

Plataforma de Dados de Saúde

Portal Institucional

Paulo Jorge da Rocha Sá

**Dissertação para obtenção do Grau de Mestre em
Engenharia Informática, Área de Especialização em
Tecnologias do Conhecimento e Decisão**

Orientador: Dr^a Fátima Rodrigues

Co-orientador: Eng. Diogo Reis

Júri:

Presidente:

Doutor José António Reis Tavares, DEI/ISEP

Vogais:

Doutor Paulo Jorge Machado Oliveira, DEI/ISEP

Doutora Maria de Fátima Coutinho Rodrigues, DEI/ISEP

Engenheiro Diogo Reis, SPMS

Porto, Outubro de 2013

Dedicatória

Dedico esta dissertação aos dois grandes mestres que me têm acompanhado durante esta longa jornada, os meus pais.

À minha esposa que mesmo tendo sido obrigada a abdicar dos preciosos momentos a dois, foi o meu suporte e fonte de motivação nas diversas fases deste projeto.

À restante família e amigos pela força, companhia e doses de incentivo.

Por fim, às minhas duas gatas que foram uma companhia incansável e inconscientemente me forneceram uma dose importante de companhia, alegria e felicidade.

“Por vezes sentimos que aquilo que fazemos não é senão uma gota de água no mar. Mas o mar seria menor se lhe faltasse uma gota.”

Madre Teresa de Calcutá

Resumo

O presente documento de dissertação retrata o desenvolvimento do projeto PDS-Portal Institucional cujo cerne é um sistema para recolha, armazenamento e análise de dados (plataforma de *Business Intelligence*). Este portal está enquadrado na área da saúde e é uma peça fundamental no sistema da Plataforma de dados da Saúde, que é constituído por quatro portais distintos. Esta plataforma tem como base um sistema totalmente centrado no utente, que agrega dados de saúde dos utentes e distribui pelos diversos intervenientes: utente, profissionais de saúde nacionais e internacionais e organizações de saúde.

O objetivo principal deste projeto é o desenvolvimento do PDS-Portal Institucional, recorrendo a uma plataforma de *Business Intelligence*, com o intuito de potenciar os utilizadores de uma ferramenta analítica para análise de dados.

Estando a informação armazenada em dois dos portais da Plataforma de dados da Saúde (PDS-Portal Utente e PDS-Portal Profissional), é necessário modular um armazém de dados que agregue a informação de ambos e, através do PDS-PI, distribua um conjunto de análises ao utilizador final. Para tal este sistema contempla um mecanismo totalmente automatizado para extração, tratamento e carregamento de dados para o armazém central, assim como uma plataforma de BI que disponibiliza os dados armazenados sobre a forma de análises específicas. Esta plataforma permite uma evolução constante e é extremamente flexível, pois fornece um mecanismo de gestão de utilizadores e perfis, assim como capacita o utilizador de um ambiente Web para análise de dados, permitindo a partilha e acesso a partir de dispositivos móveis.

Após a implementação deste sistema foi possível explorar os dados e tirar diversas conclusões que são de extrema importância tanto para a evolução da PDS como para os métodos de praticar os cuidados de saúde em Portugal.

Por fim são identificados alguns pontos de melhoria do sistema atual e delineada uma perspetiva de evolução futura. É certo que a partir do momento que este projeto seja lançado para produção, novas oportunidades surgirão e o contributo dos utilizadores será útil para evoluir o sistema progressivamente.

Palavras-chave: Plataforma de Dados da Saúde, ETL, *Business Intelligence*, Análise de dados.

Abstract

The current document essay portrays the development of the project ENP – Institutional Portal which core represents a system to collect, storage and deliver to end-user summarized data (Business Intelligence platform). This portal is checked in the healthcare business and is a fundamental piece in the eHealth National Platform which is composed of four distinct portals. This platform has the principle of a system totally centered on the patient which gather clinical data to deliver to the several stakeholders: patient, national and international health professionals and health organizations.

The primary prupose of this project is the development of the ENP – Institutional Portal, using a Business Intelligence platform to empower end-user of an analytical tool for data analysis.

Being the information stored in two portals of the eHealth National Platform (ENP – Patient Portal and ENP – Clinical Portal), it's required to modulate a data warehouse which gather data of both and with ENP– Institutional Portal, deliver a set of analysis to the end-user. For that, this system contemplates a mechanism totally automated for extracting, transforming and loading data into a data warehouse and also a Business Intelligence Platform that deliver the stored data in a form of specific analysis. This platform allows a constant evolution and is extremely flexible in providing a mechanism to manage users and profiles, as well as enables the end-user of a web environment for data analysis, sharing and access using mobile devices.

After the implementation of this system it was possible to explore data and take several conclusions that are extremely important both for the evolution of ENP as to the methods of practicing healthcare in Portugal.

Finally are identified some improvement matters about the current system and defined a perspective of its future evolution. It is certain that when this project will come up for production, new opportunities will arise and the user's feedback will be useful for the evolution of the system.

Keywords: eHealth National Platform, ETL, Business Intelligence, Data analysis.

Agradecimentos

Este projeto, PDS – Portal Institucional, só foi possível devido a um trabalho incansável da equipa da Plataforma de Dados da Saúde, que permitiu evoluir os métodos de praticar saúde em Portugal e graças ao sucesso de implementação dos portais PDS-Portal do Profissional e PDS-Portal do Utente, que originaram a necessidade deste projeto. Uma equipa exemplar tendo como pontos fortes: união, competência, eficiência e um forte carácter inovador.

Um grande obrigado e abraço ao Eng.º Diogo Reis, por acreditar no meu potencial e ter fornecido todas as condições necessárias para a realização deste projeto.

À minha orientadora Prof.ª Dr.ª Fátima Rodrigues, o meu agradecimento pelo incentivo, orientação, disponibilidade e conselhos dispensados ao longo de todo o processo.

Agradeço à SPMS – Serviços Partilhados do Ministério da Saúde, por ter confiado na minha pessoa para a realização de um projeto de extrema importância para a empresa e respetivos *stakeholders*.

Índice

1	Introdução	1
1.1	O desafio em mãos	2
1.2	Organização do documento	4
2	Contextualização do projeto	5
2.1	Influência dos SI's de <i>Business Intelligence</i> na Saúde	5
2.2	Implementações de sucesso à escala nacional	7
2.2.1	Grupo de Diagnósticos Homogéneos (GDH)	7
2.2.2	Sistema Integrado de Gestão de Inscritos para Cirurgia (SIGLIC)	11
3	Plataforma de Dados de Saúde	19
3.1	Sobre a génese	20
3.2	Portal do Utente - Aproximar o utente ao SNS	22
3.2.1	Arquitetura, Segurança e Proteção de dados	25
3.2.2	Conhecimento para análise	27
3.3	Portal do Profissional - Partilhar para melhor atender	29
3.3.1	Arquitetura, Segurança e Proteção de dados	31
3.3.2	Conhecimento para análise	33
4	Portal Institucional - Analisar para potenciar	35
4.1	Estado atual do portal	35
4.2	Arquitetura proposta	37
4.3	Análise e modelação dimensional	39
4.4	Extrair, Transformar e Armazenar (ETL)	44
4.4.1	Visão geral	45
4.4.2	Extração de dados	45
4.4.3	Implementação da área de preparação	47
4.4.4	Criação do modelo dimensional	48
4.4.5	Tabelas dimensão	48
4.4.6	Tabelas facto	52
4.4.7	Estimativa de crescimento do armazém de dados	58
4.5	Apresentação da informação	59
4.5.1	Gestão de utilizadores e perfis	62
5	Exploração de dados	65
5.1	PDS - Portal Profissional	66
5.2	PDS - Portal Utente	75
6	Conclusões	85

6.1	Limitações e contratempos	86
6.2	Perspetivar o futuro.....	87
ANEXOS.....		91
Anexo 1	- Exemplo de procura de dados numa tabela dimensão	91
Anexo 2	- Exemplos de transformação de dados	92
Anexo 3	- Exemplo de procedimento para carregamento de dimensões do tipo 2.....	93
Anexo 4	- Exemplo de procedimento para carregar dados numa dimensão <i>junk</i>	94
Anexo 5	- Tarefas de carregamento das tabelas de facto	95
Anexo 6	- Microstrategy Architect	97
Anexo 7	- <i>Microstrategy Web</i> (PDS-PI)	98

Lista de Figuras

Figura 1 – Portais da PDS.....	3
Figura 2 – Exemplo de Arquitetura de um Sistema de <i>Business Intelligence</i>	6
Figura 3 – Arquitetura do sistema de BI aplicado ao GDH.....	10
Figura 4 – Modelo em floco de neve do DW aplicado aos GDH.....	11
Figura 5 – Exemplo de análise GDH disponibilizada.....	11
Figura 6 – Efeitos perversos das LIC e tempos de espera excessivos.....	12
Figura 7 – Coerência e consistência nos sistemas de informação.....	14
Figura 8 – Funções do SIGLIC.....	15
Figura 9 – SIGLIC - Integração de dados.....	16
Figura 10 – SIGLIC – Extrações mensais.....	17
Figura 11 – SIGLIC – País: Evolução semanal da LIC, entradas e operados em 2011.....	18
Figura 12 – SIGLIC – Grupo de serviço: Características da LIC no ano de 2011.....	18
Figura 13 – Desenho da PDS como serviço para diferentes <i>stakeholders</i>	21
Figura 14 – Plataforma de Dados de Saúde - Portais.....	22
Figura 15 – PDS-PU – Novo portal lançado a 7 de Maio de 2013.....	23
Figura 16 – PDS-PU – Área pessoal para utilizadores autenticados.....	24
Figura 17 – PDS-PU – Representação da arquitetura PDS-PU.....	25
Figura 18 – PDS-PU – Aplicações protegidas pelo mecanismo SSO.....	27
Figura 19 – PDS-PU – Identificação das entidades relevantes para extração de dados.....	28
Figura 20 – PDS-PP – Mapa de instituições.....	29
Figura 21 – PDS-PP – Cronograma de contatos do utente nas diversas instituições.....	30
Figura 22 – PDS-PP – Representação da arquitetura PDS-PP.....	31
Figura 23 – PDS-PP – Processo de validação de acessos.....	32
Figura 24 – PDS-PP – Ligação segura através da RIS.....	33
Figura 25 – PDS-PP – Identificação das entidades relevantes para extração de dados.....	34
Figura 26 – PDS-PI – Primeira implementação do portal.....	36
Figura 27 – PDS-PI – Acessos de Junho de 2013 [SPMS, 2013].....	36
Figura 28 – PDS-PI – Arquitetura do armazém de dados (DW).....	38
Figura 29 – Modelação dimensional do repositório de acessos ao PDS-PP.....	40
Figura 30 – Modelação dimensional do repositório de operações no PDS-PP.....	40
Figura 31 – Modelação dimensional do repositório de contatos.....	41
Figura 32 – PDS-PU – Modelação dimensional do repositório de alergias.....	42
Figura 33 – PDS-PU – Modelação dimensional do repositório de doenças.....	42
Figura 34 – PDS-PU – Modelação dimensional do repositório de medicação.....	43
Figura 35 – ETL – Repositório do projeto PDS-PI.....	45
Figura 36 – ETL – Visão geral do processo.....	45
Figura 37 – ETL – Processo de extração de dados do PDS-PU.....	46
Figura 38 – ETL – Tarefa de extração do PDS-PU.....	46
Figura 39 – ETL – Processo de criação e carregamento da área de estágio.....	47
Figura 40 – ETL – Tarefa de carregamento das tabelas de preparação.....	47

Figura 41 – ETL – Processo de criação do modelo dimensional	48
Figura 42 – ETL – Processo de carregamento das dimensões	49
Figura 43 – ETL – Tarefa de carregamento da dimensão instituição	50
Figura 44 – ETL – Tarefa de carregamento da dimensão utente	51
Figura 45 – ETL – Tarefa de carregamento das dimensões data e hora	51
Figura 46 – ETL – Processo de carregamento das tabelas facto	52
Figura 47 – ETL – Carregamento da tabela de facto FactPDSPPEntries	53
Figura 48 – ETL – Especificidade do carregamento da tabela de facto FactPDSPPOperations ..	54
Figura 49 – ETL – Carregamento da tabela de facto FactPDSPUAllergies	56
Figura 50 – Configuração da instalação do <i>Microstrategy</i> em quatro níveis	60
Figura 51 – Componente <i>Microstrategy Desktop</i> do PDS-PI	61
Figura 52 – <i>Microstrategy Architect</i> – Criação de hierarquias	61
Figura 53 – <i>Microstrategy Web</i> - Ecrã de autenticação do Portal Institucional	62
Figura 54 – PDS - Portal Institucional – Criação de grupos de utilizadores	63
Figura 55 – PDS - Portal Institucional – Criação de utilizadores do grupo ARS	63
Figura 56 – PDS-PP – Total de acessos por ARS	66
Figura 57 – PDS-PP – Volume de acesso por tipo de instituição e categoria profissional	67
Figura 58 – PDS-PP – Total de acessos por hora e fase do dia	67
Figura 59 – PDS-PP – Total de acesso por tipo de episódio	68
Figura 60 – PDS-PP – Total de acessos por aplicações externas	68
Figura 61 – PDS-PP – Total de acessos por género e idade do utente	69
Figura 62 – PDS-PP – Total de operações executadas (restantes)	70
Figura 63 – PDS-PP – ARS Norte - rede de referência com o CHP	71
Figura 64 – PDS-PP – Restantes regiões – rede de referência com o CHP	71
Figura 65 – PDS-PP – Top 10 Normas Orientação Clínica consultadas por género do utente ...	72
Figura 66 – PDS-PP – Total de episódios registados por ARS	73
Figura 67 – PDS-PP – Total de registos/duração média de episódios por tipo de episódio	73
Figura 68 – PDS-PP – Top 10 duração média (dias) dos episódios de internamento por especialidade	74
Figura 69 – PDS-PP – Total de episódios de urgência por género e idade do utente	74
Figura 70 – PDS-PU – Evolução de utentes registados no portal	75
Figura 71 – PDS-PU – Total de utentes registados por distrito	75
Figura 72 – PDS-PU – Total de utentes registados por género e idade	76
Figura 73 – PDS-PU – Total de alergias registadas por tipo	77
Figura 74 – PDS-PU – Top 10 reações alérgicas a outras substâncias/agentes por severidade ..	77
Figura 75 – PDS-PU – Top 10 alergias medicamentosas por severidade	78
Figura 76 – PDS-PU – Alergias registadas por idade e severidade	78
Figura 77 – PDS-PU – Alergias registadas por tipo de reação alérgica	79
Figura 78 – PDS-PU – Top 10 medicação registada por utentes femininos	80
Figura 79 – PDS-PU – Top 10 medicação registada por utentes masculinos	80
Figura 80 – PDS-PU – Medicação registada por perfil de toma (exceto medicação crónica)	81
Figura 81 – PDS-PU – Medicação crónica registada por perfil do utente	82
Figura 82 – PDS-PU – Top 10 doenças mais registadas	82

Figura 83 – PDS-PU – Top 10 tipos de doença mais registados por género do utente.....	83
Figura 84 – PDS-PU – Volume de doenças do aparelho circulatório registadas.....	84
Figura 85 – PDS-PU – Total de doenças de cancro por estado e idade do utente	84
Figura 86 – ETL – Exemplo de procedimento de procura de dados numa tabela	91
Figura 87 – ETL – Exemplo de procedimento mapeamento de valores.....	92
Figura 88 – ETL – Exemplo de procedimento de seleção e normalização de campos.....	92
Figura 89 – ETL – Exemplo de procedimento para remoção de duplicados	92
Figura 90 – ETL – Exemplo de procedimento de carregamento de dimensões do tipo 2.....	93
Figura 91 – ETL – Exemplo de procedimento de carregamento de uma dimensão <i>junk</i>	94
Figura 92 – ETL – Carregamento da tabela de facto FactPDSPPOperations	95
Figura 93 – ETL – Carregamento da tabela de facto FactPDSPPContacts	95
Figura 94 – ETL – Carregamento da tabela de facto FactPDSPU Medications	96
Figura 95 – ETL – Carregamento da tabela de facto FactPDSPUPathologies	96
Figura 96 – <i>Microstrategy Architect</i> – Definição da camada de Alergias.....	97
Figura 97 – <i>Microstrategy Web</i> – Interface PDS-PI.....	98
Figura 98 – <i>Microstrategy Web</i> – Configuração de perfis.....	99

Lista de Tabelas

Tabela 1 – PDS-PU – Serviços disponibilizados.....	23
Tabela 2 – PDS-PU – Entidades principais	28
Tabela 3 – PDS-PP – Entidades principais	34
Tabela 4 – Metodologia de <i>Kimball</i> – Repositório de acessos ao PDS-PP	40
Tabela 5 – Metodologia de <i>Kimball</i> – Repositório de operações no PDS-PP	40
Tabela 6 – Metodologia de <i>Kimball</i> – Repositório de contatos.....	41
Tabela 7 – Metodologia de <i>Kimball</i> – PDS-PU repositório de alergias	41
Tabela 8 – Metodologia de <i>Kimball</i> – PDS-PU repositório de doenças	42
Tabela 9 – Metodologia de <i>Kimball</i> – PDS-PU repositório de medicação.....	43
Tabela 10 – Matriz em bus do modelo dimensional.....	43
Tabela 11 – ETL – Identificação das dimensões SCD Tipo 1	49
Tabela 12 – ETL – Tipo SCD dos atributos descritivos da dimensão DimPatient.....	50
Tabela 13 – Detalhes da tabela de facto FactPDSPPEntries.....	53
Tabela 14 – Detalhes da tabela de facto FactPDSPPOperations	54
Tabela 15 – Detalhes da tabela de facto FactPDSPPContacts	55
Tabela 16 – Detalhes da tabela de facto FactPDSPUAllergies.....	56
Tabela 17 – Detalhes da tabela de facto FactPDSPUMedications	57
Tabela 18 – Detalhes da tabela de facto FactPDSPUPathologies	57
Tabela 19 – PDS-PI – Volume de dados inicial e estimativa de crescimento mensal	58
Tabela 20 – Total de autorizações configuradas pelos utentes	76

Acrónimos e Símbolos

Lista de Acrónimos

ACES	Agrupamento de Centros de Saúde
ACSS	Administração Central do Sistema de Saúde, IP
ARS	Administração Regional de Saúde
BDNP	Base de Dados Nacional de Prescrições
BI	<i>Business Intelligence</i>
CNPD	Comissão Nacional de Proteção de Dados
CSSV	Cirurgia Segura Salva Vidas
CHP	Centro Hospitalar do Porto
CIC	Comissão de Informatização Clínica
DGS	Direção Geral de Saúde
EHR	<i>Electronic Health Record</i>
EPSOS	European Patients - Smart Open Services
ETL	Extract, Transform, Load
GDH	Grupos de Diagnósticos Homogéneos
GID	Gestão Integrada da Doença
INEM	Instituto Nacional de Emergência Médica
INFARMED	Autoridade Nacional do Medicamento e Produtos de Saúde, IP
LIC	Lista de Inscritos para Cirurgia
MCDT	Meios Complementares de Diagnóstico e Tratamento
NOC	Norma de Orientação Clínica
OCDE	Organização para a Cooperação e o Desenvolvimento Económico
PCE	Processo Clínico Eletrónico
PDS	Plataforma de Dados de Saúde

PDS-EPSOS	PDS – Portal Internacional
PDS-PI	PDS – Portal Institucional
PDS-PP	PDS – Portal Profissional
PDS-PU	PDS – Portal Utente
PECLEC	Programa Especial de Combate às Listas de Espera Cirúrgicas
PERLE	Programa Especial de Recuperação das Listas de Espera
PNV	Plano Nacional de Vacinação
PHR	<i>Patient Health Record</i>
PPA	Programa de Promoção do Acesso
PPMA	Programa de Promoção da Melhoria do Acesso
RCU²	Resumo Clínico Único do Utente
RICA	Repositório de Informação Clínica Anonimizada
RNCCI	Rede Nacional de Cuidados Continuados Integrados
RNE	Registo Nacional de Entidades
RNP	Registo Nacional de Profissionais
RNU	Registo Nacional de Utentes
SAM	Sistema de Apoio ao Médico
SAPE	Sistema de Apoio à Prática de Enfermagem
SIGIC	Sistema Integrado de Gestão de Inscritos para Cirurgia
SIGLIC	Sistema Informático de Gestão da Lista de Inscritos para Cirurgia
SIH	Sistema de Informação Hospitalar
SNS	Serviço Nacional de Saúde
SPMS	Serviços Partilhados do Ministério da Saúde
SSO	<i>Single-Sign On</i>
TCD	Tecnologias do Conhecimento e Decisão
TIC	Tecnologias Informação e Comunicação

1 Introdução

As Tecnologias da Informação e Comunicação (TIC) estão a alterar, de modo rápido e profundo, a prática das organizações de saúde e o paradigma da relação entre os profissionais e os utilizadores dos serviços de saúde [Manuel Pizarro, 2011].

Nos países desenvolvidos estão a decorrer um conjunto vasto de projetos que irão contribuir para que a incorporação das TIC nos sistemas de saúde seja potenciada.

Atualmente existem registos clínicos informatizados correspondentes a mais de sete milhões de portugueses. Neste domínio, a Plataforma de Dados de Saúde (PDS) constitui um elemento estruturante fundamental pelo serviço que presta às pessoas e aos profissionais. Esta plataforma assegura um imenso valor instrumental, pois aproxima os diferentes intervenientes no sistema e disponibiliza uma rede de referência comum que induz um progresso mais rápido e bem-sucedido na prestação de cuidados de saúde. A plataforma é constituída por quatro portais completamente distintos, que permitem a partilha de informação com diversos *stakeholders*.

No âmbito da dissertação de mestrado do curso de Engenharia Informática, na área das Tecnologias do Conhecimento e Decisão, do Instituto Superior de Engenharia do Porto, foi proposto o desenvolvimento e implementação de um sistema de *Business Intelligence* (BI) denominado PDS - Portal Institucional (PDS-PI). Este portal é uma peça fundamental na PDS, pois permite a análise de dados provenientes dos restantes portais que a constitui. Como tal, após a análise dos portais foram desenvolvidos mecanismos de extração, tratamento e carregamento de dados para repositórios centrais. O PDS-PI deverá ter interação com o utilizador final através de uma aplicação que é disponibilizada a um conjunto específico de utilizadores por organização e/ou instituição de saúde, e com diferentes níveis de acesso à informação. Após o acesso autorizado, deverão ser disponibilizadas ferramentas úteis para a análise de dados: painel de indicadores (*dashboards*), relatórios *ad-hoc* e ferramentas de análise ágil (por exemplo recorrendo a *Pivot Tables*).

Segundo [Manuel Pizarro, 2011], as linhas orientadoras para o futuro dos sistemas de informação na saúde são as seguintes:

1. Incorporar nos sistemas de informação novas ferramentas de apoio à decisão, cada vez mais refinadas. Face ao crescimento exponencial do conhecimento, deverá ser

1 Introdução

assegurada uma maior qualidade de prestação de cuidados de saúde a que o cidadão tem direito;

2. O tratamento de doenças, nomeadamente em idosos e indivíduos com doenças crónicas, constituem uma ameaça para os ganhos em saúde conseguidos nas últimas décadas. As TIC, através de programas de saúde adequados, poderá estimular a ação de cada utente e modificar a natureza da relação com os profissionais.
3. A visão das TIC como instrumento para promover a inovação. Intensificar o seu uso na saúde permitirá alargar o acesso dos cidadãos, potenciar os profissionais e administradores de saúde por forma a melhorar a qualidade dos serviços e aumentar a eficiência do Sistema Nacional de Saúde (SNS).

O projeto proposto nesta dissertação de mestrado está perfeitamente alinhado com dois dos três aspetos mencionados anteriormente:

- A partir do conhecimento gerado pelos portais da PDS é disponibilizada uma ferramenta que permite apoiar as organizações e/ou instituições de saúde, no processo de apoio à decisão e potenciar a qualidade dos serviços prestados.
- Este projeto representa um compromisso com a evolução tecnológica, para que se construa uma visão holística que suporte a complexa realidade de uma rede de sistemas integrados, garantindo que são proporcionados serviços simples e de qualidade para os profissionais e administradores do SNS.

1.1 O desafio em mãos

A Plataforma de Dados de Saúde (PDS) é um sistema de partilha de dados de saúde, que permite que a mesma informação seja fragmentada e partilhada aos diferentes agentes da prestação de cuidados de saúde (utentes, profissionais vinculados ao Serviço Nacional de Saúde (SNS) e/ou convencionados). Os dados são acedidos através de portais específicos, seguros e contextualizados, a partir das instituições locais onde permanecem guardados. Como retratado na Figura 1, os portais mencionados são os seguintes:

- PDS - Portal do Utente (PDS-PU) – Portal destinado a todos os utentes portugueses, disponibilizando um conjunto de serviços de saúde informativos e eletrónicos;
- PDS - Portal do Profissional (PDS-PP) – Portal destinado a todos os profissionais prestadores de cuidados de saúde (médicos e enfermeiros), que exerçam atividade em instituições de saúde públicas e convencionados, colocando à disposição num único local o acesso à rede de instituições do país e consulta de dados clínicos provenientes dos sistemas locais;
- PDS - Portal Internacional (PDS-EPSOS) – Portal destinado a profissionais médicos, que permite disponibilizar serviços transfronteiriços de modo a assegurar cuidados de saúde eficientes e seguros, aos cidadãos europeus quando viajam pela Europa;

- PDS - Portal Institucional (PDS-PI) – Portal destinado a instituições de saúde, administrações regionais de saúde, agrupamentos de centros de saúde e direção geral de saúde, para análise e auditoria de dados sobre os diversos portais mencionados anteriormente.



Figura 1 – Portais da PDS

Após o lançamento do PDS-PP, que decorreu em Julho de 2012, colocou-se a necessidade de analisar os acessos e a interação dos utilizadores com a aplicação, tendo como principal objetivo avaliar a utilização do portal nas diversas instituições hospitalares e de cuidados de saúde primários. Uma versão simplificada do PDS - Portal Institucional (PDS-PI), lançado em Novembro de 2012, permitiu responder a esta necessidade prioritária. Até à data, este portal é acedido por administradores da PDS e um conjunto limitado de responsáveis hospitalares, com o objetivo primordial de consulta e extração de dados.

A implementação do PDS-PI versão simplificada, foi realizado na mesma infraestrutura aplicacional do PDS-PP, e desde então tem sido utilizado em menor escala devido às seguintes limitações:

- Incapacidade de evoluir por estar limitado a dois perfis de utilizador estáticos, não permitindo a criação de novos perfis;
- Problemas de performance, pois a consulta de dados é realizada diretamente sobre as bases de dados operacionais;
- Problemas de escalabilidade, pois os dados desta aplicação estão armazenados na base de dados operacional do PDS-PP, o que impossibilita a autonomia do armazenamento de dados analíticos e a respetiva gestão.

Após a avaliação das limitações mencionadas, foi decidido em grupo de trabalho da PDS, reestruturar este projeto e descontinuar a versão simplificada, garantido numa fase inicial o acesso aos mesmos dados que a versão anterior disponibilizava. Para tal, é necessário a disponibilização de uma infraestrutura independente e virtualizada, que permita a implementação de um conjunto de repositórios centrais e respetivos mecanismos de extração, tratamento e carregamento de dados, provenientes dos diversos portais constituintes da PDS (PDS-PU e PDS-PP). O PDS – Portal Internacional (PDS-EPSOS), não deverá ser incluído na próxima versão do PDS-PI devido ao facto de se tratar de um projeto ainda numa fase piloto.

1 Introdução

O PDS-PI ao estar dotado das potencialidades mencionadas anteriormente, deverá também fornecer um conjunto de ferramentas e/ou aplicações que permitam disponibilizar as seguintes funcionalidades:

- Gestão dinâmica de utilizadores e perfis de acesso à informação;
- Painéis de indicadores (*dashboards*) / Ferramenta para análise ágil de dados (por exemplo através de *Pivot Tables*) / Relatórios a pedido (*Ad-hoc*);
- Preparado para dispositivos móveis;
- Interoperável com outros repositórios centrais (i.e.: repositórios de dados clínicos).

1.2 Organização do documento

Este documento encontra-se organizado em seis capítulos. No primeiro capítulo é redigido um enquadramento do projeto. No segundo capítulo, o projeto é contextualizado tendo em conta o levantamento do estado da arte realizado nas seguintes áreas: desenvolvimento de projetos de Business Intelligence para a Saúde; implementação de um sistema baseado em grupos de diagnósticos homogéneos e implementação do sistema integrado de gestão de inscritos para cirurgia.

O terceiro capítulo, está relacionado com a contextualização da Plataforma de Dados de Saúde e os diversos portais constituintes. Para os dois portais mais relevantes são mencionados: contextualização aprofundada; arquitetura do sistema; aspetos de segurança/proteção de dados; processo de identificação da informação essencial para análise e extração de dados.

O quarto capítulo diz respeito ao trabalho de implementação desenvolvido. Inicia com o ponto de situação anterior à implementação, mencionando as principais razões que levaram à reestruturação do projeto. Após identificadas as limitações é delineada a arquitetura que será responsável por sustentar a implementação do projeto. Relativamente ao desenvolvimento do sistema, seguidamente é apresentado o processo de modelação dimensional, extração dos dados dos sistemas operacionais, tratamento e carregamento dos mesmos para as tabelas dimensão e facto. Por fim, é descrito o processo de implementação da solução tecnológica responsável pela apresentação dos dados ao utilizador final.

No quinto capítulo é efetuado um relato da atividade de exploração de dados, disponibilizados no PDS- Portal Institucional, salientando as conclusões mais relevantes para avaliação do PDS-PP e PDS-PU.

Para finalizar, no sexto capítulo são apresentadas as conclusões, as dificuldades encontradas e sugestões para o trabalho futuro a ser desenvolvido.

2 Contextualização do projeto

Após o enquadramento e apresentação do desafio, neste capítulo é efetuada uma contextualização do projeto PDS-Portal Institucional, mencionando as principais considerações no processo de análise/conceção de um armazém de dados e de ferramentas adequadas para a análise de dados. Para tal, é importante avaliar o contexto atual dos sistemas de informação na área da saúde, e recorrendo a casos práticos, efetuar uma análise dos sistemas de *Business Intelligence* (BI) implementados com sucesso a uma escala nacional. Até à data, a temática do BI aplicado à saúde, tem sido bastante debatido junto da comunidade científica, e após um estudo exaustivo, foram considerados dois casos portugueses, que foram implementados na mesma escala que o projeto PDS – Portal Institucional. Os casos apresentados constituem uma base importante para a análise deste projeto, na medida em que poderá contribuir para um desenho adequado e enquadrado com a realidade portuguesa dos sistemas de informação para a saúde.

2.1 Influência dos SI's de *Business Intelligence* na Saúde

Um Sistema de *Business Intelligence* (BI) é uma componente da arquitetura do Sistema de Informação de uma organização para suporte à tomada de decisão através da análise dos dados de um determinado negócio. É constituído por um conjunto de tecnologias, aplicações e processos para recolher, armazenar, visualizar e analisar dados que permitem aos utilizadores tomarem decisões baseadas em evidências. [Helder Quintela, 2013]

Para o sucesso de um Sistema de BI, deve ter-se em conta as seguintes recomendações [David Stodder, 2012]:

- Desenvolvimento rápido e um acesso facilitado à informação
- Disponibilizar componentes de visualização de informação robustos e interativos
- Avaliar as diversas alternativas para a integração dos dados
- Colocar nas mãos do utilizador final o poder da descoberta de conhecimento

2 Contextualização do projeto

Os sistemas de BI podem ser simples ou mais sofisticados, por exemplo, incluindo componentes de integração de dados não estruturados e utilização de técnicas de *data-mining* (i.e.: Redes Neurais Artificiais, Algoritmos de Classificação).

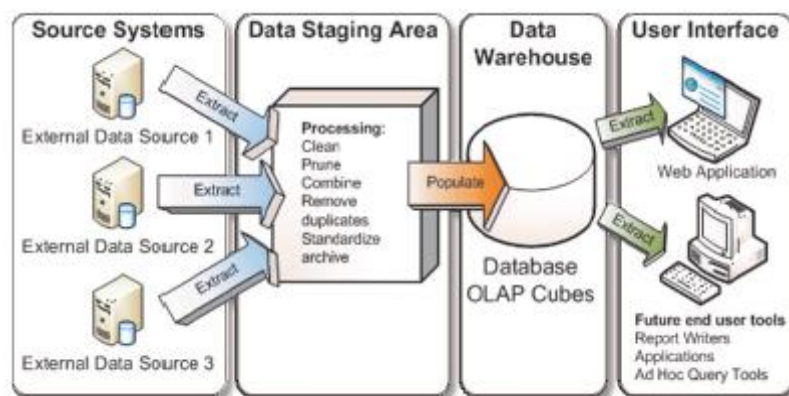


Figura 2 – Exemplo de Arquitetura de um Sistema de *Business Intelligence*
[Helder Quintela, 2013]

Como exemplificado na Figura 2, um Sistema de BI é composto tipicamente por um armazém de dados (*Data Warehouse*) para consolidação dos dados a partir de bases de dados operacionais, por um componente de carregamento e transformação de dados – ETL (que é responsável pela extração dos dados a partir das bases de dados fonte e carregamento no armazém de dados) e, por uma aplicação/plataforma analítica que permite analisar os dados (i.e.: *MicroStrategy*, *Oracle Business Intelligence*, *Microsoft Analysis Services*, *Pentaho BI*). Estas plataformas permitem a utilização de diferentes elementos para exploração de dados (i.e.: *dashboards*, relatórios tabulares, visualizações gráficas) e funcionalidades como agregação e análise multidimensional. Esta aplicação analítica é tipicamente acedida num computador, *tablet* ou dispositivo móvel, através de um *browser* ou de aplicações do tipo *Office*.

Nos últimos anos a implementação e utilização de Sistemas de BI nas organizações que prestam cuidados de saúde (*Healthcare Business Intelligence*) tem conhecido um aumento significativo, sendo que os principais fatores para a adoção de sistemas de BI [David Hatch and Michael Lock, 2008] são:

- Necessidade de melhoria na gestão de recursos;
- Melhoria na qualidade dos cuidados médicos prestados;
- Melhoria da satisfação dos pacientes;
- Necessidade de cumprir normativos legais (i.e.: Contratualização nos Cuidados Primários);
- Necessidade de atrair e reter talentos dos profissionais envolvidos na prestação de cuidados de saúde.

Segundo dados de um estudo recente da KLAS [Marianne Kolbasuk, 2012] revelam que 50% das organizações que prestam cuidados de saúde têm intenções de em curto prazo adquirir um novo sistema de BI ou substituir os sistemas atuais. Este interesse é reforçado por um inquérito promovido em 2009 pela *GARTNER* a 1500 *CIOs* (*Chief Information Officer*) que a

aquisição de sistemas de BI era a prioridade máxima ao nível dos Sistemas de Informação [Helder Quintela, 2013].

Nesta época, em que existe um domínio do conceito de "*Business Intelligence*" nos Sistemas de Informação, os repositórios centrais e os dados locais das Instituições de Saúde são dos alvos mais interessantes para a implementação e utilização de aplicações de análise de dados, pelos seguintes fatores:

1. Pela riqueza dos dados existentes;
2. Pelo potencial do conhecimento que é possível extrair dos mesmos;
3. Pelos benefícios que o conhecimento adquirido a partir da evidência dos dados, podem ter na melhoria da eficiência das instituições em diversos níveis (redução de custos, otimização de recursos, atividade e prática clínica).

Nesse sentido, existe a necessidade de olhar para os dados armazenados ao longo de vários anos, tanto nas Instituições de Saúde como nos Organismos Governamentais, e começar a tirar real partido dessa informação.

2.2 Implementações de sucesso à escala nacional

Neste ponto serão descritos dois sistemas de informação implementados em Portugal, que mesmo tendo negócios distintos, possuem repositórios centrais e ferramentas de análise de dados. Estes sistemas atuam a uma escala nacional, permitindo influenciar positivamente na rotina dos profissionais de saúde e processos das diversas instituições. Na conceção destes sistemas, foram considerados diversos exemplos internacionais de modo a enquadrar os padrões globais para a área da saúde e implementações de sucesso, por forma a auxiliar o processo de seleção das tecnologias a utilizar.

2.2.1 Grupo de Diagnósticos Homogéneos (GDH)

Os Grupos de Diagnósticos Homogéneos (GDHs) (*Diagnosis Related Groups, DRGs*) foram desenvolvidos no final da década de 1960 por investigadores americanos, em resposta aos custos crescentes dos cuidados de saúde. Os GDH proporcionam um sistema financeiro e de classificação dos doentes que utiliza o diagnóstico, tipo de tratamento, idade e outros fatores relacionados, como os critérios de triagem. É paga aos hospitais uma quantia predeterminada para o tratamento de doentes de um dado GDH, independentemente do custo real dos cuidados prestados. [CMS, 1983]

Com o aumento generalizado nos gastos com os cuidados de saúde em todo o mundo, os GDHs foram introduzidos em diversos países como estratégia de contenção de custos [Kroneman and Nagy, 2001]. Os relatórios indicam que 19 membros da Organização para a Cooperação e o Desenvolvimento Económico (OCDE) adotaram alguma forma de controlo de preços para o reembolso aos hospitais baseada no sistema GDH [Forgione DA et al., 2004]. Estes sistemas foram ainda concebidos para o planeamento, orçamento, gestão e seguimento dos cuidados prestados.

2 Contextualização do projeto

Principais vantagens de um sistema baseado no GDH:

- Os GDHs permitem uma maior transparência da gestão e financiamento do sistema hospitalar.
- Os GDHs permitem às entidades pagadoras um melhor controlo da quantidade de dinheiro gasta no reembolso aos hospitais.
- Os GDHs auxiliam as entidades pagadoras a preverem, em longo prazo, até um futuro alargado, quais serão os vencimentos financeiros dos hospitais.

Seguidamente serão descritos os requisitos fundamentais e as variáveis inerentes ao processo de classificação GDH. Este processo origina um conjunto de informação relevante para a medição do desempenho hospitalar, portanto recorrendo a um sistema de auditoria é possível avaliar e corrigir os dados. Por fim, a informação proveniente dos vários locais é resumida e armazenada num repositório central, que através de um sistema de *Business Intelligence*, permite capacitar os organismos centrais de uma ferramenta analítica para análise de dados.

Requisitos para a classificação GDH

Para o processo de classificação GDH, deverá existir um profissional codificador que necessita de possuir os seguintes conhecimentos:

- Conhecer o sistema e a estrutura de classificação CID-9-MC (Sistema de classificação de diagnósticos);
- Compreender a organização dos índices e a sua utilização na codificação das doenças e dos procedimentos;
- Aplicar, corretamente os princípios e regras da codificação da CID-9-MC;

Variáveis para a classificação GDH

No processo de classificação o codificador necessita de ter em conta as seguintes variáveis:

- Diagnóstico principal - Aquele que, após o estudo do doente, revelou ser o responsável pela sua admissão no Hospital;
- Outros diagnósticos;
- Intervenções cirúrgicas - Mesmo que um doente tenha sido submetido a múltiplas intervenções cirúrgicas relacionadas com o diagnóstico principal durante o mesmo episódio de internamento. Será agrupado num só GDH, que é definido de acordo com a hierarquia dos Procedimentos Cirúrgicos, no GDH hierarquicamente mais elevado;
- Idade - até 17 anos inclusive são aplicados GDH pediátricos, após os 17 anos são aplicados GDH adultos;
- Sexo;
- Destino após alta
 - Transferidos;
 - Saídos contra parecer do médico;
 - Falecidos

Sistema de Auditoria Interna

Através da aplicação informática "AUDITOR" [IGIF, 2005], é permitido realizar auditoria interna aos GDH's, sendo o objetivo:

- Detecção de erros de classificação (diagnósticos, procedimentos, sintomas/sinais/manifestações como diagnóstico principal) e não conformidades;
- Definição de alertas por forma a poderem ser corrigidas não conformidades detetadas;
- Amostra de dados aleatória, sendo selecionado um conjunto determinado de episódios de internamento;
- Tornar um processo sistemático;
- Dirigido a profissionais codificadores e à direção do serviço.

Informação obtida do sistema de GDH

A informação recolhida pelo sistema de classificação de doentes em GDH deve ser:

1º - Conhecida e analisada nos Serviços de Internamento pelos médicos responsáveis:

- Porque são eles que têm a sensibilidade e o conhecimento para poderem analisar, questionar e alterar os resultados;
- Detetar áreas problemáticas e integrar indicadores;
- Ajudar a melhorar a classificação dos doentes pelo rigor da informação e legibilidade do Processo Clínico.

2º - Conhecida e analisada nos Serviços Financeiros e de Gestão Hospitalar

- Análise sistemática dos indicadores de desempenho
- Divulgação dos resultados
- Reuniões de análise conjunta com todos os implicados
- Análise comparativa com outras Unidades Hospitalares comparáveis
- Detetar as áreas "Problema" e desencadear ou colaborar na resolução das mesmas
- Contratualização

Desempenho Hospitalar

Por forma a medir o desempenho hospitalar, o sistema GDH permite gerar a seguinte informação:

- Quantificação de nº de episódios de internamento e agregação por diversos campos (Região; Grupos de hospitais; GDH; Patologia; Sexo; Grupo etário);
- Indicadores de carácter clínico
 - Casos sociais;
 - Morfologia tumoral;
 - Internamentos inapropriados;
 - Percentagem de reações adversas a medicamentos;
 - Registo de todos os contactos de cada doente com o Hospital;
 - Complicações mais frequentes (Hemorragias, Infeções, entre outros);
- Indicadores para a Gestão
 - Por:
 - Serviço/Hospital/ Grupo Hospitalar;
 - Diagnóstico Principal/Procedimento/Complicação / GDH;

2 Contextualização do projeto

- Idade / Sexo / Área de residência;
- Tempo de internamento;
- Taxa ocupação / mortalidade / reinternamentos;
- Taxa de complicações (cirúrgicas, parto, recém-nascido,...);
- Indicadores para financiamento
 - GDH;
 - Índice *case-mix* - Coeficiente global de ponderação da produção, refletindo a relatividade de um Hospital face aos outros, em termos da sua maior ou menor proporção de doentes com patologias mais complexas e consequentemente mais consumidoras de recursos;
 - Nº doentes equivalentes - Nº total de episódios de internamento que se obtém, após a transformação dos dias de internamento dos doentes excepcionais e dos doentes transferidos de cada GDH, em conjuntos “equivalentes” ao tempo médio de internamento dos episódios “normais” do respetivo GDH;
- Outros indicadores
 - Dias de internamento / Demora média;
 - Nº de dias em UCI;
 - Natureza da admissão;
 - Transferência entre hospitais.

Sistema de *Business Intelligence*

Atualmente existe um sistema de *Business Intelligence* que permite realizar a extração de dados provenientes de fontes distintas e carregar informação processada para um repositório central de dados dos GDHs. Este sistema foi apresentado em 2010 à ACSS [Manuel Barrento et al., 2010], explicando todas as fases de implementação do projeto.

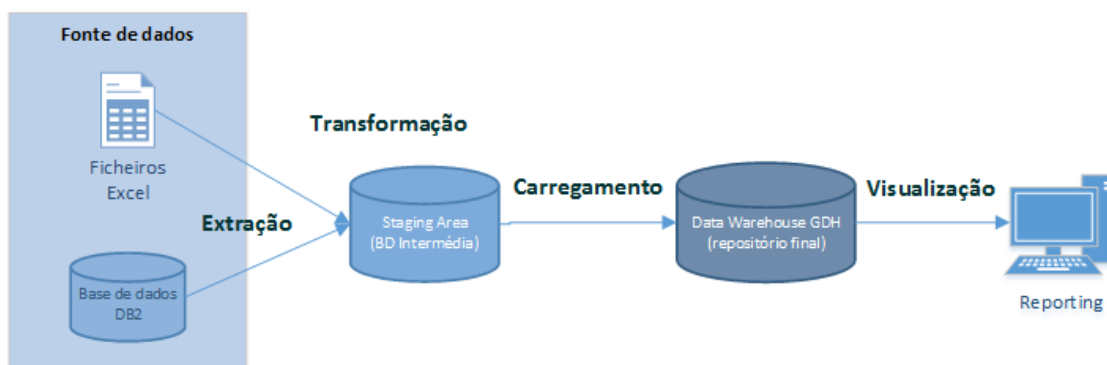


Figura 3 – Arquitetura do sistema de BI aplicado ao GDH

O repositório de dados identificado na Figura 3, denominado *Data Warehouse GDH*, é constituído por diversas tabelas de dimensões (que caracterizam os factos) e uma tabela de factos onde residem as métricas. O modelo seguido para a sua conceção foi modelo dimensional, como retratado no esquema em floco de neve na Figura 4.

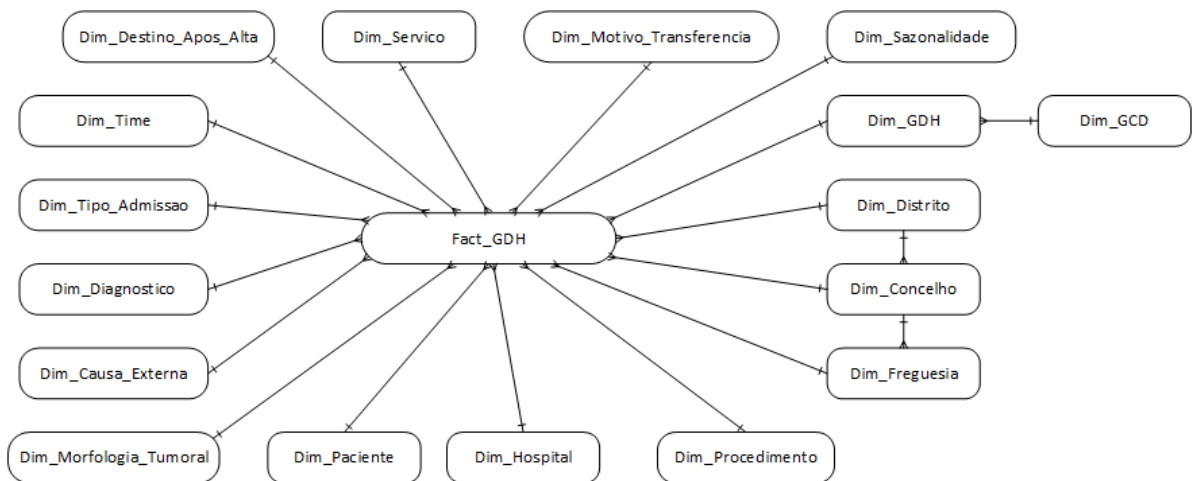


Figura 4 – Modelo em floco de neve do DW aplicado aos GDH

Para a implementação de uma aplicação de análise de dados, foi seleccionada a plataforma de BI da *Microstrategy*, com o intuito de desenvolver rapidamente relatórios *ad-hoc* e disponibilizar análises multidimensionais de forma intuitiva, como consta na Figura 5.

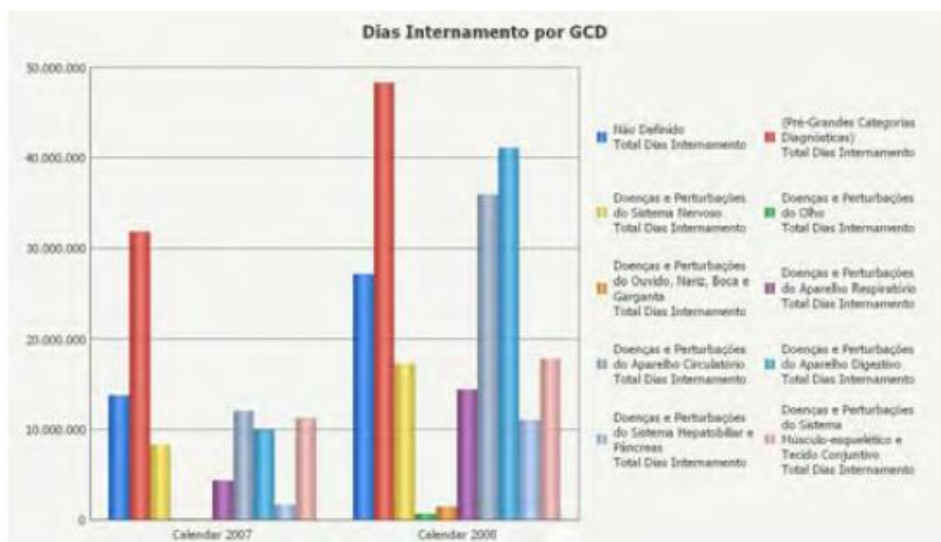


Figura 5 – Exemplo de análise GDH disponibilizada
[Manuel Barrento et al., 2010]

2.2.2 Sistema Integrado de Gestão de Inscritos para Cirurgia (SIGLIC)

Na última década, as sociedades têm vindo a demonstrar, de forma crescente, um descontentamento face ao tempo de espera para a realização de cirurgias. Em Portugal, esta

2 Contextualização do projeto

realidade deu lugar a algumas iniciativas governamentais que procuraram, através de programas especiais, diminuir o número de doentes que se encontravam em lista de espera.

Entre 1995 e 2002, foram implementados os seguintes programas [ACSS-SIGIC, 2010]:

- 1995 - Programa Especial de Recuperação das Listas de Espera (PERLE);
- 1997 - Programa de Promoção do Acesso (PPA);
- 1999 - Programa de Promoção da Melhoria do Acesso (PPMA).

Todos eles tinham genericamente em comum incluírem apenas as intervenções cirúrgicas que registavam o maior tempo de espera e o maior número de doentes em espera, até pela frequência da casuística, distinguindo-se sobretudo pela solução proposta para o efeito. Enquanto no primeiro, a solução assentava no recurso exclusivo ao sector privado mediante realização prévia de concurso público, no segundo e terceiro programa os hospitais do SNS eram os prestadores por excelência, alargando-se a prestação ao sector social só com o PPMA.

Apesar do impacto significativo destes programas, as listas de espera não paravam de aumentar e os tempos atingiam valores de espera clínica e eticamente inaceitáveis, verificando-se que os casos mais recentes rapidamente atingiam os tempos de espera dos casos mais antigos entretanto resolvidos, até com eventual agravamento futuro decorrente de uma procura em crescendo.

Na Figura 6, apresenta-se os efeitos perversos decorrentes de listas de espera de inscritos para cirurgia e tempos de espera excessivos - sobre utilização de serviços durante o tempo de espera, sofrimento acrescido, absentismo e incapacidades geradas por problemas de saúde por resolver - que podem ser minimizados com a compra de serviços ao sector privado.



Figura 6 – Efeitos perversos das LIC e tempos de espera excessivos
[ACSS-SIGIC, 2010]

Foi então aprovada pela Resolução do Conselho de Ministros nº 100/2002, de 26 de Abril a criação do Programa Especial de Combate às Listas de Espera Cirúrgicas (PECLEC), que alarga a

prestação ao sector privado mediante a celebração de convenções, com o objetivo de aumentar e melhorar o acesso dos cidadãos aos cuidados cirúrgicos.

Aproximando-se o final do PECLEC que foi criado para durar 2 anos e, nesse tempo, responder rápida e eficientemente às situações emergentes e às vozes críticas dos doentes e da sociedade, segue-se-lhe a criação do Sistema Integrado de Gestão de Inscritos para Cirurgia (SIGIC), aprovada pela Resolução do Conselho de Ministros nº 79/2004, de 3 de Junho.

Relativamente aos objetivos definidos para o SIGIC, destacam-se os seguintes:

- Reduzir o tempo de espera;
- Garantir a equidade do acesso;
- Promover a eficiência global do sistema através da otimização da gestão da Lista de Inscritos para Cirurgia e dos recursos afetos;
- Garantir a qualidade e a transparência da informação.

No período seguinte de quatro anos, o SIGIC foi implementado nas diversas regiões de saúde e por fim foi alargado às entidades do sector social e privado que prestam cuidados aos utentes do SNS ao abrigo dos acordos, contratos e convenções celebrados.

Sistema de Informação (SI)

O sistema de informação de gestão de inscritos para cirurgia, SIGLIC, foi concebido para dar resposta às necessidades de gestão de informação do SIGIC.

Para efetivar os objetivos do sistema e garantir que os serviços são prestados por profissionais credenciados no local e momento oportuno, torna-se necessário interagir nas diversas vertentes do processo:

- Gestão da procura - utentes que procuram serviços cirúrgicos;
- Registo das interações dos utentes com a instituição;
- Registo atualizado da carteira de serviços e da capacidade instalada;
- Registo dos profissionais, suas credenciais e alocação aos serviços;
- Caracterização dos utentes com determinação de riscos em programas especiais;
- Registo dos resultados obtidos em cuidados prestados;
- Acompanhamento de contratos programa e controlo de faturação;
- Observação da rede de referênciação.

Num SI existem dois sentidos principais nas preocupações: um que cumpra os objetivos para que foi concebido e outro que agregue informação com qualidade [Pedro Gomes, 2011].

2 Contextualização do projeto



Figura 7 – Coerência e consistência nos sistemas de informação
[ACSS-SIGIC, 2011a]

Assim como esquematizado na Figura 7, a consistência da informação pode ser verificada pelo SI, enquanto a coerência da mesma precisa de ser validada pelos colaboradores da instituição hospitalar que são os utilizadores da informação. O SIGLIC explora ativamente este conceito, recolhendo dados de diversos parceiros distintos (utentes, clínicos, administrativos) relativos à mesma realidade. A confrontação destes dados permite determinar a consistência da informação. No SIGLIC a consistência da informação também é observada ao longo do tempo, ou seja, a mesma realidade é observada e registada em tempos diferentes, integra-se a mutação esperada decorrente da evolução e determina-se de novo a consistência da informação. [ACSS-SIGIC, 2011a]

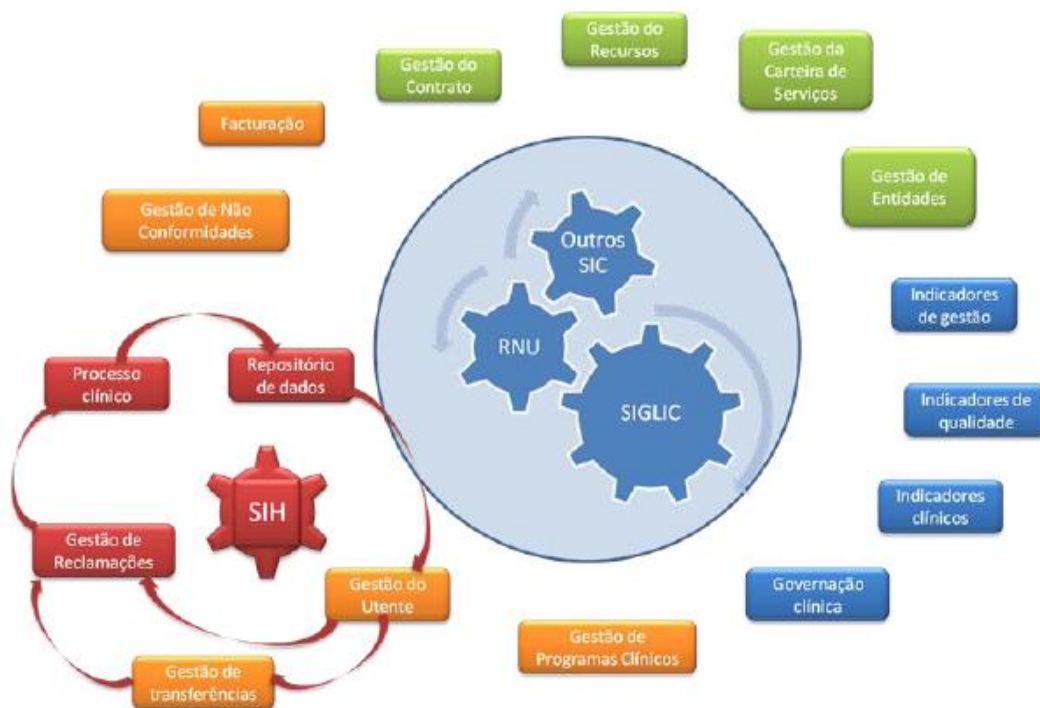


Figura 8 – Funções do SIGLIC
[ACSS-SIGIC, 2010]

O SIGLIC está projetado para se integrar com diversas bases de dados centrais da ACSS - Registos nacionais: RNU (Utente), RNE (Entidades), RNP (Profissionais). Na Figura 8, verifica-se uma sinergia que é fundamental para evitar erros e diminuir os intermediários na gestão da informação, tornando-a mais eficiente e segura.

Sistema de *Business Intelligence*

A partir do SIGLIC foi construída uma base de conhecimento para suportar as decisões de gestão, que providencia dados para investigação científica que apoie a construção de políticas de saúde e que unifique a transparência dos serviços perante o cidadão [Pedro Gomes, 2011]. A determinação dos indicadores seguintes obedece a uma rigorosa metodologia, que visa garantir a qualidade dos dados e comprometer os hospitais com a informação que disponibiliza:

- Produção de indicadores orientados para a gestão:
 - Caracterização da procura;
 - Tempos de acesso;
 - Capacidade instalada;
 - Produção;
 - Produtividade;
 - Acompanhamento dos contratos.
- Produção de indicadores de qualidade do processo de gestão de inscritos para cirurgia;
- Produção de indicadores de orientados para a governação clínica
 - Relação entre a procura discriminada por grupo de patologias e recursos disponíveis e utilizados;

2 Contextualização do projeto

- Serviços prestados e respetivos resultados contextualizados nas entidades responsáveis.
- Produção de indicadores de saúde:
 - Caracterização das necessidades em saúde e condições do hospedeiro;
 - Impacto nos hospedeiros dos tratamentos prestados.
- Produção de indicadores de desempenho do sistema de informação (SI);
- Produção de indicadores de qualidade dos dados do SI.

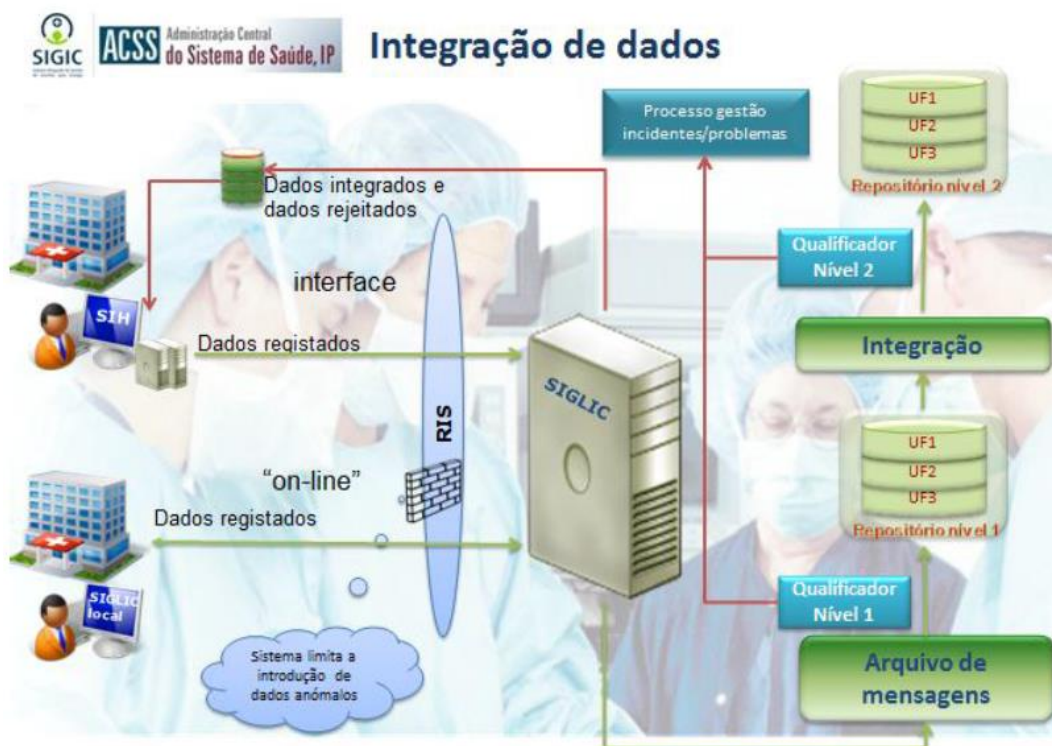


Figura 9 – SIGLIC - Integração de dados
[ACSS-SIGIC, 2011b]

Existem quatro níveis de extração / qualificação de dados e quatro repositórios [ACSS-SIGIC, 2011b], como esquematizado na Figura 9:

- Nível 1: Recolha das mensagens provenientes do processo de transferência dos hospitais para um repositório de mensagens;
- Nível 2: Tratamento do repositório de mensagens para integração de dados; Repositório da base de dados operacional;
 - Subnível a) – verifica a conformidade dos dados e rejeita dados incoerentes;
 - Subnível b) – qualifica os dados integrados em: válidos, suspeitos, inválidos no Sistema de Informação Hospitalar, inválidos no SIGLIC.
- Nível 3: Processamento dos dados do repositório operacional – Extrações mensais, exemplificado na Figura 10:
 - Recolhe e arquiva dados para cálculo de indicadores;
 - Identifica detalhes inválidos;
 - Calcula indicadores excluindo os detalhes inválidos;
 - Assinala indicadores com desvios;

- Cria avisos;
- Repositório *Data Warehouse* (DW) – Base de dados (BD): de detalhes e repositório de indicadores (inclui informação para o *QlikView* – Ferramenta de BI) e indicadores inscritos na BD operacional.

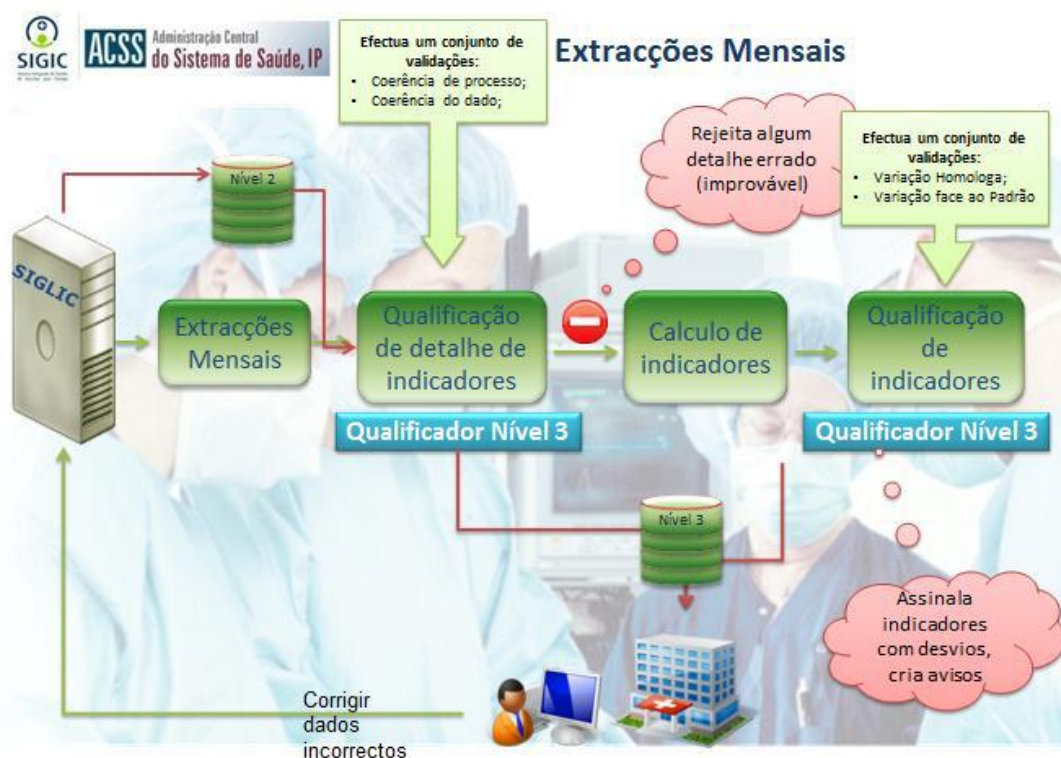


Figura 10 – SIGLIC – Extrações mensais
[ACSS-SIGIC, 2011b]

- Nível 4: Processamento de informação dos repositórios de nível 3 – duas vezes por ano (extração 1º semestre e anual):
 - Subnível a) – após o processamento da 1ª extração (equivalente ao nível 3) é efetuada uma análise manual dos dados que é reportada aos hospitais, de seguida estes analisam os dados e corrigem nos Sistemas de Informação Hospitalar (SIH) eventuais erros;
 - Subnível b) – 2ª extração sob o mesmo período três semanas após a primeira; os hospitais analisam os novos indicadores e anexam comentários, os analistas da Unidade Central de Gestão de Inscritos para Cirurgia (UCGIC) avaliam os novos indicadores integram os comentários e produzem os relatórios finais a publicar.

Nas Figura 11 e Figura 12, são evidenciadas algumas análises que poderão ser consultadas a partir da plataforma de análise de dados, *QlikView*:

2 Contextualização do projeto

- Evolução semanal das Listas de espera para internamento e cirurgia face às entradas e operações realizadas;

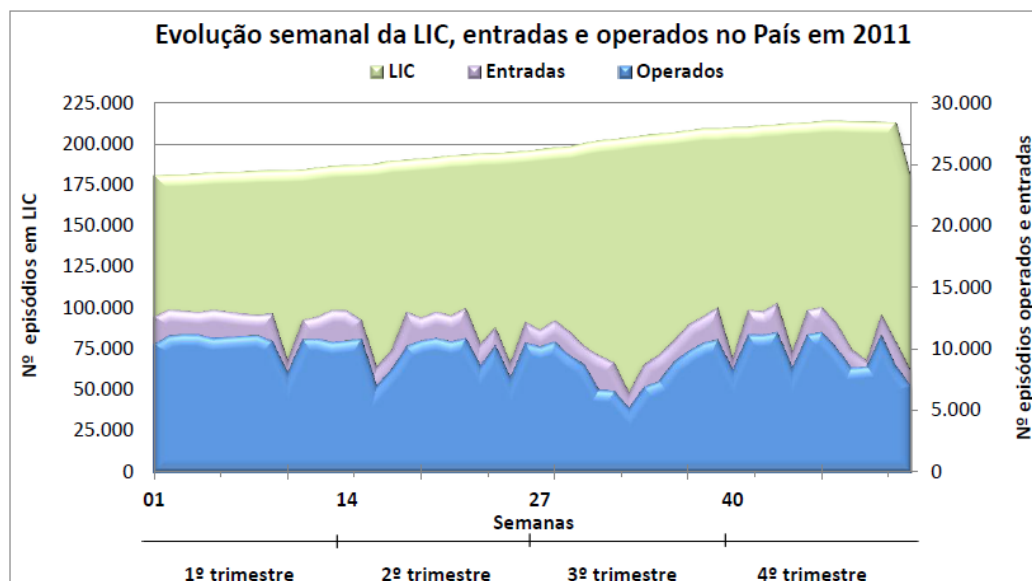


Figura 11 – SIGLIC – País: Evolução semanal da LIC, entradas e operados em 2011 [ACSS-SIGIC, 2011b]

- Características das Listas de espera para Internamento e Cirurgia (LIC) em 2011 face ao número de episódios em LIC e tipo de especialidade.

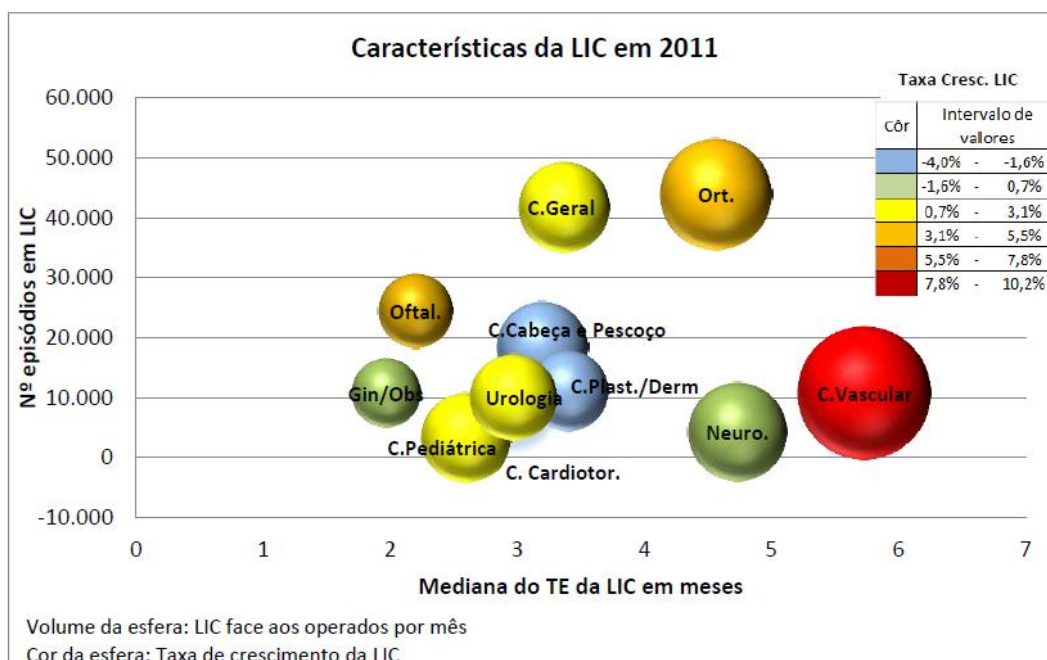


Figura 12 – SIGLIC – Grupo de serviço: Características da LIC no ano de 2011 [ACSS-SIGIC, 2011b]

3 Plataforma de Dados de Saúde

Neste capítulo será contextualizada a génese do projeto da Plataforma de Dados de Saúde (PDS), tendo por base determinados objetivos que contribuíram para a conceção e desenvolvimento da solução. Salientando o contributo que a PDS atualmente presta ao Serviço Nacional de Saúde (SNS), e de que forma contribuiu para a mudança positiva na prática da saúde em Portugal.

Nos pontos seguintes, serão aprofundados dois dos quatro portais da PDS: Portal do Profissional e Portal do Utente. Estas aplicações são consideradas os sistemas origem, sobre os quais serão extraídos os dados, que serão devidamente tratados, carregados e disponibilizados ao utilizador final através do PDS – Portal Institucional. Portanto, serão abordados os seguintes temas:

- Contextualização – Será descrito o âmbito do portal em questão, mencionando os objetivos que permitiram o desenho da solução;
- Arquitetura – Abordagem técnica sob o ponto de vista da arquitetura da solução;
- Segurança e proteção de dados – Descrição dos requisitos e solução adotada, tendo em conta aspetos essenciais para a segurança e proteção de dados, e por forma a cumprir as obrigações determinadas pela Comissão Nacional de Proteção de Dados (CNPD);
- Análise da estrutura de dados – Nesta secção será descrito o modelo de dados, que detém a informação relevante para ser armazenada no repositório central do PDS - Portal Institucional.

3.1 Sobre a génese

A partilha de informação entre organizações prestadoras de cuidados de saúde, e entre estas e o utente traduzem benefícios a vários níveis [Henrique Martins, 2013]:

- Segurança para o utente
- Apoio à boa prática clínica
- Redução de custos por maximização de recursos de informação e conhecimento

Atualmente existe informação presente nos diversos sistemas eletrónicos legados, que em conjunto constituem um registo de saúde vasto acerca dos portugueses. A disponibilização da Plataforma de Dados de Saúde (PDS) representa um somatório potencial de toda a informação registada sobre o utente, pois através da centralização de um conjunto de dados provenientes de fontes heterogéneas, permite fornecer um serviço totalmente centrado no utente.

A Plataforma surge como um projeto no âmbito da Comissão de Informatização Clínica (CIC), constituída no dia 06/Dez/2011 por despacho do Sr. Secretário de Estado da Saúde Dr. Manuel Teixeira e teve desde o seu início dois grandes pilares de definição e implementação [Diogo Reis, 2013]:

1. A realidade nacional, composta por sistemas tecnologicamente e arquitecturalmente avançados bem como de sistemas com mais de duas décadas de existência e com arquiteturas fechadas com os quais a plataforma teria que garantir a interoperabilidade de forma transparente para o utilizador;
2. Restrições de cariz financeiro com as quais a plataforma teria que ser concebida e implementada.

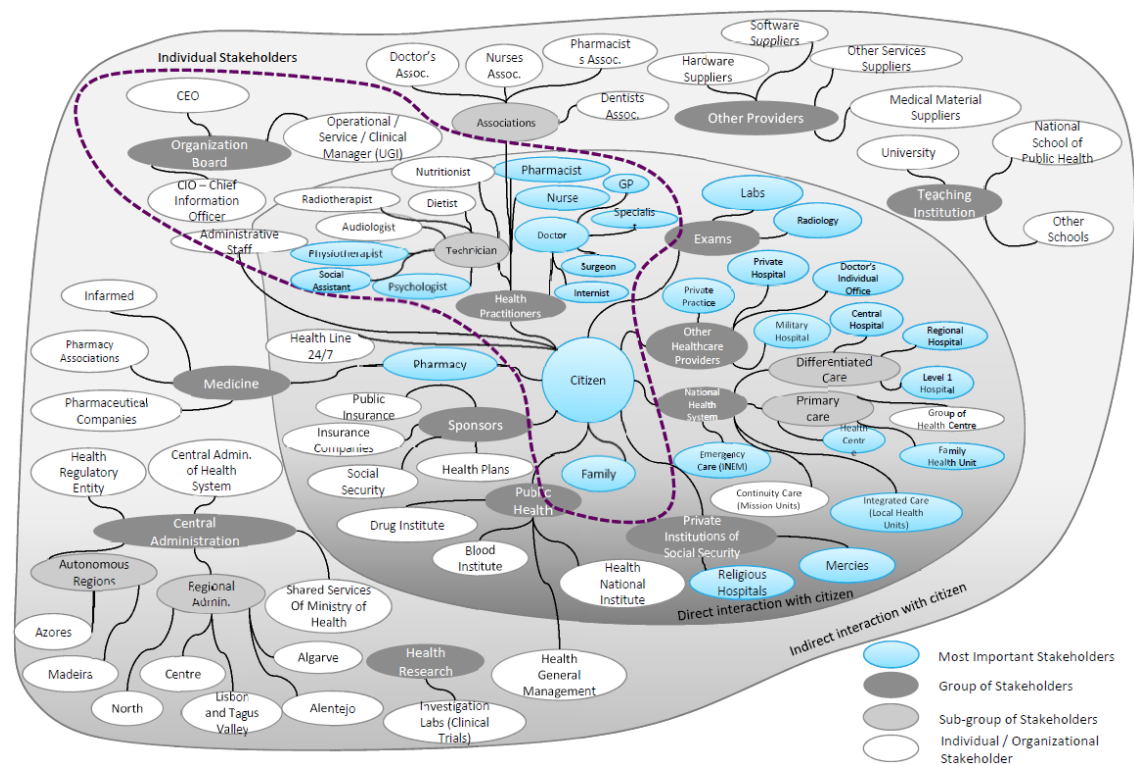


Figura 13 – Desenho da PDS como serviço para diferentes *stakeholders*

[Nelson Pinho et al., 2013]

As vantagens da implementação deste sistema foram demonstradas para os vários interessados [Luís Campos, 2012], identificados na Figura 13:

- Organizações prestadoras de cuidados de saúde aumentam a segurança dos doentes, melhoram a efetividade dos cuidados, eficiência e produtividade, reduzem a repetição desnecessária de exames e tratamentos e melhoram o cumprimento das normas de orientação clínica;
- Utentes beneficiam da diminuição dos riscos, duma melhor continuidade de cuidados e redução de gastos desnecessários em medicação e exames. Também beneficiam de uma plataforma mais orientada às suas necessidades, disponibilizando serviços consoante o respetivo perfil (Mulher, Homem, Criança, Diabético, entre outros).
- Prestadores de cuidados, pelo facto de usufruírem do acesso a toda a informação disponível sobre o utente independentemente do local de registo da mesma, usufruem da possibilidade de prestarem cuidados com mais qualidade, de forma mais efetiva e eficiente.
- Entidades pagadoras têm menos despesas administrativas, menos custos em exames desnecessários e um cumprimento mais cabal das políticas de saúde.

Na Figura 14, é esquematizada a PDS como plataforma constituída por quatro portais distintos e interoperáveis entre si. Não obstante de esclarecimento breve, visto não ser abordado seguidamente, o Portal Internacional (PDS-EPSOS) representa um canal de partilha e consulta de informação clínica entre Portugal e os países que participam no projeto EPSOS com o Resumo Clínico Único do Utente (RCU²): Áustria; Espanha; França; Itália. O RCU² é um resumo clínico que dispõe de informação mínima sobre o estado de saúde do paciente, auxiliando os profissionais de saúde, por exemplo, sempre que existe necessidade de um atendimento urgente ou programado numa entidade de saúde nacional ou estrangeira.

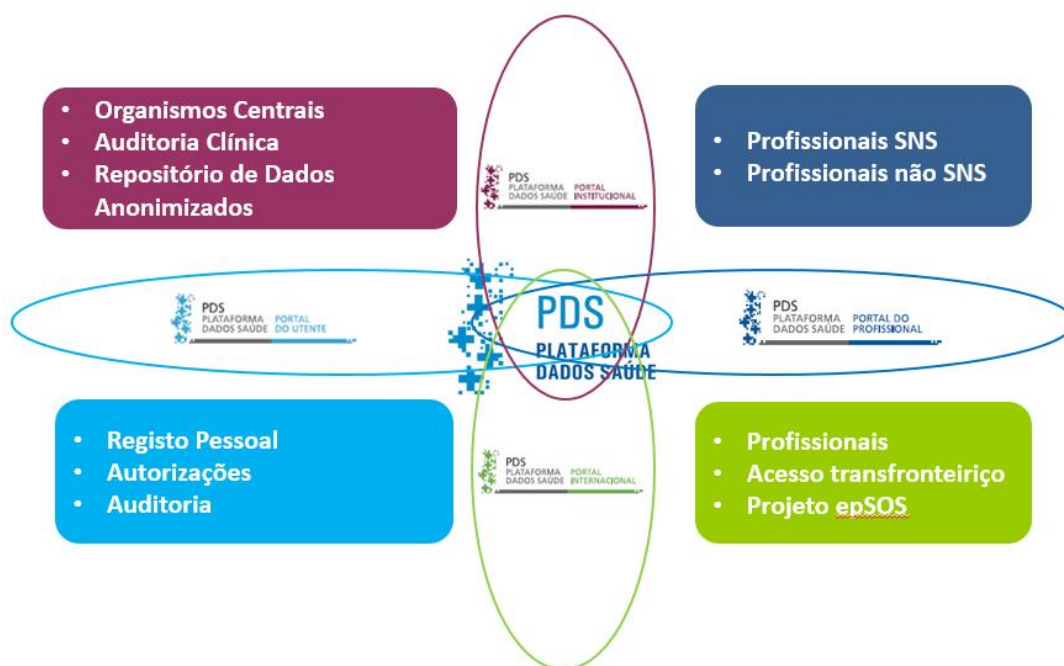


Figura 14 – Plataforma de Dados de Saúde - Portais

3.2 Portal do Utente – Aproximar o utente ao SNS

O PDS - Portal do Utente (PDS-PU) é um projeto que visa disponibilizar um conjunto de serviços e informação de acordo com as principais necessidades de saúde dos utentes. Através desta ferramenta os cidadãos podem visualizar e gerir um conjunto de informação clínica pessoal, que poderá ser relevante para si próprio como para os profissionais de saúde com os quais possa vir a interagir. Portanto, foi dada particular relevância à informação que pelo seu carácter de utilidade em âmbito de urgência, pode ter um contributo relevante para a saúde do utente, nomeadamente o registo estruturado de alergias, seus hábitos medicamentosos e informação de contactos de emergência.



Figura 15 – PDS-PU – Novo portal lançado a 7 de Maio de 2013

No dia 31 de Maio de 2012, foi lançada a primeira versão do PDS-PU, disponibilizando a partir de então, uma via de comunicação entre o utente e o Serviço Nacional de Saúde (SNS). A linha orientadora desta versão era baseada num conjunto alargado de serviços eletrónicos que o SNS disponibilizava ao utente, e que este poderia usufruir de modo a agilizar determinados processos. A 7 de Maio de 2013, foi divulgada uma nova versão com uma linha orientadora diferente, como exemplificado na Figura 15, mais centrada nas necessidades do utente, disponibilizando um conjunto de serviços informativos.

O PDS-PU pode ser acedido quer por utilizadores anónimos, que poderão usufruir de todos os serviços informativos, quer por utilizadores registados, que por sua vez através de uma área pessoal designado por PHR (*Patient Health Record*), é disponibilizado um conjunto diverso de serviços eletrónicos, representados na Tabela 1.

Tabela 1 – PDS-PU – Serviços disponibilizados

SERVIÇOS INFORMATIVOS
Informação sobre o que o utente precisa para realizar um determinado serviço do SNS, organizado por perfil do utente ou tipo de serviço (Homem, Mulher, Criança, Idoso, Consultas, Dádivas, entre outros)
Pesquisa avançada de prestadores de saúde
Dicionário do utente - informação médico-científica, em linguagem clara, sobre as doenças e situações clínicas mais frequentes, bem como sobre os exames e tratamentos mais comuns
Destaque e eventos e datas comemorativas da área da saúde
Perguntas frequentes e guias para a utilização de serviços

Contactos e linhas de atendimento da saúde

SERVIÇOS ELETRÓNICOS

Partilha de informação com os profissionais de saúde do SNS (hospitais, urgências, cuidados primários), mediante autorização prévia do utente, e possibilidade de consulta do histórico de acessos: Contactos de emergência; Hábitos, medicação, alergias e doenças; Medições de peso, altura, glicémia, tensão arterial, colesterol, triglicéridos, saturação de oxigénio e tempo de coagulação do sangue (INR); Documentos de saúde

Cronograma de contatos do utente nas diversas instituições prestadoras de cuidados de saúde com a possibilidade de visualização de receituário e guias de tratamento¹

Consulta do resumo clínico único do utente¹

Marcação de consultas e pedido de medicação crónica no centro de saúde de referência

Consulta da posição na lista de espera para cirurgia

Contatar o médico de família ou centro de saúde de referência

Na Figura 16, poderá ser visualizado o PHR, que é utilizado por utentes registados e devidamente autenticados, recorrendo a um dos seguintes métodos:

- Autenticação simplificada – Introdução do número de utente e senha de acesso;
- Autenticação forte – Leitura do cartão do cidadão e introdução do pin de acesso.

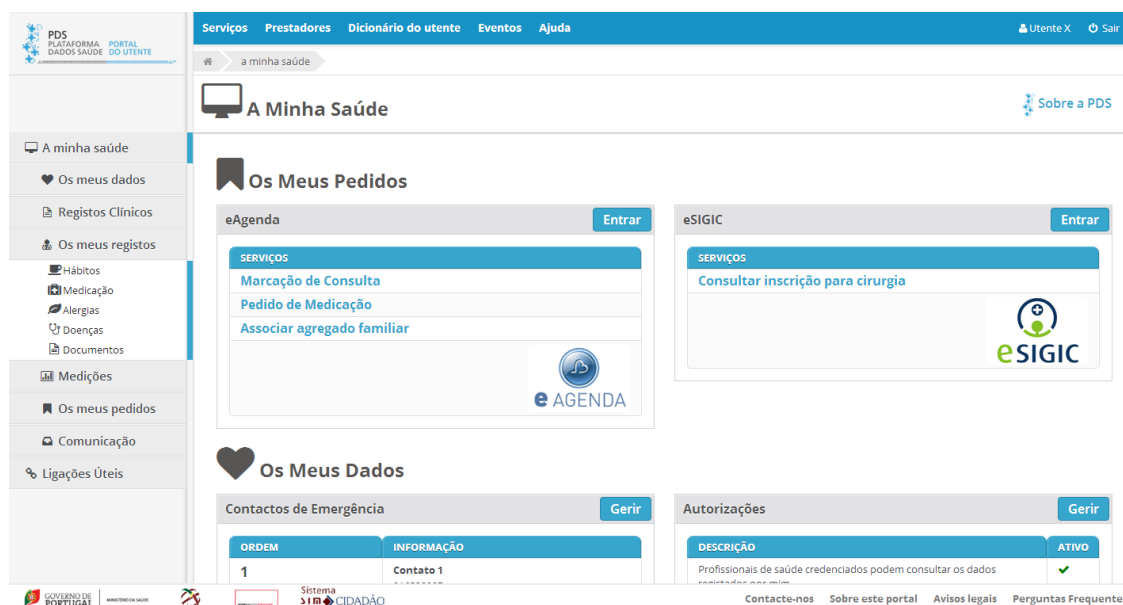


Figura 16 – PDS-PU – Área pessoal para utilizadores autenticados

¹ - Serviços disponibilizados somente com autenticação forte, ou seja, através da utilização do cartão do cidadão.

3.2.1 Arquitetura, Segurança e Proteção de dados

A arquitetura do PDS-PU, esquematizada na Figura 17, é constituída pela combinação entre uma arquitetura orientada a serviços e orientada a eventos (SOA 2.0). Este portal é constituído por um conjunto de aplicações independentes, que implementam autonomamente as respetivas regras de negócio:

- Área pública (não autenticado) – Mantem um legado de desenvolvimento em *Sharepoint*, sendo atualmente utilizado para facilitar a gestão de conteúdos dos serviços informativos do portal e disponibilizar o serviço eletrónico de pedido de isenção de taxas moderadoras;
- *eAgenda* – Aplicação responsável por fornecer os serviços eletrónicos de marcação de consultas e pedido de medicação crónica ao centro de saúde de referência;
- *eSIGIC* – Aplicação responsável por fornecer o serviços eletrónico de consulta da posição na lista de espera para cirurgia;
- Área pessoal (com autenticação) – Diverge das aplicações anteriores, recorrendo uma diferente abordagem de desenvolvimento (ASP.NET MVC4 e Web API), tendo como objetivo principal um sistema centrado no utente (PHR - *Patient Health Record*), com a disponibilização de informação orientada à situação atual de saúde (i.e.: paciente diabético) e fornecer um conjunto de diversos serviços eletrónicos. Esta aplicação agrega num único local todos os serviços eletrónicos mencionados nos pontos anteriores.

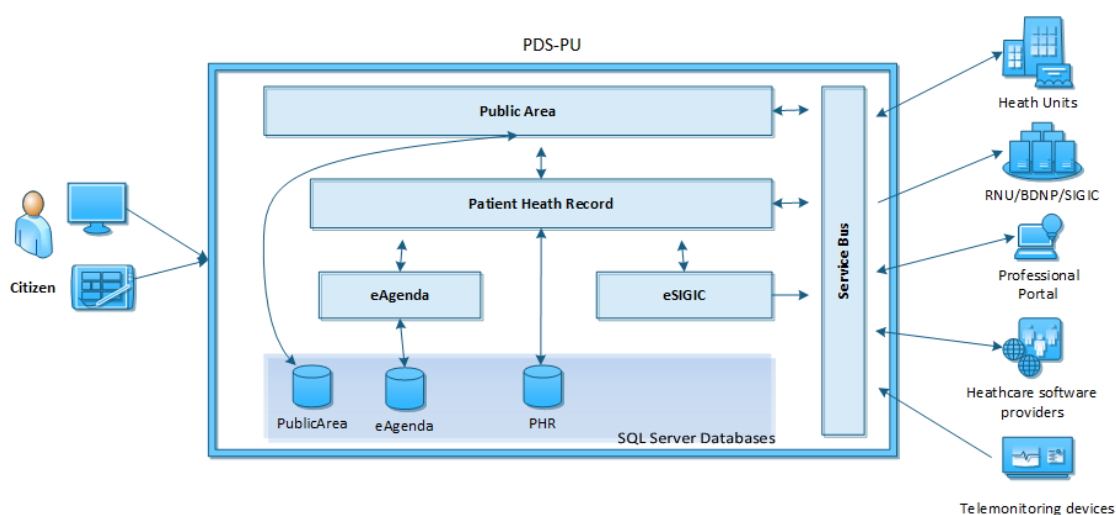


Figura 17 – PDS-PU – Representação da arquitetura PDS-PU

Um sistema PHR tem o seguinte conjunto de benefícios principais [Paul C. Tang, 2006]:

- Disponibilizar ao utente o acesso a um conjunto diverso de informação e conhecimento de saúde credível;

3 Plataforma de Dados de Saúde

- Fornecer ao utente uma ferramenta que permita gerir a sua condição de saúde, monitorizar as medições relevantes e avaliar o risco de vir a sofrer de potenciais doenças;
- Facilitar a comunicação entre o utente e as instituições prestadoras de cuidados de saúde;
- Permitir ao utente beneficiar do acesso a comunidades organizadas, para incentivar a partilha de experiências entre pessoas com os mesmos problemas/doenças.
- Partilhar a informação com os profissionais de saúde, de modo a facilitar o tratamento e diagnóstico mais apropriado, tendo em conta a condição de saúde declarada.

De modo a garantir que o utente é quem define o acesso à sua informação de saúde, podendo negar a partilha dos seus dados a outros profissionais de saúde, é disponibilizada uma área de autorizações que permite definir o acesso:

- Consulta de informação clínica, através do Portal do Profissional (PDS-PP), a profissionais de saúde de diferentes instituições da qual a informação foi originalmente registada;
- Consulta dos dados de saúde introduzidos pelo utente no PHR e visualizados pelos profissionais através do PDS-PP;
- Autorizar a partilha transfronteiriça do resumo clínico único do utente (RCU2), de modo a que quando este se desloque para um país aderente ao projeto EPSOS (Áustria; Espanha; França; Itália), este possa constituir uma base de conhecimento útil para o tratamento e diagnóstico efetuado por profissionais internacionais.

Contudo, de modo a tornar o mecanismo de autorizações transparente para o utente, é disponibilizada uma área de auditoria, onde é possível visualizar o histórico dos acessos dos profissionais que, recorrendo ao PDS-PP, acederam à informação do utente.

No que diz respeito ao processo de autenticação, retratado na Figura 18 foi implementado um mecanismo Single-Sign On (SSO), que permite o utilizador navegar entre aplicações distintas com uma única autenticação. O utente quando se regista no portal a informação é integrada com a plataforma SSO que armazena as respetivas credenciais de acesso. Posteriormente poderá ser associado ao registo o cartão do cidadão do utente, garantindo desta forma um método mais forte de autenticação e o acesso a informação reservada, através do acesso com autenticação forte (i.e.: informação clínica introduzida por profissionais de saúde nas instituições onde o utente teve contato).

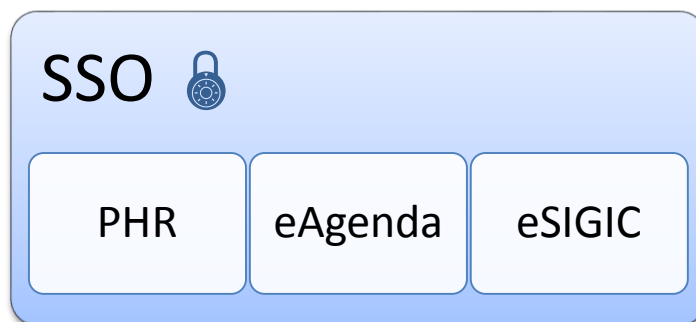


Figura 18 – PDS-PU – Aplicações protegidas pelo mecanismo SSO

3.2.2 Conhecimento para análise

A informação registada pelo utente no PDS-PU representa um alvo importante de estudo de saúde pública e de comportamento dos próprios utentes no que diz respeito ao controlo e monitorização da sua saúde. Portanto a disponibilização a um conjunto de *stakeholders*, da informação proveniente do próprio utente constitui uma mais-valia para o aumento de conhecimento que à data era somente proveniente da atividade dos profissionais de saúde. Contudo, também será possível avaliar a utilização de determinados serviços eletrónicos e avaliar a capacidade de resposta do SNS face aos pedidos efetuados por este meio.

Para a operacionalização do projeto PDS – Portal Institucional, foi identificado um conjunto inicial de informação relevante, que deverá ser disponibilizado a um conjunto distinto de utilizadores:

- Evolução de novos utilizadores registados;
- Distribuição de utilizadores por Distrito/Concelho/ARS/ACES/Instituição de referência;
- Análise de doenças, alergias e medicação registadas relacionadas por ARS/ACES/Instituição de referência;
- Resumo das autorizações configuradas pelos utentes.

Para tal, a análise do modelo de dados é um fator essencial para identificação das diversas tabelas responsáveis pelo armazenamento da informação e para a definição de uma estratégia de extração de dados. Na Figura 19, através de um diagrama Entidade-Relação (E-R) é pretendido retratar as diversas entidades envolvidas.

3 Plataforma de Dados de Saúde

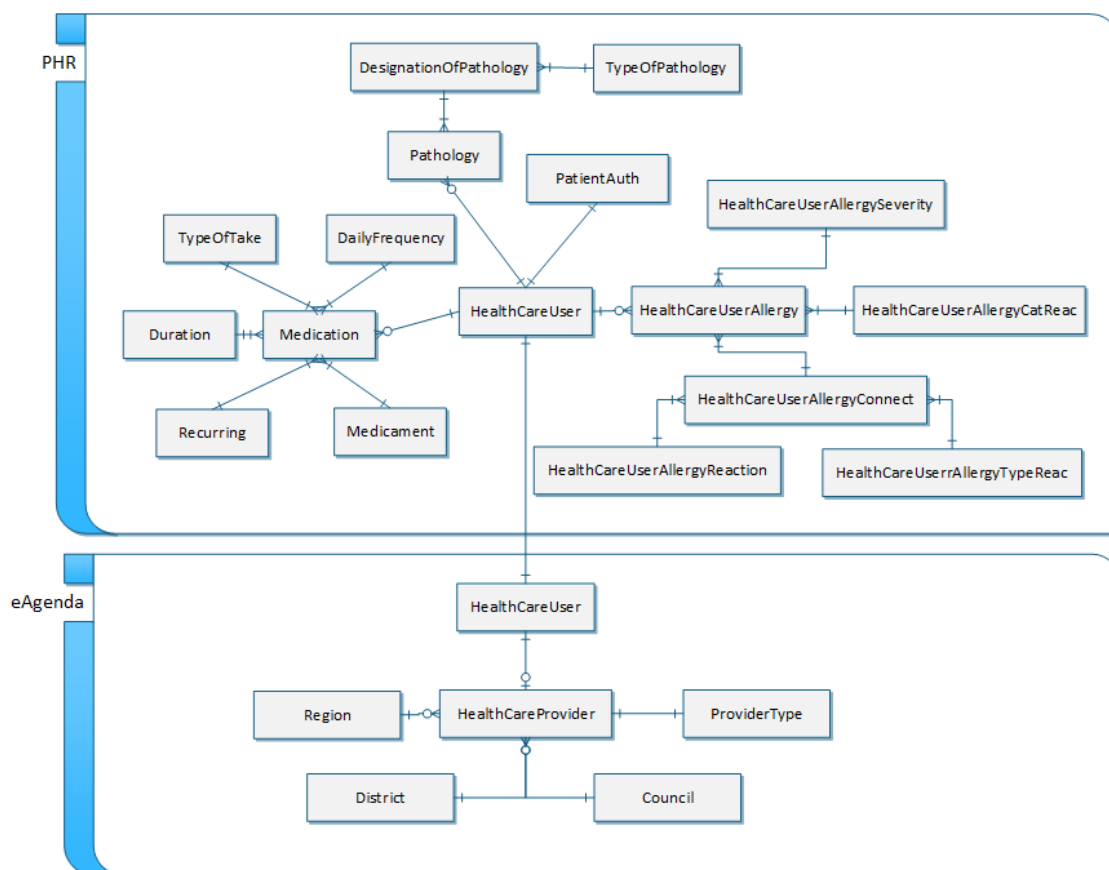


Figura 19 – PDS-PU – Identificação das entidades relevantes para extração de dados

Após a análise efetuada sobre o modelo de dados do PDS-PU, garantindo a disponibilização da informação considerada fulcral e tendo como referência o diagrama anterior, na Tabela 2 – PDS-PU – Entidades principais são identificadas as áreas de informação relevantes para o processo de extração de dados.

Tabela 2 – PDS-PU – Entidades principais

Informação	Entidades identificadas
Paciente	Base de dados PHR: (HealthCareUser; PatientAuth); Base de dados eAgenda: (HeathCareUser;HeathCareProvider; ProviderType; Region; District; Council).
Alergia	Base de dados PHR: (HealthCareUserAllergy; HealthCareUserAllergySeverity; HealthCareUserAllergyCatReac; HealthCareUserAllergyConnect; HealthCareUserAllergyReaction; HealthCareUserAllergyTypeReac)
Doença	Base de dados PHR: (Pathology; DesignationOfPathology; TypeOfPathology)
Medicamento	Base de dados PHR: (Medication; Medicament; Recurring; Duration; TypeOfTake; DailyFrequency)

3.3 Portal do Profissional – Partilhar para melhor atender

No dia 2 de Julho de 2012, foi lançado o PDS-Portal do Profissional (PDS-PP) nos hospitais e instituições de cuidados de saúde primários pertencentes à ARS Norte, e até ao início de Novembro de 2012 foi alargado às restantes instituições do país. Sendo uma aplicação que é lançada a partir do sistema clínico local (EHR – *Eletronic Health Record*), o processo de implementação contou com a participação de diversos intervenientes: infraestruturas e configuração da Rede de Informação da Saúde (RIS); atualização nacional do Sistema de Apoio ao Médico (SAM); atualização nacional do sistema clínico dos cuidados de saúde primários (SAM-CS); atualização nacional do Sistema de Apoio à Prática de Enfermagem (SAPE).

O lançamento deste portal possibilitou abrir uma janela dentro do SNS, colocando ao dispor dos profissionais de saúde o acesso a todo um manancial de dados clínicos sobre o utente que até agora estavam dispersos por diferentes sistemas de informação. Desde sistemas de âmbito nacional, como o INEM, a Rede Nacional de Cuidados Continuados Integrados (RNCCI) ou a Direção Geral de Saúde (DGS), até aos sistemas de informação (EHR – *Eletronic Health Record*) alojados localmente nos servidores de hospitais e centros de saúde.

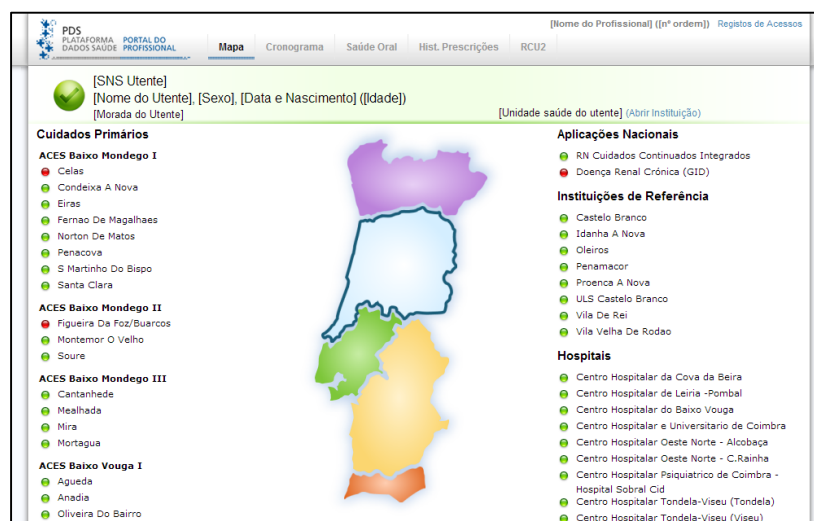


Figura 20 – PDS-PP – Mapa de instituições

Os profissionais (médicos e enfermeiros) em atividade nos hospitais e cuidados de saúde primários dispõem, através da PDS-PP, da seguinte informação:

- Visualização do Processo Clínico Eletrónico (PCE) das instituições hospitalares e cuidados de saúde primários pertencentes ao SNS, através do acesso ao mapa ilustrado na Figura 20;
- Detalhes clínicos provenientes de aplicações nacionais: INEM, Doença Renal Crónica, RNCCI e DGS;

3 Plataforma de Dados de Saúde

- Histórico do receituário do utente recorrendo à Base de Dados Nacional de Prescrição (BDNP);
- Com o consentimento do utente, consultar os dados de saúde, medições e contatos de emergência registados na área pessoal do PDS - Portal do Utente (PHR);
- Gestão do Resumo Clínico Único do Utente (RCU²) por parte do médico de família do utente, e visualização por parte dos restantes profissionais.

Atualmente, o acesso à PDS-PP é possível em praticamente todos os EHR's existentes nas instituições de saúde públicas, sejam os mesmos disponibilizados pelo Ministério da Saúde ou por fornecedores de Software para o setor da saúde. No entanto, segundo a submissão da fase II da PDS à Comissão Nacional de Proteção de Dados (CNPd), é deliberada a permissão de acesso por instituições do setor convencionado ou privado, respeitando um conjunto de normas, das quais deverão salientar-se as seguintes:

- Dispor de um processo clínico eletrónico acreditado pela própria CNPD;
- Conter o registo do utente, devidamente identificado com o nº de utente, e de acordo com o Registo Nacional de Utentes (RNU);

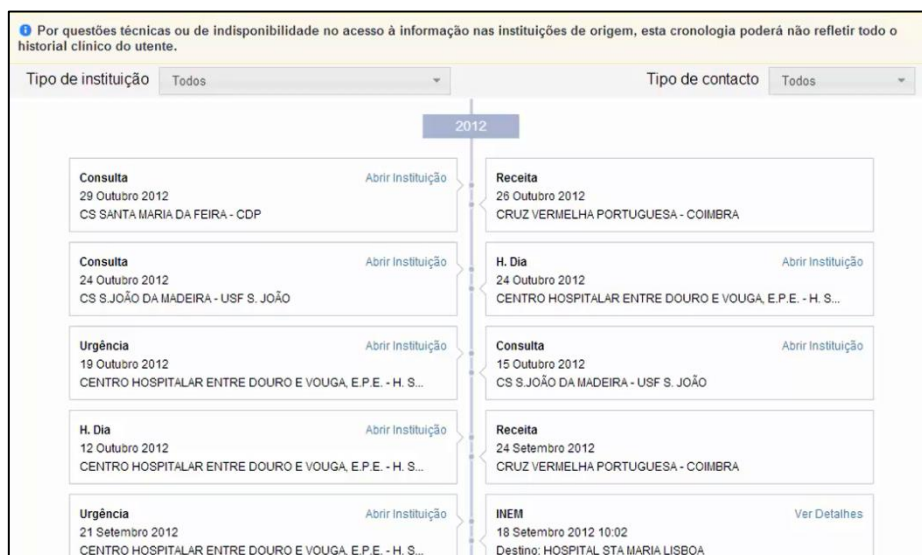


Figura 21 – PDS-PP – Cronograma de contatos do utente nas diversas instituições

Na Figura 21, é apresentado um cronograma, que disponibiliza uma perspetiva temporal dos diversos contatos, que um utente efetuou em diversas instituições de saúde, e interações em programas nacionais. A informação que consta nesta área é carregada em tempo real e os dados são provenientes de diversas fontes:

- Base Dados Nacional de Prescrição (BDNP) – Dispõe de todas as visitas a instituições, em que foi efetuada prescrição de medicação ou meios complementares de diagnóstico e tratamento.
- Serviços disponibilizados nas instituições e aplicações nacionais – Existem instituições/aplicações nacionais que disponibilizam serviços locais para consulta dos contatos ocorridos para um determinado utente.
- Repositório de contatos – Enumeras instituições hospitalares enviam diariamente os contatos de todos os utentes e marcações efetuadas no dia anterior. Desta forma, poderão ser fornecidos não só os contatos já efetuados, assim como os que estão agendados.

Mesmo recorrendo a estas fontes distintas, atualmente a base de conhecimento do cronograma não reflete na totalidade o histórico de contatos do utente, pelos seguintes fatores: contatos sem prescrição; contatos que não constam no repositório; inexistência de serviços locais, de algumas instituições públicas e privadas, para consulta de contatos.

3.3.1 Arquitetura, Segurança e Proteção de dados

O PDS-PP foi desenhado num modelo de arquitetura distribuída e de alta disponibilidade, como exemplificado na Figura 22, de modo a operar no contexto da prática clínica com o utente. A informação clínica disponibilizada reside apenas nos sistemas de origem, e é agregada e disponibilizada no momento da consulta pelo profissional.

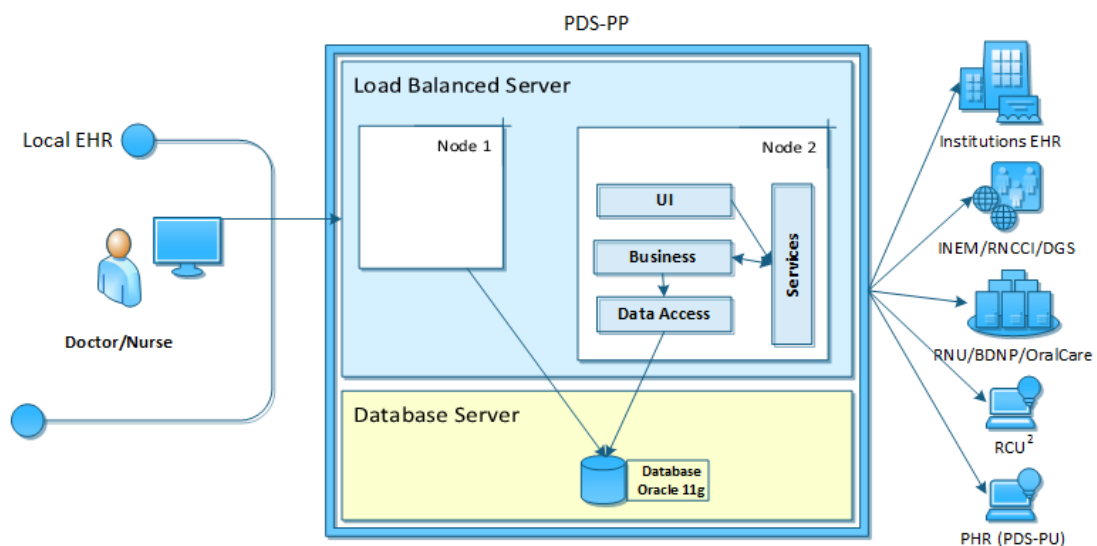


Figura 22 – PDS-PP – Representação da arquitetura PDS-PP

Na Figura 23, poder ser verificado que o acesso à PDS-PP é realizado por profissionais médicos e enfermeiros, a partir do sistema local no contexto de um determinado utente. Este

3 Plataforma de Dados de Saúde

mecanismo de acesso impossibilita a pesquisa de utentes e permite somente o acesso, se o utente estiver a ser tratado na instituição em questão. Este processo é definido por diversas etapas, como é evidenciado na figura seguinte, que asseguram um acesso seguro e autorizado ao sistema.

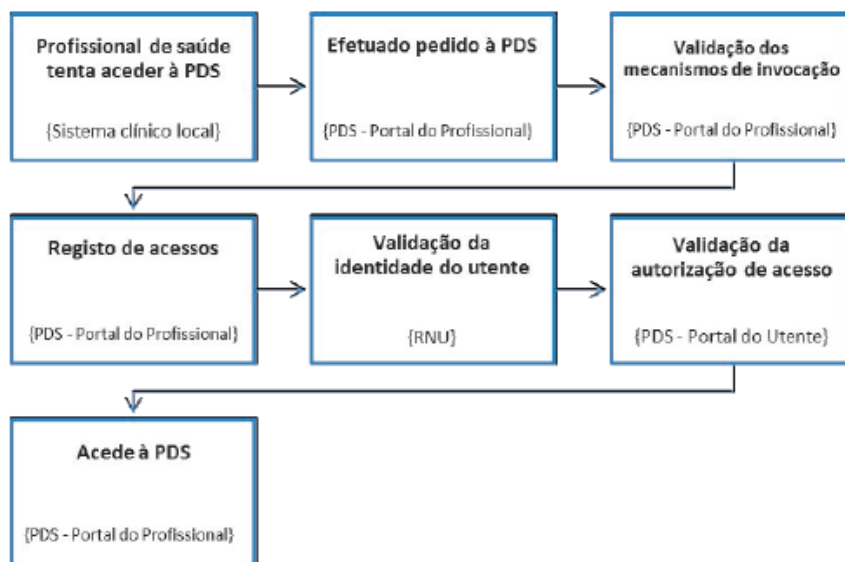


Figura 23 – PDS-PP – Processo de validação de acessos

Os acessos ao PDS-PP são registados e arquivados, podendo ser consultados com a finalidade de auditar os diversos acessos ao portal:

- Auditoria pelo profissional – Permite a visualização de todos os acessos efetuados pelo profissional ao PDS-PP;
- Auditoria pelo utente – De modo a permitir a auditoria de acessos no PDS-PU, é disponibilizado um serviço de consulta dos acessos, efetuados por profissionais no contexto do utente em questão.

Relativamente a outros aspetos relacionados com segurança, o PDS-PP está implementado num circuito protegido no âmbito da Rede de Informação da Saúde (RIS), esquematizado na Figura 24, que fornece:

- Mecanismos de segurança para o controle de acessos e para a fiabilidade das comunicações;
- Confidencialidade e privacidade dos dados;
- Rede dedicada que fornece um conjunto de serviços que trazem valor acrescentado às instituições, aos prestadores de cuidados e aos utentes do SNS.

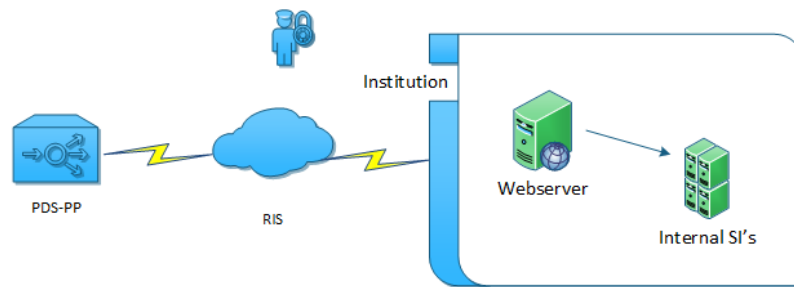


Figura 24 – PDS-PP – Ligação segura através da RIS

3.3.2 Conhecimento para análise

O armazenamento do registo de acessos e o rastreio das operações efetuadas no PDS-PP, torna possível a disponibilização a um conjunto de *stakeholders*, de informação útil sobre a utilização do portal durante a prestação de cuidados de saúde. Para além disso, o acesso ao processo clínico eletrónico (PCE) de outras instituições, permite avaliar as ligações mais comuns, estabelecendo uma perspetiva realista da rede de referência.

Para a operacionalização do projeto PDS – Portal Institucional (PDS-PI), foi identificado um conjunto inicial de informação relevante, que deverá ser disponibilizado a um conjunto distinto de utilizadores:

- Auditoria de acessos ao PDS-PP por ARS/ACES/Instituição/Profissional;
- Operações efetuadas por categoria profissional e instituição;
- Auditoria de acessos ao PCE das Instituições de referência do utente;
- Resumo dos contatos registados no repositório da PDS

Para tal, a análise do modelo de dados é um fator essencial para identificação das diversas tabelas responsáveis pelo armazenamento da informação e para a definição de uma estratégia de extração de dados. Na Figura 25, através de um diagrama Entidade-Relação (E-R) é pretendido retratar as diversas entidades envolvidas.

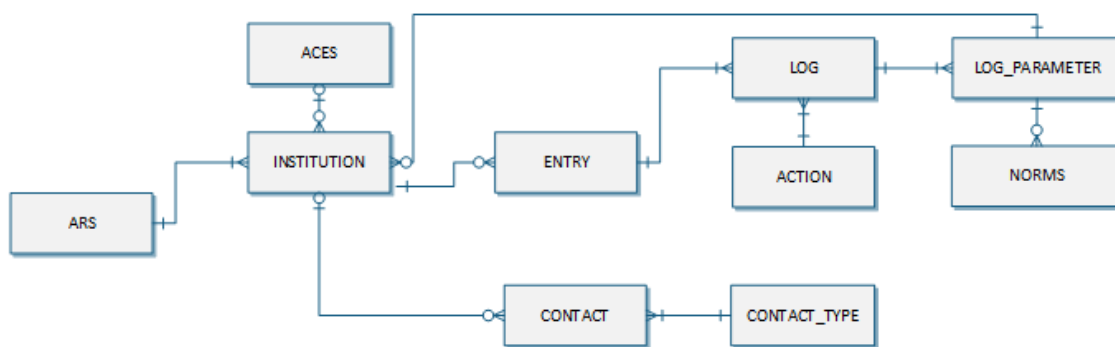


Figura 25 – PDS-PP – Identificação das entidades relevantes para extração de dados

Após a análise efetuada sobre o modelo de dados do PDS-PU, garantindo a disponibilização da informação considerada fulcral e tendo como referência o diagrama anterior, na Tabela 3 – PDS-PP – Entidades principais são identificadas as áreas de informação relevantes para o processo de extração de dados.

Tabela 3 – PDS-PP – Entidades principais

Informação	Entidades identificadas
Instituição	Institution, ACES, ARS
Normas de Orientação Clínica	Norms
Paciente	Entry
Acessos	Entry; Log; Log_Parameter; Action
Contatos	Contact; ContactType

Finalizada a contextualização e análise do PDS-PU e PDS-PP, seguidamente serão detalhadas todas as fases de desenvolvimento e implementação do PDS-PI. Tendo em conta uma inicial contextualização da atualidade, serão enumerados os fatores principais que motivaram o desenvolvimento deste projeto.

4 Portal Institucional – Analisar para potenciar

Neste capítulo, é pretendido abordar o portal PDS – Portal Institucional (PDS-PI) retratando o ponto de situação atual do projeto, os objetivos propostos e medidas para atingir as metas definidas. A reestruturação da arquitetura do PDS-PI representa uma das medidas fundamentais para a concretização do projeto, e conseqüentemente uma seleção apropriada de tecnologias.

Para a implementação do sistema de *Business Intelligence* (BI) e recorrendo a um armazém de dados central, após a análise do modelo de dados dos portais da PDS, foi fundamental delinear um modelo dimensional que suporte o conjunto de análises definidas nas secções 3.2.2 e 3.3.2. Neste sistema um dos processos mais relevantes é a extração de dados dos sistemas origem, tratamento e verificação da qualidade da informação, de modo a ser carregada no armazém de dados (denominado por ETL).

A componente de apresentação da informação aos utilizadores finais, é abordado na última secção, mencionando os diferentes níveis de acesso e a solução adotada para o *front-end* do PDS-PI.

4.1 Estado atual do portal

O PDS-PI foi concebido para responder rapidamente à necessidade de auditoria, por parte dos administradores da PDS e instituições, aos acessos ao PDS - Portal do Profissional (PDS-PP). A informação fornecida tem a finalidade de sustentar os relatórios de avaliação do projeto, e detetar grupos de profissionais potenciais para ações de divulgação do PDS-PP.

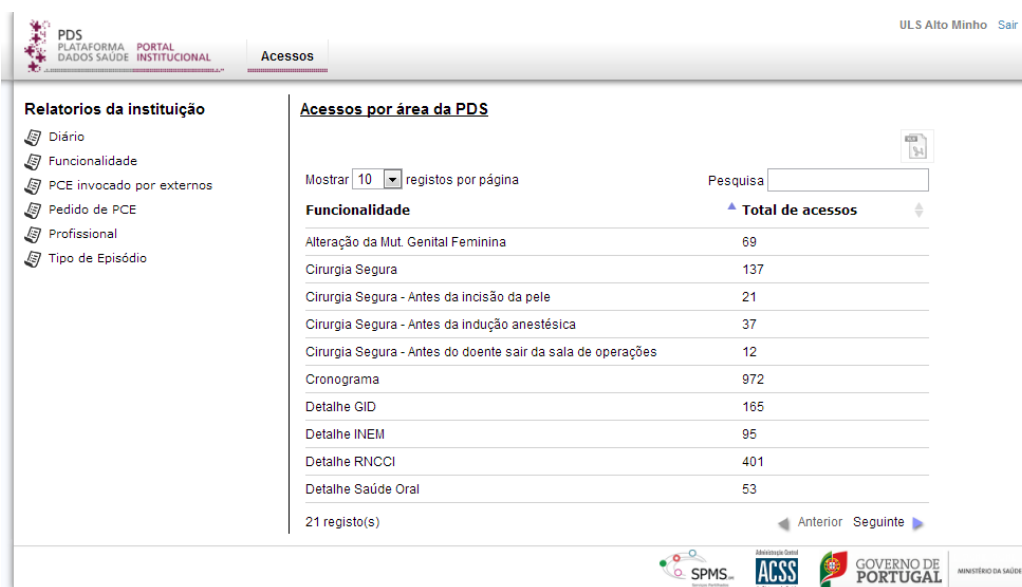


Figura 26 – PDS-PI – Primeira implementação do portal

O desenvolvimento e implementação da primeira versão do portal, ilustrado na Figura 26, é assente num interface simples e objetivo. A aplicação foi implementada sobre a infraestrutura e base de dados operacional do PDS-PP, devido à urgência da implementação que por conseguinte levou a contornar as burocracias internas para criação de uma infraestrutura independente. O modo como foi implementada esta aplicação, representa uma ameaça clara à autonomia e evolução do portal. Os dados apresentados na Figura 27, evidenciam a quantidade mensal de acessos e a totalidade de instituições abrangidas [SPMS, 2013], que traduzem uma fraca adesão e abrangência nacional.

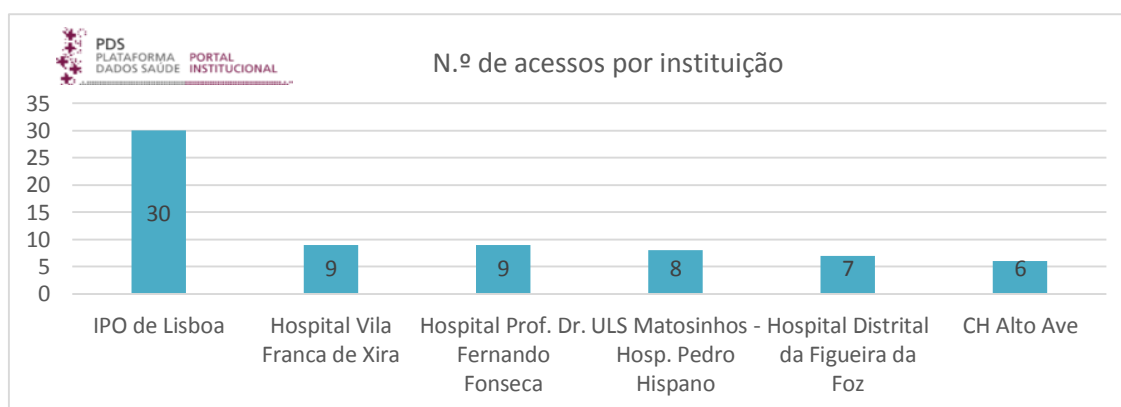


Figura 27 – PDS-PI – Acessos de Junho de 2013 [SPMS, 2013]

Outro aspeto relevante foi o facto das listagens disponibilizadas serem estáticas. Esta abordagem não permite a agilidade pretendida na análise dos dados (i.e.: agregar a informação em diferentes níveis institucionais), o que originou a existência de um processo manual de extração mensal de dados mais específicos e devidamente agregados. Contudo, foram identificados um conjunto de novos *stakeholders* com níveis de acesso distintos, que numa fase inicial não tinham sido considerados:

- Direção Geral de Saúde - DGS
- Administração Central do Sistema de Saúde – ACSS
- Administração Regional de Saúde – ARS
- Agrupamentos de Centros de Saúde – ACES

Para solucionar as problemáticas expostas, em grupo de trabalho da PDS, foi decidido descontinuar a aplicação atual. Uma nova abordagem de desenvolvimento e implementação deverá estar assente numa infraestrutura independente, e num novo sistema que contenha mecanismos eficazes de extração, tratamento e carregamento de dados para um repositório central. Por fim, deverá ser disponibilizada uma aplicação cuja função seja apresentar ao utilizador final, um conjunto de análises/indicadores adequadas ao seu perfil.

4.2 Arquitetura proposta

A arquitetura de um sistema de informação é a base que permite desenvolver uma solução que esteja de acordo com as linhas orientadoras do projeto. Portanto, para o PDS-PI as linhas principais são:

- Reconhecer a mudança como uma constante;
- Seguir uma abordagem de desenvolvimento incremental;
- Garantir a estabilidade do funcionamento de todos os portais da PDS;
- Assegurar flexibilidade e simplicidade do modelo, garantindo uma visão adequada sobre a informação.

Após a avaliação destas linhas, foi importante analisar os principais objetivos pretendidos para o PDS-PI:

- Implementar o sistema recorrendo a ambientes de virtualização;
- Disponibilização de um armazém de dados (DW - *Datawarehouse*) e um conjunto de repositórios focados no negócio em questão (DM – *Datamart*), que permitam armazenar quantidades elevadas de dados;
- Automatizar o processo de extração, tratamento e entrega de dados no DW a partir dos dados armazenados nas bases de dados transacionais dos portais (PDS-PP e PDS-PU);

- Fornecer uma aplicação que possibilite a gestão dinâmica de utilizadores/perfis de acesso, apresentando um conjunto de análises/indicadores ao utilizador;
- Aplicar uma solução que permita mobilizar o negócio do PDS-PI, recorrendo ao acesso a partir de dispositivos móveis.

Para o estudo das principais abordagens de arquitetura de DW foram analisados o paradigma de *Ralph Kimball (Data Warehouse Bus Architecture)* [Ralph Kimball, 2008] e o paradigma de *Bill Inmon (CIF – Corporate Information Factory)* [W.H. Inmon, 2001]. Tendo em conta as linhas orientadoras e os objetivos definidos anteriormente, o paradigma de arquitetura mais adequado para o PDS-PI é o de *Ralph Kimball*, pelos seguintes motivos [Claudia Imhoff, 2003]:

- Fluxo de dados *bottom-up*, permite uma evolução incremental do modelo consoante as necessidades da empresa. Ao contrário do paradigma CIF, que tem um fluxo de dados *top-down*, partindo de um modelo empresarial bem definido para disponibilização de repositórios de dados específicos;
- O esforço de implementação é reduzido face à estratégia CIF. Recorrendo a um repositório de dados inicial, é possível provar o conceito criando um piloto do portal, disponibilizando determinadas análises sobre um caso de uso.

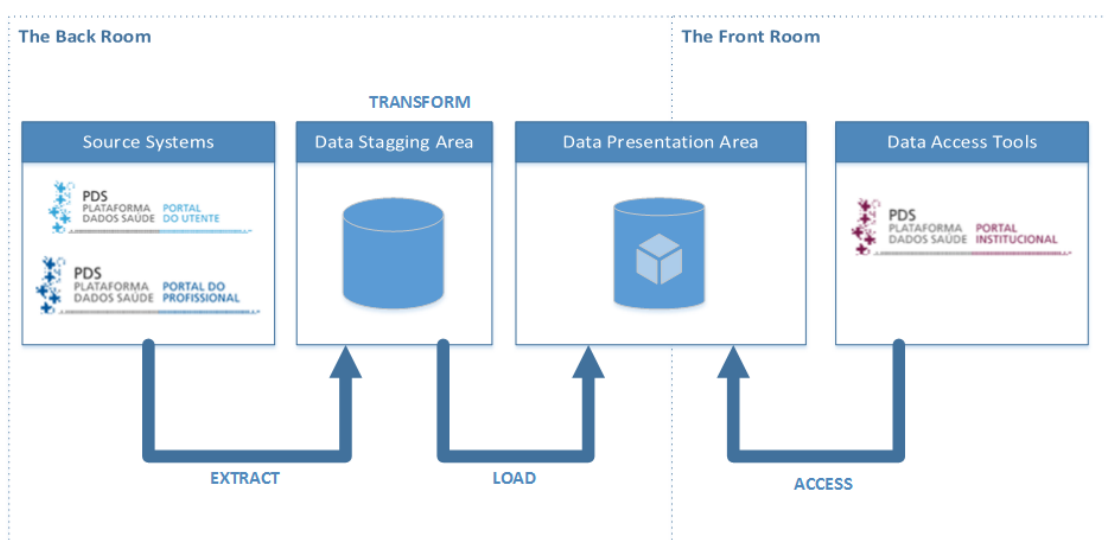


Figura 28 – PDS-PI – Arquitetura do armazém de dados (DW)

Com base no paradigma de *Ralph Kimball*, na Figura 28, é representada a arquitetura proposta para o DW que contém duas áreas principais:

- *The Back Room* – Representa o seguinte grupo de processos: Extração de dados dos sistemas locais para uma área de preparação de dados (*Data Staging Area*); Na área de preparação de dados, é analisada a qualidade dos dados e são aplicadas as transformações necessárias de modo a que a informação seja entregue convenientemente ao DW (*Data Presentation Area*);

- *The Front Room* – Área responsável por disponibilizar mecanismos de consulta sobre os diversos repositórios que constituem o armazém de dados. Recorrendo à PDS-PI, o utilizador poderá visualizar um conjunto de informação adequada ao seu perfil.

4.3 Análise e modelação dimensional

O processo de modelação dimensional exige um estudo prévio do domínio em questão. Portanto, no sentido de consolidar o conhecimento sobre o negócio e sobre as necessidades dos diversos *stakeholders* foi necessário analisar os sistemas transacionais (PDS-PP e PDS-PU), tendo em conta os modelos de dados relacionais e toda a documentação existente.

Durante o processo de análise do modelo relacional, abordado nos pontos 3.2.2 e 3.3.2, foi identificada a possibilidade de desenvolver um modelo dimensional global, devido ao facto de existirem um conjunto de tabelas e respetivos atributos, que são comuns entre os dois sistemas. De seguida, foram identificadas as tabelas dimensão e facto e estabelecidas as respetivas associações, tendo em conta uma modelação de repositórios de dados específicos (DM - *Datamarts*).

Segundo a metodologia de *Kimball* [Ralph Kimball, 2002], o processo de modelação dimensional deverá ser dividido em quatro passos consecutivos. Em cada um desses passos, pretende-se a identificação de diferentes entidades que são relevantes para os DM: Área de negócio; Nível de detalhe (granularidade); Dimensões; Factos.

A maioria das tabelas facto do PDS-PI representam registos de eventos, portanto têm a particularidade de não possuírem medições (são denominadas por *factless tables*). Contudo, também foram identificados diversos atributos descritivos que detêm um conhecimento importante do evento. Pelo facto destes atributos conterem um conjunto reduzido de valores e por forma a isolá-los devidamente, foram criadas tabelas de perfil transacional (denominadas *junk dimensions*), que permitem através de uma dimensão alternativa combinar todos os atributos descritivos.

Recorrendo à metodologia abordada anteriormente, os diversos DM seguem um esquema em estrela, devido ao facto de não existirem hierarquias entre tabelas dimensão (recorrendo a tabelas dimensão desnormalizadas) e estas estarem relacionadas exclusivamente com tabelas facto.

Na Tabela 4 e Figura 29 é descrito, segundo a metodologia de *Kimball*, o modelo dimensional aplicado ao repositório de acessos ao PDS-PP.

Tabela 4 – Metodologia de *Kimball* – Repositório de acessos ao PDS-PP

PDS-PP – Repositório de Acessos	
Área de negócio	Auditoria de acessos ao PDS - Portal do Profissional
Nível de detalhe (granularidade)	Acesso de um utilizador
Dimensões	Utente, Data, Hora, Instituição
Factos	Não contêm medições

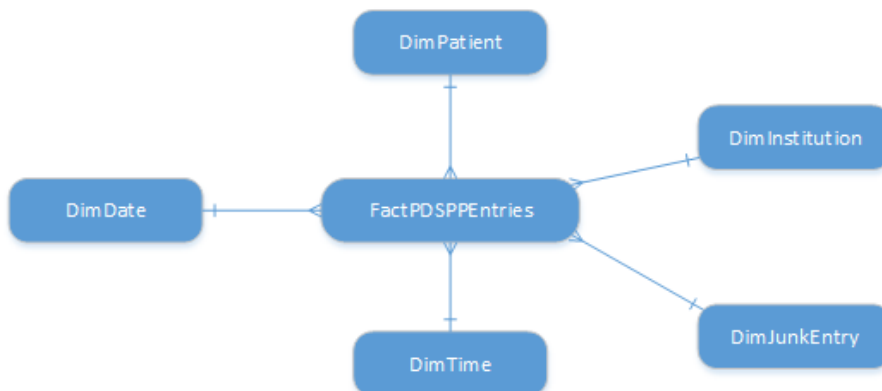


Figura 29 – Modelação dimensional do repositório de acessos ao PDS-PP

Na Tabela 5 e Figura 30 é descrito, segundo a metodologia de *Kimball*, o modelo dimensional aplicado ao repositório operações do PDS-PP.

Tabela 5 – Metodologia de *Kimball* – Repositório de operações no PDS-PP

PDS-PP – Repositório de operações	
Área de negócio	Auditoria de operações efetuadas pelos profissionais no PDS - Portal do Profissional
Nível de detalhe (granularidade)	Registo de operação por utilizador
Dimensões	Utente, Data, Hora, Instituição, Norma
Factos	Não contêm medições

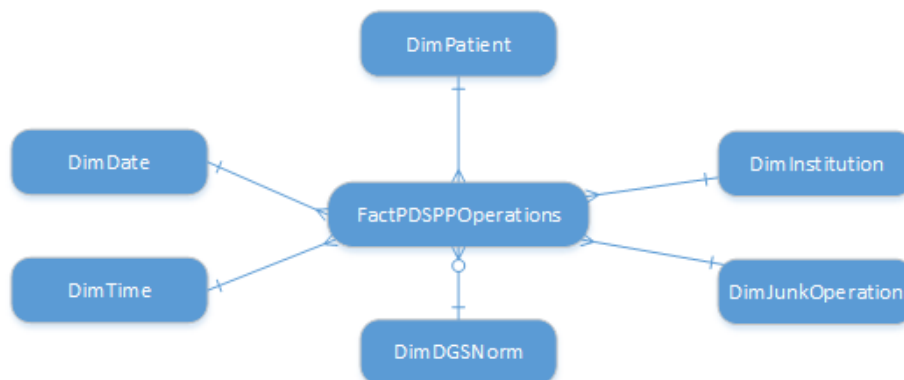


Figura 30 – Modelação dimensional do repositório de operações no PDS-PP

Na Tabela 6 e Figura 31 é descrito, segundo a metodologia de *Kimball*, o modelo dimensional aplicado ao repositório contatos do PDS-PP.

Tabela 6 – Metodologia de *Kimball* – Repositório de contatos

PDS-PP – Repositórios de contatos	
Área de negócio	Análise de contatos/episódios armazenados no repositório do PDS - Portal Profissional
Nível de detalhe (granularidade)	Registo de um episódio para um utente e instituição
Dimensões	Utente, Data, Hora, Instituição
Factos	Duração do contato (Horas)

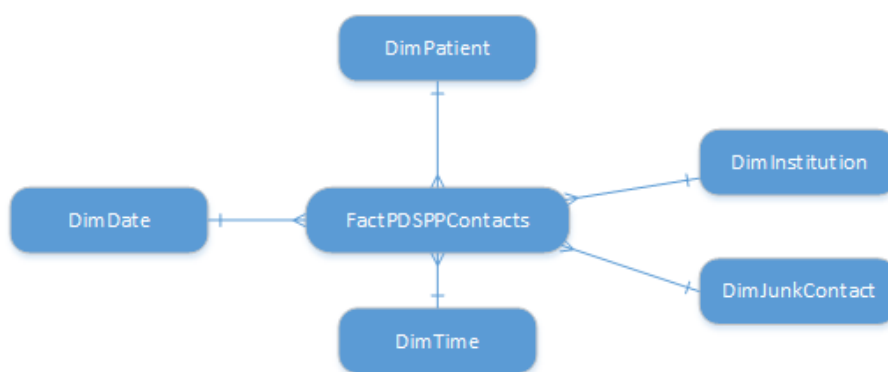


Figura 31 – Modelação dimensional do repositório de contatos

Seguidamente serão descritos os modelos propostos para armazenar os dados provenientes do PDS - Portal Utente. Cada modelo retrata um componente individual do registo pessoal de saúde (PHR): Alergias, Doenças e Medicação.

Na Tabela 7 e Figura 32 é descrito, segundo a metodologia de *Kimball*, o modelo dimensional aplicado ao repositório de registo de alergias do PDS-PU.

Tabela 7 – Metodologia de *Kimball* – PDS-PU repositório de alergias

PDS-PU – Repositório de Alergias	
Área de negócio	Análise das alergias declaradas pelos utentes no PDS – Portal Utente
Nível de detalhe (granularidade)	Evento de registo de alergia
Dimensões	Utente, Data, Alergia
Factos	Não contêm medições

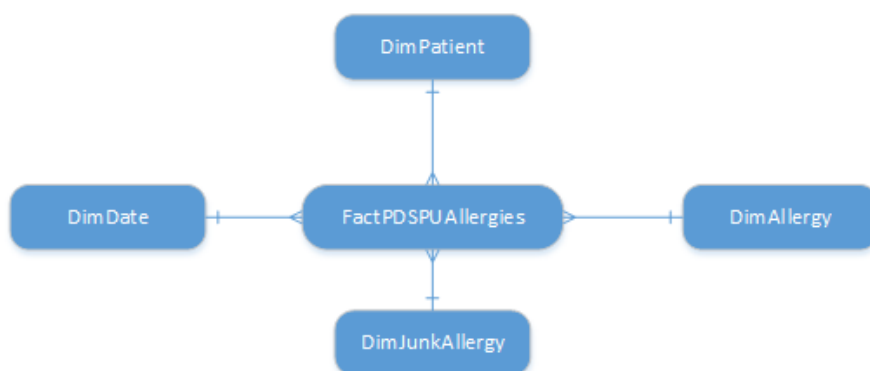


Figura 32 – PDS-PU – Modelação dimensional do repositório de alergias

Na Tabela 8 e Figura 33 é descrito, segundo a metodologia de *Kimball*, o modelo dimensional aplicado ao repositório de registo de doenças do PDS-PU.

Tabela 8 – Metodologia de *Kimball* – PDS-PU repositório de doenças

PDS-PU – Repositório de Doenças	
Área de negócio	Análise das doenças declaradas pelos utentes no PDS – Portal Utente
Nível de detalhe (granularidade)	Evento de registo de doença
Dimensões	Utente, Data, Doença
Factos	Não contêm medições

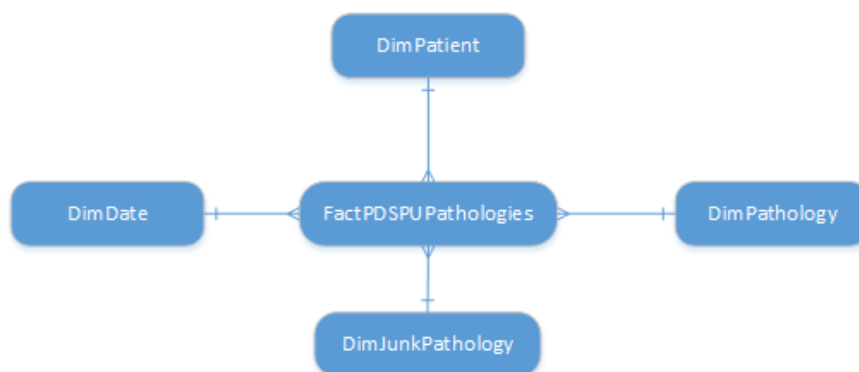


Figura 33 – PDS-PU – Modelação dimensional do repositório de doenças

Na Tabela 9 e Figura 34 é descrito, segundo a metodologia de *Kimball*, o modelo dimensional aplicado ao repositório de registo de medicação do PDS-PU.

Tabela 9 – Metodologia de *Kimball* – PDS-PU repositório de medicação

PDS-PU – Repositório de Medicação	
Área de negócio	Análise da medicação declarada pelos utentes no PDS – Portal Utente
Nível de detalhe (granularidade)	Evento de registo de medicação
Dimensões	Utente, Data, Medicamento
Factos	Não contêm medições

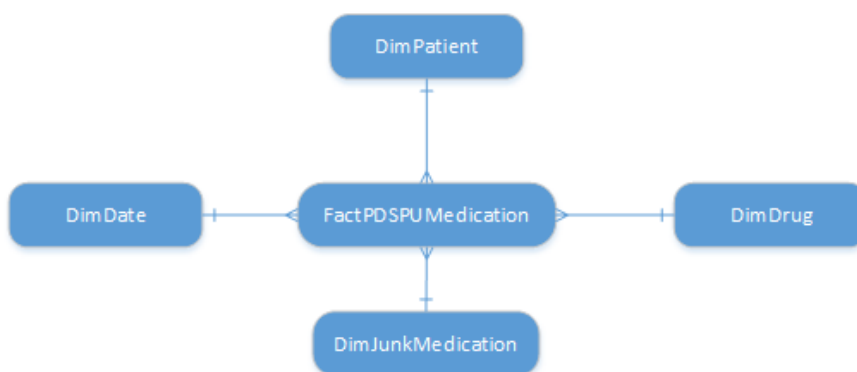


Figura 34 – PDS-PU – Modelação dimensional do repositório de medicação

Por fim, na Tabela 10 – Matriz em bus do modelo dimensional, é apresentado um quadro de resumo dos diversos repositórios que compõem o armazém de dados do PDS-PI.

Tabela 10 – Matriz em bus do modelo dimensional

Processo/Repositório (Facto)	Dimensão	DimDate	DimTime	DimPatient	DimInstitution	DimNorm	DimAllergy	DimDrug	DimJunkEntry	DimJunkOperation	DimJunkContact	DimJunkAllergy	DimJunkPathology	DimJunkMedication
PDS-PP: Acessos (FactPDSPPEntries)		X	X	X	X				X					
PDS-PP: Operações (FactPDSPPOperation)		X	X	X	X	X				X				
PDS-PP: Contatos (FactPDSPPContacts)		X	X	X	X						X			
PDS-PU: Alergias (FactPDSPUAllergies)		X		X			X					X		
PDS-PU: Doenças (FactPDSPUPathologies)		X		X				X					X	
PDS-PU: Medicação (FactPDSPU Medication)		X		X					X					X

4.4 Extrair, Transformar e Armazenar (ETL)

O processo de ETL (Extract Transform Load) é um dos principais componentes da arquitetura do PDS-PI, que apesar de ter uma descrição simples, de modo generalizado traduz uma elevada carga de processamento e complexidade de conceção. Nesta secção é demonstrada uma vista superficial sobre o processo, assim como uma descrição minuciosa das tarefas mais relevantes.

Segundo [Ralph Kimball, 2004], este processo é considerado a fundação de um armazém de dados (DW), pois é necessário garantir que os dados são extraídos dos sistemas origem, assegurar a qualidade e consistência dos dados extraídos, assim com a respetiva consolidação. Mesmo sendo provenientes de fontes diferentes, os dados podem ser conjugados e utilizados de modo a simplificar o processo de entrega dos dados ao DW. O processo de ETL é um trabalho de bastidores, que mesmo não sendo apercebido pelo utilizador final, consome em média 70% dos recursos afetados à implementação e manutenção de um DW.

Para o desenvolvimento deste processo existem duas abordagens possíveis: o desenvolvimento de rotinas recorrendo a uma linguagem de programação ou utilizar uma ferramenta apropriada para este tipo de processos. A segunda opção relevou-se claramente a mais viável pelos seguintes motivos [Ralph Kimball, 2004]:

- Desenvolvimento mais simples, rápido e barato (i.e.: recorrendo a ferramentas *open source*);
- Comunidade e documentação de apoio são recursos valiosos na fase de desenvolvimento, implementação e manutenção;
- Capacidade de ligação a bases de dados de diversos fornecedores;
- Manutenção evolutiva mais facilitada;
- Integra um conjunto de repositórios de metadados que permite o sincronismo com metadados dos sistemas origem, bases de dados de destino e outros sistemas de *Business Intelligence*.

Portanto, para o desenvolvimento do processo de ETL do PDS-PI foi selecionada uma ferramenta *open source*. O *Pentaho Data Integration – Kettle (PDI - Kettle)* para além de ter sido considerada no estudo da *BeyeNETWORK* [Mark Madsen, 2010] como líder, também possui uma comunidade ativa e uma documentação de apoio bastante completa.

4.4.1 Visão geral

Utilizando a ferramenta *PDI – Kettle* [Pentaho, 2013], inicialmente é necessário criar um repositório de projeto que armazene todos os processos (*Jobs*) e tarefas (*Transformation*) que constituem o processo de ETL do PDS-PI.

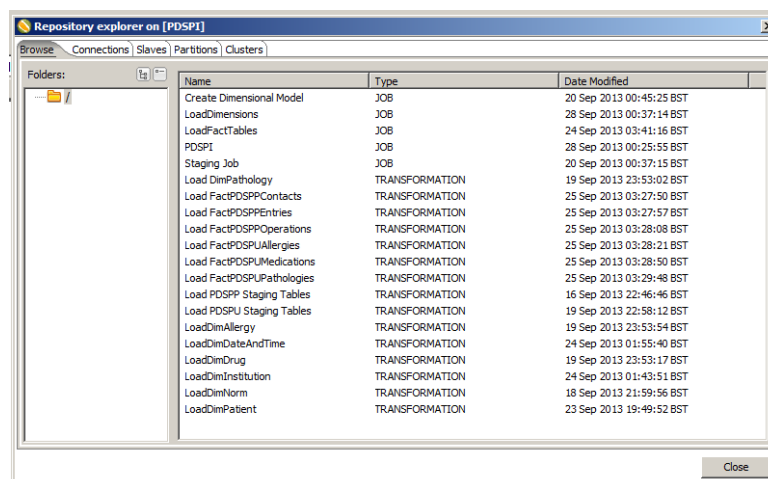


Figura 35 – ETL – Repositório do projeto PDS-PI

O processo de ETL é constituído por um conjunto processos responsáveis pela execução de cada uma das fases inerentes a esta componente.

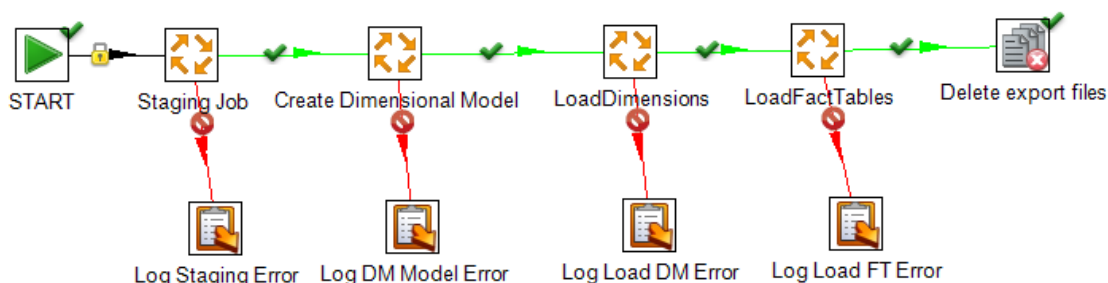


Figura 36 – ETL – Visão geral do processo

Assim, como representado no esquema da Figura 36, existem processos devidamente isolados para cada uma das fases: Processamento da área de preparação dos dados; Criação do modelo dimensional; Carregamento das tabelas dimensão; Carregamento das tabelas facto; Limpeza dos ficheiros provenientes dos sistemas origem.

4.4.2 Extração de dados

O planeamento de uma estratégia de extração é um passo importante para o desenvolvimento do processo ETL. Numa fase inicial é necessário avaliar o fluxo de dados através da criação de um mapa lógico de dados, que documenta a relação entre os atributos dos sistemas origem (PDS-PP e PDS-PU) e os atributos do sistema destino (PDS-PI).

Seguidamente é avaliada qual a latência de extração mais adequada, pelo que no PDS-PI verificou-se a necessidade de extrair os dados diariamente pelos seguintes motivos:

- O PDS-PI é um sistema que deverá ser utilizado diariamente pelos diversos utilizadores;
- Não é mandatário a consulta de dados referentes ao próprio dia;
- A transmissão de dados deverá ser planeada num horário de menor atividade;
- Garantir que a transmissão de dados não prejudica a performance da rede da empresa.

Os sistemas origem serão os responsáveis pela extração diária dos dados, através de uma tarefa agendada e recorrendo à ferramenta *Kitchen* incluída no *PDI – Kettle*, é invocado o respetivo processo. Representado na Figura 37, o processo é constituído por uma tarefa de extração e seguidamente pela transferência dos ficheiros para uma localização no servidor do PDS-PI.

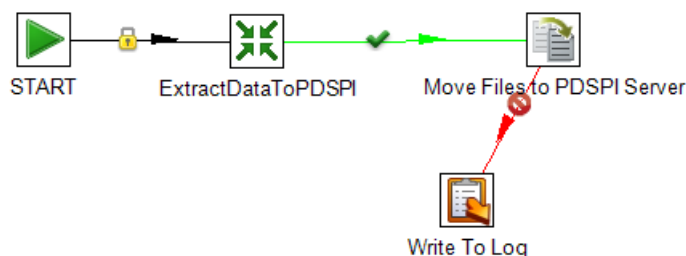


Figura 37 – ETL – Processo de extração de dados do PDS-PU

Os sistemas origem possuem ambientes de base de dados diferentes, no entanto as tarefas de extração são semelhantes. A tarefa de extração é representada na Figura 38 com procedimentos de extração de dados da base de dados local, que estabelecem ligação a uma bases de dados Oracle (PDS-PP) ou SQL Server (PDS-PU), que por fim são entregues a procedimentos que armazenam os dados localmente para ficheiros de texto tabulares (extensão .csv).

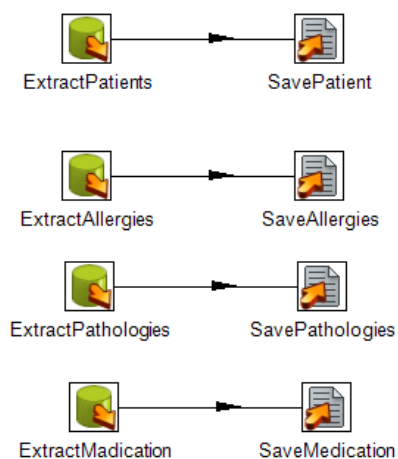


Figura 38 – ETL – Tarefa de extração do PDS-PU

4.4.3 Implementação da área de preparação

Após o armazenamento dos ficheiros provenientes dos sistemas origem, é importante decidir de que forma irão ser processados: em memória a partir da leitura direta dos ficheiros, ou através da criação de uma área de preparação de dados no sistema gestor de base de dados do PDS-PI.

A segunda alternativa revela-se a mais adequada pelos seguintes motivos:

- Capacidade de processamento em memória é inferior comparativamente ao processamento em disco;
- Capacidade de recuperação em caso de falha sem necessitar de iniciar todo o processo;
- Facilitar na auditoria do processo de ETL, recorrendo à análise dos dados extraídos.

Na Figura 39 é representado o processo de criação das tabelas de preparação, que possui procedimentos de limpeza prévia das tabelas (no caso de não ser a primeira execução) e tarefas de carregamento de dados.

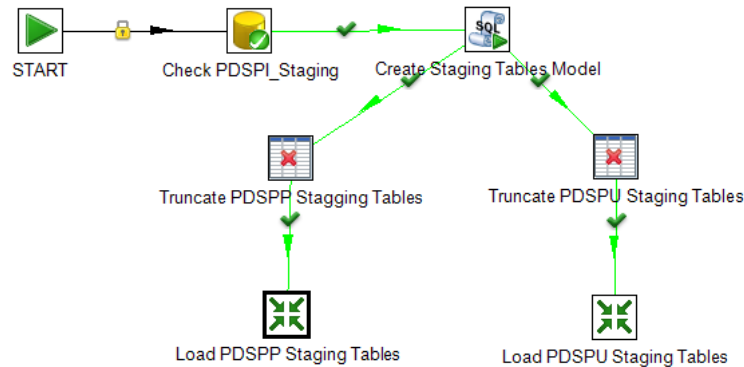


Figura 39 – ETL – Processo de criação e carregamento da área de estágio

O modelo de dados adotado para esta área foi um esquema não relacional, espelhando a estrutura de dados dos sistemas origem. Na Figura 40 é representada a tarefa de carregamento dos dados para as respetivas tabelas de preparação.

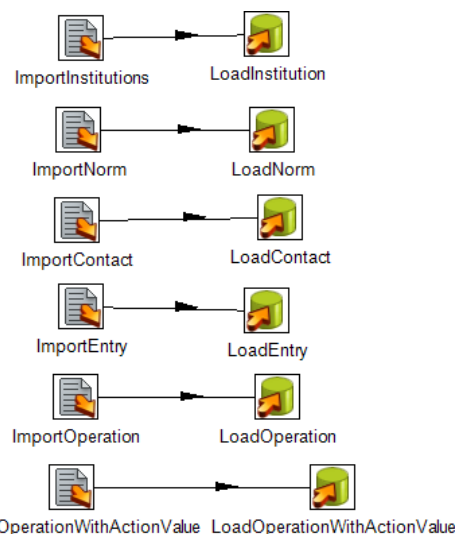


Figura 40 – ETL – Tarefa de carregamento das tabelas de preparação

4.4.4 Criação do modelo dimensional

Após a execução do processo do ponto anterior, sendo a primeira vez, é importante assegurar a criação do modelo dimensional (tabelas dimensão e facto). Na Figura 41 é possível visualizar que após a verificação da disponibilidade da ligação ao armazém de dados, são executados três procedimentos distintos: Execução das instruções SQL para criação das tabelas dimensão; Execução das instruções SQL para criação das tabelas de registo de dados inconsistentes (DQP); Execução das instruções SQL para criação das tabelas facto. As tabelas para registo de problemas de dados (DQP – *Data Quality Problems*) foram introduzidas neste modelo pelo facto de ser relevante existir uma auditoria da qualidade dos dados enviados, possibilitando a definição de uma estratégia de correção e recuperação dos dados inconsistentes.

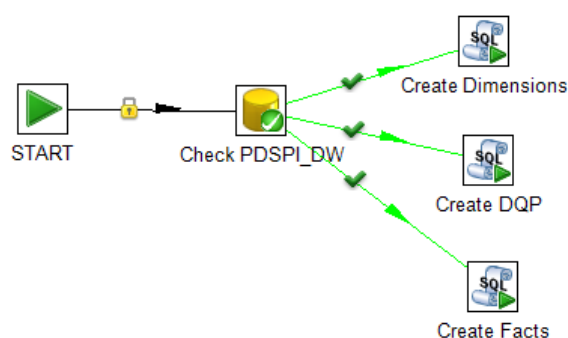


Figura 41 – ETL – Processo de criação do modelo dimensional

4.4.5 Tabelas dimensão

As tabelas dimensão são compostas pelos seguintes tipos de atributos:

- Chave primária (SK - *Surrogate Key*) – Chave única utilizada na relação com as tabelas facto. A geração de valores deste atributo é da responsabilidade do sistema gestor de base de dados (*auto increment*).
- Chave de negócio (NK - *Natural Key*) – Uma ou várias chaves que relacionam o registo em questão com os sistemas origem (i.e.: Nº de utente é utilizado no sistema origem como campo chave).
- Atributos descritivos – São os restantes atributos da tabela que descrevem as características da dimensão no domínio em questão. A quantidade de atributos não tem um limite máximo definido, portanto foi adotado um modelo desnormalizado (*flat dimension*) evitando assim a hierarquia entre dimensões que transformaria o modelo dimensional em estrela num modelo dimensional floco de neve (existência de dependências entre dimensões).

O plano de carregamento de dados para uma tabela dimensão depende do tipo de informação e da proporção da mesma, sendo importante assegurar um conjunto de fatores que garantem a qualidade da entrega dos dados:

- Eliminação de registos em duplicado;
- Limpeza e normalização de dados;
- Garantir a conformidade da informação consoante o domínio em questão.

Por fim, a entrega dos dados deverá ser efetuada seguindo o método conhecido por *slowly changing dimensions* (SCD), que permite definir o modo como os dados vão ser armazenados. Nas tabelas dimensão do PDS-PI os tipos que se aplicam são os seguintes:

- SCD Tipo 1 – Não existe a necessidade de preservar o histórico de alterações, e por esse motivo os dados são atualizados diretamente;
- SCD Tipo 2 – Existe a necessidade de preservar o registo anterior, e por esse motivo é inserido um novo registo, preservando as NK's que identificam nos sistemas origem o registo univocamente. Por conseguinte, é necessário adicionar três novos atributos:
 - Data efetiva e Data de expiração – Estes atributos permitem definir o período em que o registo está ativo.
 - Ativo – Este atributo tem somente dois valores possíveis que indicam se o registo está ativo ou expirado.

Na Figura 42, é apresentado o processo de carregamento das tabelas dimensão, que é composto por um conjunto de tarefas. Cada tarefa é responsável por assegurar a importação dos dados provenientes da área de preparação, verificar da qualidade dos dados, aplicar as transformações necessárias e entregar os dados pré-processados na dimensão correspondente.

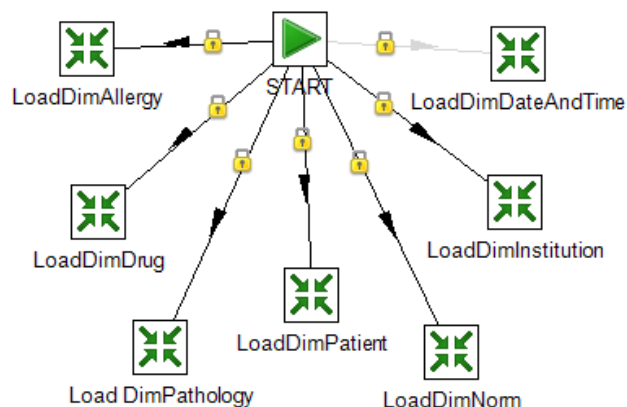


Figura 42 – ETL – Processo de carregamento das dimensões

De modo a resumir o processo anterior, serão descritas as tarefas consoante o tipo de SCD das tabelas dimensão. Na Tabela 11 são enumeradas as dimensões cujo método de armazenamento é SCD tipo 1.

Tabela 11 – ETL – Identificação das dimensões SCD Tipo 1

Área do domínio	Dimensão
Alergias	DimAllergy
Medicamento	DimDrug
Doença	DimPathology
Norma	DimNorm
Instituição	DimInstitution

As tarefas associadas ao carregamento de dimensões SCD tipo 1, como representado na Figura 43, contêm os seguintes procedimentos:

- Extração dos dados da respetiva tabela de preparação;
- Aplicação das transformações necessárias (Anexo 2 – Exemplos de transformação de dados);
- Recorrendo às NK's definidas, verificar se os registos existem na tabela dimensão de destino (Anexo 1 – Exemplo de procura de dados numa tabela dimensão);
- No caso de existir o registo na tabela de dimensão, os atributos descritivos são atualizados, caso contrário é inserido um novo registo. Estes procedimentos são executados em série (*bulk loading*) garantindo uma melhor performance na execução desta tarefa.

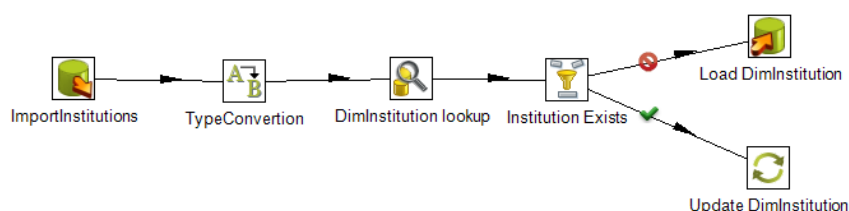


Figura 43 – ETL – Tarefa de carregamento da dimensão instituição

A dimensão utente (*DimPatient*) foi identificada como SCD tipo 2, e tem a particularidade de ter grande proporção (*big table*) e dos dados serem provenientes de diversas fontes.

Para este tipo de dimensões, quando um atributo específico é alterado, é possível configurar qual o tipo SCD a que o atributo corresponde. Na Tabela 12, são enumerados os atributos descritivos do *DimPatient* mencionando o respetivo tipo SCD.

Tabela 12 – ETL – Tipo SCD dos atributos descritivos da dimensão DimPatient

Atributos	Tipo SCD
Gender	Tipo 2
BirthDate	Tipo 1
ProviderName	Tipo 2
ACESName	Tipo 2
RegionName	Tipo 2
DistrictName	Tipo 2
CouncilName	Tipo 2
isAuthPDS	Tipo 1
isAuthPHR	Tipo 1
isAuthEPSOS	Tipo 1
PHRRegistrationDate	Tipo 1

Na Figura 44 é representada a tarefa associada ao carregamento da dimensão utente (*DimPatient*), que contém os seguintes procedimentos:

- Extração dos dados dos utentes a partir de diversas tabelas de preparação;
- Normalização dos dados de modo a uniformizar a informação dos utentes num único conjunto;
- Pelo facto dos dados serem provenientes de mais do que uma origem, é importante previamente ordenar o conjunto de dados pela NK e de seguida remover os registos duplicados;
- Aplicar as transformações necessárias aos atributos descritivos;
- Recorrendo ao identificador único de utente (NK) é verificado se já existe registo na tabela dimensão;
- Caso o registo não exista é inserido na tabela dimensão, caso contrário é aplicada a alteração através de um componente específico (*Dimension Lookup/Update* - Anexo 3 – Exemplo de procedimento para carregamento de dimensões do tipo 2), que consoante o tipo do atributo alterado permite aplicar as alterações correspondentes.

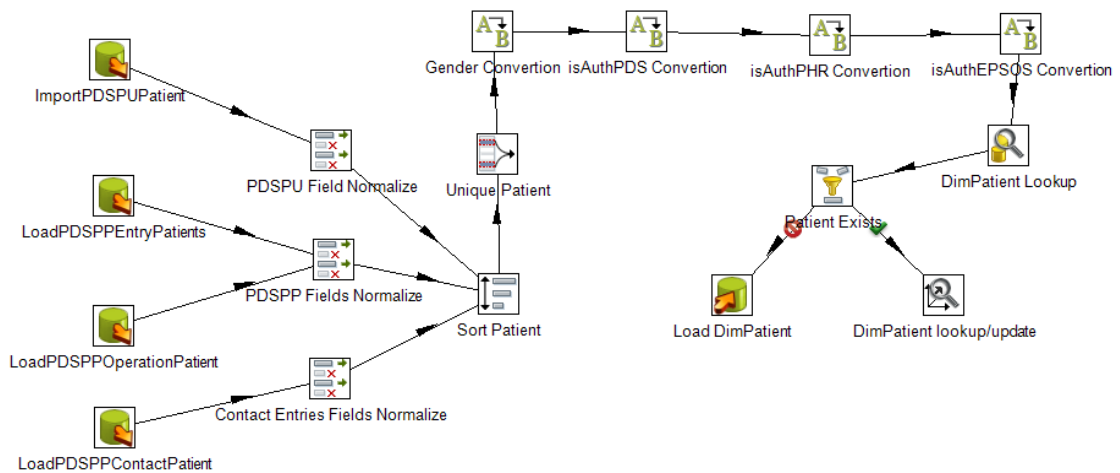


Figura 44 – ETL – Tarefa de carregamento da dimensão utente

Qualquer evento ou transação ocorre num determinado instante de tempo, portanto uma das dimensões fundamentais em qualquer armazém de dados é a dimensão data. Para um maior detalhe temporal também pode ser considerada a dimensão hora. No caso do PDS-PI, para além de armazenar o dia que ocorre um evento também é importante avaliar a hora. Portanto, estes dados após serem gerados manualmente, em ficheiros de texto tabular (.csv), na primeira execução do processo a tarefa de carregamento representada na Figura 45 é executada.

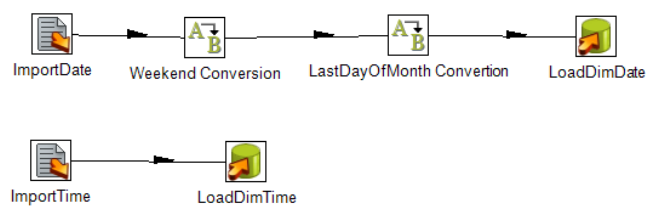


Figura 45 – ETL – Tarefa de carregamento das dimensões data e hora

4.4.6 Tabelas facto

Nos modelos dimensionais em estrela ou floco de neve, as tabelas facto representam a tabela central que contém as métricas, medições ou factos de um determinado processo de negócio. Esta tabela também é composta por um conjunto de campos chave (FK – chave estrangeira) que representam a ligação às dimensões respetivas, e por vezes por uma chave de dimensão degenerada (DD) que representa uma dimensão sem atributos descritivos, que justifique a criação de uma dimensão em separado (i.e: identificação do episódio no repositório de contactos).

No modelo proposto para o PDS-PI, existe um conjunto de tabelas facto que partilham dimensões em comum, tornando o modelo dimensional num esquema em constelação.

Na Figura 46 é apresentado o processo de carregamento das tabelas facto, que é composto por um conjunto de tarefas de carregamento de dados. Cada tarefa é responsável por assegurar um conjunto de procedimentos:

- Extração dos dados provenientes da área de preparação;
- Garantir a integridade referencial com as tabelas dimensão;
- Armazenar em tabelas de problemas de dados (DQP – *Data Quality Problems*) os registos inconsistentes;
- Avaliar a qualidade dos dados;
- Aplicar transformações a atributos específicos;
- Carregamento de atributos descritivos do facto para uma dimensão de perfil transaccional (*junk dimension*);
- Entrega dos dados pré-processados na tabela facto correspondente.

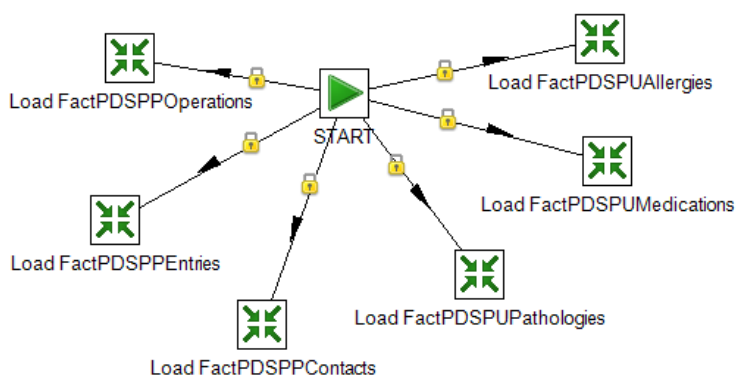


Figura 46 – ETL – Processo de carregamento das tabelas facto

Na Figura 47 é representada a tarefa de carregamento da tabela facto responsável por armazenar eventos de entrada no PDS-PP. Após a extração dos dados é verificada a integridade referencial com as tabelas dimensão, caso esta condição não seja assegurada os dados são entregues a uma área de problemas de dados (DQP). A área DQP é representada por um conjunto de tabelas, semelhantes à estrutura do sistema origem com um campo adicional que permite armazenar a descrição do problema. De seguida, são removidos todos os registos em duplicado e armazenados os atributos descritivos do facto numa tabela dimensão *junk* (exemplificado no Anexo 4 – Exemplo de procedimento para carregar dados numa dimensão *junk*). Por fim, é verificado se o facto já existe e caso seja identificado o

registo, este é descartado devido a não serem consideradas atualizações de registos no PDS-PP. Caso contrário, os registos são inseridos em série (*bulk loading*) na tabela de destino.

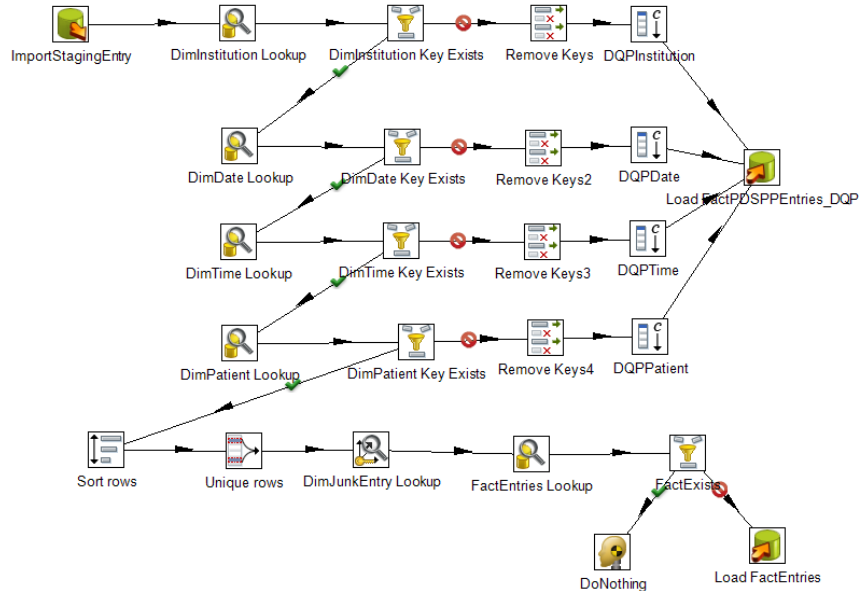


Figura 47 – ETL – Carregamento da tabela de facto FactPDSPPEntries

Na Tabela 13 – Detalhes da tabela de facto FactPDSPPEntries são descritos os atributos que constituem a tabela em questão e a dimensão *junk* gerada.

Tabela 13 – Detalhes da tabela de facto FactPDSPPEntries

Atributos	Detalhe
DateKey TimeKey InstitutionKey PatientKey	FK – Chaves estrangeiras para as dimensões
JunkKey	FK – Chave estrangeira para a dimensão <i>junk</i> que contem: <ul style="list-style-type: none"> • <i>Professional</i> – Identificação do profissional • <i>ProfessionalCategory</i> – Médico/Enfermeiro • <i>ContactType</i> – Tipo de Episódio • <i>LocalApplication</i> – Nome do sistema local utilizado
GUID	DD – Identificador único da entrada na PDS-PP

O registo de operações no PDS-PP é armazenado na tabela facto FactPDSPPOperations, recorrendo à tarefa representada na Figura 92 do Anexo 5 – Tarefas de carregamento das tabelas de facto. Nesta tabela existe a particularidade da existência de dois atributos não obrigatórios, que permitem identificar com maior detalhe a operação em questão. Após o processo de verificação da integridade referencial dos atributos obrigatórios, na Figura 48 é

possível visualizar a verificação de dois atributos, que apesar de estarem relacionados com uma tabela dimensão, não são de registo obrigatório. Os procedimentos seguintes são semelhantes à tabela facto - FactPDSPPEntries.

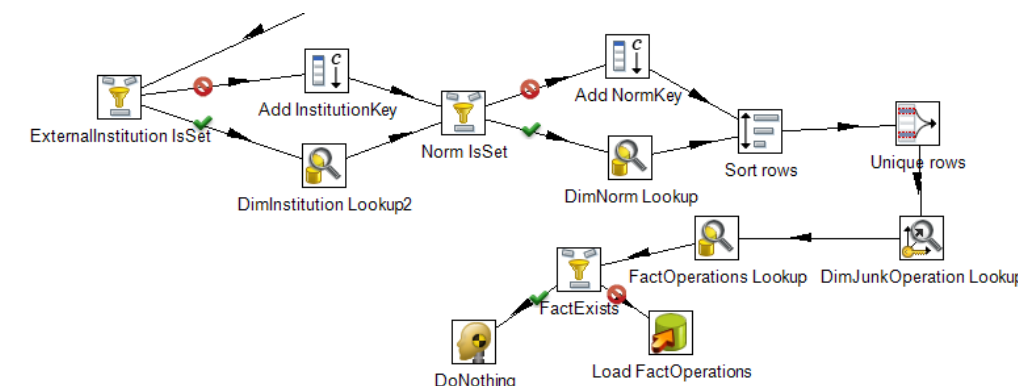


Figura 48 – ETL – Especificidade do carregamento da tabela de facto FactPDSPPOperations

Na Tabela 14 são descritos os atributos que constituem a tabela em questão e a dimensão *junk* gerada.

Tabela 14 – Detalhes da tabela de facto FactPDSPPOperations

Atributos	Detalhe
DateKey TimeKey InstitutionKey PatientKey	FK – Chaves estrangeiras para as dimensões
JunkKey	FK – Chave estrangeira para a dimensão <i>junk</i> que contem: <ul style="list-style-type: none"> • <i>Professional</i> – Identificação do profissional • <i>ProfessionalCategory</i> – Médico/Enfermeiro • <i>OperationName</i> – Operação executada
LOGID	DD – Identificador único para a operação no PDS-PP
ExternallInstitutionKey	FK – Chave estrangeira (não obrigatória) para a dimensão DimInstitution – Permite avaliar quais as instituições de referência do utente e que PCE o profissional consulta
NormKey	FK - Chave estrangeira (não obrigatória) para a dimensão DimNorm – Permite avaliar a Norma de Orientação Clínica consultada pelo profissional

O repositório de contactos existente do sistema origem contém o registo dos diversos episódios realizados numa instituição de saúde, assim como as marcações futuras. A tabela de facto FactPDSPPContacts, pretende refletir o histórico de contactos realizados nas instituições, e por esse motivo as marcações foram descartadas após a extração dos dados da área de preparação. Na Figura 93 do Anexo 5 – Tarefas de carregamento das tabelas de facto é

retratada a tarefa de carregamento da tabela *facto*, sendo esta única que contem medições. Nesta tabela um dos objetivos é fornecer a métrica de tempo de duração de um episódio (em horas), que é calculado no processo de extração dos dados através da seguinte instrução: `datediff(HH, [DATE_START], [DATE_END])`.

Na Tabela 15 são descritos os atributos que constituem a tabela em questão e a dimensão *junk* gerada.

Tabela 15 – Detalhes da tabela de facto FactPDSPPContacts

Atributos	Detalhe
DateKey TimeKey InstitutionKey PatientKey	FK – Chaves estrangeiras para as dimensões
JunkKey	FK – Chave estrangeira para a dimensão <i>junk</i> que contem: <ul style="list-style-type: none"> • <i>ContactType</i> – Tipo de episódio • <i>Speciality</i> – Especialidade do episódio • <i>HasExamResults</i> e <i>HasAnalysisResults</i> – Indica se no episódio em questão forma disponibilizados resultados de exames ou análises
ContactCode	DD – Identificador único do episódio na instituição
HoursSpent	Métrica (não obrigatória) a duração em horas desde a entrada o utente até à alta (fim do episódio). No caso de inexistência de data e hora de fim do episódio ou caso esta seja semelhante à data e hora de início, o valor considerado é nulo.

Quanto à tarefa de carregamento das tabelas *facto* do PDS-PU, os procedimentos aplicados são bastante semelhantes. Na Figura 49 é representada a tarefa de carregamento da tabela *facto* FactPDSPUAllergies, responsável pelo armazenamento dos eventos de registo de alergias. No caso de o *facto* existir, no procedimento de entrega dos dados, é possível corrigir alguns atributos (i.e.: data de fim de uma alergia).

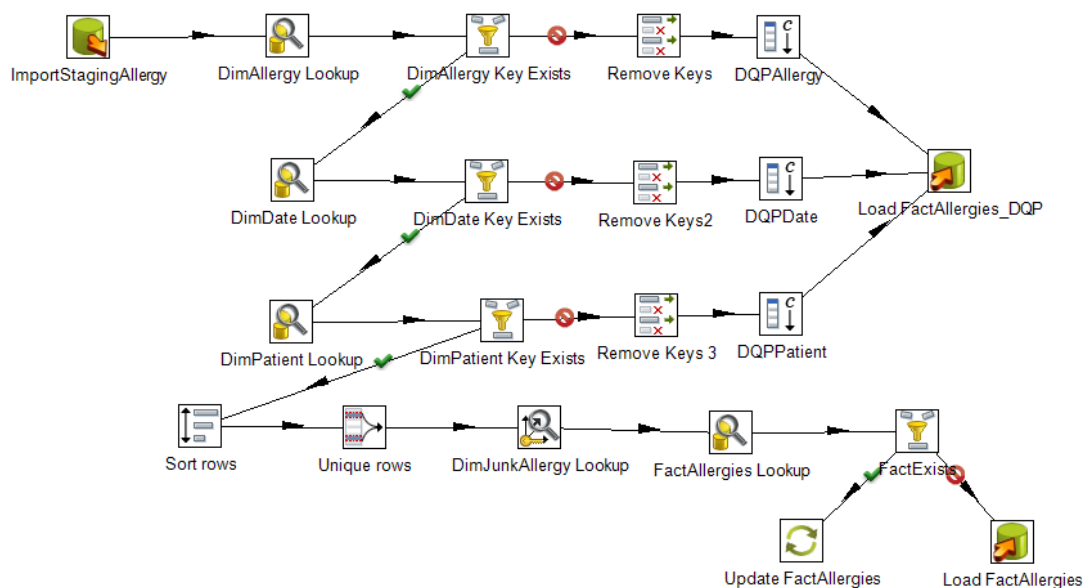


Figura 49 – ETL – Carregamento da tabela de facto FactPDSPUAllergies

Na Tabela 16 são descritos os atributos que constituem a tabela em questão e a dimensão *junk* gerada.

Tabela 16 – Detalhes da tabela de facto FactPDSPUAllergies

Atributos	Detalhe
DateKey PatientKey AllergyKey	FK – Chaves estrangeiras para as dimensões
JunkKey	FK – Chave estrangeira para a dimensão <i>junk</i> que contem: <ul style="list-style-type: none"> • <i>Reaction</i> – Tipo de reação • <i>Severity</i> – Grau de severidade da alergia
StartDate FinishDate DeleteDate	<i>SmartKeys</i> – Atributos de data (não obrigatório), que não possui relação direta com a dimensão DimDate, no entanto permite o particionamento da data e, se necessário, a ligação à tabela de dimensão através de uma <i>smartkey</i> idêntica.

Na Figura 94 do Anexo 5 – Tarefas de carregamento das tabelas de facto, é representada a tarefa de carregamento da tabela facto FactPDSPUMedications, responsável pelo armazenamento dos eventos de registo de medicação. Na Tabela 17 são descritos os atributos que a constituem e a dimensão *junk* gerada.

Tabela 17 – Detalhes da tabela de facto FactPDSPUMedications

Atributos	Detalhe
DateKey PatientKey DrugKey	FK – Chaves estrangeiras para as dimensões
JunkKey	FK – Chave estrangeira para a dimensão <i>junk</i> que contem: <ul style="list-style-type: none"> • <i>TypeOfTake</i> – Via de toma da medicação • <i>State</i> – Estado da medicação (Ativa/Inativa) • <i>Duration</i> – Duração do tratamento • <i>Frequency</i> – Frequência de toma
StartDate DeleteDate	<i>SmartKeys</i> – Atributos de data (não obrigatório), que não possui relação direta com a dimensão DimDate, no entanto permite o particionamento da data e, se necessário, a ligação à tabela de dimensão através de uma <i>smartkey</i> idêntica.

Na Figura 95 do Anexo 5 – Tarefas de carregamento das tabelas de facto, é representada a tarefa de carregamento da tabela facto FactPDSPUPathologies, responsável pelo armazenamento dos eventos de registo de doenças. Na Tabela 18 são descritos os atributos que a constituem e a dimensão *junk* gerada.

Tabela 18 – Detalhes da tabela de facto FactPDSPUPathologies

Atributos	Detalhe
DateKey PatientKey PathologyKey	FK – Chaves estrangeiras para as dimensões
JunkKey	FK – Chave estrangeira para a dimensão <i>junk</i> que contem: <ul style="list-style-type: none"> • <i>OtherPathology</i> – Descrição alternativa da doença • <i>State</i> – Estado da doença (Ativa/Inativa)
StartDate FinishDate DeleteDate	<i>SmartKeys</i> – Atributos de data (não obrigatório), que não possui relação direta com a dimensão DimDate, no entanto permite o particionamento da data e, se necessário, a ligação à tabela de dimensão através de uma <i>smartkey</i> idêntica.

4.4.7 Estimativa de crescimento do armazém de dados

O fenómeno do crescimento constante da informação, externa e interna ao armazém de dados, resulta da consequência da evolução e utilização dos sistemas operacionais. O processo de ETL deve ser devidamente dimensionado, de modo a evitar as seguintes restrições:

- Processamento (CPU)
- Armazenamento interno (Memória)
- Armazenamento externo (Disco - Input/Output)
- Recursos de rede

Devido ao facto da infraestrutura envolvida ser virtualizada, é sempre possível monitorizar esse processo progressivamente e, caso necessário expandir adicionando mais recursos. Relativamente ao tráfego de rede, a transferência de ficheiros é realizada diariamente a um horário de menor atividade, sendo que o volume de dados transferido não ultrapassa os cem Megabytes diários. Portanto, de um modo geral são asseguradas as condições necessárias que garantem escalabilidade do processo de ETL e do armazém de dados.

Seguidamente, na Tabela 19, será mencionado o volume de dados envolvido no carregamento inicial e estabelecida uma estimativa de crescimento mensal do armazém de dados. O carregamento inicial teve em conta um ano de funcionamento do PDS-PP e dois anos de funcionamento do PDS-PU. Devido ao facto do repositório de episódios ter sido implementado nos últimos quatro meses, a estimativa foi calculada com base na média de registos inseridos nesse período. Para a conversão do valor do total de registos em unidades de armazenamento (*Megabytes*), foi criada uma tabela de auditoria e foram monitorizadas todas as tabelas durante o processo de ETL, recorrendo ao procedimento de base de dados existente no *SQL Server*: `sp_spaceused '[table_name]'`.

Tabela 19 – PDS-PI – Volume de dados inicial e estimativa de crescimento mensal

Área	Volume Inicial	Estimativa Mensal
Preparação de dados (<i>Staging area</i>)	7.809.877 registos 968,78 Megabytes	Limpeza diária (sem impacto)
Armazém de dados (<i>Presentation area</i>)	8.208.785 registos 924,25 Megabytes	Aproximadamente 143 Megabytes
Tabelas Dimensão	2.258.919 registos 369,33 Megabytes	Aproximadamente 5 Megabytes
Tabelas Facto	5.949.866 registos 554,93 Megabytes	Aproximadamente 138 Megabytes

Segundo os dados apresentados anteriormente, conclui-se que o armazém de dados tem um crescimento anual de aproximadamente 1,68 *Gigabytes*. Esta estimativa é importante para o dimensionamento inicial da infraestrutura, no entanto é necessário acrescentar armazenamento adicional, visto esta previsão não ter em conta o desenvolvimento de novos repositórios de dados que poderão surgir ao longo do ano.

4.5 Apresentação da informação

A ferramenta de acesso à informação é a componente do *front-office* responsável por retribuir ao utilizador final um conjunto de dados analíticos relevantes ao processo de tomada de decisão. A ferramenta selecionada é responsável pelo interface do PDS-Portal Institucional, e deverá disponibilizar as seguintes funcionalidades:

- Gestão de utilizadores e perfis de acesso;
- Desenvolvimento de análises e relatórios *ad-hoc*;
- Consulta de relatórios adequados ao perfil em questão;
- Capacidade para expandir para cenários móveis.

O processo de seleção teve em conta o cumprimento integral do objetivos definidos para o projeto. Devido à existência de licenciamento empresarial, a ferramenta selecionada foi *Microstrategy 9.3*.

Na plataforma *Microstrategy* existe um conjunto diverso de componentes dos quais destacam-se os seguintes [Microstrategy, 2012]:

- *Desktop products* - Fornece um conjunto de funcionalidades analíticas que facilitam o processo de desenvolvimento e implementação de relatórios. Permite gerir um conjunto de objetos aplicativos como: relatórios, filtros e métricas; assim como objetos de esquema de dados criados no *Microstrategy Architect* (aplicação utilizada como ferramenta de modelação lógica e desenho do fluxo de dados do projeto);
- *Intelligence Server* – É um servidor analítico que representa a fundação da plataforma BI *Microstrategy*. Permite a execução de consulta de dados, cálculos, produção de relatórios e análise OLAP (*On-line Analytical Processing*), tendo as principais funções:
 - Partilha de objetos;
 - Partilha de dados;
 - Gere a partilha de dados e objetos através de um ambiente seguro;
 - Protege a informação no repositório de metadados.
- *Microstrategy Web* – Aplicação Web que permite fornecer aos utilizadores, um ambiente interativo para produção e consulta de relatórios/análises. Utilizando esta aplicação o utilizador final pode aceder, analisar e partilhar dados corporativos através de qualquer *browser* e sistema operativo.

- *Microstrategy Mobile* – Interface interativo da plataforma de BI que capacita os utilizadores do acesso a dados analíticos através do uso de dispositivos móveis: *BlackBerry, iPhone, Ipad e Android*.

A instalação da plataforma deverá ser executada recorrendo a uma de três configurações distintas:

- Dois níveis (*Direct* ou *Two-Tier*) – Comunicação direta entre o projeto fonte e o repositório de metadados;
- Três níveis (*Three-Tier*) – A comunicação entre o projeto fonte e repositório de dados é estabelecido através do *Intelligence Server*;
- Quatro níveis (*Four-Tier*) – É semelhante ao ponto anterior, com a adição de um servidor Web (*Microstrategy Web*) que comunica com o *Intelligence Server*.

A configuração de quatro níveis (*Four-Tier*), como representado na Figura 50, é a recomendada, garantido assim o cumprimento do objetivos delineados para o PDS-PI.

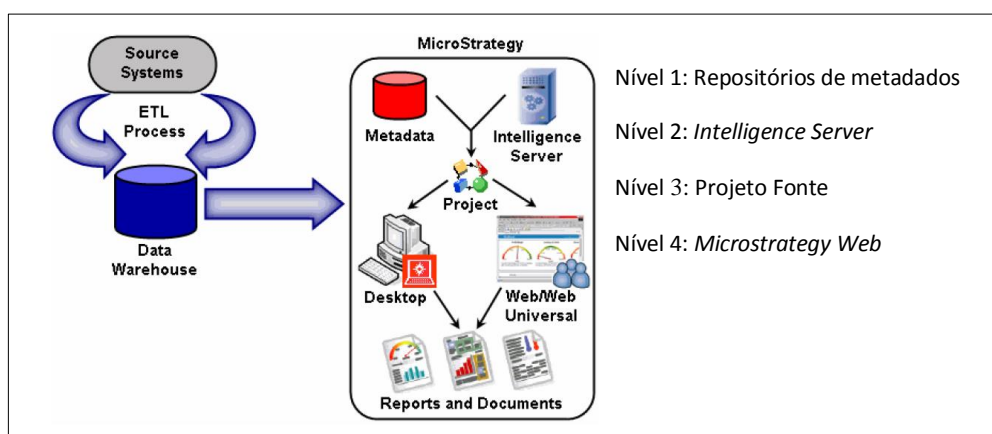


Figura 50 – Configuração da instalação do *Microstrategy* em quatro níveis

Finalizada a instalação e antes da criação do projeto PDS-PI, foram configurados os seguintes repositórios de dados:

- *Metadata* (Obrigatório) – Contém um conjunto de tabelas que armazena as definições dos objetos do *Microstrategy* (ligações e instâncias de bases de dados, definições do servidor, relatórios, atributos, factos, métricas, entre outros);
- *History List* (Opcional) – Repositório responsável por armazenar resultados de relatórios e documentos gerados pelo utilizador para uma utilização futura;
- *Statistics* (Opcional) – Conjunto de tabelas utilizadas para suportar e monitorizar a atividade e performance do sistema.

Como representado na Figura 51, recorrendo ao componente *Microstrategy Desktop*, foi criado o projeto fonte PDS – Portal Institucional referenciando o armazém de dados correspondente.



Figura 51 – Componente *Microstrategy Desktop* do PDS-PI

Durante a criação do projeto, foi necessário efetuar os seguintes passos:

1. Configurar a ligação do repositório de metadados ao armazém de dados;
2. Criar a definição do projeto fonte (PDS – Portal Institucional);
3. Criar objetos de esquema de dados (atributos e factos) – Representado no Anexo 6 – *Microstrategy Architect* recorrendo ao componente *Microstrategy Architect*;
4. Configurações adicionais ao esquema de dados
 - a. Criação de hierarquias (facilitando a navegação entre atributos) – Representado na Figura 52;
 - b. Criação de transformações.

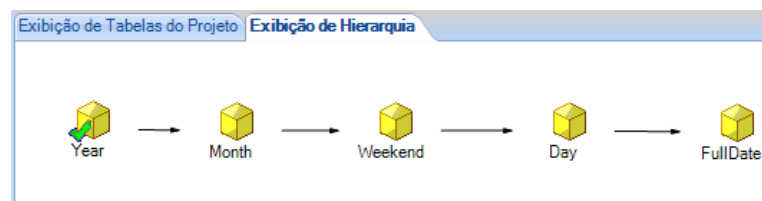


Figura 52 – *Microstrategy Architect* – Criação de hierarquias

Finalizado o processo de criação do projeto, o sistema está preparado para a criação e partilha de relatórios, análises e documentos. No ponto seguinte será contextualizada a componente Web e o mecanismo de gestão de utilizadores e perfis de acesso.

4.5.1 Gestão de utilizadores e perfis

Após a criação do projeto PDS-Portal Institucional, recorrendo ao componente *Microstrategy Web* representado na Figura 53, foi disponibilizado de imediato o utilizador administrador do portal.

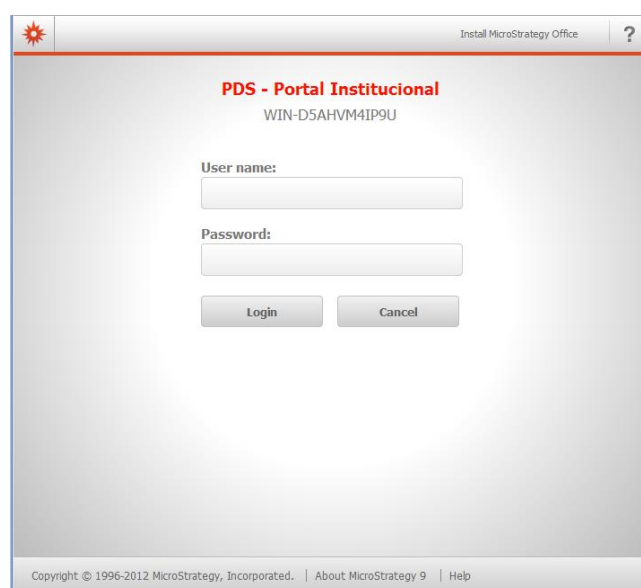


Figura 53 – *Microstrategy Web* - Ecrã de autenticação do Portal Institucional

Na Figura 97 do Anexo 7 – *Microstrategy Web* (PDS-PI), é representado o ambiente disponibilizado após autenticação que permite as seguintes funcionalidades:

- Consulta e pesquisa de relatórios;
- Criação de relatórios, documentos, filtros e análise de dados *ad-hoc*;
- Configuração do sistema e gestão do servidor *Intelligence Server*

A configuração de utilizadores é efetuada através do ambiente de gestão do *Intelligence Server*, que permite configurar:

- Perfis de acesso e autorizações – Como exemplificado na Figura 98 do Anexo 7 – *Microstrategy Web* (PDS-PI);
- Grupos de utilizadores – Criação de grupos de acesso ao PDS-PI (Figura 54);

- Utilizadores – Criação de utilizadores (exemplificado na Figura 55) que poderão ser associados a um ou mais grupos.

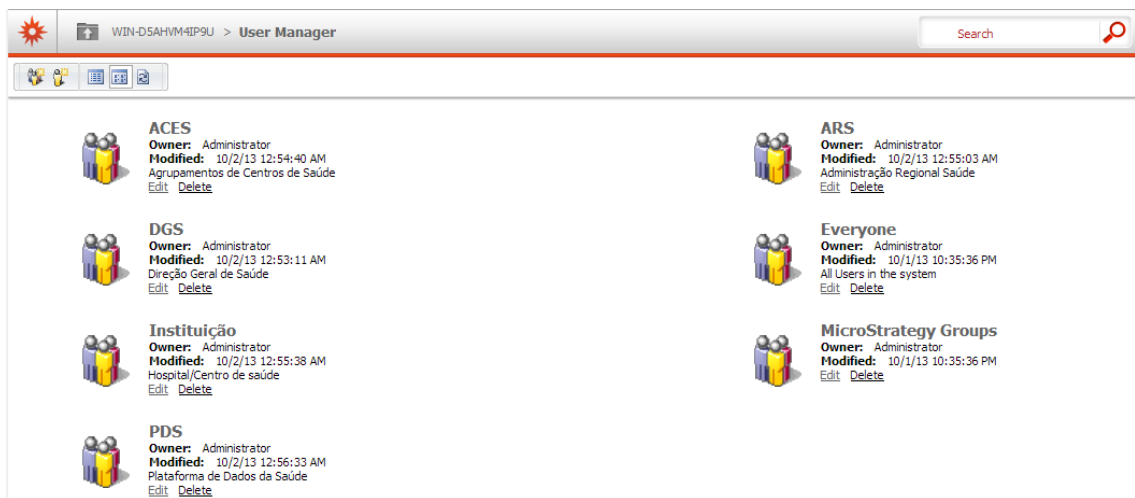


Figura 54 – PDS - Portal Institucional – Criação de grupos de utilizadores

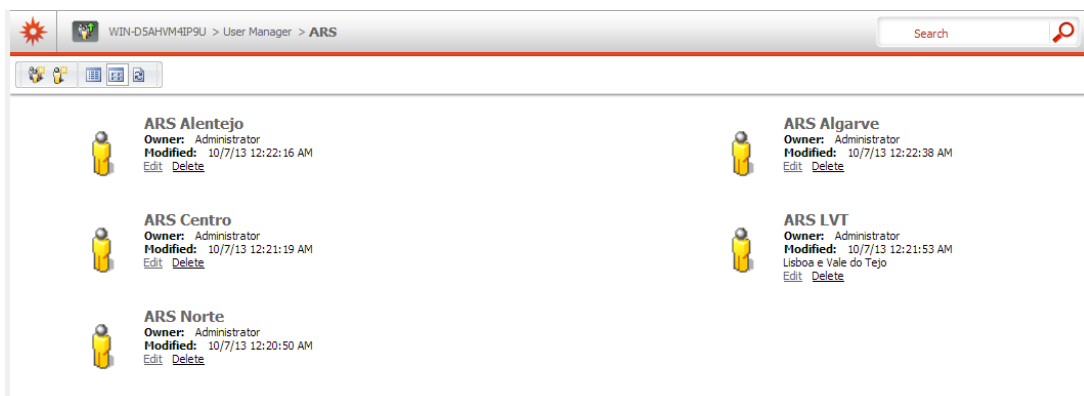


Figura 55 – PDS - Portal Institucional – Criação de utilizadores do grupo ARS

De seguida, foram explorados os dados armazenados no DW de modo a demonstrar algumas análises que poderão ser realizadas após a implementação do PDS-PI, e por conseguinte redigir algumas conclusões relevantes no âmbito da utilização dos portais (PDS-PP e PDS-PU).

5 Exploração de dados

Neste capítulo serão apresentadas diversas análises que sustentam os indicadores de desempenho e utilização dos portais PDS-Portal Profissional (PDS-PP) e PDS-Portal Utente (PDS-PU), recorrendo aos dados disponibilizados pelo PDS-Portal Institucional (PDS-PI). Os dados extraídos permitem realizar um balanço geral de um ano de projeto PDS-PP e aproximadamente dois anos de utilização do PDS-PU.

No Portal Institucional estão armazenados dados de 2.700.345 utentes e 532 instituições: 74 hospitais, 457 unidades de saúde (CSP - Cuidados de Saúde Primários) e região autónoma da Madeira, que possui um único sistema unificado à região.

Nos pontos seguintes serão aprofundados os seguintes temas por tipo de aplicação:

- PDS-PP:
 - Análise de acessos ao portal;
 - Operações realizadas:
 - Acessibilidade a determinadas áreas do portal;
 - Instituições visitadas;
 - Consulta de normas de orientação clínica.
 - Análise de episódios enviados para o repositório central.
- PDS – PU:
 - Análise do registo de utentes e autorizações;

- Análise das alergias, doenças e medicação registadas pelos utentes.

5.1 PDS - Portal Profissional

O PDS-PP está operacional desde Julho de 2012, e passado um ano acumulou o total de 1.792.800 acessos. Pelo facto da ARS Norte ter iniciado o projeto, até esta data tem sido a região que mais impulsiona a plataforma, como representado na Figura 56. Segundo estes resultados deverão ser efetuadas ações de divulgação nas restantes regiões, principalmente na região centro e Lisboa que inclui cerca de 270 instituições (51% das instituições de saúde portuguesas).

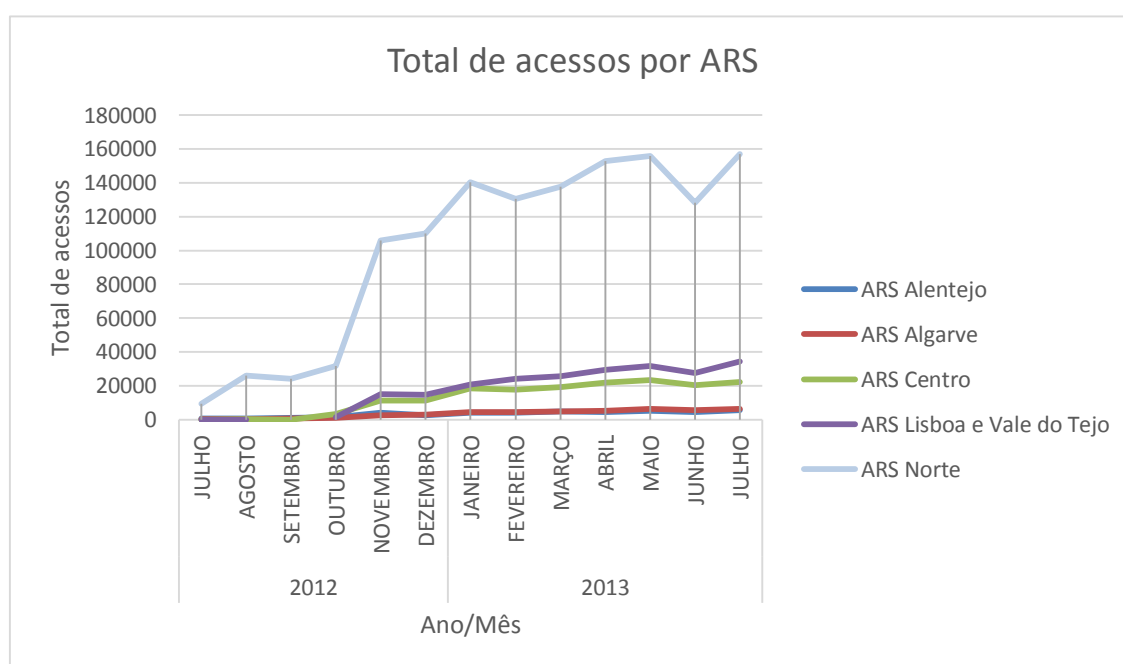


Figura 56 – PDS-PP – Total de acessos por ARS

Até Julho de 2013, já acederam 42.262 profissionais de saúde (médicos, enfermeiros) representando aproximadamente 40% do total de profissionais com licença profissional (40664 médicos² e 65467 enfermeiros³). Na Figura 57, podemos verificar que 82% dos acessos são realizados a partir de instituições prestadoras de Cuidados de Saúde Primários (CSP) e maioritariamente é acedido por médicos (72%).

² - Última estatística disponível na página da ordem dos médicos (referente ao ano de 2009)

³ - Última estatística disponível na página da ordem dos enfermeiros (referente ao ano de 2012)

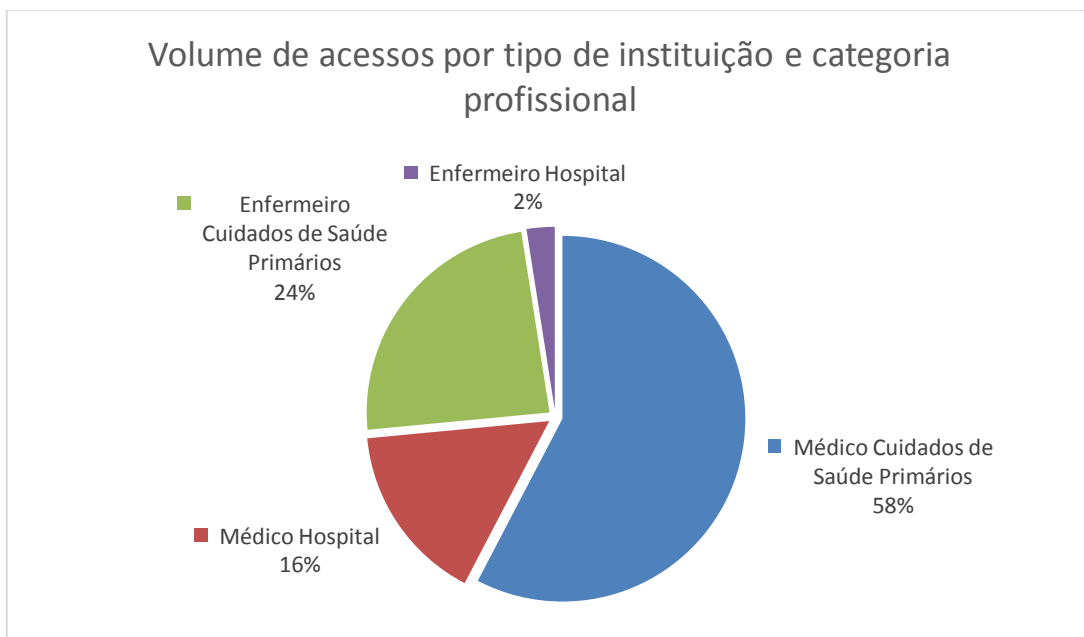


Figura 57 – PDS-PP – Volume de acesso por tipo de instituição e categoria profissional

Relativamente ao horário de consulta do PDS-PP, como representado na Figura 58, os profissionais utilizam o portal 24 horas por dia, no entanto é no horário comum de trabalho (das 8h até as 19h), em que outros serviços estão também em funcionamento, que é registada a maior afluência de acessos.

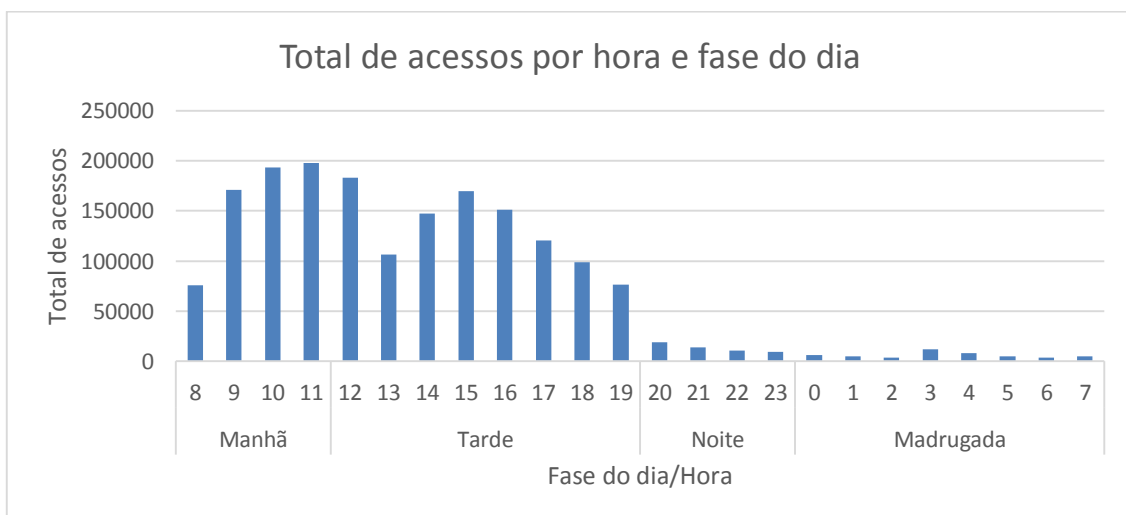


Figura 58 – PDS-PP – Total de acessos por hora e fase do dia

As instituições de saúde têm uma atividade centrada essencialmente nos utentes, portanto quando existe um contato entre o utente e profissional é gerado um registo do episódio. Na Figura 59, pode ser observado que a maioria dos acessos ao PDS-PP são realizados no âmbito de episódios de consulta (58%) ou tratamento de enfermagem (25%).

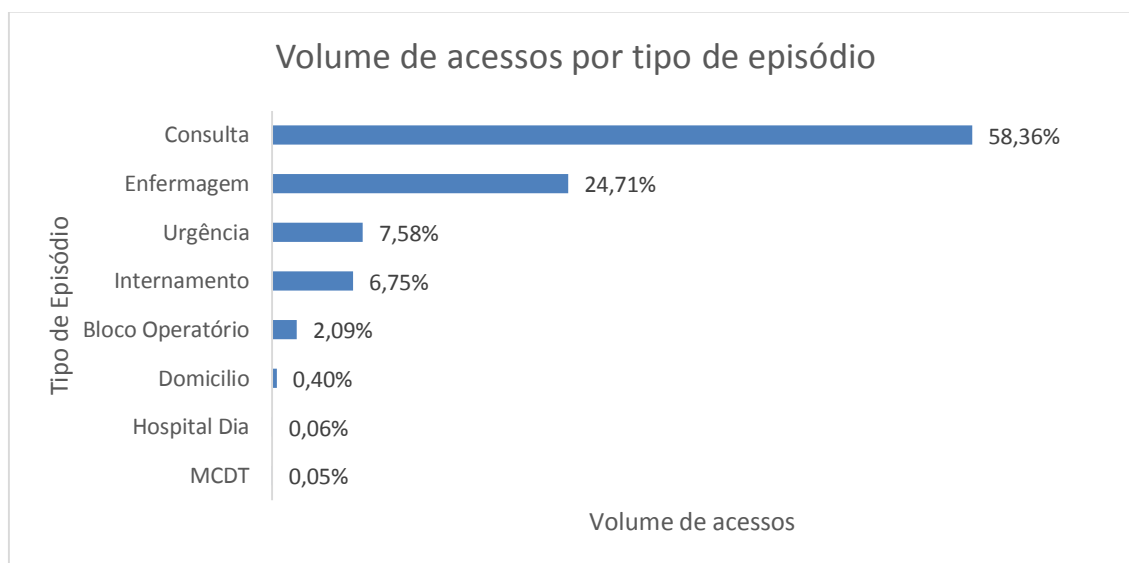


Figura 59 – PDS-PP – Total de acesso por tipo de episódio

O acesso ao PDS-PP é realizado a partir das aplicações de processo clínico eletrônico que o profissional tem acesso. As aplicações fornecidas pelo Ministério da Saúde (SAM – Sistema de Apoio ao Médico e SAPE – Sistema de Apoio à Prática de Enfermagem) são as mais utilizadas para aceder ao PDS-PP. Nos CSP o total de acessos a partir destas aplicações é de 99%, enquanto nos hospitais é de 79%. Na Figura 60 é possível verificar o conjunto de aplicações fornecidas por empresas externas que invocam o PDS-PP.

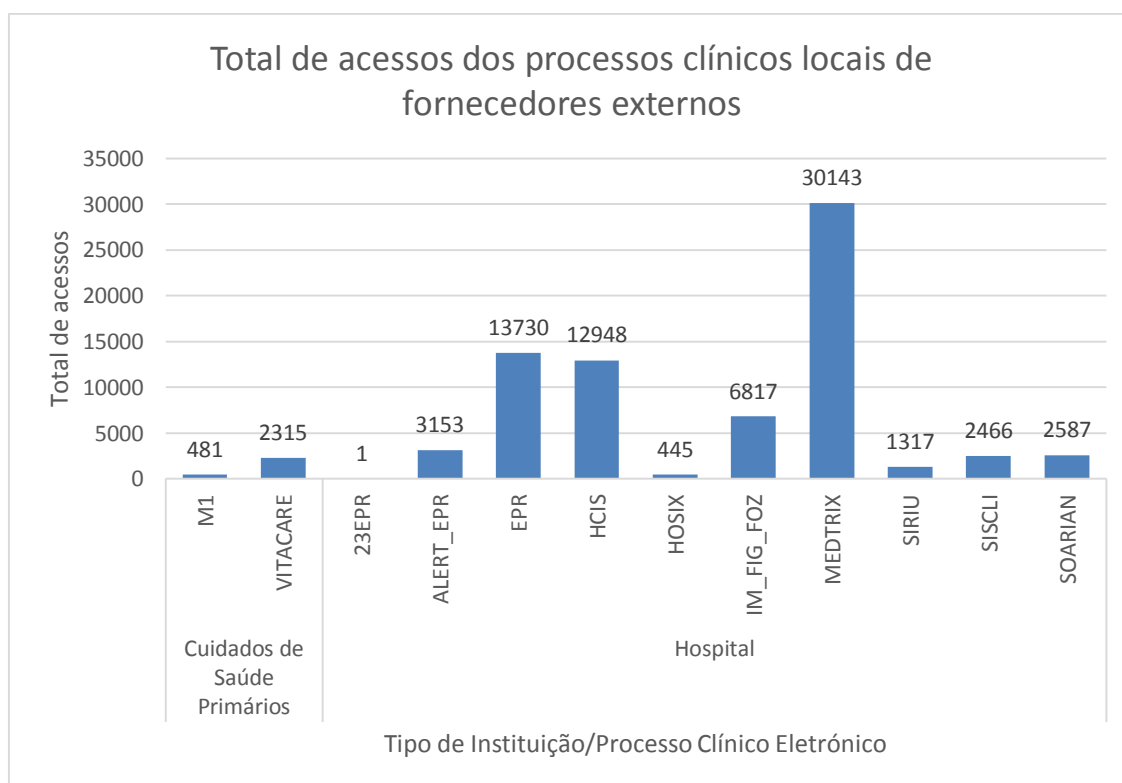


Figura 60 – PDS-PP – Total de acessos por aplicações externas

No âmbito de uma prestação de cuidados de saúde, até Julho de 2013 foi acedido ao processo clínico de 2.123.958 utentes. Na Figura 61 é possível verificar que maioritariamente os acessos são realizados a utentes do género feminino, com exceção da faixa etária dos 5 aos 15 anos.

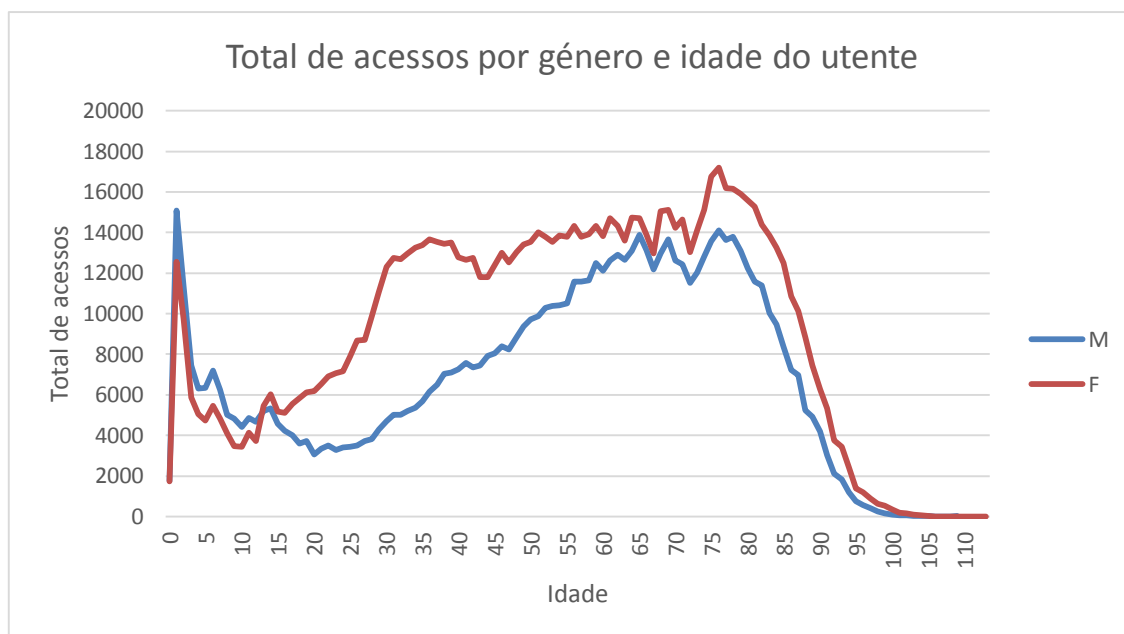


Figura 61 – PDS-PP – Total de acessos por género e idade do utente

De seguida serão analisados os resultados referentes às operações que os profissionais executam no PDS-PP. Os profissionais acedem ao PDS-PP com o principal objetivo de realizar duas operações: visualização do cronograma do utente (46%); consultar o processo clínico eletrónico (PCE) de uma instituição externa (49%); consultar a prescrição de medicamentos do utente (3%). Na Figura 62, são apresentadas os acessos às restantes as operações do PDS-PP.



Figura 62 – PDS-PP – Total de operações executadas (restantes)

A execução da operação de consulta do PCE externo, resulta da necessidade do profissional obter mais informações acerca do estado clínico, análises e avaliações realizadas noutras instituições que o utente tenha tido contato. A instituição mais requisitada para análise da situação clínica é o Centro Hospitalar do Porto (CHP) (9%) e a Unidade Local de Saúde Nordeste (8%). Este fluxo de consulta de informação entre as diversas instituições permite estabelecer uma rede de referência onde normalmente os utentes são acompanhados. Na Figura 63 pode ser observado a rede de referência das instituições com o CHP na região da ARS Norte.

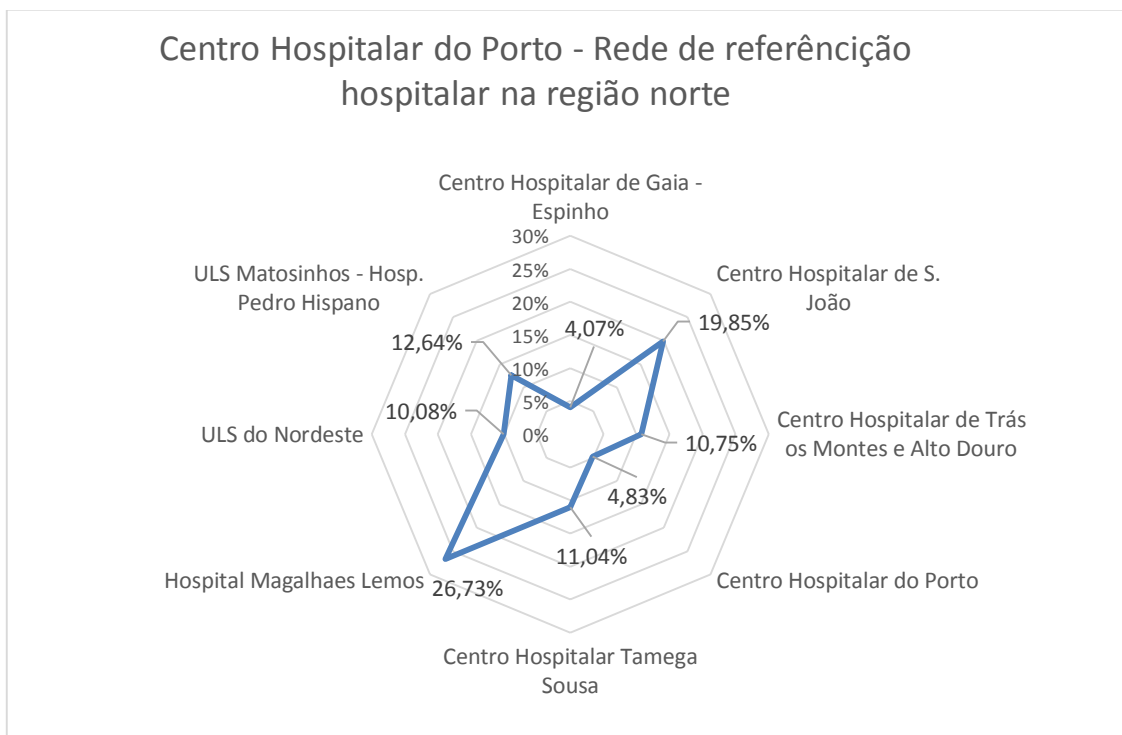


Figura 63 – PDS-PP – ARS Norte - rede de referência com o CHP

A rede do CHP, como é possível verificarmos na Figura 64, também pode abranger outras regiões de saúde do país (1% de acessos).

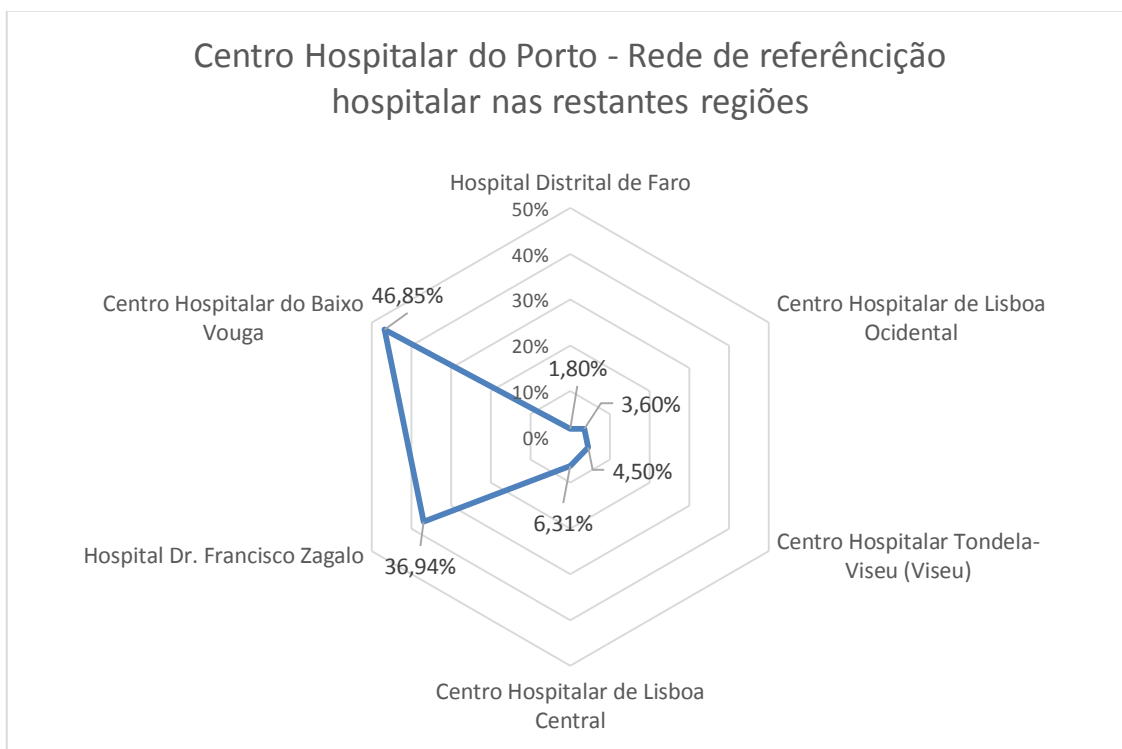


Figura 64 – PDS-PP – Restantes regiões – rede de referência com o CHP

Exploração de dados

Por fim, no âmbito da análise de dados para a Direção Geral de Saúde, é necessário obter o total de acessos às Normas de Orientação Clínica (NOC). Estas normas pretendem orientar o médico de modo a fornecer o tratamento mais adequado ao diagnóstico em questão. Na Figura 65 são apresentados as dez NOC mais consultadas relacionadas com o género do utente, e pode ser salientado que a Norma da abordagem e controlo da asma é a mais consultada (das dez mais acedidas tem um total de consulta de 53%).

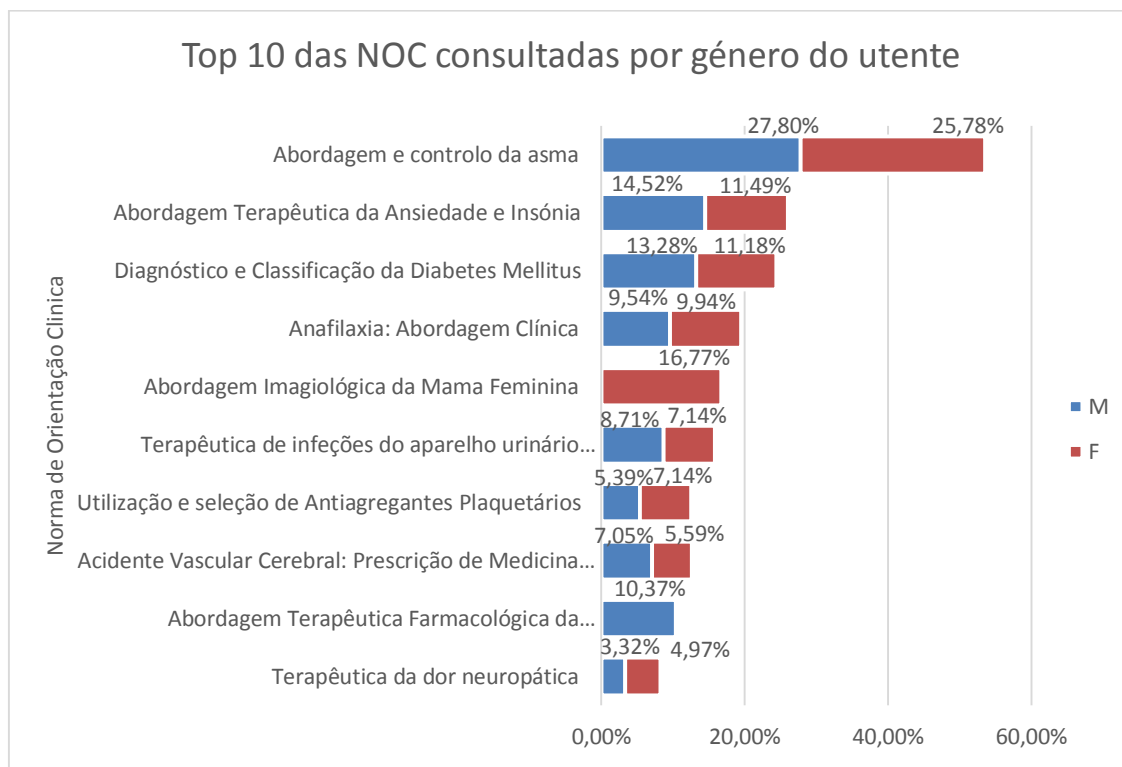


Figura 65 – PDS-PP – Top 10 Normas Orientação Clínica consultadas por género do utente

Finalizando a análise do PDS-PP, de seguida serão analisados os episódios que as instituições regularmente enviam para a PDS. Até Julho de 2013, foram registados 1.103.405 episódios provenientes de instituições hospitalares. Na Figura 66 é disponibilizado o total de registo de episódios por ARS.

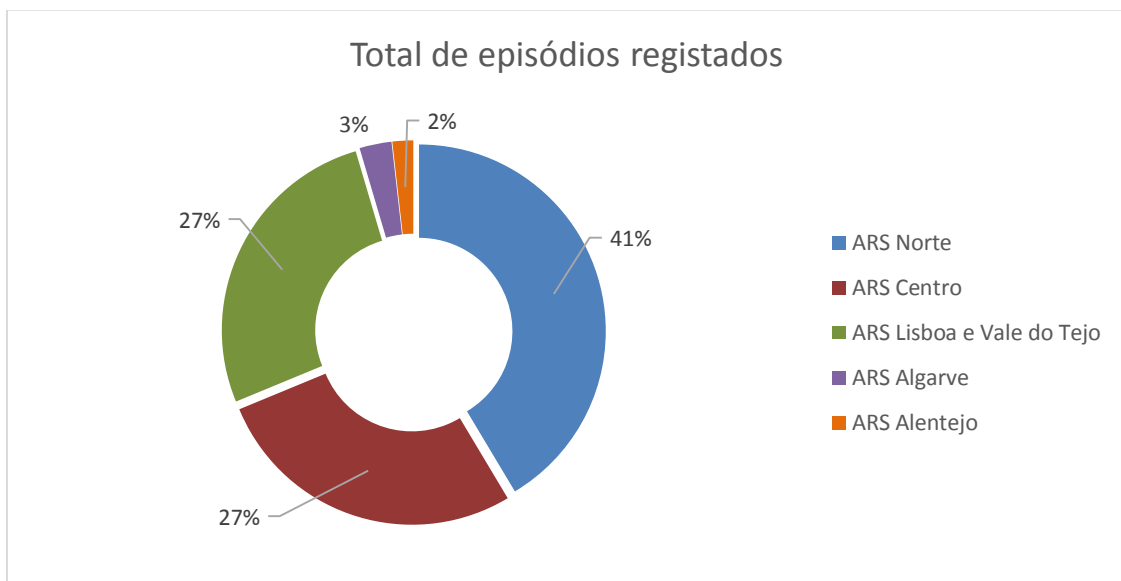


Figura 66 – PDS-PP – Total de episódios registados por ARS

De seguida, será avaliado o total de registos de episódio pelo tipo correspondente. Segundo o gráfico da Figura 67, o tipo de episódio mais registado é urgência (75%), no entanto é notório que o tipo de episódio que mais tempo consome é o internamento, em média 173 horas.

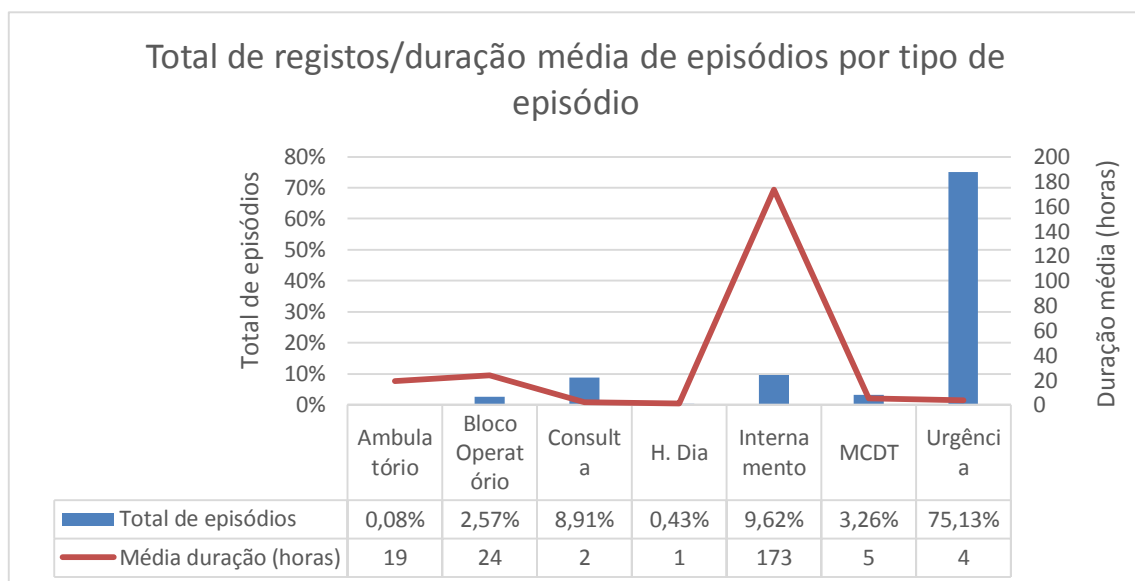


Figura 67 – PDS-PP – Total de registos/duração média de episódios por tipo de episódio

Tendo em conta a duração média observada para o internamento, seguidamente vai ser aprofundado este caso, para perceber de um ponto de vista da especialidade onde existem mais recursos consumidos (tempo). Conforme os dados expostos na Figura 68, existem especialidades a consumir em média mais de 50 dias, desde a entrada no utente até ao momento da alta.



Figura 68 – PDS-PP – Top 10 duração média (dias) dos episódios de internamento por especialidade

Por fim, é analisado o total de registos de episódios de urgência por perfil do utente. Maioritariamente a urgência é utilizada pelo utente feminino, à exceção da faixa etária dos 0 aos 15 anos.

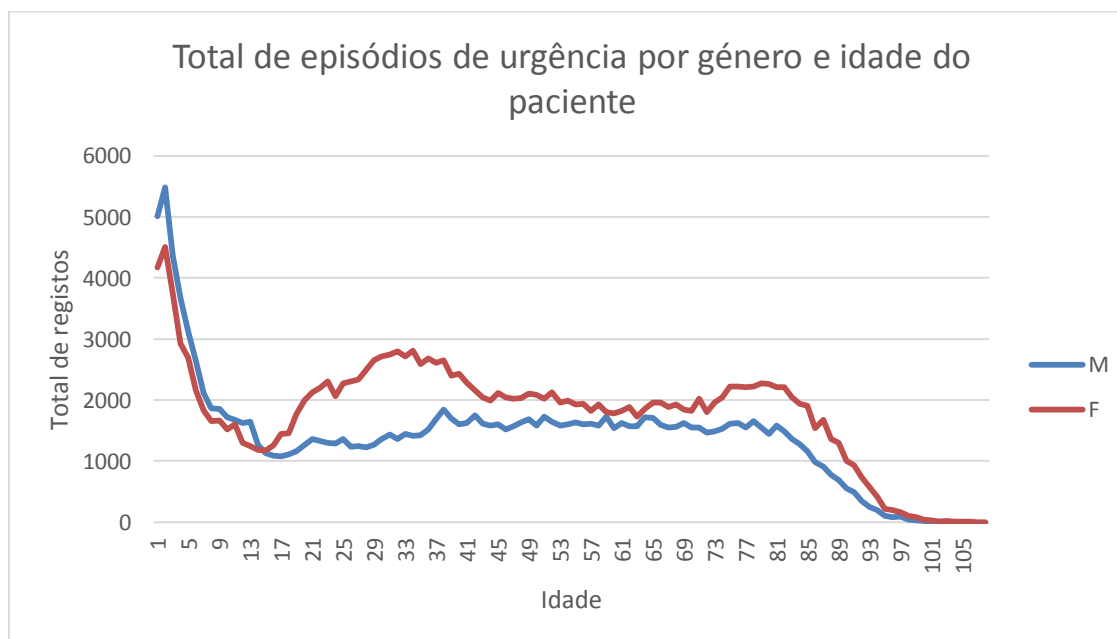


Figura 69 – PDS-PP – Total de episódios de urgência por género e idade do utente

5.2 PDS - Portal Utente

No PDS-Portal do Utente existem 789.550 utentes registados, desde o fim de 2008 até ao início do mês Setembro de 2013. Na Figura 70 é representada a evolução de utentes registados no portal.

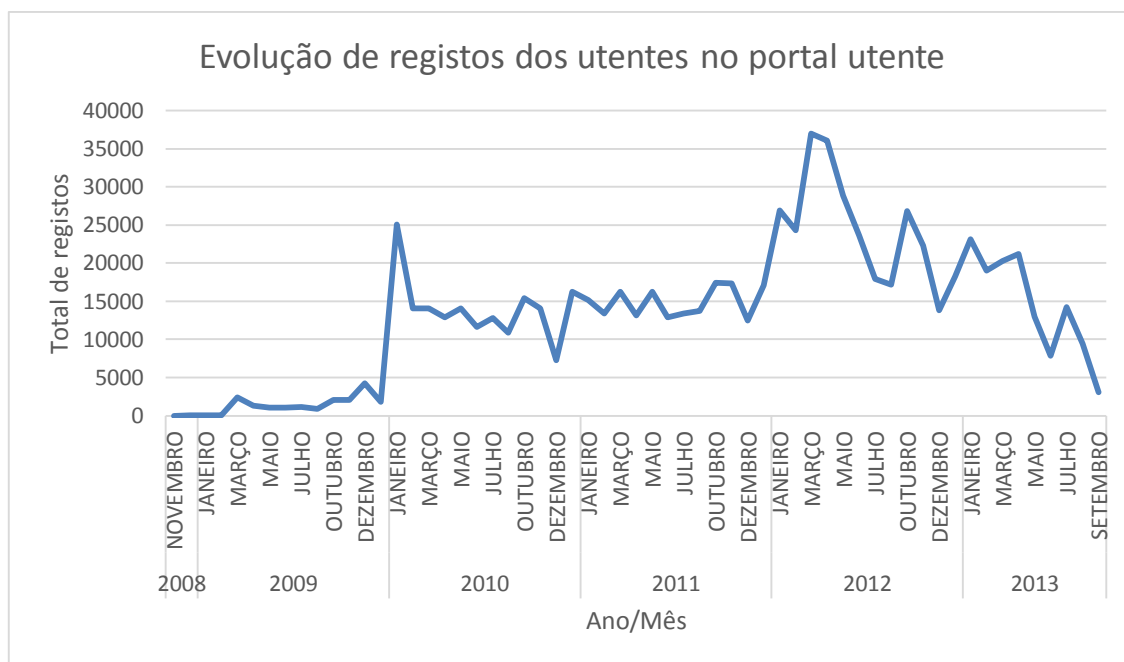


Figura 70 – PDS-PU – Evolução de utentes registados no portal

No que diz respeito à distribuição de utentes registados por distrito, segundo os dados da Figura 71, existe uma maior adesão dos utentes residentes nos grandes centros urbanos.

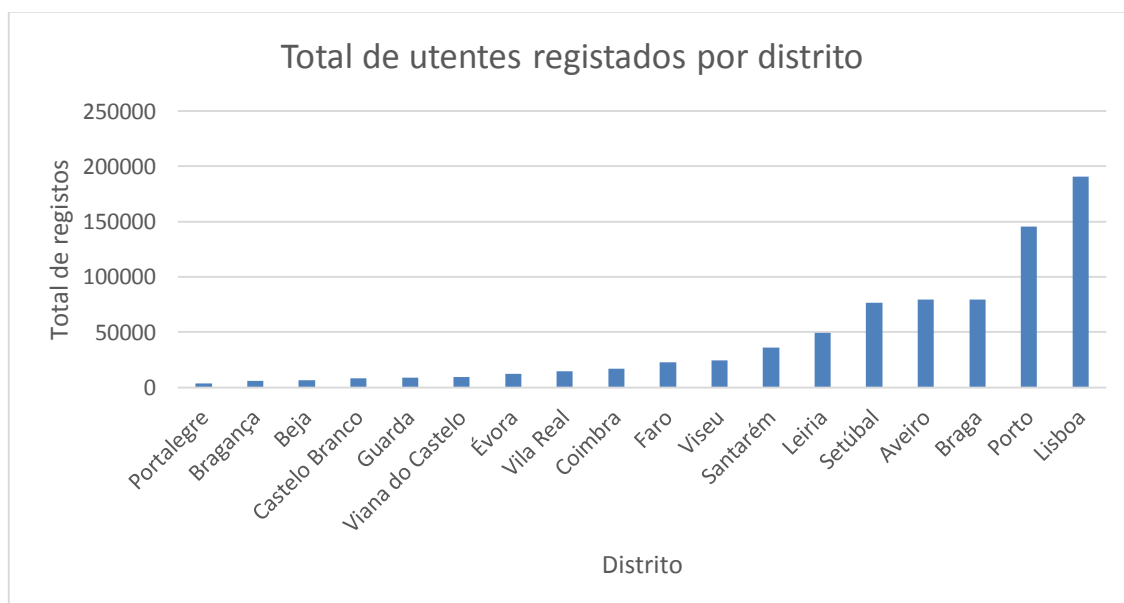


Figura 71 – PDS-PU – Total de utentes registados por distrito

Na Figura 72, é representado o perfil de utente que se regista no portal. É possível concluir que a maioria dos utilizadores tem entre 25 e 55 anos e são maioritariamente do género feminino.

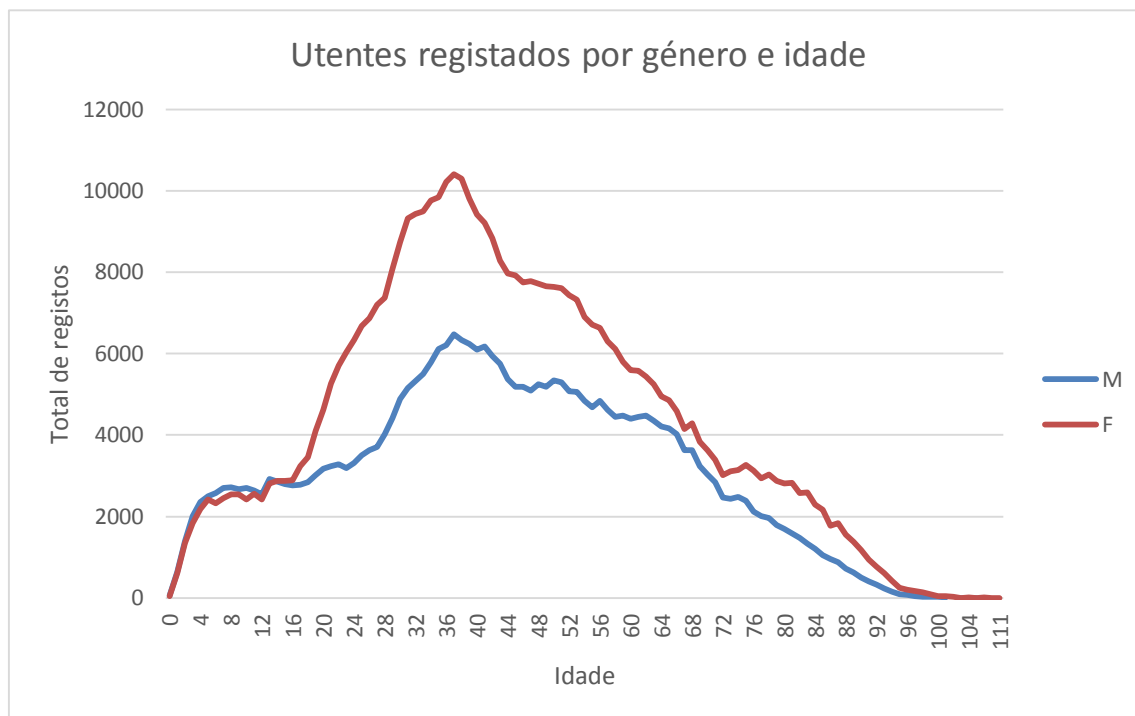


Figura 72 – PDS-PU – Total de utentes registados por género e idade

Por fim, foi constatado que até Julho de 2013 somente 2% dos utilizadores configuraram as autorizações de acesso, portanto é importante salientar o desconhecimento do utente na utilização desta área do PDS-PU e recomendar o desenvolvimento de futuras divulgações com o intuito de contrariar esta tendência. Esta área é de extrema importância pois retrata o consentimento que o utente atribui ao profissional para acesso ao respetivo perfil na PDS-PP e PDS-Portal Internacional (PDS-EPSOS), assim como a visualização dos dados registados no PHR, que capacitam o profissional de informação de saúde adicional que pode ser relevante para a prestação de cuidados de saúde.

Contudo, se o utente nunca configurou as autorizações, por defeito os profissionais podem aceder ao processo no PDS-PP (99%). Segundo a Tabela 20 podemos visualizar um resumo das autorizações configuradas.

Tabela 20 – Total de autorizações configuradas pelos utentes

Autorização	SIM	NÃO
Consulta do utente no PDS-PP	13102	628
Visualizar os registos do PHR no PDS-PP	12542	1188
Visualizar o resumo clínico no PDS-EPSOS	6589	7141

Em Maio de 2012, o PDS-PP teve a implementação de uma área pessoal para registo de dados de saúde (PHR-V1), no entanto em Maio de 2013 esta área foi reformulada de modo a disponibilizar um portal mais centrado no utente e não tanto nos serviços eletrónicos disponibilizados pelo Ministério da Saúde (PHR-V2).

De seguida serão analisados os registos de alergias inseridas no PHR. Desde da data de implementação do PHR-V1 até Setembro de 2013, foram inseridas 2.407 alergias, sendo que 75% dos registos foram inseridos a partir do momento que o PHR-V2 foi implementado. Na Figura 73, pode ser observado que 90% das alergias registadas se concentram nas reações a outra substância/agente e alergias medicamentosas.

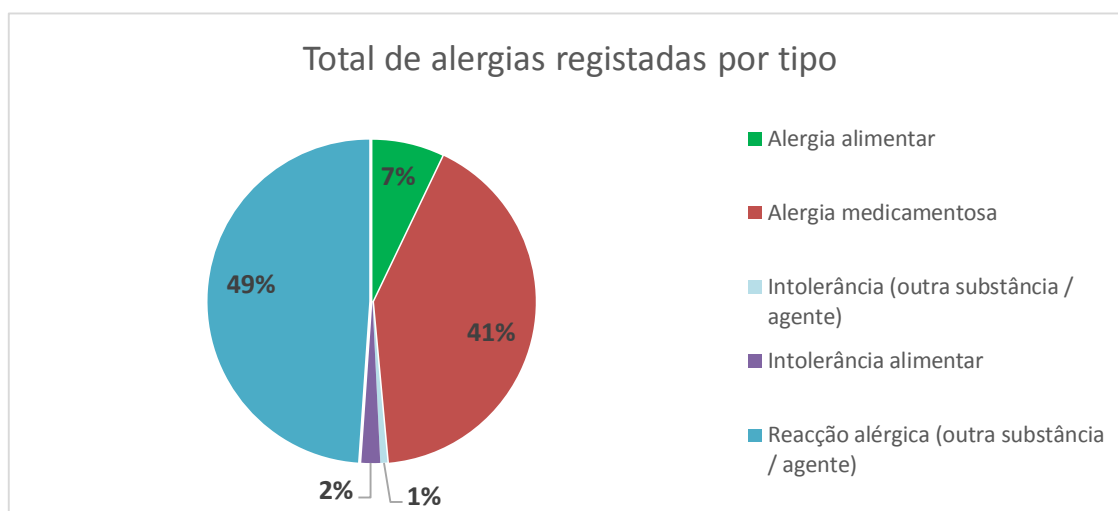


Figura 73 – PDS-PU – Total de alergias registadas por tipo

Relativamente às reações alérgicas a outras substâncias/agentes, com representado na Figura 44, 43% dos utentes declararam ter reação alérgica à maçã. Neste panorama, a severidade da reação é distribuída uniformemente entre os casos graves/moderados e ligeiros.

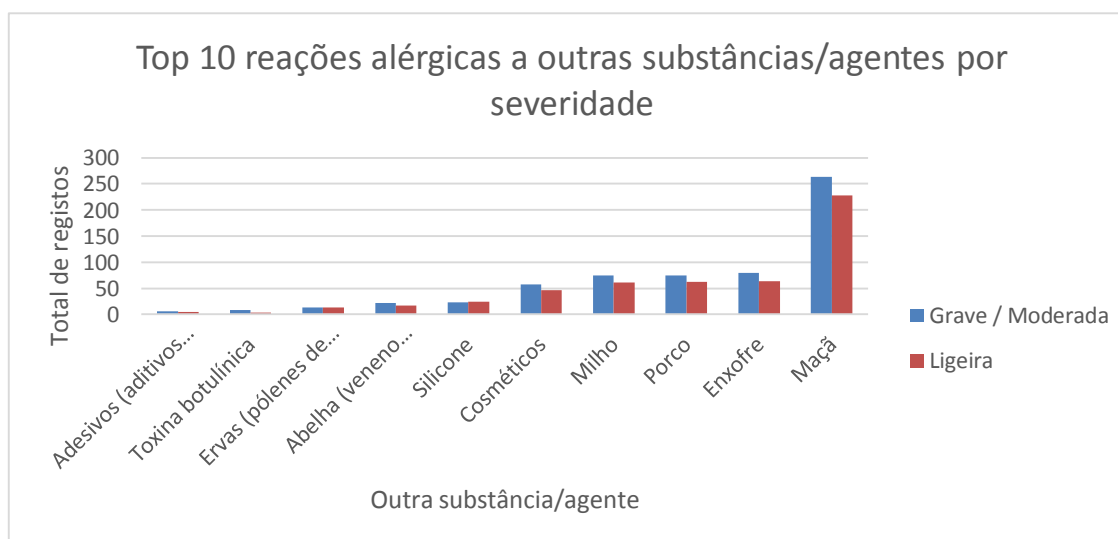


Figura 74 – PDS-PU – Top 10 reações alérgicas a outras substâncias/agentes por severidade

No que diz respeito às alergias medicamentosas, segundo o gráfico da Figura 75, existe uma disparidade quanto à severidade, tendo sido registados na maioria dos casos alergias com consequências graves ou moderadas. 64% das alergias registadas dizem respeito a antibióticos (como *Noprilam* ou *Clavamox ES*⁴).

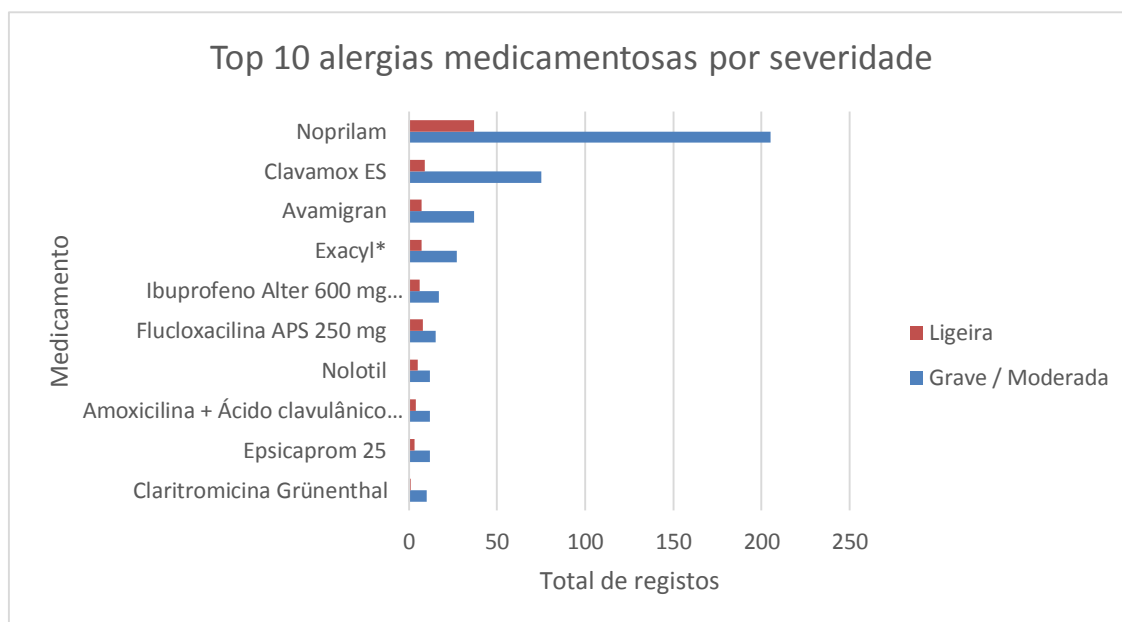


Figura 75 – PDS-PU – Top 10 alergias medicamentosas por severidade

Aprofundado a questão da disparidade de severidade alérgica, na Figura 76, é possível verificar que existe maior diferencial de casos graves e moderados na faixa etária dos 27 aos 52 anos.

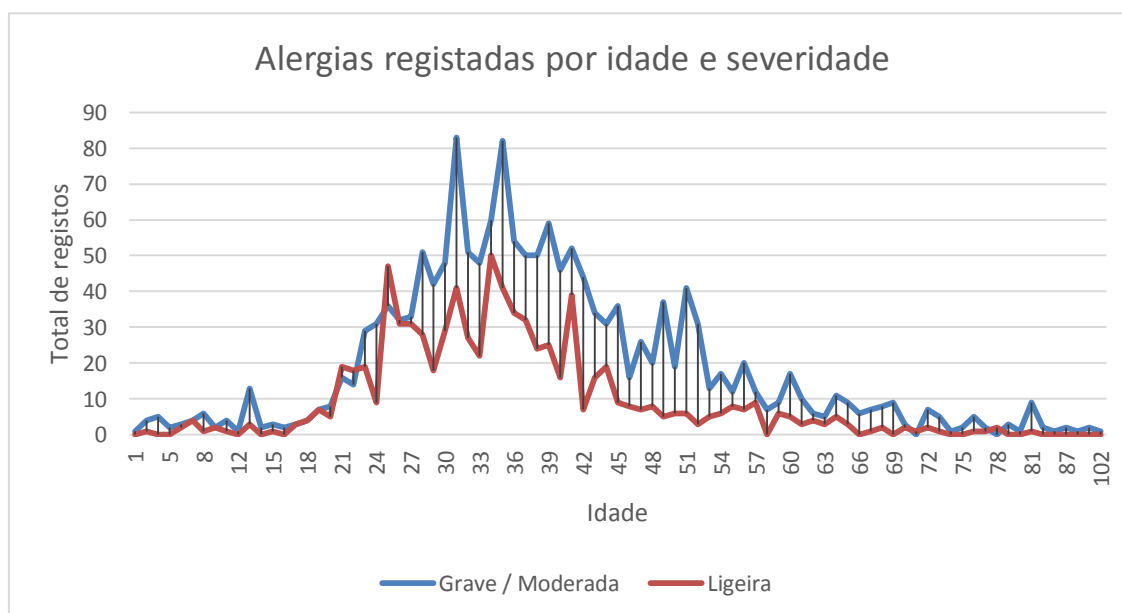


Figura 76 – PDS-PU – Alergias registadas por idade e severidade

⁴ Conhecimento adquirido através da pesquisa de medicamentos na página do Infarmed

Para finalizar a temática do registo de alergias, é importante avaliar o tipo de reação alérgica que a maioria dos utentes possui. Na Figura 77, é possível concluir que 46% das alergias registadas causa reações de rinite/conjuntivite ou dificuldade na respiração.

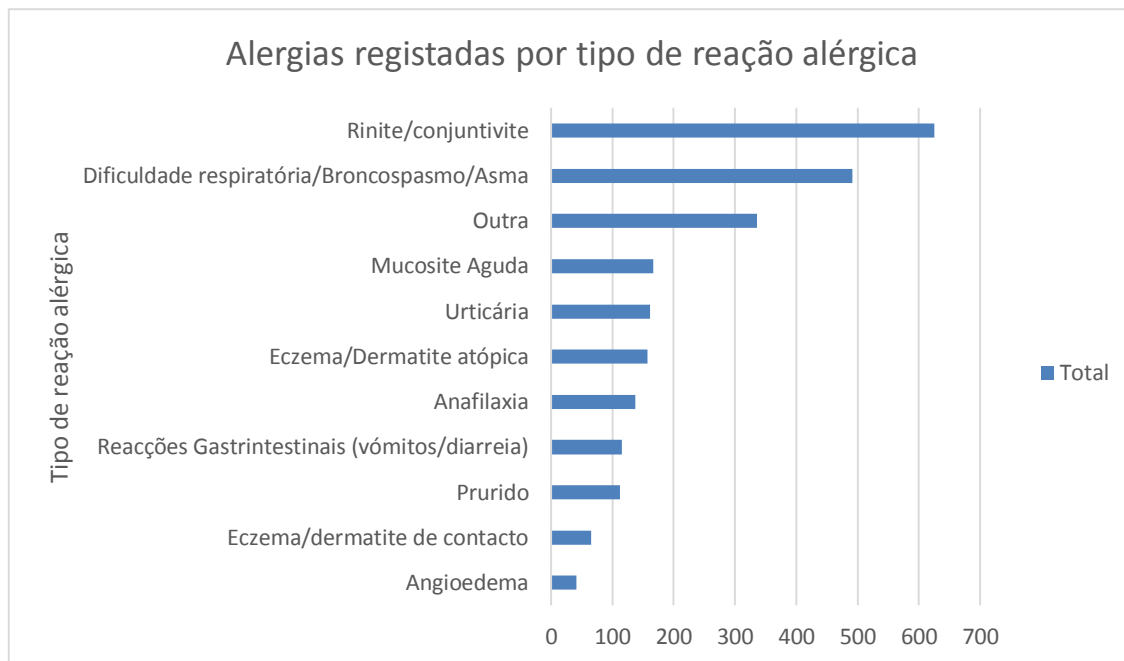


Figura 77 – PDS-PU – Alergias registadas por tipo de reação alérgica

Seguidamente serão analisados os registos de medicação registados pelos utentes no PHR. Desde da data de implementação do PHR-V1 até Setembro de 2013, foram inseridos 21.132 registos de medicação, sendo que 71% dos registos foram inseridos a partir do momento que o PHR-V2 foi implementado.

Após uma análise geral da medicação registada, observou-se um perfil de medicação distinto consoante o género do utente. Na Figura 78, pode ser verificado o top dez medicamentos mais registados cujo estado é ativo (13% do total de medicação registada por utentes femininos). E após uma avaliação deste subconjunto no catálogo do Infarmed, é possível concluir que 45% dos registos dizem respeito a anti contraceptivos e 30% a reguladores hormonais para problemas de hipertiroidismo.

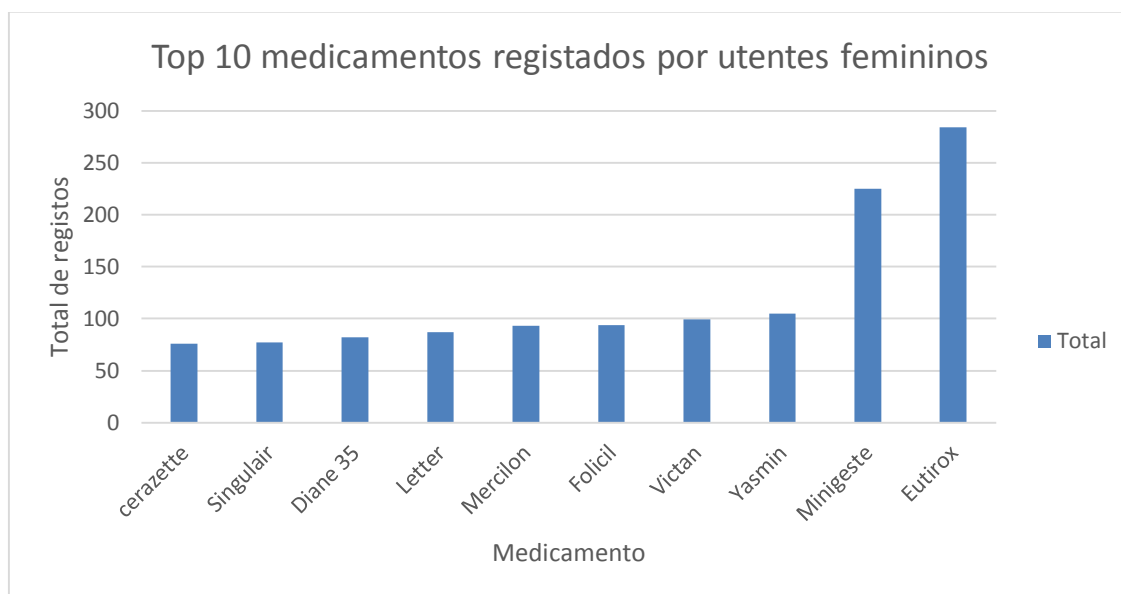


Figura 78 – PDS-PU – Top 10 medicação registada por utentes femininos

Relativamente aos utentes do género masculino, o perfil de medicação é completamente distinto. Na Figura 79, pode ser verificado o top dez medicamentos mais registados cujo estado é ativo (10% do total de medicação registada por utentes masculinos). Após uma avaliação deste subconjunto no catálogo do Infarmed, é possível concluir que 33% da medicação registada diz respeito a analgésicos e 27% para problemas de circulação do sangue.

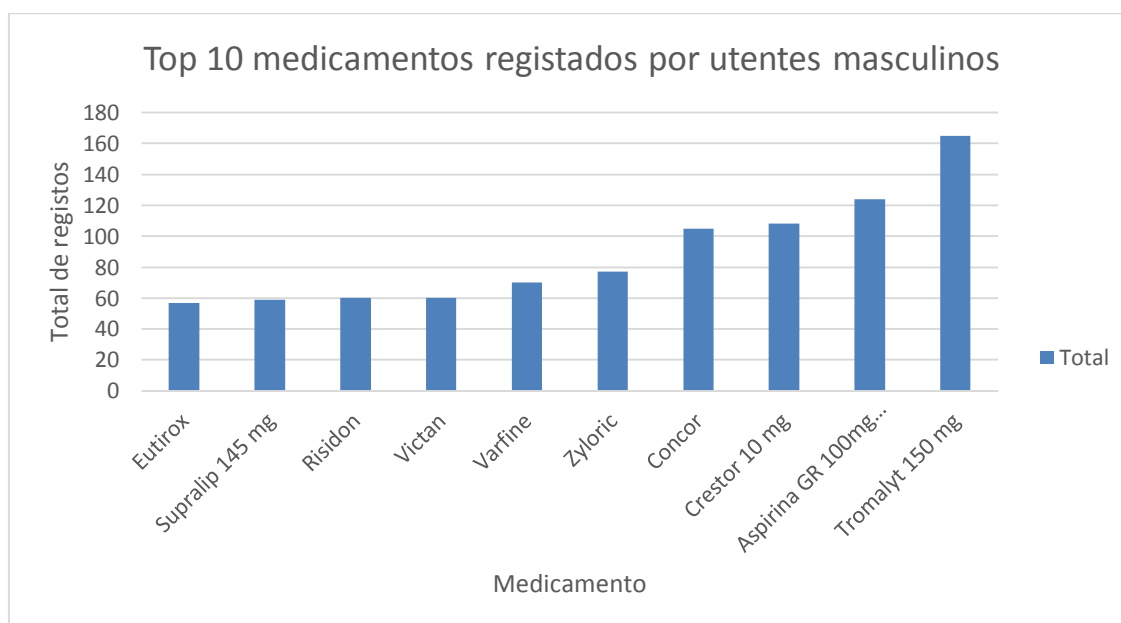


Figura 79 – PDS-PU – Top 10 medicação registada por utentes masculinos

As conclusões efetuadas anteriormente podem ser mais efetivas, se for adicionado o grupo do medicamento no modelo de dados do PDS-PU, sendo que este já contém a lista de medicamentos do Infarmed.

No que diz respeito ao perfil de toma da medicação, após análise dos dados concluiu-se que 88% da medicação registada é consumida diariamente e desta 76% dos utentes toma durante 1 ano ou mais. Este tipo de perfil está associado à medicação crónica, ou seja, medicação que é ingerida diariamente durante longos períodos de tempo (por vezes durante a vida toda). Na Figura 80 são representados os restantes 12% casos de toma de medicação, no qual 45% dos utentes toma até três dias e destes 82% em SOS.

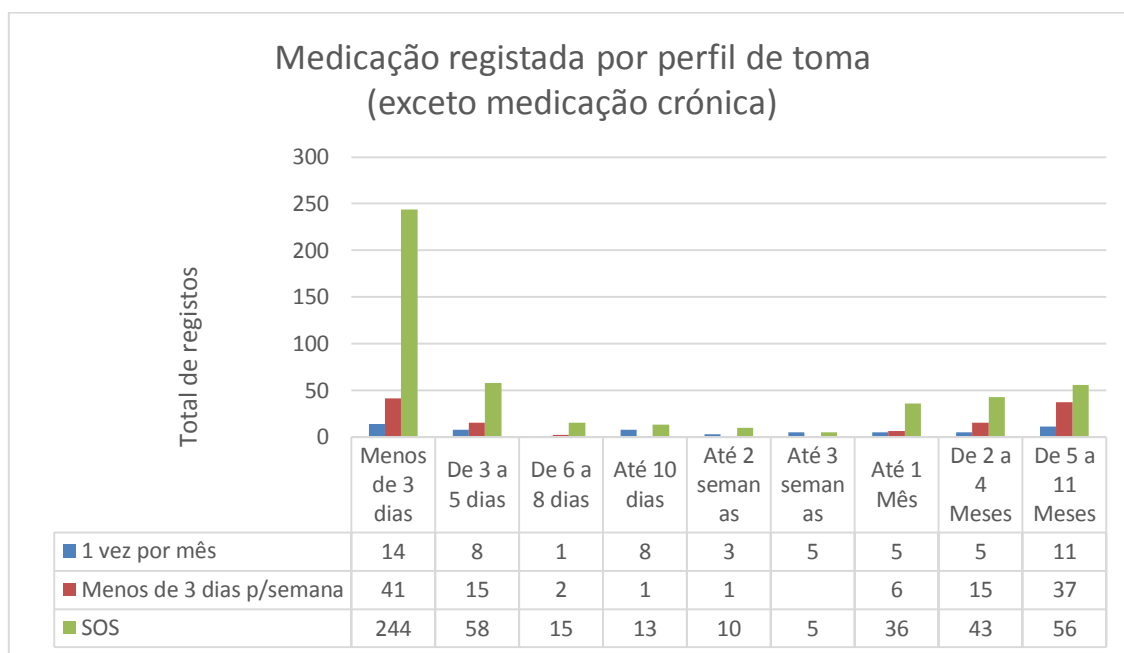


Figura 80 – PDS-PU – Medicação registada por perfil de toma (exceto medicação crónica)

Relativamente aos utentes que registaram medicação crónica, na Figura 81 é possível avaliar o perfil por idade e género do utente. Pode concluir-se então que o utente feminino cuja idade se situa na faixa etária dos 17 até aos 41 anos, possui a maioria dos registos de medicação crónica (influenciados pela toma diária de contraceptivos - pilula). A tendência inverte para o utente masculino cuja idade se situa na faixa etária dos 50 até aos 70 anos.

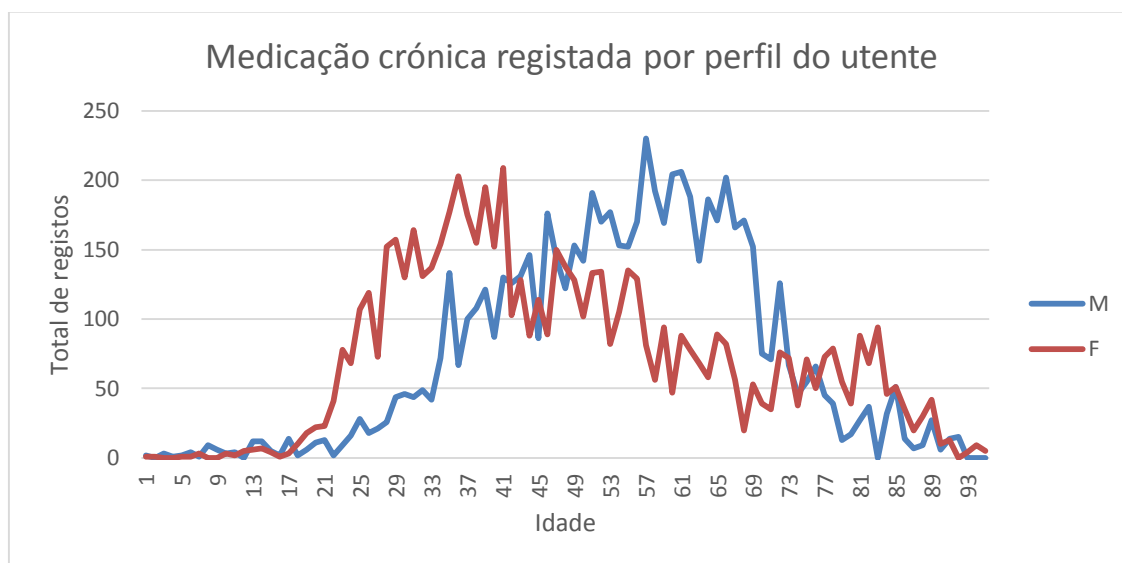


Figura 81 – PDS-PU – Medicação crónica registada por perfil do utente

Por fim, serão analisados os dados das doenças registadas pelos utentes no PHR. Desde a data de implementação do PHR-V1 até Setembro de 2013, foram inseridos 6.896 registos de doenças, sendo que 73% dos registos foram inseridos a partir do momento que o PHR-V2 foi implementado. Na Figura 82 são representadas as dez doenças mais registadas pelos utentes (52% do total de doenças registadas), das quais se destacam a Hipertensão arterial com 10%, Diabetes com 8% e Asma com 7%. Após avaliação dos registos com a opção “Outra doença” (14%), pode concluir-se que na maioria 3% declarou psoríase e 2% depressão.



Figura 82 – PDS-PU – Top 10 doenças mais registadas

Nesta ultima análise, após a exploração mais detalhadas das doenças declaradas como “outra doença” foram detetados inúmeros casos de doenças mal relacionadas. Supostamente os utentes deveriam selecionar a opção “outra doença” para registarem outros casos, no

entanto nem sempre tal acontece. A lista de outras doenças poderá ser reutilizada para avaliar casos semelhantes, assim como recorrendo a um profissional clínico, corrigir os dados existentes e adicionar novas doenças no catálogo disponibilizado no PDS-PU.

Seguidamente serão avaliados os casos corretamente tipificados. Na Figura 83, são analisados os dez tipos de doenças mais registados por género do utente (73% dos tipos de doenças registados). Portanto pode concluir-se que existe inúmeros casos de cancro registados em ambos os géneros, e os utentes masculinos têm mais doenças do aparelho circulatório e crónicas, enquanto que os utentes femininos têm mais doenças respiratórias e do sistema endócrino.

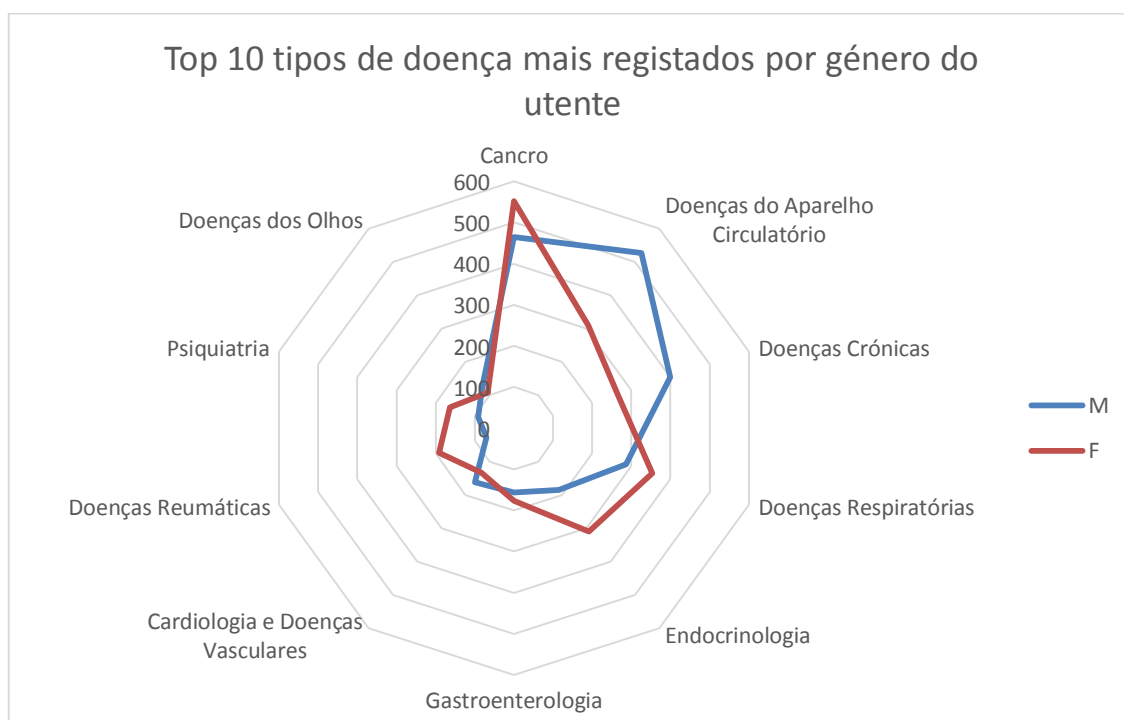


Figura 83 – PDS-PU – Top 10 tipos de doença mais registados por género do utente

Pormenorizando o grupo de doenças do aparelho circulatório (12% do total de doenças registadas), segundo o gráfico da Figura 84 podemos concluir que 75% das doenças desta categoria se devem ao problema de hipertensão arterial.

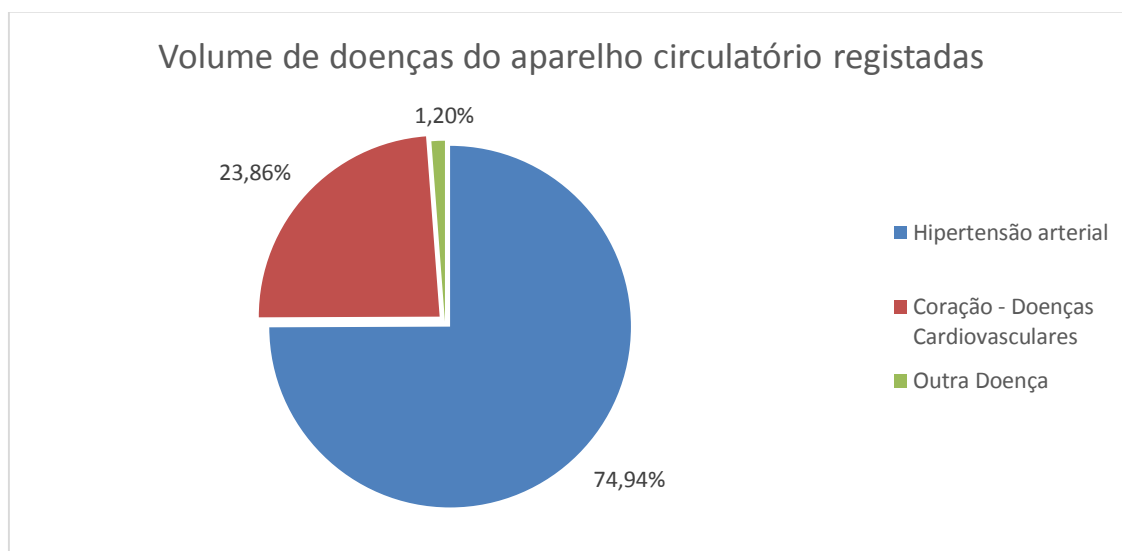


Figura 84 – PDS-PU – Volume de doenças do aparelho circulatório registadas

Para finalizar é importante avaliar os casos de cancro declarados (15% do total de doenças registadas). Segundo o gráfico da Figura 85, existe um equilíbrio entre os casos de cancro ativos e inativos, com exceção da faixa etária do 25 aos 41 anos em que a diferença entre os casos doença ativa é mais acentuado.

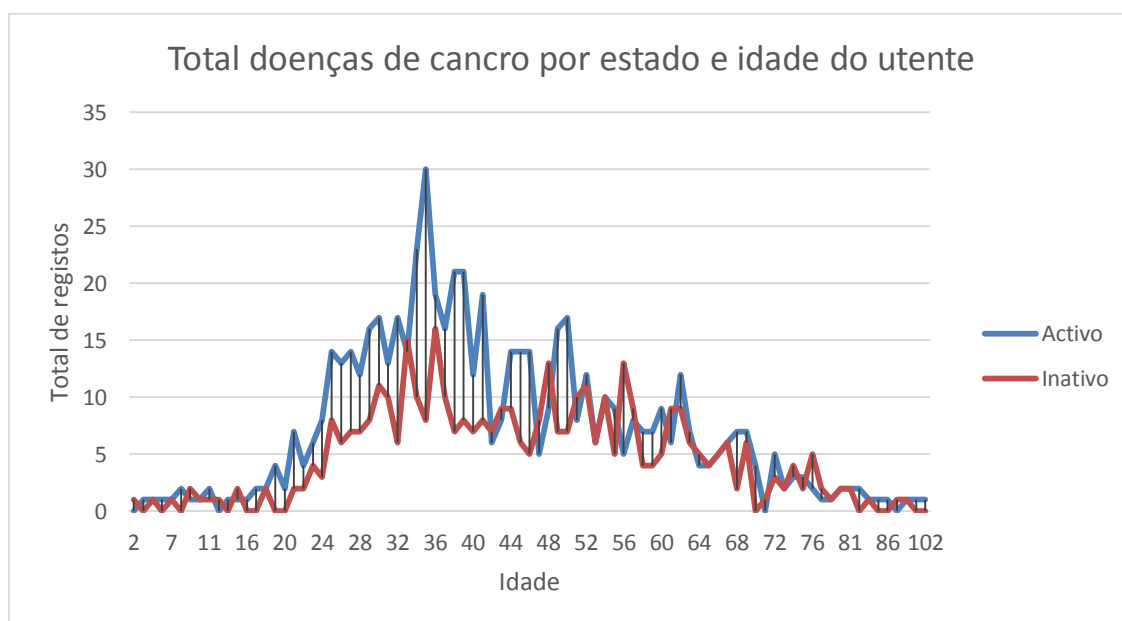


Figura 85 – PDS-PU – Total de doenças de cancro por estado e idade do utente

Em suma, a maioria dos registos de saúde dos portugueses foram inseridos a partir do mês de Maio de 2013, e devido ao número reduzido de registos, deverão ser realizadas campanhas de informação de modo a esclarecer o cidadão das vantagens do registo e partilha dos dados de saúde. A análise e investigação aprofundada desta informação pode traduzir em avanços consideráveis na forma como se pratica a saúde em Portugal.

6 Conclusões

Efetuando um balanço geral do projeto desenvolvido, de um modo geral foram atingidas as metas principais. Foi possível efetuar uma investigação aprofundada sobre diversos sistemas de saúde implementados tanto em Portugal como no exterior, e aprofundar dois exemplos de sucesso que estrategicamente representam potenciais plataformas a integrar com o PDS-Portal Institucional (PDS-PI). Por exemplo, tendo na PDS o registo de formulários de cirurgia segura salva vidas (procedimento para verificação dos parâmetros de segurança numa cirurgia) e um repositório de episódios, é possível corresponder esses mesmos episódios com os sistemas SIGLIC e GDH para obter dados clínicos e administrativos, permitindo deste modo disponibilizar novos indicadores aos diversos *stakeholders*.

Tendo este projeto a duração de sete meses, devido à alteração do projeto de dissertação, foi possível desenvolver um protótipo de PDS-Portal Institucional, tendo em conta objetivos bem definidos e uma visão estritamente delimitada sobre o domínio em questão.

Após o período inicial de investigação, foi necessário realizar uma análise aprofundada dos principais sistemas envolvidos: PDS-Portal Profissional e PDS-Portal Utente; de modo a definir os limites de informação que o protótipo do PDS-PI deveria disponibilizar. Tendo em conta, a implementação da nova versão do portal, foi tido em conta a informação essencial a ser disponibilizada aos diversos *stakeholders*. Portanto, é certo que o domínio na sua totalidade contém novos conceitos a explorar, e que deverão ser definidos novos passos permitindo uma evolução contínua do projeto.

Concluindo o ciclo de implementação dos portais que constituem a PDS, Portugal torna-se um exemplo mundial de inovação e avanço tecnológico nos sistemas de informação para a saúde. Atualmente não existe nenhum país com um sistema a nível nacional que abranja a mesma percentagem da população como o nosso país. Por exemplo, em Espanha existe um esforço

substancial em colaborar no projeto EPSOS, no entanto devido à visão da saúde ser distinta nas diversas regiões autónomas, a possibilidade de abranger este projeto a um nível nacional é praticamente impossível. Nos EUA, é um dos casos onde esta possibilidade é ainda mais remota, devido ao facto do sistema de saúde ser na sua maioria dominado pelo setor privado e por fornecedores completamente distintos e com estratégias desconcertadas. Contudo, existem esforços positivos de países como a Noruega, que vendo em Portugal um exemplo, iniciam o desenvolvimento de uma plataforma nacional recorrendo à colaboração e experiência do nosso país.

No âmbito do projeto PDS-PI, foi possível concluir a implementação do armazém de dados, que através de um mecanismo automático de extração, tratamento e carregamento de dados, disponibiliza a informação necessária para a aplicação de apresentação. Recorrendo ao *Microstrategy Web*, é disponibilizado um ambiente seguro e com capacidade para potenciar o espírito de análise dos utilizadores do PDS-PI. Por conseguinte, foi possível configurar diversos grupos, utilizadores e relatórios, sendo que existem diversas melhorias que poderão ser implementadas: personalizar o aspeto do portal (recorrendo ao *Microstrategy SDK*); desenvolver e melhorar as análises para o utilizador; potenciar a plataforma e BI recorrendo a diversos objetos disponibilizados (*Intelligent Cubes*, filtros, pesquisas personalizadas, produção de *templates*, entre outros).

6.1 Limitações e contratempos

Durante o período de desenvolvimento do PDS-PI existiram um conjunto de contratempos e limitações que atrasaram/estagnaram em determinadas fases a execução do projeto.

O desenvolvimento em ambiente virtualizado, recorrendo a um portátil pessoal como máquina física, constituiu uma enorme barreira. Existiram atrasos consideráveis no processo de ETL, instalação da plataforma de BI e exploração de dados, devido ao facto dos recursos da máquina física serem bastante limitados, principalmente em memória disponível e capacidade de processamento.

Durante a fase de investigação, existiu um acesso dificultado à informação/documentação técnica dos projetos implementados na área da saúde em Portugal. Este contratempo foi desbloqueado através de diversos contatos estabelecidos na empresa.

Para a plataforma de BI, foi selecionada uma tecnologia comercial no qual a empresa possui licenciamento. No entanto, foram necessários aproximadamente quatro meses para disponibilizar a licença e o ficheiro de instalação do *Microstrategy*, devido à decisão de aguardar o lançamento e licenciamento da versão 9.3.1, que disponibiliza um conjunto de melhorias e novas funcionalidades.

A instalação e configuração da plataforma de BI, constituiu um desafio que trouxe alguns atrasos devido ao desconhecimento da implementação deste tipo de plataforma. Foi

necessário uma análise aprofundada e cuidada da documentação disponibilizada, de modo a que o processo fosse executado com sucesso.

No entanto, face a este conjunto de limitações e contratemplos, devido à postura de persistência e determinação, foi sendo possível solucionar e avançar com o projeto.

6.2 Perspetivar o futuro

Perspetivar o futuro deste projeto é uma tarefa de imensa responsabilidade e orgulho.

Antes de implementar este projeto num ambiente de produção, deverá ser realizado um trabalho mais minucioso de tratamento e reporte de erros, principalmente no processo de ETL. Numa fase seguinte, deverá ser alargado o domínio de extração de dados a outras áreas (PDS-PP: cirurgia segura salva vidas; boletim de saúde infantil e juvenil; cartão de pessoa com doença rara; PDS-PU: registo de medições e hábitos de saúde) assim como desenvolver paralelamente uma área de *backoffice* para gestão da seguinte informação:

- Configuração de indicadores de performance das cirurgias dos hospitais (na área da cirurgia segura);
- Emissão/Impressão do cartão de pessoa com doença rara (funcionalidade somente disponível para a DGS);
- Configuração do Plano Nacional de Vacinação e definição das regras associadas (PNV);
- Gestão centralizada de stocks de vacinas nas unidades de saúde.

Por fim, é importante contextualizar um projeto na área do *big data*, que envolve uma enorme quantidade de dados clínicos complexos. O Repositório de Informação Clínica Anonimizada (RICA), é um projeto que pretende agregar os dados clínicos dos portugueses numa base de dados central e anonimizada com o objetivo de apoiar iniciativas de investigação e estabelecer protocolos com o meio universitário, potenciando a descoberta de novo conhecimento acerca da saúde em Portugal. Este projeto estará integrado com o PDS-PI, na medida em que será necessário relacionar os dados dos episódios realizados de modo a estabelecer, por exemplo, uma referência à localização geográfica de intervenção e reporte de casos.

Conclusões

Referências

- [David Hatch and Michael Lock, 2008] David Hatch and Michael Lock - Business Intelligence in Healthcare, June 2008
- [ACSS-SIGIC, 2010] ACSS - Manual de Gestão de Inscritos para Cirurgia - Volume I – Princípios Gerais
- [ACSS-SIGIC, 2011a] ACSS - Manual de Gestão de Inscritos para Cirurgia - Volume V – Organização da Informação e Sistemas de Informação
- [ACSS-SIGIC, 2011b] ACSS - Relatório da Atividade em Cirurgia Programada - Ano 2011, Cap 2
- [Claudia Imhoff, 2003] “Mastering Data Warehouse Design: Relational and Dimensional Techniques”, Claudia Imhoff, Nicholas Gallempo, Jonathan G. Geger Wiley, 2003, Cap 1 e 13.
- [CMS, 1983] CMS - Definição do Centro de Serviços Médicos, dos EUA - grupos relacionados com o diagnóstico- pode ser consultado em <http://www.cms.gov/>
- [David Stodder, 2012] David Stodder - TDWI Research - Actionable Analytics for Healthcare Providers
- [Diogo Reis, 2013] Diogo Reis - Plataforma de Dados de Saúde, “A partilha de informação de saúde já é uma realidade”, eSaúde - Magazine dos Sistemas de Informação na Saúde, Jan-Mar 2013
- [Forgione DA et al., 2004] Forgione DA, Vermeer TE, Surysekar K, Wrieden JA, Plante CA - The Impact of DRG-Based Systems on Quality of Health Care in OECD Countries, Journal of Health Care Finance, 31(1), p. 41-54
- [Helder Quintela, 2013] Helder Quintela - *Healthcare Business Intelligence*, eSaúde Magazine dos Sistemas de Informação na Saúde, 2ª ed., p. 27-30
- [Henrique Martins, 2013] Henrique Martins - Comissão para a Informatização Clínica, “Seis questões a... Henrique Martins”, eSaúde - Magazine dos Sistemas de Informação na Saúde, 2ª ed., Jan- Mar 2013
- [IGIF, 2005] IGIF - Manual AUDITOR – Auditoria às Base de dados dos GDHs, v 3.18, Nov 2005
- [Kroneman M and Nagy J, 2001] Kroneman M and Nagy J - Introducing DRG-based Financing in Hungary: A study into the relationship between supply of hospital beds and the use of these beds under changing institutional circumstances, Health Policy, 55, 19-36.
- [Luis Campos, 2011] Luís Campos - Sistemas de Informação na Saúde – O registo de saúde eletrónico em Portugal, Cap. 13

Conclusões

- [Manuel Barrento et al., 2010] Manuel Barrento, Maria Martins, Miguel Neto, Sara Dias - Business Intelligence applied to Homogeneous Diagnostic Groups
- [Manuel Pizarro, 2011] Manuel Pizarro - Sistemas de Informação na Saúde - Perspetivas e desafios em Portugal, prefácio.
- [Marianne Kolbasuk, 2012] Marianne Kolbasuk - 11 BI Tools To Analyze Healthcare Operations, Information Week Healthcare, 25 May 2012
- [Mark Madsen, 2010] Mark Madsen - BeyeNETWORK RESEARCH REPORT – “Open Source Solutions: Managing, Analyzing and Delivering Business Information”, March 2010
- [Microstrategy, 2012] “Microstrategy 9 Project Design Guide”, Microstrategy, 2012, 14ª Edição, versão 9.3
- [Nelson Pinho et al.,2013] Nelson Pinho, Lia Patrício, Raymond Fisk – “Designing the first version of the PDS as a service for the different stakeholders”, FEUP Final presentation, July 12th 2013
- [Paul C. Tang, 2006] Paul C. Tang - “Personal Health Records: Definitions, Benefits, and Strategies for Overcoming Barriers to Adoption”, J Am Med Inform Assoc., Mar-Apr 2006
- [Pedro Gomes, 2011] Pedro Gomes, Sistemas de Informação na Saúde - Sistemas e repositórios de âmbito nacional - o caso SIGIC, Cap. 15
- [Pentaho, 2013] Pentaho - “Pentaho Data Integration (aka Kettle) documentation” V.80, Pentaho, última alteração por Jens Bleuel a 27 de Agosto 2013
- [Ralph Kimball, 2002] Ralph Kimball - “The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling”, Ralph Kimball, Margy Ross, Second Edition Wiley, 2002, Pag 30-32
- [Ralph Kimball, 2004] Ralph Kimball - “The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming , and Delivering Data”, Ralph Kimball, Joe Caserta, Second Edition Wiley, 2004, Pag 30-32
- [Ralph Kimball, 2008] Ralph Kimball - “The Data Warehouse Lifecycle Toolkit”, Ralph Kimball, Margy Ross, Warren Thornthwaite, Joy Mundy, Bob Becker, 2ª Edição, Janeiro 2008, Cap 8 a 10
- [SPMS, 2013] SPMS - “PDS - Plataforma de Dados da Saúde - Ponto de situação”, SPMS, Julho 2013
- [W. H. Inmon, 2001] W. H. Inmon - “Corporate Information Factory”, W. H. Inmon, Claudia Imhoff, Ryan Sousa, Janeiro 2001

ANEXOS

Anexo 1 – Exemplo de procura de dados numa tabela dimensão

Step name: DimInstitution lookup

Connection: PDSPI_DW

Lookup schema:

Lookup table: DimInstitution

Enable cache?

Cache size in rows (0=cache everything): 0

Load all data from table

The key(s) to look up the value(s):

#	Table field	Comparator	Field1	Field2
1	InstitutionID	=	IDINSTITUTION	
2				

Values to return from the lookup table :

#	Field	New name	Default	Type
1	InstitutionID	InstitutionID_OLD		Integer
2	ARS	ARS_OLD		String
3	ACES	ACES_OLD		String
4	NAME	NAME_OLD		String
5	TYPE	TYPE_OLD		String

Do not pass the row if the lookup fails

Fail on multiple results?

Order by:

Buttons: OK, Cancel, Get Fields, Get lookup fields

Figura 86 – ETL – Exemplo de procedimento de procura de dados numa tabela

Anexo 2 – Exemplos de transformação de dados

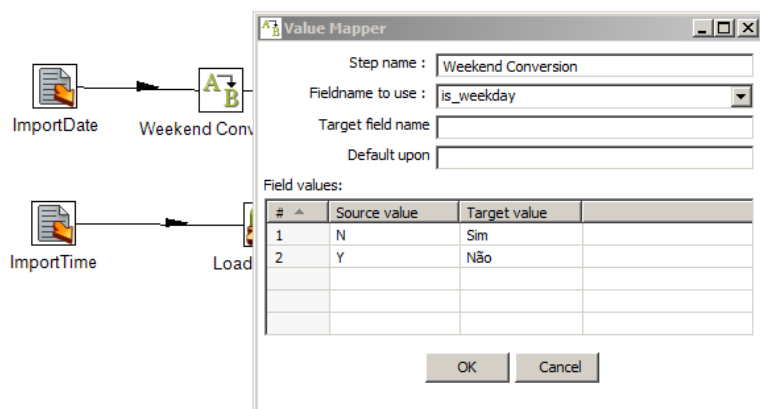


Figura 87 – ETL – Exemplo de procedimento mapeamento de valores

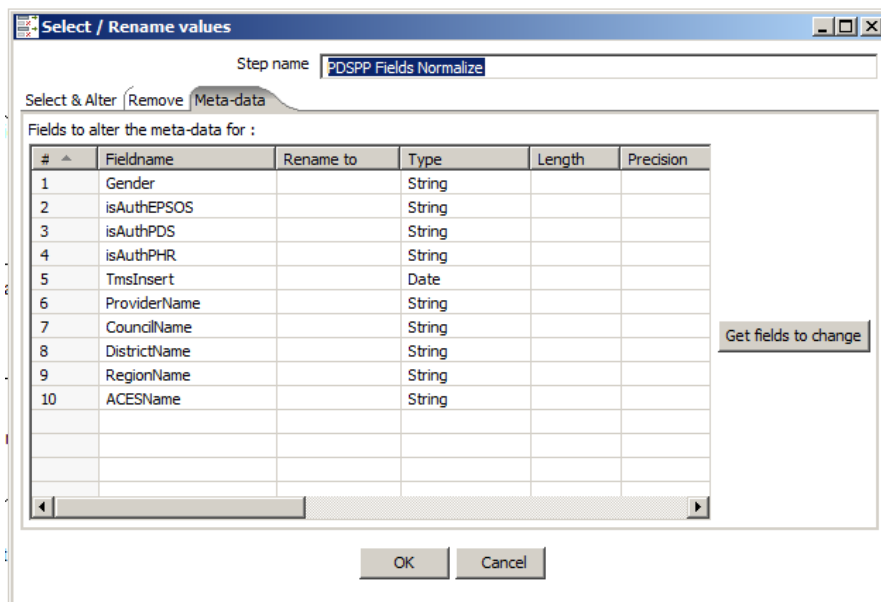


Figura 88 – ETL – Exemplo de procedimento de seleção e normalização de campos

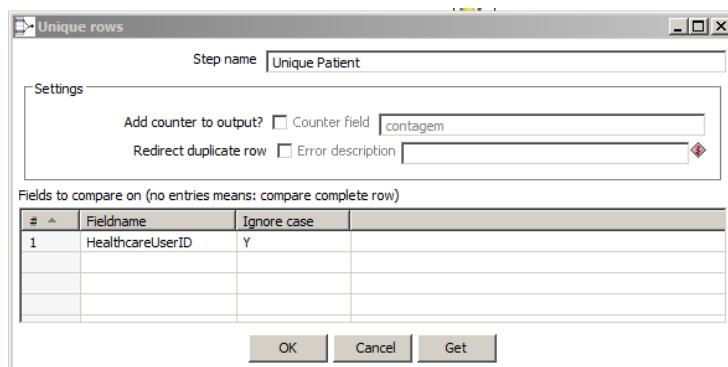


Figura 89 – ETL – Exemplo de procedimento para remoção de duplicados

Anexo 3 – Exemplo de procedimento para carregamento de dimensões do tipo 2

Step name: DimPatient lookup/update

Update the dimension?

Connection: PDSPI_DW

Target schema: dbo

Target table: DimPatient

Commit size: 100

Enable the cache?

Pre-load the cache?

Cache size in rows (0 = cache all): 5000

#	Dimension field	Stream field to compare with	Type of dimension update
1	Gender	Gender	Insert
2	BirthDate	BirthDate	Insert: Kimball Type II
3	ProviderName	ProviderName	Punch Through: Kimball Type I
4	ACESName	ACESName	Update: Correct error in last version
5	RegionName	RegionName	Update
6	DistrictName	DistrictName	Update
7	CouncilName	CouncilName	Insert
8	isAuthPDS	isAuthPDS	Punch through

Technical key field: PatientKey

Creation of technical key:

- Use table maximum + 1
- Use sequence
- Use auto increment field

Version field: Version

Stream Datefield:

Date range start field: EffectiveDate (Min. year: 1900)

Use an alternative start date? System date

Table daterange end: ExpireDate (Max. year: 2199)

Buttons: OK, Cancel, Get Fields, SQL

Figura 90 – ETL – Exemplo de procedimento de carregamento de dimensões do tipo 2

Anexo 4 – Exemplo de procedimento para carregar dados numa dimensão *junk*

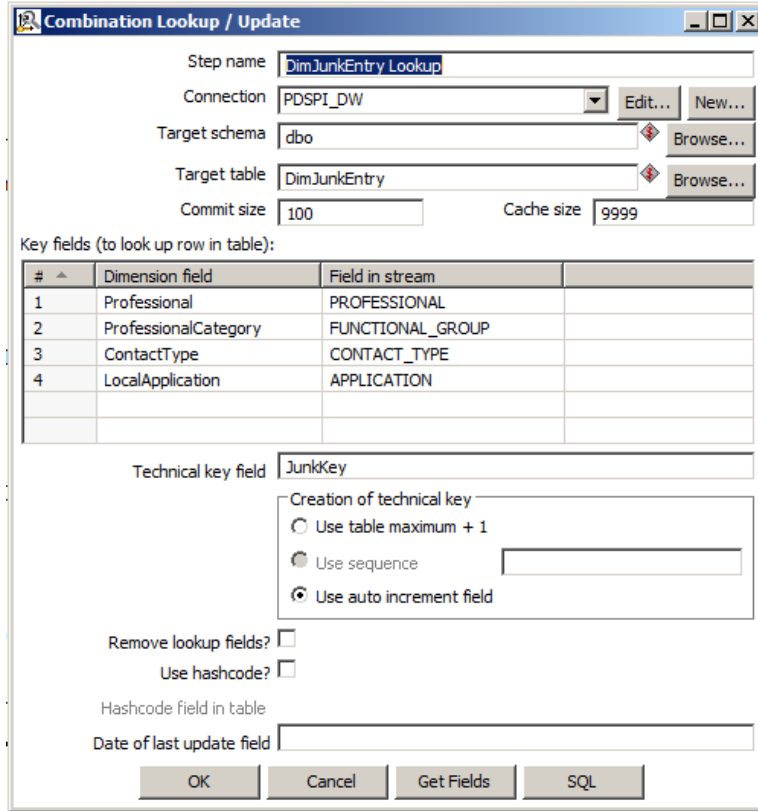


Figura 91 – ETL – Exemplo de procedimento de carregamento de uma dimensão *junk*

Anexo 5 – Tarefas de carregamento das tabelas de facto

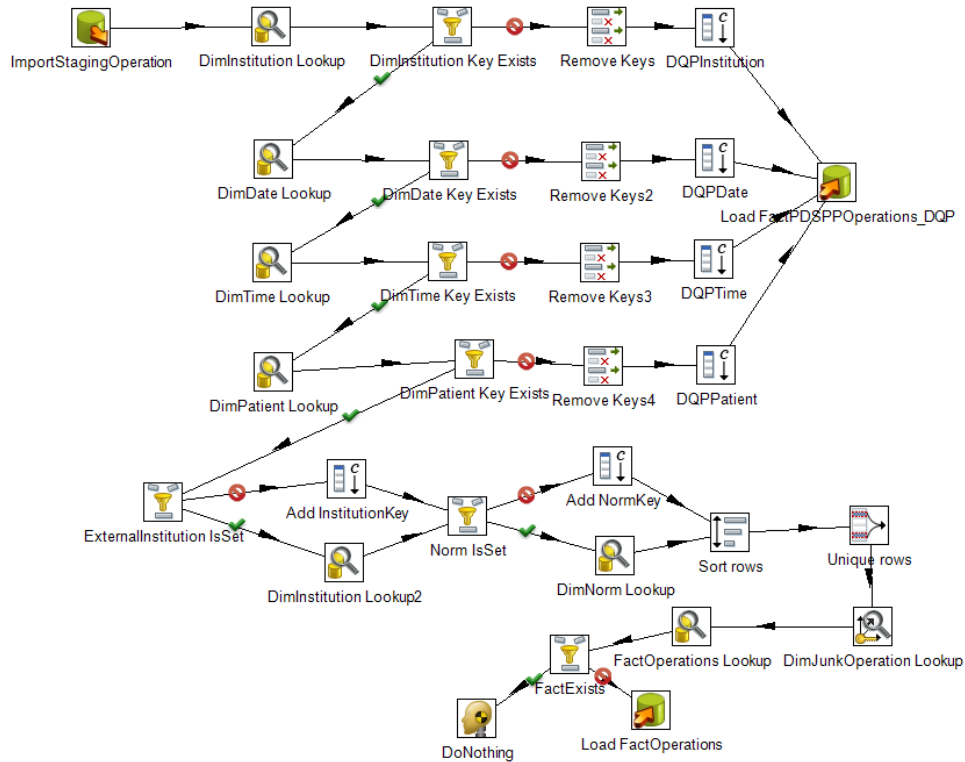


Figura 92 – ETL – Carregamento da tabela de facto FactPDSPPOperations

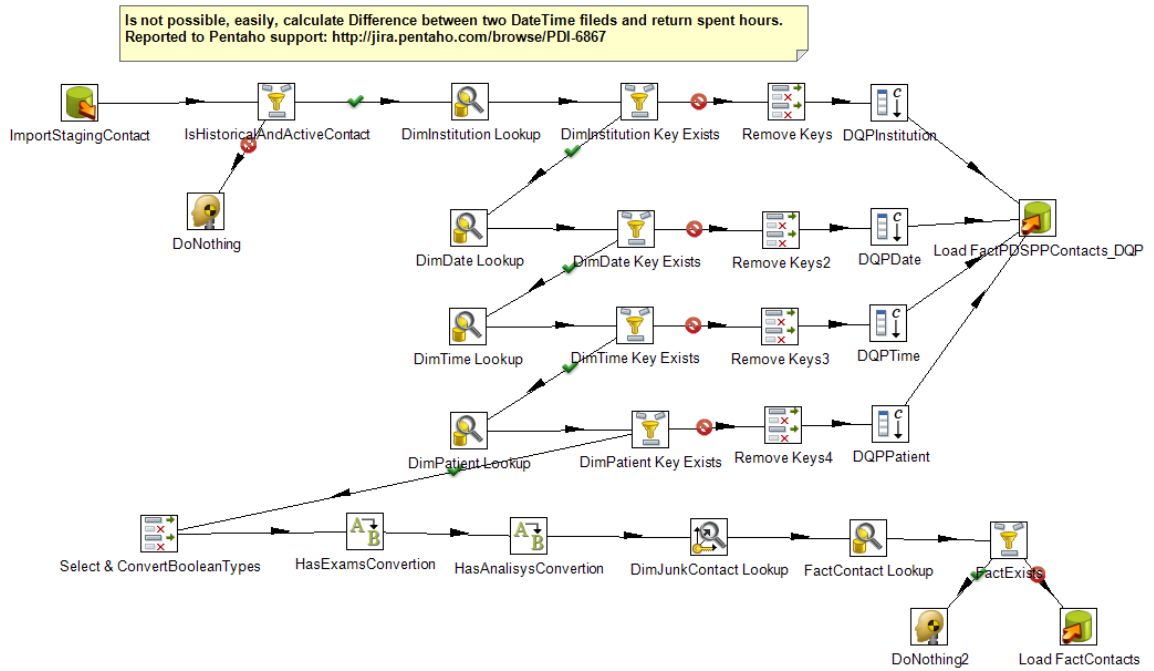


Figura 93 – ETL – Carregamento da tabela de facto FactPDSPPContacts

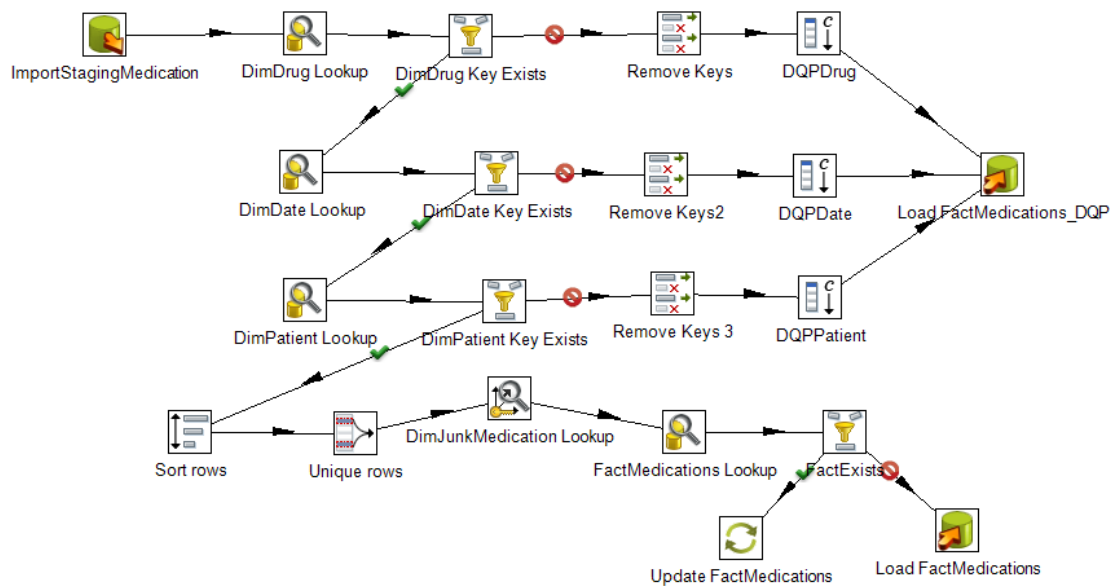


Figura 94 – ETL – Carregamento da tabela de facto FactPDSPUMedications

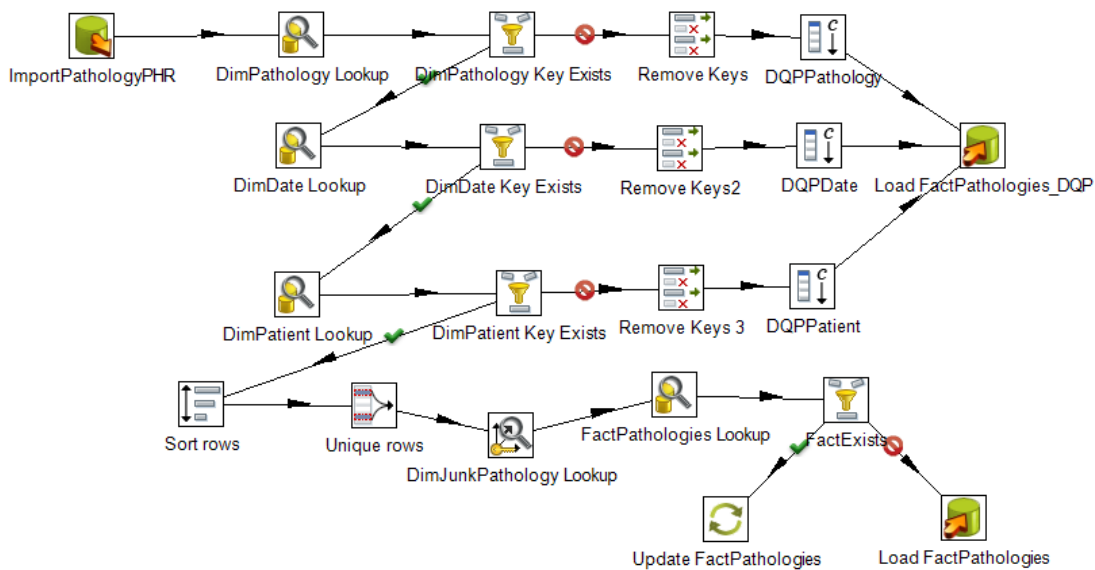


Figura 95 – ETL – Carregamento da tabela de facto FactPDSPUPathologies

Anexo 7 – Microstrategy Web (PDS-PI)

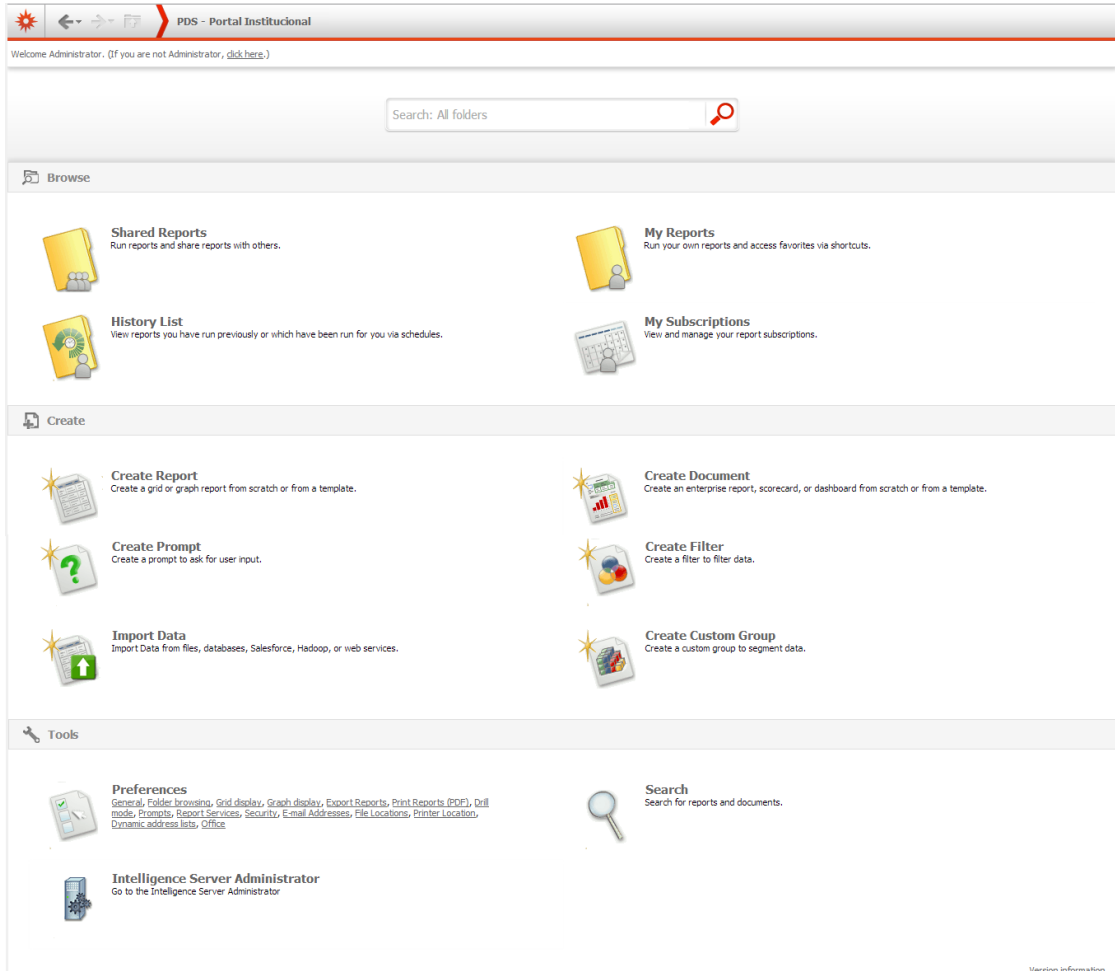


Figura 97 – Microstrategy Web – Interface PDS-PI

WIN-D5AHVM4IP9U > User Manager > ARS >

General **Project Access** Security Filters Groups Authentication Addresses Contacts ?

Security Role Selection	PDS - Portal Institucional
Privilege	Inherited Access
☒ <input checked="" type="checkbox"/> 1. Web Reporter	
☒ <input type="checkbox"/> 2. Web Analyst	
☒ <input type="checkbox"/> 3. Web Professional	
☒ <input checked="" type="checkbox"/> 4. Common Privileges	
☒ <input type="checkbox"/> 5. Office	
☒ <input type="checkbox"/> 6. Mobile	
☒ <input type="checkbox"/> 7. Deliveries	
☒ <input type="checkbox"/> 8. Multi-source option	
☒ <input type="checkbox"/> 9. Desktop Analyst	
☒ <input type="checkbox"/> 10. Desktop Designer	
☒ <input type="checkbox"/> 11. Architect	
☒ <input type="checkbox"/> 12. MicroStrategy Object Manager	
☒ <input type="checkbox"/> 13. MicroStrategy Command Manager	
☒ <input type="checkbox"/> 14. MicroStrategy Integrity Manager	
☒ <input type="checkbox"/> 15. Administration	

Privileges Obtained From:

- ✔ Security Role
- ✘ No project access granted
- * OLAP Services
- ** Report Services
- *** Distribution Services
- ✔ User/Group
- ✘ No access granted
- N/A - Not applicable per project

Figura 98 – Microstrategy Web – Configuração de perfis