



IPS Instituto
Politécnico de Setúbal
**Escola Superior de
Tecnologia de Setúbal**

Carlos Tavares *Data Warehouse
& Business Intelligence*
SAPA

Sistema de Atribuição de
Produtos de Apoio - SAPA

Trabalho de Projeto, submetido como requisito
parcial para obtenção do grau de **Mestre em
Informática de Gestão.**

Júri

Presidente (Doutor, José Cordeiro)

Orientador (Doutor, Cláudio Sapateiro)

Vogal (Doutor, Jacinto Estima)

Vogal (Doutor, Cláudio Sapateiro)

Dezembro de 2017

*Dedico este trabalho a quem diariamente me alimenta de vida
e me permite sonhar, a minha Família,
da qual destaco a minha esposa
e os meus filhos André e o Diogo, por tudo o que me proporcionam como Pai*

Agradecimentos

Ao meu orientador, Doutor Cláudio Sapateiro, pelos ensinamentos orientação e presença sempre que precisei.

Aos meus colegas e Diretor do Departamento de Análise e Gestão de Informação, tenho a plena consciência que foram fulcrais na orientação e disponibilização ao longo do projeto, com toda a paciência e partilha de conhecimento que me almejaram.

Resumo

A entidade gestora do sistema de atribuição de produtos de apoio, o Instituto Nacional para Reabilitação, identificou a necessidade de possuir uma ferramenta que permita efetuar análises e extrair diversa informação estatística sobre os processos de atribuição de qualquer produto, instrumento, equipamento, tecnologia e *software* usado por uma pessoa com deficiência, vulgo, produtos de apoio.

O presente projeto é a concretização do acima referenciado, desenvolvimento de um *Data Warehouse* e de relatórios, como uma ferramenta de *Business Intelligence*.

Inicialmente foi efetuado um estudo aprofundado das regras do processo de negócio e do modelo relacional, que suporta o sistema de gestão de base de dados operacionais. Foram efetuadas várias reuniões com os representantes da entidade gestora e um estudo dos requisitos do projeto, através da criação de *queries* ao sistema fonte de dados.

Após refinamento dos requisitos da informação estatística e conclusão da fase de análise, que inclui as atividades como identificação de fontes, destino de dados e modelização, foi desenvolvido o *Data Warehouse*. Trabalho envolvendo os inerentes processos de extração, transformação e carregamento, seguido dos testes, automatização de processos e colocação em produção.

O *Data Warehouse*, desenvolvido permitiu concretizar a implementação da infraestrutura de análise, efetuado com a ferramenta de *Business Intelligence* e assim atingir o objetivo final, possuir informação estatística, atualizada e precisa.

Palavras-chave: *Data Warehouse*, *Business Intelligence*, Produtos de Apoio.

Abstract

The managing body responsible for allocating disability equipment, the National Institute for Rehabilitation (*Instituto Nacional para Reabilitação*), has identified the need for a tool that allows the analysis and extraction of all the statistical information related to the process of granting disability equipment.

In order to achieve this objective, the current project developed a Data Warehouse and reporting system using a Business Intelligence tool.

The first step was to conduct an in-depth study focusing on the business process rules, and the relational model that supports the database management system. In order to determine the project's requirements, various meetings were held with the representatives of the managing body, as well as data source queries.

The Data Warehouse was created after refining the requirements of the statistical information and the completion of the analysis step, which included tasks such as source identification, data destination, and model-building.

This involved the extraction, transformation and loading of the data, which was followed by testing, process automation, and the production stage.

After loading the Data Warehouse, the data is organised in a way that allows the deployment of the analysis infrastructure, via the Business Intelligence tool, in order to achieve the final goal – to have updated and accurate statistical information.

Keywords: Data Warehouse, Business Intelligence, Disability Equipment.

Índice

Agradecimentos	4
Resumo	5
Abstract	6
Índice	7
Lista de Figuras	11
Lista de Tabelas	13
Lista de Siglas e Acrónimos	15
Capítulo 1	17
Introdução	17
1.1. Enquadramento	18
1.2. Objetivos do projeto	19
1.3. Metodologia de trabalho	21
1.4. Estrutura da dissertação	22
Capítulo 2	24
Visão global do sistema e processo de negócio SAPA	24
2.1. O processo de negócio	25
2.2. Conceitos	27
2.2.1. <i>Produtos de apoio</i>	27
2.2.2. <i>Agrupamento escolar</i>	27
2.2.3. <i>Equipa técnica multidisciplinar</i>	27
2.2.4. <i>Classificação Internacional de funcionalidade, incapacidade e saúde</i>	28
2.3. Atores	30
2.3.1. <i>Entidade gestora do SAPA (INR,I.P.)</i>	30
2.3.2. <i>Entidades financiadoras</i>	30
2.3.3. <i>Entidades prescritoras</i>	31
2.4. Principais funcionalidades do SAPA	31
2.4.1. <i>Tratar informação dos beneficiários</i>	31
2.4.2. <i>Tratar processos e prescrições</i>	32
Capítulo 3	35
Análise dos requisitos estatísticos	35

3.1. Perceção dos requisitos estatísticos.....	36
3.2. Bus Matrix	37
3.3. Granularidade temporal	38
3.4. Estudo direcionado para responder aos <i>reports</i>.....	40
Capítulo 4	41
Análise do modelo de dados SAPA	41
4.1. Interface com outros sistemas.....	42
4.2. Análise das tabelas operacionais	43
4.2.1. <i>Tabela PROCESSO_SAPA.....</i>	46
4.2.2. <i>Tabela PRODUTO_APOIO</i>	48
4.2.3. <i>Tabela PRESCRICAO.....</i>	49
4.3. Modelo de entidade e relacionamento.....	50
4.3.1. <i>Caracterização das entidades</i>	51
4.3.2. <i>Caracterização das associações</i>	51
Capítulo 5	55
Desenho e criação do <i>Data Warehouse</i>.....	55
5.1. Modelos do <i>Data Warehouse</i>.....	56
5.1.1. <i>Modelo processos</i>	56
5.1.2. <i>Modelo utilizadores.....</i>	58
5.1.3. <i>Modelo prescrições</i>	60
5.1.4. <i>Modelo produtos de apoio.....</i>	61
5.1.5. <i>Modelo entidades</i>	63
5.1.6. <i>Modelo fator ambiental.....</i>	65
5.1.7. <i>Modelo deficiências.....</i>	67
5.1.8. <i>Modelo atividades em que participa.....</i>	69
5.2. Tabelas de Dimensões	71
5.2.1. <i>Dimensões comuns a outros projetos.....</i>	74
5.2.2. <i>Dimensões manuais.....</i>	75
5.2.3. <i>Dimensões automáticas.....</i>	77
5.2.4. <i>Dimensões híbridas</i>	84
5.3. Tabelas de Factos	89
5.3.1. <i>FAC_SAPA_ENTIDADE_SAPA</i>	90
5.3.2. <i>FAC_SAPA_PRESCRICAO</i>	91
5.3.3. <i>FAC_SAPA_UTILIZADOR.....</i>	93

Capítulo 6.....	96
Cadeia ETL.....	96
6.1. Tecnologia usada.....	97
6.2. Caracterização dos processos ETL.....	98
6.2.1. <i>Parâmetros.....</i>	<i>101</i>
6.2.2. <i>ETL.....</i>	<i>102</i>
6.2.3. <i>Extração.....</i>	<i>102</i>
6.2.4. <i>Transformação.....</i>	<i>103</i>
6.2.5. <i>Carregamento.....</i>	<i>104</i>
6.3. Jobs.....	105
6.4. Work Jobs.....	106
6.4.1. <i>Data Sets.....</i>	<i>107</i>
6.4.2. <i>Dimensões manuais.....</i>	<i>111</i>
6.4.3. <i>Dimensões automáticas.....</i>	<i>112</i>
6.4.4. <i>Dimensões híbridas.....</i>	<i>115</i>
6.4.5. <i>Tabelas de factos.....</i>	<i>121</i>
6.5. Run Jobs.....	124
6.6. Chain Jobs.....	125
Capítulo 7.....	127
Relatórios BI.....	127
7.1. Tecnologia usada.....	128
7.2. Relatórios base do projeto.....	129
7.2.1. <i>Utilizadores.....</i>	<i>129</i>
7.2.2. <i>Processos.....</i>	<i>131</i>
7.2.3. <i>Entidades prescritoras.....</i>	<i>132</i>
7.2.4. <i>Prescrições.....</i>	<i>133</i>
7.2.5. <i>Produtos de apoio.....</i>	<i>135</i>
7.2.6. <i>Beneficiários.....</i>	<i>136</i>
7.2.7. <i>Classificação internacional de funcionalidade.....</i>	<i>137</i>
7.3. Relatórios não efetuados.....	139
Capítulo 8.....	141
Validações.....	141
8.1. Processo ETL.....	142
8.2. Validação dos dados dos relatórios.....	144

8.3. Validação do cliente	145
Capítulo 9	152
Conclusões	152
9.1. Conclusão e trabalhos futuros	153
Bibliografia.....	156
Anexo I.....	1
Tabelas do sistema operacional SAPA	1
Anexo II.....	2
Estudo direcionado para responder aos requisitos do INR	2
Anexo III.....	3
Requisitos do INR- Informação estatística.....	3
Anexo IV	4
Tabelas do Data Warehouse.....	4
Anexo V	5
Scripts da extração dos processos ETL.....	5
Anexo VI	6
Validação dos <i>reports</i> - testes finais	6
Anexo VII	7
Validação e percepção do grau de satisfação e relevância dos relatórios	7

Lista de Figuras

Figura 2-1	Visão Global do SAPA.....	26
Figura 2-2	Automatismo preenchimento campos em SAPA	33
Figura 4-1	Interações SAPA, com outros sistemas/subsistemas.....	43
Figura 4-2	Modelo entidade e relacionamento	52
Figura 4-3	Associações, chaves primárias e estrangeiras	53
Figura 5-1	Esquema do modelo Processos	57
Figura 5-2	Esquema do modelo Utilizadores	59
Figura 5-3...	Esquema do modelo Prescrições	60
Figura 5-4	Esquema do modelo Produtos de Apoio.....	62
Figura 5-5	Esquema do modelo Entidades	64
Figura 5-6	Esquema do modelo Fator ambiental	66
Figura 5-7	Esquema do modelo deficiências	68
Figura 5-8	Esquema do modelo atividades em que participa.....	70
Figura 5-9	Visão global das dimensões	73
Figura 6-1	Login com a ferramenta ETL usada.....	97
Figura 6-2	Processos ETL da categoria DIMS.....	99
Figura 6-3	Processos ETL da categoria FACS	100
Figura 6-4	Definição dos parâmetros do processo ETL	101
Figura 6-5	Diagrama geral do processo ETL	102
Figura 6-6	Programação do SQL, no processo de extração	103
Figura 6-7	Processo de transformação das categorias DIM's.....	103
Figura 6-8	Input de processo de carregamento	104
Figura 6-9	<i>Work jobs</i> da categoria DIMS	106
Figura 6-10	<i>Work jobs</i> da categoria FACS.....	106
Figura 6-11	Diagrama ETL do Data Set.....	107
Figura 6-12	Processo transformação do DS_ENTIDADE_INFO.....	108
Figura 6-13	Processo transformação do DS_PROCESSO_INFO	109
Figura 6-14	Processo transformação do DS_NISS_INFO	110
Figura 6-15	Propriedades dimensões manuais.....	111
Figura 6-16	Propriedades dimensões automáticas	113

Figura 6-17	Caracterização do processo ETL das tabelas híbridas	116
Figura 6-18	Processo extração da DIM_SAPA_PRODUTO_APOIO_PRESC	117
Figura 6-19	Processo transformação da DIM_SAPA_ENTIDADE	117
Figura 6-20	Processo ETL da DIM_SAPA_UTILIZADOR	118
Figura 6-21	Diagrama do processo update da DIM_SAPA_PRESCRICAO	120
Figura 6-22	Processo extração da FAC_SAPA_PRESCRICAO	123
Figura 6-23	<i>Run jobs</i> da categoria DIMS.....	124
Figura 6-24	<i>Run jobs</i> da categoria FACS.....	125
Figura 6-25	<i>Chain jobs</i> da categoria DIMS.....	125
Figura 6-26	<i>Chain jobs</i> da categoria FACS	126
Figura 7-1	Framework do projeto SAPA no MicroStrategy	128
Figura 7-2	Filtro definido pelos utilizadores <i>MicroStrategy</i>	128
Figura 7-3	Relatório 1.03, inerente aos utilizadores SAPA.....	130
Figura 7-4	Relatório 2.04, inerente aos processos SAPA	132
Figura 7-5	Relatório 3.01, inerente às entidades prescritoras	133
Figura 7-6	Relatório 4.03, inerente às prescrições	134
Figura 7-7	Relatório 5.01, inerente aos produtos de apoio.....	135
Figura 7-8	Desenvolvimento do relatório 6.03, inerente aos beneficiários	137
Figura 7-9	Relatório 7.08, inerente à CIF	138
Figura 8-1	Tempo de execução da cadeia ETL.....	142
Figura 8-2	Tempo de execução dos jobs FAC	142
Figura 8-3	Tempo de execução dos jobs DIM	143
Figura 8-4	Questão n.º1 do inquérito de validação dos <i>reports</i>	146
Figura 8-5	Questão n.º2 do inquérito de validação dos <i>reports</i>	147
Figura 8-6	Questão n.º3 do inquérito de validação dos <i>reports</i>	147
Figura 8-7	Questão n.º4 do inquérito de validação dos <i>reports</i>	148
Figura 8-8	Questão n.º5 do inquérito de validação dos <i>reports</i>	149
Figura 8-9	Questão n.º6 do inquérito de validação dos <i>reports</i>	149
Figura 8-10	Questão n.º7 do inquérito de validação dos <i>reports</i>	150
Figura 8-11	Questão n.º8 do inquérito de validação dos <i>reports</i>	151

Lista de Tabelas

Tabela 2-1	Visão geral do CIF	29
Tabela 3-1	Bus Matrix SAPA.....	38
Tabela 4-1	Associações entre as tabelas SAPA.....	45
Tabela 4-2	Campos da tabela PROCESSO_SAPA	47
Tabela 4-3	Campos da tabela PRODUTO_APOIO.....	49
Tabela 4-4	Campos da tabela PRESCRICAO	50
Tabela 5-1	Dimensões do modelo processos	58
Tabela 5-2	Dimensões do modelo utilizadores	60
Tabela 5-3	Dimensões do modelo prescrições	61
Tabela 5-4	Dimensões do modelo produtos de apoio.....	63
Tabela 5-5	Dimensões do modelo entidades	65
Tabela 5-6	Dimensões do modelo fator ambiental.....	67
Tabela 5-7	Dimensões do modelo deficiências.....	69
Tabela 5-8	Dimensões do modelo atividades em que participa.....	71
Tabela 5-9	Valores da FAC_SAPA_ENTIDADE_SAPA.....	91
Tabela 5-10	Valores da FAC_SAPA_PRESCRICAO.....	93
Tabela 5-11	Valores da FAC_SAPA_UTILIZADOR.....	95
Tabela 6-1	Processos ETL versus Data Sets.....	108
Tabela 6-2	Processos ETL versus dimensões manuais	111
Tabela 6-3	Processos ETL versus dimensões automáticas.....	112
Tabela 6-4	Tabela fonte usada na extração das dimensões automáticas	114
Tabela 6-5	Processos ETL versus dimensões híbridas	115
Tabela 6-6	Processos ETL versus tabelas de factos	122
Tabela 7-1	Utilizadores, requisitos versus <i>reports</i>	131
Tabela 7-2	Processos, requisitos versus <i>reports</i>	132
Tabela 7-3	Entidades prescritoras, requisitos versus <i>reports</i>	133
Tabela 7-4	Prescrições versus <i>reports</i>	134
Tabela 7-5	Produtos de apoio versus <i>reports</i>	136
Tabela 7-6	Beneficiários versus <i>reports</i>	137
Tabela 7-7	CIF versus <i>reports</i>	139

Tabela 8-1	Dimensão dos dados alvo do processo ETL	144
Tabela 9-1	Documento de requisitos versus <i>reports</i> desenvolvidos.....	154

Lista de Siglas e Acrónimos

ACSS	Administração Central do Sistema de Saúde
BI	<i>Business Intelligence</i>
CID	Classificação Internacional de Deficiência
CIF	Classificação Internacional de funcionalidade, incapacidade e saúde
DW	<i>Data Warehouse</i>
ETL	<i>Extract, Transform, Load</i>
GMD	Gestão de Modelos de Documentos
GUS	Gestão de Utilizadores e Segurança
GRS	Gestão de Regras do Sistema
IEFP	Instituto de Emprego e Formação Profissional
INR	Instituto Nacional de Reabilitação
IDQ	Identificação e Qualificação
ISS	Instituto da Segurança Social
II	Instituto de Informática
ME	Ministério da Educação
MER	Modelo Entidade e relacionamento
MIA	Modelo de Informação Auxiliar
MS	Ministério da Saúde

MTSS Ministério do Trabalho, Solidariedade e Segurança Social

NISS Número de Identificação da Segurança Social

RNU Registo Nacional de Utentes

SAPA Sistema de Atribuição de Produtos de Apoio

SISS Sistema de Informação da Segurança Social

SSN Serviço de Suporte de Negócio

SQL *Structured Query Language*

Capítulo 1

Introdução

Neste primeiro capítulo, é feito um enquadramento e exposto qual a motivação desta dissertação de mestrado, a identificação e definição do problema abordado e qual a finalidade e objetivos do projeto. É sumarizado o conjunto de informação que irá ser tratada para fins estatísticos, resultante de vários processos do Sistema de Atribuição de Produtos de Apoio - SAPA. Posteriormente é apresentada a metodologia usada na elaboração do projeto e termina, apresentando um breve resumo de cada um dos capítulos.

1.1. Enquadramento

Este projeto é mais um passo para a concretização do transposto no Decreto-Lei 93/2009, a criação de um *data warehouse* (DW) e de relatórios, com uma ferramenta de *Business Intelligence* (BI), áreas em que se começa a apostar fortemente na Administração Pública, através das tecnologias de informação, por forma a obter uma melhor gestão dos seus recursos e uma capacidade de visão e análise de todas as medidas e aplicações.

A revisão do sistema supletivo de financiamento, prescrição e atribuição de ajudas técnicas deu origem ao novo sistema de informação denominado Sistema de Atribuição de Produtos de Apoio (SAPA), desenvolvido muito recentemente para a gestão desta medida.

A entidade gestora do SAPA, o Instituto Nacional de Reabilitação (INR), é um instituto público integrado e atualmente tutelado pelo Ministério do Trabalho, Solidariedade e Segurança Social (MTSS), e dotado de autonomia administrativa e património próprio.

O INR pretendia possuir uma ferramenta que lhe permita efetuar análises e extrair toda a informação estatística que necessita sobre os processos de atribuição de produtos. É de elevada importância possuir informações atualizadas e precisas sobre o estado e resultados deste tema, pois é com base nestes dados que são tomadas as decisões de gestão.

A resposta à necessidade do INR é o desenvolvimento de um processo de recolha e tratamento de informação, com uma ferramenta de *Business Intelligence*, alimentada através de um grande e eficiente repositório de dados, um *Data Warehouse*.

O desenvolvimento do DW, com uma arquitetura que visa agilizar e acelerar a produção dos relatórios estatísticos, através da ferramenta de BI, relatórios esses, referentes aos processos de atribuição de produtos de apoio a nível nacional. Vai permitir ao INR otimizar os processos de tomada de decisão, desbloquear os estrangulamentos (que existem na transformação dos dados em informação) e que se não forem resolvidos podem reduzir a produtividade, aumentar os custos e no limite causar perdas financeiras preocupantes.

O projeto *Data Warehouse & Business Intelligence* SAPA, foi desenvolvido no Instituto de Informática - Ministério do Trabalho, Solidariedade e Segurança Social, no Departamento de Análise e Gestão de Informação, onde estive integrado numa equipa do departamento, composta por mais dois elementos.

O Instituto de Informática, I.P. (II,I.P.) é um instituto público integrado na administração indireta do Estado, dotado de autonomia administrativa e financeira e património próprio. É um organismo central com intervenção sobre todo o território nacional.

1.2. Objetivos do projeto

A capacidade de identificar possíveis necessidades, melhorar o processo de negócio vigente, gerir da melhor forma os recursos humanos e financeiros disponíveis são de forma geral o objetivo que o presente projeto visa dar resposta. Este permitira ter um controlo, da atribuição de produtos de apoio e da gestão de reembolsos/pagamentos de prescrições a serem financiadas pelo Instituto da Segurança Social e a troca de informação entre as entidades atoras no SAPA.

Os requisitos base deste projeto foram disponibilizados pelo INR, num documento que destaca os relatórios/informação que pretendiam ter como output final, documento que pode ser consultado no Anexo III – Requisitos do INR, informação estatística. Este documento foi a base para desenvolvimento do *Data Warehouse* e assunto foco posteriormente, no capítulo 3-Análise dos requisitos estatísticos.

O conjunto dos requisitos a implementar no sistema, consistem em indicadores e contagens a elaborar sobre os dados do SAPA. O conjunto de informação que ira ser tratada para fins estatísticos, abrange dados resultantes de vários processos do sistema SAPA, destacando de forma sucinta:

- Informações dos beneficiários SAPA.
- Dados dos processos de prescrição através dos números de identificação da segurança social (NISS), nº utente do serviço nacional de saúde ou outro que se ache necessário.
- Histórico de prescrições de produtos de apoio por beneficiário.

- Informações relativas às prescrições de produtos de apoio, por todas as entidades envolvidas.
- O controlo da execução orçamental por distrito, relativa à Segurança Social.
- Informação de prescrições de produtos de apoio concedidos a beneficiários, pelo subsistema financiador da Administração Central do Sistema de Saúde (ACSS), Ministério da Saúde.
- Informação dos diversos subsistemas/organismos financiadores;
- Informação sobre os produtos de apoio/beneficiários caracterizando as várias vertentes enquadradas na classificação internacional de deficiência (CID).
- Informação das prescrições por subsistema;

A construção de uma base centralizada de informação para exploração e análise estatística, irá permitir aos organismos envolvidos, a caracterização e consulta da informação de beneficiários SAPA, bem como dos respetivos produtos de apoio atribuídos, a gestão de prescrições de produtos de apoio, e a gestão da lista de produtos de apoio a atribuir, de acordo com a competência estabelecida para cada organismo.

A análise qualitativa e quantitativa dos processos de atribuição de produtos de apoio a nível nacional, gerida pelo Instituto Nacional de Reabilitação (INR) e traduzida na informação estatística, face ao seu conteúdo pode-se caracterizar por ajudar em três grandes e relevantes áreas, estratégica, de gestão e funcional.

➤ Estratégica

- Melhoria dos níveis de controlo e acompanhamento.
- Agilidade na produção de indicadores.
- Aumento da capacidade de análise.

➤ De gestão

- Disponibilização de indicadores da gestão do sistema do SAPA.
- Existência de uma ferramenta de auxílio à análise da evolução temporal.

➤ Funcional

- Maior fluidez de informação.
- Redução do tempo utilizado na gestão da informação.

1.3. Metodologia de trabalho

Relativamente à metodologia utilizada no presente projeto a sua sustentação e estudo passou por uma criteriosa e exaustiva recolha de elementos que permitem o conhecimento sobre as características do processo de negócio sustentado pelo sistema aplicacional SAPA, e demais ferramentas tecnológicas que vão servir para o desenvolvimento do *Data Warehouse & Business Intelligence SAPA*. Sucintamente é caracterizado pelas seguintes etapas:

1. Estudo das tecnologias usadas; *Toad for Oracle*, *IBM InfoSpher DataStage* e *Microstrategy*.
2. Estudo aprofundado do processo de negócio, associado ao sistema SAPA.
3. Estudo do modelo relacional que suporta o Sistema de Gestão de Base de Dados (SGBD) SAPA.
4. Análise exaustiva das tabelas operacionais da base de dados SAPA.
5. Estudo e identificação das necessidades de informação do INR, sustentado principalmente em reuniões.
6. Identificação dos requisitos iniciais do *Data Warehouse*.
7. Definição e desenvolvimento do modelo dimensional do *Data Warehouse*.
8. Desenvolvimento e execução do processo de extração, transformação e carregamento para o *Data Warehouse* SAPA.
9. Desenvolvimento dos relatórios/*dashboards*, respondendo às necessidades do INR.
10. Avaliação e feedback do INR dos relatórios/*dashboards* finais.

As ações acima transcritas estão numeradas de certa forma pela ordem lógica com que foram executadas, mas existem *nuances* pois nem sempre tal aconteceu, especificando tais pontos:

- O frisado no ponto 1; com exceção do *Toad for Oracle*, ferramenta com que

estou familiarizado, foi dado um maior ênfase e conseqüentemente estudo dirigido às necessidades identificadas no desenvolvimento do projeto, à medida que as ferramentas usadas foram necessárias.

- O frisado nos pontos 3,4 e 5; são ações que se complementaram na aquisição do conhecimento necessário, pelo que foram alvo de atenção muitas vezes, paralelamente.

Saliento que a tecnologia usada para o processo ETL, carregamento das tabelas do DW, a ferramenta da IBM, o *Infosphere Datastage and Infosphere Qualitystage*, é a usada no Instituto de Informática, pelo que de forma natural foi a escolha para o desenvolvimento do projeto SAPA, igual razão pela opção da tecnologia de BI, usada para desenvolver relatórios estatísticos, o *MicroStrategy Developer*.

1.4. Estrutura da dissertação

Para além da presente introdução onde se contextualiza o problema e se especificam os objetivos a atingir, este documento conta com mais nove capítulos.

No segundo capítulo é dado a conhecer o processo de negócio inerente ao projeto SAPA. É feita uma abordagem direcionada, para se familiarizar o leitor, com as regras de negócio, conceitos, processos, atores e principais funcionalidades do core que é suportado pelo sistema SAPA.

O capítulo três aborda e dá a conhecer os requisitos solicitados pela entidade gestora do projeto e que são a principal bússola orientadora do projeto.

O estudo e conhecimento sobre o modelo de dados que suporta o sistema operacional do SAPA, é o cerne do capítulo quatro, assim como a identificação de quais os outros sistemas e subsistemas, que interagem com o sistema SAPA.

No capítulo cinco é descrito o desenho do *data warehouse* SAPA, dando a conhecer os modelos desenvolvidos. É feita uma caracterização destes, identificando-se a sua granularidades, tabelas fonte e indicadores que se propõem a responder e claro inerentes tabelas de factos e dimensões

O capítulo seis caracteriza e explica os processos ETL, dando a conhecer as

nomenclaturas adotadas, categorias e características comuns a toda cadeia ETL, assim como dos *Jobs* que compõem a cadeia de carregamento do *data warehouse* SAPA.

Os relatórios BI desenvolvidos, são o tema do capítulo 7, seccionando estes de acordo com as suas características. É feito também o paralelismo dos requisitos respondidos por estes versus os relatórios desenvolvidos.

A validação final sobre os relatórios estatísticos é transposta no capítulo 8, contemplando a eficiência do processo ETL, a qualidade dos dados, da informação dos relatórios e por fim ao nível dos utilizadores finais.

No nono e ultimo capitulo, é feita uma sumarização dos objetivos respondidos pelo presente projeto e principais desafios e oportunidades que desde provieram, assim como as perspectivas de trabalhos futuros que deste podem advir.

Capítulo 2

Visão global do sistema e processo de negócio SAPA

Este segundo capítulo tem como objetivo, dar a conhecer as regras de negócio, conceitos, processos, atores e principais funcionalidades do sistema, de forma a familiarizar o leitor com os mesmos e permitir assim a compreensão, respetivas motivações e decisões tomadas no desenvolvimento do DW.

O transposto neste capítulo é parte do esforço desenvolvido para responder aos requisitos do projeto. Foi necessário efetuar um estudo aprofundado das regras do processo de negócio SAPA. O presente conhecimento foi adquirido essencialmente através da leitura de legislação e nas reuniões com o INR.

2.1. O processo de negócio

O Sistema de Atribuição de Produtos de Apoio tem como finalidade implementar um conjunto de processos que permitem de uma forma integrada e centralizada, dar resposta às necessidades das pessoas com deficiências ou incapacidade, permanente ou temporária, através da atribuição de produtos indispensáveis à sua autonomia, integração e funcionalidade.

No âmbito do programa de simplificação administrativa e legislativa, para criar as condições necessárias à implementação das medidas SIMPLEX, desenvolveu-se o Sistema de Atribuição de Produtos de Apoio, denominado SAPA, conforme Decreto-Lei n.º 93/2009, de 16 de abril (Decreto-lei n.º 93-2009) (Decreto-Lei n.º 42-2011), substituiu à data o então sistema supletivo de prescrição e financiamento de ajudas técnicas e tecnologias de apoio. O SAPA veio responder à necessidade de se proceder à revisão do sistema supletivo de financiamento, prescrição e atribuição de ajudas técnicas e conceção de um novo sistema integrado.

Este permitiu proceder a uma reformulação do sistema que vigorava, identificar as dificuldades existentes e adotar medidas necessárias para garantir a igualdade de oportunidades de todos os cidadãos, promover a integração e participação das pessoas com deficiência e em situação de dependência na sociedade e promover uma maior justiça social.

Assim, o decreto-lei 93/2009 visa o enquadramento específico para o Sistema de Atribuição de Produtos de Apoio, que veio substituir o então sistema supletivo de ajudas técnicas e tecnologias de apoio, designadas agora por produtos de apoio nos termos da nomenclatura utilizada na Norma ISO 9999:2007.

Consideram-se produtos de apoio, qualquer produto (incluindo dispositivos, equipamento, instrumentos, tecnologia e software) especialmente produzido ou geralmente disponível, para prevenir, compensar, monitorizar, aliviar ou neutralizar as incapacidades, limitações das atividades e restrições na participação.

Antes do SAPA existir, os processos de atribuição de produtos de apoio eram realizados na sua maioria manualmente, com exceção de alguns tratamentos automáticos, existentes no Ministério da Saúde e no Instituto Emprego e Formação

Profissional (IEFP, IP.) Estes tratamentos automáticos não comunicavam entre si, não permitiam a consolidação de informação e não tinham integração com o Sistema de Informação da Segurança Social (SISS).

O SAPA veio permitir uma gestão da atribuição desses produtos de forma a garantir, por um lado, a eficácia do sistema, através da operacionalidade e eficiência dos seus mecanismos e a sua aplicação criteriosa e, por outro, a desburocratização do sistema ao simplificar as formalidades exigidas pelos serviços prescritores criando uma base de dados de registo de pedidos.

O SAPA permite que as entidades que o compõem estejam interligadas por um sistema informático centralizado cuja gestão da informação é da competência do Instituto Nacional para a Reabilitação, I. P. (INR, I. P.), permitindo obter a consolidação e contextualização da informação a nível nacional.

O sistema é assente na conceção de uma base de dados de registo, com o objetivo de garantir a eficácia do sistema, a operacionalidade e a eficiência dos mecanismos do SAPA, promovendo a comunicação de forma integrada, com diversas estruturas do Ministério do Trabalho e da Solidariedade Social (MTSS), do Ministério da Saúde (MS), do Ministério da Educação (ME) e outras entidades como centros prescritores com equipas multidisciplinares, que venham a integrar/interagir com o sistema.

CONFIDENCIAL

Figura 2-1 Visão Global do SAPA

São igualmente partilhadas as atuais infraestruturas do SISS, otimizando os recursos existentes, reduzindo a redundância da informação, a duplicação de tarefas, os tempos gastos nos processos e assim obter uma maior rapidez na consolidação da informação.

Este sistema informático centralizado, permite a desburocratização, desmaterialização e simplificação do processo de negócio, pois o que estava disperso por várias entidades e sem comunicação é congregado no SAPA.

2.2. Conceitos

O Para entender o sistema de informação é fundamental conhecer conceitos que serão frisados posteriormente, possibilitando a compreensão e interpretação do SAPA (Despacho n.º 6133/2012),(Despacho n.º 6133/2012).

2.2.1. Produtos de apoio

No âmbito do SAPA, entende-se como produtos de apoio, qualquer produto, instrumento, equipamento ou sistema técnico usado por uma pessoa com deficiência, especialmente produzido ou disponível que previne, compensa, atenua ou neutraliza a limitação funcional ou de participação nas atividades diárias (Despacho n.º 16313/2012),(Despacho n.º 14278/2014).

2.2.2. Agrupamento escolar

Entidade que ao nível do Ministério da Educação é responsável por adquirir e entregar o produto de apoio. Um agrupamento de escolas (de vários níveis de ensino) congrega um conjunto de escolas de uma determinada zona geográfica, facilitando a gestão dos recursos humanos e materiais.

2.2.3. Equipa técnica multidisciplinar

A equipa técnica multidisciplinar é constituída, no mínimo, por dois técnicos com saberes transversais das várias áreas de intervenção em reabilitação. Integra, designadamente, médico, enfermeiro, fisioterapeuta, terapeuta ocupacional, terapeuta da

fala, psicólogo, docente, recorrendo quando necessário a outros técnicos em função de cada uma das situações, nomeadamente técnicos de serviço social, protésicos, engenheiros e ergonomistas, para que a identificação dos produtos de apoio seja a mais adequada à situação concreta, no contexto de vida da pessoa.

2.2.4. Classificação Internacional de funcionalidade, incapacidade e saúde

A classificação internacional de funcionalidade, incapacidade e saúde (CIF) engloba todos os aspetos da saúde humana e alguns componentes relevantes para a saúde relacionados com o bem-estar e descreve-os em termos de domínios de saúde e domínios relacionados com a saúde. Permite descrever situações relacionadas com a funcionalidade do ser humano e as suas restrições e serve como enquadramento para organizar esta informação.

Saliento que é traduzido no modelo de dados SAPA a forma como é estruturada a informação CIF, refletindo-se como é óbvio no entendimento de componentes do DW e inerentes relatórios estatísticos, pelo que é aqui aprofundado o tema.

A CIF estrutura a informação de maneira útil, integrada e facilmente acessível, organizando a informação em duas partes; 1) Funcionalidade e da incapacidade; 2) Fatores contextuais (CIF, 2004).

1. Funcionalidade e da Incapacidade

O componente corpo inclui duas classificações, uma para as funções dos sistemas orgânicos e outra para as estruturas do corpo. Nas duas classificações os capítulos estão organizados de acordo com os sistemas orgânicos.

O componente atividades e participação cobrem a faixa completa de domínios que indicam os aspetos da funcionalidade, tanto na perspetiva individual como social.

2. Fatores Contextuais

Os fatores ambientais têm um impacto sobre todos os componentes da funcionalidade e da incapacidade e estão organizados de forma sequencial, do ambiente mais imediato do indivíduo até ao ambiente geral.

Os fatores pessoais também são um componente dos fatores contextuais, mas eles não estão classificados na CIF devido à grande variação social e cultural associada aos mesmos.

No contexto da saúde e de forma a cimentar este tópico, salienta-se que:

- Funções do corpo são as funções fisiológicas dos sistemas orgânicos (incluindo as funções psicológicas).
- Estruturas do corpo são as partes anatómicas do corpo, tais como, órgãos, membros e seus componentes.
- Deficiências são problemas nas funções ou nas estruturas do corpo, tais como, um desvio importante ou uma perda.
- Atividade é a execução de uma tarefa ou ação por um indivíduo.
- Participação é o envolvimento de um indivíduo numa situação da vida real.
- Limitações da atividade são dificuldades que um indivíduo pode ter na execução de atividades.
- Restrições na participação, são problemas que um indivíduo pode enfrentar quando está envolvido em situações da vida real.
- Fatores ambientais constituem o ambiente físico, social e atitudinal em que as pessoas vivem e conduzem sua vida.

Tabela 2-1 Visão geral do CIF

Componentes	Parte 1: Funcionalidade e Incapacidade		Parte 2: Factores Contextuais	
	Funções e Estruturas do Corpo	Actividades e Participação	Factores Ambientais	Factores Pessoais
Dominios	Funções do Corpo Estruturas do Corpo	Áreas Vitais (tarefas, acções)	Influências externas sobre a funcionalidade e a incapacidade	Influências internas sobre a funcionalidade e a incapacidade
Constructos	Mudança nas funções do corpo (fisiológicas) Mudança nas estruturas do corpo (anatômicas)	Capacidade Execução de tarefas num ambiente padrão Desempenho/Execução de tarefas no ambiente habitual	Impacto facilitador ou limitador das características do mundo físico, social e atitudinal	Impacto dos atributos de uma pessoa
Aspectos positivos	Integridade funcional e estrutural	Actividades Participação	Facilitadores	Não aplicável
	Funcionalidade			
Aspectos negativos	Deficiência	Limitação da actividade Restrição da participação	Barreiras	Não aplicável
	Incapacidade			

2.3. Atores

O Para entender o sistema de informação é fundamental conhecer os seus atores, pelo que neste subcapítulo é dado a conhecer os intervenientes no processo de negócio, conhecimento, essencial para interpretação do presente projeto.

2.3.1. Entidade gestora do SAPA (INR, I.P.)

O Instituto Nacional para Reabilitação, IP, (INR, I.P.) é um instituto público integrado e atualmente tutelado pelo Ministério do Trabalho, Solidariedade e Segurança Social, e dotado de autonomia administrativa e património próprio (Decreto-lei n.º 93-2009), (INR 2017).

A garantia de igualdade de oportunidades, o combate à discriminação e a valorização das pessoas com deficiência, numa perspetiva de promoção dos seus direitos fundamentais, são os princípios que norteiam a atuação do Instituto Nacional para Reabilitação, IP.

Compete ao INR, I.P, como entidade gestora do SAPA:

- Constituição e a atualização de um catálogo indicativo de produtos de apoio, que são propostos pelas entidades financiadoras.
- Gestão da informação do SAPA.
- Apresentação, às entidades financiadoras, de um relatório anual de execução do SAPA.

2.3.2. Entidades financiadoras

Entidades financiadoras, são as entidades que comparticipam a aquisição do produto de apoio com base na prescrição passada por uma entidade prescritora (Despacho n.º 5128/2013).

As entidades que constam em SAPA são:

- As entidades hospitalares, através da Administração Central do Sistema de Saúde, I. P. (ACSS, I.P.).
- Os Centros Distritais de Segurança Social, através do ISS, I. P..

- Os serviços financiadores de produtos de apoio para a formação profissional e ou emprego, através do IEFP, I. P.

2.3.3. Entidades prescritoras

Designa-se de entidades prescritoras, as entidades, serviços, organismo ou centro de referência à qual pertence uma equipa técnica multidisciplinar ou o médico que procede à prescrição de um produto de apoio (Despacho n.º 5128/2013).

Os produtos de apoio são prescritos pelos centros de saúde, unidades hospitalares e equipas multidisciplinares. No caso de prescrição médica obrigatória, os produtos de apoio são prescritos apenas por médicos. Os produtos de apoio prescritos do âmbito dos centros especializados, designados pela entidade prescritora, são prescritos por uma equipa técnica multidisciplinar. Com vista ao financiamento dos produtos de apoio as entidades intervenientes no SAPA devem obrigatoriamente preencher uma ficha de prescrição (Portaria n.º 78/2015).

2.4. Principais funcionalidades do SAPA

O objetivo deste subcapítulo é identificar e descrever de forma muito genérica, as funcionalidades que identifico como serem as mais relevantes do sistema de informação SAPA.

2.4.1. Tratar informação dos beneficiários

Esta funcionalidade permite identificar os beneficiários, através de dados que caracterizam os mesmos e assim, de forma automática, identificar os beneficiários, permitindo de forma mais ágil, atribuir processos, consultar o histórico de prescrições de produtos de apoio dos beneficiários e o acesso a registo de prescrições.

- Tratar informação dos beneficiários:
 1. Consultar histórico das prescrições de beneficiários.
 2. Integrar atribuição NISS do beneficiário.
 3. Identificar beneficiário.

1. Consultar histórico das prescrições de beneficiários

Tem como objetivo permitir ao utilizador consultar o histórico de prescrições de um dado beneficiário:

- O sistema permite ao utilizador identificar o beneficiário.
- O sistema permite ao utilizador consultar o histórico de prescrições, através do NISS ou do número de utente do beneficiário.
- O sistema permite ao utilizador consultar a informação detalhada de cada prescrição.

2. Integrar Atribuição NISS do beneficiário

Tem como objetivo a atribuir NISS's, a processos que tenham ficado pendentes por falta de NISS.

- Identifica um dado beneficiário através de dados mínimos.
- Verifica se existe registo do beneficiário que satisfaça os dados mínimos recolhidos, caso afirmativo efetua a atribuição de NISS a uma prescrição pendente.

3. Identificar Beneficiário

Quando o utilizador sapa pretende identificar um beneficiário, através de dados mínimos.

2.4.2. Tratar processos e prescrições

Agrega uma serie de funcionalidades ao nível das prescrições e processos, nomeadamente ações como registar, alterar, pesquisar, validar e consultar a prescrição e alterar estado da prescrição, validar nº utente, consultar o processo de prescrições de um beneficiário, consultar os produtos prescritos para um beneficiário.

As funcionalidades base acima referenciadas, são as típicas de um sistema de informação, existindo no entanto pequenas particularidades inerentes ao processo de negócio que passo a destacar.

1. Identificação de beneficiário ao tratar um processo/prescrição.
2. Registrar um processo.
3. Alterar um processo.

1 - Identificação de beneficiário ao tratar um processo/prescrição

A identificação de um beneficiário é efetuada pelo campo NISS, no entanto existe a possibilidade de tratar um processo/prescrição através do nº utente no serviço nacional de saúde, possível devido a ligação de SAPA com o sistema da ACSS.

2 – Registrar um processo SAPA

Esta funcionalidade engloba o registar de processos e da prescrição dos produtos de apoio, tendo em conta:

- O subsistema a que o utilizador está associado.
- As características do utilizador.

Em suma, ao ser efetuado um registo e de acordo com o utilizador que efetuou login no sistema, é atribuído de forma automaticamente a entidade SAPA a que o utilizador pertence, assim como a respetiva entidade financiadora associada ao processo. Da mesma forma é automaticamente identificado e atribuído ao campo que identifica o prescriptor, os dados de quem tem a sessão aberta.

CONFIDENCIAL

Figura 2-2 Automatismo preenchimento campos em SAPA

3 – Alterar processo/prescrição

Esta funcionalidade permite ao utilizador alterar dados do processo e da prescrição, de acordo com o seu subsistema e a sua área de responsabilidade, sendo automaticamente identificado e atribuído ao campo que identifica o prescriptor e ao campo inerente à entidade responsável, com os dados de quem tem a sessão aberta.

Capítulo 3

Análise dos requisitos estatísticos

É feita uma abordagem com objetivo de dar a conhecer os requisitos solicitados pela entidade gestora do projeto, o INR. É abordado um facto fundamental para a construção do DW, a granularidade pretendida para a visão histórica dos indicadores.

Por fim, tendo em conta o modelo de dados operacional versus requisitos, é dado a conhecer o estudo efetuado e orientado para a perceção das escolhas e razões na definição da arquitetura do DW.

3.1. Perceção dos requisitos estatísticos

Os indicadores estatísticos que o INR pretende ter como output final são “alimentados” pelo grande repositório de dados, materializado na criação de um *Data Warehouse*, pelo que a sua conceção visa organizar os dados de forma a facilitar e acelerar a produção dos relatórios estatísticos pela ferramenta de BI.

Os requisitos foram disponibilizados pela entidade gestora do projeto, o INR, num documento que destaca os relatórios/informação que pretendiam ter como output final, documento que pode ser consultado no Anexo III – Requisitos do INR, informação estatística. Este que foi a base para desenvolvimento do DW e assunto foco, de esclarecimentos contínuos nas reuniões.

Sintetizo, com o objetivo de identificar e caracterizar de forma genérica, quais os dados a serem disponibilizados e tratados para fins estatísticos, e que caracterizam assim a definição da arquitetura do DW:

- Informações dos beneficiários SAPA.
- Dados dos processos de prescrição através dos números de identificação da segurança social (NISS), nº utente do serviço nacional de saúde ou outro que se ache necessário.
- Histórico de prescrições de produtos de apoio por beneficiário.
- Informações relativas às prescrições de produtos de apoio, por todas as entidades envolvidas.
- O controlo da execução orçamental por distrito, relativa à Segurança Social.
- Informação de prescrições de produtos de apoio concedidos a beneficiários, pelo subsistema financiador do Ministério da Saúde para a ACSS.
- Informação dos diversos subsistemas/organismos financiadores.
- Informação sobre os produtos de apoio/beneficiários caracterizando as várias vertentes enquadradas na classificação internacional de deficiência (CID).
- Informação das prescrições por subsistema.

O desenvolvimento e definição da arquitetura do DW teve sempre como foco e foi orientado, pelos requisitos estatísticos solicitados pelo INR e pelo “porque?”, quais as

razões da necessidade de tal informação, respostas consolidadas e assunto foco nas reuniões e nas trocas de msg com o grupo de trabalho responsável pelo acompanhamento e desenvolvimento do INR, assim como no estudo exposto no capítulo 3.

Fazer igualmente o controlo de duplicação de atribuição de produtos de apoio, a gestão de reembolsos/pagamentos de prescrições a serem financiadas pelo Instituto da Segurança Social e disponibilizar ainda um conjunto vasto de informação que auxiliara numa gestão eficaz da atribuição dos produtos de apoio mediante, designadamente, a simplificação de procedimentos exigidos pelas entidades, respetivo financiamento simplificado dos produtos de apoio e o controlo de duplicação de atribuição de produtos de apoio, garantindo uma melhor gestão de pagamentos de prescrições a serem financiadas.

A compreensão e estudo do acima exposto, permite identificar e traduzir de forma muito mais clara os processos de negócio do SAPA, os requisitos base, e como tal perceber as opções tomadas na modelagem do DW.

3.2. Bus Matrix

O conjunto de requisitos implementados, consistem em indicadores e contagens a elaborar sobre os dados do SAPA, outputs pretendidos em termos de informação a gerar a partir dos dados operacionais fruto dos processos de negócio. A criação de um esquema matricial, vulgo *bus matrix* (Sapateiro, 2016), traduz os principais processos de negócio e a forma com estes cruzam com as propriedades características do seu funcionamento (dimensões), permite identificar pontos comuns e é um valioso auxílio na perceção da definição dos *data marts* que foram desenvolvidos, possibilitando inclusive ter a perceção da razão da arquitetura implementada.

Analisando a tabela abaixo, representativa do *bus matrix*, é possível afirmar que os processos de negócio podem ser separados em quatro áreas distintas, processos, prescrições, produtos e utilizadores. Existe uma forte relação entre estas, ao ponto de haver áreas que tem dimensões associadas, que por sua vez também são identificadas claramente nos processos de negócio.

Tabela 3-1 *Bus Matrix* SAPA

DIMENSÕES \ PROCESSOS NEGOCIO	DIMENSÕES											
	Data	Localização	Beneficiário	Utilizador	Entidade	Processo	Produto de apoio	Evolução/estado	Organismo(s)	Classificação ISO	CIF	Incapacidade
Criar processo sapa	x	x	x	x	x			x	x			x
Prescrever produto apoio	x		x	x	x	x	x	x	x	x	x	x
Criar produto apoio	x			x			x		x	x	x	
Criar utilizador SAPA	x			x	x				x			
Criar entidade SAPA	x	x			x				x			

Exemplificando, o processo de negócio, prescrição de um produto de apoio tem como dimensões associadas o processo, o produto de apoio e a entidade. O processo de negócio, criar processo SAPA tem com dimensão associada a entidade.

Ao nível das dimensões destaco o utilizador a entidade e o organismo como sendo as que tem uma frequência mais visível. A informação identificada como necessária, e respetiva granularidade que caracteriza o DW, não é totalmente espelhada na tabela *bus matrix*, pois a totalidade das dimensões que são criadas não são diretamente identificadas, mas estão sem dúvida relacionadas com as características associadas às nomenclaturas usadas na identificação das dimensões na *bus matrix*.

3.3. Granularidade temporal

Um aspecto que é de todo pertinente, e que foi debatido e definido nas reuniões com o INR, qual a granularidade pretendida para a visão histórica dos indicadores estatísticos, traduzida na dimensão TEMPO. Ficou estabelecido que pretendiam informação unicamente ao nível do ANO.

Face à existência de mais de um campo do tipo data em determinadas tabelas do modelo operacional, foi também alvo de atenção e definido quais os campos a serem usadas na obtenção dos indicadores estatísticos referentes aos utilizadores SAPA, processos, beneficiários, entidades prescritoras e prescrições efetuadas. Ficou estabelecido que as datas a ter em consideração, são data de criação e data do estado,

que consoante o tipo de relatório, aplica-se da seguinte forma:

- Data de criação da(o):
 - Prescrição, para obter o número de utilizadores que prescreveram. O número de entidades prescritoras existentes que prescreveram, por região, concelho, distrito, entidade financiadora.
 - Do processo, para obter o número de processos e beneficiários por estado, entidade financiadora por ano e entidade prescritora por ano.
 - Da entidade, para obter o número de entidades prescritoras existentes, por região, concelho, distrito, entidade financiadora.

- Data do estado da:
 - Prescrição, para obter o número de prescrições por entidade prescritora, estado e entidade financiadora.
 - Processo para obter o número de beneficiários com processos por estado, entidade prescritora e entidade financiadora.
 - Processo para obter o número de beneficiários face ao sexo, escalão etário, distrito, conselho, ano de escolaridade, situação face ao emprego, finalidade de ajuda técnica e beneficiários com incapacidade multiuso.

Com influência na definição da arquitetura do DW, principalmente no que concerne a tabelas de factos, destaco o esclarecimento sobre quais os indicadores estatísticos que se definiu ter uma visão histórica anual, em suma, que fica-se registado todas as alterações e mudanças dos valores/estado face aos diferentes anos, foram estes:

- Dados sobre as entidades sapa.
- Dados sobre as prescrições efetuadas.
- Dados sobre os utilizadores sapa.

Saliento ainda assim, que o DW foi desenvolvido, de forma a facilmente se conseguir alterar esta decisão e abranger no futuro a possibilidade de ter visão histórica anual de outras vertentes identificadas como sendo pertinentes na criação dos

indicadores estatísticos, nomeadamente sobre:

- Dados sobre os processos.
- Dados sobre os produtos de apoio.
- Dados sobre classificação internacional de funcionalidade, incapacidade e saúde.

3.4. Estudo direcionado para responder aos *reports*

Esta ação foi fundamental para efetuar um mapeamento, definir e identificar os modelos e as dimensões a serem criadas na construção do DW. Ação que vem na sequência do estudo exposto nos anteriores subcapítulos, permitindo cimentar o conhecimento do modelo de dados SAPA.

Este consistiu em efetuar *querys* à base de dados SAPA, com o objetivo de ter como outputs a informação estatística solicitada pelo INR, código transcrito no Anexo II – Estudo direcionado para responder aos requisitos do INR.

Destaco desta ação, aquando do desenvolvimento do código SQL, dúvidas que surgiram e foram posteriormente esclarecidas e definidas em reuniões. A sua maior parte devia-se a omissão no documento que foi entregue pelo INR (Anexo III – Requisitos do INR – Informação estatística.) das características/valores dos atributos que definiam os requisitos de informação estatística. Saliento duas das dúvidas que considero serem as mais importantes:

- A ausência de definição dos períodos temporais a que os *reports* deviam ter em consideração, e quais das colunas das tabelas com valores de datas, ter em conta para a definição do período.
- A existência de mais de um processo SAPA, por beneficiário, quando tinha sido transmitido que existiria somente um beneficiário por processo.

Face a esta realidade, os indicadores estatísticos solicitados pelo INR, referentes a processos passaram a ter novos requisitos, sobre os respetivos beneficiários. Para a totalidade dos relatórios inerentes aos processos é feita a contabilização dos totais de beneficiários a que correspondem os processos.

Capítulo 4

Análise do modelo de dados SAPA

Dá a conhecer toda a informação sobre o modelo de dados que suporta o sistema operacional do SAPA, matéria alvo de profunda análise face ao impacto e influencia que tem na modelagem e desenvolvimento do DW.

Num primeiro subcapítulo é identificado com que os outros sistemas e subsistemas, interage o sistema SAPA.

É dado a conhecer características das tabelas, os seus campos, tipos de valor e respetivas associações e por fim é feito um estudo do modelo entidade e relação, que permite aprofundar o conhecimento da estrutura logica da base de dados.

4.1. Interface com outros sistemas

Um dos objetivos principais, aquando do desenvolvimento da aplicação SAPA, foi obter uma solução integrada e centralizada para suporte à gestão de atribuição de produtos de apoio, pelo que o sistema contemplar interações com alguns dos subsistemas internos da Segurança Social através do SISS e com os sistemas externos ao SISS, que com ele interagem.

Interface com o IEFEP, sistema financiador de produtos de apoio aos Beneficiários, sendo responsável pela receção e envio da informação do IEFEP, estando em sincronismo com o sistema SAPA.

Interface com sistemas de informação do Ministério da Saúde, para permitir a validação do número de utente registado em SAPA, com o número de utente existente no Registo Nacional de Utentes (RNU) – ACSS e obter dados de identificação do sistema RNU, que permitam a identificação dos beneficiários no SISS. Por sua vez o SAPA fornece à ACSS, a informação relativa aos produtos de apoio atribuídos pelo Ministério da Saúde, enquanto financiador de produtos de apoio aos beneficiários.

Interface com seis subsistemas, do sistema de informação da Segurança Social, nomeadamente;

CONFIDENCIAL

A figura seguinte identifica as principais interações do sistema SAPA, com os outros sistemas e subsistemas.

CONFIDENCIAL

Figura 4-1 Interações SAPA, com outros sistemas/subsistemas

De todos os subsistemas e sistemas acima referenciados, cabe destacar dois, pois serão os únicos modelos de dados, além do SAPA, em que é necessário beber informação das suas tabelas para a construção do DW.

CONFIDENCIAL

4.2. Análise das tabelas operacionais

Pretende-se com este subcapítulo dar a conhecer toda a informação, sobre as tabelas da base de dados que suportam o sistema operacional do SAPA, matéria considerada relevante para o desenvolvimento do DW e como tal, alvo de profunda análise.

São ■ as tabelas que compõem a base de dados SAPA:

CONFIDENCIAL

É possível consultar ao pormenor no Anexo I - Tabelas do sistema operacional SAPA, para cada uma das tabelas informação sobre

- Nome da tabela.
- Descrição, comentários e notas.
- Total de registos à data do presente estudo (Janeiro2017).
- Para cada uma das colunas da Tabela:
 - Nome da coluna.
 - Tipo de dados.

- Informação se permite valores nulos.
- *Constraint*.
- Descrição da tabela no sistema operacional.
- Comentários e notas.

Construiu-se uma tabela com a identificação das tabelas que compõem a base de dados SAPA, destacando as associações/ligações, conseguida através das chaves primárias e das chaves estrangeiras e assim ter uma perspetiva visual das associações

Tabela 4-1 Associações entre as tabelas SAPA

CONFIDENCIAL

No seguimento da análise das tabelas operacionais constatou-se que existem

tabelas com 0 registros. Foi apurado que a sua existência tem como finalidade, responder a uma segunda fase do desenvolvimento da aplicação SAPA, que vai contemplar dados contábilístico/financeiros. Tabelas essas que são:

CONFIDENCIAL

Existem três tabelas que se destacam, face à sua pertinência no processo de negócio e relevância no desenvolvimento do DW; PROCESSO_SAPA ,PRODUTO_APOIO e PRESCRICAO

4.2.1. Tabela PROCESSO_SAPA

Reflete as propriedades do processo de atribuição de produtos de apoio aos beneficiários, identificando os beneficiários e respectivas deficiências ou incapacidades permanentes ou temporárias. à data do presente estudo a tabela tinha cerca de [REDACTED] registros.

A tabela tem cerca de [REDACTED], dos quais um é a chave primária, [REDACTED], e como chaves estrangeiras tem dois campos, [REDACTED] [REDACTED].

Esta tabela é uma das mais relevantes do SAPA, facilmente constatável pois a sua chave primária, [REDACTED], é chave estrangeira, em cerca [REDACTED] tabelas do modelo operacional, sendo o campo que existe em mais tabelas.

Nesta consta toda a informação pessoal que identifica o beneficiário a quem pertence o processo, o seu número de identificação da segurança social, nome, morada, contactos e respectivas deficiências ou incapacidades. Realço que contem igualmente informação sobre o estado do processo, permitindo assim identificar em que fase de este

se encontra.

Na tabela abaixo é possível consultar a totalidade dos campos que compõem o PROCESSO_SAPA, o seu tipo de valor e uma breve descrição sobre a sua caracterização.

Tabela 4-2 Campos da tabela PROCESSO_SAPA

CONFIDENCIAL

CONFIDENCIAL

4.2.2. Tabela PRODUTO_APOIO

Tabela que reflete as propriedades da lista de produtos homologada pela ISO 9999:2002 - classificação internacional de ajudas técnicas, que constam nas prescrições. à data do presente estudo a tabela tinha cerca de ■ registros.

A tabela tem cerca de ■ campos, dos quais o CODIGO_ISO é a chave primária. Realço os campos ■, que além de permitirem identificar a classe e subclasse de um produto de apoio, são parte integrante da chave primária, pelo que se desagregarmos o CODIGO_ISO, obtemos essa informação. De forma a ser mais explícito, abaixo exemplifico, para o produto com a designação, “Cadeiras de rodas elétricas com comando de direção elétrico” e com:

CODIGO_ISO : 122306

CODIGO_CLASSE_PRODUTO : 12

CODIGO_SUBCLASSE_PRODUTO : 1223

CONFIDENCIAL

Nesta consta toda a informação da lista de produtos de apoio disponíveis e devidamente homologada pela ISO 9999:2002, sendo possível identificar a que classe e subclasse os produtos de apoio pertencem, assim como se este é possível de ser reutilizado e quais as entidades que podem prescrever o mesmo

Na tabela abaixo é possível consultar a totalidade dos campos que compõem o PRODUTO_APOIO , o seu tipo de valor e uma breve descrição sobre a sua caracterização.

Tabela 4-3 Campos da tabela PRODUTO_APOIO

CONFIDENCIAL

4.2.3. Tabela PRESCRICAO

Tabela que reflete as propriedades da ficha de prescrição, de teor obrigatório sempre que se procede à prescrição de um produto de apoio a um beneficiário. à data do presente estudo a tabela tinha cerca de ■■■■■ registos.

A tabela tem cerca de ■■ campos, dos quais a chave primária é

██████████ e tem uma chave estrangeiras ██████████. Um processo SAPA pode ter uma ou mais prescrições, facilmente identificadas nesta tabela.

Nesta consta, qual o motivo por que existe a necessidade de prescrever o(s) produto(s) de apoio, em que estado esta se encontra e em que data foi revisto estado da mesma.

Na tabela abaixo é possível consultar a totalidade dos campos que compõem a PRESCRICAO, o seu tipo de valor e uma breve descrição sobre a sua caracterização.

Tabela 4-4 Campos da tabela PRESCRICAO

CONFIDENCIAL

4.3. Modelo de entidade e relacionamento

O estudo do MER permitiu adquirir o conhecimento da estrutura logica da base de dados, cimentou-se a análise das regras de negócio, das entidades que compõem o SAPA e das relações que existem entre estas (*Ralph Kimball, Margy Ross,2013*).

4.3.1. Caracterização das entidades

É normal que a nomenclatura dada às tabelas do sistema operacional SAPA, identifiquem de forma direta as entidades que a compõem, a maior parte, já referenciadas ao longo do documento. Estudando a “figura 3.1-Modelo entidade e relacionamento”, facilmente se identifica as entidades e associação entre estas, destaco algumas atribuindo-lhes a respetiva classificação dadas as suas características.

As que se caracterizam por serem entidades estruturantes, ou básicas são as entidades fundamentais, como por exemplo:

CONFIDENCIAL

Entidades cuja existência apenas se justifica pela necessidade de descodificação, são as entidades descodificadoras, como por exemplo:

CONFIDENCIAL

4.3.2. Caracterização das associações

É normal que em qualquer conjunto consolidado de informação, cada entidade esteja relacionada com quase todas as outras (*Ralph Kimball, Margy Ross, 2013*). No

entanto não é boa prática na modelação contemplar de forma direta todas as associações possíveis no modelo de dados, facto constatado no modelo de dados SAPA.

CONFIDENCIAL

Figura 4-2 Modelo entidade e relacionamento

Ao invés de esquematizar o tipo de associações existentes (1:1 ; 1:N ; N:M), facilmente identificável e de leitura direta, ao se consultar a figura abaixo, opta-se por

abordar as associações existentes/particularidades destas, mas não identificadas ao se estudar o MER.

CONFIDENCIAL

Figura 4-3 Associações, chaves primárias e estrangeiras

Esta abordagem tem como intuito dar a conhecer melhor o processo de negócio e particularidades que foram sendo descobertas, à medida que se efetuou o presente estudo. São essas as associações que abaixo destaco, e que mereceram um pedido de esclarecimento junto do INR.

UTILIZADORES

Um utilizador registado na tabela UTILIZADOR_SAPA, identificado pela chave primária da tabela, está sempre associado a [REDACTED] e a uma [REDACTED]. Um utilizador também está sempre associado [REDACTED] de um

CONFIDENCIAL

PROCESSO SAPA

Face ao *core* do negócio, a tabela que PROCESSO_SAPA é sem dúvida uma das principais do modelo de dados e composta por campos que são determinantes na elaboração dos relatórios estatísticos solicitados pelo INR.

CONFIDENCIAL

Ao se efetuar o estudo prévio para desenvolver o DW, foi identificada esta situação e que mereceu uma abordagem junto do INR, e que teve como é obvio um impacto nos relatórios inicialmente solicitados, (Anexo III, Requisitos do INR – Informação estatística), passando a ter novas premissas. [REDACTED]

[REDACTED]

[REDACTED]

Capítulo 5

Desenho e criação do *Data Warehouse*

É feita uma abordagem, para cada uma das tabelas factos e de dimensões criadas. As dimensões são agrupadas em três grupos, caracterizadas de acordo com o processamento das fontes de informação, dimensões manuais, automáticas e híbridas. É dado a conhecer os oito modelos desenvolvidos, sendo feito uma caracterização destes, identificando-se a sua granularidades, tabelas fonte e indicadores que se propõem a responder.

Saliento que foram utilizadas na conceção e desenho do DW, as ferramentas *Toad for Oracle* versão 12.1 e o *Platinum erwin* versão 3.5.2. .

5.1. Modelos do *Data Warehouse*

Os modelos desenvolvidos caracterizam-se por possuir um conjunto de indicadores e dimensões de análise logicamente relacionados entre si num determinado contexto do negócio SAPA.

Os modelos respondem de forma direta ao necessário para alimentar a ferramenta de BI, ou seja um modelo apenas conterá informação necessária para determinar o que os utilizadores do INR, realmente necessitam obter sem ter que recorrer a diferentes análises (*Jiawei Han, Micheline Kamber,2006*).

Para cada modelo é feita uma caracterização (*Jorge Sá, 2009*), identificando-se a sua granularidades, tabelas fonte e indicadores que se propõe a responder, são oito os modelos;

- Modelo processos.
- Modelo utilizadores.
- Modelo prescrições.
- Modelo produtos de apoio.
- Modelo entidades.
- Modelo fator ambiental.
- Modelo deficiências.
- Modelo atividades em que participa.

5.1.1. *Modelo processos*

Este modelo tem por objetivo observar indicadores sobre os processos SAPA,

CONFIDENCIAL

CONFIDENCIAL

Figura 5-1 Esquema do modelo Processos

Este modelo responde a indicadores estatísticos que se identificou não haver a necessidade de ter uma “visão histórica”, que realce alterações ao longo do tempo, pelo que é sempre tido em conta os dados que vigoram de igual modo, no modelo operacional e refletido na tabela principal deste, a [REDACTED].

Face à adoção do esquema em “flocos de neve”, a escolha das dimensões do modelo processos, reflete-se em todas as ligações com as chaves estrangeiras que irão estar presentes na tabela representativa deste, a tabela DIM_SAPA_PROCESSO. Cobre assim, todas as necessidades dos relatórios a serem efetuados, abaixo numerados;

CONFIDENCIAL

A numeração acima atribuída, tem como objetivo, permitir a interpretação da tabela abaixo, “Dimensões do modelo processos “, correspondendo à numeração das suas colunas, necessidades respondidas.

Tabela 5-1 Dimensões do modelo processos

CONFIDENCIAL

5.1.2. Modelo utilizadores

Este modelo tem por objetivo observar indicadores sobre os utilizadores SAPA, nomeadamente contagens, existirá um registo unívoco para cada conjunto [REDACTED]

A definição do presente modelo teve em consideração a necessidade de responder a relatórios com indicadores e contagens sobre [REDACTED]

CONFIDENCIAL

CONFIDENCIAL

Figura 5-2 Esquema do modelo Utilizadores

Este modelo responde a indicadores estatísticos que se identificou ser necessário ter uma visão histórica anual, em suma, [REDACTED]

[REDACTED]

CONFIDENCIAL

A numeração acima atribuída, tem como objetivo, permitir a interpretação da tabela abaixo, “Dimensões do modelo utilizadores “, correspondendo à numeração das suas colunas, necessidades respondidas.

Tabela 5-2 Dimensões do modelo utilizadores

CONFIDENCIAL

5.1.3. Modelo prescrições

Este modelo tem por objetivo observar indicadores sobre as prescrições efetuadas, nomeadamente contagens, existirá um registo unívoco para cada conjunto [REDACTED]

CONFIDENCIAL

Figura 5-3...Esquema do modelo Prescrições

A definição do presente modelo teve em consideração a necessidade de responder a relatórios com indicadores e contagens sobre o [REDACTED]

Face à adoção do esquema em “flocos de neve”, a escolha das dimensões do modelo prescrições, reflete-se em todas as ligações com as chaves estrangeiras que irão estar presentes na tabela representativa deste, [REDACTED]

CONFIDENCIAL

A numeração acima atribuída, tem como objetivo, permitir a interpretação da tabela abaixo, “Dimensões do modelo prescrições“, correspondendo à numeração das suas colunas, necessidades respondidas.

Tabela 5-3 Dimensões do modelo prescrições

CONFIDENCIAL

5.1.4. Modelo produtos de apoio

Este modelo tem por objetivo observar indicadores sobre os produtos de apoio que

foram prescritos, nomeadamente contagens, existirá [REDACTED]

CONFIDENCIAL

Figura 5-4 Esquema do modelo Produtos de Apoio

A definição do presente modelo teve em consideração a necessidade de responder a relatórios com indicadores e contagens sobre [REDACTED]

CONFIDENCIAL

CONFIDENCIAL

A numeração acima atribuída, tem como objetivo, permitir a interpretação da tabela abaixo, “Dimensões do modelo produtos de apoio “, correspondendo à numeração das suas colunas, necessidades respondidas.

Tabela 5-4 Dimensões do modelo produtos de apoio

CONFIDENCIAL

5.1.5. Modelo entidades

Este modelo tem por objetivo observar indicadores sobre as entidades registradas em SAPA, nomeadamente contagens, existirá um [REDACTED]

A definição do presente modelo teve em consideração a necessidade de responder a relatórios com indicadores e contagens sobre [REDACTED]

CONFIDENCIAL

Figura 5-5 Esquema do modelo Entidades

CONFIDENCIAL

CONFIDENCIAL

Figura 5-6 Esquema do modelo Fator ambiental

A definição do presente modelo teve em consideração a necessidade de responder a relatórios com indicadores e contagens sobre a caracterização dos [REDACTED]

CONFIDENCIAL

Tabela 5-6 Dimensões do modelo fator ambiental

CONFIDENCIAL

5.1.7. Modelo deficiências

Este modelo tem por objetivo observar indicadores sobre a CIF, particularizando as deficiências dos beneficiários, [REDACTED]

[REDACTED]

[REDACTED]

CONFIDENCIAL

CONFIDENCIAL

Figura 5-7 Esquema do modelo deficiências

Este modelo responde a indicadores estatísticos que se identificou não haver a necessidade de ter uma “visão histórica”, que realce alterações ao longo do tempo, pelo que é sempre tido em conta os dados que [REDACTED]

CONFIDENCIAL

Tabela 5-7 Dimensões do modelo deficiências

CONFIDENCIAL

5.1.8. Modelo atividades em que participa

Este modelo tem por objetivo observar indicadores sobre a CIF, caracterizando as capacidades e desempenho inerentes às deficiências dos beneficiários, nomeadamente contagens, [REDACTED]

[REDACTED]

CONFIDENCIAL

CONFIDENCIAL

Figura 5-8 Esquema do modelo atividades em que participa

CONFIDENCIAL

Tabela 5-8 Dimensões do modelo atividades em que participa

CONFIDENCIAL

5.2. Tabelas de Dimensões

O objetivo deste subcapítulo é dar a conhecer as dimensões sobre as quais é possível efetuar uma análise, sendo estas usadas nos vários modelos definidos no DW SAPA.

Cada dimensão poderá ser caracterizada de acordo com o processamento das fontes de informação, nomeadamente:

- Manual, os dados necessários para o processamento da dimensão estão contidos no processo de carregamento, pois não possuem tabela fonte no sistema operacional SAPA. Estas são carregadas através de processos semelhantes às outras, com a diferença que a fonte dos dados está codificada manualmente no processo de DataStage. Deste modo, qualquer alteração aos valores destas dimensões será efetuado ao nível do processo de DataStage.
- Automática, os dados necessários para o processamento da dimensão estão contidos nas tabelas da base de dados que suportam [REDACTED], [REDACTED], pertencente ao sistema de informação [REDACTED]

- Híbrida, combinação dos dois tipos anteriores, onde existe uma tabela fonte no sistema operacional SAPA, mais informação criada e mantida manualmente.

Abordo desde já, algo que facilmente se constata na figura seguinte, visão global das dimensões, existe um número elevado de relacionamento entre dimensões, traduzido nos modelos desenvolvidos, [REDACTED], assunto que é ainda mais clarividente nos subcapítulos que aborda as dimensões automáticas, subcapítulos 5.2.3 e 6.4.3

Alvo de abordagem mais à frente neste capítulo, a figura permite ter uma visão global e perceção da relação entre as dimensões, sendo possível identificar hierarquias e esquemas [REDACTED] que caracterizam a criação de determinadas dimensões.

CONFIDENCIAL

CONFIDENCIAL

Figura 5-9 Visão global das dimensões

5.2.1. Dimensões comuns a outros projetos

Existe um conjunto de dimensões, identificadas à posteriori na descrição das dimensões e modelos SAPA, que derivam de outros projetos já existente no [REDACTED], sendo estas utilizadas no DW SAPA, realço, sem prejuízo para as necessidades do presente projeto. São estas, as dimensões:

- Tempo
- Geografia
- Sexo
- Idade

Reforço que estas dimensões já existem e não foram alvo de desenvolvimento, mas como estão presentes nos modelos SAPA é feita uma pequena abordagem explicativa sobre as mesmas.

DIM TEMPO

Esta dimensão é utilizada para permitir a visão histórica dos indicadores estatísticos, sendo constituída pelos níveis ano, semestre, trimestre, mês e dia. O tipo de processamento desta dimensão é manual.

DIM IDQ IDADE

Esta dimensão permite fazer a associação da idade das Pessoas Singulares ao seu Escalão Etário. Esta informação é calculada para vários períodos de análise, sendo que em cada período a idade corresponde à idade que a PS tem no fim desse período. O tipo de processamento desta dimensão é manual.

DIM GEOGRAFIA

Esta dimensão descreve a localização física de um conceito geográfico, com uma hierarquia que abrange deste a identificação do continente até à freguesia do local. O tipo de processamento desta dimensão é híbrido.

DIM IDQ SEXO

Esta dimensão permite Identificar o sexo da pessoa singular, o tipo de processamento desta dimensão é manual.

████████████████████

Esta dimensão caracteriza as pessoas ██████████ registrada ██████████. Existe um registo para ██████████ o seu processamento ██████████.

5.2.2. Dimensões manuais

Face à inexistência de uma tabela fonte no sistema operacional SAPA, foram criadas nove dimensões manuais, nomeadamente:



É feita uma apresentação, para cada uma das dimensões acima referenciadas abordando, o objetivo da sua existência, chaves primárias, campos, hierarquia, exemplos reais de valores que podem os campos ter como registo e outras considerações que se determinou pertinentes. A informação referenciada, face à sua dimensão é remetida para o anexo IV - Tabelas do *Data Warehouse*, tendo-se optado em apresentar de seguida, uma pequena abordagem explicativa sobre as mesmas.

CONFIDENTIAL

CONFIDENCIAL

5.2.3. Dimensões automáticas

Os dados necessários para o processamento das dimensões aqui referenciadas, estão contidos nas tabelas da base de dados que suportam o sistema operacional SAPA ou no [REDACTED], pertencente ao sistema de informação [REDACTED], foram criadas vinte e oito dimensões automáticas, nomeadamente:

CONFIDENCIAL

É feita uma apresentação, para cada uma das dimensões acima referenciadas abordando, o objetivo da sua existência, chaves primárias, campos, hierarquia, exemplos reais de valores que podem os campos ter como registo e outras considerações que se determinou pertinentes. A informação referenciada, face à sua dimensão é remetida para o anexo IV - Tabelas do *Data Warehouse*, tendo-se optado em apresentar de seguida, uma pequena abordagem explicativa sobre as mesmas.

CONFIDENCIAL

CONFIDENTIAL

CONFIDENTIAL

CONFIDENTIAL

CONFIDENCIAL

CONFIDENTIAL

5.2.4. Dimensões híbridas

Os dados necessários para o processamento das dimensões aqui referenciadas, estão contidos nas tabelas da base de dados que suportam o sistema operacional SAPA ou nos subsistemas ██████████, pertencente ao ██████, em conjunto com informação criada e mantida manualmente. Foram criadas oito dimensões híbridas, nomeadamente:



É feita uma apresentação, para cada uma das dimensões acima referenciadas abordando, o objetivo da sua existência, chaves primárias, campos, hierarquia, exemplos reais de valores que podem os campos ter como registo e outras considerações que se determinou pertinentes. A informação referenciada, face à sua dimensão é remetida para o anexo IV - Tabelas do *Data Warehouse*, tendo-se optado em apresentar de seguida, uma pequena abordagem explicativa sobre as mesmas.

CONFIDENTIAL

CONFIDENTIAL

CONFIDENTIAL

CONFIDENTIAL

CONFIDENCIAL

5.3. Tabelas de Factos

Os indicadores estatísticos que o INR solicitou necessitar de uma [REDACTED], são a principal razão da criação das três tabelas de factos abaixo referenciadas:

- FAC_SAPA_ENTIDADE_SAPA
- FAC_SAPA_PRESCRICAO
- FAC_SAPA_UTILIZADOR

É feita uma apresentação, para cada uma das dimensões acima referenciadas abordando, o objetivo da sua existência, chaves primárias, campos, hierarquia, exemplos reais de valores que podem os campos ter como registo e outras considerações que se determinou pertinentes. A informação sobre tabelas no DW, colunas e tipo de dados e tamanho das tabelas factuais podem ser consultadas no anexo IV - Tabelas do *Data Warehouse*.

5.3.1. FAC_SAPA_ENTIDADE_SAPA

A presente tabela factual, permite ter uma visão histórica das [REDACTED], com indicadores estatísticos ao nível [REDACTED]. Tem um registo unívoco para cada conjunto dos [REDACTED] que traduz assim a sua granularidade, nomeadamente:

CONFIDENCIAL

A métrica traduzida nesta tabela de factos é registada na [REDACTED].

CONFIDENCIAL

Tendo em perspectiva os relatórios que se pretendia efetuar e para simplificar o acesso aos dados, foram ainda criados campos obtidos através das dimensões que ligam à tabela de factos, nomeadamente:

CONFIDENCIAL

Os exemplos de alguns valores possíveis atualmente, são os abaixo apresentados.

Tabela 5-9 Valores da FAC_SAPA_ENTIDADE_SAPA

CONFIDENCIAL

5.3.2. FAC_SAPA_PRESCRICAO

A presente tabela factual, permite ter uma visão histórica das prescrições efetuadas

em SAPA, com indicadores estatísticos [REDACTED]. Tem um registo unívoco para [REDACTED] que constituem a sua chave primária, que traduz assim a sua granularidade, nomeadamente:

CONFIDENCIAL

As métricas traduzidas nesta tabela de factos são registadas nas colunas

CONFIDENCIAL

Tendo em perspetiva os relatórios que se pretendia efetuar e para simplificar o acesso aos dados, foram ainda criados campos obtidos através das dimensões que ligam à tabela de factos, nomeadamente:

Os exemplos de alguns valores possíveis atualmente, são os abaixo apresentados.

Tabela 5-10 Valores da FAC_SAPA_PRESCRICAO

CONFIDENCIAL

5.3.3. FAC_SAPA_UTILIZADOR

A presente tabela factual, permite ter uma visão histórica dos utilizadores que tiveram parte ativa [REDACTED] e/ou estejam registados em SAPA, com indicadores [REDACTED]. Tem um registo unívoco para cada conjunto [REDACTED] granularidade, nomeadamente:

CONFIDENCIAL

A forma como as ligações da tabela factual são concretizadas com as dimensões

CONFIDENCIAL

Tendo em perspetiva os relatórios que se pretendia efetuar e para simplificar o acesso aos dados, foram ainda criados campos obtidos através das dimensões que ligam à tabela de factos, nomeadamente:

Os exemplos de alguns valores possíveis atualmente, são os abaixo apresentados.

Tabela 5-11 Valores da FAC_SAPA_UTILIZADOR

CONFIDENCIAL

Capítulo 6

Cadeia ETL

Inicialmente é feita uma breve referência à tecnologia usada para a extração, transformação e carregamento dos dados para o *Data Warehouse SAPA*, o *Infosphere Datastage and Infosphere Qualitystage*, versão 9.1 da IBM.

De seguida é feita uma caracterização geral dos processos ETL, dando a conhecer as nomenclaturas adotadas, categorias e características comuns a toda cadeia ETL. É igualmente feita uma abordagem sobre os três tipos de Jobs que compõem a cadeia de SAPA.

É dado a conhecer todos os processos ETL, separados de acordo com as características do tipo de tabela que se carrega, data sets, tabelas de dimensão e tabelas de factos.

6.1. Tecnologia usada

Para efetuar o carregamento e processamento dos dados para o *Data Warehouse* SAPA, foram desenvolvidos um conjunto de processos e jobs. A tecnologia usada para o efeito é da IBM, o *Infosphere Datastage and Infosphere Qualitystage*, versão 9.1,(doravante designada de ferramenta ETL) [REDACTED] a escolha para o desenvolvimento do projeto SAPA (*IBM Knowledge Center*).



Figura 6-1 Login com a ferramenta ETL usada

É uma das ferramentas ETL conhecidas do mercado, permitindo a integração sólida das informações corporativas, suportando operações sobre grandes quantidades de dados, permitindo efetuar com grande eficiência e eficácia a extração, transformação, carregamento e validação da qualidade dos dados.

Caracteriza-se por ter uma estrutura gráfica, facto que auxilia no desenvolvimento dos processos, destaco os seguintes componentes desta ferramenta ETL, com uma pequena nota explicativa:

- *DataStage Administrator*, componente onde se adiciona e configura os novos projetos e são definidas as políticas de utilização.
- *DataStage Designer*, componente onde, através de visualização gráfica se cria e se efetua a manutenção dos processos ETL assim como a ligações entre estes.

- *DataStage Director*, Componente que permite a execução validação e monitoramento dos processos ETL compilados pelo *DataStage Designer*. Possibilita o agendamento e execução dos jobs e permite identificar erros com facilidade, pois cada job cria um log.

6.2. Caracterização dos processos ETL

Os processos existentes na cadeia SAPA, caracterizam-se pela sua função, facilmente identificada pelas convenções de nomes que foi adotado de forma a tornar óbvia a sua função. Os processamentos que são efetuados ao nível dos dados, também são assim identificados facilmente nos correspondentes *WorkJobs*.

Os processos no projeto encontram-se divididos em duas categorias, consoante a sua funcionalidade e refletido na convenção adotada no nome, nomeadamente:

- DIMS, corresponde ao carregamento das tabelas de dimensão caracterizadas de acordo com o processamento das fontes de informação, como automáticas ou manuais.
- FACS, corresponde ao carregamento das tabelas de factos, de tabelas auxiliares ao nível do DW (*Data Sets*) e tabelas dimensão caracterizadas de acordo com o processamento das fontes de informação, como sendo híbridas.

A adoção de uma nomenclatura, tem como objetivo tornar mais transparente a função de um processo. Estabeleceu-se um conjunto de convenções para a definição do nome do processo, alias todos os nomes dos Jobs e *WorkJobs* de *DataStage* possuem tipicamente dois formatos, consoante os tipo de categoria em cima referenciados.

Se corresponde a categoria DIMS;

CONFIDENCIAL

CONFIDENCIAL

Figura 6-2 Processos ETL da categoria DIMS

Ao nível da categoria DIMS, foram criados cerca de trinta e seis processos ETL, facilmente identificados na figura acima

Se corresponde a categoria FACS;

CONFIDENCIAL

CONFIDENCIAL

Figura 6-3 Processos ETL da categoria FACS

Ao nível da categoria FACS, foram criados cerca de quinze processos ETL, facilmente identificados na figura acima.

6.2.1. Parâmetros

Cada um dos processos desenvolvidos na cadeia do SAPA possui vários parâmetros, definidos de uma forma centralizada, num conjunto de ficheiros que residem [REDACTED], numa diretoria criada para o projeto de SAPA [REDACTED]. Sempre que se cria um processo com a ferramenta ETL, os parâmetros necessários ao processo são definidos, e guardados [REDACTED] [REDACTED] seus valores comuns a todos os processos. As diferenças que destaco e que serão posteriormente alvo de referência, consoante o processo que se esta a desenvolver são os parâmetros, [REDACTED] [REDACTED]

Na figura a seguir, o *layout* da configuração dos parâmetros do processo ETL da tabela do DW, [REDACTED], exemplo explicativo do frisado. A edição na coluna “*Default Value*”, e consoante o tipo de processo que foi alvo de criação/edição, os parâmetros que ficam disponíveis para edição podem diferir. Facilmente se constata que os nomes dos parâmetros são elucidativos quanto à sua funcionalidade.

CONFIDENCIAL

Figura 6-4 Definição dos parâmetros do processo ETL

6.2.2. ETL

O diagrama do processo ETL abaixo representado é similar para todos os processos das categorias DIM's, sendo o utilizado com mais frequência, inerente à grande quantidade de tabelas de dimensão automáticas que compõem o SAPA DW. Todos os processos que não tenham esta estrutura, serão alvo de abordagem e identificadas posteriormente.

A nomenclatura usada na definição do diagrama ETL, identifica:



Figura 6-5 Diagrama geral do processo ETL

6.2.3. Extração

Destaco na fase de extração e que foi alvo de atenção, os scripts SQL efetuados sobre as tabelas fonte do sistema operacional, refletem uma característica comum a todos os processos. Os scripts dos processos de extração encontram-se disponíveis no anexo V - Scripts da extração dos processos ETL.

Abaixo imagem exemplo do local na ferramenta ETL, onde é programado o código SQL.

CONFIDENCIAL

Figura 6-6 Programação do SQL, no processo de extração

6.2.4. Transformação

O processo de transformação é similar para todos os processos das categorias DIM's, Todos os processos que não tenham a estrutura retratada na figura a seguir, serão alvo de abordagem e identificadas posteriormente.

CONFIDENCIAL

Figura 6-7 Processo de transformação das categorias DIM's

Na figura em cima é possível identificar no lado esquerdo, com a designação ■■■, os campos resultantes do processo da extração e do lado direito os campos finais da

transformação, com a designação [REDACTED], assim como os respectivos *match* entre estes.

Em [REDACTED], as colunas [REDACTED] são fruto de instruções sobre os parâmetros do job, com a finalidade de identificar este. Esta ação é comum a todos os processos de transformação para todas as tabelas do DW, de forma a se obter os dois campos de controlo do job, permitindo identificar este e a data em que este foi despoletado pela última vez.

6.2.5. Carregamento



Figura 6-8 Input de processo de carregamento

6.3. Jobs

Começo por salientar a extração dos dados/execução dos jobs referenciados neste capítulo e inerente cadeia SAPA, é executada [REDACTED], conforme acordado [REDACTED].

Os Jobs no projeto encontram-se divididos em três categorias, consoante a sua funcionalidade, nomeadamente:

CONFIDENCIAL

Existem três tipos de Jobs que compõem a cadeia de SAPA, os *ChainJobs*, os *RunJobs* e os *WorkJobs*. A convenção adotada para todos os nomes dos Jobs é herdada da nomenclatura atribuída nos processos ETL (subcapítulo 6.3), sendo adicionado ao início da mesma e identificados estes da seguinte forma;

6.4. Work Jobs

Os *WorkJobs*, são os *Jobs* de *DataStage*, que efetuam as operações de extração, tratamento, processamento e carregamento dos dados inerentes aos processos SAPA criados.

Ao nível da categoria DIMS, foram criados cerca de trinta e seis *work jobs*, facilmente identificados na figura abaixo.



Figura 6-9 *Work jobs* da categoria DIMS

Ao nível da categoria FACS, foram criados cerca de quinze *work jobs*, facilmente identificados na figura abaixo.



Figura 6-10 *Work jobs* da categoria FACS

No desenho das tabelas do projeto SAPA, foram identificados grupos com semelhanças substanciais entre si, nomeadamente, no formato da tabela ou no modo de carregamento. Por forma a simplificar a caracterização dos vários processos criados, são apresentadas as características, as semelhanças e as diferenças identificadas para cada tipo de processo ETL, consoante são:

CONFIDENCIAL

6.4.1. Data Sets

Foi identificado na fase de desenho e criação do DW, um grupo de informação comum e necessária para o processo ETL das tabelas de dimensão híbridas e das tabelas de factos. A solução encontrada para tornar os demais processos, mais ágeis e eficientes foi a criação de três *data sets*, de acordo com as necessidades identificadas, nomeadamente os *data sets*;

CONFIDENCIAL

Figura 6-11 Diagrama ETL do Data Set

Este grupo de *data sets* caracterizam-se por terem uma tabela fonte num dos dois sistemas operacionais, ou no [REDACTED]. Deste modo todas são alimentadas através de [REDACTED], na fase da extração. São três os processos ETL com estas características, correspondendo cada um a um data set, conforme quadro abaixo.

Tabela 6-1 Processos ETL versus Data Sets

CONFIDENCIAL

DS_ENTIDADE_INFO

Este *data set* permite caracterizar todas [REDACTED] associadas a um

CONFIDENCIAL

Figura 6-12 Processo transformação do DS_ENTIDADE_INFO

Na figura acima, o diagrama da ferramenta ETL que permite a edição e manipulação do processo de transformação, é perceptível quais as colunas que alimentam o *data set*, resultado do processo de extração (██████████) e correspondentes colunas que vão alimentar o processo de carregamento (██████████).

Para este *data set* foi definida uma chave primária (██████████).

DS_PROCESSO_INFO

Este *data set* permite caracterizar todos os ██████████. Contem os dados com informação sobre o estado/fase em que se encontra ██████████ e razão justificativa

CONFIDENCIAL

Figura 6-13 Processo transformação do DS_PROCESSO_INFO

Para este *data set* foi definida uma chave primária



DS_NISS_INFO

Este *data set* permite caracterizar os



associados aos



CONFIDENCIAL

Figura 6-14 Processo transformação do DS_NISS_INFO

O processo de extração caracteriza-se, por não existir uma tabela fonte no sistema operacional SAPA, pelo que os *scripts* SQL de carregamento destas tabelas na fase de extração, refletem esta característica comum. Os *scripts* destes processos encontram-se disponíveis no anexo V - *Scripts da extração dos processos ETL*, subcapítulo dimensões manuais.

O processo de transformação das dimensões manuais é muito direta, e de fácil percepção face à origem única e de fácil entendimento que caracteriza a fonte dos dados.

6.4.3. *Dimensões automáticas*

Este grupo de dimensões caracterizam-se por terem uma tabela fonte num dos dois sistemas operacionais, [REDACTED]. Deste modo todas são alimentadas através de *scripts* SQL às tabelas fonte, na fase da extração. São vinte e oito os processos ETL com estas características, correspondendo cada um a uma tabela DIM, conforme quadro abaixo.

Tabela 6-3 Processos ETL versus dimensões automáticas

CONFIDENCIAL

O diagrama do processo ETL é similar para todas dimensões, sendo retratado no subcapítulo 6.2.2. onde é possível identificar todo o fluxo ETL.

Na definição das propriedades de cada processo o item “tabela dimensão” é singular para cada uma, como é óbvio, identificando a tabela no DW. Consoante o

sistema operacional fonte, o valor a colocar no item tabela origem será preenchido com o valor:

CONFIDENCIAL

O referido é exemplificado na imagem seguinte, com dois quadros distintos das propriedades de dois processos ETL. O processo [REDACTED], cuja fonte é do sistema [REDACTED] seguido do processo [REDACTED], cuja fonte é do sistema [REDACTED].

CONFIDENCIAL

Figura 6-16 Propriedades dimensões automáticas

No processo de extração o aspecto a salientar e que é diferente para cada processo existente é a tabela fonte utilizada nos *scripts* SQL, que pode ser do sistema operacional [REDACTED]. No caso de ser uma tabela do sistema [REDACTED] teve que ser identificado qual o [REDACTED] correspondente à tabela [REDACTED], conforme tabela seguinte.

Tabela 6-4 Tabela fonte usada na extração das dimensões automáticas

CONFIDENCIAL

Os scripts destes processos encontram-se disponíveis no anexo V - *Scripts* da extração dos processos ETL, subcapítulo dimensões automáticas.

O processo de transformação das dimensões automáticas, tal como as manuais é muito direta, e de fácil perceção face à origem única e de fácil entendimento que caracteriza a fonte dos dados.

6.4.4. Dimensões híbridas

Este grupo de dimensões caracterizam-se por, com exceção de uma tabela, usarem os *data set* criados no início da cadeia DW SAPA e utilizarem tabelas fonte nos sistemas operacionais [REDACTED]. Os processos ETL com estas características, inerentes tabela [REDACTED], são identificados na tabela abaixo, assim como os data sets fonte de dados.

Tabela 6-5 Processos ETL versus dimensões híbridas



O diagrama do processo ETL abaixo representado é exemplificativo dos processos acima referenciados, pois a única alteração a efetuar é a exclusão do [REDACTED] [REDACTED] em alguns deles.

A nomenclatura usada na definição do diagrama ETL, identifica:



CONFIDENCIAL

Figura 6-17 Caracterização do processo ETL das tabelas híbridas

Todos os scripts criados na fase de extração, destes processos, encontram-se disponíveis no anexo V - Scripts da extração dos processos ETL, subcapítulo dimensões híbridas.

CONFIDENCIAL

CONFIDENCIAL

Figura 6-18 Processo extração da DIM_SAPA_PRODUTO_APOIO_PRESC

CONFIDENCIAL

Figura 6-19 Processo transformação da DIM_SAPA_ENTIDADE

O processo ETL responsável pela cadeia que alimenta esta tabela dimensão, tem a nomenclatura [REDACTED]

CONFIDENCIAL

Figura 6-20 Processo ETL da DIM_SAPA_UTILIZADOR

Saliento dois aspetos particulares, no processo de extração tidos em conta:

CONFIDENTIAL

CONFIDENCIAL

Figura 6-21 Diagrama do processo *update* da DIM_SAPA_PRESCRICAO

CONFIDENCIAL

6.4.5. Tabelas de factos

O carregamento das tabelas de factos do DW SAPA caracterizam-se todas por evidenciarem o espaço temporal que o INR indicou como requisito, uma visão [REDACTED]

[REDACTED]. Saliento que não foi solicitado pelo INR, informação que [REDACTED]

[REDACTED].

Os processos ETL e inerentes tabelas factuais, são identificados na tabela abaixo.

Tabela 6-6 Processos ETL versus tabelas de factos

CONFIDENCIAL

Todos os scripts criados na fase de extração, destes processos, encontram-se disponíveis no anexo V - Scripts da extração dos processos ETL, subcapítulo tabelas de factos.

CONFIDENCIAL

CONFIDENCIAL

Figura 6-22 Processo extração da FAC_SAPA_PRESCRICAO

6.5. Run Jobs

Os *RunJobs* tem como função executar os *WorkJobs* e para cada *WorkJob*, existe obrigatoriamente um *RunJob*. Um *Runjob* não possui *design*, tendo apenas duas características, possui um parâmetro onde se menciona a diretoria base do projeto, diretoria SAPA, parâmetro [REDACTED] e possui uma chamada à rotina [REDACTED] que se encarrega de todas as tarefas de lançamento do *WorkJob*, onde se identifica este.



Figura 6-23 *Run jobs* da categoria DIMS

Ao nível da categoria DIMS, foram criados cerca de trinta e seis *Run jobs*, um por cada *Work Jobs*, facilmente identificados na figura acima.

Ao nível da categoria FACS, foram criados cerca de quinze *Run jobs*, um por cada *Work Jobs*, facilmente identificados na figura abaixo.

CONFIDENCIAL

Figura 6-24 *Run jobs* da categoria FACS

6.6. Chain Jobs

Os *ChainJobs*, tem como função executar um conjunto de *Jobs* com uma ordem pré-definida. Essa ordem de execução pode ser a execução de *Jobs* em sequência, um após outro, ou todos simultaneamente em paralelo.

Cada *ChainJob*, não possui qualquer *design*, sendo a sua única programação, a chamada no *JobControl* à rotina que controla a execução, a rotina *ChainConfig*. Este facto permite isolar nesta rotina toda a configuração da sequência de jobs que cada *ChainJob* executa.

Ao nível da categoria DIMS, foram criados cerca de seis *chain jobs* facilmente identificados na figura abaixo.

CONFIDENCIAL

Figura 6-25 *Chain jobs* da categoria DIMS

Ao nível da categoria FACS, foram criados cerca de seis *chain jobs* facilmente identificados na figura abaixo.



Figura 6-26 *Chain jobs* da categoria FACS

Capítulo 7

Relatórios BI

Inicialmente é feita uma breve referência à tecnologia de BI, usada para desenvolver relatórios estatísticos, o *MicroStrategy Developer*, versão 10.3.

De seguida é feita uma caracterização geral dos relatórios desenvolvidos, seccionando estes em sete grandes grupos face às suas características. É feito o paralelismo dos requisitos respondidos, que constam no documento enviado pelo INR, com os relatórios desenvolvidos. É feita referência aos relatórios que não foram ainda desenvolvidos, em virtude de faltar informação do INR, que possibilite a sua concretização

7.1. Tecnologia usada

Para concretizar o objetivo final, disponibilizar *reports* com a informação estatística atualizada e precisa, de forma a responder aos requisitos do INR, a tecnologia usada para o efeito foi o *MicroStrategy Developer*, versão 10.3. É a ferramenta usada no [REDACTED] para a criação dos relatórios estatísticos.

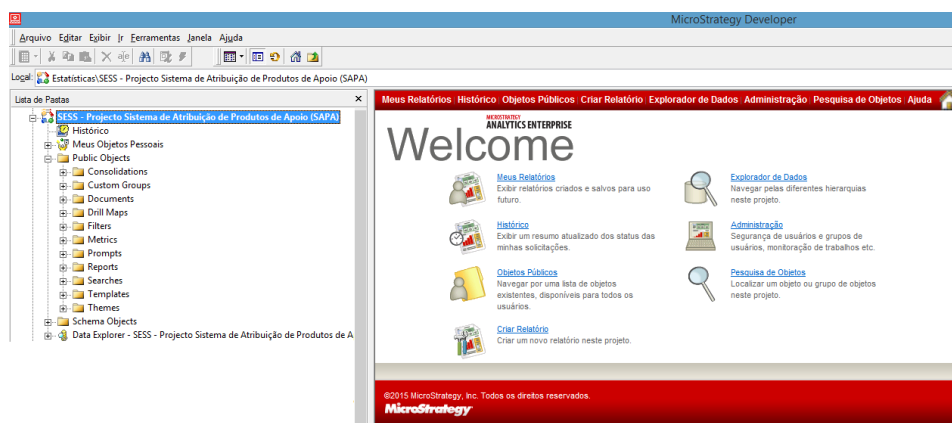


Figura 7-1 Framework do projeto SAPA no MicroStrategy

O *MicroStrategy* (*MicroStrategy Product Documentation*, 2017) é uma das ferramentas de *reporting*, usadas em ambientes empresariais. Caracteriza-se por possibilitar o desenho e criação de *reports* em modo gráfico, através de *drag-and-drop* na escolha dos seus componentes, da visualização e edição do código. Está diretamente ligado ao DW SAPA, fornecendo a informação requerida pelo utilizador e que se encontra armazenada no inerente [REDACTED]

CONFIDENCIAL

Figura 7-2 Filtro definido pelos utilizadores *MicroStrategy*

Na figura acima, esta exemplificado o exposto, um filtro com a opção de escolha

da região sobre qual se quer efetuar a consulta dos dados estatísticos.

7.2. Relatórios base do projeto

O conjunto de requisitos implementados no sistema, vão ao encontro do solicitado pelo INR e tem a sua parte visível nos indicadores e contagens a elaborar sobre os dados do SAPA, refletidos nos outputs gerados pela ferramenta de BI.

Neste capítulo identifica-se os relatórios base, relatórios esses que se podem agrupar em sete grandes grupos face às suas características:

1. Utilizadores.
2. Processos.
3. Entidades prescritoras.
4. Prescrições.
5. Produtos de apoio.
6. Beneficiários.
7. Classificação Internacional de Funcionalidade.

Saliento que na fase de análise, após o estudo efetuado constatou-se que existia a

CONFIDENCIAL

7.2.1. Utilizadores

É possível gerar relatórios que permitem caracterizar os utilizadores de acordo

CONFIDENCIAL

São cerca de três os relatórios que se caracterizam por responder a indicadores e contagens sobre dados inerentes aos utilizadores SAPA, estes são identificados no *MicroStrategy* com as seguintes nomenclaturas:

CONFIDENCIAL

Figura 7-3 Relatório 1.03, inerente aos utilizadores SAPA

Os relatórios respondem a três pontos do capítulo I, dos requisitos solicitados pelo INR identificados no documento que serviu de base para estes, disponível no Anexo III – Requisitos do INR, informação estatística. A tabela abaixo faz o paralelismo entre os referenciados pontos e os *reports* criados.

Tabela 7-1 Utilizadores, requisitos versus *reports*

CONFIDENCIAL

7.2.2. Processos

É possível gerar relatórios que permitem caracterizar os processos existentes em

CONFIDENCIAL

CONFIDENCIAL

Figura 7-4 Relatório 2.04, inerente aos processos SAPA

São cerca de quatro os relatórios que se caracterizam por responder a indicadores e contagens sobre dados inerentes aos processos SAPA, estes são identificados no *MicroStrategy* com as seguintes nomenclaturas:

Os relatórios respondem a um ponto do capítulo II, dos requisitos solicitados pelo INR identificados no documento que serviu de base para estes, disponível no Anexo III – Requisitos do INR, informação estatística. A tabela abaixo faz o paralelismo entre os referenciados pontos e os *reports* criados.

Tabela 7-2 Processos, requisitos versus *reports*

CONFIDENCIAL

7.2.3. Entidades prescritoras

É possível gerar relatórios que permitem caracterizar as entidades prescritoras existentes em SAPA de acordo com [REDACTED]

CONFIDENCIAL

Figura 7-5 Relatório 3.01, inerente às entidades prescritoras

Os relatórios respondem a dois pontos do capítulo I, dos requisitos solicitados pelo INR identificados no documento que serviu de base para estes, disponível no Anexo III – Requisitos do INR, informação estatística. A tabela abaixo faz o paralelismo entre os referenciados pontos e os *reports* criados.

Tabela 7-3 Entidades prescritoras, requisitos versus *reports*

CONFIDENCIAL

7.2.4. Prescrições

É possível gerar relatórios que permitem caracterizar as prescrições efetuadas em SAPA, de acordo com o seu [REDACTED]

CONFIDENCIAL

Figura 7-6 Relatório 4.03, inerente às prescrições

Os relatórios respondem a um ponto do capítulo IV, dos requisitos solicitados pelo INR identificados no documento que serviu de base para estes, disponível no Anexo III – Requisitos do INR, informação estatística. A tabela abaixo faz o paralelismo entre os referenciados pontos e os *reports* criados.

Tabela 7-4 Prescrições versus *reports*

CONFIDENCIAL

7.2.5. Produtos de apoio

É possível gerar relatórios que permitem caracterizar os produtos de apoio prescritos, [REDACTED]

CONFIDENCIAL

Figura 7-7 Relatório 5.01, inerente aos produtos de apoio

Os relatórios respondem a dois pontos do capítulo IV, dos requisitos solicitados pelo INR identificados no documento que serviu de base para estes, disponível no Anexo III – Requisitos do INR, informação estatística. A tabela abaixo faz o paralelismo entre os referenciados pontos e os *reports* criados.

Tabela 7-5 Produtos de apoio versus *reports*

CONFIDENCIAL

7.2.6. *Beneficiários*

É possível gerar relatórios que permitem caracterizar os beneficiários, de acordo com as [REDACTED]

CONFIDENCIAL

CONFIDENCIAL

Figura 7-8 Desenvolvimento do relatório 6.03, inerente aos beneficiários

Os relatórios respondem a quatro pontos do capítulo II, dos requisitos solicitados pelo INR identificados no documento que serviu de base para estes, disponível no Anexo III – Requisitos do INR, informação estatística. A tabela abaixo faz o paralelismo entre os referenciados pontos e os *reports* criados.

Tabela 7-6 Beneficiários versus *reports*

CONFIDENCIAL

7.2.7. Classificação internacional de funcionalidade

É possível gerar relatórios que permitem caracterizar a CIF, de acordo com as deficiências, incapacidades e fatores ambientais inerentes aos beneficiários. [REDACTED]

[REDACTED]

CONFIDENCIAL

Figura 7-9 Relatório 7.08, inerente à CIF

São cerca de oito os relatórios que se caracterizam por responder a indicadores e contagens sobre dados inerentes à classificação internacional de funcionalidade, estes são identificados no *MicroStrategy* com as seguintes nomenclaturas:

CONFIDENCIAL

Os relatórios respondem a três pontos do capítulo III, dos requisitos solicitados pelo INR identificados no documento que serviu de base para estes, disponível no Anexo III – Requisitos do INR, informação estatística. A tabela abaixo faz o paralelismo entre os referenciados pontos e os *reports* criados.

Tabela 7-7 CIF versus *reports*

CONFIDENCIAL

7.3. Relatórios não efetuados

Foram muitas as questões levantadas junto do INR, a fim de se clarificar por completo os requisitos deste projeto e que tiveram sempre como ponto de orientação o documento entregue pelo INR, disponível no Anexo III – Requisitos do INR, informação estatística.

Face ao *timing* delineado para a entrega do projeto e como houve questões levantadas nas reuniões realizadas, para a completa clarificação dos relatórios finais, que não obtivemos a resposta necessária em tempo útil para concretizar os *reports* no *Microstrategy*, estes não foram desenvolvidos no âmbito deste trabalho.

CONFIDENCIAL

Alem dos cinco pontos acima referenciados, existe ainda quatro pontos, em que a informação solicitada no documento de requisitos entregue pelo INR, face às suas características, não vai ser disponibilizado [REDACTED]. Ficou estabelecido, [REDACTED]

Os quatro pontos acima referenciados no documento de requisitos entregues pelo INR, são o ponto 5 do capítulo I, e os pontos 4,5 e 6 do capítulo IV.

Capítulo 8

Validações

É abordada a forma com se optou por efetuar a validação dos relatórios desenvolvidos. Numa primeira fase contemplando a eficiência do processo ETL, numa segunda fase é validada a qualidade dos dados, da informação dos relatórios e por fim é tido um feedback ao nível dos utilizadores finais sobre os relatórios estatísticos.

8.1. Processo ETL

A conclusão de todas as ações inerentes à execução dos processos ETL, inerente atualização do DW é de [REDACTED]. O tempo distendido não tem qualquer constrangimento, impacto negativo nas atividades a “jusante”, que dependem do seu carregamento. Na seguinte figura é possível constatar o tempo de execução acima referenciado, conforme valores do job que despoleta toda a cadeia o job [REDACTED].



Figura 8-1 Tempo de execução da cadeia ETL

O tempo de execução dos jobs inerentes ao carregamento dos *data set*, factuais e tabelas designadas de dimensões híbridas é de cerca de [REDACTED], conforme job acima [REDACTED]. Na figura a seguir é possível identificar a sequência da cadeia de jobs que este despoleta e respetivos tempos.



Figura 8-2 Tempo de execução dos jobs FAC

O tempo de execução dos jobs inerentes ao carregamento das tabelas designadas de dimensões manuais e automáticas é de cerca de [REDACTED], conforme job acima [REDACTED]. Na figura a seguir é possível identificar a sequência da cadeia de jobs que este despoleta e respetivos tempos.

CONFIDENCIAL

Figura 8-3 Tempo de execução dos jobs DIM

Conforme se pode constatar a cadeia é efetuada de [REDACTED], pois não houve a necessidade de reduzir o tempo de execução. Saliento que este tempo pode ser reduzido para menos, mas reforço, não se identificou necessidade em se efetuar tal. Os jobs são executados [REDACTED], quando na sua maioria podiam ser [REDACTED] redução do tempo de execução total.

CONFIDENCIAL



8.2. Validação dos dados dos relatórios

Para todos os relatórios produzidos pela ferramenta de BI, foram efetuados teste de forma a confirmar e validar a qualidade dos dados.

Os testes efetuados consistiram em comparar os valores dos outputs, resultantes das seguintes ações:

- A. Desenvolvimento de queries SQL sobre as tabelas do modelo operacional SAPA, de forma a ter como *output*, cada um dos relatórios gerados pelo *microstrategy*.
- B. Desenvolvimento de queries SQL sobre as tabelas do DW SAPA, de forma a ter como *output* cada um dos relatórios gerados pelo *microstrategy*.
- C. Gerar os relatórios no *microstrategy*.

Para cada relatório desenvolvido foram efetuados os passos acima identificados como A, B e C, comparando os valores finais de forma a identificar erros no carregamento do DW ou dos relatórios. Saliento que durante este processo foi identificado um erro no desenvolvimento de um dos *reports*, tendo sido identificado e corrigido desta forma.

As queries desenvolvidas para concretizar, os acima referenciados passos A e B estão disponível para consulta no anexo VI- Validação dos *reports*-testes finais.

8.3. Validação do cliente

Com o objetivo de avaliar o nível de satisfação dos utilizadores finais, sobre os relatórios estatísticos desenvolvidos e colocados em produção, foi elaborado um questionário com 8 perguntas, disponíveis no Anexo VII - Validação e perceção do grau de satisfação e relevância dos relatórios, focados no grau de utilidade, resposta e eficiência que estes proporcionam.

O questionário foi construído com a ferramenta “*Google Forms*” e disponibilizado via *link* de acesso através de correio eletrónico.

CONFIDENCIAL

Passo a
fazer uma análise a cada uma das respostas obtidas ao questionário.

Questão n.º1 - acesso, pesquisa e navegabilidade

Na questão que pretende obter indicadores sobre o nível de acesso, pesquisa e navegabilidade do sistema de BI, teve uma avaliação positiva, [REDACTED]

[REDACTED]

Deduzo que não será alheio à avaliação positiva, o facto de muitos dos utilizadores já estarem familiarizados com esta ferramenta. Atrevo-me inclusive a afirmar que os [REDACTED] que responderam médio serão utilizadores que tiveram um contacto, pela primeira vez com a mesma.

CONFIDENCIAL

Figura 8-4 Questão n.º1 do inquérito de validação dos *reports*

Questão n.º2 - inteligibilidade dos *layouts* e facilidade de interpretação dos relatórios

Na questão que pretende obter indicadores sobre o nível de inteligibilidade dos layouts e facilidade de interpretação dos relatórios, teve uma avaliação positiva, conforme [REDACTED]

[REDACTED] é mediana.

Pese existam relatórios com grandes quantidades de informação, depreendo que estes resultados se devem ao facto de as necessidades dos inquiridos não passarem pela necessidade de layouts gráficos e interativos, pese estes estejam contemplados, mas sem grande tempo dedicado ao seu desenvolvimento e criação.

CONFIDENCIAL

Figura 8-5 Questão n.º2 do inquérito de validação dos *reports*

Questão n.º3 - tempo de resposta do sistema

Na questão que pretende obter indicadores sobre o nível de tempo de resposta do sistema, teve uma avaliação positiva, conforme figura em baixo, [REDACTED]

[REDACTED] Saliento que existem fatores que entram na “equação” desta avaliação, que nada tem a ver com responsabilidades inerentes ao trabalho desenvolvido e ferramenta de BI e sim a características das máquinas usadas pelos utilizadores e velocidade e configuração das redes locais

CONFIDENCIAL

Figura 8-6 Questão n.º3 do inquérito de validação dos *reports*

Questão n.º4 - frequência com que são efetuadas as atualizações dos dados

Na questão que pretende obter indicadores sobre a frequência com que são efetuadas as atualizações dos dados, teve uma avaliação que se pode [REDACTED]

[REDACTED]

[REDACTED] responderam que é mediana.

CONFIDENCIAL

Figura 8-7 Questão n.º4 do inquérito de validação dos *reports*

Questão n.º5 - relevância da informação no trabalho que desempenham

Na questão que pretende obter indicadores sobre a relevância da informação no trabalho que desempenham teve uma avaliação positiva, conforme figura em baixo,

[REDACTED]

[REDACTED]

CONFIDENCIAL

CONFIDENCIAL

Figura 8-8 Questão n.º5 do inquérito de validação dos *reports*

Questão n.º6 - relevância na capacidade de visão e análise do processo de negócio

Na questão que pretende obter indicadores sobre a relevância na capacidade de visão e análise do processo de negócio SAPA, teve uma avaliação que se pode considerar positiva, conforme figura em baixo, [REDACTED]

CONFIDENCIAL

Figura 8-9 Questão n.º6 do inquérito de validação dos *reports*

Questão n.º7 - avaliação global dos relatórios

Na questão que pretende obter indicadores sobre a avaliação global dos relatórios SAPA teve uma avaliação positiva, conforme figura em baixo, [REDACTED]
[REDACTED]
[REDACTED]

CONFIDENCIAL

Figura 8-10 Questão n.º7 do inquérito de validação dos *reports*

Questão n.º8 - melhoramentos dos relatórios ou mesmo novos desenvolvimentos

Na questão que pretende obter indicadores sobre que área dos indicadores estatísticos, os utilizadores consideram relevante um melhoramento ou mesmo um desenvolvimento novo, sobre indicadores que não são visados atualmente, as respostas ilustram claramente que a componente financeiro/contabilística, [REDACTED]
[REDACTED], face ao intuito de haver a vontade, de numa segunda fase do projeto implementar dados sobre esta. A segunda área com maior percentagem de respostas foi na [REDACTED]
[REDACTED].

CONFIDENCIAL

Figura 8-11 Questão n.º8 do inquérito de validação dos *reports*

Capítulo 9

Conclusões

É feito uma sumarização dos objetivos respondidos pelo presente projeto e principais desafios e oportunidades que desde provieram.

São apresentadas as perspectivas de trabalhos futuros, destacando as mais-valias que resultam dos mesmos.

9.1. Conclusão e trabalhos futuros

O objetivo final deste projeto consistiu em disponibilizar informação estatística, atualizada e precisa ao Instituto Nacional de Reabilitação, de forma eficiente e dotar este, de mecanismos para suportar o processo de tomada de decisão de gestão de forma eficaz. Objetivo concretizado através do uso de uma ferramenta de *Business Intelligence* e inerentes *reports*, que traduz a parte visível de um longo e árduo trabalho.

Ao longo do desenvolvimento dos relatórios (*reports*), do DW, do estudo aprofundado das regras do processo de negócio, conceitos e atores inerentes ao Sistema de Atribuição de Produtos de Apoio, foram encontradas soluções que respondem aos requisitos, mas foram muitos os desafios e dificuldades com que me confrontei, próprios de um projeto desta natureza, abrangência e complexidade.

O conhecimento adquirido sobre o processo de negócio SAPA, teve como fonte principal de informação os interlocutores do Instituto Nacional de Reabilitação. O conhecimento dos mesmos espelhava o *core business* do INR, direcionado para assegurar a todas as pessoas a proteção social, garantia de igualdade de oportunidades, o combate à discriminação e a valorização das pessoas com deficiência. O SAPA abrange a realidade do *core business* do INR, mas é complementada com uma realidade que abrange a necessidade de conhecimentos no âmbito da medicina e das estruturas que são atoras no SAPA. [REDACTED]

[REDACTED] o conhecimento através de leitura de legislação, pesquisa de documentação e leitura de conteúdos no âmbito da medicina. Saliento, inclusive ter recorrido a amigos pessoais que são profissionais na área da medicina, tendo com eles cimentado conhecimento. Mas ainda assim não foi o suficiente em determinados conteúdos, conforme traduzido no subcapítulo 7.3 .

Além de ser fundamental para o desenvolvimento da arquitetura do DW, saliento que a análise do processo de negócio em conjunto com o conhecimento tecnológico, também permitiu definir e auxiliar os interlocutores do INR, na definição de indicadores relevantes.

O projeto *Data Warehouse* SAPA congregou dados dispersos pelos sistemas [REDACTED] num único repositório de dados, algo que não existia, permitindo obter uma visão única e integral dos dados, quer através da solução BI e em pedidos

futuros ao DAGI. Essa integração obrigou a um esforço complementar, principalmente ao nível do [REDACTED], pois o número de tabelas da base de dados do sistema operacional SAPA com que foi necessário congrega, é considerável.

Foram desenvolvidos cerca de trinta e um relatórios estatísticos com a ferramenta de *Business Intelligence*. No quadro abaixo é possível identificar os quatro capítulos e inerentes pontos do conjunto de requisitos constantes no documento entregue pelo Instituto Nacional para a Reabilitação, disponível no Anexo III – Requisitos do INR, informação estatística e a identificação dos *reports* desenvolvidos que dão resposta aos mesmos.

Da totalidade dos relatórios, em cerca de dez identificou-se a necessidade de contemplar indicadores distintos para o total de processos SAPA e de beneficiários, conforme última coluna da tabela. Decisão tomada pelo [REDACTED]
[REDACTED]. Os quatro pontos identificados na tabela com “Questão por clarificar”, serão desenvolvidos assim que o INR responder às dúvidas colocadas e cujas respostas são essenciais para o término dos relatórios solicitados.

Tabela 9-1 Documento de requisitos versus *reports* desenvolvidos

CONFIDENCIAL

Aquando da concretização pelo INR da segunda fase do desenvolvimento da aplicação SAPA, que vai contemplar dados contábilístico e financeiros é uma oportunidade de enriquecer ainda mais o conhecimento proporcionado com os novos

dados. O desenvolvimento de novos modelos no DW SAPA ou mesmo alguma redefinição dos existentes é de certeza um projeto futuro promissor, com uma enorme mais-valia traduzido no desenvolvimento de novos *reports BI*, com indicadores contabilísticos e financeiros.

Pessoalmente este projeto foi um enorme desafio, contribuindo para a minha valorização profissional e pessoal. Foi imprescindível o auxílio, apoio e transmissão de conhecimento que os dois colegas da equipa de trabalho do Departamento de Análise e Gestão de Informação, me prestaram. Destaco o contributo precioso destes, no ponto que mais dificuldades tive e que mais exigência me obrigou, trabalhar com a tecnologia usada para a extração, transformação e carregamento dos dados para o *Data Warehouse SAPA*, o *Infosphere Datastage and Infosphere Qualitystage* da IBM. Foi o meu primeiro contacto com esta ferramenta, fiquei ciente das suas potencialidades mas igualmente ciente, que muito tenho a aprender para poder tirar o máximo proveito desta.

Por fim, destaco o feedback positivo tido do cliente e inerente impacto que o projeto DW/BI SAPA tem, ressaltando a vantagem estratégica que o INR pode agora retirar da nova e preciosa fonte de informação.

Bibliografia

CIF, (2004) Classificação internacional de funcionalidade, incapacidade e saúde
(http://www.inr.pt/uploads/docs/cif/CIF_port_%202004.pdf)

Cláudio Sapateiro, (2016) documentação de apoio à disciplina de Sistemas de Apoio a Decisão EST- IPS.

Diário da Republica - Decreto-lei n.º 93-2009, Decreto de Lei que visa criar o Sistema de Atribuição de Produtos de Apoio.

Diário da Republica - Decreto-Lei nº 42-2011, alteração do Decreto-Lei n.º 93/2009, de 16 de Abril.

Diário da Republica - Despacho n.º 6133/2012, definição de conceitos no universo dos Produtos de Apoio.

Diário da Republica - Despacho n.º 16313/2012, homologa lista de Produtos de Apoio.

Diário da Republica - Despacho n.º 5128/2013, normas reguladoras da execução do e definição de procedimentos das entidades prescritoras e financiadoras de Produtos de Apoio.

Diário da Republica - Portaria n.º 192/2014, Portaria que regula a criação e manutenção da base de dados de registo do Sistema de Atribuição de Produtos de Apoio.

Diário da Republica - Despacho n.º 14278/2014, homologa lista de Produtos de Apoio reutilizáveis.

Diário da Republica - Portaria n.º 78/2015, aprovação do modelo da ficha de prescrição de produtos de apoio, no âmbito do Sistema de Atribuição de Produtos de Apoio.

Diário da Republica - Despacho n.º 7225/2015, aprova os procedimentos gerais inerentes à atribuição de produtos de apoio.

IBM Knowledge Center, IBM InfoSphere Information Server Version 9.1 documentation.

(https://www.ibm.com/support/knowledgecenter/en/SSZJPZ_9.1.0)

INR (2017), site do Instituto Nacional de Reabilitação, Instituto Público tutelado pelo Ministério do Trabalho, Solidariedade e Segurança Social.

(<http://www.inr.pt/>)

Jiawei Han University of Illinois at Urbana-Champaign Micheline Kamber (2006), Data Mining: Concepts and Techniques, Second Edition.

(http://ccs1.hnue.edu.vn/hungtd/DM2012/DataMining_BOOK.pdf)

Jorge Vaz de Oliveira e Sá, Tese de doutoramento (2009), Metodologia de sistemas Data Warehouse, Universidade do Minho.

MicroStrategy Product Documentation.

(<https://microstrategyhelp.atlassian.net/wiki/spaces/MSTRDOCS/overview?mode=global>)

Ralph Kimball, Margy Ross(2013), The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, Third Edition.

(<http://www.essai.rnu.tn/Ebook/Informatique/The%20Data%20Warehouse%20Toolkit,%203rd%20Edition.pdf>)

Anexo I

Tabelas do sistema operacional SAPA

Informação pormenorizada sobre as tabelas da base de dados que suportam o sistema operacional do SAPA

CONFIDENCIAL

Anexo II

Estudo direcionado para responder aos requisitos do INR

Estudo/transcrição do SQL efetuado à base de dados do SAPA, com o objetivo de ter como outputs a informação estatística solicitada pelo INR

CONFIDENCIAL

Anexo III

Requisitos do INR- Informação estatística

Documento entregue pelo INR, com a informação estatística que pretendiam visualizar como output, através da ferramenta de *Business Intelligence* .

CONFIDENCIAL

Anexo IV

Tabelas do Data Warehouse

Informação pormenorizada sobre as tabelas, que constituem as ■ tabelas do *Data Warehouse* SAPA.

CONFIDENCIAL

Anexo V

Scripts da extração dos processos ETL

Informação pormenorizada sobre todos os scripts desenvolvidos no processo de extração, sendo possível identificar;

- Tipo processamento ETL
- Nome da tabela DW ou do Data Set
- Nome do processo ETL

CONFIDENCIAL

Anexo VI

Validação dos *reports* - testes finais

Testes finais efetuados para garantir a validação dos *reports* gerados pelo *microstrategy*, comparando estes com os outputs gerados pelas queries à BD operacional e ao DW SAPA,, abaixo transpostas.

CONFIDENCIAL

Anexo VII

Validação e percepção do grau de satisfação e relevância dos relatórios

O questionário abaixo contempla o elenco de questões enviadas para obter o grau de satisfação e relevância dos relatórios com a informação estatística SAPA.

CONFIDENCIAL