telecomunicações, serviços públicos e de saúde.
Utilizadores	- Utilizadores comuns, funcionários das organizações, administrador de base de dados	- Gestores, executivos, analistas, cientistas, profissionais de <i>marketing</i>
Tipos de tarefas de dados	- Tarefas operacionais e de negócios	- Elaboração de relatórios e análise de dados
Atualização de dados	- Operações curtas e rápidas de inserir, atualizar e eliminar	- Atualizações de grandes quantidades de dados; - Atualizações demoradas e periódicas
Modelo de dados	- Modelo de Entidade-Relacionamento (ER)	- Modelo de dados multidimensional
Esquema da base de dados	- Esquemas normalizados. - Muitas tabelas e relacionamentos	- Esquemas mais utilizados: Estrela (<i>Star</i>), Floco de Neve (<i>Snowflake</i>) e Constelação (<i>Constellation</i>). - Usualmente menos tabelas e não normalizadas
Consultas	- Consultas simples e mais rápidas, retornando os resultados esperados para a atividade do sistema.	- Consultas complexas dos dados, a fim de agregar as informações envolvidas.
Cópia de segurança	- <i>Backups</i> regulares e completos	- <i>Backups</i> simples ou recarregamento dos mecanismos que suportam a inserção dos dados
Intervalo de tempo	- Dia-a-dia, semanas, meses	- Dados de tempos longos
Velocidade de processamento	- Rápido	- Lento, no entanto, depende da quantidade de dados
Espaço de armazenamento	- Normalmente espaços de armazenamento de dados operacionais pequenos	- Necessidade de grande espaço de armazenamento, para grandes conjuntos de dados de informações históricas
Acessos a dados	- Leitura, escrita, eliminação de dados	- Normalmente, leitura de dados

2.2.2.1 Abordagens para o desenvolvimento de *Data Warehouse*

O sucesso do desenvolvimento de um *Data Warehouse* depende, principalmente, da abordagem e estratégia definidas atendendo ao propósito e às necessidades da organização. Existem várias abordagens e arquiteturas que se podem adotar no desenvolvimento e implementação de um *Data Warehouse*. As duas abordagens mais utilizadas na projeção de um *Data Warehouse* são: *top-down* e *bottom-up*.

A abordagem *top-down*, apoiada por Inmon, é realizada de cima para baixo, isto é, numa primeira fase são abordadas a definição e a construção do *Data Warehouse* organizacional. E numa segunda fase, desta abordagem, é efetuada a implementação dos diversos *Data Marts* existentes para serem carregados com os dados. Assim, os *Data Marts* são criados a partir do *Data Warehouse* da organização. Inmon em [56] afirma que o *Data Warehouse* é a fonte de dados de todos os *Data Marts*. Cada *Data Mart* de um departamento da organização será implementado de acordo com o esquema projetado para o *Data Warehouse* [62].

Embora *top-down* permita uma manutenção mais fácil, o facto é que a implementação desta abordagem é mais complexa. Esta abordagem permite uma manutenção mais fácil, uma vez que os *Data Marts* têm origem a partir do mesmo *Data Warehouse*. Isto conduz à necessidade de somente um conjunto de ferramentas para o processo de ETL e para a integração dos dados. No entanto, *top-down* é uma abordagem que implica uma implementação mais complexa, devido à dificuldade na realização de um esquema geral que englobe toda a informação da organização. Posto isto, a implementação deste tipo de abordagem é mais demorada, apresentando uma taxa de risco mais elevado, pois não existem garantias para o investimento significativo neste tipo de ambiente [63]. Além disso, esta metodologia é inflexível à evolução das necessidades dos departamentos da organização durante a fase de implementação.

A figura 2.6, que se segue, apresenta a abordagem *top-down* proposta por Inmon, na qual primeiro é definido o esquema do *Data Warehouse*, e posteriormente, realizada a implementação de cada um dos *Data Marts*.

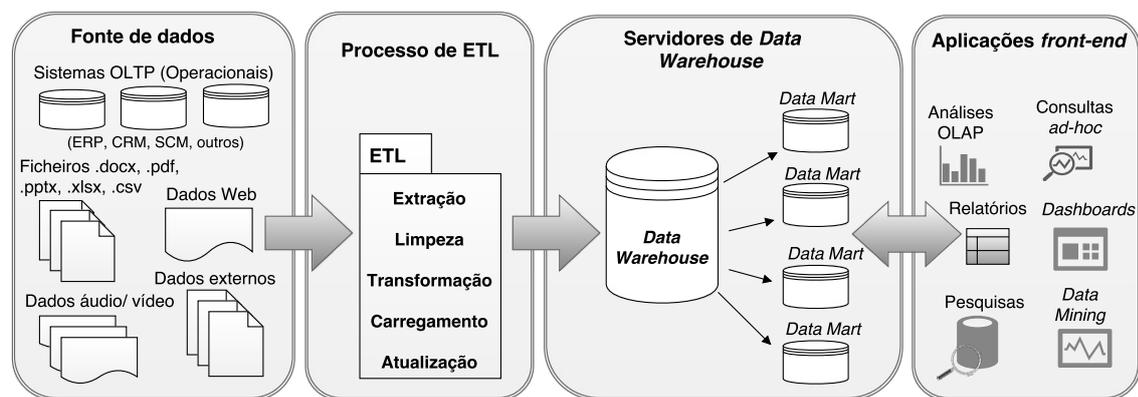


Figura 2.6: Abordagem *top-down* da arquitetura de um sistema de *Data Warehouse*

A abordagem *bottom-up*, apoiada por Kimball, tem como objetivo a implementação de um *Data Warehouse* organizacional incremental a partir do desenvolvimento de vários *Data Marts* independentes. Cada *Data Mart* constitui uma unidade de negócio de cada departamento da organização [62].

Esta abordagem permite que o planeamento e a implementação dos vários *Data Marts* sejam realizados antes da definição da infraestrutura do *Data Warehouse*.

Bottom-up apresenta um desenvolvimento mais simples e rápido, com menores custos, e resultados em pouco tempo, uma vez que a estrutura pode ser utilizada pela organização mesmo antes da finalização da construção do *Data Warehouse*. Esta abordagem permite que um *Data Warehouse* seja expandido facilmente de modo a conter novas unidades de negócio, bastando criar novos *Data Marts* e integrá-los com os outros. Contudo, a integração de *Data Marts* independentes torna-se complexa numa estrutura muito variável [62]. A abordagem *Bottom-up* exige mais esforços e recursos na implementação e uma maior necessidade no controlo dos negócios da organização, de modo a evitar muitos conjuntos de dados que dificultam integrações futuras [63].

A figura 2.7, que se segue, mostra a abordagem *bottom-up* proposta por Kimball, no qual é implementado, em primeiro, cada um dos *Data Marts* independentes, que conduzem à construção de um *Data Warehouse*.

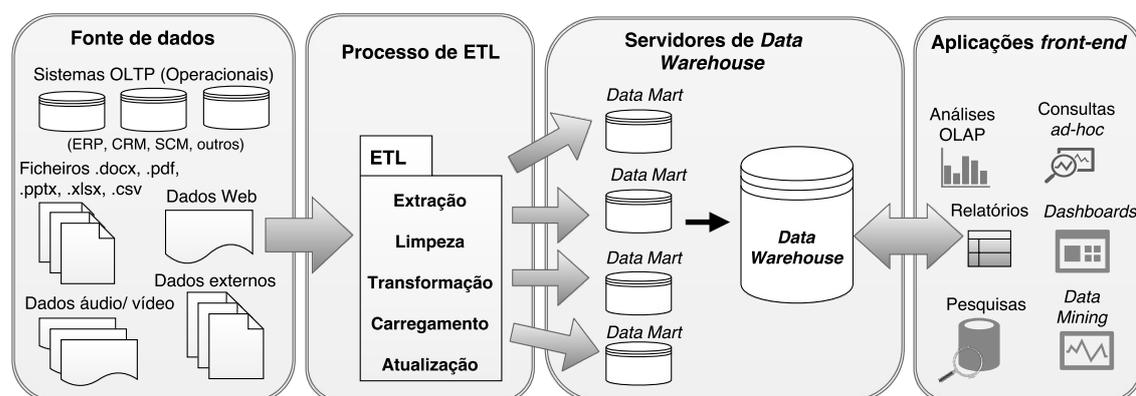


Figura 2.7: Abordagem *bottom-up* da arquitetura de um sistema de *Data Warehouse*

2.2.3 Modelação de dados

A técnica de modelação de Entidade-Relacionamento é adotada para modelação de dados em sistemas operacionais, conhecidos por sistemas OLTP. Esta técnica baseia-se em diagramas para descrever as entidades envolvidas num determinado domínio organizacional, os atributos (caraterísticas) e a forma como se relacionam entre si (relacionamentos). Este modelo destina-se, sobretudo, a sistemas transacionais, não sendo adequado a sistemas analíticos devido às caraterísticas de um *Data Warehouse*. O modelo de Entidade-Relacionamento evita a redundância dos dados, pelo facto de utilizarem a normalização das tabelas, permitindo uma otimização do processamento de atualização de dados [61]. Em contrapartida, um *Data Warehouse* por ser um repositório de dados não-volátil, faz com que os dados sejam redundantes. No entanto, essa redundância otimiza o desempenho, permitindo a publicação e exploração de grandes quantidades de dados para que a organização possa analisá-los eficientemente [64].

O modelo de base de dados assente em diagramas de Entidade-Relacionamento é normalmente eficiente para sistemas que manipulam dados operacionais. Porém, não é o mais apropriado para sistemas de apoio à decisão (SAD), onde a eficiência na consulta e no carregamento de dados são fundamentais. Desse modo, um *Data Warehouse* requer um esquema que facilite a análise de dados [5, 3]. De salientar que, embora a modelação relacional não seja a mais conveniente, o modelo de dados multidimensional pode derivar de esquemas existentes que utilizam o modelo de dados Entidade-Relacionamento [65].

A estrutura do conteúdo de uma base de dados operacional, suportada por siste-

mas OLTP, é apresentada de modo diferente de um *Data Warehouse*. Enquanto as bases de dados de ambientes OLTP se baseiam em modelos de Entidade-Relacionamento e em técnicas de normalização [5], os *Data Warehouses* recorrem, normalmente, a dois tipos de modelo de dados: modelo relacional e modelo multidimensional [66]. O modelo multidimensional pode ser aplicado tanto em bases de dados relacionais como multidimensionais.

Neste sentido, normalmente, num *Data Warehouse* os dados são modelados com base na modelação multidimensional. Esta modelação permite proporcionar pesquisas, análises e visualizações dos dados sob diversas perspetivas, mais rápidas e eficientes. Além do mais, apresenta a informação de uma forma simples e fácil de entender. Para tal, é necessário definir um modelo de dados que consiste nos seguintes principais componentes: factos, dimensões, medidas, atributos e hierarquias. Estes modelos de dados são compostos por tabelas de factos e tabelas de dimensões.

A **tabela de factos** é formada pelas chaves estrangeiras originadas das tabelas de dimensões associadas e por medidas e/ou indicadores calculados no processo de transformação [47]. Os factos correspondem, geralmente, ao eventos que ocorrem dinamicamente numa organização [65]. As **medidas**, também denominadas por métricas, são os atributos numéricos que representam um facto. Cada medida é obtida da interseção de todas as dimensões, ou seja, pela combinação das tabelas de dimensão envolvidas num facto. As medidas são os valores que os utilizadores vão visualizar durante a análise dos dados [61]. Estas medidas podem ser referidas como medidas aditivas, medidas semi-aditivas e medidas não aditivas. As medidas aditivas são aquelas em que o operador soma pode ser utilizado para agregar valores de atributos ao longo de todas as dimensões, incluindo hierarquias. As medidas semi-aditivas, podem ser somadas através de apenas uma parte das dimensões. E as medidas são denominadas por não aditivas no caso de não poderem ser adicionadas por alguma das dimensões [65].

A **tabela de Dimensão** é uma unidade de análise que agrupa dados de negócio relacionados, ou seja, é descrita por um conjunto de atributos textuais. Os **atributos** de uma dimensão podem estar relacionados por meio de uma relação hierárquica [5]. Por exemplo, a dimensão Produto poder ter uma relação hierárquica com a tabela Categoria. Um **Membro** é um subconjunto de uma dimensão, que pertence a um nível hierárquico apropriado. Exemplo disso é uma dimensão ‘Tempo’, constituída por um nível hierárquico ‘Ano’ com os membros ‘1º Trimestre’ e ‘2º Trimestre’, sendo o primeiro composto por membros como ‘Janeiro’ e ‘Fevereiro’. As tabelas

de dimensão são tabelas utilizadas para visualizar as medidas em diferentes ângulos [61]. A dimensão também determina a granularidade adotada para representar os factos [65]. A granularidade determina o nível de detalhe ou de resumo dos dados armazenados no *Data Warehouse*, sendo definida consoante as necessidades de uma organização [17]. A definição adequada da granularidade de cada esquema é crucial para a utilidade e o custo do *Data Warehouse*, para além de influenciar no volume de dados armazenados [17]. Quanto mais detalhado for o dado, mais baixo é o nível de granularidade. Quanto menor o detalhe do dado, maior é o nível de granularidade [56].

A **Hierarquia** é composta por cada nível de uma dimensão, definindo uma sequência de mapeamentos de um conjunto de conceitos muito específicos de baixo nível para conceitos mais gerais de nível mais alto. Por exemplo uma dimensão geográfica pode conter rua-cidade-estado-país e a dimensão tempo pode ser formado por dia-semana-mês-ano [16]. A hierarquia pode ser balanceada ou não balanceada. Numa hierarquia balanceada, os níveis mais baixos são equivalentes. No entanto, isto não ocorre numa hierarquia não balanceada, no qual não possui uma equivalência hierárquica. Exemplo disso, pode-se encontrar numa dimensão geográfica, o nível ‘país’ pode não possuir o sub-nível estado’ relativamente a um determinado membro, e pode possuir para outro. No caso específico, o país Portugal não possui qualquer estado, ao contrário dos Estados Unidos da América que possuem vários estados.

2.2.3.1 Abordagens utilizadas no modelo multidimensional

As abordagens mais utilizadas para representar o modelo de dados multidimensional são o esquema em Estrela (*Star schema*), o esquema em Floco de Neve (*Snowflake schema*) e o esquema em Constelação de factos (*Facts Constellation schema*).

- **Esquemas em Estrela (*Star Schema*)**

O esquema em Estrela é o que apresenta o modelo mais simples, sendo caracterizado por conter uma tabela central de factos ligada a várias tabelas de dimensão, no qual estas não estão interligadas. Além disso, ao contrário do esquema em Floco de Neve, em que as dimensões são feitas de várias tabelas adjacentes, este tipo de esquema não utiliza, portanto, a normalização dos dados [47].

Os principais benefícios deste esquema incluem a facilidade de compreensão, uma vez que o esquema é simples. Este esquema proporciona uma redução no número de junções (*joins*) necessárias para recuperar dados [16]. No entanto, não suporta

o conceito de hierarquias de atributos ou sub-dimensões. O facto de este esquema não ser normalizado permite alcançar um desempenho nas consultas relativamente melhor que os restantes esquemas.

A figura 2.8 apresenta o esquema em Estrela composto pela tabela de factos ao centro, a qual está ligada a várias tabelas de dimensão. As tabelas de dimensão, como se pode observar na imagem, não possuem qualquer relação entre si.

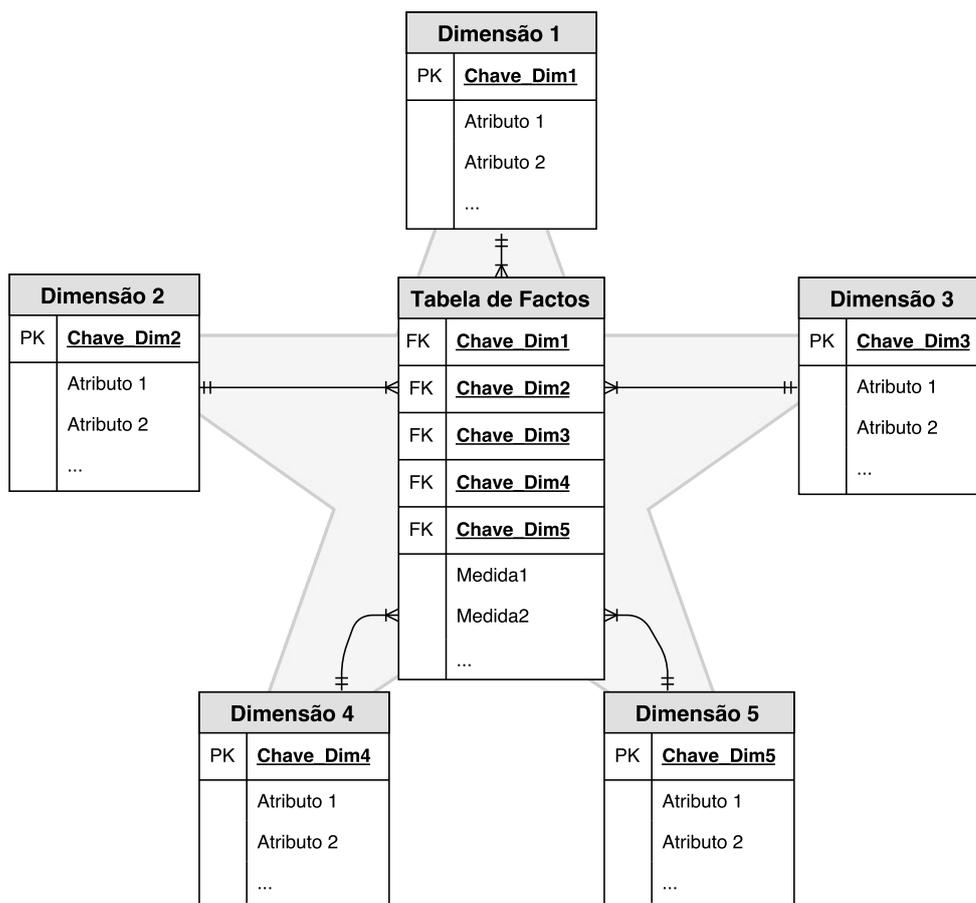


Figura 2.8: Esquema em Estrela

- **Esquema em Floco de Neve (*Snowflake Schema*)**

O esquema em Floco de Neve é um refinamento do esquema em Estrela. Este esquema é caracterizado por conter uma tabela de factos e tabelas de dimensão normalizadas em tabelas de dimensão mais pequenas, apresentando uma configuração semelhante a um floco de neve [16].

As tabelas são normalizadas em várias tabelas de pesquisa, representando cada uma um nível hierárquico da dimensão. Este modelo evita a redundância dos dados nas tabelas de dimensão, uma vez que se encontra normalizado. Esta normalização das tabelas de dimensão melhora a facilidade na manutenção das próprias tabelas e permite economizar o espaço de armazenamento [16]. Além disso, este esquema é adequado para o conceito de hierarquias. Porém, requer um número maior de operações de junção (*joins*) para fornecer respostas para a maioria das consultas, diminuindo o desempenho da recuperação de dados [16]. Apesar deste modelo não ser aconselhado em relação ao modelo em Estrela, pela possibilidade de problemas de desempenho nas consultas, existem algumas situações em que se pode considerar este esquema. Exemplo disso, é o facto do esquema em Estrela não favorecer a extração um subconjunto de atributos de uma dimensão, pois este esquema não proporciona suporte, de forma explícita, para hierarquias de atributos. Além disso, um esquema em Floco de Neve simplifica relações complexas das tabelas, como casos de muitos-para-muitos.

A figura 2.9 representa um esquema em Floco de Neve, que demonstra ao centro a tabela de factos e as diversas tabelas de dimensão. Nesta imagem verifica-se a existência de hierarquia dimensional, no qual as dimensões estão normalizadas.

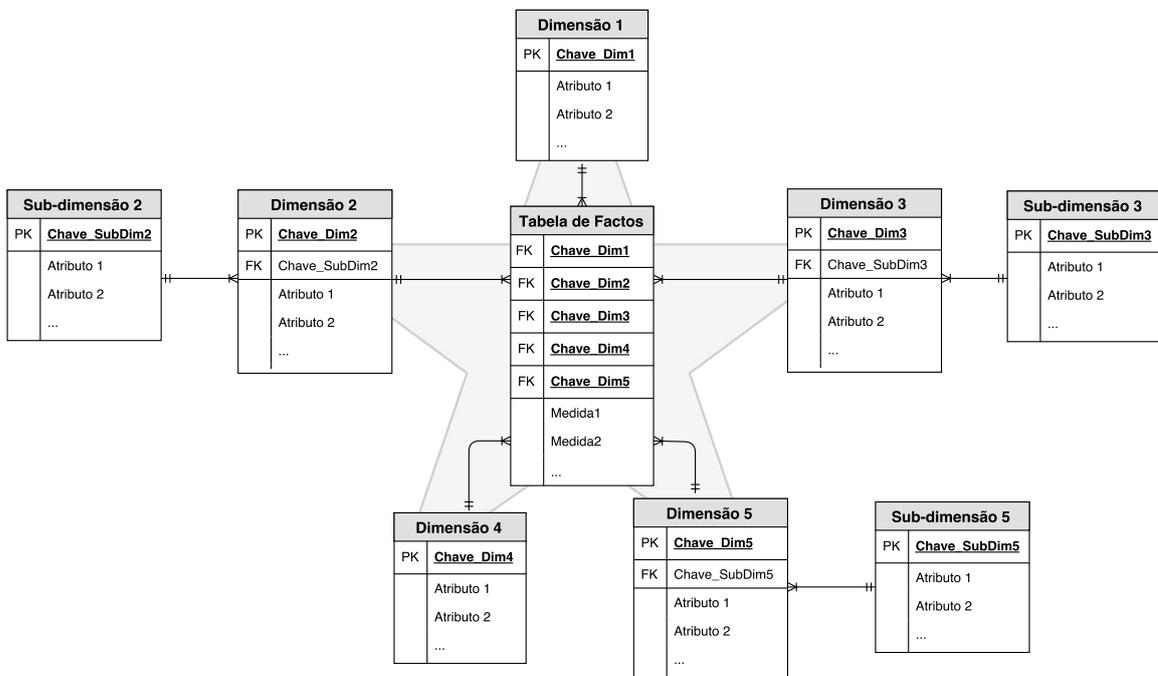


Figura 2.9: Esquema em Floco de Neve

- **Esquema em Constelação de Factos (*Facts Constellation schema*)**

O esquema em Constelação de Factos é caracterizado por apresentar estruturas mais complexas, em que várias tabelas de factos compartilham tabelas de dimensão em comum [5]. É um conjunto de esquemas em Estrela onde as tabelas de factos estão ligadas através da partilha de uma ou várias tabelas de dimensão [16]. A figura 2.10 demonstra um exemplo de esquema em Constelação de Factos, apresentando dois esquemas em Estrela ligados através de duas tabelas de dimensão que têm em comum.

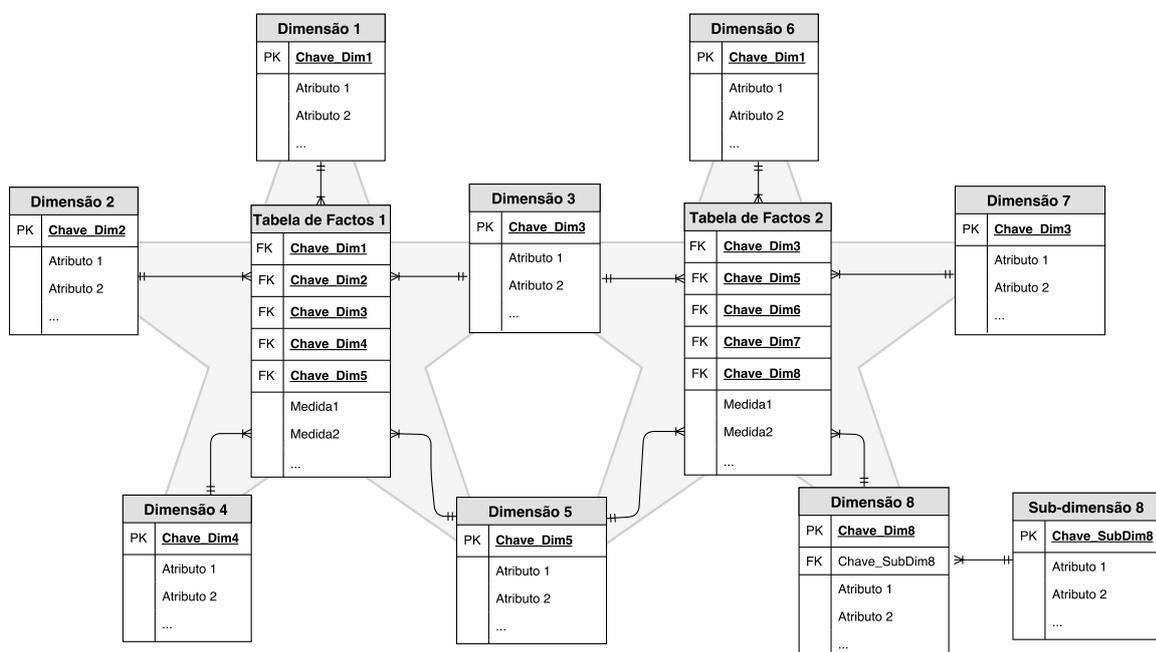


Figura 2.10: Esquema em Constelação de Factos

A análise multidimensional é uma das grandes utilidades dos sistemas OLAP, que permite obter informações rápidas e eficientes através de vários ângulos e perspectivas diferentes por intermédio de cubos multidimensionais.

Inmon e Kimball, além de defenderem diferentes abordagens, filosofias e estratégias na implementação de um *Data Warehouse*, também adotam técnicas de modelação de dados diferentes. Inmon defende que o projeto de um *Data Warehouse* que cumpra uma abordagem onde é garantida a normalização é a mais correta [56]. Inmon em [56] afirma que existem boas razões para o qual a normalização produza um projeto ideal para um *Data Warehouse*. O facto da normalização proporcionar flexibilidade, adaptar-se bem ao modelo de dados e a dados muito granulares (detalhados)

e, ainda, não se otimizar para qualquer conjunto de requisitos de processamento, são exemplos dessas boas razões. No entanto, Inmon refere que as estruturas de um *Data Mart*, que variam de acordo com as necessidades de cada área de uma organização, são tipicamente baseadas numa abordagem multidimensional [56]. Kimball, por sua vez, defende um modelo de dados multidimensional para apoiar a análise analítica dos dados e *Data Marts*, conceituando os esquemas mencionados anteriormente [64].

2.2.4 Cubos multidimensionais

O cubo de dados é uma estrutura que armazena os dados de negócio no formato multidimensional, facilitando a análise dos dados. Esta estrutura expressa a forma através da qual os tipos de informações se relacionam entre si.

No contexto de OLAP, um cubo é um modelo que é projetado para trabalhar, especificamente, com dados armazenados em diversas dimensões [47].

Os cubos permitem realizar uma análise multidimensional de dados e fornecem a capacidade de cálculos complexos, padrões e tendências importantes para o auxílio de tomadas de decisão nas organizações. Um conjunto de medidas agregadas, de acordo com as dimensões constituintes e com a tabela de factos, abrange um cubo de dados [67]. Assim, um cubo, além de ser composto por medidas, também é constituído pela tabela de factos e pelas tabelas de dimensão, que representam as possíveis formas de visualizar e consultar os dados.

A figura 2.11, que se segue, apresenta um exemplo de um cubo com três dimensões: Produtos, Tempo e Localização. Nesta imagem, também, é possível visualizar a medida denominada por 'Quantidade vendida', que representa a quantidade de produtos vendidos. Esta medida permite obter informações sobre essas quantidades por produtos, tempo e localização (dimensões).

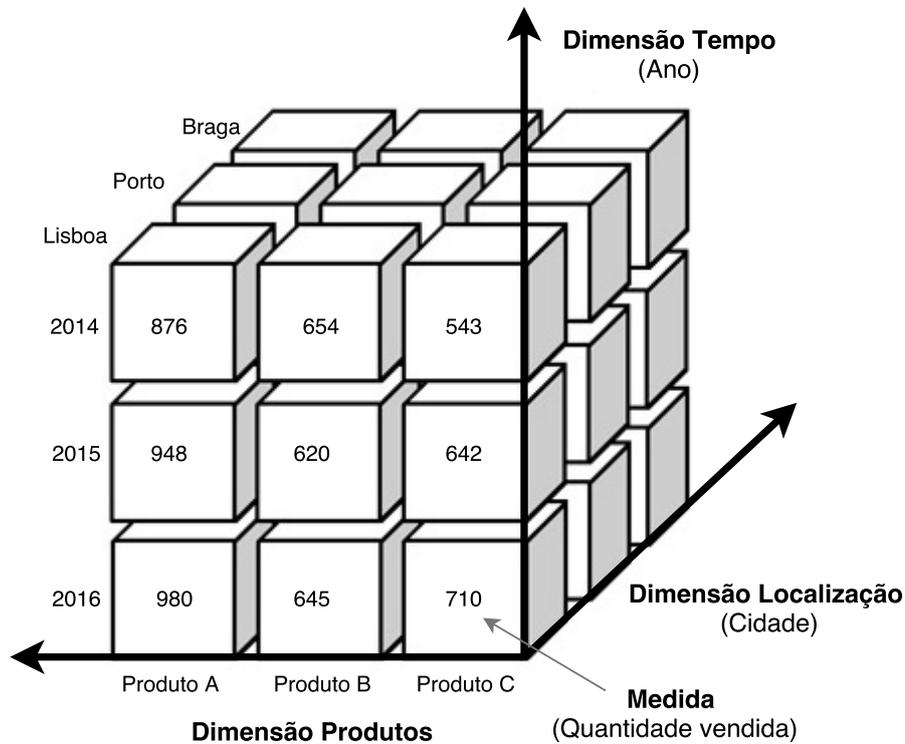


Figura 2.11: Exemplo de um cubo com três dimensões: Produtos, Tempo e Localização

Estes cubos, também conhecidos por cubos OLAP, fornecem suporte a análises mais detalhadas e pormenorizadas dos dados através de várias funcionalidades, tais como [55, 16]:

- **Roll up:** também denominado por *Drill up*, é utilizado para navegar para níveis mais baixos de detalhe de um determinado cubo de dados. O utilizador pode diminuir o nível de detalhe dos dados, o que corresponde a um aumento do nível da agregação dos dados, isto é, da granularidade;
- **Drill Down:** é o oposto de *Roll up* e é utilizado para navegar para níveis mais elevados de detalhe. O utilizador pode aumentar o nível de detalhe ao longo de uma ou mais hierarquias da dimensão, correspondendo a uma diminuição da granularidade;
- **Dice:** fornece apenas uma célula do cubo;

- **Slice:** permite selecionar uma determinada parte do cubo de dados, possibilitando aos utilizadores concentrarem-se numa dada parte específica dos dados dentro do cubo;
- **Pivot:** reorienta a visão multidimensional dos dados, ou seja, gira o cubo para mudar a perspetiva.

2.2.5 OLAP (*Online Analytical Processing*)

Após o processo de ETL e do armazenamento da informação nos repositórios de dados, *Data Warehouse* ou *Data Mart*, as ferramentas e metodologias de análise de dados vão permitir a pesquisa e análise de informações úteis para a tomada de decisão, destacando-se as ferramentas OLAP e *Data Mining* [44].

O termo OLAP, significa, em português, Processamento *Online* Analítico, e é uma poderosa ferramenta de *Business Intelligence*, altamente interativa. Esta ferramenta permite aos utilizadores realizar análises rápidas e dinâmicas de dados agregados, visualizar informações de múltiplas perspetivas ou dimensões, efetuar análises de tendências ao longo de períodos de tempo significativos, entre outros [47].

O OLAP é um processo ou arquitetura que possibilita análises complexas, multidimensionais, geralmente, através de *interfaces* gráficas de dados, a partir de bases de dados tradicionais ou de um *Data Warehouse* [44].

A tecnologia OLAP viabiliza às organizações uma competitividade mais acrescida. E isto, devido ao facto destas ferramentas permitem explorar rápida e eficientemente grandes volumes de dados e obter informações úteis para tomadas de decisão. Assim, tomadas de decisão rigorosas e imediatas, beneficiam a gestão dos negócios, conduzindo a empresas mais inovadoras e competitivas.

2.2.5.1 Arquiteturas OLAP

A arquitetura de sistemas OLAP está estruturada em três componentes principais [68]:

1. **Base de Dados** - é a fonte de dados utilizada para uma análise OLAP. Pode ser utilizado uma base de dados relacional, uma base de dados para assegurar o armazenamento multidimensional (*Data Warehouse*), uma estrutura de base de dados multidimensional, entre outras.

2. **Servidor OLAP** - representa a gestão da estrutura multidimensional dos dados e realiza, simultaneamente, a ligação entre a base de dados e um cliente OLAP.
3. **Cliente OLAP** - é representado pelas aplicações que garantem a exploração de dados e o suporte para a geração de resultados, tais como gráficos, relatórios e *dashboards*.

A arquitetura de um sistema OLAP refere-se como atuam os servidores quanto ao modo como os dados são armazenados numa aplicação OLAP. A arquitetura pode variar consoante o tipo de base de dados utilizado para as análises e consultas de dados. Através dessas variações surgiram as principais arquiteturas dos servidores dos sistemas OLAP que se seguem [69]:

- **ROLAP** (*Relational Online Analytical Processing*)

Nesta arquitetura, os servidores assumem que os dados são armazenados em bases de dados relacionais. Além disso, estes servidores suportam extensões para SQL e métodos de acesso e de execução especiais para implementar eficientemente o modelo de dados multidimensional e as operações pretendidas [5]. O modelo multidimensional e as suas operações têm de ser mapeados em relações e consultas SQL. Nesta arquitetura as consultas são enviadas para o servidor de base de dados relacional, sendo processadas no mesmo, ou seja, o cubo de dados é mantido no servidor. Os projetos de bases de dados utilizados em ROLAP são otimizados para consultas e carregamentos de dados eficientes.

A maioria destes sistemas utilizam como modelo de dados multidimensional um esquema em Estrela ou um esquema em Floco de Neve [2].

A principal vantagem de ROLAP é a escalabilidade em lidar com grandes volumes de dados. Porém, apresenta um baixo desempenho, podendo as consultas serem demoradas caso exista um número elevado de utilizadores a acederem em simultâneo [16]. O tamanho de dados limite neste tipo de sistema depende da limitação do tamanho da base de dados. Exemplo de motores ROLAP incluem o comercial *IBM Informix Metacube*⁴ e o servidor *Microstrategy DSS*⁵, e o produto *open source Mondrian*⁶ [16].

⁴<http://www.ibm.com>

⁵<http://www.microstrategy.com>

⁶<http://www.mondrian.sourceforge.net>

- **MOLAP** (*Multidimensional Online Analytical Processing*)

Nesta arquitetura, os servidores armazenam diretamente dados multidimensionais em estruturas de dados especiais como matrizes (*arrays*) e implementam as operações OLAP sobre estas estruturas de dados [5].

As ferramentas que utilizam esta arquitetura acedem a dados armazenados em bases de dados multidimensionais, os quais armazenam a informação em cubos dimensionais. Normalmente é utilizado grandes cubos de dados para acelerar o processo das consultas. O utilizador trabalha e manipula os dados do cubo diretamente no servidor. Além disso, têm a vantagem de possuírem excelentes propriedades de indexação e tempos de resposta de consultas rápidas. No entanto, proporcionam uma utilização do armazenamento relativamente pobre, sobretudo quando o conjunto de dados é escassa [2].

A arquitetura de MOLAP concentra-se no pré-cálculo dos dados transacionais para as agregações. Isto resulta numa execução de consulta rápida, ou seja, pré-calcula e armazena medidas agregadas de todos os níveis hierárquicos no momento de carregamento. Além de que, também, armazena e indexa esses valores para uma recuperação imediata [16].

Esta arquitetura apresenta um elevado desempenho e permite a execução de cálculos complexos. Contudo, apresenta baixa escalabilidade, sendo limitada em relação à quantidade de dados, e exige investimentos altos, sobretudo na projeção de cubos. Exemplo de produtos MOLAP são o comercial *Hyperion Ebase*⁷, o *Applix TM1*⁸ e o produto *open source Palo*⁹ [16];

- **HOLAP** (*Hybrid Online Analytical Processing*)

Nesta arquitetura, os servidores combinam a arquitetura dos servidores MOLAP e ROLAP, de modo a alcançar um equilíbrio entre a escalabilidade do ROLAP e o desempenho de consulta de MOLAP. Neste caso, o utilizador pode decidir que parte dos dados é armazenada no ROLAP e no MOLAP em separado [16]. Esta arquitetura apresenta alto desempenho e uma elevada escalabilidade. No entanto, detém entre as arquiteturas um custo mais elevado. Exemplo de um produto que suporta os três tipos de arquiteturas é o *Microsoft OLAP Services*¹⁰, a qual faz parte da empresa *SQL Server* [16].

⁷<http://www.hyperion.com>

⁸<http://www.applix.com>

⁹<http://www.opensourceolap.org>

¹⁰<http://www.microsoft.com>

- **DOLAP** (*Desktop Online Analytical Processing*)

A arquitetura DOLAP é uma arquitetura *desktop* do OLAP. Esta arquitetura permite ao utilizador emitir uma consulta de uma máquina cliente para o servidor, sendo que este retorna um micro-cubo de dados respetivo de volta, possibilitando a análise do mesmo.

Com um sistema deste tipo, o utilizador pode transferir para o seu computador uma cópia da base de dados multidimensional na sua totalidade ou em partes, bem como aceder a um repositório de dados e trabalhar esses dados localmente [70]. Um sistema com este tipo de arquitetura é mais fácil de implementar. Todavia não oferece tantas funcionalidades em relação aos outros sistemas [71]. Além disso, esta arquitetura reduz o tráfego na rede e melhora o desempenho do servidor de base de dados, pois o processamento da informação é realizado na máquina cliente. Esta arquitetura apresenta um tamanho do cubo de dados limitado. Um cubo neste tipo de sistemas não deve ser muito grande, pois caso contrário, a análise pode ser demorada e pôr em causa a capacidade de suporte da máquina cliente [70].

- **WOLAP** (*Web Online Analytical Processing*)

A arquitetura WOLAP refere-se a uma *interface web* para aceder à aplicação OLAP, ou seja, os dados OLAP são acedidos através de um *web browser*. Os sistemas que apresentam esta arquitetura são compostos por três componentes: um cliente, um *middleware* e um servidor de base de dados [71].

As ferramentas com esta arquitetura permitem ao utilizador emitirem uma consulta através de um *web browser* para o servidor, retornando a respetiva informação por páginas *web* para ser analisada. A evolução das tecnologias *web* têm possibilitado às empresas de *software*, que desenvolvem ferramentas para OLAP, integrarem nas suas *suites* uma plataforma *web*, de modo os clientes poderem ter acesso aos dados através de um *web browser* [72]. As principais vantagens que este tipo de arquitetura apresenta são o baixo custo do sistema e da sua manutenção, maior acessibilidade e facilidade de instalação, de configuração e de implementação [71]. Além do mais, não é necessário a instalação de um *software* cliente na máquina cliente. Sendo isto devido ao facto das ferramentas com esta arquitetura permitem que o acesso aos dados e o processamento analítico sejam realizados através de qualquer computador com acesso de rede ao servidor OLAP [72]. Exemplos de ferramentas WOLAP são: *SpagoBI*, *Pentaho BI suite*, *Vanilla* e *Palo*. Estas ferramentas são analisadas e comparadas mais adiante.

No caso de o ambiente de apresentação ser baseado numa base de dados relacional, em seguida, essas tabelas são modeladas dimensionalmente, assentes num dos três principais esquemas de modelação (Estrela, Floco de Neve e Constelação de Factos), como o caso de ROLAP. Se o ambiente de apresentação é baseado na base de dados multidimensional, então os dados são armazenados em cubos, como o caso de MOLAP [64].

2.2.6 *Data Mining*

O processo de *Data Mining* consiste na exploração e análise de grandes volumes de informação de modo a permitir descobrir padrões de dados, relações entre conjuntos de dados e regras relevantes para a resolução de um determinado problema. *Data Mining* refere-se à extração de conhecimento a partir de uma grande quantidade de dados ou de um *Data Warehouse*. A realização da extração dos dados nos sistemas de base de dados é um processo complexo, no qual *Data Mining* combina inteligência artificial, técnicas, análises e algoritmos estatísticos para tentar extrair conhecimento a partir dos dados armazenados [15]. Os resultados são apresentados através de ferramentas *front-end* que permitem a visualização dos mesmos. Os algoritmos frequentemente implementados incluem redes neurais, árvores de decisão, *clustering*, algoritmos genéticos e estatística [17].

Esta técnica permite descobrir relações e padrões existentes em vastas informações auxiliando a tomar melhores decisões. Além disso, *Data Mining* é muito eficaz para previsões de tendências e de comportamentos futuros, fornecendo factos aos gestores para tomarem decisões mais acertadas.

A diferença básica entre ferramentas OLAP e *Data Mining* está na maneira como a exploração dos dados é abordada. Com ferramentas OLAP a exploração dos dados é realizada na base da verificação, no qual o utilizador conhece a questão, elabora uma hipótese e utiliza a ferramenta para confirmá-la. Com ferramentas *Data Mining*, a questão é parcial ou totalmente desconhecida, sendo a exploração dos dados conduzida pela ferramenta utilizada para a busca de conhecimento, com o mínimo de intervenção por parte do utilizador.

2.2.7 Aplicações *Front-End*

As aplicações *front-end* constituem a *interface* entre o utilizador e o sistema. Estas aplicações permitem aos utilizadores aceder, explorar e visualizar as análises a

dados para obterem novos conhecimentos essenciais para tomadas de decisão. Desta forma, proporcionam, aos decisores, gestores e analistas, eficácia na interpretação de factos ocorridos e na decisão sobre estratégias futuras.

Estas aplicações são responsáveis pela seleção dos dados necessários, pelo cálculo e manipulação desses dados e pela apresentação das informações. As aplicações *front-end* permitem consultas *ad-hoc*, criar diversos gráficos, relatórios e *dashboards*. Os sistemas de *Data Warehouse*, OLAP e *Data Mining* permitem manipular e obter dados, no qual através das aplicações *front-end* os utilizadores podem ter acesso a esses dados de forma rápida, simples e eficaz.

Os sistemas de *Business Intelligence* devem fornecer aplicações intuitivas e com *interfaces* acessíveis de utilizarem de modo a facilitarem o acesso aos dados por parte dos utilizadores.

Existem várias aplicações *front-end* através das quais os utilizadores executam tarefas de *Business Intelligence*. Por exemplo, folhas de cálculo, portais de empresa para pesquisas e ferramentas para consultas *ad-hoc*. E ainda, ferramentas de visualização de modelos de *Data Mining*, ferramentas de análise de dados (OLAP) e aplicações de gestão de desempenho, que permitem aos gestores acompanhar os indicadores de desempenho (KPIs) através de *dashboards*, entre outros [2, 52].

2.3 Conclusões

Atualmente, o rápido crescimento da população nas zonas urbanas tem levando à necessidade de criar soluções e estratégias inteligentes e sustentáveis face aos novos desafios que surgem constantemente. Estas soluções e estratégias são imprescindíveis para promoverem cidades mais habitáveis e mais sustentáveis, denominadas por Cidades Inteligentes. Para tal, é fundamental as cidades focarem-se em diversos parâmetros como proporcionarem uma urbanização sustentável, um desenvolvimento económico atrativo, melhor qualidade de vida e do ambiente, igualdade social, entre outros.

Uma Cidade Inteligente proporciona muitos benefícios, pois é um princípio que potencia a cidade como um ambiente ideal para inovações e incubação de novos negócios, ideias e criação de novos serviços, promovendo um crescimento económico e social atrativo. Além disso, uma Cidade Inteligente pode ocasionar reduções nas despesas públicas, especialmente, pela prestação de serviços públicos *online*, aperfeiçoar a eficiência e a qualidade dos serviços através de uma gestão ágil de recursos utilizando

ferramentas de TIC, e ainda, disponibilizar informações da cidade em tempo real para apoiar tomadas de decisão. A utilização das TIC nas diversas práticas em Cidades Inteligentes viabiliza o melhoramento das condições de vida e o desenvolvimento da sustentabilidade.

As iniciativas em Cidades Inteligentes surgem como modelo para atenuar e retificar os problemas atuais nas áreas urbanas, tornando as cidades mais habitáveis, mais sustentáveis, com melhores serviços e com melhor desenvolvimento económico e social. Estas iniciativas procuram soluções que integram os aspetos mais essenciais como energia, mobilidade, ambiente, economia, transporte, comunicações, tecnologias de informação e comunicação, saúde, água potável e cidadãos.

Dentro desta temática, existe um projeto denominado por “*Smart Cities for Sustainable Development*”, desenvolvido pelos investigadores Elsa Estevez, Nuno Vasco Lopes e Tomasz Janowski, com a parceria da UNU-EGOV e IDRC [1]. Este projeto teve a finalidade de explorar a tese de que Cidades Inteligentes permitem alcançar o desenvolvimento sustentável. Neste projeto foi elaborado um documento intitulado por “*Smart Cities for Sustainable Development - State of Practice Survey*”, fundamentado num modelo conceptual. O documento mencionado tem o intuito de obter descrições uniformes das diversas iniciativas em Cidades Inteligentes para o desenvolvimento sustentável. Este documento juntamente com as 21 iniciativas descritas neste projeto referenciado serviram de base para a elaboração da plataforma *online*.

No contexto dos setores públicos e empresarial, em virtude das progressivas evoluções, como o rápido aumento do volume dos dados, desenvolvimento tecnológico e aumento da concorrência, as empresas estão à procura de soluções capazes de corresponder a esses progressos. As empresas procuram implementar sistemas de informação mais complexos, dada à gradual importância de tomar decisões fundamentais e à necessidade de obter grandes quantidades de informação de forma eficaz, confiável, num curto espaço de tempo. Desta forma, cada vez mais é frequente, as empresas utilizarem ferramentas de *Business Intelligence* para o armazenamento e análise de grandes quantidades de dados. Destas ferramentas, têm-se evidenciado as ferramentas *open source* e de *software* livre, sobretudo, para pequenas e médias empresas, devido ao facto de, muitas das vezes, possuírem um custo reduzido, ou até mesmo, inexistente. Os termos *Data Warehousing* e sistemas OLAP são dois conceitos que se destacam neste projeto, de entre diversas práticas de *Business Intelligence*. *Data Warehousing* é a projeção e a implementação de sistemas que permitem reunir, integrar, manipular dados provenientes de diversas fontes, os quais são armazenados

em repositórios de armazenamento, denominados por *Data Warehouses*. Os sistemas OLAP permitem, com base na recolha de informação no *Data Warehouse*, a obtenção de gráficos, *dashboards*, relatórios, que possibilitam análises eficientes e rápidas de grandes volumes de dados através de diversas perspetivas. Os sistemas OLAP e *Data Warehouse* utilizam modelos de dados multidimensionais para definir a estrutura de dados em Cubos, os quais permitem a análise rápida de grandes informações em diversas perspetivas.

Este capítulo apresentou os conceitos teóricos relacionados com o desenvolvimento desta dissertação. Inicialmente, foram demonstradas as principais noções pertinentes a Cidades Inteligentes, bem como referida a importância em tornar uma cidade em Cidade Inteligente face ao crescimento populacional. Além disso, por último, neste capítulo descreveu-se os conteúdos mais relevantes associados a *Business Intelligence*. Deste tema, evidenciou-se os sistemas de *Data Warehouse*, como repositórios de dados e os sistemas OLAP, para análises eficientes de grandes volumes de dados.

Capítulo 3

Estudo de ferramentas de BI *open source*

Este capítulo apresenta uma visão geral sobre o estudo de algumas ferramentas *open source* disponíveis para *Business Intelligence*. Este estudo teve o objetivo de selecionar uma ferramenta de BI *open source* para a criação de análises estatísticas sobre os dados das diversas iniciativas em Cidades Inteligentes inseridos através da plataforma *online*. Primeiramente, na secção 3.1 é disponibilizada uma visão geral sobre os critérios conhecidos para avaliação de produtos OLAP. Em seguida, na secção 3.2 é apresentada uma descrição das funcionalidades e capacidades de cinco ferramentas *Business Intelligence open source* do tipo *suite* denominadas por: *SpagoBI*, *Pentaho*, *JasperSoft*, *Palo/Jedox* e *Vanilla*. Estas ferramentas suportam o conceito OLAP e são as mais utilizadas pela comunidade científica e empresarial. Posteriormente, na secção 3.3 é demonstrada uma tabela comparativa dessas ferramentas *open source*. Finalizando, este capítulo, com uma revisão literária sobre o estudo de ferramentas *Business Intelligence open source* e uma conclusão sobre a seleção de uma das ferramentas mencionadas anteriormente (secção 3.4 e 3.5, respetivamente).

3.1 Critérios para avaliação de ferramentas OLAP

O termo OLAP foi referido pela primeira vez por E. F. Codd, em 1993, num artigo, no qual definiu doze regras (às quais, posteriormente, foram adicionadas mais seis regras), que caracterizam uma ferramenta OLAP (ver tabela 3.1).

A tabela 3.1 retrata essas doze regras inicialmente validadas para avaliação de produtos OLAP.

Tabela 3.1: Doze regras de Codd para a avaliação de produtos OLAP, baseado em [17, 18]

Regras	Caraterísticas das ferramentas OLAP
1ª: Visão conceptual multidimensional	Os utilizadores podem manipular os modelos multidimensionais de dados de forma fácil e intuitiva, incluindo operações como <i>slice</i> e <i>dice</i> .
2ª: Transparência	Interação fácil com os <i>front-ends</i> , permitindo a inclusão de uma ferramenta analítica onde o utilizador desejar.
3: Acessibilidade	Tratamento de dados heterogéneos de forma lógica, possibilitando uma apresentação aos utilizadores de forma única, coerente e consistente.
4: Desempenho coerente de fornecimento de informações	O utilizador não deve observar reduções significantes no desempenho de fornecimento de informações, independentemente do tamanho da base de dados.
5: Arquitetura cliente-servidor	É importante que uma ferramenta seja capaz de operar num ambiente cliente-servidor.
6: Dimensionamento genérico	A dimensão dos dados não deve influenciar a estrutura dos dados nem o formato dos relatórios.
7: Manipulação dinâmica de matrizes esparsas	A ferramenta deve permitir o ajuste do esquema físico de acordo com a densidade dos dados para a obtenção de desempenho máximo.
8: Suporte a multi-utilizadores	Deve permitir o acesso simultâneo sem prejudicar a segurança e desempenho dos dados.
9: Operações de cruzamento dimensional sem restrições	Qualquer conjunto de dados pode ser acedido a qualquer momento para obtenção de cálculos.
10: Manipulação de dados intuitiva	Deve permitir a realização dos cálculos e a manipulação dos dados de forma mais intuitiva possível.
11: Relatórios flexíveis	Os relatórios devem apresentar os dados de forma lógica e sintetizada.
12: Dimensões e níveis de agregações ilimitados	Deve possibilitar a acomodação de diversas dimensões de dados num modelo analítico, sendo que cada dimensão deve permitir um número sem limites de níveis de agregação.

Porém, estas regras acabaram por tornarem-se numa forma inadequada e indesejável de detetar a “obediência OLAP”. Este facto ocorreu devido à verificação de uma imparcialidade nessa definição, devido à descoberta de financiamento a Codd, por uma empresa fornecedora de um produto OLAP. Posto isto, estas regras foram muito criticadas, levando à criação de uma alternativa por parte de outros investigadores. Assim, surgiu uma teoria mais simples, fácil de memorizar e independente de produtos já existentes de empresas diferentes, denominada por teste “FASMI” (*Fast*

Analysis of Shared Multidimensional Information) [17].

FASMI é um termo alternativo para a avaliar produtos OLAP e foi inventado por Nigel Pendse, em 1995. O investigador Pendse criou este conceito uma vez que sentia que as doze regras utilizadas por Codd, além de serem numerosas, eram demasiadas controversas e tendenciosas. Esta sintetização de OLAP em cinco palavras, denominado, também, por teste FASMI, caracterizam as aplicações OLAP de uma forma específica sem ditar a sua implementação, de fácil compreensão, permitindo avaliar os produtos OLAP. Essas palavras constituintes de FASMI são as seguintes [17, 16]:

- **Fast (Rápida).** O sistema deve estar orientado para oferecer respostas rápidas. Tendencialmente o sistema deve estar voltado para dar respostas aos utilizadores entre um a vinte segundos, dependendo da complexidade, para com análises mais simples. Existem estudos que mostram que os utilizadores finais após 30 segundos consideram que o sistema entrou em erro e cancelam a operação, a menos que o sistema avise que o relatório vai demorar mais tempo. Embora os utilizadores sejam avisados dessa demora, estes estão suscetíveis de se distraírem e perder a linha de pensamento prejudicando a qualidade de análise. Esta velocidade não é fácil de conseguir com grandes quantidades de dados e os fornecedores recorrem a uma grande variedade de técnicas para atingir esse objetivo.
- **Analysis (Analítica).** O sistema pode lidar com qualquer lógica de negócio e análise estatística que é indispensável para a aplicação e o utilizador. Além disso, o sistema deve disponibilizar uma ferramenta de análise suficientemente fácil de ser usada, funcional e intuitiva para o utilizador final. Também, deve permitir ao utilizador definir novos cálculos *ad-hoc* e apresentar relatórios sobre os dados, que satisfaçam as necessidades de flexibilidade e cálculo orientado dos utilizadores, sem necessidade de programação.
- **Shared (Partilhada).** O sistema deve implementar todos os requisitos de segurança e de confidencialidade e, se caso várias permissões de escrita sejam necessárias, deve existir cuidados adaptados a atualizações simultâneas. Nem todas as aplicações necessitam que os utilizadores escrevam dados de volta. Porém para números crescentes, o sistema deve ser capaz de lidar com várias atualizações num tempo adequado. Em muitos produtos OLAP, esta é uma das principais áreas de fraqueza e que tendem a assumir que todas as aplicações

OLAP adotam controlos de segurança simplistas e que necessitam apenas de permissões de leitura.

- ***Multidimensional (Multidimensional)***. É o principal requisito de um sistema OLAP. O sistema deve fornecer uma visão conceptual multidimensional dos dados, incluindo o suporte completo para hierarquias e múltiplas hierarquias, sendo este certamente a forma mais lógica para analisar grandes dados de empresas e organizações.
- ***Information (Informação)***. Os sistemas devem ser capazes de lidar com grandes informações e quantidades de dados, independentemente do modelo de servidor. Está-se a medir a capacidade de vários produtos em termos de quantidade de dados de entrada que podem segurar. As capacidades dos produtos diferem muito – os maiores produtos OLAP podem manter, pelo menos, mil vezes mais dados do que o menor produto. Há muitas considerações neste ponto, como a necessidade de replicação de dados, memória RAM necessária, a utilização do espaço em disco, o desempenho, a integração com outros sistemas que sejam necessários à análise de dados e afins.

Investigações afirmam que o teste FASMI é uma definição razoável e compreensível dos objetivos e propósitos de sistemas OLAP, procurando, também, incentivar os utilizadores e os fornecedores a adotar esta definição.

The OLAP Council, um concelho fundado, em 1995, para padronizar a tecnologia OLAP, referenciou OLAP como sendo “uma categoria de tecnologias de *software*, que permite, aos analistas, gestores e executivos, obter conhecimento sobre os dados, através de um acesso rápido, consistente e interativo, de uma grande variedade de possíveis visualizações de informações, que foram transformadas a partir de dados brutos, para refletir a dimensionalidade real da empresa como é entendida pelo utilizador” [73].

Embora não seja o único requisito para OLAP, o principal conceito que todos os produtos OLAP têm em comum é a multidimensionalidade. Este conceito inclui a noção de múltiplas dimensões hierárquicas, podendo ser utilizado por qualquer pessoa capaz de pensar em termos de um mundo multidimensional [74].

As ferramentas OLAP constituem, assim, um poderoso instrumento de auxílio às empresas, num mundo cada vez mais competitivo, onde é necessário conseguir informações mais precisas em períodos de tempo relativamente curtos. Desta forma,

estas ferramentas permitem, aos utilizadores finais, obter conhecimento por via da transformação de informação em dados essenciais para tomadas de decisão benéficas e inovadoras.

3.2 Descrição das ferramentas de BI *open source* OLAP

No mundo de *Business Intelligence* existe, no mercado atual, uma vasta variedade de ferramentas, tanto comerciais como *open source* (código aberto). Estas ferramentas possuem uma grande utilidade para as organizações, na medida que permitem aperfeiçoar a gestão de negócios fundamental para que estas possam adaptar-se às constantes evoluções e suprimir a competitividade com inovação. As ferramentas de BI fornecem um conjunto de funcionalidades para análises de dados e previsão de tendências que permitem melhorar e aumentar as potencialidades das organizações.

As ferramentas de *Business Intelligence*, do tipo *open source*, têm vindo, cada vez mais, a tornarem-se relevantes e utilizadas, sobretudo em pequenas e médias empresas, existindo atualmente, no mercado, competição com as soluções comerciais. A adoção de ferramentas *open source* tem sido, particularmente, associada a ganhos económicos nas organizações, uma vez que apresentam custos muito baixos ou nulos e, simultaneamente, constituem produtos estáveis e de qualidade [75].

Open source é um método de desenvolvimento de *software* que promete melhor qualidade, maior fiabilidade e flexibilidade, segurança, bom desempenho, menores custos e vantagens competitivas. Por estes motivos é que as empresas têm aplicado este tipo de soluções [24]. Uma ferramenta *open source* deve disponibilizar o seu código fonte, bem como permitir a sua distribuição [76]. É a licença de *software open source* que define e estabelece o que é permitido fazer com o código fonte disponibilizado. Uma vez que o código fonte encontra-se disponível e licenciado, enquanto detentor dos direitos de autor, é permitido aos utilizadores a visualização, a leitura, a verificação e a correção de erros, a modificação, melhorias e redistribuição do código. A licença que mais se destaca nas ferramentas *open source* é a GNU GPL (*GNU General Public License*) [77]. O conceito de *open source* pode ser visto como um fenómeno revolucionário capaz de proporcionar a indústria de *software* como uma forma alternativa e competitiva de fazer negócios [76].

Contrariamente ao *software* comercial, o *software open source* aproveita a capacidade das comunidades (desenvolvedores, colaboradores e programadores volun-

tários) para melhorar a qualidade dos produtos e reduzir o custo de desenvolvimento [77].

O estudo de um conjunto de ferramentas *Business Intelligence open source*, que incluíssem a tecnologia OLAP, surge no sentido de analisar e selecionar uma ferramenta robusta para a exploração e análises de dados relacionados com iniciativas em Cidades Inteligentes.

Embora existam ferramentas do tipo servidor OLAP e cliente OLAP, optou-se pela análise de ferramentas que possuem servidor e cliente OLAP em simultâneo, sendo designadas por *suite*. As ferramentas do tipo *suite* agrupam vários componentes e funcionalidades de BI. E, uma vez que apresentam configurações pré-estabelecidas, têm como principal característica a facilidade na instalação, configuração e utilização das mesmas.

A seguir, na próxima secção, é descrito o conjunto de ferramentas analisadas de *Business Intelligence open source* do tipo *suite*. A escolha destas ferramentas, para o estudo e comparação das mesmas, está relacionada com a sua forte utilização, aprovação e influência na comunidade científica e empresarial. Deste modo, as ferramentas analisadas são as seguintes: *SpagoBI*¹, *Pentaho*², *JasperSoft*³, *Palo*⁴ e *Vanilla*⁵.

3.2.1 *SpagoBI Suite*

A ferramenta *suite SpagoBI*, desenvolvida pela *SpagoWorld* (fundada e gerida pela empresa de consultoria italiana *Engineering Group*⁶), é um *software open source/free* (código aberto/livre) de nível empresarial. Este *software* caracteriza-se por ser versátil e adaptável face às necessidades dos utilizadores finais. *SpagoBI* encontra-se alojado numa organização internacional, independente e sem fins lucrativos, a *OW2 Forge*. Esta *suite* tem o objetivo de garantir uma transparência permanente, uma sustentabilidade e a disponibilidade de *software*, e ser aberto [78].

Como é um *software* de código aberto/livre pode ser estudado, utilizado, modificado e distribuído em cópias modificadas deste *software*, sob os termos da licença MPL na versão 2 (v2). *SpagoBI* é fornecido apenas na versão (*Community*) (Comunidade), completamente gratuita, tal como é e sem qualquer tipo de garantia.

¹<http://www.spagobi.org>

²<http://www.pentaho.com>

³<http://www.jaspersoft.com>

⁴<http://www.jedox.com>

⁵<http://www.bpm-conseil.com>

⁶<http://www.eng.it>

A ferramenta até à versão 3.4 foi disponibilizada através da licença GNU LGPL, permitindo aos utilizadores o uso das suas funcionalidades na sua totalidade. A licença MPL v2, utilizada nesta ferramenta, apresenta um direito de cópia *copyleft*⁷, e caracteriza-se por ser simples, compreensível e compatível com GNU (L)GPL. Esta licença foi projetada para incentivar contribuidores a partilhar as modificações que fazem ao código de licença MPL, e permite que os utilizadores criem projetos combinados entre licenças MPL sob outras licenças, quer seja de código aberto ou proprietário [79].

A ferramenta, que é desenvolvida em Java, possibilita diversas funcionalidades. Algumas dessas funcionalidades fornecidas são: a criação e a exportação de relatórios e gráficos, a integração de dados através de processos ETL, a análise OLAP, a criação de filtros de consulta e a utilização de técnicas *Data Mining*. Além disso, permite a visualização de indicadores-chave de desempenho (KPIs) em tempo real, a obtenção de consultas simples a bases de dados e a disponibilização de técnicas de georreferência [21, 80].

As ferramentas *Eclipse BIRT*, *JasperReport*, *Accessible Report* e *Business Object* são utilizadas, nesta *suite*, para criar e visualizar relatórios. Os motores *JPivot/Mondrian*, *JPalo/Mondrian* e *JPivot/XMLA Server* permitem o suporte dos serviços de OLAP [21].

SpagoBI tem uma visão de plataforma de integração e não de produto, ou seja, o servidor *SpagoBI* pode atender a vários mecanismos para trabalhar ou apresentar diferentes elementos de uma solução BI. Por exemplo, o serviço de OLAP pode ser suportado pelos motores *JPivot/Mondrian*, *JPalo/Mondrian* e *JPivot/XMLA Server* e a geração de relatórios pode ser efetuado através do *Birt*, *Jasper* ou *Business Objects* [21, 80].

Esta ferramenta suporta toda a escalabilidade necessária numa organização, quer a nível de arquitetura, como a nível de funcionalidades e segurança [22].

SpagoBI é um sistema de *Business Intelligence open source* completo, que pode ser implementado e utilizado sem pagamento de licenças [21].

A figura 3.1 apresenta um exemplo do ambiente que a ferramenta *suite SpagoBI* fornece, na qual se observa alguns exemplos de gráficos disponíveis nesta ferramenta.

⁷ *Copyleft* é um direito de permissão de cópia por outros utilizadores, que permite a liberdade de copiar, modificar e redistribuir, desde que esse direito seja preservado em todas as versões modificadas

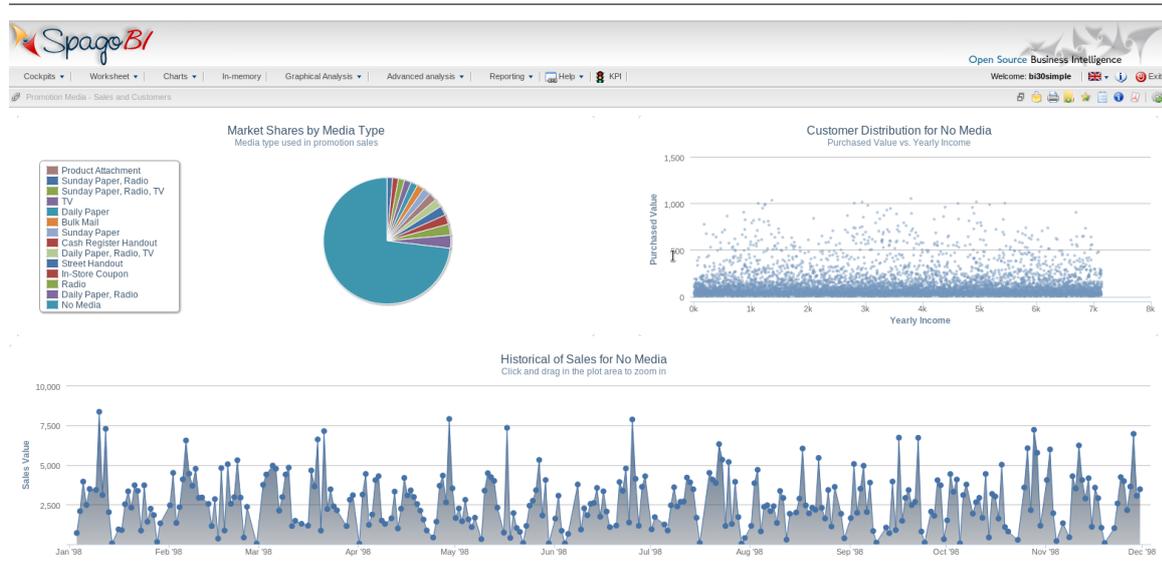


Figura 3.1: Exemplo de gráficos interativos na *suite SpagoBI*, retirado de [6]

3.2.2 JasperSoft

A ferramenta *open source JasperSoft* surgiu em 2001, sendo inicialmente designada por *Panscopic*. Após algum tempo, o fundador da *JasperReports*, Teodor Danciu, associou-se à equipa *Panscopic*, surgindo, em 2004, a *JasperSoft* [21]. A ferramenta *suite* de *JasperSoft* foi criada em 2006, logo após a empresa ter-se dedicado apenas à criação de várias ferramentas individuais [22].

JasperSoft é desenvolvida em linguagem *Perl* e *Java* e possui uma licença GNU GPL [24]. Existem duas versões, a Comercial, distribuída por três edições (*Express*, *Professional* e *Enterprise*) e a *Community* disponibilizada em três produtos individuais: servidor *JasperReports*, *JasperSoft OLAP* e *JasperSoft ETL* [22, 21]. A versão *Community* é *open source*, gratuita e bastante limitada relativamente às versões comerciais, sendo distribuída através de uma licença GNU GPL [22]. A ferramenta *JasperSoft BI* do tipo *suite* encontra-se integrada por um servidor, o *JasperServer*. Este servidor possibilita várias funcionalidades, tais como criar e estruturar relatórios interativos e diferentes tipos de gráficos, consultas *ad-hoc* (eventualidade de estas serem em tempo real) e simplicidade no acesso à informação proveniente de bases de dados. Além do mais, esta ferramenta permite obter análises OLAP, integrar processos de ETL, fornece a funcionalidade *query builder* (Construtor de consultas do tipo SQL) e de georreferenciação suportada por vários mapas de escala mundial. Esta ferramenta, também, é composta pelos módulos *JasperReports Server* (servidor de

relatórios), *JasperReports Library* (criação de relatórios mais popular), *JasperSoft ETL* (integração de dados), *JasperSoft Studio* (ambiente de desenho de relatórios baseado em *Eclipse*) e *iReport Designer* (ambiente de desenho de relatórios baseado em *NetBeans*) [22].

JasperSoft caracteriza-se por ser robusta, confiável e atualizável, permitindo aos utilizadores sem conhecimentos técnicos sobre bases de dados ou linguagens de consulta, obterem, sem grandes dificuldades, análises e consultas de dados. Embora esta ferramenta seja um pacote BI aceitável, possuindo as principais funcionalidades de BI, apresenta algumas lacunas, tais como a inexistência de um módulo de *Data Mining* e de funcionalidades de KPIs [21].

A figura 3.2 demonstra um exemplo de gráfico que a ferramenta *suite JasperSoft* oferece.

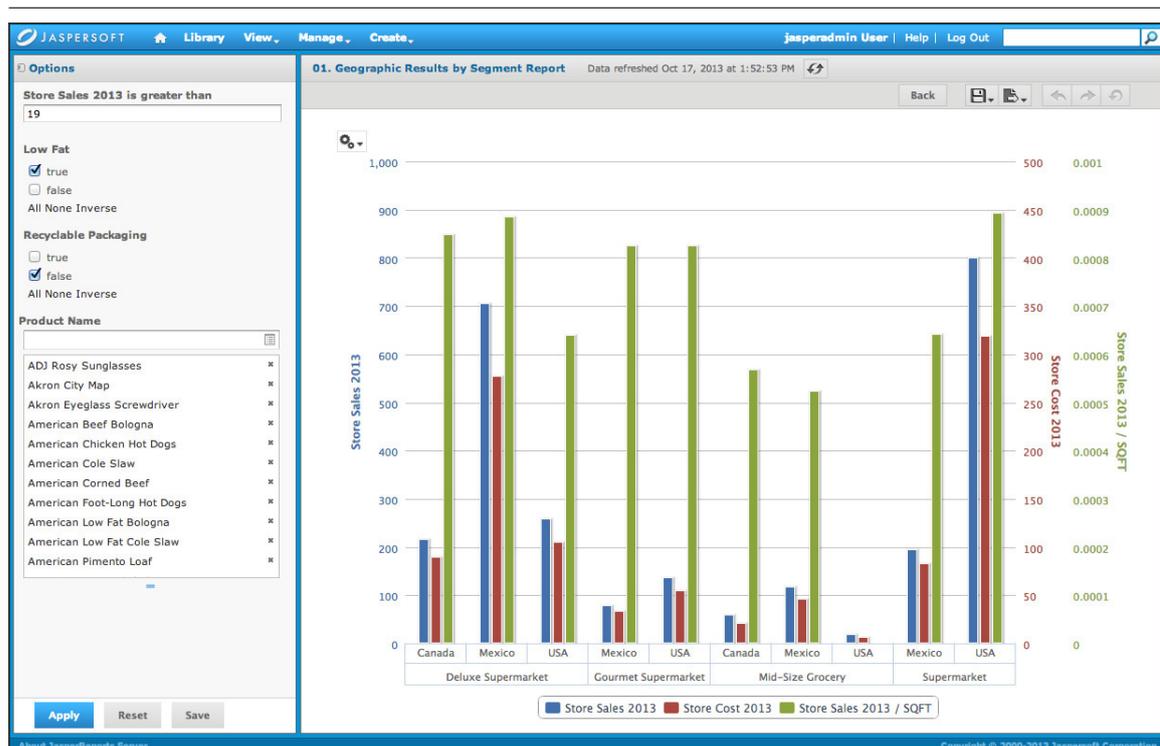


Figura 3.2: Exemplo de análise interativa na *suite JasperSoft*, retirado de [7]

3.2.3 *Vanilla*

Vanilla é uma ferramenta BI *open source* do tipo *suite*, desenvolvida e suportada pelo consórcio francês *BPM-Conseil*.

Esta ferramenta difere das outras, uma vez que foi desenvolvida em PHP e à semelhança das restantes possui uma licença GNU GPL [21].

Vanilla é uma plataforma que agrega várias ferramentas com uma grande variedade de componentes, possibilitando definir a sequência completa do negócio [22]. A ferramenta *Vanilla* possui várias funcionalidades como possibilitar a análise de dados, a criação de gráficos, a definição de KPIs, a aplicação de processos de ETL e *Data Mining*, e desenho de processos de negócios complexos [21, 22]. Além disso, permite alocar diferentes projetos em apenas um servidor e disponibilizar o acesso aos vários tipos de relatórios e a um livro de histórico no seu portal.

A ferramenta *Vanilla* fornece um conjunto completo de módulos de desenvolvimento: *OLAP Schema Designer*, *BI Metadata Designer*, *BI Dashboard Designer*, *BI Workflow Designer*, *KPI Metrics Designer*, *KPI Maps Designer* e *Report Designer Plugins* para *iReports* e *BIRT*. Para possibilitar a integração de metadados, esta *suite*, recorre às ferramentas *BIRT* e *iReports*. Esta plataforma disponibiliza também uma versão para sistemas operativos *Android*. *Vanilla*, destaca-se, assim, por, além de oferecer suporte a todas as funcionalidades de *Business Intelligence*, também fornecer uma versão para dispositivos móveis. Por conseguinte, esta ferramenta torna-se num instrumento com grandes perspetivas de evolução, devido a esta migração para este tipo de dispositivos [21].

A ferramenta *Vanilla* é, portanto, uma plataforma robusta, completa e com a vantagem de ser completamente *open source*.

A figura 3.3 expõe exemplos de gráficos e *dashboards* disponíveis na ferramenta *suite Vanilla*.

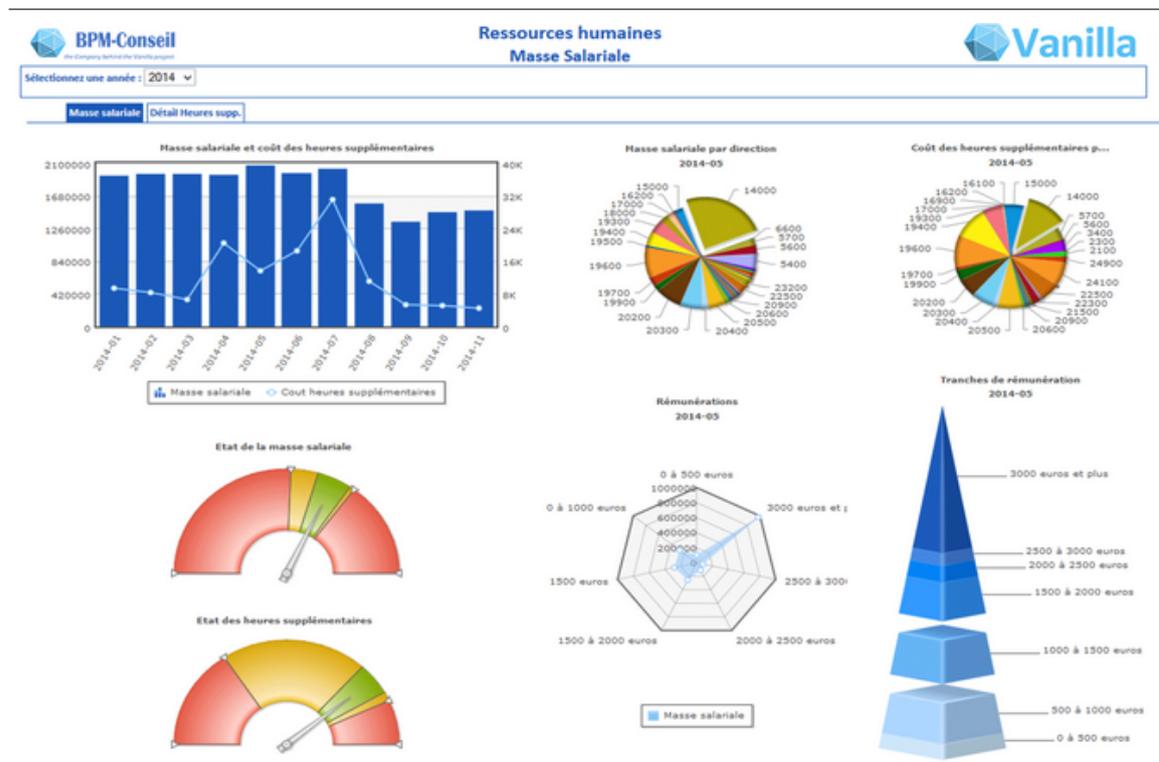


Figura 3.3: Exemplo de *dashboards* e gráficos interativos na *suite Vanilla*, retirado de [8]

3.2.4 Palo/Jedox

A ferramenta *Palo*, também denominada por *JPalo* (*Java Palo*), foi concebida pela empresa alemã *Jedox AG*.

Palo é uma plataforma desenvolvida em linguagem Java, totalmente gratuita e distribuída sob uma licença GNU GPL. Esta ferramenta é disponibilizada em duas versões, a *Enterprise* (comercial) e a *Community* (*open source*).

A *suite* é constituída pelos seguintes componentes: *Palo OLAP Server*, *Palo Web*, *Palo ETL Server* e *Palo para Excel* [23]. *Palo OLAP Server* é um servidor multidimensional OLAP, que permite trabalhar em tempo real num ambiente multi-sessão e oferece estabilidade, desempenho e algoritmos de lógica inovadores. A *interface Palo Web* possibilita a criação e a gestão de relatórios OLAP, bem como a visualização de processos ETL. *Palo ETL Server* possibilita a criação e o controlo de projetos de integração de dados. Existem *plugins Palo* para *Excel* da *Microsoft Office* e para *Calc* da *Open Office* [22]. Estes *plugins* juntamente com um servidor de base de dados OLAP constituem uma solução completa de BI [22, 21]. Esta *suite Palo*,

que é baseada em folhas de cálculo, permite integrar dados resultantes de folhas de cálculo de *Excel* e de *Calc*. As funcionalidades que se destacam nesta plataforma são a integração de dados ETL e as consultas OLAP [21].

A figura 3.4 apresenta um exemplo de um relatório interativo com diversos gráficos efetuado na ferramenta *Palo*.

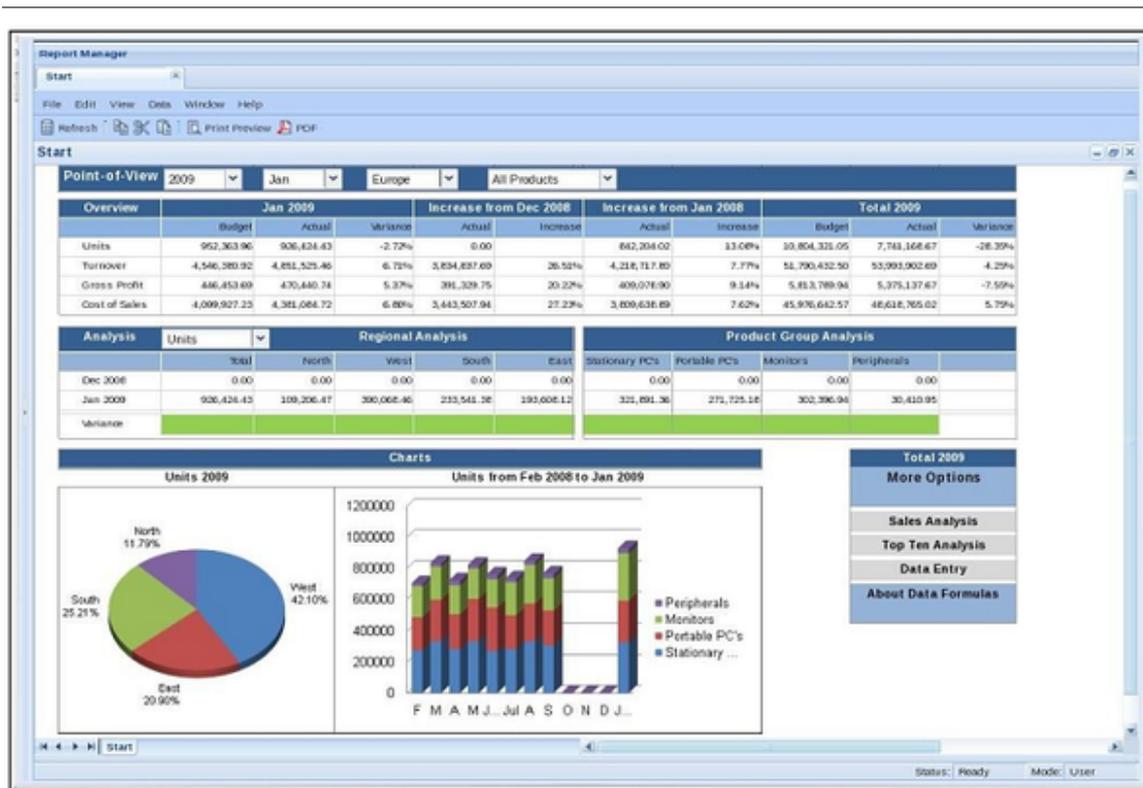


Figura 3.4: Exemplo de relatório e de gráficos interativos na *suite Palo/Jedox*, retirado de [9]

3.2.5 Pentaho Business Intelligence

A ferramenta Pentaho foi desenvolvida pela empresa, fundada em 2004, *Pentaho Corporation*.

Esta ferramenta, que é desenvolvida em linguagem Java, detém uma licença GNU GPL. *Pentaho* é uma ferramenta disponibilizada em duas edições: a *Community* e a *Enterprise*. A versão *Community* é completamente gratuita, disponibiliza apenas parte das funcionalidades em relação à outra versão, e é distribuída sob as licenças GNU LGPL, GNU GPL e MPL (versão *open source*). A edição *Enterprise* (versão

comercial) contém todas as funcionalidades, sob uma licença proprietária com custos para o utilizador [21, 22].

A ferramenta *Pentaho* permite realizar análises, criar e visualizar diversos gráficos, relatórios e *dashboards*. As aplicações *Mondrian* suportam o cruzamento de dados OLAP e o módulo *Weka* suporta as propriedades de *Data Mining*.

A figura 3.5, que se segue, apresenta a arquitetura que ilustra o conjunto de funcionalidades disponibilizadas pela ferramenta *Pentaho*. Estas funcionalidades estão agrupadas em 4 conjuntos distintos: Relatórios (produção, operacional e *ad-hoc*), Análises (*Data Mining*, OLAP e *drill/explore*), *Dashboards* (métricas, KPIs e alertas) e Processos de Gestão (integração, definição e execução).

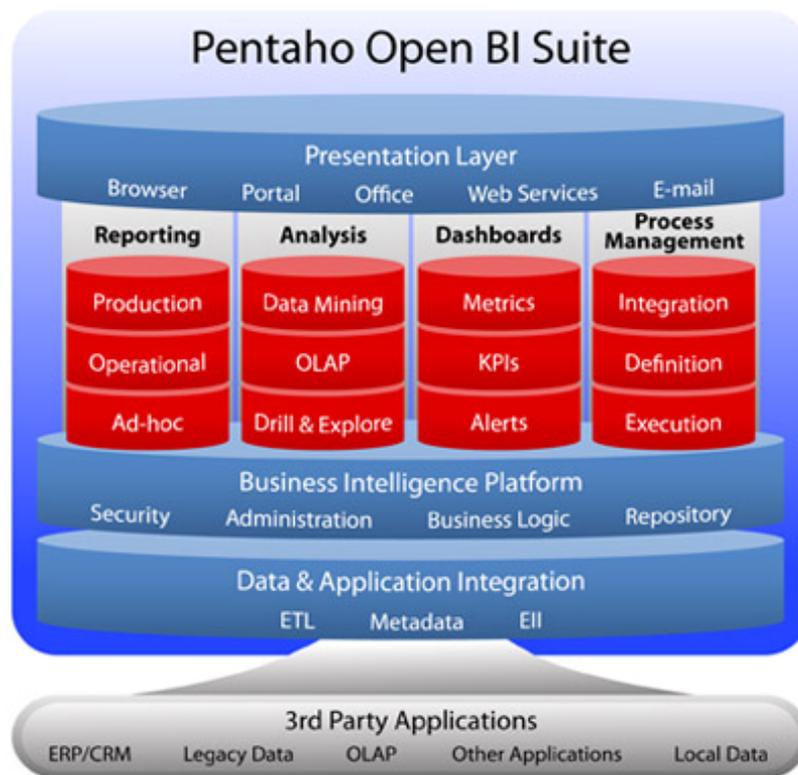


Figura 3.5: Arquitetura da ferramenta *suite* Pentaho BI, retirado de [10]

A *Pentaho BI Community Edition (CE)*⁸ inclui processos de ETL, análises OLAP, metadados, *Data Mining*, relatórios e *dashboards*. Esta versão é composta pelos módulos *Pentaho BI Server* (servidor), *Pentaho Data Integration* (ETL), *Pen-*

⁸<http://community.pentaho.com/>

taho Reporting (Relatórios), *Pentaho Dashboards* (Painéis de instrumentos), *Pentaho Analysis Services* (OLAP) e *Pentaho Data Mining* (Mineração de dados) [22].

Esta ferramenta apresenta todas as funcionalidades de um pacote de BI e consiste nos seguintes componentes [22]:

- ***Pentaho Data Integration (PDI)***: também conhecido como *Kettle*, é uma ferramenta poderosa que permite a execução do processo ETL de forma fácil e intuitiva, para repositórios de dados como *Data Warehouse*. Esta componente suporta as diversas fontes de dados devido a uma grande biblioteca de objetos de mapeamento, e permite o armazenamento de dados em *Data Warehouses*, bem como em outras bases de dados ou ficheiros. Este módulo é composto pelos seguintes blocos de construção: *Step*, *Transformation* e *Job*. O *Step* tem uma entrada de dados e uma saída de dados, e cada um apresenta uma determinada função como a de alterar ou filtrar dados. Uma *Transformation* é um conjunto de *steps* ligados entre si, sendo obrigatório um *step* de entrada e outro de saída para ocorrer a transformação dos dados. Um *Job* constitui pelo menos uma transformação a ser executada pela ordem definida;
- ***Pentaho Analysis Services (PAS)***: é um conjunto de serviços composto por 4 componentes: *Mondrian OLAP Engine*, *Pentaho Schema Workbench (PSW)*, *JPivot* e *Pentaho Aggregation Designer*.

O ***Mondrian OLAP Engine*** é um motor OLAP que permite obter análises a cubos de informação, no qual recebe consultas MDX provenientes de, por exemplo, *JPivot* e *Saiku Analytics*, e retorna um conjunto de resultados multidimensionais. Este motor está incluído no servidor *Pentaho (Pentaho BI Server)* e é baseado numa arquitetura ROLAP. PAS inclui a ferramenta **PSW** que disponibiliza uma *interface* gráfica para criar os esquemas de cubos OLAP, de acordo com o esquema multidimensional definido numa fase anterior, sendo armazenados no formato XML. Os cubos gerados nesta ferramenta são utilizados como suporte ao *Mondrian*, que interpreta MDX e traduz essas consultas em consultas SQL dos principais SGBD. A linguagem MDX é uma linguagem que fornece uma sintaxe específica para a consulta e a manipulação de dados multidimensionais armazenados em bases de dados OLAP. Esta linguagem também retrata as seleções, os cálculos e as definições de dados. Este conjunto de serviços é composto, também, é composto por análises *front-end JPivot* que é uma ferramenta de análise de cubos OLAP, baseada em *Java*. Como alter-

nativa a esta ferramenta, tem-se o *plugin Saiku Analytics*. Por último, PAS é constituído pela ferramenta ***Pentaho Aggregation Designer***, que permite gerar tabelas agregadas de modo a acelerar o mecanismo analítico.

- ***Pentaho Reporting***: é um módulo que suporta diversas fontes de dados, bem como OLAP e a linguagem de marcação XML. Este componente é constituído por uma ferramenta que permite gerar relatórios, denominada por *JFreeReport*, e por uma ferramenta para gerar metadados, que possibilita criar relatórios *ad-hoc*, através de um ambiente *web* e exportá-los em vários formatos;
- ***Pentaho Data Mining***: também conhecido como *Weka*. Este módulo é um conjunto de ferramentas relacionadas com a mineração de dados. Sendo assim, este componente fornece uma análise preditiva e um estudo de dados de modo as organizações obterem uma melhor compreensão do negócio e a melhorarem o seu desempenho futuro. Estas análises e estudos são fundamentados num conjunto de regras de classificação, regressão, segmentação, associação e algoritmos de *clustering*;
- ***Pentaho Dashboard***: permite criar painéis de controlo e reunir na mesma janela os principais indicadores da organização ou de um determinado departamento, com base em indicadores de desempenho (KPIs). Este módulo permite uma monitorização contínua, disponibilizando alertas de notificação de problemas;

A ferramenta *Pentaho* disponibiliza o *Pentaho User Console* (PUC) que é uma ferramenta muito poderosa e completa para a criação, edição e visualização de gráficos. O PUC é uma *interface web* disponibilizada pelo Servidor BI do *Pentaho*. Existem alguns clientes *web* disponíveis na forma de *plugin* para este servidor do *Pentaho*. Estes *plugins* são explorados através da aplicação *front-end* PUC. De entre eles, destacam-se os seguintes:

- ***Community Dashboard Edition***⁹ (CDE): é um *plugin* que permite um desenvolvimento e implementação de *dashboards*. CDE foi projetado para simplificar a criação, o *design* e os processos de tratamento de *dashboards* de *Pentaho CTools Dashboards*¹⁰. Este *plugin* é composto por três menus de configuração:

⁹<http://www.webdetails.pt/ctools/cde/>

¹⁰*Pentaho CTools Dashboards* é um conjunto de ferramentas e componentes que permitem criar *dashboards avançados*

Layout, *Datasources* e *Components*. O menu *Layout* permite definir o *layout* do gráfico/*dashboard*. O menu *Components* serve para adicionar e configurar os diferentes componentes que compõem os gráficos/*dashboards*. E o menu *Datasources* possibilita definir as fontes de dados utilizadas pelos componentes definidos em *Components*. CDE proporciona a integração com o *Pentaho Reporting* e o *Pentaho Analysis*. Porém, CDE opera independentemente do *Pentaho Dashboard Designer*, no qual os *dashboards* não são compatíveis uns com os outros. Este módulo permite gerar gráficos com suporte à biblioteca *JFreeChart*, que contém diversos gráficos como em 2D, 3D e em barras. Além disso, permite o acesso a diversas fontes de dados como bases de dados, cubos OLAP, metadados, ficheiros XML e transformações (*Kettle*).

- ***Saiku Analytics***¹¹ é uma *interface* gráfica fácil de entender. Este *plugin* é utilizado para efetuar análises de dados multidimensionais e para criar gráficos com recurso aos cubos OLAP. *Saiku Analytics* utiliza o motor OLAP *Mondrian* para facilitar o recurso de cubos OLAP de forma simples para o utilizador. *Saiku Analytics* permite a conexão a uma grande variedade de fontes de dados como *MySQL*, *MongoDB*, *Oracle Database*, *Microsoft SQL Server* e *PostgreSQL*, permitindo, assim, explorar os dados em tempo real diretamente da fonte.

De referir, que para executar o *Pentaho BI Server* é necessária uma instância *Tomcat*, o *Apache Tomcat* já pré-configurado, que atua como servidor *web* de sistemas desenvolvidos em *Java*.

A ferramenta de *BI open source Pentaho* é caracterizada por ser fácil de utilizar, bastante flexível, fiável, segura e eficiente. Além disso, esta ferramenta fornece um conjunto prestigioso de funcionalidades essenciais para análises rápidas e eficientes de dados.

A ferramenta *Pentaho* é uma das ferramentas de *Business Intelligence* mais poderosas do mercado de BI e, atualmente, das mais utilizadas [22, 21].

A figura 3.6 apresenta alguns exemplos de gráficos e *dashboards* disponíveis na ferramenta *suite Pentaho BI*.

¹¹<http://www.meteorite.bi/products/saiku>

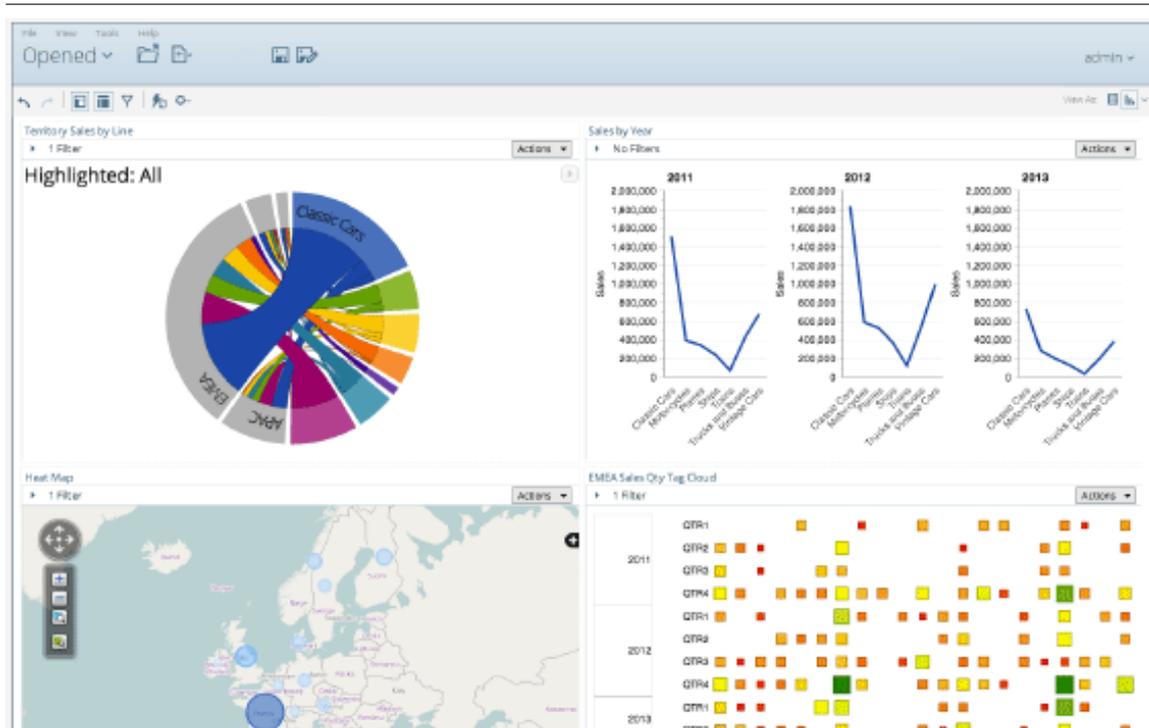


Figura 3.6: Exemplo de gráficos e *dashboards* interativos na *suite* Pentaho BI, retirado de [11]

3.3 Tabela Comparativa das Ferramentas do tipo OLAP

Após a realização de uma análise literária de ferramentas de BI *open source* e de uma descrição de algumas dessas ferramentas do tipo *suite*, procedeu-se a uma comparação entre estas mesmas.

Devido à grande diversidade de ferramentas de BI *open source* existentes no mercado de BI, optou-se, primeiramente, por ferramentas *suites* por possuírem aplicações servidor e cliente. Em seguida, tornou-se essencial a seleção de determinados critérios a ter em conta para a escolha de apenas uma ferramenta. Estes critérios selecionados, consoante alguns fatores, foram utilizados para uma comparação demonstrada através de uma tabela (tabela 3.2). Esta tabela serve para facilitar a escolha da ferramenta a ser utilizada nesta dissertação.

Durante a análise e a descrição das ferramentas verificaram-se alguns aspetos frequentemente referenciados. Destes aspetos destacam-se os seguintes: os sistemas

operativos que suportam as ferramentas, as licenças que possuem, as linguagens de programação utilizadas para o seu desenvolvimento, características como OLAP, ETL e *Data Mining*. Aspectos como estes tornaram-se como critérios para a comparação e seleção de uma ferramenta de BI *open source* do tipo *suite*.

Os dados da tabela 3.2 foram obtidos através de informações constituintes de diversas fontes durante a revisão literária [19, 20, 21, 22, 23, 24, 25].

A tabela 3.2, que se segue, apresenta uma comparação das principais características e funcionalidades relativas às seguintes ferramentas de BI *open source*: *SpagoBI*, *Pentaho*, *JasperSoft*, *Palo/Jedox* e *Vanilla*. É de evidenciar que, esta tabela, consistiu numa estrutura de apoio, utilizada para simplificar a seleção de uma destas ferramentas, pelo que contém as informações mais essenciais. Porém, esta tabela pode não conter a informação completa relativamente a uma determinada característica e/ou funcionalidade.

Tabela 3.2: Funcionalidades e características de ferramentas *open source*, baseado em [19, 20, 21, 22, 23, 24, 25]

Funcionalidades/Caraterísticas	SpagoBI ¹	Pentaho ¹	Jasper-Soft ¹	Palo/Jedox ¹	Vanilla ¹
Fontes de dados	Várias, ex.: Oracle, PostgreSQL, DB2.	Várias, ex.: MySQL, PostgreSQL, folhas de cálculo.	Várias, ex: bases de dados relacionais, bases de dados <i>flat-file</i> ¹² .	Várias, ex.: SGBD relacional, SAP, serviços web, Microsoft Excel e Open Office Calc.	Várias, ex.: fontes de dados SQL e Microsoft Access, e fontes dados externas (XML, XLS,...)

¹✓- Possui funcionalidade/caraterística | ✗- Não Possui funcionalidade/caraterística | N/A - Não Aplicável ou disponível | W. - *Windows* | L. - *Linux* | U. - *Unix*

¹²Bases de dados *flat-files* são bases de dados constituídas por apenas uma tabela

Informação de output (formatos)	relatórios (ex.: PDF, TXT, XLS, HTML, XML, CSV e RTF) e gráficos (ex.: SWF)	relatórios, <i>dashboards</i> , gráficos (ex.: HTML, PDF, XLS, CSV, RTF e plain text)	relatórios (ex.: PDF, HTML, XLS, SWF, DOC, RTF e ODS/ODT), XHTML, ficheiros de texto, Excel, Word, OpenOffice	relatórios (ex.: PDF)	N/A
Sistemas Operativos suportados¹	W., L. e U.	W., L. e U.	W. e L.	W. e L	W, L e U
Licença	GNU LGPL e MPL	GNU GPLv2 e EPLv1	GNU GPL	GNU GPL	GNU GPL e MPL
SGBD suportado/-Conetividade	MySQL, Oracle, PostgreSQL, etc.	MySQL, PostgreSQL, Oracle, Hypersonic, etc.	MySQL, Oracle, PostgreSQL, MongoDB, Hadoop-Hive e Bean, etc.	MySQL, ect.	Base de dados SQL, incluindo Microsoft Access
Conetividade	JDBC e JNDI	JDBC, JNDI, ODBC, OCI	JDBC, JNDI	JDBC, ODBC	JDBC
Método de Armazenamento/análises	ROLAP	ROLAP	ROLAP	MOLAP	N/A
Acesso a dados	XML/A	XML/A	XML/A, Hibernate, EJB, POJO ou tipos personalizados	XML/A	XML

Linguagem de Consulta	MDX	MDX	MDX padronizado para <i>web</i> ou <i>interface</i> em folhas de cálculo Excel	MDX	MDX
Usabilidade	Amigável	Muito amigável	Amigável	Amigável	N/A
Linguagens de desenvolvimento	Java	Java	Java, Perl	Java	PHP
Relatórios	✓	✓	✓	✓	✓
Dashboards	✓	✓	✓	✓	✓
Consultas Ad-hoc	✓	✓	✓	✓	✓
OLAP	✓	✓	✓	✓	✓
Data Mining	✓	✓	✗	✗	✓
Exportação de dados	✓	✓	✓	✓	✓
Versão móvel	✓	✓	✓	✓	✓
Servidor OLAP	motores OLAP, Mondrian, XMLA Server	Mondrian	JasperReports Server	Palo OLAP Server	-
Cliente OLAP	Jpivot e Jpalo	Jpivot, Saiuku Analytics, CDE	-	Palo Client	N/A
ETL/ferramenta	✓(TOS (Talend Open Studio))	✓(PDI)	✓(Talend)	✓(Palo ETL Server)	✓(BI Gateway)
GEO	✓	✓	✓	✗	✗
KPIs	✓	✓	✗	✗	✓

¹✓- Possui funcionalidade/caraterística | ✗- Não Possui funcionalidade/caraterística | N/A - Não Aplicável ou disponível | W. - *Windows* | L. - *Linux* | U. - *Unix*

3.4 Revisão literária de ferramentas de BI *open source*

De modo a apoiar a realização da pesquisa e análise comparativa de ferramentas de *Business Intelligence open source* foi realizada uma revisão literária de trabalhos que abordaram esta temática, realçando os que se seguem.

Inicialmente no contexto *Business Intelligence*, as primeiras ferramentas eram individuais. C. Thomsen e T. B. Pedersen em [81] fizeram uma pesquisa em ferramentas *open source* que não são consideradas plataformas. Nesta investigação os autores realizaram um estudo em três ferramentas de ETL, em três servidores OLAP e dois clientes OLAP. De acordo com determinados critérios definidos pelos autores, os mesmos selecionaram um conjunto de ferramentas. Os principais critérios para as ferramentas ETL foram as fontes de dados que a ferramenta suporta, o tipo de sistemas (ROLAP e MOLAP) a que se destinam os carregamentos de dados, sendo selecionadas as ferramentas *Bee*, *CloverETL* e *Octopus*. Para a escolha de um cliente OLAP deve ser considerado quais os servidores OLAP que podem ser utilizados com o cliente. Quanto ao servidor deve-se ter em conta as APIs e as linguagens de consulta que o cliente OLAP suporta. Como servidores OLAP consideraram os servidores *Bee*, *Lemur* e *Mondrian* e como clientes OLAP selecionaram *Bee* e *JPivot*. Os clientes *Bee* e *JPivot* são para serem utilizados com os servidores *Bee* e com o *Mondrian*, respetivamente. Neste artigo, os investigadores evidenciaram quatro SGBD *open source*: *MonetDB*, *MySQL*, *MaxDB* e *PostgreSQL*. Para eles, estas SGBDs são as mais notáveis e bem conhecidas pelo elevado desempenho.

Atualmente, a evolução da tecnologia relativamente a sistemas de BI, tem possibilitado o desenvolvimento de plataformas de BI completas, que integram um conjunto de funcionalidades numa solução única. Matteo Golfarelli em [25], procura entender se a utilização de ferramentas *open source* será apta de ser uma alternativa válida em relação a ferramentas comerciais num contexto de BI. Neste artigo é realizada uma avaliação comparativa entre três plataformas de *Business Intelligence open source*: *JasperSoft*, *Pentaho* e *SpagoBI*. O estudo destas ferramentas teve como objetivo a compreensão das características, potencialidades futuras e os limites quando adotadas em projetos reais. Os autores não se focam no estudo profundo das características de cada uma. Eles basearam-se num conjunto de critérios para comparação, tal como a filosofia da plataforma, tipo de licença e edições disponibilizadas (aspectos

não técnicos). Além disto, fundamentaram-se na arquitetura global e nos módulos constituintes e nas suas relações, nos sistemas operativos suportados e nas linguagens de programação. Os autores, ainda, basearam-se na segurança e na autenticação de perfis dos utilizadores e de dados; na usabilidade (tanto do ponto de vista do utilizador em relação à utilização da ferramenta como do ponto de vista da complexidade na instalação e na administração). Nesta investigação, os autores realizaram uma comparação entre as principais funcionalidades disponibilizadas pelas plataformas, tais como fornecimento de relatórios *ad-hoc*, do processo ETL e de análises OLAP. Os autores concluíram que as ferramentas *open source* do tipo *suite* trazem vantagens relativamente a ferramentas de BI individuais. E isto devido ao facto de as ferramentas *open source* permitirem diversas funcionalidades para serem acedidas de forma transparente, e possuírem um conjunto de processos que as tornam centralizadas e simplificadas, reduzindo a administração e o esforço de desenvolvimento.

Em [21], M. Tereso e J. Bernardino efetuaram uma avaliação e comparação de seis ferramentas de *Business Intelligence open source*: *SpagoBI*, *OpenI*, *Pentaho*, *Jaspersoft*, *Palo* e *Vanilla*. Os autores referem a importância de ferramentas de BI para o apoio na tomada de decisão nas organizações. Eles procuraram soluções exequíveis de ferramentas *BI open source* para pequenas e médias empresas. Este artigo focou-se na descrição de cada uma das ferramentas e numa comparação das mesmas em relação às suas funcionalidades, características e sistemas operativos. Neste estudo, os autores referem que *Pentaho* é uma ferramenta consistente e bastante completa a par da *SpagoBI*. Como conclusão salientam o facto do recurso a ferramentas de BI permitir aos gestores e analistas, obterem conhecimento extraído a partir da informação das empresas. Ainda salientam, que é importante estipular quais as necessidades do processo de negócio antes da escolha da ferramenta [21].

3.5 Conclusão do estudo e análise das ferramentas de BI

Este subcapítulo apresentou uma visão geral sobre critérios para avaliação de produtos OLAP, demonstrando um conjunto de doze regras definidas por Codd, que acabaram por serem uma forma inadequada e imparcial de avaliar as ferramentas OLAP. Por conseguinte, isto levou à otimização do conceito de OLAP por Nigel Pendse, surgindo assim o teste FASMI. Este conceito é mais simples, fácil de memorizar e apresenta uma maior conformidade entre as ferramentas do tipo OLAP

pertencentes a empresas distintas.

Em seguida, foi apresentada uma descrição de um conjunto de ferramentas de BI *open source* do tipo OLAP mais utilizadas pela comunidade científica e empresarial encontradas durante a revisão literária. Por último, foi demonstrada uma comparação entre as principais funcionalidades e características dessas ferramentas através de uma tabela comparativa (3.2). Esta tabela demonstra que todas as ferramentas analisadas têm em comum determinadas funcionalidades como a possibilidade de criar relatórios, *dashboards*, consultas *ad-hoc* e a exportação de dados. Em relação à técnica de *Data Mining*, as ferramentas *JasperSoft* e *Palo/Jedox* não possuem, enquanto que as restantes ferramentas detêm essa capacidade. Relativamente às informações de *input* e *output* todas as ferramentas permitem reunir informações a partir de várias fontes, e obter dados em diversos formatos, respetivamente. Além disso, esta tabela apresenta os sistemas operativos que suportam cada uma das ferramentas, no qual se verifica que os sistemas operativos *Windows* e *Linux* possuem a capacidade de suportar todas as ferramentas analisadas. Porém as ferramentas *SpagoBI*, *Pentaho* e *Vanilla* são também suportadas por sistemas operativos *Unix*. Relativamente ao tipo de licença que cada ferramenta *open source* possui, observou-se que a maioria das ferramentas detêm licenças do tipo GNU GPL, havendo ferramentas possuidoras de outras licenças como GNU LGPL, GNU MPL. Através das diversas pesquisas efetuadas, verificou-se que as ferramentas suportam diversas bases de dados, e que todas apresentam, pelo menos, conectividade JDBC, a qual permite o envio de instruções SQL para qualquer base de dados relacional. Quanto ao método de armazenamento, observou-se que os servidores da maioria das ferramentas utilizam a arquitetura ROLAP. Quanto ao acesso a dados, todas as ferramentas permitem aceder a dados através de XML e/ou XMLA¹³. A linguagem de consulta que o servidor de cada ferramenta suporta é a MDX, a qual permite a consulta a cubos de dados multidimensionais. A tabela também compara as linguagens de programação utilizadas no desenvolvimento de cada ferramenta, verificando-se que à exceção da ferramenta *Vanilla* desenvolvida em PHP, as restantes são implementadas em Java, sendo que *JasperSoft* é também desenvolvida em *Perl*. Quanto ao processo de ETL, verificou-se que todas as ferramentas possuem esta funcionalidade. A tabela permite verificar que as ferramentas *Palo/Jedox* e *Vanilla* não permitem análise geográfica (GEO), ao contrário das restantes. Por último, a tabela mostra que as ferramentas

¹³XMLA é um padrão para acesso a dados em sistemas analíticos como OLAP e *Data Mining*, permitindo às aplicações cliente comunicarem com bases de dados dimensionais

Jaspersoft e *Palo/Jedox* não suportam KPIs.

Após uma análise às ferramentas de BI *open source* do tipo *suite* optou-se pela ferramenta *Pentaho*. A seleção desta ferramenta teve em conta, diversos aspetos como a escolha de uma ferramenta compatível com o sistema operativo utilizado (*Windows*), as funcionalidades disponíveis, as bases de dados suportadas, a facilidade de configuração e utilização. Além disso, esta seleção da ferramenta de análise OLAP foi fundamentada de acordo com os requisitos e necessidades pretendidos para o desenvolvimento da plataforma.

Pentaho é uma ferramenta de grande potencial e bastante poderosa. Esta ferramenta é muito utilizada pela comunidade empresarial, na medida em que serve como instrumento de suporte para a definição de estratégias de competitividade nos negócios de uma empresa ou organização. O *Pentaho* suporta as principais características e funcionalidades que se pretendem, com uma configuração relativamente simples. Além disso, possui uma grande comunidade de utilizadores e desenvolvedores, o que a torna com maior capacidade em adaptar-se a novas tecnologias, podendo eventuais *bugs* serem corrigidos mais rapidamente. A ferramenta *Pentaho* permite organizar e analisar uma grande quantidade de dados de forma rápida e eficiente. *Pentaho* possui um conjunto de componentes bastante completos que contemplam todas as fases de um projeto de *Business Intelligence*.

Capítulo 4

Especificação do sistema

Este capítulo demonstra os requisitos levantados durante a fase de análise e identifica os requisitos para a implementação do sistema, dos quais estão incluídos os requisitos não funcionais e os requisitos funcionais (secção 4.1). O levantamento e identificação de requisitos teve em conta aspetos, tais como: os objetivos, as necessidades para a elaboração da plataforma *online* e as características/funcionalidades disponíveis nas ferramentas do tipo *Business Intelligence open source*. Em seguida, após o levantamento de requisitos, são apresentados os diagramas de casos de uso que demonstram a interação entre os atores (intervenientes) e o sistema (secção 4.2). Posteriormente, são expostos os principais esboços efetuados para uma melhor compreensão desses requisitos e dos serviços/funcionalidades a disponibilizar na plataforma *online* (secção 4.3). Além disso, por último, neste capítulo é apresentada uma visão geral sobre as componentes pertencentes ao sistema relativo à plataforma *online* (secção 4.4).

4.1 Levantamento e especificação de requisitos

O levantamento de requisitos é fundamental no início de qualquer atividade de desenvolvimento de *software*, de modo a dar conhecimento das principais funcionalidades, características, propriedades, restrições e atributos de qualidade associados ao sistema a desenvolver, auxiliando no planeamento e implementação do mesmo. O levantamento e especificação dos requisitos vai de acordo com as necessidades e com o tipo de tecnologias requeridas para o sistema. A especificação dos requisitos encontra-se dividida em: (i) requisitos não funcionais, que determinam a qualidade do sistema, tais como desempenho, segurança e disponibilidade, e em (ii) requisitos

funcionais, que fornecem as diversas funcionalidades do sistema.

4.1.1 Requisitos Não Funcionais

Os requisitos não funcionais, também denominados por atributos de qualidade do sistema, estão relacionados com a utilização da aplicação em termos de desempenho, usabilidade, fiabilidade, segurança, disponibilidade, tecnologias envolvidas, entre outros, que são fatores que determinam a qualidade do sistema desenvolvido. Estes requisitos não estão diretamente relacionados com as funcionalidades, porém, possuem um grau de importância durante o desenvolvimento de um sistema.

Assim, o sistema deve obedecer a um conjunto de requisitos, tais como:

- A aplicação deverá ter um desempenho favorável e satisfazer as necessidades do utilizador perante a utilização do sistema. O sistema deverá ser fácil de aprender e de utilizar, possibilitando uma execução rápida das tarefas;
- O sistema poderá sofrer modificações, mesmo depois da sua disponibilização, no qual engloba tanto atividades de reparação/correção de eventuais erros/defeitos detetados, por exemplo bugs, como atividades de alteração/melhoramento de determinadas características ou adição de novas funcionalidades não previstas na fase inicial do projeto;
- O tempo de resposta do sistema deve ser relativamente curto e aceitável pelos utilizadores;
- O sistema deve apresentar uma *interface* amigável, intuitiva e de fácil utilização. As funcionalidades e informações devem estar bem visíveis e disponíveis;
- O sistema deve estar preparado para manter a integridade da informação que circula pelo sistema;
- O sistema deverá ser rápido na resolução de falhas como a baixa em servidores;
- O sistema deverá ser capaz de se adaptar aos diversos *web browsers* e diferentes sistemas operativos, nomeadamente a *Windows*, *Mac OS* e *Linux*;
- O sistema deverá garantir a segurança dos dados dos utilizadores, não permitindo o acesso a dados não autorizados. Apenas pessoas registadas e autenticadas poderão ter acesso às informações que lhe são destinadas. Assim, a integridade do sistema, também, é garantida;

- O sistema deverá assegurar um algoritmo de encriptação dos dados, nomeadamente, das *passwords* do utilizador, de modo a garantir a segurança dos seus dados de *login*.
- O sistema deverá estar disponível a qualquer hora e disponibilizar as funcionalidades de acordo com o tipo de utilizador (com autenticação e sem autenticação);
- O carregamento e atualização do DW, do cubo OLAP e da base de dados devem ser instantâneos;
- O sistema deve ser desenvolvido com recurso a ferramentas *Business Intelligence* do tipo *suite open source*, que permitam funcionalidades OLAP, e ferramentas *open source* para a base de dados e para a programação *web*;

4.1.2 Requisitos Funcionais

Os requisitos funcionais descrevem as funções e serviços que o sistema deve facultar, a forma de como o sistema deve reagir perante entradas específicas e, também, o modo como se deve comportar perante determinadas situações. Estes requisitos representam uma interação entre o sistema e o seu ambiente. Sendo assim, estes requisitos determinam as funcionalidades, serviços, características e ações que o sistema deve proporcionar.

Os requisitos funcionais gerais do sistema são os seguintes:

(i) **Sistema do utilizador sem *login* efetuado:**

- O sistema deve permitir ao utilizador efetuar o registo para a autenticação (*login*) na plataforma *online*;
- O sistema enviará um *e-mail* automático de confirmação após o utilizador efetuar o registo. Em seguida o administrador, através da aplicação para o efeito, efetuará a autorização ou não autorização do *login* do utilizador, sendo também enviado um *e-mail* automático com o resultado do pedido do registo;
- O sistema permitirá ao utilizador enviar mensagens para o administrador;
- O sistema apenas disponibilizará alguns dos gráficos analíticos relativos às iniciativas em Cidades Inteligentes;

(ii) Sistema do utilizador com *login* efetuado:

- O sistema deve permitir somente a autenticação no sistema a utilizadores autorizados pelo administrador a efetuar *login*, e que, além disso, inseriram corretamente os dados (*username: e-mail* e *password*). Somente os utilizadores autorizados pelo administrador e devidamente autenticados é que possuem acesso às várias funcionalidades oferecidas, como a inserção de iniciativas através e um formulário *online*. Na autenticação, caso os dados inseridos estejam corretos, será iniciado a sessão no sistema, caso contrário, será mostrada uma mensagem de erro, porém, poderá efetuar novamente o *login*;
- O sistema deve permitir o *login* e *logout* dos utilizadores devidamente registados e autorizados;
- O sistema deverá associar uma sessão logo após o utilizador se autenticar, de modo a garantir a segurança da aplicação do utilizador;
- O sistema deve permitir a inserção de iniciativas através de um formulário *online*;
- O sistema deve contar com uma base de dados para armazenar os dados do formulário *online* e dos utilizadores registados;
- O sistema deve permitir listar todos os casos de estudo/iniciativas em Cidades Inteligentes e visualizar em detalhe cada um, perante a realização do *login*;
- O sistema permitirá ao utilizador enviar mensagens para o administrador;
- O sistema deve mostrar uma análise do número de iniciativas por:
 - ano;
 - *Responsible Institution* (Instituição Responsável);
 - categoria do atributo *Concepts* (Conceitos);
 - tipo do atributo *Innovations* (Inovações);
 - categoria do atributo *Lessons Learnt* (Aprendizagens);
 - tipo do atributo *Benefits* (Benefícios);
 - tipo do atributo *Drivers* (Fatores impulsionadores);
 - categoria do atributo *Values* (Valores);

- tipo de *Approach* (Abordagem);
 - categoria do atributo *Challenges* (Desafios);
 - tipo de atributo *Risks* (Riscos);
 - categoria do atributo *Tools* (Ferramentas);
 - tipo do atributo *Governance* (Governança);
 - tipo do atributo *Maturities* (Maturidade);
 - categoria do atributo *Technologies* (Tecnologias).
- O sistema deve mostrar um mapa de distribuição das cidades inteligentes;

(iii) **Sistema do Administrador:**

- O sistema do administrador deverá permitir efetuar o *login* e *logout*;
- O sistema deverá associar uma sessão logo após o administrador se autenticar, de modo a garantir a segurança da aplicação do administrador;
- O sistema do administrador deverá permitir a inserção de novos administradores;
- O sistema do administrador deve contar com uma base de dados para armazenar os dados dos administradores;
- O sistema do administrador deverá permitir efetuar a administração de utilizadores, nomeadamente, a gestão do acesso dos utilizadores ao conteúdo da plataforma *online*;
- O sistema do administrador deve permitir editar ou remover os diversos formulários armazenados na base de dados.

4.2 Descrição dos casos de uso

Nesta secção é apresentada uma descrição interativa do funcionamento da plataforma *online*, bem como do sistema administrativo, através de diagramas de casos de uso, de modo a retratar as ações permitidas a cada um dos atores (intervenientes) do sistema (utilizadores e administradores). No sistema relativo à plataforma *online* existem dois tipos de utilizadores: utilizador não autenticado (subsecção 4.2.1) e utilizador autenticado (subsecção 4.2.2). Quanto ao sistema do administrador, somente existe como interveniente o administrador (subsecção 4.2.3).

Os diagramas de casos de uso permitem descrever as funcionalidades a serem implementadas no sistema de acordo com o tipo de utilizador, ou seja, a interação entre o utilizador e o sistema. Assim, com recurso a estes diagramas procura-se especificar um conjunto de ações realizadas entre um ator e o sistema, sendo o resultado destas observáveis.

4.2.1 Utilizador sem *login* efetuado no sistema

O diagrama de caso de uso visualizado na figura 4.1 retrata as funcionalidades associadas a um utilizador sem *login* efetuado. Este tipo de utilizador apenas pode usufruir de algumas das funcionalidades cedidas pela plataforma *online*, tais como visualizar parcialmente algumas das análises estatísticas referentes às Cidades Inteligentes. Além disso, este tipo de utilizador pode efetuar o seu registo, necessitando de confirmar o seu *e-mail*/registo através de um *e-mail* enviado pelo sistema para o efeito. Assim que o administrador receber a confirmação do *e-mail*/registo do utilizador, poderá proceder à autorização para que possa efetuar o *login*. Também, é possível a este tipo de utilizador contactar o administrador do sistema, enviando-lhe uma mensagem, via *e-mail*, através do preenchimento de um formulário disponibilizado na plataforma *online*.

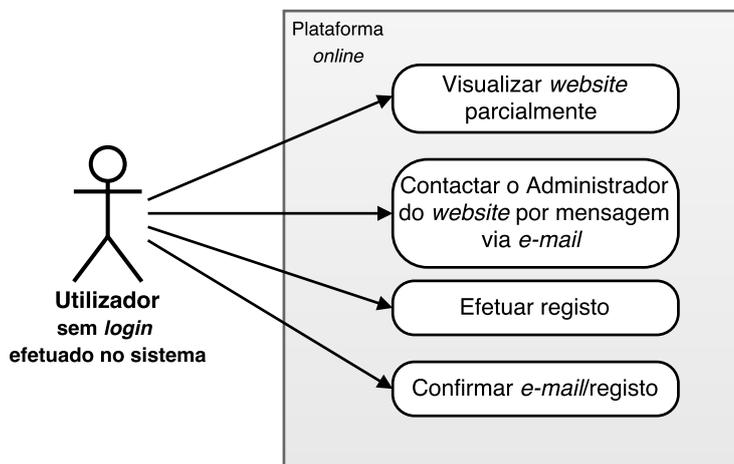


Figura 4.1: Diagrama de caso de uso - Utilizador sem *login* efetuado no sistema

4.2.2 Utilizador com autenticação no sistema

O diagrama de caso de uso relativo à figura 4.2, que se segue, expõe as funcionalidades referentes a um utilizador que tenha efetuado o seu registo e que poste-

riormente foi autorizado a realizar *login* pelo administrador. Este tipo de utilizador após efetuar a autenticação no sistema através do *login*, poderá usufruir de todas as funcionalidades proporcionadas pela plataforma *online*. Tal como o tipo de utilizador descrito anteriormente (subsecção 4.2.1), este utilizador pode contactar o administrador através da plataforma *online* com o envio de uma mensagem via *e-mail*. Além disso, este utilizador poderá preencher o formulário *online*, visualizar a listagem dos dados relativos aos casos de estudo/iniciativas inseridos na base de dados, consultar as análises estatísticas (gráficos) sobre esses dados e visualizar o mapa com a indicação das Cidades Inteligentes armazenadas nessa base de dados.

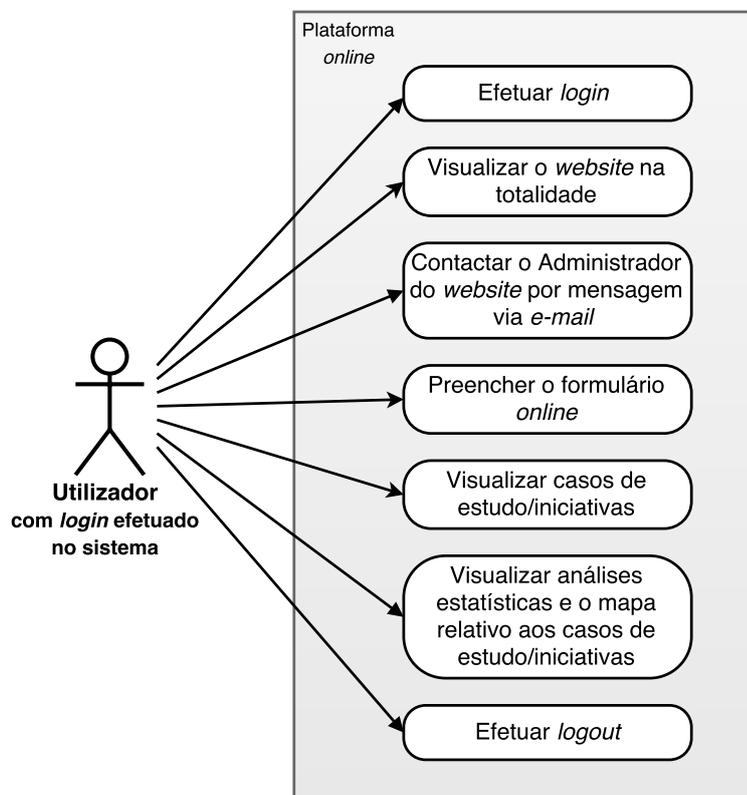


Figura 4.2: Diagrama de caso de uso - Utilizador autenticado no sistema

4.2.3 Administrador

O próximo diagrama de caso de uso (figura 4.3) descreve as funcionalidades referentes ao administrador do sistema, no qual este possui um *website* exclusivo à administração da plataforma *online*. O administrador após se autenticar no sistema poderá inserir outros administradores e ativar/desativar a permissão dos pedidos de

registro dos utilizadores. Também, é da responsabilidade dos administradores editar/eliminar a pedido dos utilizadores, e em concordância, os dados inseridos através do formulário *online* de modo a obter análises mais concisas.

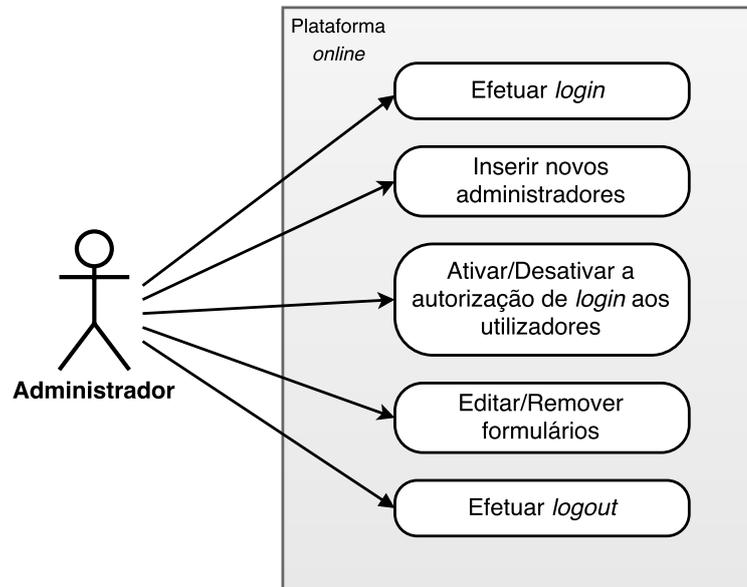


Figura 4.3: Diagrama de caso de uso - Administrador

4.3 Esboço da Plataforma *Online*

Posteriormente ao levantamento de requisitos, para uma percepção mais consolidada e coerente sobre a plataforma *online* para casos de estudo e iniciativas em Cidades Inteligentes, recorreu-se à elaboração de esboços sobre o sistema a desenvolver. Assim, estes esboços ajudarão a visualizar a *interface web* do sistema final e a auxiliar no desenvolvimento do mesmo, uma vez que permite planejar e simular a hierarquia de informações, o *layout* do *website* e as possíveis interações.

Nas figuras que se seguem serão expostos os principais exemplos dos esboços efetuados durante o planeamento da plataforma *online*.

A figura 4.4 apresenta o planeamento da *interface web* destinada a todos os utilizadores que não efetuam *login*. A página inicial abrangerá uma descrição do objetivo da plataforma, sendo possível através dos restantes separadores aceder a outros serviços/funcionalidades, tais como: efetuar o registo ou *login*, visualizar algumas análises referentes aos casos de estudo e às iniciativas em Cidades Inteligentes, e contactar o administrador através de mensagens via *e-mail*.

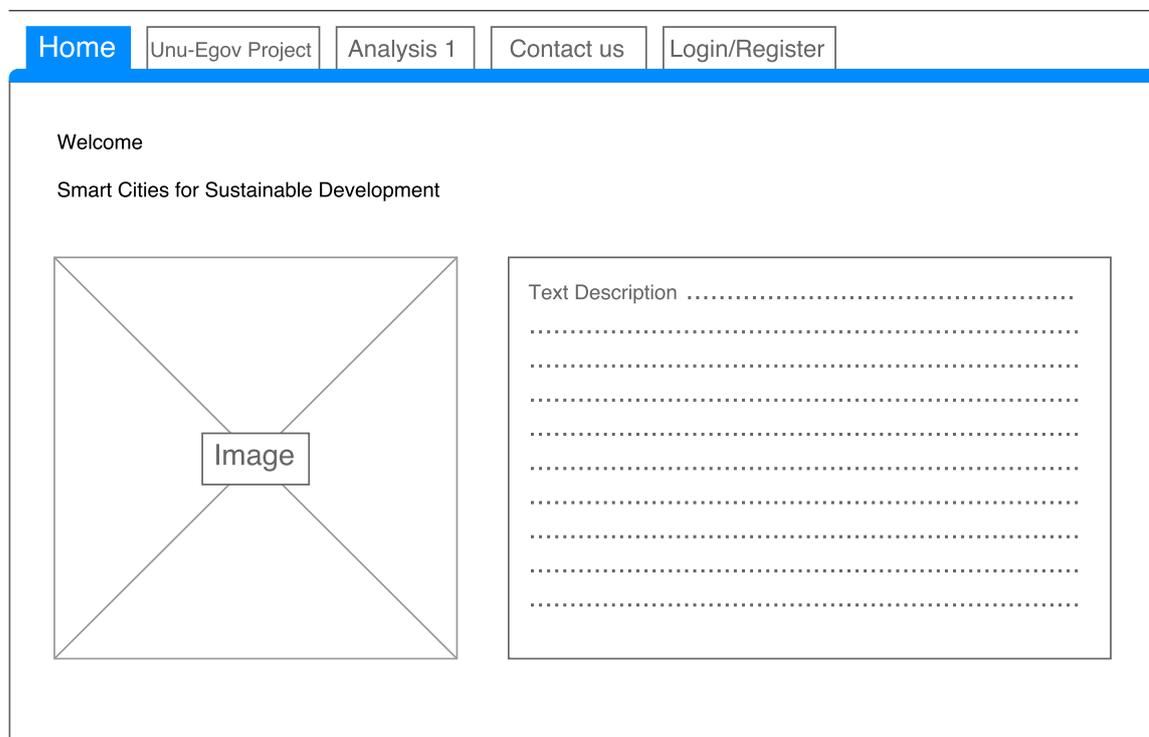
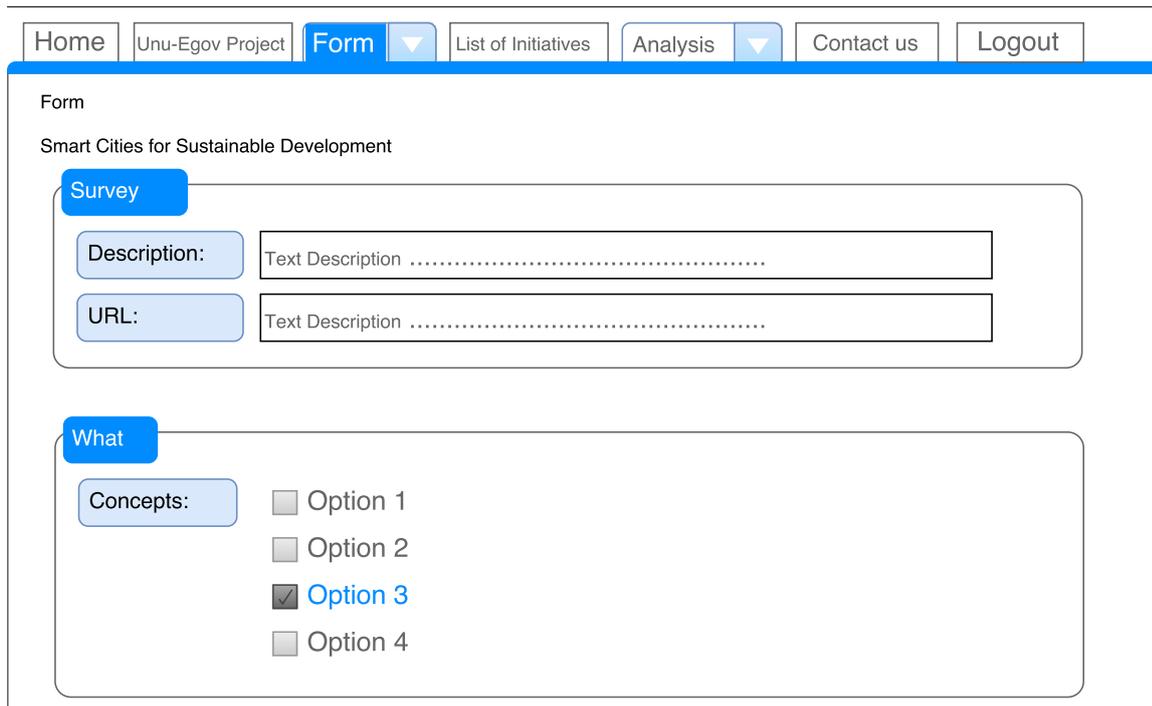


Figura 4.4: Esboço da plataforma *online* - Inicial

A figura 4.5 mostra o esboço da *interface web* de uma fração do formulário *online* destinado aos utilizadores que tenham efetuado o *login* com sucesso. Nesta imagem é possível observar, também, o planeamento dos separadores da plataforma apenas para utilizadores com sessão iniciada, no qual, apresenta mais serviços/funcionalidades para este tipo de utilizadores.



The image shows a web interface for a platform. At the top, there is a navigation bar with several buttons: 'Home', 'Unu-Egov Project', 'Form' (with a dropdown arrow), 'List of Initiatives', 'Analysis' (with a dropdown arrow), 'Contact us', and 'Logout'. Below the navigation bar, the main content area is titled 'Form' and 'Smart Cities for Sustainable Development'. There are two main sections: 'Survey' and 'What'. The 'Survey' section contains two text input fields: 'Description:' and 'URL:'. The 'What' section contains a 'Concepts:' label and four radio button options: 'Option 1', 'Option 2', 'Option 3' (which is selected and highlighted in blue), and 'Option 4'.

Figura 4.5: Esboço da plataforma *online* - *Login* efetuado: Formulário

A figura 4.6 expõe o esboço para a *interface web* referente aos gráficos analíticos, relativos aos dados obtidos do formulário *online*, desenvolvidos através da ferramenta OLAP *open source* selecionada. Nos diversos gráficos poderá ser visualizado o número de iniciativas referente aos vários domínios em cada ano.

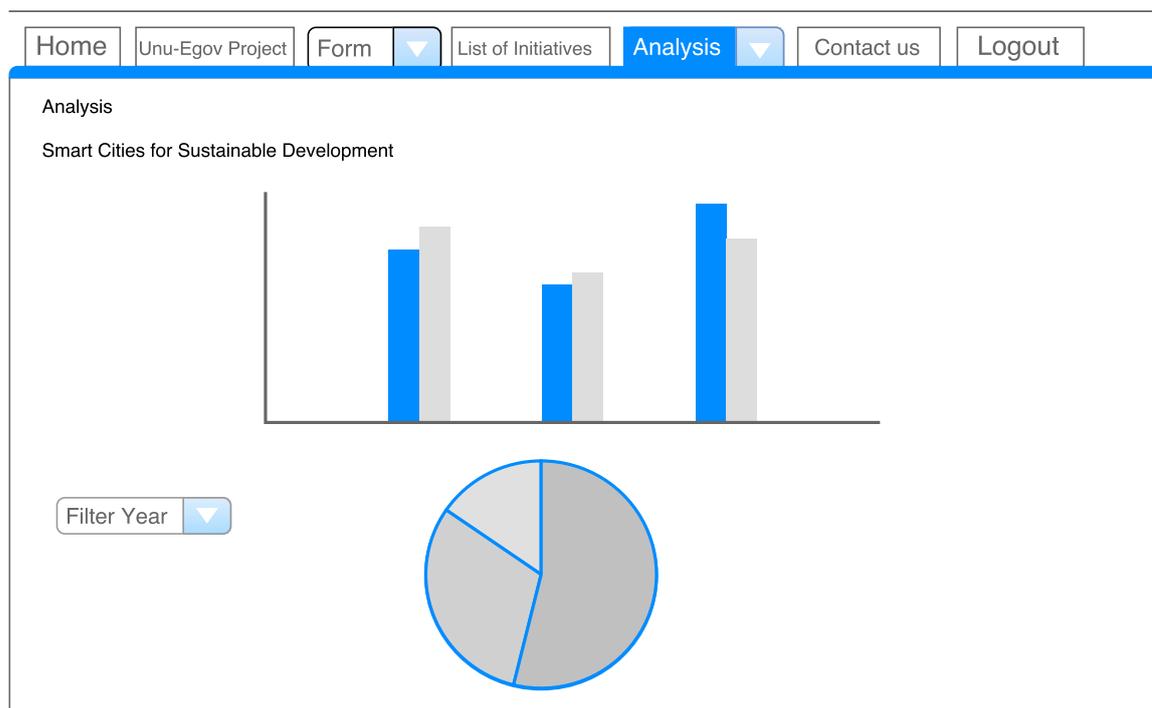


Figura 4.6: Esboço da plataforma *online* - *Login* efetuado: Análises (Gráficos)

4.4 Arquitetura do sistema

De modo a desenvolver o sistema proposto, que permitirá reunir e analisar informações acerca de cada iniciativa em Cidade Inteligente, têm sido fundamentais um conjunto de tarefas. Numa primeira fase foram estabelecidos os objetivos e a realização de uma análise e definição de requisitos e restrições, de acordo com as necessidades do sistema a desenvolver. Além disso, realizou-se uma análise de todo o sistema que permitirá uma melhor idealização do seu funcionamento e serviços a disponibilizar na plataforma *online*, e ainda o estudo de tecnologias e ferramentas possíveis a serem aplicadas para o desenvolvimento do sistema. Em consequência dessas fases foi realizada toda a especificação do sistema, nomeadamente, a arquitetura do sistema de acordo com as ferramentas e tecnologias a utilizar na elaboração do sistema.

A arquitetura da plataforma *online*, que tem por objetivo reunir, organizar e fornecer gráficos analíticos e análises de dados de casos de estudo e iniciativas em Cidades Inteligentes, é apresentada na figura 4.7, bem como as tecnologias utilizadas para a implementação e exploração do sistema. Como já foi referido na secção de le-

vantamento e definição de requisitos (secção 4.1) para este sistema, um dos requisitos é a utilização de ferramentas *open source* tanto para as análises OLAP, como para o desenvolvimento da plataforma *online*. Sendo assim, a arquitetura proposta para o desenvolvimento do sistema reflete a arquitetura geral de um sistema de *Business Intelligence* demonstrada no capítulo 2 (secção 2.2), e de acordo também, com as ferramentas *open source* e tecnologias definidas na fase de análise do sistema.



Figura 4.7: Arquitetura da plataforma *online*

A arquitetura apresentada na figura 4.7 representa a arquitetura proposta para o desenvolvimento da plataforma *online*, na qual estão incluídas as diversas componentes, em que cada uma abrange diferentes tecnologias. A arquitetura demonstrada, encontra-se dividida nas seguintes componentes:

- Recolha de dados da fonte de dados (formulário *online*);

- Carregamento e armazenamento dos dados no *Data Warehouse* implementado;
- Criação do cubo OLAP;
- Exploração e Análise OLAP dos dados;
- Visualização das análises OLAP na plataforma *online*.

Assim, primeiramente, esta arquitetura será formada pelo *website* que possuirá, de entre diversas funcionalidades, um formulário *online* destinado a iniciativas e casos de estudo em Cidades Inteligentes, o qual irá constituir a fonte de dados. De modo a armazenar os dados preenchidos de cada formulário, ter-se-á uma base de dados. A base de dados destinada para o armazenamento dos dados obtidos através do formulário será efetuada de acordo com o modelo multidimensional, constituindo, portanto, o *Data Warehouse* através do servidor *MySQL*. O *Data Warehouse* fornecerá suporte ao sistema OLAP, uma vez que integra os dados provenientes de diversas fontes numa única estrutura, permitindo o acesso a sistemas de análise como OLAP. Este tipo de sistemas fornece mecanismos sofisticados para a exploração e análise aos dados através de diversas perspetivas.

Em seguida, apresenta-se o cubo OLAP relacional que será elaborado de acordo com a implementação do modelo multidimensional. O cubo OLAP será implementado através da ferramenta *Pentaho Schema Workbench* (PSW), a qual permitirá a publicação do mesmo no servidor BI do *Pentaho*. O servidor BI da ferramenta *Pentaho* utilizará o cubo multidimensional para interpretar e executar as consultas (*queries*) MDX criadas através da aplicação *front-end* para a exploração dos dados denominada por PUC (com os *plugins Saiku* e CDE). Após a criação do cubo e através de uma *interface web* como aplicação *front-end*, disponibilizada pelo servidor de BI da ferramenta de análise *Pentaho Community Edition*, será possível a obtenção de análises de dados eficientes e rápidas. Estas análises e gráficos obtidos através da ferramenta *Pentaho* serão disponibilizados no *website* na totalidade para utilizadores devidamente registados e autorizados pelo administrador.

Os dados dos utilizadores registados são armazenados numa base de dados desenvolvida para o efeito.

No final, é esperado que esta arquitetura permita o desenvolvimento de uma plataforma *online* que proporcione uma utilização de forma amigável e intuitiva, oferecendo diversos serviços e funcionalidades que auxiliem os gestores e todas as entidades relacionados com as iniciativas em Cidades Inteligentes. Sendo isto viável

devido ao facto de estes poderem analisar a evolução de diversos aspetos dos principais domínios retratados no formulário *online* e assim, ajudar a melhorar essas iniciativas, ou obter informação que apoie no planeamento e desenvolvimento de novas iniciativas.

4.5 Conclusões

Este capítulo apresentou a definição de requisitos funcionais e não funcionais essenciais para a elaboração da plataforma *online* destinada a iniciativas e casos de estudo em Cidades Inteligentes. De seguida, foram apresentados os principais esboços elaborados após o levantamento de requisitos, que permitem obter uma compreensão mais completa do sistema a desenvolver. Posteriormente, foram apresentados os diagramas de casos de uso referentes aos atores do sistema, neste caso, utilizador não autenticado no sistema, utilizador com autenticação e administrador. Além disso, foi demonstrada uma visão geral da arquitetura com os componentes constituintes do sistema da plataforma *online* essenciais para a elaboração do mesmo. Esta arquitetura permite dar a conhecer tanto os constituintes do sistema, como as tecnologias utilizadas.

Capítulo 5

Implementação do sistema

Este capítulo descreve os principais procedimentos da implementação da plataforma *online* e da aplicação *web* de administração. Inicialmente, neste capítulo, são descritas todas as ferramentas, tecnologias e linguagens utilizadas durante o desenvolvimento do sistema pretendido (secção 5.1). De seguida, este capítulo expõe a elaboração do sistema de *Data Warehouse*, o qual integra e armazena os dados provenientes de diversas fontes numa estrutura de dados denominado por Cubo multidimensional. Posteriormente, é apresentado como se elaborou o esquema do Cubo multidimensional, tendo em conta a modelação multidimensional dos dados. Este cubo é utilizado pelo sistema OLAP, que permitirá a análise eficiente de grandes quantidades de dados sob diversas perspetivas. Além disso, é descrito o modo como foi realizada a exploração dos dados através da aplicação *front-end* da ferramenta *Pentaho*. Esta exploração dos dados engloba a elaboração de gráficos analíticos e de um mapa ilustrativo das localizações das iniciativas referentes às Cidades Inteligentes. Por último, segue a descrição da elaboração da plataforma *online* e de uma aplicação *web* para a administração (secção 5.3 e secção 5.4, respetivamente).

5.1 Ferramentas e tecnologias utilizadas

Nesta secção são apresentadas as várias ferramentas, tecnologias e linguagens utilizadas para o desenvolvimento da plataforma *online* e da aplicação *web* destinada ao(s) administrador(es). Após a conclusão do estudo das ferramentas de *Business Intelligence open source* com a funcionalidade OLAP, optou-se pela utilização da ferramenta *Pentaho*, na versão *suite*, denominada por *Pentaho Community Edition*. Esta ferramenta funciona com vários tipos de Sistemas de Gestão de Base de Dados,

do qual se optou por utilizar o *MySQL*, na versão, também, *open source*, denominada por *MySQL Community Edition*. O sistema de *Data Warehouse* foi elaborado através desta ferramenta.

A plataforma *online*, bem como a aplicação *web* de administração, são composta por um *website* desenvolvido com recurso às tecnologias *HTML*, *PHP*, *JavaScript*, *Ajax* e *CSS*. De modo a armazenar os dados preenchidos de cada formulário, bem como o registo de utilizadores ou de administradores, têm-se, para cada um dos casos, bases de dados para o efeito.

A próxima imagem (figura 5.1) retrata as ferramentas, tecnologias e linguagens utilizadas em cada componente constituinte da arquitetura da plataforma *online*, as quais são descritas com detalhe nas próximas subsecções (subsecções 5.1.1 e 5.1.2).



Figura 5.1: Ferramentas e tecnologias/linguagens utilizadas no desenvolvimento da plataforma *online*

De seguida, é explicado com mais detalhe o propósito da utilização de cada uma dessas ferramentas, tecnologias e linguagens no desenvolvimento deste sistema, mais especificamente, na elaboração das análises sobre os dados e na implementação da plataforma *online* e da aplicação *web* de administração.

5.1.1 Ferramentas para análises estatísticas

Nesta secção é retratado um conjunto de ferramentas utilizadas para elaborar os gráficos analíticos e o mapa relacionados com os dados obtidos do formulário *online*. Essas ferramentas pertencem à gama disponibilizada pela ferramenta de *Business Intelligence open source Pentaho*.

5.1.1.1 *Pentaho Community Edition*

Uma das ferramentas utilizadas durante a implementação da plataforma *on-line*, mais especificamente para a elaboração dos gráficos analíticos e de um mapa, foi a ferramenta *open source Pentaho*, tal como já foi referido anteriormente. Esta versão *open source* oferece diversos *plugins* e ferramentas que foram fundamentais para a exploração e análises eficazes aos dados. De seguida, é descrito o propósito da utilização da cada um desses *plugins* e ferramentas:

- ***Pentaho Schema Workbench (PSW)***: Esta ferramenta permitiu o desenvolvimento do esquema *Mondrian* do cubo OLAP, que oferece suporte à análise de dados em diversas perspetivas. O servidor OLAP denominado por *Mondrian*, permite que o sistema responda a consultas (*queries*) MDX de forma rápida e eficaz o suficiente para possibilitar uma exploração interativa de grandes volumes de dados. O esquema desenvolvido é um esquema ROLAP, composto pela tabela de factos, pelas diversas tabelas de dimensão, pelas hierarquias e pela(s) medida(s). O cubo OLAP é armazenado no formato XML.

A imagem que se segue (figura 5.2) mostra o ambiente de trabalho da ferramenta *Pentaho Schema Workbench*.

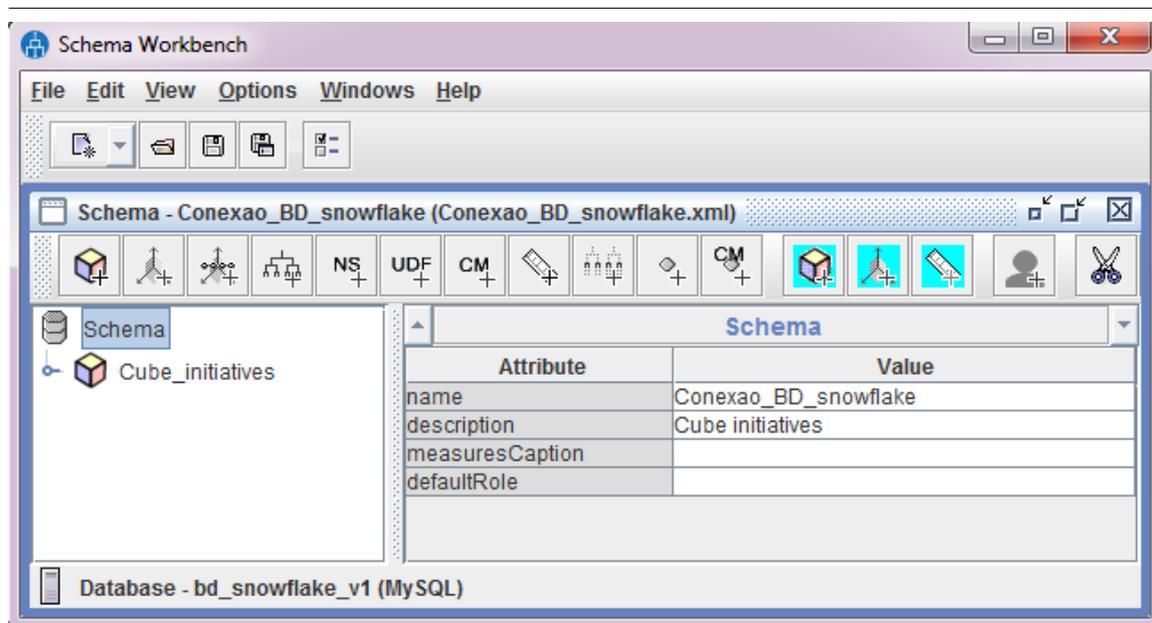


Figura 5.2: Ambiente de trabalho da ferramenta *Pentaho Schema Workbench*

- ***Pentaho User Console (PUC)***: Este componente constitui uma *interface*

gráfica e dinâmica que permite a criação, edição e visualização de gráficos e de mapas, entre outros. Esta aplicação *front-end* é disponibilizada pelo servidor da *Pentaho*, denominado por *BI Server*. O PUC permitiu a elaboração dos diversos *dashboards* e do mapa ilustrativo das cidades inteligentes, disponibilizados na plataforma *online*.

A imagem seguinte (figura 5.3) demonstra o ambiente de trabalho da aplicação *front-end* da *Pentaho*, denominada por PUC.

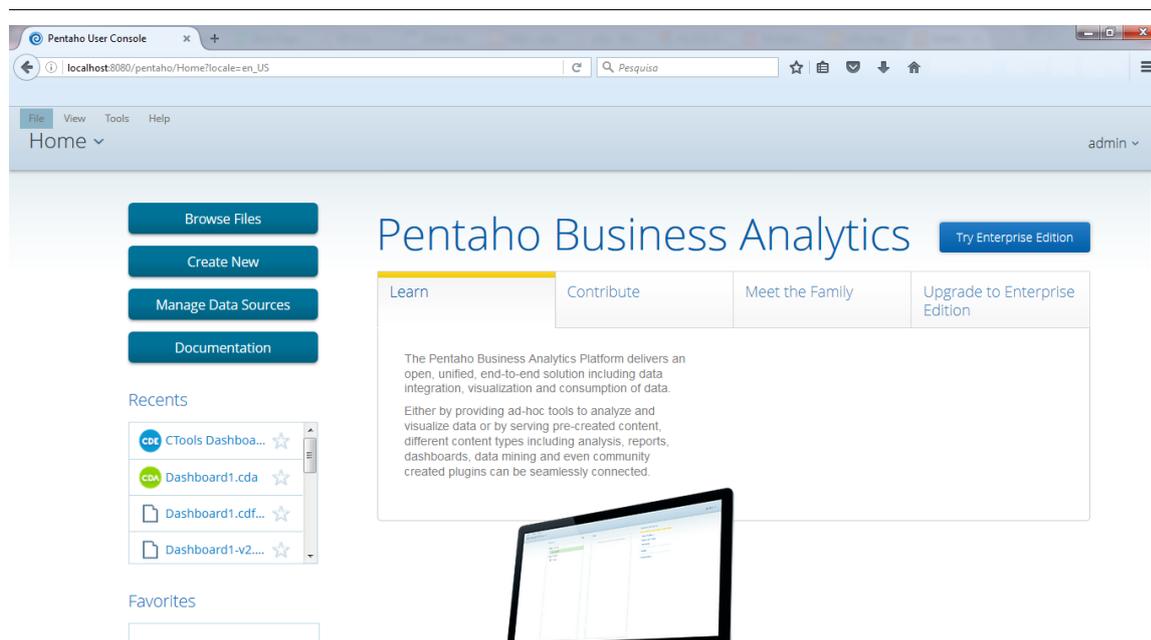


Figura 5.3: Ambiente de trabalho da aplicação *front-end* PUC

- **Saiku Analytics:** É um *plugin*, instalado na ferramenta PUC, que permite visualizar os dados sobre diversas perspectivas do cubo desenvolvido, e obter assim, diversos tipos de gráficos e mapas disponíveis nesse *plugin*.

A imagem que se segue (figura 5.4) apresenta o ambiente de trabalho do *plugin Saiku Analytics* da aplicação PUC.

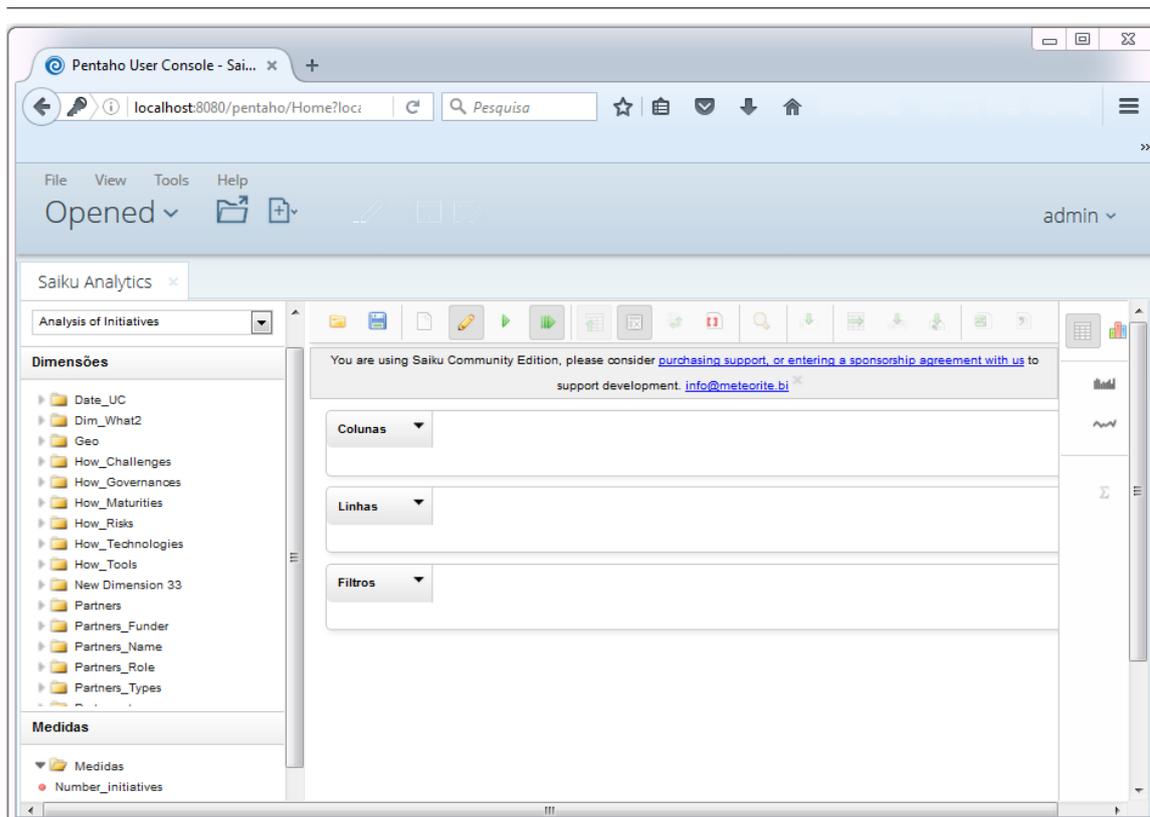


Figura 5.4: Ambiente de trabalho do *plugin Saiku Analytics*

- **Community Dashboard Edition (CDE):** Este *plugin*, instalado na ferramenta PUC, permite a elaboração e edição de *dashboards*, de diversos gráficos e de mapas.

A próxima imagem (figura 5.5) apresenta o ambiente de trabalho do *plugin CDE* da aplicação PUC.

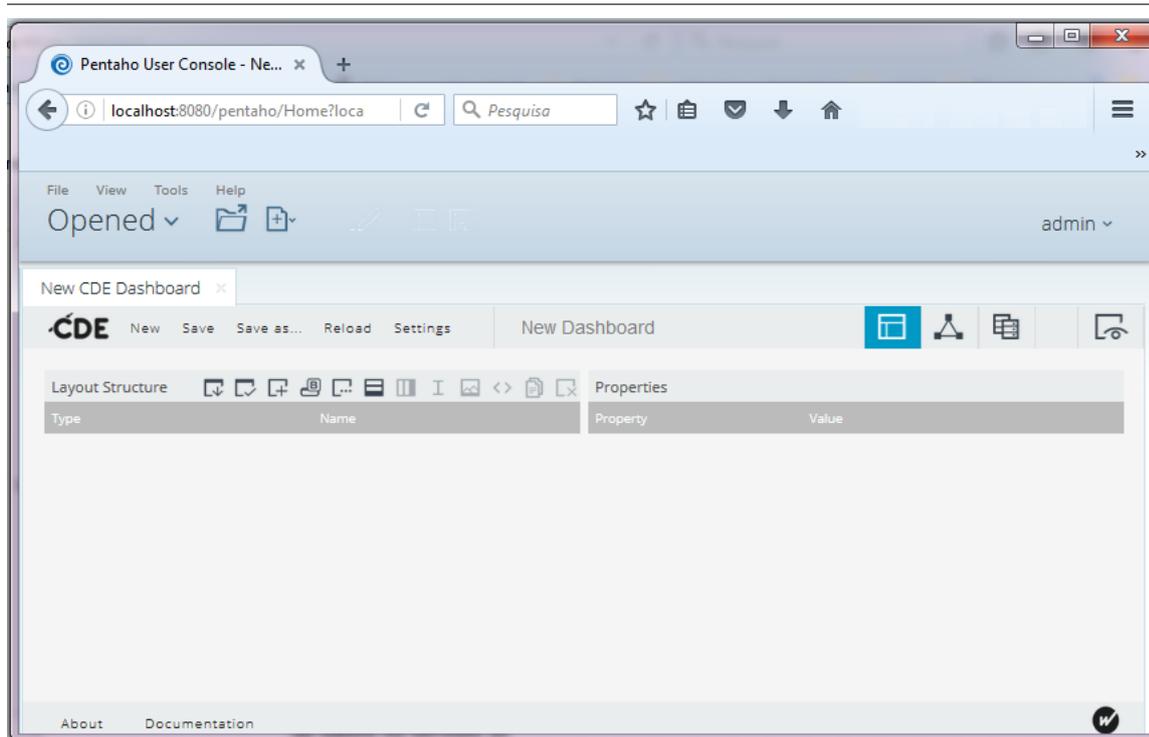


Figura 5.5: Ambiente de trabalho do *plugin* CDE

Especificação e Configuração da ferramenta *Pentaho*:

Para poder explorar a versão *open source* da ferramenta *Pentaho* foi necessário um conjunto de procedimentos e de configurações. De modo a demonstrar esses procedimentos segue-se uma breve descrição das etapas:

- Instalação e configuração do *Pentaho BI Server*: Inicialmente procedeu-se à instalação e configuração do *Pentaho BI Server*. A instalação do *Pentaho* já inclui alguns *drivers* pré-instalados para motores de base de dados, tal como HSQLDB (*Hyper Structured Query Language Database*) e PostgreSQL. No entanto, pelo facto de se ter optado pela utilização do SGBD *MySQL*, para o sistema de *Data Warehouse*, foi necessário proceder à migração para este SGBD.
- Instalação e configuração da ferramenta *Pentaho Schema Workbench*: De seguida efetuou-se a instalação e configuração da ferramenta PSW para a realização do esquema do cubo *Mondrian*. Nesta ferramenta foi necessário, também, criar conexão com o servidor do *MySQL*.

- Publicação do esquema do cubo *Mondrian* no servidor de BI *Pentaho*: Para ter acesso ao esquema do cubo OLAP *Mondrian* na *interface web* PUC, é necessário a publicação do mesmo no servidor do *Pentaho*. Para tal, foi utilizada a ferramenta PSW para publicar o ficheiro XML do esquema criado diretamente no servidor OLAP da ferramenta *Pentaho*.
- Exploração de dados e criação de gráficos analíticos na aplicação *front-end Pentaho User Console*: Após a publicação do esquema do cubo OLAP, recorreu-se à aplicação de *front-end* PUC para o procedimento da exploração e obtenção das análises estatísticas e de mapas.

De salientar, que para proceder à instalação do servidor *Java* do *Pentaho* que fornece a componente *Pentaho User Console* e os diversos *plugins* é necessário ter em conta um conjunto de requisitos de *Hardware* e de *Software*. A tabela 5.1, que se segue, descreve esses requisitos, sobretudo relacionados com a versão utilizada do Servidor BI do *Pentaho* (versão 5.4).

Tabela 5.1: Descrição de requisitos de *Hardware* e de *Software* do Servidor da ferramenta *Pentaho*, adaptado de [26]

<i>Hardware - 64 bits</i>	<i>Software - 64 bits</i>
<ul style="list-style-type: none"> • Processador: - <i>Apple Macintosh Pro Quad-Core</i> ou <i>Macintosh Mini Quad-Core</i> - <i>Intel EM64T</i> ou <i>AMD64 Dual-Core</i> • RAM: 8 GB com 4 GB dedicado aos servidores <i>Pentaho</i> • Espaço de disco: 20 GB livres após a instalação 	<ul style="list-style-type: none"> • <i>Apple Macintosh OS X Server 10.9 e 10.10</i> • <i>CentOS Linux 5 e 6</i> • <i>Microsoft Windows 2008 Server R2, 2012 Server R2</i> • <i>Red Hat Enterprise Linux 5 e 6</i> • <i>Solaris 10</i> • <i>Ubuntu Server 12.04 LTS e 14.04 LTS</i>

Quanto à instalação da ferramenta *Pentaho Schema Workbench* é necessário ter em conta os seguintes requisitos de *Hardware* e de *Software*:

Tabela 5.2: Descrição de requisitos de *Hardware* e de *Software* da ferramenta *Pentaho Schema Workbench*, adaptado de [26]

<i>Hardware - 32 ou 64 bits</i>	<i>Software - 32 ou 64 bits</i>
<ul style="list-style-type: none"> • Processador: <ul style="list-style-type: none"> - <i>Apple Macintosh Dual-Core</i>, 64 bits apenas - <i>Intel EM64T</i> ou <i>AMD64 Dual-Core</i> • RAM: 2 GB RAM • Espaço de disco: 2 GB livres após a instalação • Tamanho mínimo do ecrã: 1280 x 960 	<ul style="list-style-type: none"> • <i>Apple Macintosh OS 10.9 (Maverick)</i> e 10.10 (<i>Yosemite</i>), 64 bits apenas • <i>Microsoft Windows 7</i> e 8 • <i>Ubuntu Desktop 12.04 LTS</i> e <i>14.04 LTS</i> • <i>iPad 2, 3</i> e <i>Air (iOS 7.x e 8.x)</i>

Os procedimentos para a instalação e configuração do servidor da *Pentaho*, especialmente para o funcionamento da aplicação *front-end*, isto é, da *interface web* disponibilizado pelo seu servidor é algo complexo. A instalação englobou um conjunto de etapas que foram seguidas minuciosamente para um correto funcionamento da ferramenta, tal como a alteração de variáveis de ambiente e a criação da estrutura de diretorias específicas. Além disso, a migração para o SGBD *MySQL* exigiu especial atenção pois é necessário proceder à modificação de *scripts* de base de dados e de determinados ficheiros XML contidos nas diversas diretorias existentes na ferramenta. Posteriormente a esses procedimentos, é que se efetuou a instalação do *Pentaho Schema Workbench* e dos *plugins Saiku Analytics* e CDE na aplicação *front-end*.

A adaptação à ferramenta *Pentaho Community* foi um pouco complexa, uma vez que esta ferramenta *open source* é constituída por diversos *plugins* e componentes desenvolvidos pela comunidade do *Pentaho*, que carecem de informação, sobretudo informação atualizada, sobre o seu funcionamento. Devido a essa lacuna, foi complicado de proceder à escolha dos *plugins* e componentes que mais se adaptavam à implementação do sistema desenvolvido, bem como entender o funcionamento de cada um. Deste modo, requereu uma dedicação especial e demorada.

5.1.2 Ferramentas e tecnologias/linguagens para o desenvolvimento *web*

Esta secção descreve as ferramentas e linguagens utilizadas para o desenvolvimento da plataforma *online* e para o *website* exclusivo ao(s) administrador(es) desta plataforma.

- **Editor *Sublime Text 3***

Esta ferramenta é um editor de texto e de código-fonte multiplataforma, e foi utilizada para desenvolver todo o código da implementação da plataforma *online* e da aplicação *web* de administração. A versão utilizada do editor foi a versão 3 gratuita, que se caracteriza pela sua simplicidade, robustez, desempenho e exibição de uma *interface* amigável.

A imagem que segue (figura 5.6) apresenta o ambiente de trabalho da ferramenta *Sublime Text 3*.

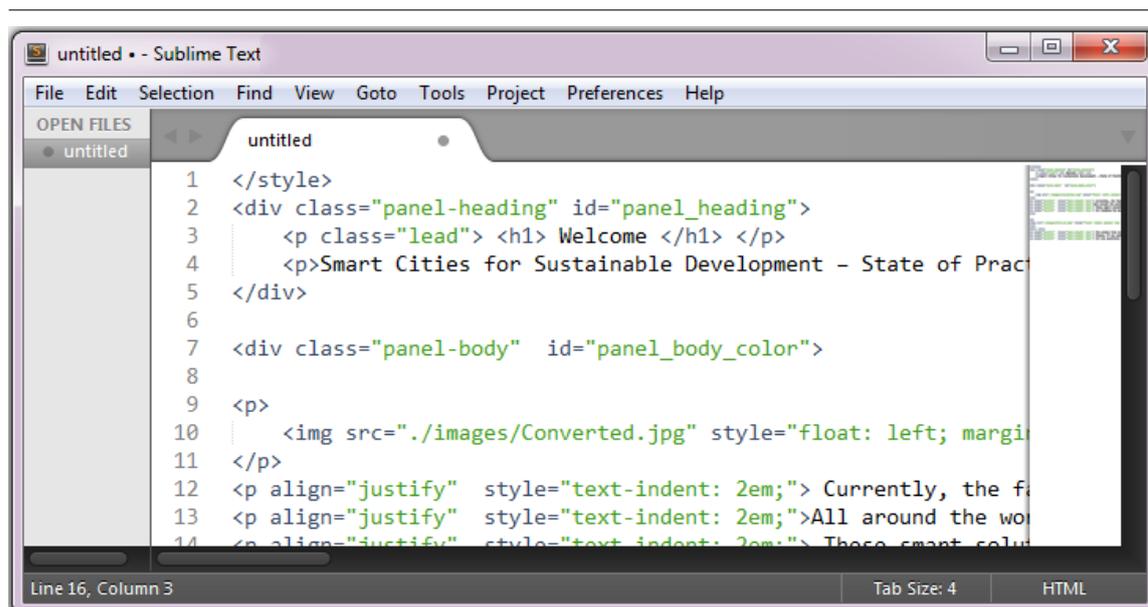


Figura 5.6: Ambiente de trabalho do Editor *Sublime Text 3*

- ***XAMPP Control Panel***

XAMPP é um pacote de distribuição *Apache* (servidor *web* livre), que contém um conjunto de servidores, nomeadamente o servidor *HTTP Apache*, o servidor

MySQL e o interpretador de PHP, fundamentais para o desenvolvimento da plataforma *online* e do *website* de administração.

A próxima imagem (figura 5.7) apresenta a ferramenta *XAMPP Control Panel*, no qual se pode observar o servidor *Apache* ligado.

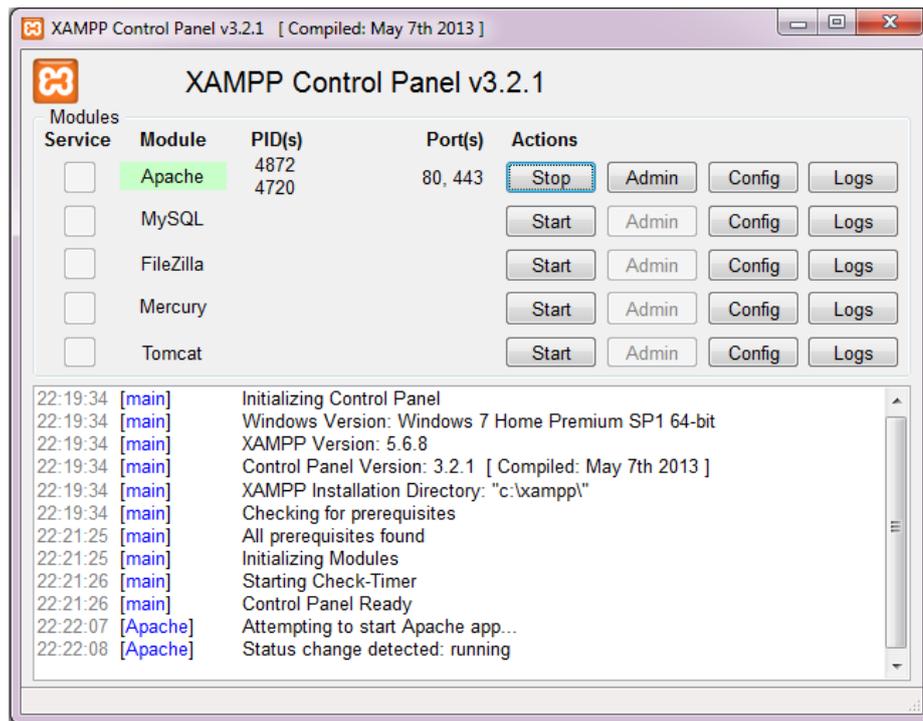


Figura 5.7: Ambiente da ferramenta *XAMPP Control Panel*

- ***MySQL Workbench 6.3 CE (Community Edition):***

O *MySQL* é um sistema de gestão de base de dados (SGBD), que utiliza linguagem *SQL*. A ferramenta *MySQL Workbench* é uma ferramenta cliente destinada à criação de diagramas de entidade e relacionamento e à conceção de bases de dados. Esta ferramenta foi utilizada para o desenho do modelo multidimensional e para a implementação da base de dados, tanto para armazenar a informação relativa aos dados das iniciativas em Cidades Inteligentes como para o registo de utilizadores e administradores.

A imagem que se segue (figura 5.8) apresenta o ambiente de trabalho inicial da ferramenta *MySQL Workbench*, utilizada na versão 6.3 da edição *Community Edition*.

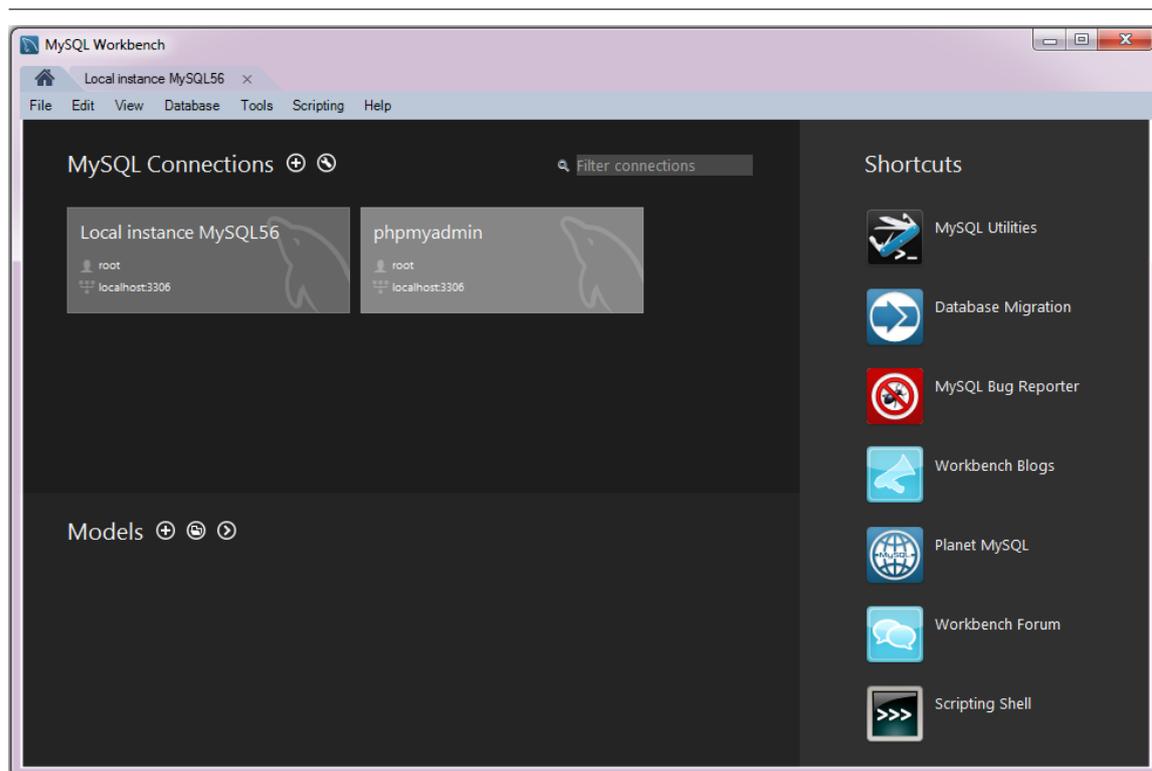


Figura 5.8: Ambiente de trabalho da ferramenta *MySQL Workbench* 6.3 CE

Quanto às tecnologias e linguagens de programação utilizadas para a implementação dos *websites* relativos à plataforma *online* e ao sistema de administração, são as seguintes:

- **Tecnologias e Linguagens utilizadas do lado do cliente:**

- ***Framework Bootstrap:***

A *framework Bootstrap* é uma *framework front-end*, de código aberto (*open source*) desenvolvida em HTML, CSS, e JS (*JavaScript*). Esta *framework* foi concebida para facilitar o desenvolvimento *web*, tornando-o mais rápido e acessível. Além disso, a utilização de *Bootstrap* permite a adaptação dos *websites* a múltiplos tamanhos de ecrã, tornando-os, assim, compatíveis com a visualização em diversos dispositivos.

– **HTML:**

HTML é uma linguagem de marcação de hipertexto utilizada para o desenvolvimento de páginas *web* do lado cliente. Esta linguagem é interpretada por interpretadores contidos nos *web browsers*.

– **JavaScript:**

JavaScript é uma linguagem de programação interpretada, controlada pelo *web browser*. Esta linguagem permite inserir dinamismo nas páginas *web*, efetuando ações que não necessitam de ligação ao servidor, no próprio *browser*. Dessa forma, isto torna a resposta da página *web* mais rápida devido ao menor número de dados trocados com o servidor. Esta linguagem foi diversas vezes utilizada durante a implementação do sistema, tais como, para gerar conteúdos dinâmicos e para efetuar a validação dos valores inseridos durante o preenchimento dos formulários, assegurando um correto preenchimento desses valores antes de serem enviados ao servidor.

– **CSS:**

É uma linguagem de folhas de estilo que permite controlar e definir a aparência (o estilo) de documentos escritos em linguagem de marcação, como o HTML. Este tipo de linguagem tem como principal característica o facto de permitir a separação entre a especificação da formatação e a especificação do documento, sendo necessário apenas a criação de uma ligação (*link*) no documento para essa formatação. Assim, isto permite obter uma maior liberdade na formatação dos conteúdos, maior produtividade e maior facilidade na alteração/atualização dessa formatação.

– **AJAX:**

AJAX é um conjunto de técnicas utilizadas no desenvolvimento *web* no lado do cliente. O AJAX utiliza padrões já existentes das tecnologias *JavaScript* e XML fornecidas pelos *web browsers*, para efetuar solicitações de informações de forma assíncrona ao servidor. Desta forma, as páginas *web* tornam-se mais rápidas e dinâmicas. Este conjunto de tecnologias permite uma troca de dados com o servidor, sem necessitar de recarregar a página *web* na sua totalidade, podendo assim, recarregar apenas fragmentos de

uma página *web*. Portanto, este tipo de técnica foi utilizado durante o desenvolvimento dos *websites* com o objetivo de os tornarem mais rápidos e interativos. A comunicação assíncrona com o servidor, disponibilizada pelo AJAX, permite, também, a não interrupção das funcionalidades da página *web* em utilização. Devido a essa vantagem, AJAX foi utilizada no envio dos dados de cada formulário existente no sistema para o servidor.

- **Tecnologias e Linguagens utilizadas do lado do servidor:**

- **PHP:**

PHP é uma linguagem interpretada (de *script*) *open source*, que significa Pré-Processador de Hipertexto. O código PHP é interpretado e executado do lado do servidor através do módulo PHP que o servidor possui. O PHP também é responsável por gerar HTML (a página *web*) que é enviado para o *web browser*, ficando, portanto, visível no lado do cliente. Esta linguagem pode ser executada a partir de diversos servidores *web*, tais como o *Apache*. Este tipo de linguagem possibilita, também, o acesso à base de dados, tendo sido utilizada, por exemplo, para um conjunto das seguintes operações na base de dados: inserção, alteração, remoção e visualização de dados.

De referir, ainda, que o formato de dados utilizado para a resposta do servidor ao cliente, acerca do envio dos dados dos diversos formulários existentes no sistema para a base de dados, foi o JSON. JSON é um formato leve utilizado para armazenamento de dados e troca de dados em rede.

5.2 Sistema de *Data Warehousing*

O processo de *Data Warehousing* é o planeamento e a implementação de um sistema que possibilita reunir e integrar dados oriundos de diversas fontes, que são armazenados em repositórios de armazenamento de dados, denominados por *Data Warehouse*. O *Data Warehouse* permite armazenar grandes quantidades de dados, sendo a exploração desses dados vulgarmente realizada através de ferramentas de análise OLAP.

Nesta secção é exposta a modelação multidimensional dos dados, que foi utilizada para a elaboração do sistema *Data Warehouse*.

5.2.1 Modelo de dados do *Data Warehouse*

A modelação de dados do *Data Warehouse* é das etapas mais relevantes na conceção e elaboração de um sistema de *Data Warehouse*. Usualmente, num *Data Warehouse* os dados são modelados com base na modelação multidimensional, como foi referido na secção 2.2.3. É de elevada importância efetuar uma boa modelação multidimensional de dados, de modo a possibilitar um bom desempenho, eficácia e escalabilidade de um sistema de *Data Warehouse*, e ainda, a permitir uma análise OLAP fiável e coerente.

Visto que este trabalho assenta no conceito OLAP, foi estudado o modelo multidimensional associado à tecnologia OLAP que melhor se adapte ao sistema de *Data Warehouse* desenvolvido. A modelação de dados deste sistema foi definida de acordo com as necessidades de informação impostas nos objetivos da elaboração da plataforma *online*. Assim, o modelo de dados apresentado está, ainda, em conformidade com a principal medida a obter (o número de iniciativas) e com os requisitos e objetivos definidos para as análises pretendidas. Este modelo de dados serviu como suporte à base de dados elaborada através da linguagem de programação SQL, que armazenará os dados referentes aos projetos de Cidades Inteligentes.

A implementação do *Data Warehouse* foi realizada com base em OLAP Relacional (ROLAP), de modo a permitir um processamento analítico consistente e rápido dos dados. Por conseguinte, o modelo de dados multidimensional é armazenado numa base de dados relacional, no qual as operações são expressas, normalmente, em SQL.

O modelo de dados foi adaptado do conceito de modelação em Floco de Neve (*Snowflake*), encontrando-se otimizado para facilitar e realizar carregamentos e consultas dos dados de forma eficiente e rápida. Optou-se por ter como base este modelo multidimensional dado que, parcialmente possui normalização, e assim sendo, a maioria das tabelas de dimensão foram normalizadas, o que reduz o custo de manutenção das tabelas de dimensão. Além disso, a opção deste modelo tem a vantagem de indicar explicitamente a estrutura das suas dimensões e de evitar que se armazene informações redundante, sendo por isso, esquemas mais fáceis de manter, gerir e alterar.

A figura 5.9, que se segue, mostra o modelo multidimensional do sistema de *Data Warehouse*, o qual foi projetado com recurso à ferramenta *MySQL Workbench*. Este modelo relaciona a tabela de factos (*Fact_Use_Case*) com as diversas tabelas de dimensão baseadas em cinco questões definidas no modelo conceptual apresen-

tado na secção 2.1.3 para a elaboração do formulário: “Who?”, “Where?”, “What?”, “Why?” e “How?”. Alguns exemplos de tabelas de dimensão são os seguintes: *Dim_Responsible_Institutions*, *Dim_Partners*, *Dim_Sources*, *Dim_Drivers*, *Dim_Benefits* e *Dim_Risks*.

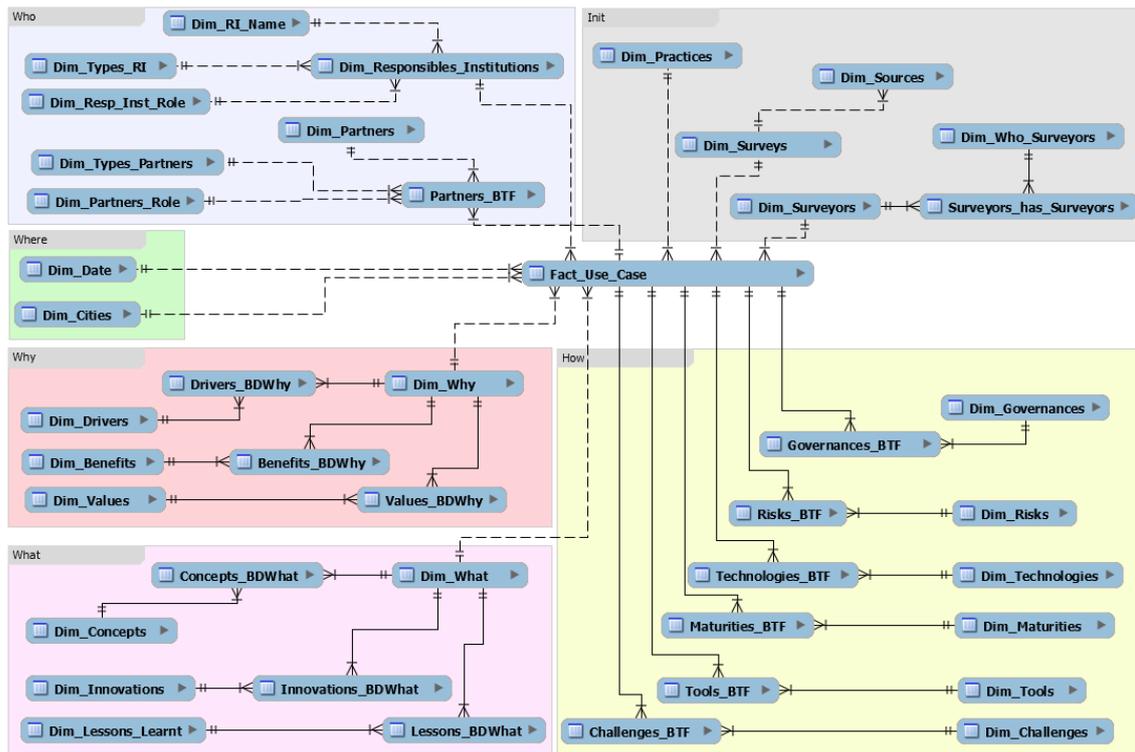


Figura 5.9: Modelo multidimensional do sistema de *Data Warehouse*

As tabelas de dimensão encontram-se, como se pode observar na imagem 5.9, agrupadas no esquema de modelação de dados de acordo com o grupo de questão a que pertencem. Cada grupo de questão tem atribuído uma cor diferente com o propósito de facilitar a visualização dessas tabelas.

Sendo assim, segue-se uma breve descrição desses grupos:

- **Cor Cinzenta (“Init”)** - agrupa as tabelas de dimensão referentes aos dados iniciais do formulário que são: o nome da prática, as pesquisas e os investigadores. Neste grupo evidenciam-se as seguintes tabelas como exemplos: *Dim_Practises*, *Dim_Surveys*, *Dim_Sources*, *Dim_Surveyors* e *Dim_Who_Surveyors*;

- **Cor Azul (“Who?”)** - agrupa as tabelas de dimensão relativas à questão “Who?” (“Quem?”), a qual referencia a instituição e o(s) parceiro(s) relacionados com o desenvolvimento da iniciativa. Neste grupo destacam-se as tabelas *Dim_Partners* e *Dim_RI_Name* como exemplos;
- **Cor Verde (“Where?”)** - reúne as tabelas de dimensão relacionadas com a questão “Where?” (“Onde?”), a qual determina o país, estado/província, cidade, município/região, continente onde a iniciativa tem sido desenvolvida, bem como o período de tempo do início do seu desenvolvimento. Este grupo é constituído pelas tabelas de dimensão *Dim_Cities* e *Dim_Date*;
- **Cor Vermelha (“Why?”)** - agrega as tabelas de dimensão associadas à questão “Why?” (“Porquê?”), sendo esta que determina o impacto da prática em termos da sua contribuição para as dimensões sociais, económicas ou ambientais. Neste grupo destacam-se as tabelas de dimensão *Dim_Drivers*, *Dim_Benefits* e *Dim_Values* como exemplos;
- **Cor Lilás (“What?”)** - agrupa as tabelas de dimensão pertencentes à questão “What?” (“O quê/Qual?”), em que descreve o objetivo da prática, e resume as principais características e aprendizagens. As tabelas de dimensão *Dim_Concepts*, *Dim_Innovations* e *Dim_Lessons_Learnt* são exemplos de tabelas pertencentes a este grupo;
- **Cor Amarela (“How?”)** - reúne as tabelas de dimensão relativas à questão “How?” (“Como?”), a qual retrata o mecanismo utilizado para transmitir a prática, tais como instrumentos de política, iniciativas governamentais, ferramentas, mecanismos de governação, entre outros. Neste grupo mencionam-se como exemplo de tabelas constituintes as seguintes: *Dim_Governances* e *Dim_Technologies*.

A imagem seguinte (figura 5.10) apresenta a tabela de factos, podendo observar-se as chaves pelo qual esta tabela é formada.



Figura 5.10: Tabela de factos *Fact_Use_Case*

A tabela de factos *Fact_Use_Case*, demonstrada na figura 5.10 e descrita na tabela 5.3, permite armazenar a informação sobre os casos de estudo e iniciativas inseridos através do formulário *online*. Esta tabela de factos está relacionada com as várias tabelas de dimensão, possibilitando explorar análises do número de iniciativas relativamente a diversas perspetivas.

Tabela 5.3: Descrição da tabela de factos *Fact_Use_Case*

Atributo	Tipo	Chave	Descrição
id_Fact_Use_Case	Int - Not Null	PK	Identificador único de uma linha referente ao caso de estudo/iniciativa
Dim_Surveys_idSurvey	Int - Not Null	FK	ID da tabela de dimensão Dim_Surveys
Dim_Surveyors_idSurveyor	Int - Not Null	FK	ID da tabela de dimensão Dim_Surveyors
Dim_Date_idDate	Int - Not Null	FK	ID da tabela de dimensão Dim_Date
Dim_Responsibles_Institutions_idResp_Institution	Int - Not Null	FK	ID da tabela de dimensão Dim_Responsibles_Institutions
Dim_Practices_idDim_Practices	Int - Not Null	FK	ID da tabela de dimensão Dim_Practices

Dim_Why_idDim_Why	Int - Not Null	FK	ID da tabela de dimensão Dim_Why
Dim_What_idDim_What	Int - Not Null	FK	ID da tabela de dimensão Dim_What
Dim_Cities_idCity	Int - Not Null	FK	ID da tabela de dimensão Dim_Cities que contém a localização
Number_initiatives	Int - (<i>distinct- count</i>)	-	Medida para o cálculo do número de iniciativas (id_Fact_Use_Case)

Tal como se pode observar na tabela anterior (tabela 5.3), a tabela de factos *Fact_Use_Case* possui uma Chave Primária (PK - *Primary Key*) que identifica um caso de estudo/iniciativa em particular, servindo, também, para determinar a medida pretendida neste sistema, que é o número de iniciativas (*Number_initiatives*). Além disso esta tabela de factos contém as diversas chaves estrangeiras (FK - *Foreign Key*) relativas às tabelas de dimensão interligadas.

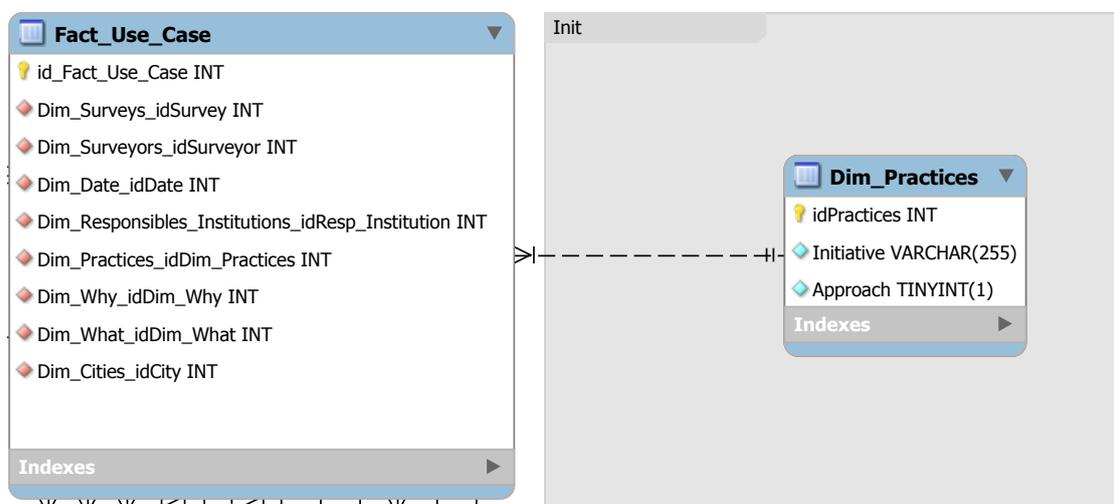
Por vezes, neste tipo de modelação pode existirem, também, relações de muitos-para-muitos entre as tabelas de factos e as tabelas de dimensão, ou mesmo entre os valores das várias dimensões. Por isso, torna-se necessário modelar, como o caso deste trabalho, essas relações de muitos-para-muitos. De modo a solucionar este problema, foram criadas tabelas denominadas por tabelas de ponte (*bridge tables*). As tabelas de pontes foram elaboradas entre a tabela de factos e algumas tabelas de dimensão, e também, entre algumas tabelas de dimensão.

De seguida, sucede-se a descrição das diversas tabelas de dimensão interligadas à tabela de factos (*Fact_Use_Case*) e das tabelas de ponte.

- **Descrição das tabelas referentes a ‘*Init*’**

Este conjunto de tabelas é composto pelas tabelas de dimensão *Dim_Practices*, *Dim_Surveys*, *Dim_Sources*, *Dim_Surveyors* e *Dim_Who_Surveyors*, e pela tabela de ponte *Dim_has_Surveyors*.

A próxima imagem (figura 5.11) mostra a tabela de dimensão *Dim_Practices* do sistema de *Data Warehouse*, no qual se pode observar a chave pelo qual é composta, os atributos e a sua relação com a tabela de factos.

Figura 5.11: Tabela de dimensão *Practices*

A tabela de dimensão *Dim_Practices*, tal como se observa na figura 5.11, contém as informações relativas ao nome do caso de estudo/iniciativa de cada Cidade Inteligente e ao tipo de abordagem que segue (*Approach: Top-Down* ou *Bottom-up*).

A tabela 5.4 descreve os atributos, indicando o tipo de dados de cada um, bem como, as chaves constituintes da tabela de dimensão *Dim_Practices*.

Tabela 5.4: Descrição da tabela de dimensão *Dim_Practices*

Atributo	Tipo	Chave	Descrição
idPractices	<i>Int - Not Null</i>	PK	Identificador único da tabela de dimensão
Initiative	<i>Varchar - Not Null</i>	-	Nome do caso de estudo/iniciativa
Approach	<i>Tinyint - Not Null</i>	-	Tipo de abordagem, referente à questão “How?”, tem dois valores: 0 referente a <i>Top-Down</i> e 1 a <i>Bottom-up</i>

A próxima imagem (figura 5.12) retrata as tabelas de dimensão *Dim_Surveys* e *Dim_Sources* e a relação com a tabela de factos *Fact_Use_Case*.

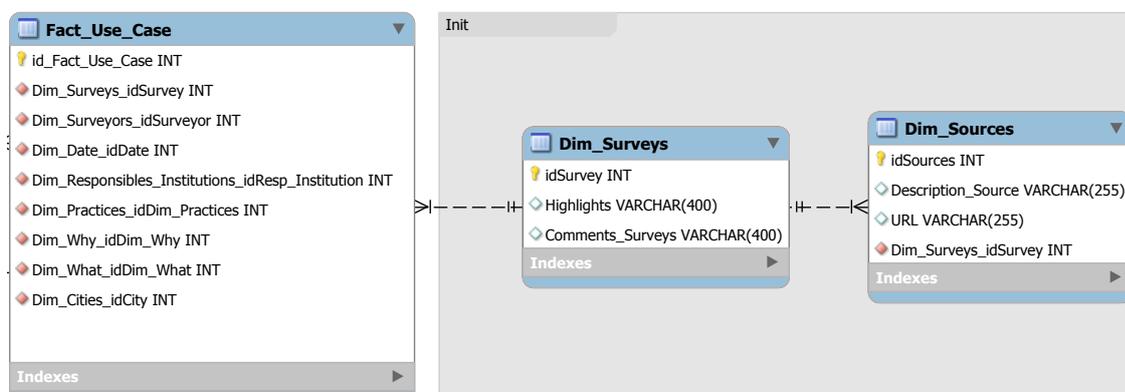


Figura 5.12: Tabelas de dimensão *Practices Surveys* e *Sources*

Estas tabelas de dimensão armazenam os dados relativos à pesquisa de cada caso de estudo/iniciativa, nomeadamente os principais destaques, comentários acerca da pesquisa sobre o caso de estudo/iniciativa (*Dim_Surveys*) e as descrições relevantes e endereços de URL (*Dim_Sources*).

A tabela seguinte (tabela 5.5) descreve os atributos, indicando o tipo de dados de cada um e as chaves constituintes, referentes à tabela de dimensão *Dim_Surveys*.

Tabela 5.5: Descrição da tabela de dimensão *Dim_Surveys*

Atributo	Tipo	Chave	Descrição
idSurvey	<i>Int</i>	PK	Identificador único da tabela de dimensão
Highlights	<i>Varchar</i>	-	Descrição de principais destaques relativos à pesquisa sobre o caso de estudo/iniciativa
Comments_Surveys	<i>Varchar</i>	-	Principais comentários acerca da pesquisa

Uma vez que, uma pesquisa pode estar associada a mais que um endereço de URL, existe a tabela *Dim_Sources* para armazenar essa informação, bem como uma descrição breve sobre esses endereços de URL.

Na próxima tabela (tabela 5.6) pode-se observar a descrição de cada atributo referente à tabela de dimensão *Dim_Sources*, bem como o tipo de dados e as chaves constituintes.

Tabela 5.6: Descrição da tabela de dimensão *Dim_Sources*

Atributo	Tipo	Chave	Descrição
idSources	<i>Int</i>	PK	Identificador único da tabela de dimensão
Description_Source	<i>Varchar</i>	-	Descrição breve acerca dos endereços de URL
URL	<i>Varchar</i>	-	Endereço de URL
Dim_Surveys_idSurvey	<i>Int</i>	FK	Chave Estrangeira que interliga à tabela de dimensão <i>Dim_Surveys</i>

A figura 5.13, que se segue, representa as tabelas de dimensão *Dim_Surveyors* e *Dim_Who_Surveyors* e a tabela de ponte *Surveyors_has_Surveyors*, bem com a relação com a tabela de factos *Fact_Use_Case*.

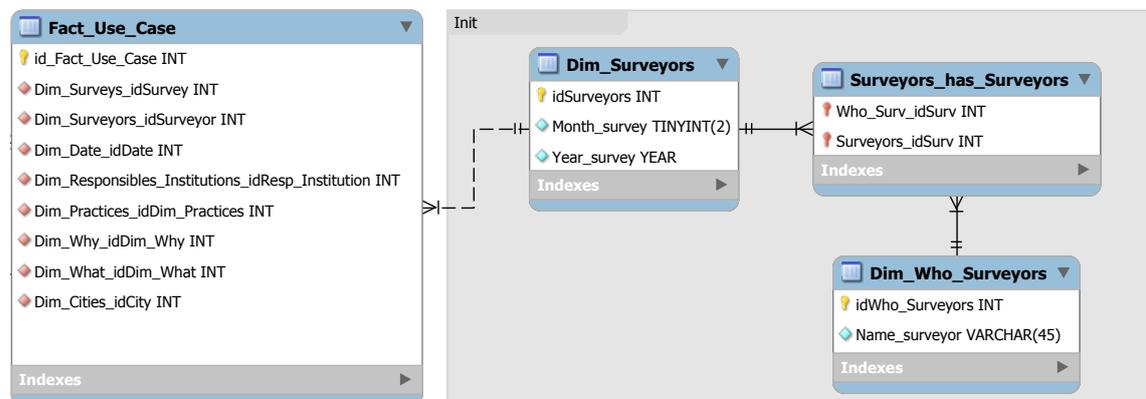


Figura 5.13: Tabelas de dimensão *Dim_Surveyors* e *Dim_Who_Surveyors* e tabela de ponte *Surveyors_has_Surveyors*

Estas tabelas armazenam a informação relativa aos investigadores *Dim_Who_Surveyors*, nomeadamente o nome de quem efetuou a pesquisa e quando estes realizaram-na (*Dim_Surveyors*). A descrição de cada uma delas segue-se nas próximas tabelas 5.7 e 5.8, respetivamente. Como uma pesquisa pode ser efetuada por mais que um investigador, existe uma tabela de ponte para criar uma relação de muitos-para-muitos, denominada por *Surveyors_has_Surveyors* e descrita na tabela 5.9.

A tabela 5.7, que se segue, descreve os atributos, indicando o tipo de dados de cada um, bem como as chaves constituintes da tabela de dimensão *Dim_Surveyors*.

Tabela 5.7: Descrição da tabela de dimensão *Dim_Surveyors*

Atributo	Tipo	Chave	Descrição
idSurveyors	<i>Int - Not Null</i>	PK	Identificador único da tabela de dimensão
Month_survey	<i>Tinyint - Not Null</i>	-	Mês referente à pesquisa
Year_survey	<i>Int - Not Null</i>	-	Ano referente à pesquisa

A próxima tabela 5.8 mostra a descrição de cada um dos atributos da tabela de dimensão *Dim_Who_Surveyors*, incluindo a indicação do tipo de dados e das chaves constituintes.

Tabela 5.8: Descrição da tabela de dimensão *Dim_Who_Surveyors*

Atributo	Tipo	Chave	Descrição
idWho_Surveyors	<i>Int - Not Null</i>	PK	Identificador único da tabela de dimensão
Name_surveyor	<i>Varchar - Not Null</i>	-	Nome do pesquisador

A tabela 5.9 que se segue descreve cada atributo da tabela de ponte *Surveyors_has_Surveyors*, no qual indica, também, o tipo de dados e as chaves constituintes.

Tabela 5.9: Descrição da tabela de dimensão *Surveyors_has_Surveyors*

Atributo	Tipo	Chave	Descrição
Who_Surv_idSurv	<i>Int - Not Null</i>	FK	Chave Estrangeira para armazenar o identificador da tabela de dimensão <i>Dim_Surveyors</i>

Surveyors_idSurv	Int - Not Null	FK	Chave Estrangeira para armazenar o identificador da tabela de dimensão Dim_Who_Surveyors
------------------	----------------	----	---

- Descrição das tabelas referentes à questão “Who?”

Este grupo de tabelas, referente à questão “Who?”, é composto pelas tabelas de dimensão *Dim_Responsibles_Institutions* e *Dim_Partners_BTF*, e por um conjunto de tabelas com relação a cada uma dessas.

A figura 5.14, que se segue, mostra as tabelas de dimensão *Dim_Responsibles_Institutions*, *Dim_RI_Name*, *Dim_Types_RI* e *Dim_Resp_Inst_Role*, bem com a relação com a tabela de factos *Fact_Use_Case*.

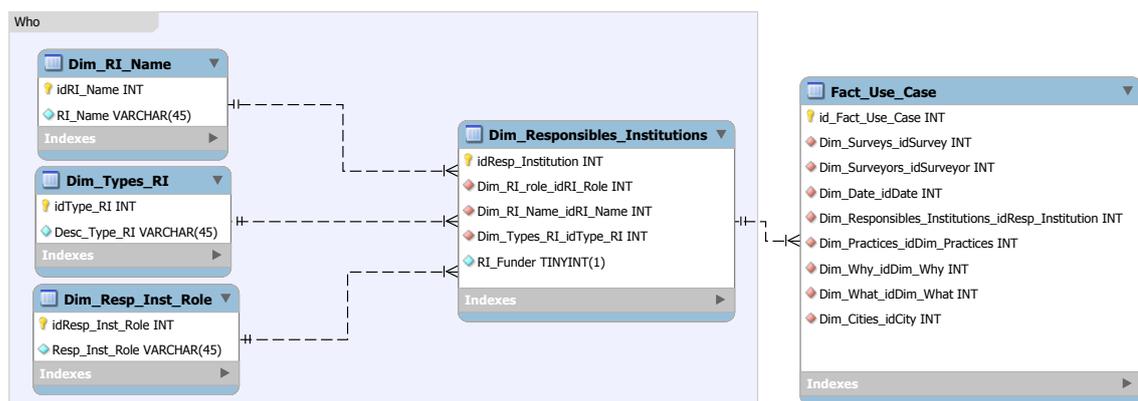


Figura 5.14: Tabelas de dimensão *Dim_Responsibles_Institutions*, *Dim_RI_Name*, *Dim_Types_RI* e *Dim_Resp_Inst_Role*

Estas tabelas de dimensão têm como finalidade armazenarem informações sobre as instituições responsáveis por cada caso de estudo/iniciativa em Cidades Inteligentes. Cada caso de estudo/iniciativa está associado a uma instituição.

A tabela de dimensão *Dim_Responsibles_Institutions* armazena a informação que indica se a instituição é financiadora ou não, sendo por isso este atributo do tipo *tinyint*, para armazenar. Desta forma, é armazenado 0 em caso de a instituição não ser financiadora e 1 em caso contrário. Além disso, esta tabela serve também de tabela de ponte entre as tabelas de dimensão que lhe estão interligadas, sendo essas as tabelas *Dim_RI_Name*, *Dim_Types_RI* e *Dim_Resp_Inst_Role*.

A tabela seguinte (tabela 5.10) apresenta a descrição da tabela de dimensão *Dim_Responsibles_Institutions*, nomeadamente a identificação de cada atributo, o tipo de dados e as chaves constituintes.

Tabela 5.10: Descrição da tabela de dimensão *Dim_Responsibles_Institutions*

Atributo	Tipo	Chave	Descrição
idResp_Institution	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Responsibles_Institutions</i>
Dim_RI_Name_idRI_Name	<i>Int - Not Null</i>	FK	Chave Estrangeira para armazenar o identificador da tabela de dimensão <i>Dim_RI_Name</i>
Dim_Types_RI_idType_RI	<i>Int - Not Null</i>	FK	Chave Estrangeira para armazenar o identificador da tabela de dimensão <i>Dim_Types_RI</i>
Dim_RI_role_idRI_Role	<i>Int - Not Null</i>	FK	Chave Estrangeira para armazenar o identificador da tabela de dimensão <i>Dim_Resp_Inst_Role</i>
RI_Funder	<i>Tinyint - Not Null</i>	-	Armazena se a instituição é ou não financiadora da iniciativa, contendo dois valores: 0 para caso negativo ou 1 para caso afirmativo

A próxima tabela (tabela 5.11) descreve os atributos, o tipo de dados de cada um e as chaves constituintes da tabela de dimensão *Dim_RI_Name*, a qual armazena os nomes referentes a cada instituição.

Tabela 5.11: Descrição da tabela de dimensão *Dim_RI_Name*

Atributo	Tipo	Chave	Descrição
idRI_Name	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_RI_Name</i>

RI_Name	Varchar - Not Null	-	Descrição do nome da instituição responsável pelo iniciativa
---------	--------------------	---	--

A tabela que se segue (tabela 5.12) descreve a tabela de dimensão *Dim_Types_RI*, nomeadamente os atributos, o tipo de dados de cada um e as chaves constituintes. A tabela de dimensão *Dim_Types_RI* armazena a informação sobre o tipo de instituição responsável por cada iniciativa. Nesta tabela, por defeito, estão armazenados três tipos de instituição: *Government* (Governo), *Industry* (Indústria), e *Non-Governmental Organization* (Organização Não Governamental). Um destes tipos terá de ser selecionado durante o preenchimento do formulário *online*.

Tabela 5.12: Descrição da tabela de dimensão *Dim_Types_RI*

Atributo	Tipo	Chave	Descrição
idType_RI	Int - Not Null	PK	Identificador da tabela de dimensão <i>Dim_Types_RI</i>
Desc_Type_RI	Varchar - Not Null	-	Descrição do tipo de instituição: <i>Government</i> , <i>Industry</i> e <i>Non-Governmental Organization</i>

A seguinte tabela (tabela 5.13) apresenta a descrição de cada atributo, o tipo de dados e as chaves constituintes da tabela de dimensão *Dim_Resp_Inst_Role*. Esta tabela armazena a função desempenhada por cada instituição responsável por determinada iniciativa, tal como coordenador (*coordinator*), administrador (*administrator*) e consultor (*consultant*).

Tabela 5.13: Descrição da tabela de dimensão *Dim_Resp_Inst_Role*

Atributo	Tipo	Chave	Descrição
idResp_Inst_Role	Int - Not Null	PK	Identificador da tabela de dimensão <i>Dim_Resp_Inst_Role</i>
Resp_Inst_Role	Varchar - Not Null	-	Descrição do tipo de papel desempenhado por cada instituição

A próxima imagem (figura 5.15), que se segue, mostra as tabelas de dimensão *Partners_BTF*, *Dim_Partners*, *Dim_Types_Partners*, *Dim_Partners_Role*, bem com a relação com a tabela de factos *Fact_Use_Case*.

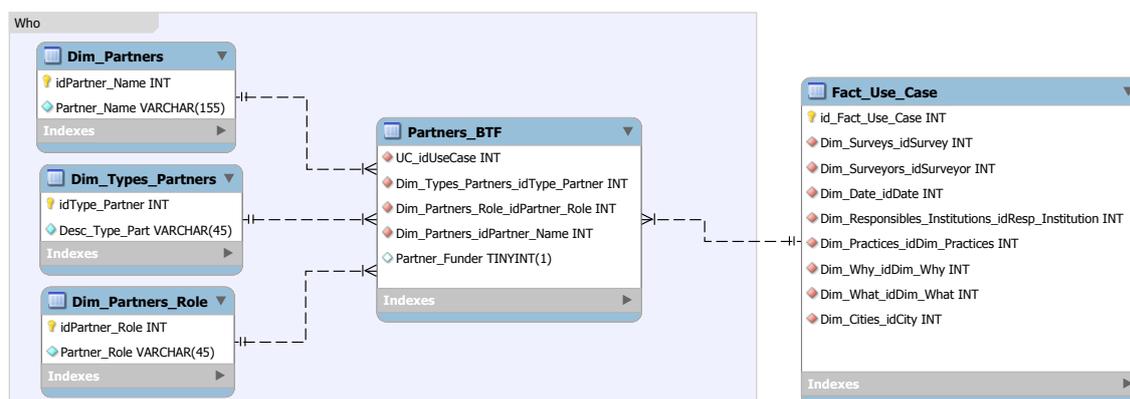


Figura 5.15: Tabelas de dimensão *Partners_BTF*, *Dim_Partners*, *Dim_Types_Partners*, *Dim_Partners_Role*

Estas tabelas de dimensão armazenam informações sobre os parceiros responsáveis pela realização de cada caso de estudo/iniciativa em Cidades Inteligentes. Um caso de estudo/iniciativa pode não estar associado a algum parceiro, como estar associado apenas a um ou ainda a vários parceiros.

A tabela de dimensão *Partners_BTF* armazena a informação que indica se cada parceiro associado é financiador ou não da iniciativa em causa, sendo por isso este atributo do tipo *tinyint*. Desta forma, é armazenado 0 em caso de o parceiro não ser financiador e 1 em caso contrário. Além disso, esta tabela serve também de tabela de ponte entre as tabelas de dimensão que lhe estão interligadas, sendo essas as tabelas *Dim_Partners*, *Dim_Types_Partners* e *Dim_Partners_Role*.

A tabela que se segue (tabela 5.14) apresenta a descrição de cada atributo, bem como a identificação do tipo de dados de cada um e das chaves constituintes da tabela de dimensão *Partners_BTF*.

Tabela 5.14: Descrição da tabela de dimensão *Partners_BTF*

Atributo	Tipo	Chave	Descrição
UC_id_UseCase	<i>Int - Not Null</i>	PK	Chave estrangeira para armazenar o identificador da tabela de factos Fact_Use_Case
Dim_Partners_Role_idPartner_Role	<i>Int - Not Null</i>	FK	Chave Estrangeira para armazenar o identificador da tabela de dimensão Dim_Partners_Role
Dim_Types_Partners_idType_Partner	<i>Int - Not Null</i>	FK	Chave Estrangeira para armazenar o identificador da tabela de dimensão Dim_Types_Partners
Dim_Partners_idPartner_Name	<i>Int - Not Null</i>	FK	Chave Estrangeira para armazenar o identificador da tabela de dimensão Dim_Partners
Partner_Funder	<i>Tinyint</i>	-	Armazena se o parceiro é ou não financiador da instituição contendo dois valores: 0 para caso não afirmativo ou 1 para caso afirmativo

A seguinte tabela (tabela 5.15) descreve os atributos da tabela de dimensão *Dim_Partners*, bem como o tipo de dados de cada um e as chaves constituintes. Esta tabela de dimensão armazena os nomes relativos a cada parceiro associado a um determinado caso de estudo/iniciativa.

Tabela 5.15: Descrição da tabela de dimensão *Dim_Partners*

Atributo	Tipo	Chave	Descrição
idPartner_Name	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Partners</i>
Partner_Name	<i>Varchar - Not Null</i>	-	Nome do parceiro associado à iniciativa

A próxima tabela (tabela 5.16) descreve a tabela de dimensão

Dim_Types_Partners, e, ainda, indica o tipo de dados de cada um e as chaves constituintes. A tabela de dimensão *Dim_Types_Partners* armazena a informação sobre o tipo de cada parceiro associado a um dado caso de estudo. Nesta tabela, por defeito, estão armazenados três tipos de parceiro: *Government* (Governo), *Industry* (Indústria), e *Non-Governmental Organization* (Organização Não Governamental). Um destes tipos terá de ser selecionado para cada parceiro associado à iniciativa, durante o preenchimento do formulário *online*.

Tabela 5.16: Descrição da tabela de dimensão *Dim_Types_Partners*

Atributo	Tipo	Chave	Descrição
idType_Partner	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Types_Partners</i>
Desc_Type_Part	<i>Varchar - Not Null</i>	-	Descrição do tipo de parceiro: <i>Government, Industry e Non-Governmental Organization</i>

A seguinte tabela (tabela 5.17) apresenta a descrição da tabela de dimensão *Dim_Partners_Role*, relativamente a cada atributo, ao tipo de dados de cada um e às chaves constituintes. Esta tabela de dimensão descreve o papel desempenhado por cada parceiro associado a uma iniciativa, como por exemplo: gestor (*manager*), assistente (*assistant*) e consultor (*consultant*).

Tabela 5.17: Descrição da tabela de dimensão *Dim_Partners_Role*

Atributo	Tipo	Chave	Descrição
idPartner_Role	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Partners_Role</i>
Partners_Role	<i>Varchar - Not Null</i>	-	Designação do tipo de papel desempenhado por cada parceiro associado à iniciativa

- **Descrição das tabelas de dimensão referentes à questão “*Where?*”**

Este grupo de tabelas, que se referem à questão “*Where?*”, integra as tabelas de dimensão *Dim_Date* e *Dim_Cities*.

A figura 5.16 apresenta a tabela de dimensão *Dim_Date*, bem com a relação com a tabela de factos *Fact_Use_Case*.

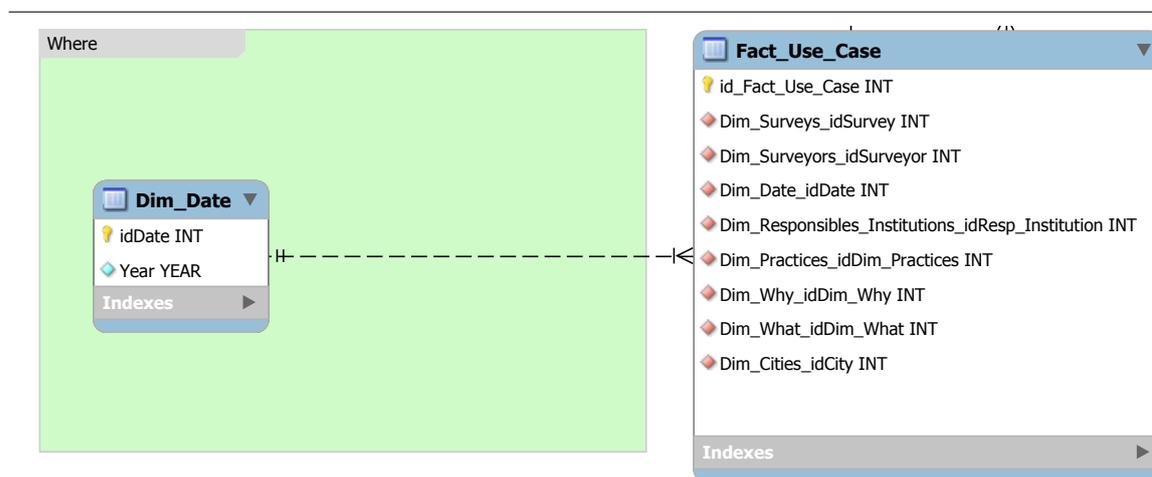


Figura 5.16: Tabela de dimensão *Dim_Date*

A tabela de dimensão *Dim_Date* tem o objetivo de armazenar os dados referentes à data, mais especificamente ao ano, do desenvolvimento do caso de estudo/iniciativa.

Na tabela que se segue (tabela 5.18) é apresentada a descrição da tabela de dimensão *Dim_Date*, nomeadamente a identificação de cada atributo, o tipo de dados de cada um e as chaves constituintes.

Tabela 5.18: Descrição da tabela de dimensão *Dim_Date*

Atributo	Tipo	Chave	Descrição
idDate	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Date</i>
Year_ce	<i>Year - Not Null</i>	-	Ano de início do desenvolvimento da iniciativa

A figura 5.17, que se segue, mostra a tabela de dimensão *Dim_Cities*, bem com a sua relação com a tabela de factos *Fact_Use_Case*.

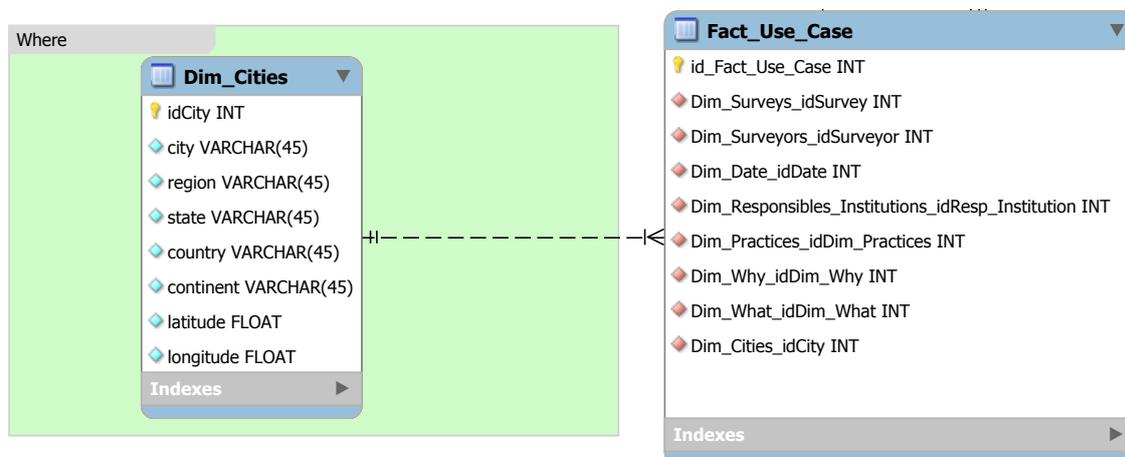


Figura 5.17: Tabela de dimensão *Dim_Cities*

A tabela *Dim_Cities* armazena a informação relativa à localização do caso de estudo/iniciativa em Cidades Inteligentes, nomeadamente, a cidade, o/a município/-região, o/a estado/província, o país, o continente, e ainda, a latitude e a longitude.

Em seguida, a tabela 5.19 descreve a tabela de dimensão *Dim_Cities* relativamente a cada atributo, ao tipo de dados de cada um e às chaves constituintes.

Tabela 5.19: Descrição da tabela de dimensão *Dim_Cities*

Atributo	Tipo	Chave	Descrição
idCity	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Cities</i>
city	<i>Varchar - Not Null</i>	-	Nome da cidade referente à iniciativa
region	<i>Varchar - Not Null</i>	-	Nome do município ou região referente à iniciativa
state	<i>Varchar - Not Null</i>	-	Nome da província ou estado referente à iniciativa
country	<i>Varchar - Not Null</i>	-	Nome do país referente à iniciativa
continent	<i>Varchar - Not Null</i>	-	Nome do continente referente à iniciativa
latitude	<i>Float - Not Null</i>	-	Latitude referente à iniciativa
longitude	<i>Float - Not Null</i>	-	Longitude referente à iniciativa

- **Descrição das tabelas de dimensão referentes à questão ‘What?’**

As tabelas, que estão relacionadas com a questão “What?” (“O quê/Qual?”), incluem algumas informações sobre a própria iniciativa. Exemplo dessas informações são: uma breve descrição sobre a cidade na qual a iniciativa foi realizada; os principais conceitos subjacentes à iniciativa; o objetivo da prática; as principais inovações ocasionadas; e as experiências adquiridas.

Este grupo de tabelas é constituído pelas tabelas de dimensão *Dim_What*, *Dim_Concepts*, *Dim_Innovations* e *Dim_Lessons_Learnt*. Também é composto pelas tabelas de ponte *Concepts_BDWhat*, *Innovations_BDWhat* e *Lessons_BDWhat*.

A tabela de dimensão *Dim_What* armazena informações relativas a uma breve descrição da cidade pertencente à iniciativa, às suas inovações e experiências adquiridas, e outros comentários sobre a mesma.

A tabela seguinte (tabela 5.20) expõe a descrição de cada atributo, bem como a identificação do tipo de dados de cada um e das chaves constituintes da tabela de dimensão *Dim_What*.

Tabela 5.20: Descrição da tabela de dimensão *Dim_What*

Atributo	Tipo	Chave	Descrição
idDim_What	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_What</i>
City_Background	<i>Varchar - Not Null</i>	-	Descrição breve sobre a cidade referente à iniciativa
What_is_about_summary	<i>Varchar</i>	-	Resumo breve sobre os conceitos relativos à iniciativa
Desc_Innovations	<i>Varchar</i>	-	Descrição breve sobre as inovações (<i>Innovations</i>) trazidas pela iniciativa
Desc_Lessons_Learnt	<i>Varchar</i>	-	Descrição breve sobre as experiências obtidas (<i>Lessons Learnt</i>)
Comments	<i>Varchar</i>	-	Breves comentários sobre o objetivo da iniciativa

Uma vez que cada iniciativa pode estar relacionada a mais que um conceito, inovação ou lição aprendida, existe para cada um destes casos uma tabela de ponte associ-

ada. Assim, as tabelas de dimensão *Dim_Concepts*, *Dim_Innovations* e *Dim_Lessons_Learnt* estão interligadas à tabela de dimensão *Dim_What* através das tabelas de ponte *Concepts_BDWhat*, *Innovations_BDWhat* e *Lessons_BDWhat*, respetivamente.

A próxima figura (figura 5.18) retrata a tabela de dimensão *Dim_What* e *Dim_Concepts*, a tabela de ponte *Concepts_BDWhat* e, ainda, a relação com a tabela de factos *Fact_Use_Case*.

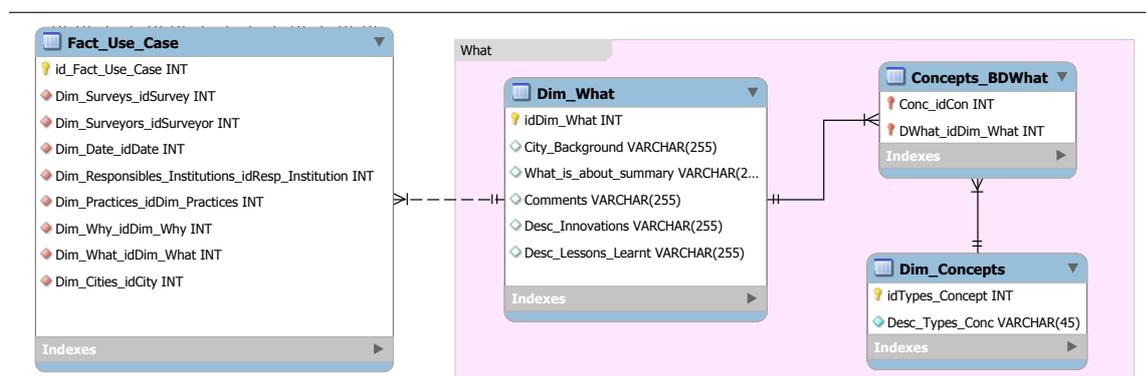


Figura 5.18: Tabelas de dimensão *Dim_What* e *Dim_Concepts* e tabela de ponte *Concepts_BDWhat*

A tabela 5.21, que se segue, descreve os atributos da tabela de ponte *Concepts_BDWhat*, a qual existe entre as tabela de dimensão *Dim_Concepts* e *Dim_What*.

Tabela 5.21: Descrição da tabela de ponte *Concepts_BDWhat*

Atributo	Tipo	Chave	Descrição
Conc_idCon	Int - Not Null	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_Concepts</i>
DWhat_idDim_What	Int - Not Null	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_What</i>

A tabela de dimensão *Dim_Concepts* armazena um conjunto de tipos de conceitos já pré-definidos, pelo qual terá de ser selecionado pelo menos um durante o preenchimento do formulário *online*. Desta lista definida constam os seguintes ti-

pos: *Energy City* (Cidade da Energia), *Digital City* (Cidade Digital), *Intelligent City* (Cidade Inteligente), *Knowledge City* (Cidade do Conhecimento), *Eco City* (Cidade Ecológica), *Ubiquitous City* (Cidade Ubíqua) e *Smart City* (Cidade Inteligente).

A tabela 5.22, que se segue, descreve os atributos, o tipo de dados de cada um e as chaves constituintes da tabela de dimensão *Dim_Concepts*.

Tabela 5.22: Descrição da tabela de dimensão *Dim_Concepts*

Atributo	Tipo	Chave	Descrição
idType_Concept	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Concepts</i>
Desc_Type_Conc	<i>Varchar - Not Null</i>	-	Designação de cada tipo de conceito

A figura seguinte (figura 5.19) expõe as tabelas de dimensão *Dim_What* e *Dim_Innovations* e a tabela de ponte *Innovations_BDWhat*, bem como, a relação com a tabela de factos *Fact_Use_Case*.

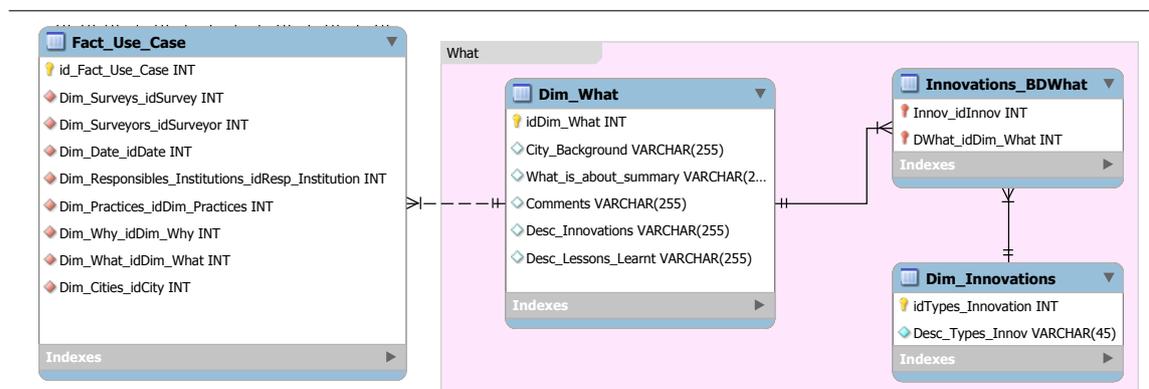


Figura 5.19: Tabelas de dimensão *Dim_What* e *Dim_Innovations* e tabela de ponte *Innovations_BDWhat*

A próxima tabela (tabela 5.23) descreve os atributos da tabela de ponte *Innovations_BDWhat*. Esta tabela de ponte existe entre as dimensões *Dim_What* e *Dim_Innovations*.

Tabela 5.23: Descrição da tabela de ponte *Innovations_BDWhat*

Atributo	Tipo	Chave	Descrição
Innov_idInnov	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_Innovations</i>
DWhat_idDim_What	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_What</i>

A tabela de dimensão *Dim_Innovations* armazena um conjunto de tipos de inovações já pré-definidos, na qual terá de ser selecionado pelo menos um durante o preenchimento do formulário *online*. Desta lista definida constam os seguintes tipos: *Legal/Regulatory* (Legar/Regulamentar), *Planning* (Planeamento), *Technology* (Tecnologia), *Participation* (Participação), *Service Delivery* (Serviço de entrega).

A tabela seguinte (tabela 5.24) descreve os atributos, indicando o tipo de dados de cada um, bem como, as chaves constituintes da tabela de dimensão *Dim_Innovations*.

Tabela 5.24: Descrição da tabela de dimensão *Dim_Innovations*

Atributo	Tipo	Chave	Descrição
idType_Innovation	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Innovations</i>
Desc_Type_Innov	<i>Varchar - Not Null</i>	-	Designação de cada tipo de inovação

A figura 5.20, que se segue, apresenta as tabelas de dimensão *Dim_What* e *Dim_Lessons_Learnt*, a tabela de ponte *Lessons_BDWhat*, bem como, a relação com a tabela de factos *Fact_Use_Case*.

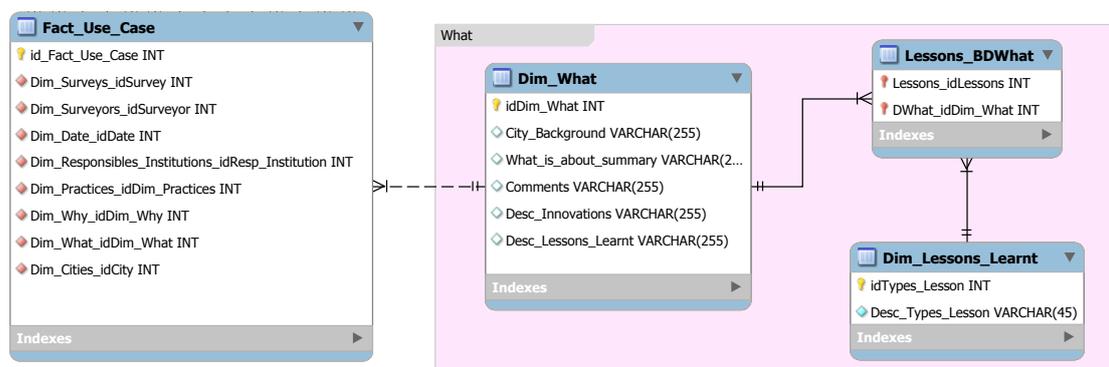


Figura 5.20: Tabelas de dimensão *Dim_What* e *Dim_Lessons_Learnt* e tabela de ponte *Lessons_BDWhat*

A seguinte tabela (tabela 5.25) descreve os atributos da tabela de ponte *Lessons_BDWhat*, sendo esta existente entre as tabelas de dimensão *Dim_Lessons_Learnt* e *Dim_What*.

Tabela 5.25: Descrição da tabela de ponte *Lessons_BDWhat*

Atributo	Tipo	Chave	Descrição
Lessons_idLessons	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_Lessons_Learnt</i>
DWhat_idDim_What	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_What</i>

A tabela de dimensão *Dim_Lessons_Learnt* armazena um conjunto de tipos de experiências/aprendizagens adquiridas com o desenvolvimento da iniciativa já definidos. Estes tipos já são pré-definidos, tendo de ser selecionado pelo menos um durante o preenchimento do formulário *online*. Desta lista constam os seguintes tipos: *City Participation* (Participação da Cidade), *Tools* (Ferramentas), *Economic Development/Opportunities* (Desenvolvimento Económico/Oportunidades), *Consortium of Stakeholders* (Consórcio de Partes interessadas) e *Innovations* (Inovações).

A tabela seguinte (tabela 5.26) mostra a descrição de cada um dos atributos da tabela de dimensão *Dim_Lessons_Learnt*, incluindo a indicação do tipo de dados de cada um e as chaves constituintes.

Tabela 5.26: Descrição da tabela de dimensão *Dim_Lessons_Learnt*

Atributo	Tipo	Chave	Descrição
idType_Lesson	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Lessons_Learnt</i>
Desc_Type_Lesson	<i>Varchar - Not Null</i>	-	Designação de cada tipo de experiência/aprendizagem

- **Descrição das tabelas de dimensão referentes à questão “*Why?*”**

As tabelas, que estão relacionadas com a questão “*Why?*” (“Porquê?”), contêm a informação sobre a razão para a realização da iniciativa e para que dimensões a iniciativa contribui, sendo estas: o desenvolvimento económico, a governação, a mobilidade, o ambiente, o desenvolvimento social (incluindo capital social e humano) e a qualidade de vida. Além disso, também procura reunir informação sobre os benefícios e os valores promovidos pela iniciativa.

Este grupo de tabelas é constituído pelas tabelas de dimensão *Dim_Why*, *Dim_Drivers*, *Dim_Benefits* e *Dim_Values*. Além disso, este grupo é composto pelas tabelas ponte *Drivers_BDWhy*, *Benefits_BDWhy* e *Values_BDWhy*.

A tabela de dimensão *Dim_Why* armazena informações como breves descrições sobre os fatores que levaram ao impulso da iniciativa (*Drivers*), os benefícios (*Benefits*) e os valores *Values* fomentados pela iniciativa.

A tabela seguinte (tabela 5.27) apresenta a descrição da tabela de dimensão *Dim_Why*, relativamente aos atributos, ao tipo de dados de cada um e às chaves constituintes.

Tabela 5.27: Descrição da tabela de dimensão *Dim_Why*

Atributo	Tipo	Chave	Descrição
idDim_Why	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Why</i>
Outcome_Drivers	<i>Varchar</i>	-	Descrição breve sobre os fatores (<i>Drivers</i>) que levaram ao impulso da iniciativa
Desc_Benefits	<i>Varchar</i>	-	Resumo breve sobre os benefícios (<i>Benefits</i>) da prática da iniciativa

Desc_Values	Varchar	-	Descrição breve sobre os valores (<i>Values</i>) promovidos pela iniciativa
-------------	---------	---	---

Uma vez que cada iniciativa pode estar relacionada a mais que um fator, a diversos benefícios e à promoção de vários valores, existe para cada uma destes casos uma tabela de ponte associada. Desta forma, as tabelas de dimensão *Dim_Drivers*, *Dim_Benefits* e *Dim_Values* estão interligadas à tabela de dimensão *Dim_Why* através das tabelas de ponte *Drivers_BDWhy*, *Benefits_BDWhy* e *Values_BDWhy*, respetivamente.

A figura 5.21, que se segue, expõe as tabelas de dimensão *Dim_Why* e *Dim_Drivers*, a tabela de ponte *Drivers_BDWhy*, bem como, a relação com a tabela de factos *Fact_Use_Case*.

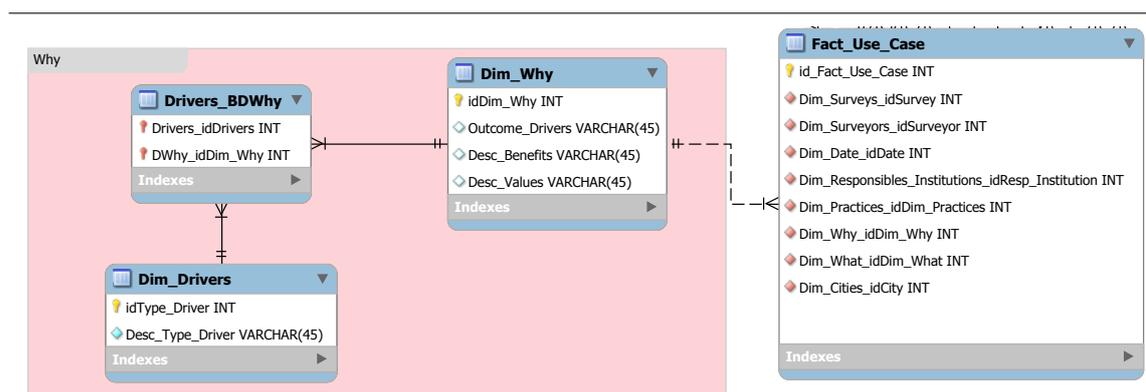


Figura 5.21: Tabelas de dimensão *Dim_Why* e *Dim_Drivers* e tabela de ponte *Drivers_BDWhy*

A próxima tabela (tabela 5.28) descreve os atributos da tabela de ponte *Drivers_BDWhy*, a qual existe entre as tabelas de dimensão *Dim_Why* e *Dim_Drivers*.

Tabela 5.28: Descrição da tabela de ponte *Drivers_BDWhy*

Atributo	Tipo	Chave	Descrição
Drivers_idDrivers	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_Drivers</i>
DWhy_idDim_Why	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_Why</i>

A tabela de dimensão *Dim_Drivers* armazena um conjunto de categorias de fatores que motivam, fornecem e permitem a criação e implementação de iniciativas em Cidades Inteligentes. Estas categorias já se encontram pré-definidas, pelo qual terá de ser selecionada pelo menos uma durante o preenchimento do formulário *online*. Desta lista constam as seguintes categorias: *Economic Development* (Desenvolvimento Económico), *Governance* (Governança), *Mobility* (Mobilidade), *Environment* (Ambiente), *Social Development* (Desenvolvimento Social) e *Quality of Life* (Qualidade de Vida).

A tabela 5.29, que se segue, expõe a descrição da tabela de dimensão *Dim_Drivers*, nomeadamente a identificação dos atributos, do tipo de dados de cada um e das chaves constituintes.

Tabela 5.29: Descrição da tabela de dimensão *Dim_Drivers*

Atributo	Tipo	Chave	Descrição
idType_Driver	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Drivers</i>
Desc_Type_Driver	<i>Varchar - Not Null</i>	-	Designação de cada categoria de fatores impulsionadores (<i>Drivers</i>)

A figura 5.22 apresenta as tabelas de dimensão *Dim_Why* e *Dim_Benefits* e a tabela de ponte *Benefits_BDWhy*, bem como, a relação com a tabela de factos *Fact_Use_Case*.

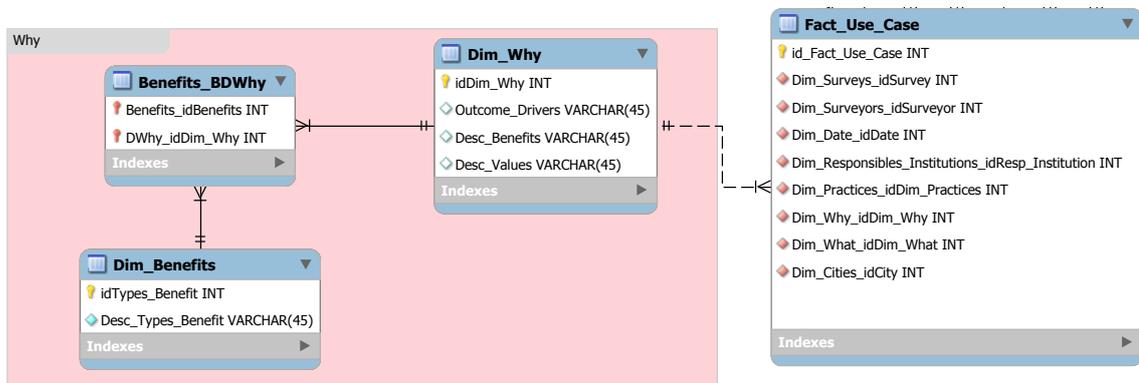


Figura 5.22: Tabelas de dimensão *Dim_Why* e *Dim_Benefits* e tabela de ponte *Benefits_BDWhy*

A seguinte tabela (tabela 5.30) descreve os atributos da tabela de ponte *Benefits_BDWhy*, a qual existe entre as tabelas de dimensão *Dim_Benefits* e *Dim_Why*.

Tabela 5.30: Descrição da tabela de ponte *Benefits_BDWhy*

Atributo	Tipo	Chave	Descrição
Benefits_idBenefits	Int - Not Null	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_Benefits</i>
DWhy_idDim_Why	Int - Not Null	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_Why</i>

A tabela de dimensão *Dim_Benefits* armazena um conjunto de categorias, relativas aos benefícios previamente definidas, pelo qual terá de ser selecionada pelo menos uma durante o preenchimento do formulário *online*. Desta lista definida constam as seguintes categorias: *City Planning* (Planeamento da Cidade), *New and Enhanced Digital Services* (Novos e aperfeiçoados Serviços Públicos), *Environment Protection* (Proteção do Ambiente), *Human Development* (Desenvolvimento Humano), *Infrastructure for Public Services* (Infraestrutura para Serviços Públicos), *Governance and Participation* (Governança e Participação), *Economic Growth* (Crescimento Económico) e *Improving Quality of Life* (Melhoramento da Qualidade de Vida).

A tabela seguinte (tabela 5.31) descreve cada atributo da tabela de dimensão

Dim_Benefits, no qual indica, também, o tipo de dados de cada um e as chaves constituintes.

Tabela 5.31: Descrição da tabela de dimensão *Dim_Benefits*

Atributo	Tipo	Chave	Descrição
idType_Benefit	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Benefits</i>
Desc_Type_Benefit	<i>Varchar - Not Null</i>	-	Designação de cada categoria de benefícios (<i>benefits</i>)

A figura 5.23 retrata as tabelas de dimensão *Dim_Why* e *Dim_Values* e a tabela de ponte *Values_BDWhy*, bem como, a relação com a tabela de factos *Fact_Use_Case*.

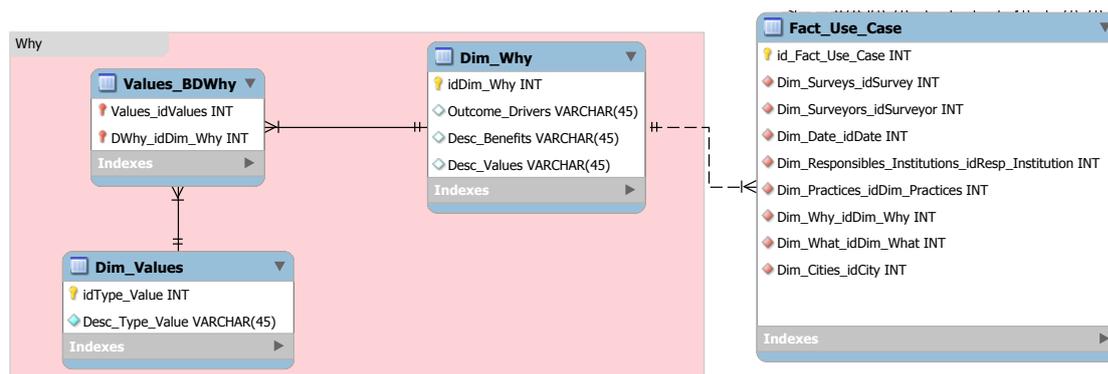


Figura 5.23: Tabelas de dimensão *Dim_Why* e *Dim_Values* e tabela de ponte *Values_BDWhy*

A próxima tabela (tabela 5.32) descreve os atributos da tabela de ponte *Values_BDWhy*. Esta tabela de ponte existe entre as tabelas de dimensão *Dim_Why* e *Dim_Values*.

Tabela 5.32: Descrição da tabela de ponte *Values_BDWhy*

Atributo	Tipo	Chave	Descrição
Values_idValues	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_Values</i>

DWhy_idDim_Why	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão Dim_Why
----------------	-----------------------	----	--

A tabela de dimensão *Dim_Values* armazena um conjunto de categorias, relativas aos valores promovidos pela iniciativa, previamente definidas, sendo que terá de ser selecionada pelo menos uma durante o preenchimento do formulário *online*. Desta lista definida constam as seguintes categorias: *Citizens Quality of Life* (Qualidade de Vida dos Cidadãos), *Economic* (Económico), *Region* (Região), *City Attributes* (Atributos da Cidade), *Society Social Values* (Valores sociais da Sociedade), *Environment* (Ambiente), *Human Capital* (Capital Humano), *City Infrastructure* (Infraestrutura da Cidade), *City Services* (Serviços da Cidade) e *Governance* (Governança).

A tabela seguinte (tabela 5.33) apresenta a descrição dos atributos da tabela de dimensão *Dim_Values*, incluindo a identificação do tipo de dados de cada um e as chaves constituintes.

Tabela 5.33: Descrição da tabela de dimensão *Dim_Values*

Atributo	Tipo	Chave	Descrição
idType_Value	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão Dim_Values
Desc_Type_Value	<i>Varchar - Not Null</i>	-	Designação de cada categoria dos valores (<i>values</i>)

- **Descrição das tabelas de dimensão referentes à questão “How?”**

As tabelas, que estão relacionadas com a questão “How?” (“Como?”), documentam como a iniciativa foi implementada, nomeadamente o modelo de governação, os riscos, as tecnologias, os estados de maturidade, as ferramentas e os desafios.

Neste grupo de tabelas estão integradas as tabelas de dimensão *Dim_Governances*, *Dim_Risks*, *Dim_Technologies*, *Dim_Maturities*, *Dim_Tools* e *Dim_Challenges*. Além disso, este grupo é composto pelas tabelas de ponte *Governances_BTF*, *Risks_BTF*, *Technologies_BTF*, *Maturities_BTF*, *Tools_BTF* e *Challenges_BTF*.

A figura 5.24 expõe a tabela de ponte *Governances_BTF* e a tabela de dimensão *Dim_Governances*, bem como, a relação com a tabela de factos *Fact_Use_Case*.

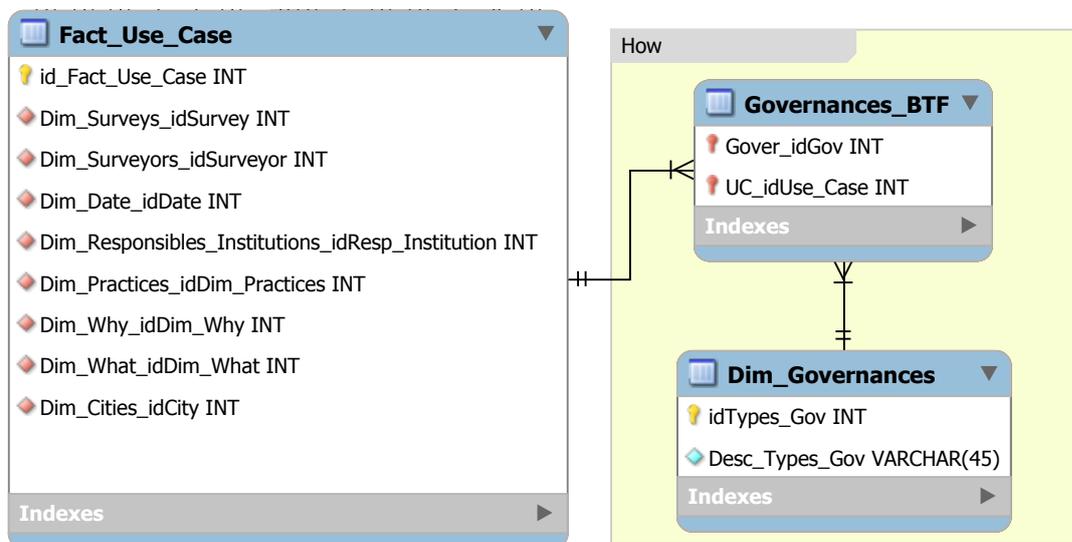


Figura 5.24: Tabela de ponte *Governances_BTF* e tabela de dimensão *Dim_Governances*

A próxima tabela (tabela 5.34) descreve os atributos da tabela de ponte *Governances_BTF*, a qual existe entre a tabela de dimensão *Dim_Governances* e a tabela de factos (*Fact_Use_Case*).

Tabela 5.34: Descrição da tabela de ponte *Governances_BTF*

Atributo	Tipo	Chave	Descrição
Gover_idGov	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_Governances</i>
UC_idUse_Case	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de factos <i>Fact_Use_Case</i>

A tabela de dimensão *Dim_Governances* armazena um conjunto de categorias relativas à forma como o governo opera sobre a Cidade Inteligente. Estas categorias estão relacionadas ao modo como são geridos os fundos públicos, a infraestrutura e os serviços públicos disponibilizados; como o desenvolvimento de cidade sustentável é suportado; e como os cidadãos são envolvidos nos processos de tomada de decisão. Estas categorias estão já definidas, pelo qual terá de ser seleccionada pelo menos

uma durante o preenchimento do formulário *online*. Desta lista definida constam as seguintes categorias: *Principles* (Princípios), *Model* (Modelo), *Vision* (Visão) e *Resources* (Recursos).

A tabela seguinte (tabela 5.35) descreve os atributos da tabela de dimensão *Dim_Governances*, indicando o tipo de dados de cada um, bem como, as chaves constituintes.

Tabela 5.35: Descrição da tabela de dimensão *Dim_Governances*

Atributo	Tipo	Chave	Descrição
idType_Gov	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Governances</i>
Desc_Type_Gov	<i>Int - Not Null</i>	-	Designação de cada categoria de governação (<i>governance</i>)

A próxima figura (figura 5.25) apresenta a tabela de ponte *Risks_BTF* e a tabela de dimensão *Dim_Risks*, bem como, a relação com a tabela de factos *Fact_Use_Case*.

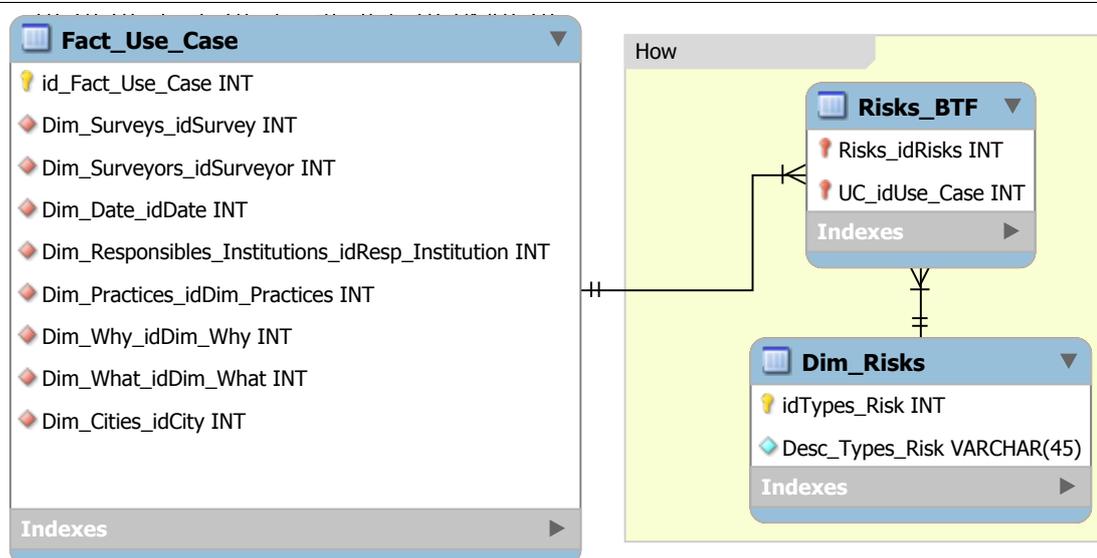


Figura 5.25: Tabela de ponte *Risks_BTF* e tabela de dimensão *Dim_Risks*

A tabela 5.36, que se segue, mostra a descrição dos atributos da tabela de ponte *Risks_BTF*. Esta tabela de ponte existe entre a tabela de dimensão *Dim_Risks* e a tabela de factos (*Fact_Use_Case*).

Tabela 5.36: Descrição da tabela de ponte *Risks_BTf*

Atributo	Tipo	Chave	Descrição
Risks_idRisks	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão Dim_Risks
UC_idUse_Case	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de factos Fact_Use_Case

A tabela de dimensão *Dim_Risks* armazena um conjunto de tipos de riscos que cada Cidade Inteligente enfrenta. Estes tipos estão previamente definidos, e terá de ser selecionado pelo menos um durante o preenchimento do formulário *online*. Desta lista constam os seguintes tipos de riscos: *Economic* (Económico), *Environmental* (Ambiental), *Social* (Social), *Financial* (Financeiro) e *Technical* (Técnico).

A tabela seguinte (tabela 5.37) apresenta a descrição da tabela de dimensão *Dim_Risks*, relativamente a cada atributo, ao tipo de dados e às chaves constituintes.

Tabela 5.37: Descrição da tabela de dimensão *Dim_Risks*

Atributo	Tipo	Chave	Descrição
idType_Risk	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão Dim_Risks
Desc_Type_Risk	<i>Int - Not Null</i>	-	Designação de cada tipo de risco (<i>risk</i>)

A próxima figura (figura 5.26) retrata a tabela de ponte *Technologies_BTf* e a tabela de dimensão *Dim_Technologies* bem como, a relação com a tabela de factos *Fact_Use_Case*.

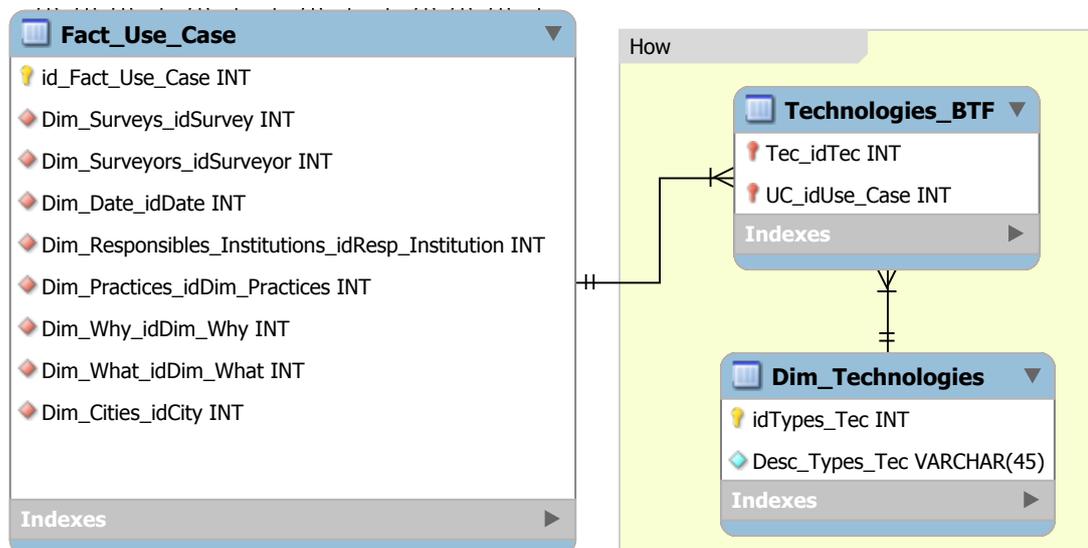


Figura 5.26: Tabela de ponte *Technologies_BTF* e tabela de dimensão *Dim_Technologies*

A próxima tabela (tabela 5.36) expõe a descrição dos atributos da tabela de ponte *Technologies_BTF*, a qual existe entre a tabela de dimensão *Dim_Technologies* e a tabela de factos (*Fact_Use_Case*).

Tabela 5.38: Descrição da tabela de ponte *Technologies_BTF*

Atributo	Tipo	Chave	Descrição
Tec_idTec	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_Technologies</i>
UC_idUse_Case	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de factos <i>Fact_Use_Case</i>

A tabela de dimensão *Dim_Technologies* armazena um conjunto de categorias relativas às tecnologias que cada Cidade Inteligente faz uso para promover o seu desenvolvimento. Estas categorias estão já pré-definidas, e terá de ser selecionada pelo menos uma durante o preenchimento do formulário *online*. Desta lista constam as seguintes categorias de tecnologias: *Basic Infrastructure Technologies* (Infraestrutura básica para Tecnologias), *ICT Infrastructure Technologies* (Infraestrutura de TIC

para Tecnologias), *ICT Hardware Tools* (Ferramentas de TIC para *Hardware*), *ICT Software Tools* (Ferramentas de TIC para *Software*) e *ICT Techniques* (Técnicas de TIC).

A tabela seguinte (tabela 5.39) descreve os atributos da tabela de dimensão *Dim_Technologies*, incluindo o tipo de dados de cada um e as chaves constituintes.

Tabela 5.39: Descrição da tabela de dimensão *Dim_Technologies*

Atributo	Tipo	Chave	Descrição
idType_Tec	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Technologies</i>
Desc_Type_Tec	<i>Int - Not Null</i>	-	Designação de cada categoria de tecnologia (<i>technology</i>)

A seguinte figura (figura 5.27) apresenta a tabela de ponte *Maturities_BTF* e a tabela de dimensão *Dim_Maturities*, bem como, a relação com a tabela de factos *Fact_Use_Case*.

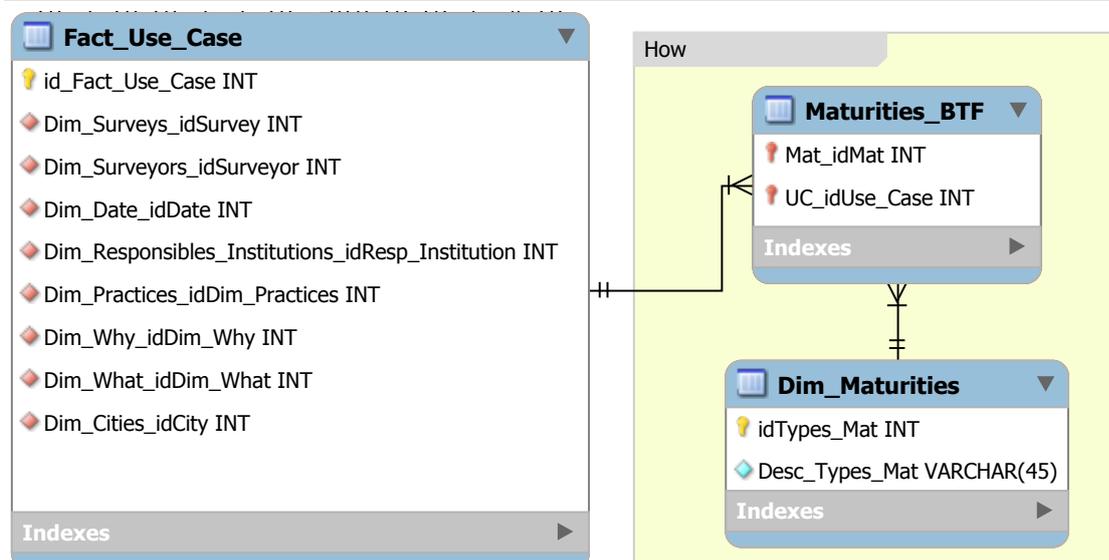


Figura 5.27: Tabela de ponte *Maturities_BTF* e tabela de dimensão *Dim_Maturities*

A seguinte tabela (tabela 5.40) expõe a descrição dos atributos da tabela de ponte *Maturities_BTF*. Esta tabela de ponte existe entre a tabela de dimensão *Dim_Maturities* e a tabela de factos (*Fact_Use_Case*).

Tabela 5.40: Descrição da tabela de ponte *Maturities_BTTF*

Atributo	Tipo	Chave	Descrição
Mat_idMat	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_Maturities</i>
UC_idUse_Case	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de factos <i>Fact_Use_Case</i>

A tabela de dimensão *Dim_Maturities* armazena um conjunto de categorias relativas aos modelos de maturidade que representam o progresso do desenvolvimento de uma Cidade Inteligente, ao longo de uma série de estados de maturidade discretos. Estas categorias estão já pré-definidas, pelo qual terá de ser seleccionada pelo menos uma durante o preenchimento do formulário *online*. Desta lista constam as seguintes categorias de modelos de maturidade: *Measurement* (Medição), *Practices* (Práticas), *Approach* (Abordagem), *Scope* (Âmbito) e *Analysis Tools* (Ferramentas de Análise).

A tabela seguinte (tabela 5.41) apresenta a descrição dos atributos da tabela de dimensão *Dim_Maturities*, incluindo a identificação do tipo de dados de cada um e das chaves constituintes.

Tabela 5.41: Descrição da tabela de dimensão *Dim_Maturities*

Atributo	Tipo	Chave	Descrição
idType_Mat	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Maturities</i>
Desc_Type_Mat	<i>Int - Not Null</i>	-	Designação de cada categoria de maturidade (<i>Maturity</i>)

A próxima figura (figura 5.28) apresenta a tabela de ponte *Tools_BTTF* e a tabela de dimensão *Dim_Tools*, bem como, a relação com a tabela de factos *Fact_Use_Case*.

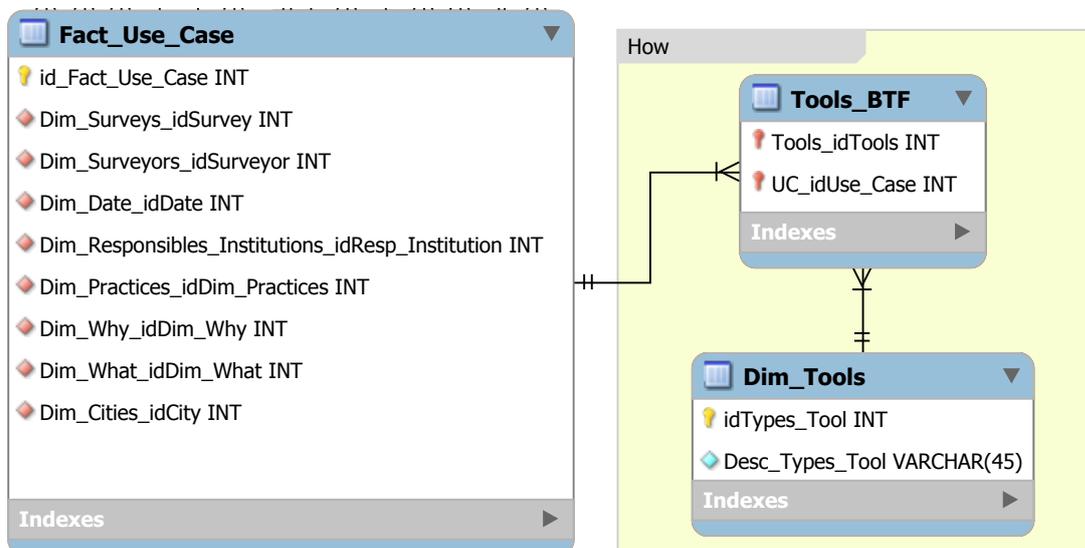


Figura 5.28: Tabela de ponte *Tools_BTF* e tabela de dimensão *Dim_Tools*

A tabela 5.42, que se segue, descreve os atributos da tabela de ponte *Tools_BTF*, a qual existe entre a tabela de dimensão *Dim_Tools* e a tabela de factos (*Fact_Use_Case*).

Tabela 5.42: Descrição da tabela de ponte *Tools_BTF*

Atributo	Tipo	Chave	Descrição
Tools_idTools	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_Tools</i>
UC_idUse_Case	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de factos <i>Fact_Use_Case</i>

A tabela de dimensão *Dim_Tools* armazena um conjunto de categorias que representam os diversos guias e instrumentos conceituais, técnicos e metodológicos, sendo que muitos destes encontram-se disponíveis em plataformas digitais, para apoiarem o planeamento e a implementação de iniciativas. Estas categorias estão já pré-definidas, pelo qual terá de ser selecionada pelo menos uma durante o preenchimento do formulário *online*. Desta lista constam as seguintes categorias de ferramentas:

Monitoring (Monitorização), *Governance and Management* (Governança e Gestão), *Operations* (Operações), *Regulatory/Legal* (Regulamentação/Legal) e *Planning and Implementation* (Planeamento e Implementação).

A tabela seguinte (tabela 5.43) apresenta a descrição de cada um dos atributos da tabela de dimensão *Dim_Tools*, incluindo a indicação do tipo de dados de cada um e das chaves constituintes.

Tabela 5.43: Descrição da tabela de dimensão *Dim_Tools*

Atributo	Tipo	Chave	Descrição
idType_Tool	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Tools</i>
Desc_Type_Tool	<i>Int - Not Null</i>	-	Designação de cada categoria de ferramenta (<i>Tool</i>)

A figura 5.29, que se segue, apresenta a tabela de ponte *Challenges_BTF*, a tabela de dimensão *Dim_Challenges* e, ainda, a relação com a tabela de factos *Fact_Use_Case*.

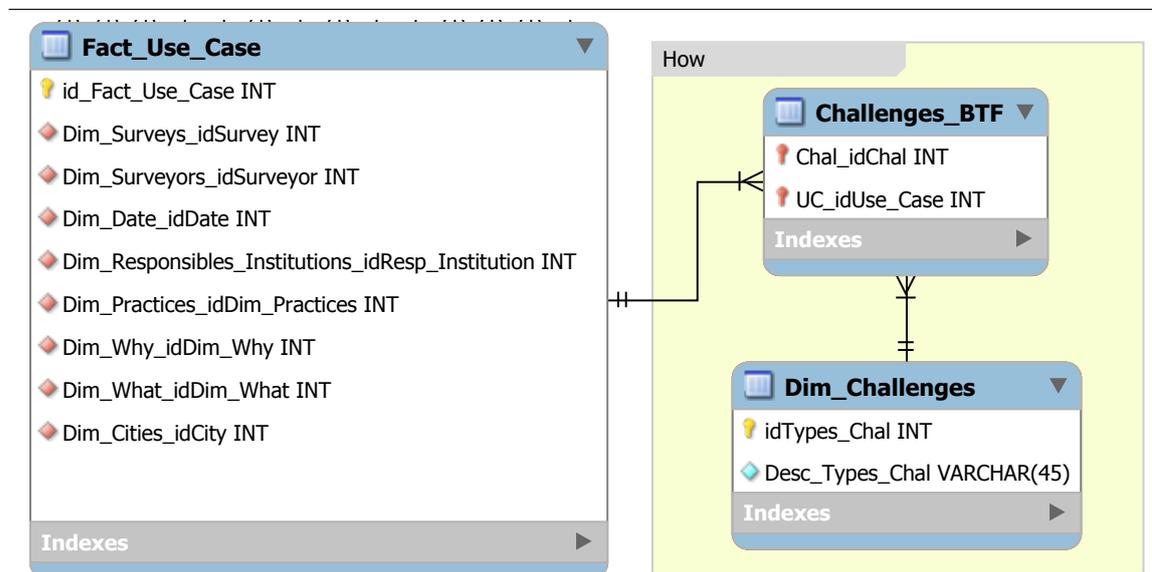


Figura 5.29: Tabela de ponte *Challenges_BTF* e tabela de dimensão *Dim_Challenges*

A próxima tabela (tabela 5.44) descreve os atributos da tabela de ponte *Challenges_BTF*. Esta tabela de ponte existe entre a tabela de dimensão *Dim_Challenges*

e a tabela de factos (*Fact_Use_Case*).

Tabela 5.44: Descrição da tabela de ponte *Challenges_BTF*

Atributo	Tipo	Chave	Descrição
Chal_idChal	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de dimensão <i>Dim_Challenges</i>
UC_idUse_Case	<i>Int - Not Null</i>	FK	Chave estrangeira que armazena o identificador da tabela de factos <i>Fact_Use_Case</i>

A tabela de dimensão *Dim_Challenges* armazena um conjunto de categorias relativas aos diversos desafios no qual uma Cidade Inteligente enfrenta para ultrapassar barreiras e obstáculos relacionados com a criação e implementação da iniciativa ou para fornecer soluções. Estas categorias estão já pré-definidas, e terá de ser selecionada pelo menos uma durante o preenchimento do formulário *online*. Desta lista constam as seguintes categorias de desafios: *Economic* (Económico), *Human Capital* (Capital Humano), *Plan and Implementation* (Planeamento e Implementação), *Quality* (Qualidade), *Governance* (Governança), *Financial* (Financeiro), *Technical* (Técnico), *Social* (Social) e *Environmental* (Ambiental).

A tabela 5.45, que se segue, descreve cada um dos atributos da tabela de dimensão *Dim_Challenges*, no qual indica, também, o tipo de dados de cada um e as chaves constituintes.

Tabela 5.45: Descrição da tabela de dimensão *Dim_Challenges*

Atributo	Tipo	Chave	Descrição
idType_Chall	<i>Int - Not Null</i>	PK	Identificador da tabela de dimensão <i>Dim_Challenges</i>
Desc_Type_Chall	<i>Int - Not Null</i>	-	Designação de cada categoria de desafios (<i>Challenges</i>)

5.2.2 Cubo OLAP

Após a apresentação do modelo multidimensional dos dados do sistema de *Data Warehouse*, esta secção mostra como foi elaborado o esquema do cubo OLAP. Este cubo OLAP servirá de suporte aos sistemas OLAP, de modo a obter as análises pretendidas neste trabalho. A vantagem de um cubo multidimensional assenta numa obtenção rápida de um grande volume de dados sobre diversas dimensões.

A fonte de dados do cubo multidimensional tem origem no formulário *online* desenvolvido para recolher dados sobre as diversas iniciativas em Cidades Inteligentes.

A elaboração do esquema do cubo OLAP foi desenvolvido com recurso à ferramenta *Pentaho Schema Workbench*. O esquema do cubo elaborado descreve a tabela de factos, as diversas dimensões e respetivas tabelas associadas, e ainda, a medida que permitirá o cálculo do número de iniciativas. Esta ferramenta é vantajosa pois possui integração com a ferramenta *Pentaho*, permitindo assim publicar o esquema criado diretamente no servidor OLAP do *Pentaho* (*Pentaho BI Server*).

A figura 5.30, que se segue, demonstra o esquema do cubo desenvolvido na ferramenta *Pentaho Schema Workbench*.

O cubo OLAP foi desenvolvido de acordo com o modelo multidimensional dos dados apresentado anteriormente na secção 5.2. Uma vez que existem relações de muitos-para-muitos, no qual a ferramenta *Pentaho Schema Workbench* não suporta este tipo de relações, a solução encontrada assentou na elaboração de *views*¹ na base de dados *MySQL*.

A *view* elaborada tem por base a tabela de factos (*Fact_Use_Case*), e as diversas tabelas de ponte às quais está interligada (*Partners_BTF*, *Governances_BTF*, *Risks_BTF*, *Technologies_BTF*, *Maturities_BTF*, *Tools_BTF* e *Challenges_BTF*). De modo a simplificar, juntou-se nesta *view* a conjugação das tabelas de dimensão *Dim_What* e *Dim_Why* juntamente com as tabelas de ponte interligadas às mesmas (*Drivers_BDWhy*, *Benefits_BDWhy*, *Values_BDWhy*, *Concepts_BDWhat*, *Innovations_BDWhat* e *Lessons_Learnt_BDWhat*).

¹Uma *view* pode ser considerada uma tabela virtual/vista, representando uma tabela que fisicamente não existe. Uma *view* pode ser elaborada a partir de uma ou da conjugação de várias tabelas, ou até mesmo a partir de outras *views* pertencentes a uma base de dados.

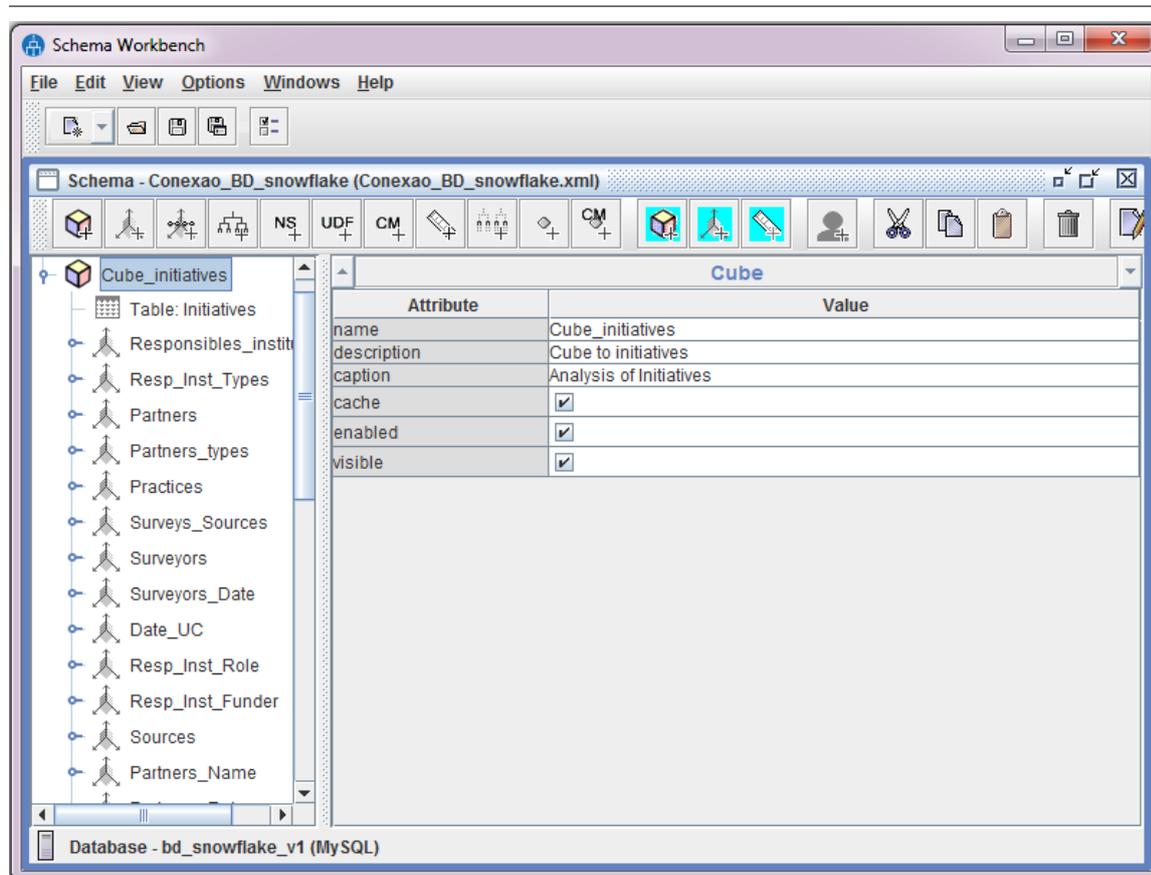


Figura 5.30: Esquema do cubo OLAP desenvolvido na ferramenta PSW

A medida pretendida para este projeto é o número de iniciativas e foi definida durante a elaboração do cubo. Para uma medida ser definida no cubo necessita de um nome, uma coluna da tabela de factos e o tipo de agregação (*aggregator*). O tipo de agregação pode ser: contagem distintiva (*distinct-count*), soma (*sum*), mínimo (*min*), média (*avg*), entre outros. Neste caso, para a definição da medida foi utilizado o nome *Number_initiatives*, selecionada a coluna respetiva à chave Primária (*id_Fact_Use_Case*) da tabela de factos (*Fact_Use_Case*) e definido o tipo de agregação como *distinct-count*, o qual permite a contagem do número de iniciativas.

5.2.3 Criação de Gráficos com a ferramenta analítica *Pentaho Community Edition*

Após a criação do esquema do cubo OLAP e a realização da publicação do mesmo no servidor *Pentaho BI*, procedeu-se à exploração dos dados provenientes

do preenchimento do formulário *online*, de forma a obter as consultas e análises pretendidas na elaboração desta plataforma.

O servidor BI da ferramenta *Pentaho CE* inclui a aplicação de *front-end Pentaho User Console (PUC)*. Esta aplicação PUC é uma *interface web* de interação com o utilizador que possibilita efetuar consultas, relatórios interativos, criar gráficos analíticos e *dashboards* a partir de cubos OLAP.

Inicialmente, foi utilizado o *plugin Saiku Analytics*, que permitiu, através da seleção do cubo OLAP elaborado, efetuar consultas aos dados e obter tabelas e diversos gráficos, necessitando apenas da seleção da medida *Number_initiatives* e das dimensões pretendidas. No entanto, este *plugin* não permite criar análises dinâmicas e nem possibilita a exportação das consultas e dos gráficos no formato pretendido, HTML. Contudo, o *Saiku Analytics* revelou-se importante, na medida que permitiu uma compreensão sobre a exploração dos dados e o funcionamento das consultas de forma mais intuitiva e clara.

Por conseguinte, procedeu-se à utilização de outro *plugin* na ferramenta PUC, denominado por CDE. Este *plugin* permite a criação, edição e visualização de vários *dashboards* dinâmicos, possibilitando exportá-los no formato HTML.

Tal como já foi referido na secção 3.2.5, o CDE é composto por três menus de configuração: *Layout*, *Components* e *Datasources*, que necessitam da aplicação de um conjunto de definições e configurações de modo a obter os gráficos pretendidos.

A imagem que se segue (figura 5.31) apresenta o ambiente de trabalho da ferramenta *Pentaho* com o *plugin CDE*, no menu de configuração *Layout*.

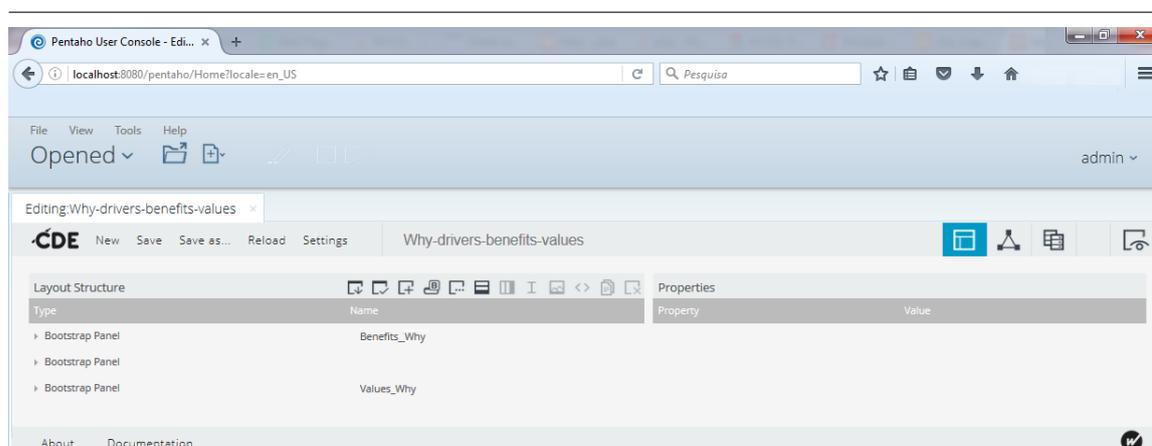


Figura 5.31: Ambiente de trabalho do *plugin CDE* - Menu de configuração *Layout*

A secção *Layout* é o primeiro menu de configuração e foi utilizada para definir, em cada *dashboard* elaborado, o seu *layout*, ou seja, a aparência do painel. Para isso, podem ser adicionados elementos HTML como texto, imagens, adicionar ficheiros CSS e/ou de *JavaScript*, e ainda, alguns painéis pertencentes à *framework Bootstrap*.

A imagem seguinte (figura 5.32) demonstra o ambiente de trabalho da ferramenta *Pentaho* com o *plugin CDE*, do menu de configuração *Components*.

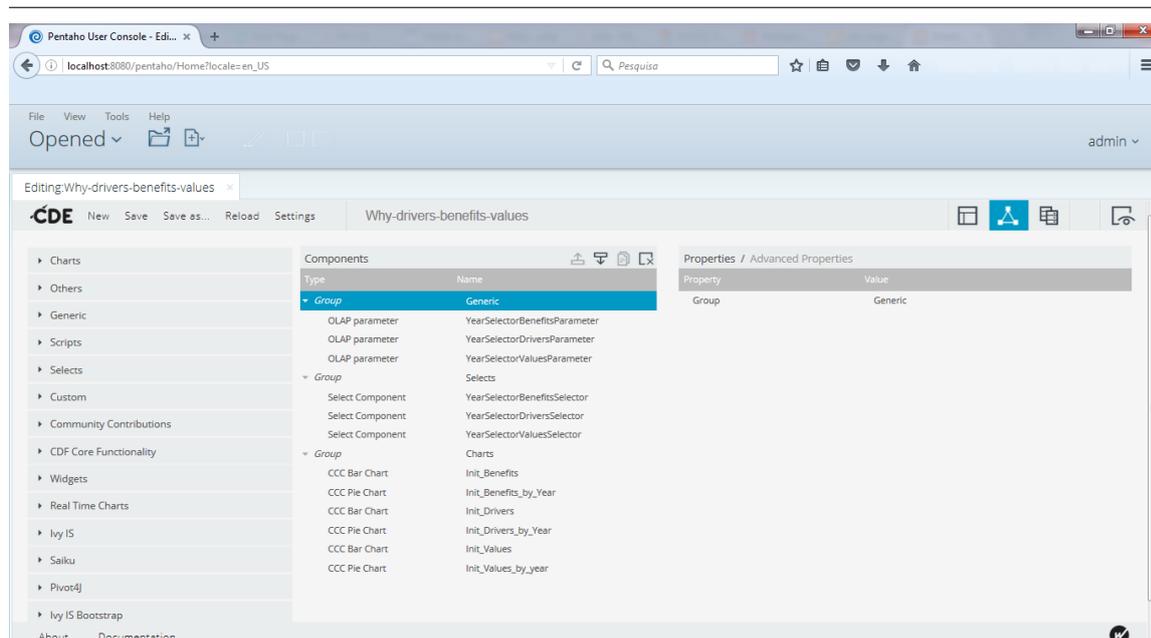


Figura 5.32: Ambiente de trabalho do *plugin CDE* - Menu de configuração *Components*

A secção *Components*, que constitui o segundo menu de configuração, foi utilizada para adicionar e configurar os diversos componentes visuais do *dashboard*, tais como: os elementos de texto, os elementos de seleção, as tabelas, os tipos de gráficos, entre outros. E além disso, este menu foi utilizado para definir parâmetros e *scripts* necessários à elaboração dos gráficos. De modo a tornar os componentes funcionais e visuais na *interface web*, é necessário associar aos mesmos um elemento HTML. Relativamente à definição de parâmetros, estes servem para que haja interatividade entre os gráficos implementados e, por exemplo, filtros de dados. Assim, no caso da aplicação de filtros em gráficos, quando o valor da caixa de seleção de um filtro é alterado, esse mesmo valor fica armazenado num parâmetro. Por sua vez, este parâmetro, também, encontra-se associado a um determinado gráfico. Deste modo,

quando ocorre uma mudança do valor da caixa de seleção significa uma nova consulta (*query*) e, portanto, uma atualização do gráfico.

A próxima imagem (figura 5.33) expõe o ambiente de trabalho da ferramenta *Pentaho* com o *plugin* CDE, no menu de configuração *Datasources*.

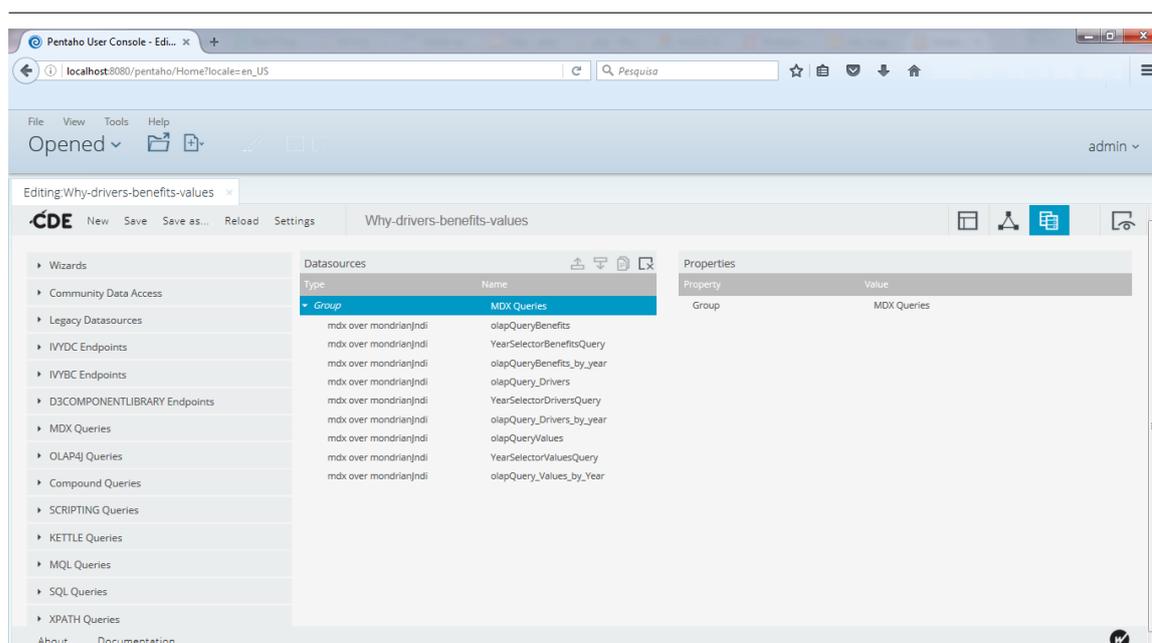


Figura 5.33: Ambiente de trabalho do *plugin* CDE - Menu de configuração *Datasources*

Na secção *Datasources*, que é o terceiro menu de configuração, procedeu-se à definição das fontes de dados utilizadas pelos diversos componentes. Nesta secção, para cada *dashboard* desenvolvido foram estabelecidas as diversas consultas MDX executadas sobre o cubo OLAP implementado.

5.3 Plataforma *Online*

O principal objetivo desta dissertação assentou na elaboração de uma plataforma *online* para a obtenção de informações e análises relativas às iniciativas em Cidades Inteligentes existentes por todo o mundo.

Nas etapas anteriores foram efetuados um conjunto de procedimentos, tais como: o estabelecimento dos objetivos, requisitos e restrições conforme as necessidades do sistema proposto; a análise e especificação desse sistema e dos serviços a disponibilizar; e a implementação do sistema de *Data Warehouse* e do Cubo OLAP.

Nesta etapa procedeu-se então, ao desenvolvimento e implementação da plataforma *online* e da aplicação *web* de administração.

Na figura 5.34, que se segue, é ilustrado o mapa da plataforma *online* desenvolvida, de modo a possibilitar uma compreensão mais intuitiva da estrutura e da interligação das diversas páginas *web* que constituem a plataforma.

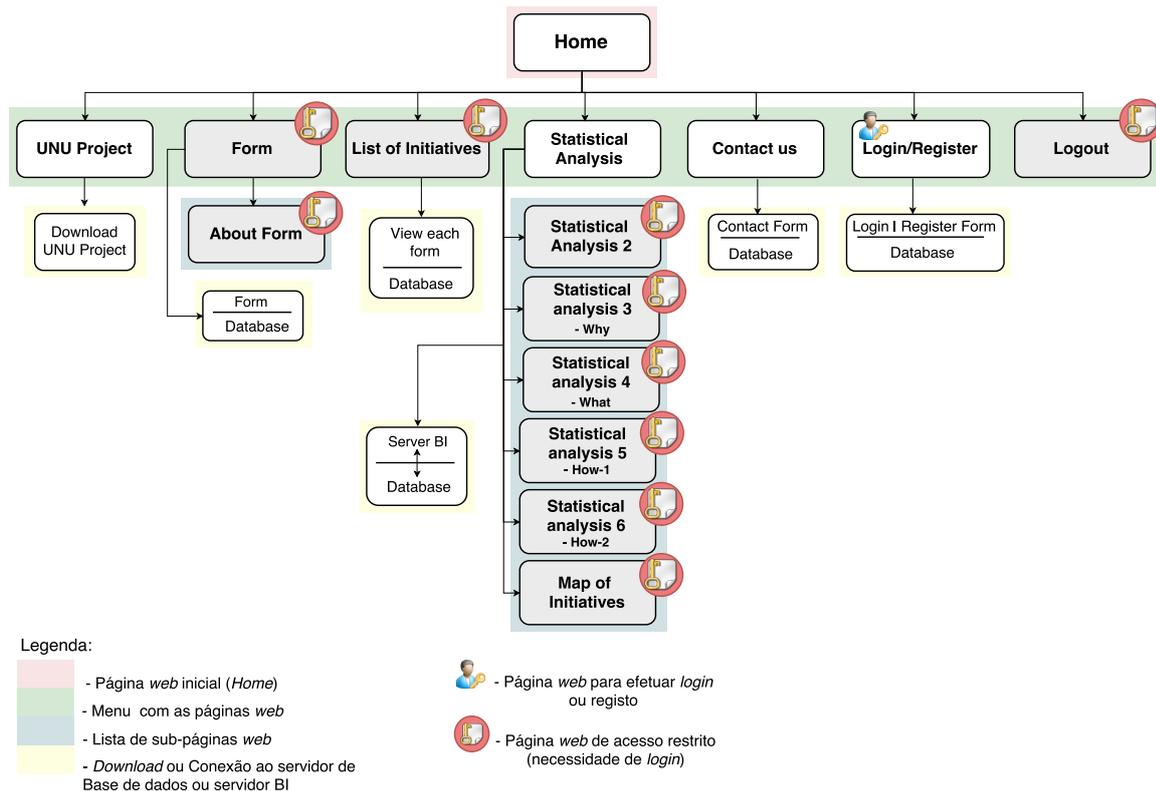


Figura 5.34: Mapa do *website* relativo à plataforma *online*

Tal como se pode verificar pela ilustração da figura 5.34, a plataforma *online* é destinada a dois tipos de utilizadores, os que não efetuaram o *login*, e os que se autenticaram no sistema através do *login*. Consoante o tipo de utilizador, é disponibilizado um determinado leque de funcionalidades, sendo que estas encontram-se acessíveis na totalidade para o tipo de utilizadores com *login* efetuado. As páginas *web* disponíveis para a generalidade dos utilizadores (utilizadores que não efetuaram *login*) estão mapeadas numa caixa de texto de cor branca. Por outro lado, as páginas *web* acessíveis apenas para utilizadores autenticados, estão mapeadas em caixas de texto de cor cinzenta, com um símbolo que caracteriza a necessidade de *login* para a visualização das mesmas.

Esta plataforma segue uma arquitetura cliente-servidor, no qual o servidor recebe solicitações HTTP (*Hypertext Transfer Protocol*) e devolve as respostas HTTP ao cliente. O HTTP é, portanto, um protocolo de transferência de hipertexto encarregue das solicitações de recursos e das respostas entre o cliente e o servidor em aplicações *web*. Os protocolos TCP (*Transmission Control Protocol*) e IP (*Internet Protocol*) são fundamentais para o estabelecimento de ligação entre o cliente e o servidor, de modo que seja possível a transferência de dados entre os intervenientes. O processo tem início quando o cliente HTTP solicita uma requisição, neste momento é estabelecida uma ligação TCP para uma porta específica do servidor HTTP. Este servidor encontra-se constantemente à escuta, na porta conectada, de uma mensagem de solicitação por parte do cliente (*web browser*). Assim que recebe um pedido, este é processado e, em seguida, o recurso solicitado é retornado.

A figura 5.35, que se segue, mostra o procedimento de uma solicitação HTTP de um cliente e da resposta do servidor *web*, ilustrado através de um diagrama de sequência.

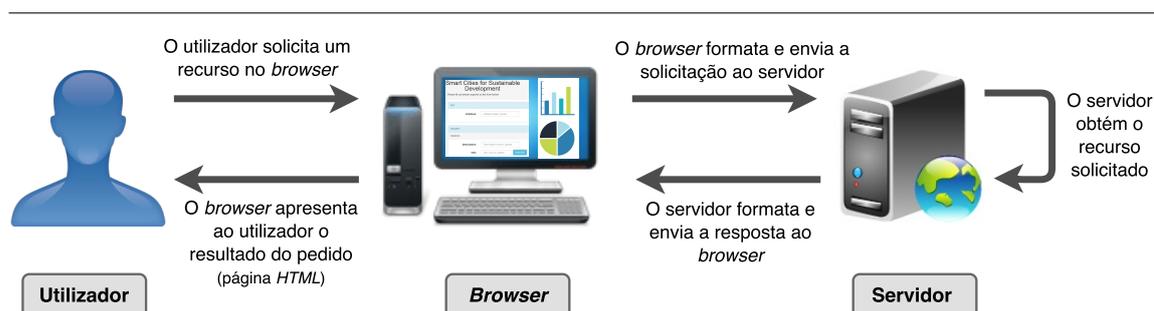


Figura 5.35: Esquema do procedimento de uma solicitação HTTP de um cliente e de uma resposta HTTP do servidor

Existem diversos métodos definidos no protocolo HTTP como *CONNECT*, *DELETE*, *GET*, *HEAD*, *OPTIONS*, *POST*, *PUT* e *TRACE*. Cada um destes métodos estabelece a ação que tem de ser efetuada perante o recurso solicitado. De entre os diversos métodos que este protocolo define, foram utilizados, para as requisições de recursos, os métodos *GET* e *POST*.

O método *GET* solicita que o recurso pretendido seja apresentado, devendo ser utilizado somente para casos de recuperação de dados, no qual os parâmetros são especificados na parte do endereço de URL do recurso pedido. As solicitações realizadas através deste método, geralmente, requisitam dados, não devendo ser utilizadas para

efetuar qualquer outro tipo de ação sobre os dados. Na implementação da plataforma *online* este método foi utilizado para a navegação entre as páginas *web*. Além disso, o método *GET* é utilizado aquando a seleção na listagem de alguma iniciativa para visualização e/ou edição dos dados.

O método *POST* é utilizado no envio de dados para o servidor. Neste caso, foram utilizados no envio dos dados referentes aos diversos formulários existentes no sistema implementado: formulário para a inserção das iniciativas, do registo de utilizadores, do *login*, e o formulário de inserção de novos administradores na aplicação *web* de administração.

De modo a obter segurança nos dados das *password* dos utilizadores e dos administradores foi utilizada a codificação/criptografia de *passwords* MD5. MD5 é um algoritmo que gera um *hash* de 128 bits unidirecional, o que significa que não é possível transformar um *hash* MD5 novamente na *password* que o ocasionou.

Os diversos formulários existentes, tanto na plataforma *online* como na aplicação *web* de administração, utilizam validação dos dados inseridos, tanto no lado cliente através de *JavaScript* como no lado do servidor por PHP. Do lado do cliente, a validação dos dados é realizada através do *plugin BootstrapValidator* que requer *JQuery* e *Bootstrap 3*. Este *plugin* de validação efetua a verificação de cada campo preenchido. Desta forma, é exibida uma mensagem de confirmação de sucesso nos campos preenchidos corretamente, e uma mensagem de erro nos campos preenchidos incorretamente ou nos campos não preenchidos, caso esse campo seja de preenchimento obrigatório. A validação de cada campo de entrada é efetuada via *JavaScript*, ou seja, a verificação do correto funcionamento de cada campo não leva a que seja necessário recarregar a página *web* por completo. Além disso, o botão de submeter fica desativado até se verificar que todos os campos obrigatórios estão preenchidos e de forma válida. Do lado do servidor, a validação é realizada através de PHP, no qual se verifica se a variável existe e/ou está vazia ou não.

5.3.1 Formulário *online*: “*Smart Cities for Sustainable Development*”

O desenvolvimento e a implementação do formulário *online* teve por base o documento referente ao projeto “*Smart Cities for Sustainable Development*”, referido anteriormente na secção 2.1.3 e disponibilizado pela UNU-EGOV.

O formulário *online* permite automatizar a inserção de casos de estudo e inici-

ativas em Cidades Inteligentes no sistema de *Data Warehouse*. O sistema de *Data Warehouse* armazena numa primeira fase a informação relativa aos 21 casos de estudo disponibilizados nesse documento. No entanto, com a disponibilização da plataforma *online* é esperado a adição de novas informações acerca de outras iniciativas.

O formulário *online* é composto por vários elementos HTML, sendo estes: (i) elementos *input* para permitir entradas do tipo texto (*text*) ou do tipo rádio (*radio*); (ii) elementos *select* para a criação de listas de seleção; e (iii) caixas de seleção (*checkbox*); (iv) botões e respetivas entradas de texto para adicionar a determinadas listas de seleção; e (v) botões de adicionar/remover determinados campos.

O elemento de formulário de entrada (*input*) do tipo texto (`<input type="text">`) foi utilizado para o utilizador inserir informações como o nome da iniciativa (*Initiative*), e diversas descrições e comentários, como nos campos *Highlights* e *Comments*.

A imagem que se segue (figura 5.36) demonstra um exemplo, implementado, da utilização do elemento de entrada do tipo texto para o caso dos campos *Highlights* e *Comments*, extraído do formulário *online*.

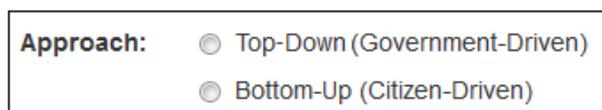


The image shows a rectangular box containing two text input fields. The first field is labeled "Highlights:" and contains the text "Highlights, please". The second field is labeled "Comments:" and contains the text "Comments, please".

Figura 5.36: Exemplo da utilização do elemento de entrada de texto

O elemento *input* do tipo rádio (`<input type="radio">`) foi utilizado para definir botões de rádio para a seleção de uma opção de um número limitado de opções, como o caso do atributo *Approach*.

Na imagem que se segue (figura 5.37) é apresentado um exemplo, implementado, da utilização do elemento de entrada do tipo rádio para o caso do campo *Approach*, retirado do formulário *online*.



The image shows a rectangular box containing a label "Approach:" followed by two radio button options. The first option is "Top-Down (Government-Driven)" and is selected. The second option is "Bottom-Up (Citizen-Driven)" and is unselected.

Figura 5.37: Exemplo da utilização do elemento de entrada do tipo rádio

O elemento *input* do tipo caixa de seleção (`<input type="checkbox">`) foi utilizado nas diversas caixas de seleção existentes no formulário. Estas caixas de seleção permitem ao utilizador selecionar mais que uma opção de um número limitado de opções, como por exemplo no caso dos atributos *Risks* e *Benefits*.

A imagem que se segue (figura 5.38) retrata um exemplo, implementado, da utilização do elemento de entrada do tipo caixa de seleção para o caso do campo *Benefits*, extraído do formulário *online*.



The image shows a rectangular box containing the label 'Benefits:' followed by a vertical list of eight items, each with an unchecked checkbox to its left. The items are: City Planning, New and Enhanced Digital Services, Environment Protection, Human Development, Infrastructure for Public Services, Governance and Participation, Economic Growth, and Improving Quality of Life.

Figura 5.38: Exemplo da utilização do elemento de entrada do tipo caixa de seleção

O elemento *select* foi utilizado em situações como selecionar o mês (*Month*) e o ano (*Year*) da pesquisa da iniciativa, o nome da instituição responsável (*Responsible institution*), o nome dos parceiros (*Partners*), a localização, entre outros.

Na figura 5.39 é apresentado um exemplo, implementado, da utilização do elemento *select* para o caso dos campos *Month* e *Year*, retirado do formulário *online*.



The image shows two dropdown menus. The first is labeled 'Month:' and has a 'Select' option visible. The second is labeled 'Year:' and has a 'Select' option visible, with a dropdown menu open showing a list of years from 2016 down to 2007.

Figura 5.39: Exemplo da utilização do elemento *select*

Para o envio de dados para o servidor foi utilizado, como referido anteriormente, o método *POST*, através do atributo *method* (*method="post"*) que serve para especificar o método HTTP utilizado no momento da submissão do formulário.

De modo a obter uma lista o mais completa possível com os continentes, países, estados/províncias, municípios/regiões e cidades relativos às hierarquias da dimensão *Dim_cities* foi utilizado um serviço *web open source* denominado por *Geonames*. Este serviço facilita o preenchimento do formulário acerca da localização de cada iniciativa. A partir deste serviço é possível obter as informações sobre cada local como uma estrutura em árvore, no qual na hierarquia superior se encontra a lista de continentes. Desta forma, assim que se seleciona um continente, aparecerá a lista de países referentes a esse continente e assim sucessivamente, até ser obtido a lista das cidades, sendo este o último da hierarquia. No caso da localização da iniciativa não se encontrar incluída na lista, é possível ao utilizador inserir através de um elemento de entrada (*input*) de texto.

A imagem que se segue (figura 5.40) retrata a secção do formulário *online* destinado à introdução da localização da iniciativa, como referido anteriormente.



O formulário apresenta cinco campos de seleção hierárquica, cada um com um ícone de seta para baixo:

- Continent:** com o texto "Select One" e uma seta para baixo.
- Country:** com uma seta para baixo.
- State/Province:** com uma seta para baixo.
- County/Region:** com uma seta para baixo.
- City:** com uma seta para baixo.

Figura 5.40: Elementos para inserção da localização da iniciativa

5.3.2 Análises Estatísticas

A plataforma *online* disponibiliza gráficos e análises sobre os dados mais relevantes obtidos a partir do formulário *online*. Para tal, foram desenvolvidos *dashboards* dos dados e mapas referentes à localização das iniciativas em Cidades Inteligentes, resultantes da inserção da informação nessa plataforma. A elaboração deste conjunto de análises foi obtido através do recurso da tecnologia OLAP, incluída na ferramenta CDE, tal como se descreveu na secção 5.2.3.

As principais análises apresentadas são constantemente atualizadas à medida que novos casos de estudo/iniciativas são inseridos. Estas análises permitirão analisar a evolução dos vários aspetos dos principais domínios referentes a uma Cidade Inteligente. Exemplo dessas análises são o número de iniciativas por: ano; tipo de Conceito (*Concept*); Instituição Responsável (*Responsible Institution*); tipo de Desafios (*Challenges*); tipo de Riscos (*Risks*); tipo de Benefícios (*Benefits*), entre outras.

Desta forma, esta plataforma *online* constitui um instrumento útil para promover novas iniciativas que viabilizem o desenvolvimento sustentável. É com base no conhecimento partilhado, com experiências e resultados obtidos noutras iniciativas, na inovação, na criatividade e no apoio económico que é possível atualmente o desenvolvimento de iniciativas em Cidades Inteligentes.

5.4 Aplicação *web* de Administração

Após o desenvolvimento e implementação da plataforma *online* para casos de estudo em Cidades Inteligentes, procedeu-se à elaboração de uma aplicação *web* destinada a administrador(es) desta plataforma. Esta aplicação *web* destina-se, sobretudo, a gerir as autorizações de *login* dos utilizadores, e ainda, à alteração de dados dos formulários a pedido dos utilizadores, ou mesmo à eliminação de formulários.

A figura 5.41 que se segue, ilustra o mapa da aplicação *web* desenvolvida para o(s) administrador(es).

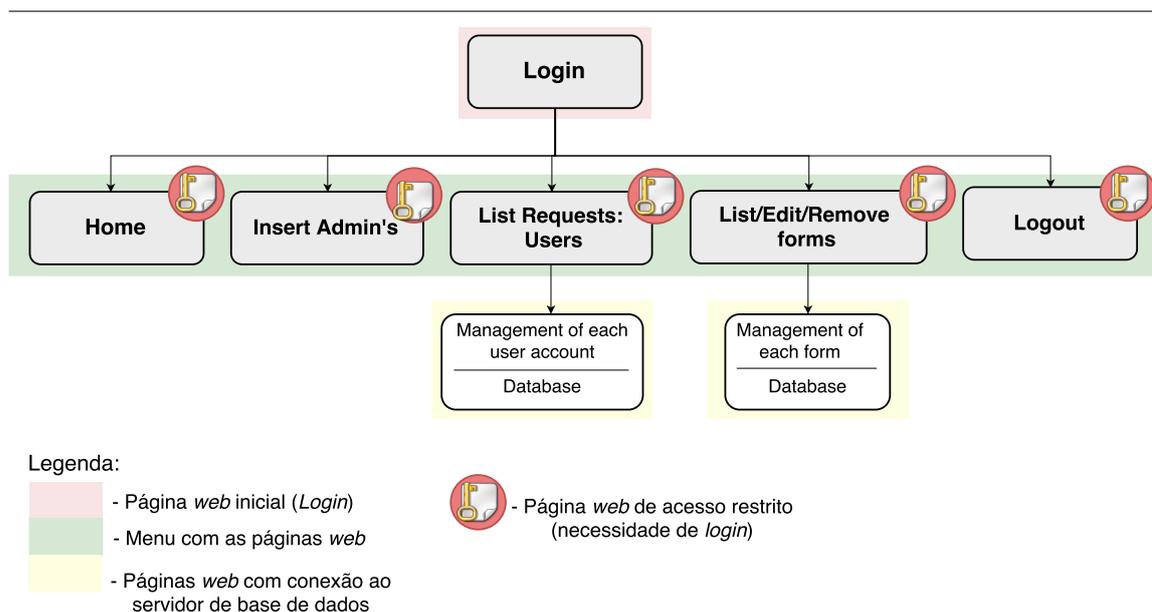


Figura 5.41: Mapa do *website* relativo à aplicação *web* de administração

Tal como se pode observar na figura 5.41, é necessário efetuar *login* de modo a poder efetuar a administração destinada à plataforma *online*. O administrador pode listar todos os pedidos dos utilizadores para possuírem autorização de efetuar *login* na plataforma. Aquando a aceitação do pedido por parte do administrador, o utilizador receberá de imediato uma informação através de *e-mail* com a indicação dessa permissão; o mesmo acontece com a desativação da autorização. Nesta listagem é exibida, também, a indicação de que o utilizador confirmou o seu registo/*e-mail*, assim, o administrador pode confirmar a veracidade dos pedidos. Além disso, o administrador principal pode inserir, neste sistema de administração, mais colaboradores, ou seja, utilizadores com funções de administrar.

O(s) administrador(s) pode(m) a pedido dos utilizadores, e em concordância entre ambos, efetuar alterações aos dados de cada formulário, bem como proceder à eliminação dos mesmos.

Esta aplicação *web* segue, tal como a plataforma *online*, uma arquitetura cliente-servidor, em que o cliente envia solicitações HTTP ao servidor, e este processa-os e devolve as respostas.

5.5 Conclusões

Este capítulo destinou-se a apresentar um conjunto de ferramentas, tecnologias e linguagens de programação utilizadas, bem como todo o processamento da implementação do sistema proposto.

Inicialmente, nas primeiras secções foram descritas as ferramentas *open source*, as tecnologias e linguagens de programação utilizadas na elaboração das análises OLAP, na elaboração da plataforma *online* e da aplicação *web* de administração.

Em seguida, foram descritas a modelação dos dados e a elaboração do sistema de *Data Warehouse*. Este sistema permite armazenar, reunir, e aceder aos dados relacionados com as diversas iniciativas em Cidades Inteligentes.

O modelo de dados implementado para o *Data Warehouse* foi o modelo de dados multidimensional. Este modelo permite através da criação, demonstrada também, do cubo OLAP, explorar e obter dados através de diversas perspetivas. Além disso, o modelo de dados multidimensional foi desenvolvido com base no conceito do esquema em Floco de Neve.

O cubo OLAP fornece suporte ao sistema OLAP, que permite obter análises rápidas e eficientes dos dados. Neste trabalho, foram elaboradas diversas análises estatísticas relevantes, que permitem adquirir conhecimento útil, para o auxílio de tomadas de decisão. Estas análises servirão de apoio aos gestores e decisores políticos que pretendam o melhoramento de iniciativas em Cidades Inteligentes ou, mesmo, o desenvolvimento de novas iniciativas.

Nas seguintes secções procedeu-se à especificação do modo como se realizou a exploração dos dados através da aplicação *front-end* PUC da ferramenta *Pentaho*, conhecida, também, por *BI Server*. Esta ferramenta de *interface web* permitiu a elaboração e exportação dos diversos gráficos e análises para a plataforma *online*. Além disso, esta ferramenta possibilitou o desenvolvimento de um mapa geográfico, que ilustra as localizações dos casos de estudo e iniciativas em Cidades Inteligentes.

A exploração da ferramenta *Pentaho* foi resultado da seleção de uma ferramenta *Business Intelligence open source*, a partir de um estudo de comparação entre algumas ferramentas deste tipo, referido no capítulo 3. Uma vez que a utilização desta ferramenta resultou de uma seleção a partir de um estudo, sendo completamente desconhecida, foi fundamental dedicar bastante tempo para a exploração e compreensão dos componentes utilizados pertencentes à *suite Pentaho*. Ao longo da pesquisa para a perceção sobre o funcionamento da ferramenta, por vezes se deparava

com informação disponibilizada não atualizada, dado que esta ferramenta *open source* está em constantes evoluções e melhoramentos pela Comunidade *Pentaho*. Isto, incitou, então, a uma exploração algo demorada sobre o propósito, o funcionamento e a integração dos diversos componentes disponibilizados pela ferramenta *Pentaho*. A configuração e instalação do servidor BI do *Pentaho*, bem como a migração para o SGBD *MySQL*, foi algo minucioso, que exigiu um conjunto de etapas de caráter rigoroso, pois caso contrário comprometeria o correto funcionamento da ferramenta. Além disso, a informação e documentação técnica de suporte disponível sobre o funcionamento de cada componente dos menus de configuração (*Layout*, *Components* e *Datasource*) do *plugin* CDE, que permitiu a criação de *dashboards*, é relativamente escassa e incompleta por ainda consistir num *plugin* recente.

No entanto, toda a exploração desta ferramenta foi enriquecedora, permitindo obter gráficos e mapas que possibilitarão, aos gestores e entidades interessadas por iniciativas em Cidades Inteligentes, tomadas de decisão eficientes. Além do mais, esta ferramenta, disponibilizada na versão *open source*, é bastante completa e fornece os vários componentes essenciais para o desenvolvimento de um projeto de *Business Intelligence*.

Com base em cinco questões foi desenvolvido um formulário *online* de modo a obter informações sobre as iniciativas. Algumas dessas informações são sobre: as instituições, os parceiros, os objetivos e principais características, os impactos sobre as dimensões sociais, económicas ou ambientais, a localização e o período de tempo relativamente a cada prática. Estes dados adquiridos com o formulário permitem sustentar os gráficos e mapas disponibilizados na plataforma. A plataforma centra-se num sistema OLAP, sendo este um dos componentes de *Business Intelligence*, o qual se tem tornado cada mais importante e amplamente utilizado pelas organizações para tomadas de decisão. A análise OLAP constituiu uma fase fundamental no desenvolvimento deste sistema, permitindo a exploração, análise e cálculo de diversos dados levando à criação de vários gráficos e de um mapa (*GeoMap*) do tipo *Google Maps*, que estão disponibilizados na plataforma *online*.

A implementação da plataforma *online* será uma mais valia para a promoção de iniciativas em Cidades Inteligentes. Esta plataforma consistirá num instrumento significativo para todos os que estão relacionados com este tipo de iniciativas, disponibilizando elementos e informações relevantes sobre esta temática.

A criação desta plataforma permitirá, assim, auxiliar em estratégias de modo a dinamizar iniciativas existentes em Cidades Inteligentes e, ainda, a incentivar novas

práticas, de modo a aumentar o desenvolvimento sustentável. A conceção da aplicação *web* para administração permitirá, além da inserção de administradores, a gestão das permissões de autenticação dos utilizadores na plataforma, e a edição/remoção de formulários.

Capítulo 6

Resultados do sistema

Neste capítulo são apresentados os principais resultados obtidos com a elaboração da plataforma *online* para casos de estudo e iniciativas em Cidades Inteligentes, bem como da aplicação *web* de administração. Inicialmente, são demonstrados os principais resultados referentes à plataforma *online*, nomeadamente o formulário, os diversos *dashboards* e um mapa obtidos através da tecnologia OLAP (secção 6.1). Por fim, são expostos os resultados relativos à aplicação *web* de administração (secção 6.2). Esta aplicação permite alterar as permissões dos utilizadores relativas à realização de *login* na plataforma *online* e, possibilita, também, a edição/remoção de formulários.

6.1 Plataforma *Online*

A plataforma *online* proposta para iniciativas em Cidades Inteligentes será uma mais valia para todos os que estão relacionados com esta matéria, pois permite reunir e disponibilizar um conjunto de informações devidamente estruturadas. O formulário *online* é o meio utilizado para recolher os dados sobre as iniciativas, e, também, constitui a fonte de dados do sistema de *Data Warehouse*.

Esta informação captada por cada formulário, preenchido e enviado, será disponibilizada através de uma secção na plataforma destinada à listagem de iniciativas, e de análises estatísticas aos dados mais relevantes, fornecidas numa secção para o efeito. As análises correspondem a um conjunto de vários *dashboards* e mapas geográficos elaborados por técnicas OLAP. O sistema OLAP constitui um dos componentes do sistema de *Business Intelligence*, que tem vindo a destacar-se, cada vez mais, no setor público e empresarial, como forma de extração, monitorização e análise de dados para

apoio a tomadas de decisão.

Desta forma, a plataforma *online* fornece, aos gestores e decisores políticos associados a Cidades Inteligentes, um método de apoio na determinação de estratégias para o desenvolvimento sustentável, a partir de iniciativas desta natureza. A disponibilização de gráficos analíticos, assentes nos principais dados obtidos das iniciativas, permite analisar a evolução de diversos atributos nos principais domínios referentes a uma Cidade Inteligente. Assim, a plataforma serve de apoio ao aperfeiçoamento de iniciativas existentes, promovendo, também, a execução de novas iniciativas.

A plataforma *online* inicialmente, na sua página *web* de início, denominada por *Home*, apresenta uma breve descrição sobre o âmbito da mesma, indicando ao utilizador que para usufruir da plataforma na totalidade deverá proceder ao registo, sendo necessário o administrador autorizar a realização do *login*.

A imagem seguinte (figura 6.1) mostra a página *web* inicial denominada por *Home*.

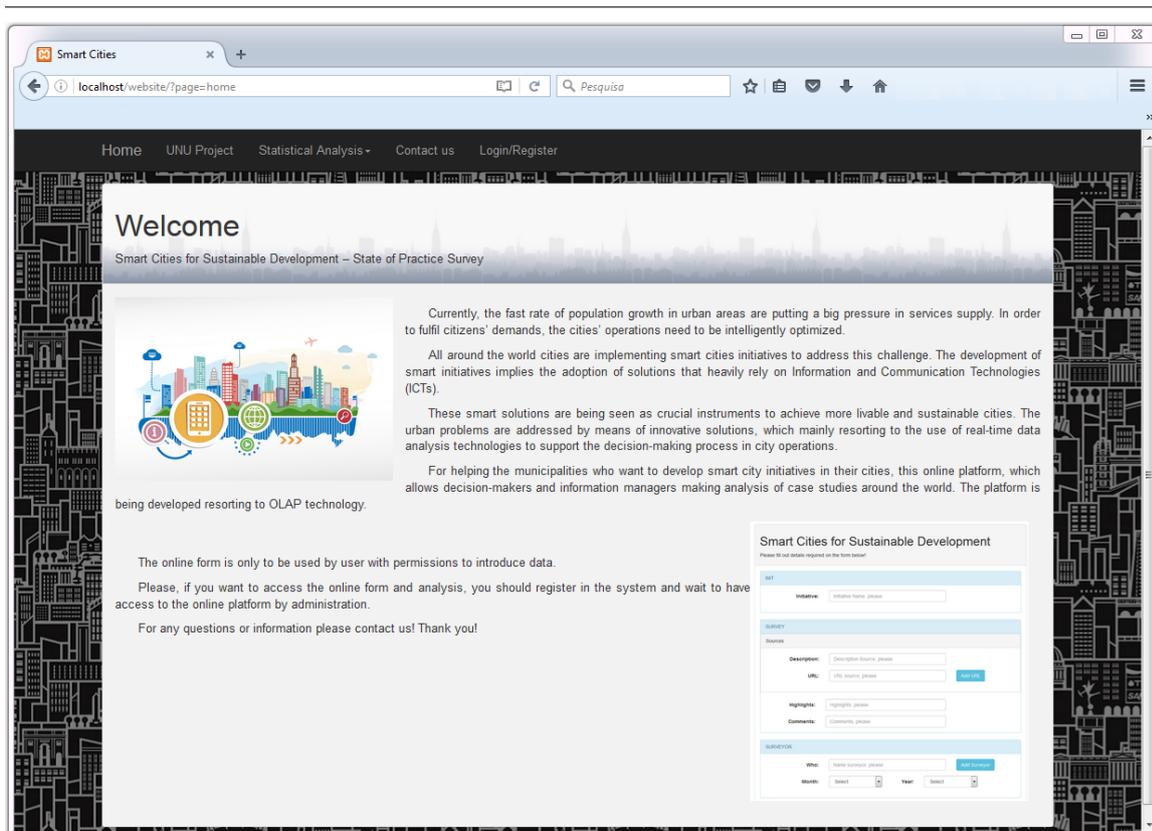


Figura 6.1: Demonstração da página *web* inicial - *Home*

Em seguida, a plataforma *online* apresenta sucintamente o projeto relacionado com Cidades Inteligentes denominado por “*Smart Cities for Sustainable Development*” referenciado na secção 2.1.3. Nesta página *web* denominada por *UNU Project*, o utilizador poderá efetuar o *download* do documento que descreve todo esse projeto, dando a conhecer de forma mais detalhada todos os campos relativos ao formulário *online*.

Na seguinte imagem (figura 6.2) é possível visualizar a página *web* denominada por *UNU Project*.

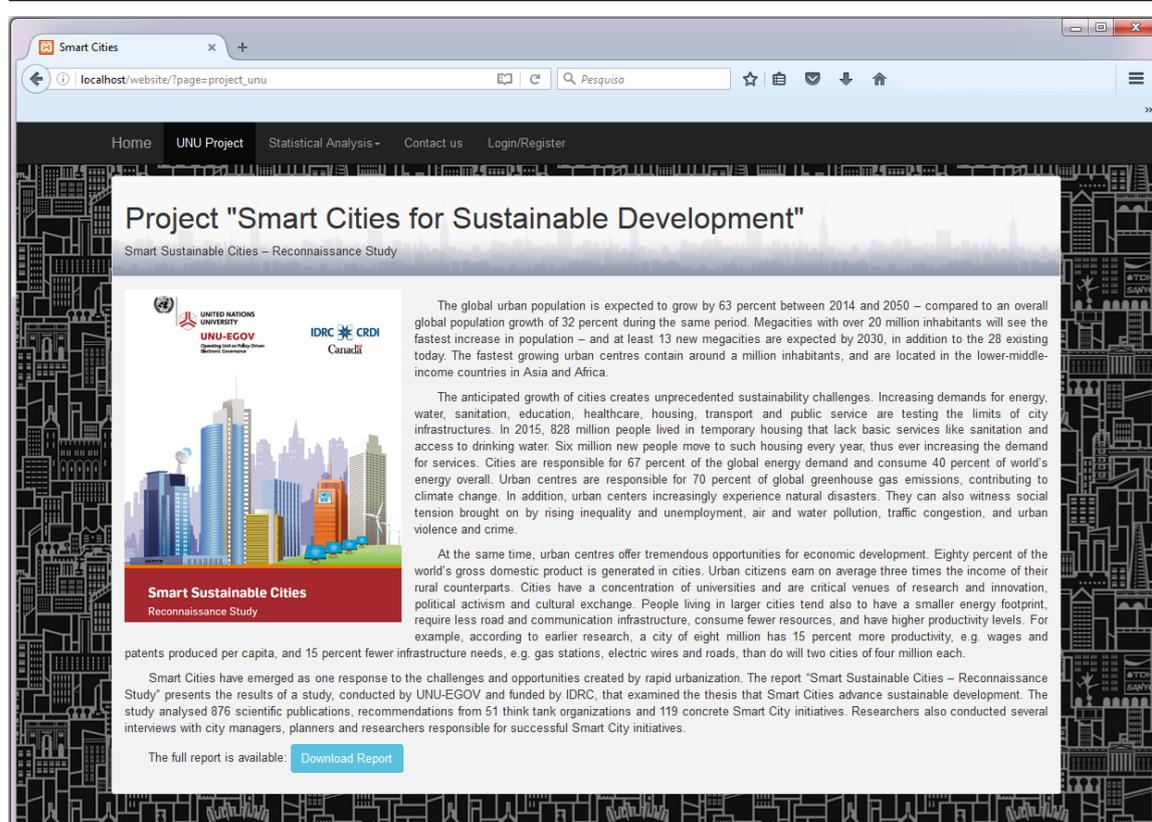
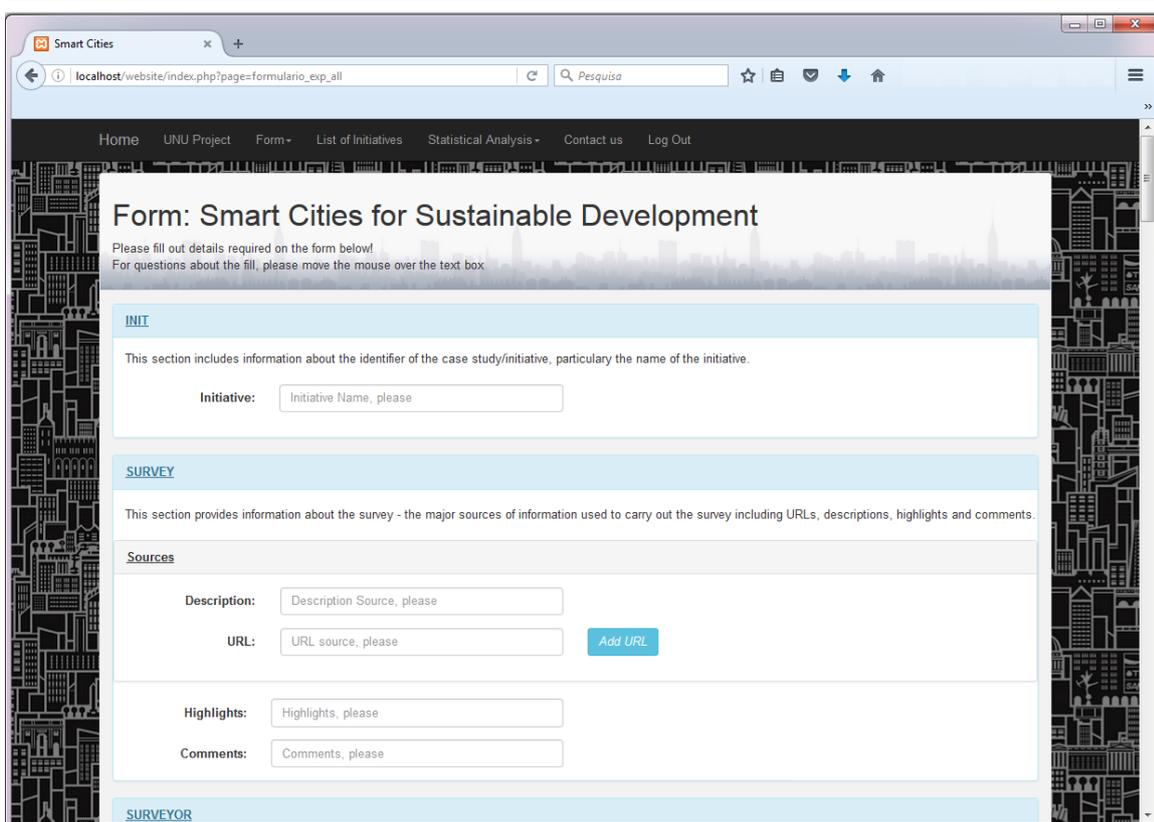


Figura 6.2: Demonstração da página *web* *UNU Project*

Os utilizadores com *login* devidamente efetuado, têm a possibilidade de preencherem o formulário *online* destinado às iniciativas em Cidades Inteligentes. Este formulário é composto por diversos elementos (atributos) relacionados com as diversas dimensões associadas a Cidade Inteligente, tal como a dimensão social, económica e ambiental. Este formulário visa captar informações, nomeadamente sobre: a pesquisa e os investigadores; a localização e o ano de início de desenvolvimento; os principais

riscos, fatores impulsionadores, mudanças, desafios, valores, ferramentas tecnológicas entre outros. Estes atributos estão associados às cinco questões: “*Quem?*”, “*Onde?*”, “*O quê?*”, “*Porquê?*” e “*Como?*” já mencionadas no capítulo anterior (capítulo 5).

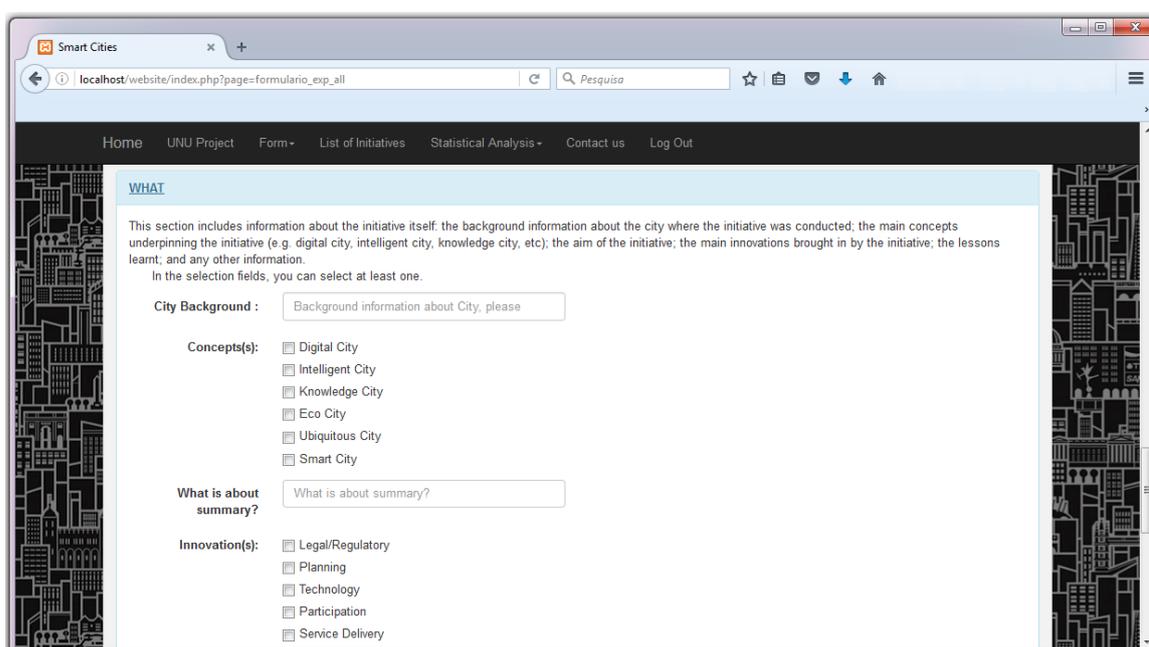
A figura que se segue (figura 6.3) apresenta a parte inicial do formulário, no qual se podem observar os primeiros campos a serem preenchidos.



The image shows a web browser window displaying a form titled "Form: Smart Cities for Sustainable Development". The browser's address bar shows the URL "localhost/website/index.php?page=formulario_exp_all". The form is divided into several sections. The first section is labeled "INIT" and contains a text input field for "Initiative Name, please". The second section is labeled "SURVEY" and contains a "Sources" section with three text input fields: "Description Source, please", "URL source, please", and "Highlights, please". There is also a "Comments, please" field. A blue button labeled "Add URL" is positioned next to the "URL source, please" field. The form is set against a background of a city skyline.

Figura 6.3: Demonstração da página *web Form* - Parte inicial

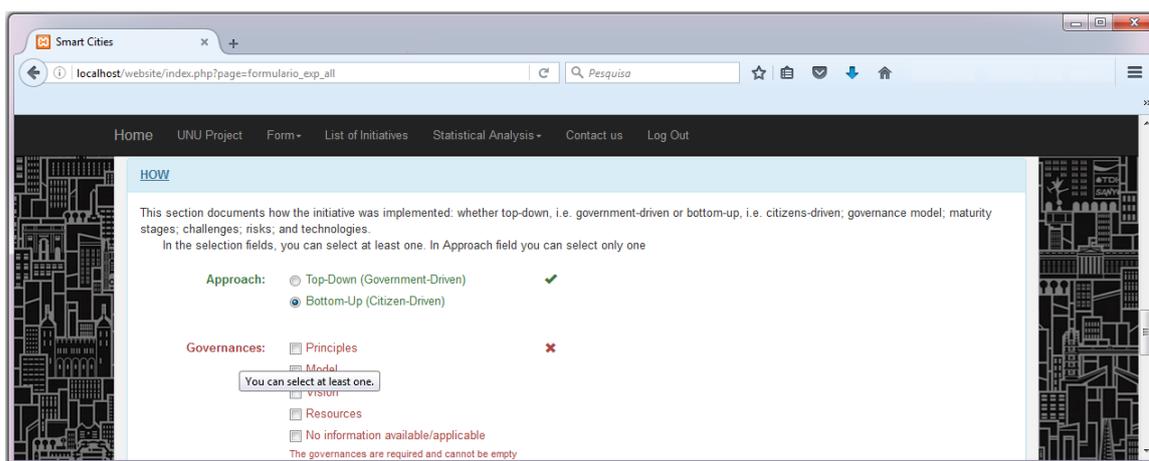
A próxima imagem (figura 6.4) retrata outra secção de campos a serem preenchidos no formulário da plataforma *online*, no qual se pode observar os atributos pertencentes à questão “*What?*”.



The screenshot shows a web browser window with the URL `localhost/website/index.php?page=formulario_exp_all`. The page title is "Smart Cities". The navigation menu includes "Home", "UNU Project", "Form", "List of Initiatives", "Statistical Analysis", "Contact us", and "Log Out". The main content area is titled "WHAT" and contains the following text: "This section includes information about the initiative itself: the background information about the city where the initiative was conducted; the main concepts underpinning the initiative (e.g. digital city, intelligent city, knowledge city, etc), the aim of the initiative, the main innovations brought in by the initiative, the lessons learnt, and any other information. In the selection fields, you can select at least one." Below this text are several form fields: "City Background:" with a text input field containing "Background information about City, please"; "Concepts(s):" with a list of checkboxes for "Digital City", "Intelligent City", "Knowledge City", "Eco City", "Ubiquitous City", and "Smart City"; "What is about summary?" with a text input field containing "What is about summary?"; and "Innovation(s):" with a list of checkboxes for "Legal/Regulatory", "Planning", "Technology", "Participation", and "Service Delivery".

Figura 6.4: Demonstração da página *web Form* - questão “What”

A figura 6.5, que se segue, mostra a validação durante o preenchimento do formulário *online*, no qual que pode observar casos de preenchimento correto e incorreto dos dados. Além disso, também é possível visualizar, nesta imagem, uma dica sobre como se pode preencher determinado campo.



The screenshot shows the same web browser window, but the page title is "Smart Cities" and the URL is `localhost/website/index.php?page=formulario_exp_all`. The navigation menu is the same. The main content area is titled "HOW" and contains the following text: "This section documents how the initiative was implemented: whether top-down, i.e. government-driven or bottom-up, i.e. citizens-driven; governance model; maturity stages; challenges; risks; and technologies. In the selection fields, you can select at least one. In Approach field you can select only one". Below this text are several form fields: "Approach:" with two radio buttons, "Top-Down (Government-Driven)" and "Bottom-Up (Citizen-Driven)", where "Bottom-Up (Citizen-Driven)" is selected and has a green checkmark; "Governances:" with a list of checkboxes for "Principles", "Model", "Vision", "Resources", and "No information available/applicable", where "Principles" is selected and has a red 'x' mark. Below the "Governances:" field, there is a text box containing "You can select at least one." and a red error message: "The governances are required and cannot be empty".

Figura 6.5: Demonstração da página *web Form* - Validação

Nesta secção existe, também, uma página *web* para um esclarecimento mais

detalhado sobre os diversos aspetos do formulário, denominada por *About Form*, tal como se pode observar na figura 6.6.

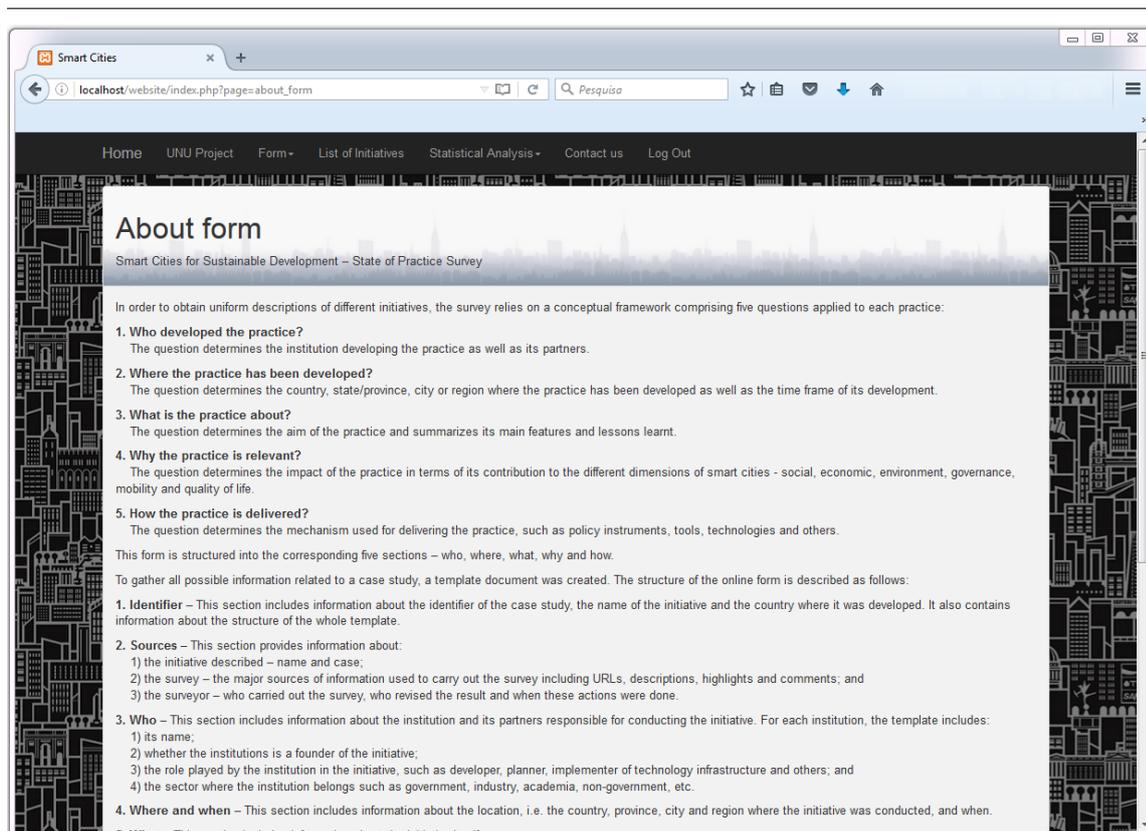
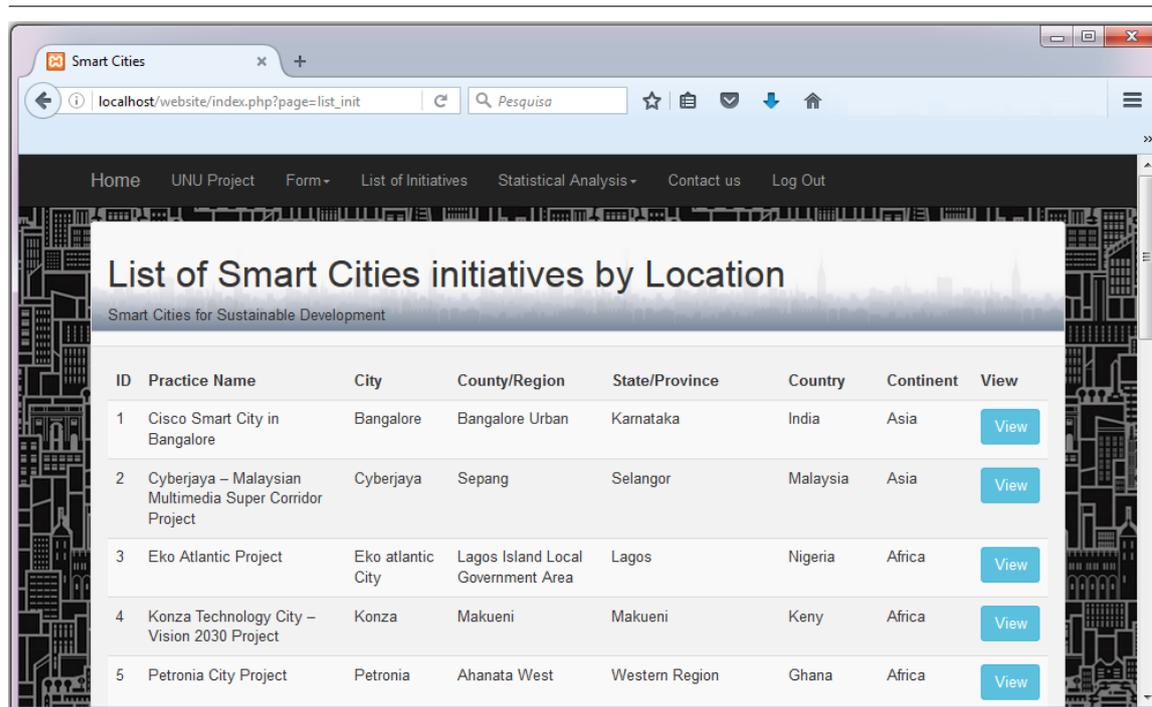


Figura 6.6: Demonstração da página *web About form*

A plataforma *online* apresenta, inclusive, uma página *web* destinada à listagem dos diversos casos de estudo/iniciativas armazenados na base de dados. Esta listagem possui o nome da iniciativa e a informação relativa à localização da mesma. Ao clicar no botão de visualização denominado por *View* o utilizador pode observar, com detalhe, o preenchimento do *formulário* referente à iniciativa selecionada. Esta página *web* encontra-se na secção *List of Initiatives* e está estabelecida apenas para utilizadores com *login* efetuado corretamente.

A próxima imagem (figura 6.7) apresenta a página *web List of Initiatives*, que permite a listagem e visualização do caso de estudo/iniciativa pretendido(a) em detalhe.



Smart Cities

localhost/website/index.php?page=list_init

Home UNU Project Form - List of Initiatives Statistical Analysis - Contact us Log Out

List of Smart Cities initiatives by Location

Smart Cities for Sustainable Development

ID	Practice Name	City	County/Region	State/Province	Country	Continent	View
1	Cisco Smart City in Bangalore	Bangalore	Bangalore Urban	Karnataka	India	Asia	View
2	Cyberjaya – Malaysian Multimedia Super Corridor Project	Cyberjaya	Sepang	Selangor	Malaysia	Asia	View
3	Eko Atlantic Project	Eko atlantic City	Lagos Island Local Government Area	Lagos	Nigeria	Africa	View
4	Konza Technology City – Vision 2030 Project	Konza	Makueni	Makueni	Kenya	Africa	View
5	Petronia City Project	Petronia	Ahanata West	Western Region	Ghana	Africa	View

Figura 6.7: Demonstração da página *web List of Initiatives*

Posteriormente, a plataforma apresenta uma secção para a demonstração das análises estatísticas titulada de *Statistical Analysis*. Esta secção expõe os diversos *dashboards* elaborados na aplicação *front-end* PUC do *Pentaho*. Os *dashboards* desenvolvidos referem-se a dados relativos aos atributos da questão “*Why?*” (*Benefits, Drivers e Values*), “*What?*” (*Concepts, Innovations e Lessons Learnt*) e “*How?*” (*Challenges, Risks, Tools, Governance, Maturities, Technologies e Approaches*). E ainda, existem *dashboards* relativos ao número de iniciativas por ano e por tipo de cada instituição responsável (*responsible institution*). Contudo, os utilizadores não registados ou sem realização de *login* podem visualizar um dos *dashboards* elaborados. Somente, os utilizadores, registados e autorizados, com *login* efetuado é que possuem acesso a todos os *dashboards* e, ainda, a um mapa geográfico com a indicação da localização das Cidades Inteligentes inseridas na base de dados.

A próxima imagem (figura 6.8) demonstra um dos vários *dashboards* desenvolvidos. Nesta imagem pode-se observar a categoria *Benefits* (Benefícios), representada pelo número de iniciativas por cada tipo e por cada ano. O utilizador poderá seleccionar o ano pretendido e analisar o número de iniciativas por cada tipo de *Benefits* nesse ano.

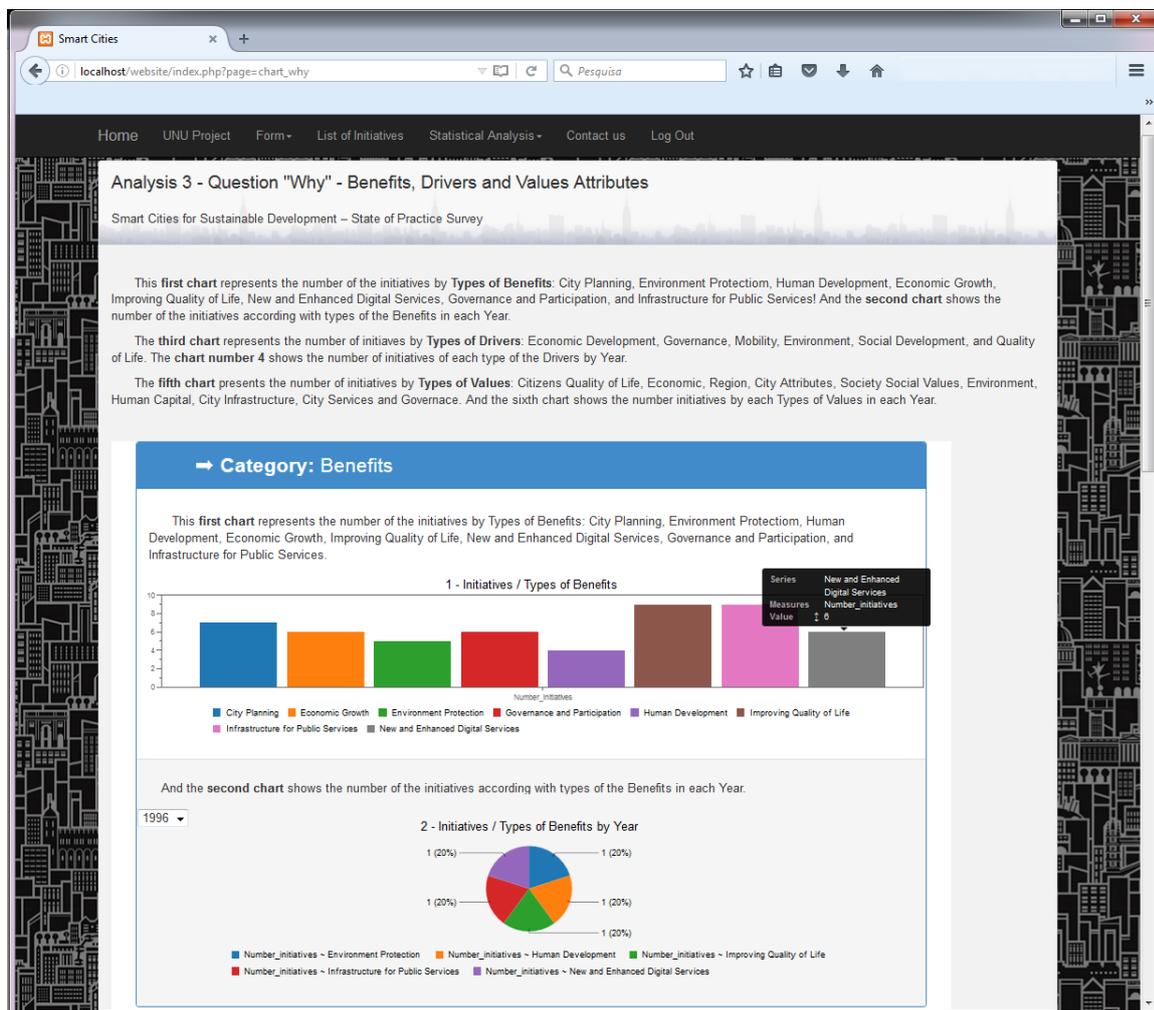


Figura 6.8: Demonstração da página *web Statistical Analysis 3 - Why* - Análises estatísticas

A figura 6.9, que se segue, apresenta o mapa geográfico, elaborado com recurso da ferramenta *Pentaho*, que ilustra a localização das iniciativas inseridas no formulário *online*. Ao clicar em cada marcação são apresentados todos os detalhes da localização relativos à Cidade Inteligente, mais especificamente, cidade, região, estado/província, país, continente, latitude e longitude. Nesta imagem observa-se a localização dos 21 casos de estudo inseridos nesta fase inicial.

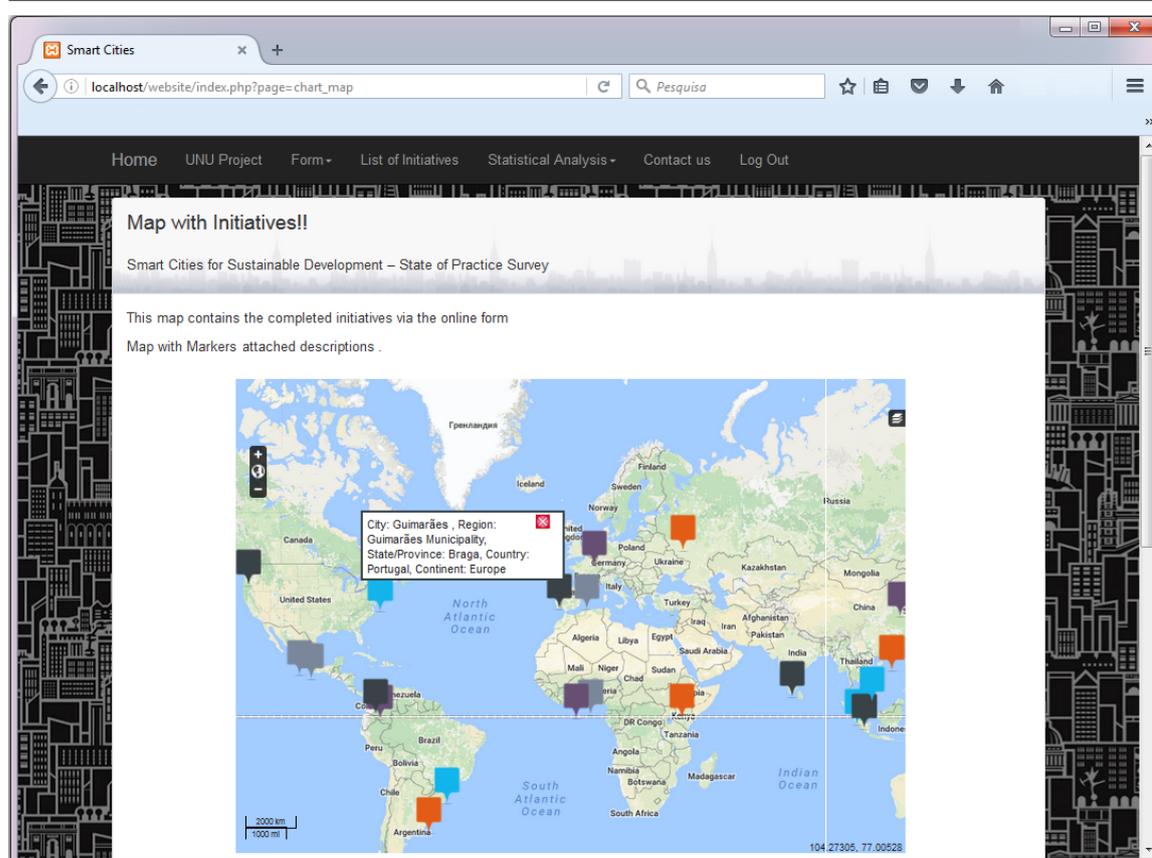


Figura 6.9: Demonstração da página *web* *Map of Initiatives* - Mapa

A plataforma *online* dispõe de uma secção na qual qualquer tipo de utilizador pode contactar o administrador via *e-mail*.

A próxima imagem (figura 6.10) expõe a página *web* denominada por *Contact us*, destinada ao envio de mensagens através de *e-mail* para o administrador. Tal como se observa nesta imagem, o utilizador pode seleccionar um dos tipos de assunto (“*Subject*”) da mensagem. Desse conjunto de opções de assunto, apenas algumas existem para os utilizadores com *login* efetuado, como, por exemplo, para assuntos relacionados com o editar/remover formulários.

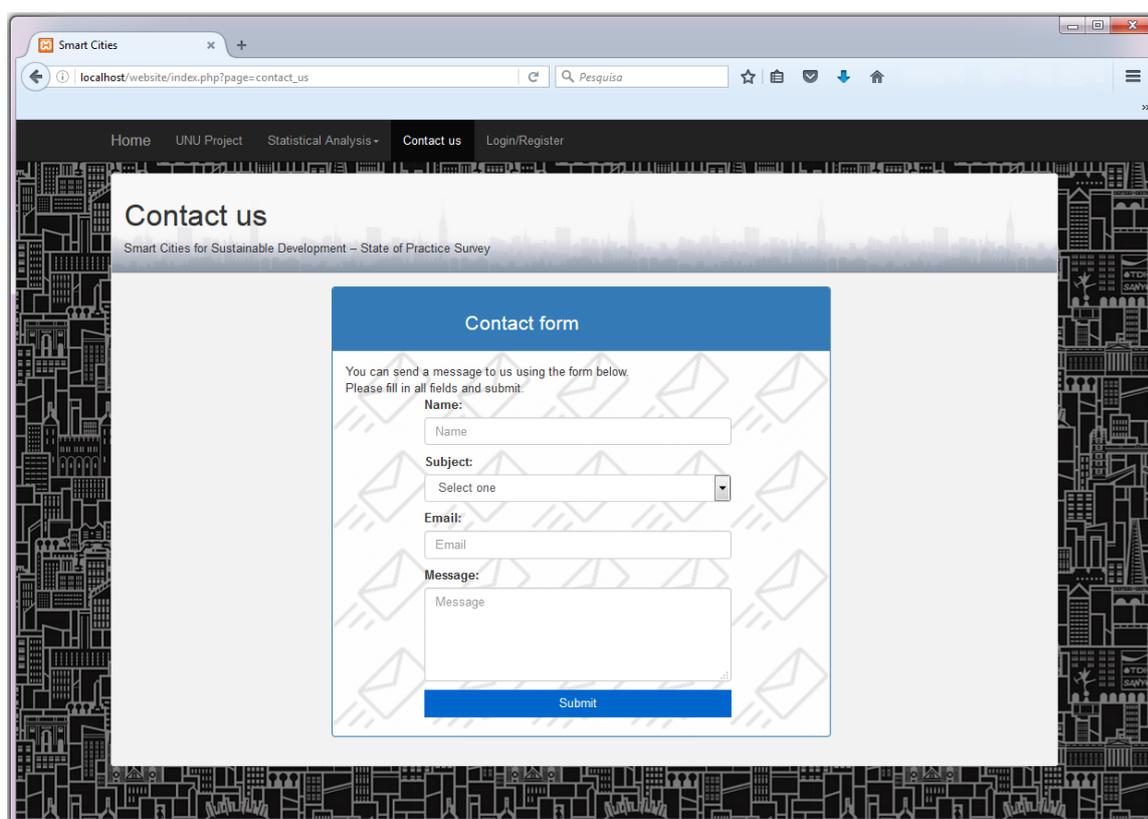


Figura 6.10: Demonstração da página *web* *Contact us*

De seguida, a última secção da plataforma *online* destina-se ao registo do utilizador ou à realização de *login*. No caso do utilizador ter efetuado o *login* devidamente, então a última secção será destinada para o efeito de término da sessão, designada por *Logout*. Esta secção está dividida em duas partes, a primeira destina-se ao *login* do utilizador e a segunda ao registo de um utilizador. Na secção atribuída para o registo, o utilizador preenche os campos solicitados e, assim que submete os dados, recebe um *e-mail* do sistema. O conteúdo deste *e-mail* inclui os dados que preencheu e o pedido de confirmação de *e-mail*, bastando para tal, clicar na hiperligação determinada para o efeito. Logo após o utilizador clicar na hiperligação para confirmar o *e-mail*, é indicado na aplicação *web* do administrador essa confirmação. Deste modo, o administrador antes de autorizar um utilizador a efetuar *login*, saberá que o mesmo é detentor do contacto eletrónico de *e-mail* que forneceu, servindo este para *username* na plataforma. Todos os campos são devidamente validados, quer durante o preenchimento dos dados para *login* como durante o preenchimento do formulário de registo, permitindo apenas o envio dos dados após o correto preenchimento dos

mesmos.

A imagem que se segue (figura 6.11) apresenta a página *web* da plataforma atribuída à realização de *login* por parte de utilizadores registados e autorizados para o efeito pelo(s) administrador(es) do sistema.

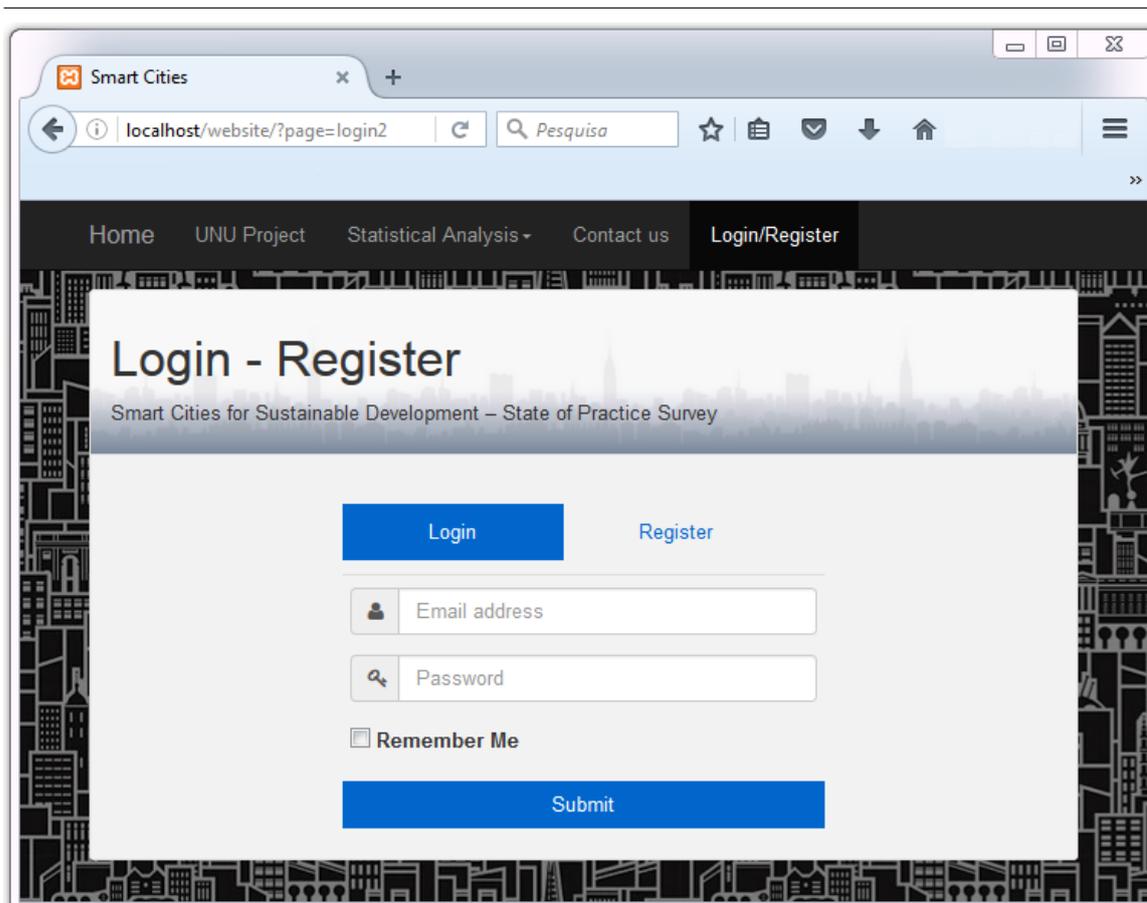


Figura 6.11: Demonstração da página *web* *Login/Register* - Secção *Login*

A imagem seguinte (figura 6.12) expõe a página *web* da plataforma delineada para que os utilizadores possam efetuar o seu registo.

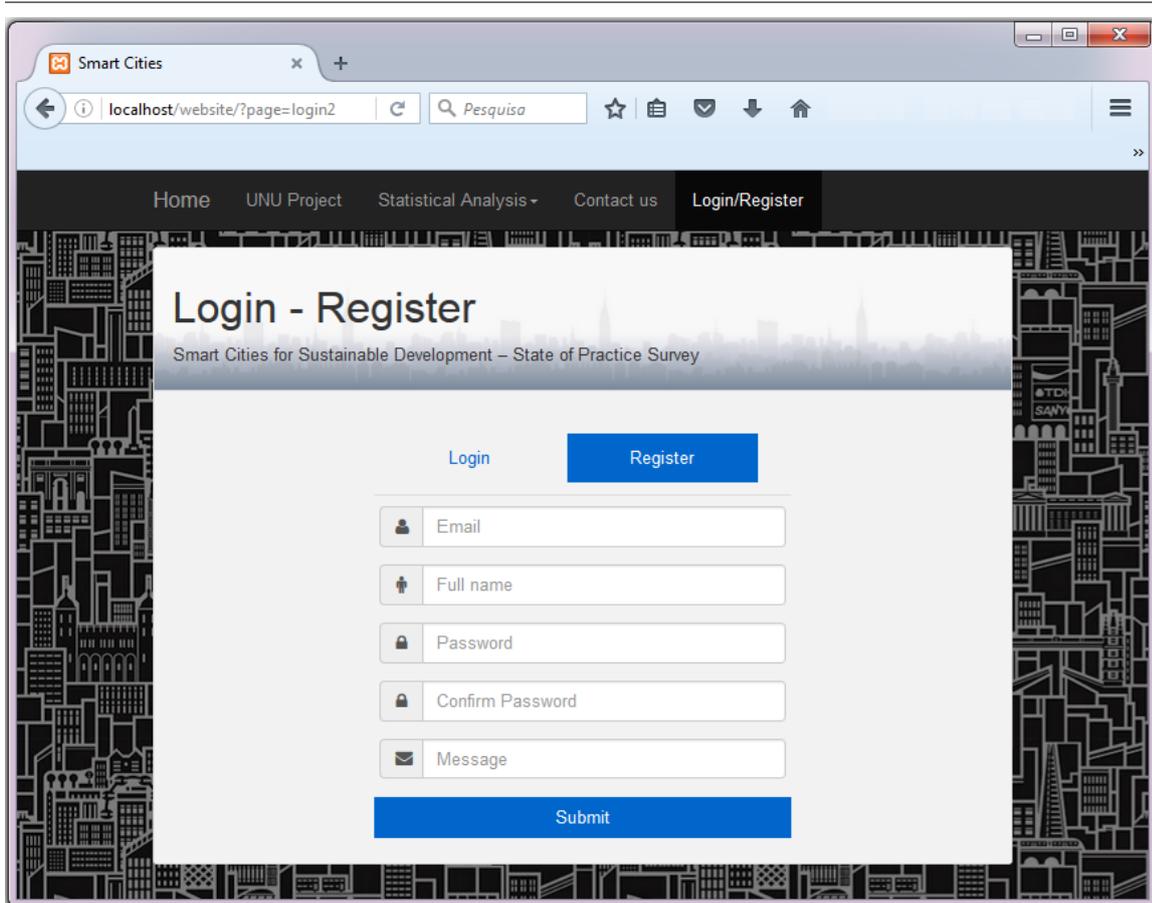


Figura 6.12: Demonstração da página *web Login/Register* - Secção *Register*

Assim que o utilizador submete corretamente os dados no formulário de registo, é enviado, para o *e-mail* do mesmo, uma mensagem sobre o aviso de registo e pedido de confirmação de *e-mail*, tal como se pode observar na figura 6.13.

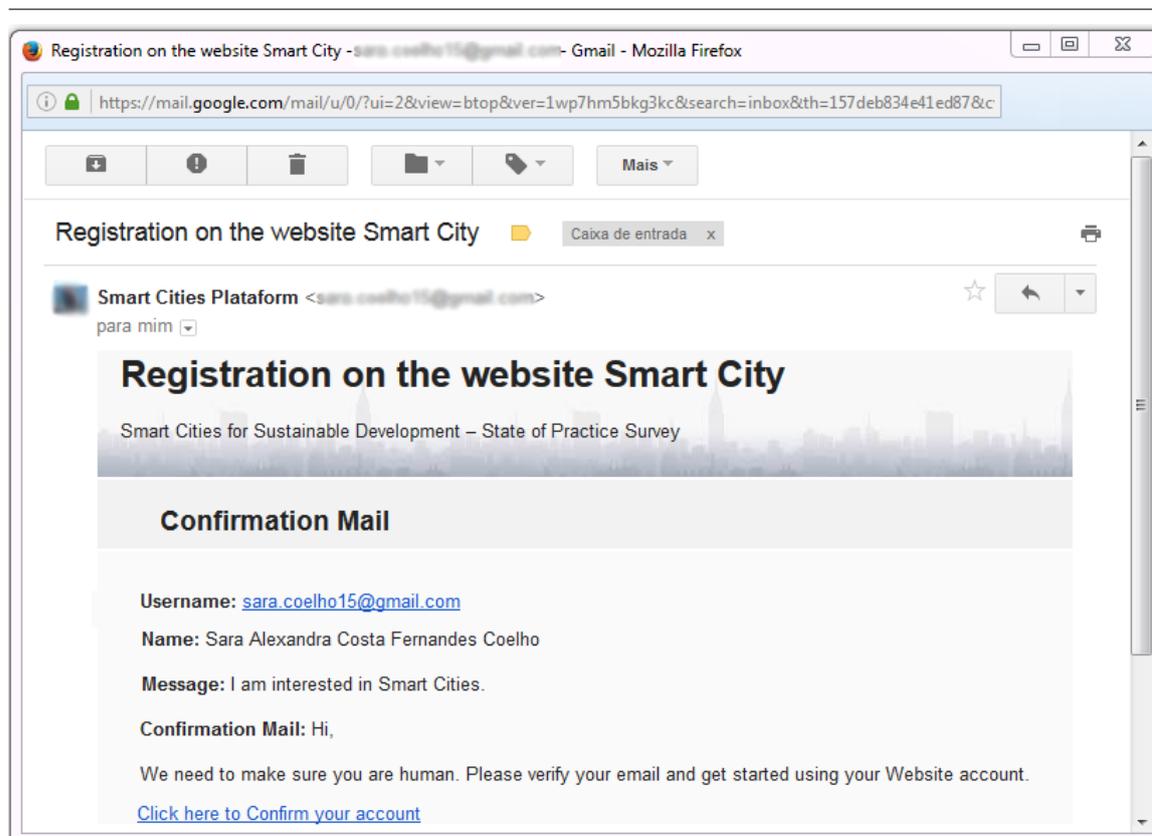


Figura 6.13: Demonstração do *e-mail* enviado ao utilizador após o registo

Após o utilizador clicar na hiperligação para confirmar a conta, é apresentada uma mensagem de confirmação no *web browser*, como se pode observar na figura 6.14.

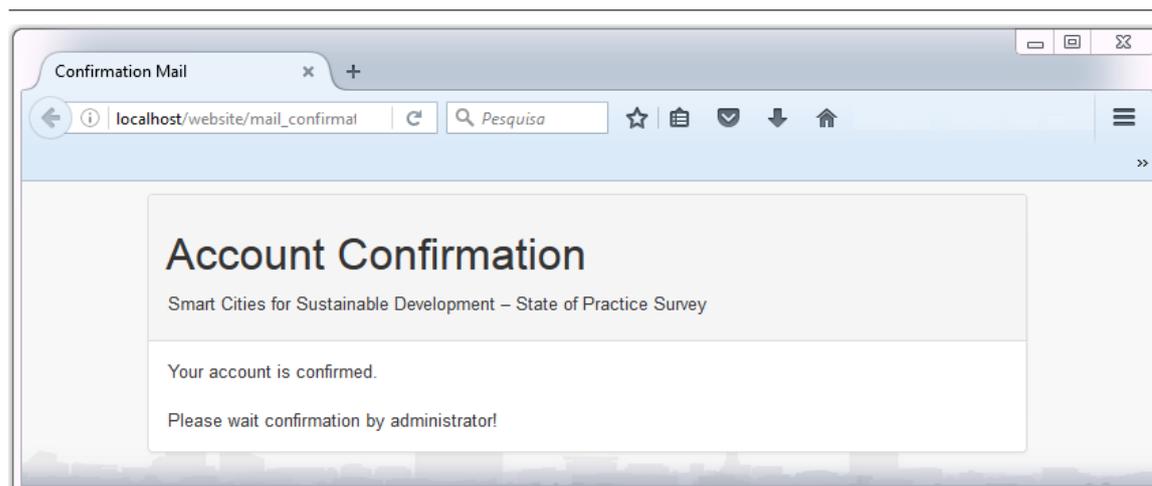


Figura 6.14: Mensagem de confirmação via *web browser*

6.2 Aplicação *web* de Administração

A aplicação *web* de administração tem a finalidade de coordenar os acessos à plataforma *online* na sua totalidade e de proceder a alterações nos formulários. Inicialmente, a primeira página *web* desta aplicação destina-se ao *login* do administrador, tal como se observa na próxima imagem (figura 6.15).

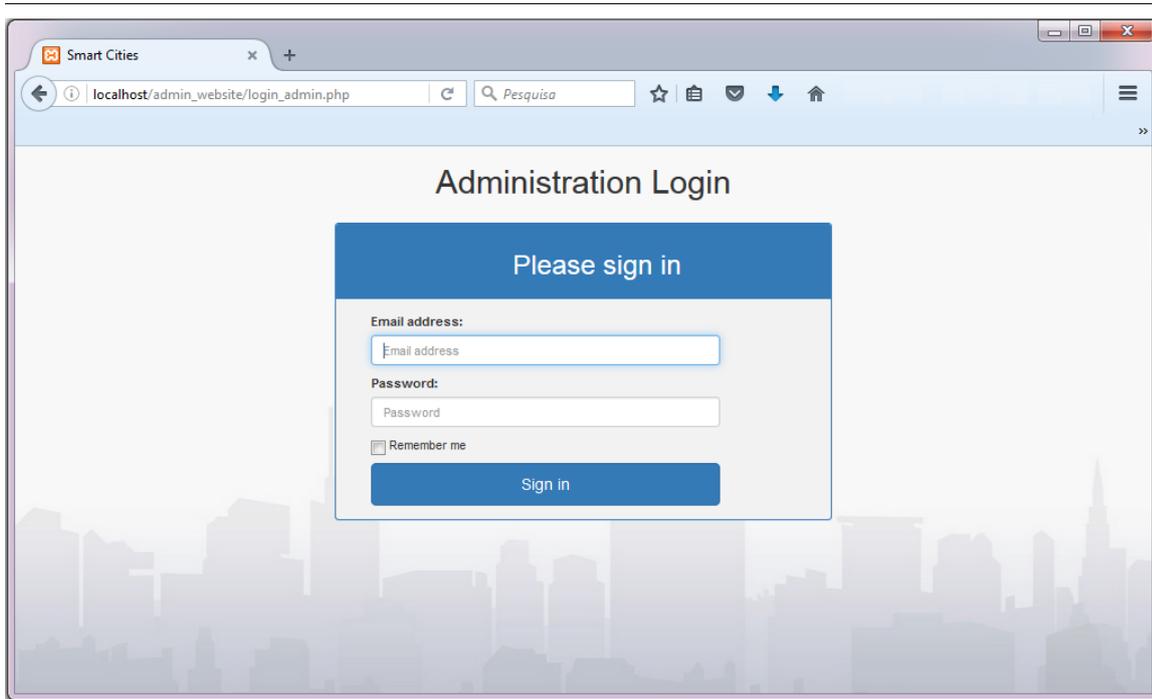


Figura 6.15: Demonstração da página *web Login* - Administração

Assim que o *login* é efetuado devidamente, o administrador é redirecionado para a página de entrada denominada por *Home*, como se pode observar na imagem que se segue (figura 6.16).

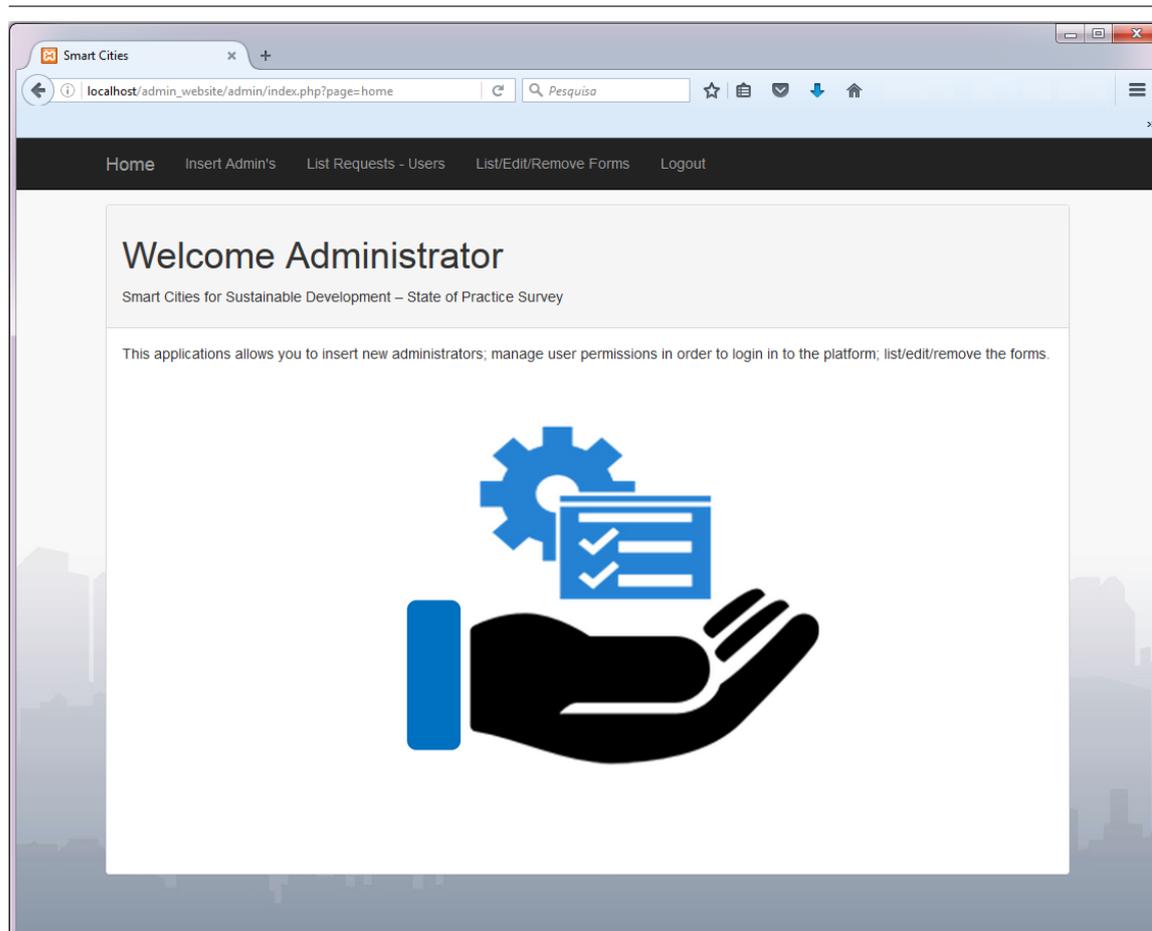


Figura 6.16: Demonstração da página *web Home* - Administração

Em seguida, esta aplicação *web* apresenta uma secção atribuída à inserção de novos administradores. Todos os campos são devidamente validados durante o preenchimento do formulário de inserção, permitindo o envio dos dados somente após o correto preenchimento dos mesmos.

A seguinte imagem (figura 6.17) apresenta a secção designada por *Insert Admin's*, que permite inserir novos administradores.

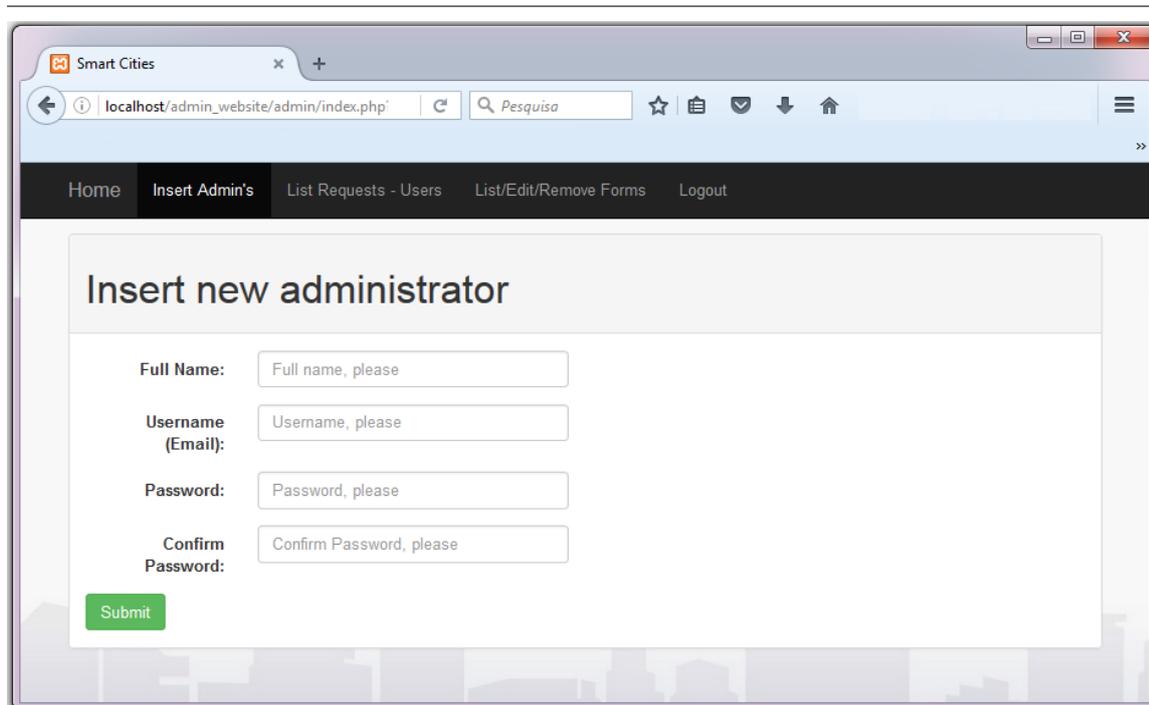


Figura 6.17: Demonstração da página *web Insert Admin's* - Administração

A próxima secção da aplicação *web* destina-se à administração dos pedidos dos utilizadores que efetuaram o registo. Nesta secção o administrador, com base na confirmação do contacto de *e-mail* do utilizador, poderá proceder à autorização para o utilizador estar possibilitado de efetuar o *login* na plataforma *online*. Assim, que o administrador autorize a realização de *login* ou cancele essa autorização, o utilizador receberá, automaticamente, um *e-mail* com o aviso consoante for o caso.

A imagem que se segue (figura 6.18) mostra a página *web* denominada por *List Requests - Users*, a qual apresenta a listagem das solicitações dos utilizadores para a realização de *login* na plataforma *online*.

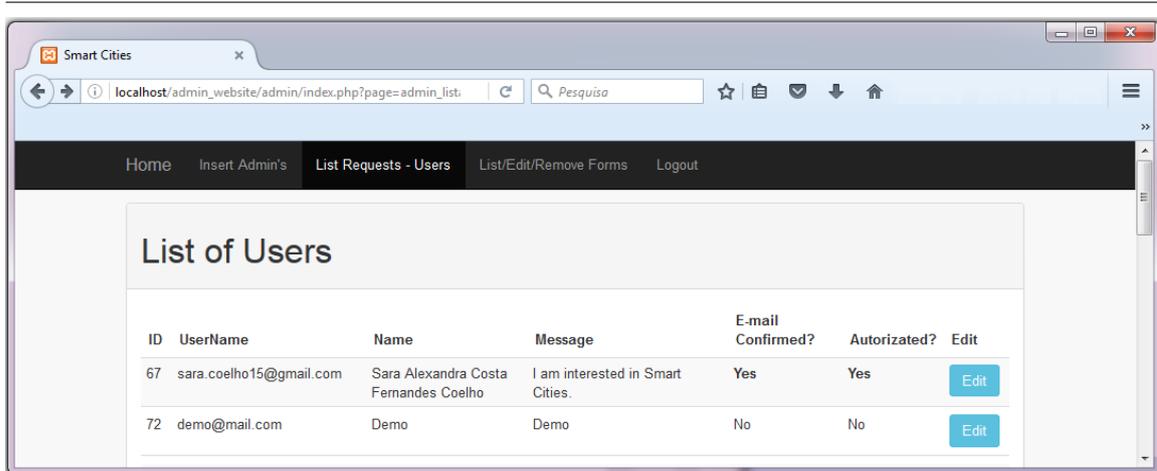


Figura 6.18: Demonstração da página *web* *List Requests - Users* - Administração

A imagem que se segue (figura 6.19) apresenta a página *web* destinada à gestão da permissão dos utilizadores acerca da realização do *login* na plataforma *online*.

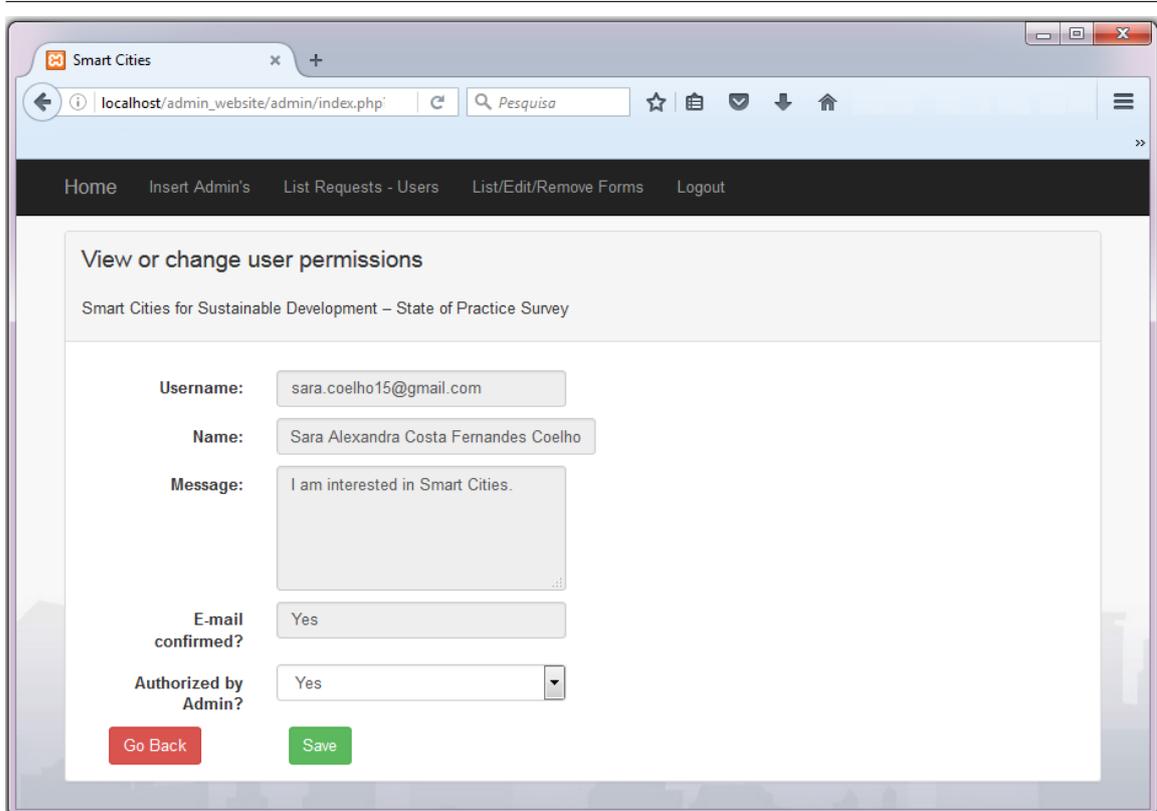


Figura 6.19: Demonstração da página *web* para gestão da permissão dos utilizadores - Administração

A imagem que se segue (figura 6.20) exhibe o formato do *e-mail* enviado ao utilizador, assim que o administrador autorizar o mesmo a poder efetuar o *login* na plataforma *online*.

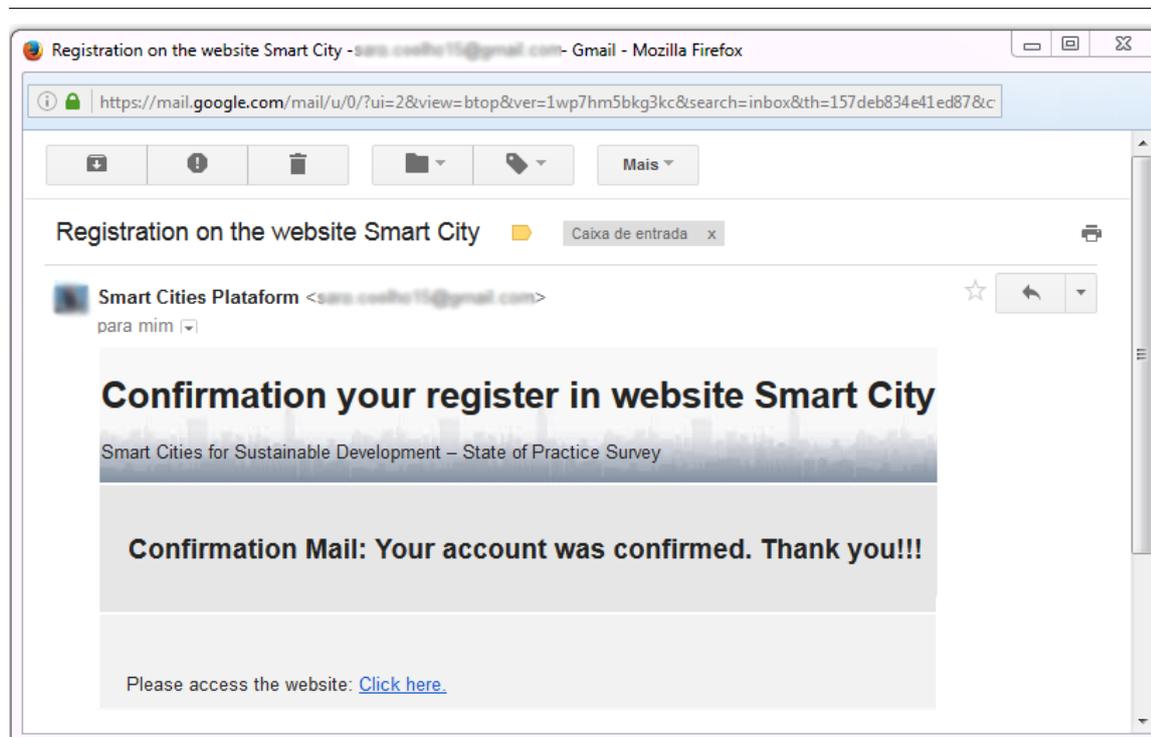


Figura 6.20: Demonstração do *e-mail* de autorização de *login* ao utilizador

A secção anterior à de término de sessão (*Logout*), denomina-se por *List/Edit/-Remove Forms* e destina-se à listagem, edição e remoção dos formulários por parte do(s) administrador(es). O administrador pode visualizar os vários casos de estudo por nome, cidade, município/região, estado/província e país, e seleccionar o que pretende editar ou remover, tal como se pode observar na figura 6.21 que se segue. Ao clicar em editar (*Edit*) serão apresentadas todas as informações do formulário previamente preenchido, podendo-se proceder às alterações pretendidas. Por outro lado, ao clicar em remover (*Remove*) é apresentada uma mensagem a questionar se o administrador tem a certeza que pretende efetuar a remoção do formulário, e em caso afirmativo o sistema procede à remoção, e em caso contrário é mantida a página *web* da listagem.

A figura 6.21 apresenta esta secção de listagem de casos de estudo/iniciativas existentes, com a possibilidade da sua edição e remoção de formulários.

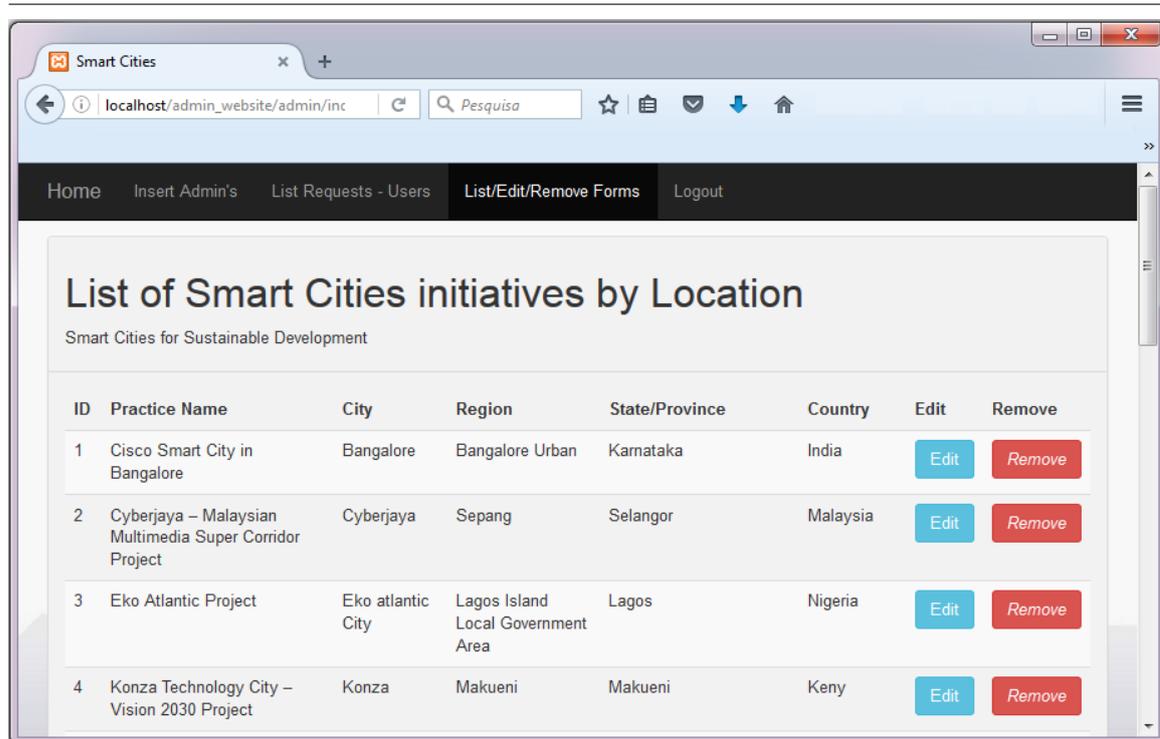


Figura 6.21: Demonstração da página *web List/Edit/Remove Forms* - Administração

6.3 Conclusões

Este capítulo apresentou os principais resultados da elaboração da plataforma *online* para casos de estudo em Cidades Inteligentes e da aplicação *web* para administração.

Na primeira parte foram demonstradas as diversas páginas *web* que constituem a plataforma *online*. Ao longo desta demonstração são apresentadas as diversas funcionalidades disponibilizadas, quer para utilizadores sem autenticação no sistema, como para utilizadores com *login* efetuado. De salientar que, neste capítulo, é exposto o formulário *online*, parte dos diversos *dashboards* elaborados, e ainda, o mapa geográfico, sendo estes últimos elaborados com recurso à tecnologia OLAP.

Na segunda parte deste capítulo é apresentada a aplicação *web* destinada ao(s) administrador(es), que permite inserir novos administradores; gerir as permissões dos pedidos dos utilizadores para realização de *login* na plataforma *online*; e ainda, listar, editar e remover os formulários a pedido dos utilizadores e em concordância de ambos.

A plataforma *online*, juntamente com a aplicação *web*, pretendem constituir um

sistema relevante para tomadas de decisão dos diversos decisores políticos e gestores relacionados com iniciativas sobre Cidades Inteligentes.

Capítulo 7

Conclusões

Este capítulo apresenta uma descrição sucinta do trabalho elaborado, nesta dissertação, de acordo com os objetivos inicialmente traçados (secção 7.1). Também é neste capítulo que são descritas as conclusões retiradas pelo autor, acerca de todas as etapas de desenvolvimento da plataforma elaborada nesta dissertação. De seguida, são apresentadas, neste capítulo, as contribuições deste projeto de dissertação (secção 7.2). No final deste capítulo, são descritas algumas sugestões de expansão à plataforma desenvolvida, de modo a dar continuidade ao desenvolvimento da mesma (secção 7.3).

7.1 Síntese

No início deste documento foi apresentado o âmbito, a motivação, os objetivos e a finalidade da elaboração desta dissertação. As cidades estão em constantes evoluções face ao crescimento da população urbana, tornando-as mais complexas e, em simultâneo, essenciais para o crescimento económico-social. Isto leva ao surgimento de muitos fatores associados ao rápido aumento da população nas cidades, como, por exemplo, as alterações climáticas, o esgotamento de recursos naturais e os riscos económicos e sociais. Estes fatores levam a uma necessidade urgente em encontrar soluções inteligentes devido a estes desafios que constantemente estão a surgir. Muitas dessas soluções e estratégias inteligentes procuram, assim, tornar as cidades mais habitáveis e sustentáveis, tendo em conta diversos parâmetros. De entre esses parâmetros destacam-se os seguintes: proporcionar uma urbanização sustentável, obter uma economia atrativa, fomentar a inovação, melhorar a eficiência, promover a resiliência e estimular a igualdade social. Estas cidades que procuram e

adotam este tipo de soluções são denominadas por Cidades Inteligentes. Muitas das práticas associadas às Cidades Inteligentes têm sido realizáveis devido à utilização das TIC, que possibilitam a obtenção de melhores condições de vida e promovem a sustentabilidade.

Relativamente aos setores públicos e empresariais, devido às constantes evoluções, tais como o rápido aumento do volume da informação, o desenvolvimento tecnológico e a concorrência gradual, estes procuram implementar soluções capazes de corresponder a esses progressos para tornarem-se mais competitivos e inovadores.

A necessidade crescente em transformar dados, oriundos de diversas fontes e sistemas, em conhecimento útil, para suporte de tomadas de decisão, tem levado levado à aplicação de sistemas de informação mais eficientes, capazes de armazenar e procurar grandes volumes de dados. Esta necessidade tem conduzido as empresas a adotarem o modelo de *Business Intelligence*, de modo a melhorarem a sua eficiência na gestão de negócios, facilitando a definição de estratégias de competitividade. As empresas utilizam ferramentas de *Business Intelligence* para reunir, armazenar, manipular, fornecer o acesso e analisar dados de grande volume.

Um dos instrumentos de *Business Intelligence* é o sistema de *Data Warehouse*, que permite o armazenamento, a integração e a manipulação de dados de uma ou mais fontes, criando, assim, repositórios de dados. De modo a analisar os grandes volumes de dados, normalmente são utilizadas, neste tipo de sistemas, ferramentas OLAP. Estas ferramentas OLAP permitem, com base na recolha de informação, obter análises eficientes e rápidas dos dados através de diversas perspetivas.

Assim, neste âmbito surgiu o projeto de dissertação “Plataforma Online para casos de estudo em Cidades Inteligentes”. Este projeto tem a finalidade de elaborar uma plataforma *online*, através de ferramentas *open source* e de técnicas OLAP. Esta plataforma *online* permite reunir um conjunto de informações sobre casos de estudo e iniciativas em Cidades Inteligentes através de um formulário *online*. O formulário *online* encontra-se dividido em cinco questões, que são: “Quem?”, “Onde?”, “O quê?”, “Porquê?” e “Como?”. Estas questões permitem obter diversas informações sobre cada iniciativa, tais como: quais as instituições responsáveis, que parceiros estão envolvidos, quais os objetivos e principais características, quais os impactos sobre as dimensões sociais, económicas ou ambientais, onde fica localizada e quando foi desenvolvida. Com base nos principais dados, obtidos a partir do formulário, são apresentadas algumas análises através de *dashboards* e um mapa geográfico, elaborados com ferramentas *open source* do tipo OLAP. Esta plataforma inicialmente reúne

informação acerca de 21 casos de estudos referenciados no projeto “*Smart Cities for Sustainable Development*”, desenvolvido em parceria com a UNU-EGOV e o IDRC.

A elaboração deste projeto de dissertação assentou fundamentalmente em dois pontos essenciais: um de caráter teórico e outro prático. Primeiramente, a fase teórica, que consistiu numa revisão literária sobre os aspetos mais relevantes que englobam o tema. E, posteriormente, a fase prática, que englobou a análise, o esboço e a implementação do sistema, com recurso a ferramentas *open source*.

A primeira parte da revisão literária, expôs a descrição dos vários conceitos de Cidade Inteligente e das principais dimensões interligadas a este termo. Ainda, foi referenciada, nesta revisão, a importância da utilização das TIC e a relevância das iniciativas em Cidades Inteligentes, para promover a sustentabilidade. Por último, foi apresentado, de forma breve, o projeto “*Smart Cities for Sustainable Development*”. A segunda parte, consistiu na revisão do estado da arte relativo ao conceito de *Business Intelligence*. Deste conceito, destacam-se as noções de *Data Warehouse*, modelação multidimensional dos dados, Cubo OLAP e sistemas de processamento analítico dos dados (sistemas OLAP) para a manipulação e análise rápida a grandes volumes de dados.

De seguida, foi apresentada uma visão geral acerca dos critérios conhecidos para avaliação de produtos OLAP. De modo a selecionar uma ferramenta do tipo OLAP, para a criação das análises aos dados captados a partir do formulário, foi realizado um estudo de cinco ferramentas de BI *open source* mais utilizadas pela comunidade científica e empresarial: *SpagoBI*, *Pentaho*, *JasperSoft*, *Palo/Jedox* e *Vanilla*. A seleção da ferramenta de BI *open source* foi fundamentada no que era pretendido para as análises da plataforma e nas funcionalidades disponibilizadas por essas ferramentas estudadas. Esse estudo foi enriquecido por uma tabela de comparação entre essas cinco ferramentas e uma revisão literária sobre diversas ferramentas de BI *open source*.

Posteriormente, foram identificados os requisitos levantados para a implementação do sistema, sendo que estes são constituídos pelos requisitos funcionais e não funcionais. Este levantamento de requisitos teve por base diversos aspetos, nomeadamente os objetivos e as necessidades para a elaboração do sistema, bem como o tipo de tecnologias requerido. Além disso, foram apresentados os diagramas de casos de uso, que demonstram a interação entre os atores (utilizadores com e sem *login* efetuado e administradores) e o sistema. Também foram apresentados os principais esboços elaborados para uma melhor compreensão dos requisitos e serviços/funcionalidades a disponibilizar na plataforma *online*. Ainda, foi apresentada uma visão

geral sobre as várias componentes constituintes da arquitetura do sistema relativo à plataforma *online*.

Após o levantamento de requisitos e da arquitetura, procedeu-se à descrição da modelação dos dados e implementação da plataforma *online*. Primeiramente, foram descritas todas as ferramentas, tecnologias e linguagens utilizadas durante a implementação do sistema. Em seguida, foi exposta a modelação dos dados e o desenvolvimento do sistema de *Data Warehouse*, que permite recolher e armazenar os dados provenientes do formulário *online*, servindo, também, de suporte ao sistema OLAP. Posteriormente, foi demonstrado o procedimento da elaboração do Cubo multidimensional, utilizado pelo sistema OLAP para a exploração e análise multidimensional dos dados. Além disso, foi exposto o modo como foram criados os diversos *dashboards* e o mapa geográfico das Cidades Inteligentes na aplicação *front-end* PUC do *Pentaho*. Seguidamente, descreveu-se os principais aspetos na criação e implementação da plataforma *online* e da aplicação *web* para a administração da mesma.

Durante o processo de implementação do protótipo do sistema foram ultrapassados alguns obstáculos, nomeadamente a exploração da ferramenta *open source Pentaho*, utilizada para a análise dos dados. A utilização desta ferramenta surgiu da opção resultante de um estudo de algumas ferramentas de *Business Intelligence*, sendo por isso completamente desconhecida. Portanto, foi essencial dedicar um longo período de tempo para a exploração e compreensão de toda a ferramenta e das componentes que a constitui. Ao longo da pesquisa para a instalação/configuração, funcionamento, exploração e resolução de problemas da ferramenta *Pentaho*, deparou-se com informações escassas e não atualizadas, incluindo a documentação técnica de suporte ao funcionamento dos *plugins*. Isto é devido sobretudo, ao facto de que esta ferramenta *open source* encontra-se em constantes evoluções e melhoramentos pela Comunidade *Pentaho*. Além disso, a configuração e a instalação desta ferramenta e a migração para o SGBD *MySQL* constituíram processos de carácter minucioso, pois caso contrário o correto funcionamento da mesma estaria comprometido. Todavia, a exploração da ferramenta *Pentaho* foi enriquecedora, viabilizando a obtenção de gráficos analíticos e de um mapa que possibilitam, aos decisores e às entidades responsáveis, a tomada de decisões mais eficiente.

Por último, foram demonstrados os resultados obtidos com a elaboração da plataforma *online* e da aplicação *web* de administração. De entre as diversas funcionalidades disponibilizadas na plataforma *online*, destacam-se o formulário *online*, para a inserção de casos de estudo/iniciativas em Cidades Inteligentes; a listagem dos

dados dos diversos formulários preenchidos; e a disponibilização de diversos *dashboards* e de um mapa associados aos principais dados retirados desse formulário. Nesta plataforma, assim que um utilizador efetua o registo, recebe no *e-mail* um pedido de confirmação de *e-mail*. Alguns dos conteúdos, disponibilizados pela plataforma, apenas poderão ser acedidos por utilizadores que estejam registados e devidamente autenticados. Quanto à aplicação *web* salienta-se, como principais funcionalidades disponibilizadas, a administração da permissão de autenticação dos utilizadores na plataforma *online* e a edição/remoção dos diversos formulários. Nesta aplicação *web* destinada à administração da plataforma *online*, sempre que o administrador autoriza ou cancela o pedido de realização de *login* a um utilizador, o mesmo receberá um *e-mail* com a indicação devida consoante for a situação.

A criação da plataforma *online* possibilita reunir informações de práticas existentes em Cidades Inteligentes por todo o mundo, através do preenchimento de um formulário. Nesta plataforma são disponibilizados um robusto conjunto de gráficos e um mapa geográfico que permitem obter análises dessas informações captadas. Estas informações, que estão relacionadas com os vários atributos relativos aos domínios, permitem analisar os aspetos tidos em conta e os resultados obtidos das diversas iniciativas inseridas na plataforma.

Desta forma, a plataforma poderá contribuir para que os responsáveis pelas iniciativas possam aperfeiçoar determinados aspetos nas práticas existentes, conduzindo, assim, a uma progressão eficaz do desenvolvimento sustentável e socioeconómico. Além do mais, a plataforma *online*, também, poderá promover novas iniciativas, na medida que permite, aos decisores políticos e aos responsáveis por iniciativas em Cidades Inteligentes, servirem-se desta como ferramenta de suporte, para o estudo e implementação de iniciativas futuras.

A plataforma fornece, assim, apoio para tomadas de decisão estratégicas, dos gestores e decisores políticos, que pretendam dinamizar as iniciativas existentes. E ainda, esta plataforma, pretende incentivar, os mesmos, à implementação de novas iniciativas em Cidades Inteligentes.

7.2 Contribuições

As cidades estão em constante progresso, devido, sobretudo, ao crescimento da população nas zonas urbanas, o que provoca diversos problemas como esgotamento de recursos naturais, conduzindo à adoção de soluções inteligentes. Muitas das soluções

inteligentes assentam no favorecimento da sustentabilidade, no desenvolvimento social e económico, tornando-as apelidadas por Cidades Inteligentes. No que concerne ao setor empresarial e público, as empresas face ao aumento contínuo da concorrência, do desenvolvimento tecnológico e da informação, procuram cada vez mais soluções aptas para corresponder a esses desafios. A necessidade de tomadas de decisão fundamentais e instantâneas juntamente com a procura de obtenção de grandes volumes de dados, têm conduzido a empresas a adotarem sistemas de *Business Intelligence*.

Posto isto surgiu a oportunidade de desenvolver uma plataforma *online* para reunir informações sobre as iniciativas em Cidades Inteligentes, e com base nesses dados, disponibilizar gráficos analíticos. Assim, o projeto desenvolvido consistiu num sistema para análise de iniciativas, contribuindo para facilitar o processo de tomadas de decisão por parte de decisores políticos e interessados nesta matéria. A elaboração desta plataforma inovadora, vai contribuir para que todos os responsáveis por iniciativas em Cidades Inteligentes, possam analisar e até melhorar as práticas existentes, e em simultâneo, incentivar a implementação de novas iniciativas.

Este trabalho desenvolvido teve a colaboração da UNU-EGOV, tendo por base, para a elaboração do formulário, o documento “*Smart Cities for Development - State of Practice Survey*”, incluído no projeto “*Smart Cities for Sustainable Development*”, desenvolvido pelos investigadores Elsa Estevez, Nuno Vasco Lopes e Tomasz Janowski, em parceria com a UNU-EGOV e o IDRC. Esta plataforma, inicialmente, está composta pelos 21 casos de estudo em Cidades Inteligentes disponibilizados também neste projeto.

Além do mais, sendo este projeto relacionado com a área de *Business Intelligence*, consistiu inicialmente numa área ainda inexplorada pelo autor, constituindo, portanto, o primeiro trabalho realizado nesta matéria. Contudo, o projeto revelou-se uma mais valia em termos pessoais, o qual facultou a oportunidade de exploração de diversos conhecimentos quer nesta área de BI, fortemente utilizada no setor empresarial para tomadas de decisão, quer em termos de desenvolvimento *web*.

De referir também, que, como trabalho futuro, este protótipo de sistema elaborado, apresenta uma elevada contingência em proporcionar continuidade ao desenvolvimento e aperfeiçoamento do mesmo.

Por fim, salienta-se o facto deste trabalho ter dado origem à elaboração e publicação científica em conferência de âmbito internacional: “*SpliTech2016 - International Multidisciplinary Conference On Computer and Energy Science*”, decorrida entre 13 a 14 de Julho de 2016. Esta publicação é intitulada por “*Online Platform for*

Case Studies in Smart Cities”, com os autores Sara Coelho e Nuno Vasco Lopes, e editor *University of Split, FESB* [82]. A referida publicação encontra-se publicada na Ata da respetiva conferência e disponibilizada em *IEEE Xplore*, bem como em outros serviços de bases de dados destinados a este tipo de publicações.

7.3 Trabalho Futuro

Os objetivos traçados inicialmente para este projeto foram cumpridos, sendo obtido no final um sistema completamente operacional. No entanto, é possível adicionar novas funcionalidades, tanto na plataforma *online* como na aplicação *web* de administração, de modo a refinar e a estender a gama de funcionalidades disponibilizadas neste sistema, tornando-o, assim, mais completo e aprimorado. De entre essas novas funcionalidades, destacam-se as seguintes:

- Associar ao *Data Warehouse* uma tabela de dimensão para o armazenamento de informação relativa aos utilizadores que preencheram os formulários - ao inserir esta tabela permitirá associar um formulário ao utilizador que o preencheu;
- Fornecer a possibilidade dos utilizadores editarem os dados dos formulários que preencheram, e/ou até mesmo removê-los na plataforma *online*, somente com a autorização do administrador - o sistema permite aos administradores a edição/remoção de formulários a pedido dos utilizadores, no entanto, seria vantajoso facultar aos utilizadores essa funcionalidade;
- Introdução de filtros na listagem das iniciativas, e na listagem de utilizadores na aplicação *web* de administração - de modo a facilitar a pesquisa de iniciativas ou de utilizadores para realização de *login* seria útil desenvolver filtros de pesquisa;
- Exploração analítica dos dados noutras ferramentas - Os *dashboards* e o mapa ilustrativo das localizações das iniciativas foram elaborados com recurso à ferramenta *Pentaho* na versão *open source*. No entanto, seria interessante fazer uma comparação por exemplo com a ferramenta proprietária, ou até mesmo com outras ferramentas *open source* estudadas e referenciadas neste projeto.

Bibliografia

- [1] E. Estevez, N. V. Lopes, and T. Janowski, “Smart Sustainable Cities - Reconnaissance Study,” Tech. Rep., 2016. [Online]. Available: https://joinup.ec.europa.eu/sites/default/files/smart_cities_report.pdf
- [2] S. Chaudhuri, U. Dayal, and V. Narasayya, “An overview of business intelligence technology,” *Communications of the ACM*, vol. 54, no. 8, p. 88, 2011.
- [3] A. Singhal, “An Overview of Data Warehouse, OLAP and Data Mining Technology,” in *Data Warehousing and Data Mining Techniques for Cyber Security*. Boston, MA: Springer US, 2007, vol. 31, ch. Chapter 1, pp. 1–23. [Online]. Available: http://dx.doi.org/10.1007/978-0-387-47653-7_1
- [4] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, 2nd ed., J. Gray and M. Reserach, Eds. San Francisco, CA: Morgan Kaufmann Publishers, 2006.
- [5] S. Chaudhuri and U. Dayal, “An overview of data warehousing and OLAP technology,” *ACM Sigmod Record*, vol. 26, no. 1, pp. 65–74, 1997.
- [6] Infosoporte’s, “Instalar SpagoBi, Business Intelligence Open Source,” p. 1, 2012. [Online]. Available: <https://infosoporte.wordpress.com/2012/11/22/instalar-spagobi-business-intelligent-open-source/>
- [7] JasperSoft, “Self-Service Reporting & Analysis Solutions for Any User,” p. 1, 2016. [Online]. Available: <http://www.jaspersoft.com/self-service-business-intelligence-improves-your-app>
- [8] BPM-Conseil, “Visualization : Choose your Chart type, merge KPI with Open Street Maps, Google Maps ...,” p. 1, 2013. [Online]. Available: <http://www.bpm-conseil.com/content/big-data>
- [9] S. O. S. Solutions, “Palo BI Suite,” p. 1. [Online]. Available: <http://decisionnel-open-source.smile.fr/Les-suites-decisionnelles/Palo-bi-suite>
- [10] Xoomtrainings, “Introduction and Architecture of Pentaho,” p. 1, 2015. [Online]. Available: <http://www.xoomtrainings.com/blog/introduction-and-architecture-of-pentaho>

- [11] Pentaho, “Graphical and Responsive Dashboards,” p. 1, 2013. [Online]. Available: <http://www.pentaho.com/product/business-visualization-analytics>
- [12] J. Hedlund, “The Smart City Using IT to Make Cities More Livable,” pp. 1–22, 2011. [Online]. Available: <http://smartcitiescouncil.com/resources/smart-city-using-it-make-cities-more-livable-0>
- [13] Polsku, “OLTP vs. OLAP,” p. 1, 2009. [Online]. Available: <http://datawarehouse4u.info/OLTP-vs-OLAP.html>
- [14] DataOnFocus, “OLTP vs OLAP,” p. 1, 2015. [Online]. Available: <http://www.dataonfocus.com/oltp-vs-olap/>
- [15] G. S. Reddy, R. Srinivasu, M. P. C. Rao, and S. R. Rikkula, “Data Warehousing, Data Mining, OLAP and OLTP Technologies are essential elements to support decision-making process in industries,” *International Journal on ...*, vol. 02, no. 09, pp. 2865–2873, 2010. [Online]. Available: <http://enggjournals.com/ijcse/doc/IJCSE10-02-09-010.pdf>
- [16] K. J. Cios, R. W. Swiniarski, W. Pedrycz, and L. A. Kurgan, “Databases, Data Warehouses, and OLAP,” in *Data Mining: A Knowledge Discovery Approach*. Boston, MA: Springer US, 2007, ch. 6, pp. 95—131. [Online]. Available: http://dx.doi.org/10.1007/978-0-387-36795-8_6
- [17] S. Nagabhushana, *Data Warehousing OLAP and Data Mining*. New Delhi: New Age International (P) Ltd, 2008.
- [18] “Codd’s Paper,” p. 1. [Online]. Available: <http://olap.com/learn-bi-olap/codds-paper/>
- [19] M. M. Cruz-Cunha, F. Moreira, and J. a. Varajão, *Handbook of Research on Enterprise 2.0: Technological, Social, and Organizational Dimensions (2 Volumes)*, J. Gamon, Ed. United States of America: Business Science Reference, 2013.
- [20] M. R. Information Resources, *Open Source Technology: Concepts, Methodologies, Tools and Applications*, K. Wolfe, C. Henning, and A. DeMarco, Eds. United States of America: Information Science Reference, 2015.
- [21] M. Tereso and J. Bernardino, “Open source business intelligence tools for SMEs,” *Information Systems and Technologies (CISTI), 2011 6th Iberian Conference on*, pp. 1 – 4, 2011. [Online]. Available: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5974187>
- [22] A. Marinheiro and J. Bernardino, “Analysis of Open Source Business Intelligence Suites,” *Analysis of open source Business Intelligence suites*, p. 7, 2013. [Online]. Available: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6615764>

- [23] —, “Experimental Evaluation of Open Source Business Intelligence Suites using OpenBRR,” *IEEE Latin America Transactions*, vol. 13, no. 3, pp. 810–817, 2006. [Online]. Available: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=7069109>
- [24] J. Bernardino, “Open Source Business Intelligence Platforms for Engineering Education,” in *SEFI Annual Conference 2011*, vol. 1. Lisboa: European Society for Engineering Education (SEFI), 2011, pp. 693–698. [Online]. Available: <http://www.sefi.be/wp-content/papers2011/T13/164.pdf>
- [25] M. Golfarelli, “Open source bi platforms: A functional and architectural comparison,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5691 LNCS, 2009, pp. 287–297.
- [26] Pentaho, “Components Reference,” p. 1, 2016. [Online]. Available: <https://help.pentaho.com/Documentation/5.4/OD0/160/000>
- [27] T. Nam and T. A. Pardo, “Conceptualizing smart city with dimensions of technology, people, and institutions,” *Proceedings of the 12th Annual International Digital Government Research Conference on Digital Government Innovation in Challenging Times - dg.o '11*, p. 282, 2011. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2037556.2037602>
- [28] R. Giffinger, C. Fertner, H. Kramar, R. Kalasek, N. Pichler-Milanovic, and E. Meijers, “Smart Cities - Ranking of European medium-sized cities,” University of Technology, Centre of Regional Science, Vienna, Tech. Rep. 1, 2007. [Online]. Available: http://www.smart-cities.eu/download/smart_cities_final_report.pdf
- [29] R. G. Hollands, “Will the real smart city please stand up?” *City*, vol. 12, no. 3, pp. 303–320, 2008.
- [30] K. Gama, A. Alvaro, and E. Peixoto, “Em Direção a um Modelo de Maturidade Tecnológica para Cidades Inteligentes,” *VIII Simpósio Brasileiro de Sistemas de Informação*, no. Sbsi, pp. 150–155, 2012.
- [31] E. Woods and N. Goldstein, “Navigant Research Leaderboard Report: Smart City Suppliers. Assessment of Strategy and Execution for 15 Smart City Suppliers,” Navigant Consulting, Inc., Boulder, USA, Tech. Rep., 2014. [Online]. Available: http://www.ibm.com/smarterplanet/global/files/us__en__us__smarter_cities__navigant_research_report.pdf
- [32] R. E. Hall, B. Bowerman, J. Braverman, J. Taylor, H. Todosow, and U. von Wimmersperg, “The vision of a smart city,” *2nd International Life Extension Technology Workshop*, p. 7, 2000. [Online]. Available: ftp://24.139.223.85/Public/Tesis_2011/Paper_Correction_4-15-09/smartycitypaperpdf.pdf

- [33] D. Schuurman, B. Baccarne, L. De Marez, and P. Mechant, “Smart ideas for smart cities: Investigating crowdsourcing for generating and selecting ideas for ICT innovation in a city context,” *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 7, no. 3, pp. 49–62, 2012.
- [34] I. T. Iso/Iec Jtc 1, “Information technology Smart cities,” *Smart cities - Preliminary Report 2014*, pp. 1–71, 2014.
- [35] S. Dirks and M. Keeling, “A vision of smarter cities: How cities can lead the way into a prosperous and sustainable future,” *IBM Global Business Services*, 2009.
- [36] A. Caragliu, C. Del Bo, and P. Nijkamp, “Smart Cities in Europe,” *Journal of Urban Technology*, vol. 18, no. 2, pp. 65–82, 2011.
- [37] D. F. Anderle and V. F. Junior, “A utilização da Tecnologia da Informação nas “Smart Cities” – um estudo bibliométrico,” in *Proceeding of the 3rd International Congress of Knowledge and innovation (ciKi)*. Porto Alegre, Brasil: Instituto Federal Catarinense, 2013, pp. 1–20. [Online]. Available: <http://www.congresociki.org/wp-content/uploads/2013/08/ATAS-III-ciKi-2013.pdf>
- [38] Fundación Telefónica, “Smart Cities: un primer paso hacia la Internet de las cosas,” *Informatica Economica*, vol. 16, no. 4, pp. 57–67, 2011. [Online]. Available: http://www.socinfo.es/contenido/seminarios/1404smartcities6/01-TelefonicaSMART_CITIES-2011.pdf
- [39] M.-L. Marsal-Llacuna, J. Colomer-Llinàs, and J. Meléndez-Frigola, “Lessons in urban monitoring taken from sustainable and livable cities to better address the Smart Cities initiative,” *Technological Forecasting and Social Change*, vol. 90, no. PB, pp. 611–622, 2015.
- [40] T. Nam and T. a. Pardo, “Smart City as Urban Innovation: Focusing on Management, Policy, and Context,” *Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance - ICEGOV '11*, p. 185, 2011. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2072069.2072100>
- [41] H. Chourabi, T. Nam, S. Walker, J. R. Gil-Garcia, S. Mellouli, K. Nahon, T. a. Pardo, and H. J. Scholl, “Understanding smart cities: An integrative framework,” *Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 2289–2297, 2012.
- [42] C. Selada and C. Silva, “As Cidades Inteligentes na Agenda Europeia: Oportunidades para Portugal,” in *II Conferência de Planeamento Regional e Urbano, VIII ENPLAN e XVIII Workshop APDR: “Europa 2020: retórica, discursos, política e prática”*, U. Editora, Ed., no. 1, Aveiro, Portugal,

- 2013, pp. 1–31. [Online]. Available: http://www.inteli.pt/uploads/documentos/documento_1373454640_1255.pdf
- [43] A. Kramers, M. Höjer, N. Lövehagen, and J. Wangel, “Smart sustainable cities - Exploring ICT solutions for reduced energy use in cities,” *Environmental Modelling and Software*, vol. 56, pp. 52–62, 2014.
- [44] F. R. Botelho and R. E. Filho, “Conceituando o termo Business intelligence: origem e principais objetivos,” pp. 1–7, 2014.
- [45] H. Luhn, “A Business Intelligence System,” *IBM Journal of Research and Development*, vol. 2, no. 4, pp. 314–319, 1958. [Online]. Available: <http://altaplana.com/ibmrd0204H.pdf>
- [46] B. Nedelcu, “Business Intelligence Systems,” *Database Systems Journal*, vol. IV, no. 4, pp. 12–20, 2012.
- [47] D. Agiu, V. Mateescu, and J. Mutean, “Business Intelligence overview,” *Database Systems Journal*, vol. V, no. 3, pp. 23–36, 2014.
- [48] D. Power, “A Brief History of Decision Support Systems,” *DSSResources.com*, p. Version 2.8, 2003. [Online]. Available: <http://dssresources.com/history/dsshistoryv28.html>
- [49] M. Golfarelli, S. Rizzi, and I. Cella, “Beyond Data Warehousing : What’s Next in Business Intelligence?” *7th ACM international workshop on Data warehousing and OLAP*, pp. 1–6, 2004.
- [50] S. Negash, “Business intelligence,” pp. 1–21, 2004. [Online]. Available: <http://aisel.aisnet.org/cais/vol13/iss1/15>
- [51] L. Zeng, L. Xu, Z. Shi, M. Wang, and W. Wu, “Techniques, process, and enterprise solutions of business intelligence,” *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, vol. 6, pp. 4722–4726, 2006.
- [52] R. Khan and S. Quadri, “Business intelligence: an integrated approach,” *Business Intelligence Journal*, vol. 5, no. 1, pp. 64–70, 2012. [Online]. Available: http://www.saycocorporativo.com/saycouk/BIJ/journal/Vol5No1/BJournal_January2012.pdf#page=68
- [53] J. Ranjan, “Business Intelligence: Concepts, Components, Techniques and Benefits,” *Journal of Theoretical and Applied Information Technology*, vol. 9, p. 60, 2009. [Online]. Available: <http://jatit.org/volumes/research-papers/Vol9No1/9Vol9No1.pdf>
- [54] V. T. U. o. T. Pirttimäki, “Business Intelligence Conceptual Model,” *South African Journal of Information Management*, vol. 9, no. 2, pp. 48–66, 2007.

- [55] D. Subotić, P. Pošćić, and V. Slavuj, “Olap Tools in Education,” *Media, Culture and Public Relations*, vol. 4, no. 1, pp. 34–44, 2013.
- [56] W. H. Inmon, *Building the Data Warehouse*, 3rd ed., J. . Wiley and I. Sons, Eds. New York, NY, USA: Robert Ipsen, 2002.
- [57] D. J. Power, *Decision Support Systems: Concepts and Resources for Managers*, DSSResources.COM, Ed. Westport, CN: Quorum Books division Greenwood Publishing, 2002.
- [58] 1keydata.com, “Data Warehouse Definition.” [Online]. Available: <https://www.1keydata.com/datawarehousing/data-warehouse-definition.html>
- [59] R. Kimball, L. Reeves, M. Ross, and W. Thornthwaite, *The Data Warehouse Lifecycle Toolkit: Expert Methods for Designing, Developing, and Deploying Data Warehouses*, 1st ed. New York, NY, USA: John Wiley & Sons, Inc., 1998. [Online]. Available: <http://books.google.com/books?hl=fr&lr=&id=abEwJJLewDAC&pgis=1>
- [60] M. Shariat and R. Hightower, “Conceptualizing Business Intelligence Architecture,” *Marketing Management Journal*, pp. 40–47, 2007. [Online]. Available: http://department.utcc.ac.th/library/images/stories/file/Database/EBSCO_BSC/Database39.pdf
- [61] I. H. Seyrek, “The use of Online Analytical Processing (OLAP) for Business Intelligence,” *Lex ET Scientia International Journal*, vol. 86, pp. 229–238, 2007. [Online]. Available: <https://doaj.org/article/ad489e899c7546779649dba8bc6ee3a9>
- [62] M. Alsqour, K. Matouk, and M. Owoc, “A survey of data warehouse architectures - preliminary results,” *Proceedings of the Federated Conference on Computer Science and Information Systems*, pp. 1121–1126, 2012.
- [63] A. Palitot, “Análise Comparativa de Arquiteturas de distribuição de Data Warehouse,” Ph.D. dissertation, Universidade Federal de Pernambuco, 2007.
- [64] R. Kimball and M. Ross, *The Data Warehouse Toolkit: The Complete guide to Dimensional Modeling*, 2nd ed., R. Elliott, Ed. New York, NY, USA: John Wiley & Sons, Inc., 2002.
- [65] M. Golfarelli, D. Maio, and S. Rizzi, “Conceptual design of data warehouses from E/R schemes,” *Proceedings of the Thirty-First Hawaii International Conference on System Sciences*, vol. 7, pp. 334–343, 1998. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=649228>
- [66] M.-C. Wu and A. P. Buchmann, “Research Issues in Data Warehousing.” *Btw*, pp. 61–82, 1997. [Online]. Available: <http://dblp.uni-trier.de/db/conf/btw/btw97.html#WuB97>

- [67] V. Saxena and A. Pratap, “Olap Cube Representation for Object-Oriented Database,” *International Journal of Software Engineering & Applications*, vol. 3, no. 2, pp. 109–117, 2012.
- [68] D. E. Codreanu, “Modelling Financial-Accounting Decisions by Means of OLAP Tools,” *Database Systems Journal*, vol. II, no. 1, pp. 23–32, 2011. [Online]. Available: http://www.dbjournal.ro/archive/3/3_Diana_Codreanu.pdf
- [69] R. Cramer, “Estudo analítico de ferramentas open source para Ambientes OLAP,” Ph.D. dissertation, Universidade do Extremo Sul Catarinense - UNESC, 2006. [Online]. Available: http://www.ambientelivre.com.br/downloads/doc_details/85-estudo-analitico-de-ferramentas-open-source-para-ambientes-olap.html
- [70] C. M. Santos, F. J. Marcondes, and R. V. Moravia, “Tool Adoption Management Online Analytical Processing to Assist in Business Management,” *Revista Pensar Tecnologia*, vol. v.4, p. 20, 2015. [Online]. Available: http://revistapensar.com.br/tecnologia/pasta_upload/artigos/a90.pdf
- [71] K. Tate, Z. Lili, and X. Zhong, “A Comparison of OLAP and Coherence Large Scale Aggregations,” Faculty of Worcester Polytechnic Institute, Worcester, Estados Unidos, Tech. Rep. January, 2015. [Online]. Available: https://www.wpi.edu/Pubs/E-project/Available/E-project-032315-163852/unrestricted/Final_Report_-_BNP_with_Appendix.pdf
- [72] a. Arriaga and P. Marques, “Web OLAP, Exploração de Sistemas OLAP via Web,” Departamento de Informática, Universidade do Minho, Braga, Portugal, Tech. Rep., 2009. [Online]. Available: <http://www3.di.uminho.pt/~prh/uce15-0809/g15.pdf>
- [73] O. Council, “OLAP AND OLAP Server Definitions,” 1997. [Online]. Available: <http://www.olapcouncil.org/research/glossaryly.htm>
- [74] E. Thomsen, *OLAP Solutions: Building Multidimensional Information Systems, Second Edition*, 2nd ed. New York, NY, USA: John Wiley & Sons, Inc., 2002.
- [75] A. Sukhoo, M. Soobron, and R. Soodin, “Open Source Software Adoption in Mauritius,” *IIMC International Information Management Corporation*, pp. 1–10, 2013.
- [76] D. Xia, X. Xie, and Y. Xu, “Web GIS server solutions using open-source software,” *OSSC-2009 - Proceedings of 2009 IEEE International Workshop on Open-source Software for Scientific Computation*, pp. 135–138, 2009.
- [77] Y. H. Tung, C. J. Chuang, and H. L. Shan, “A framework of code reuse in open source software,” *APNOMS 2014 - 16th Asia-Pacific Network Operations and Management Symposium*, 2014.

- [78] Spagobi, “SpagoBI - Vision.” [Online]. Available: <http://www.spagobi.org/homepage/opensource/vision/>
- [79] Engineering Ingegneria Informatica S.p.A., “SpagoBI - License,” 2016. [Online]. Available: <http://www.spagobi.org/homepage/opensource/license/>
- [80] M. K. O. de Lacy, “SpagoBI - Plataforma BI livre e aberta,” *Revista espírito livre*, p. 95, Jun. 2010. [Online]. Available: http://revista.espiritolivres.org/pdf/Revista_EspiritoLivre_015_junho2010.pdf
- [81] C. Thomsen and T. B. Pedersen, “A Survey of Open Source Tools for Business Intelligence,” *International Journal of Data Warehousing and Mining*, vol. 5, no. 3, pp. 56–75, 2008.
- [82] S. Coelho and N. V. Lopes, “Online platform for case studies in smart cities,” in *2016 International Multidisciplinary Conference on Computer and Energy Science (SpliTech)*, University of Split and FESB, Eds. IEEE, 2016, pp. 1–6. [Online]. Available: <http://ieeexplore.ieee.org/document/7555941/>