

UNIVERSIDADE TÉCNICA DE LISBOA
INSTITUTO SUPERIOR DE ECONOMIA E GESTÃO



MESTRADO EM GESTÃO DE SISTEMAS
DE INFORMAÇÃO

*Concepção de um repositório de Master Data
de Entidades numa Seguradora*

Por:

Rodrigo Manuel Gonçalves Pereira Catarino

Orientação:

Ana Marques Ribeiro dos Santos Lucas

Setembro 2011

Resumo

Para empresas do ramo financeiro como bancos e seguradoras o seu principal activo são os dados. São os dados consistentes, seguros, fidedignos e disponíveis na altura certa que vão potenciar o conhecimento e possibilitar a tomada de decisões correctas por parte do negócio.

Qualquer organização possui dados sobre clientes, fornecedores, localizações, contratos, etc chamados de *Master Data*, que são críticos para o funcionamento dos principais processos de negócio.

Quando numa companhia de seguros estes dados estão espalhados por vários sistemas, existindo várias cópias diferentes, e sem qualidade, do mesmo objecto de negócio então a organização não conseguirá competir adequadamente num mercado tão maduro e regulamentado.

Este trabalho procura explicar os conceitos base do processo, ferramentas, métodos e tecnologias envolvidas na aquisição, correcção, integração, manutenção, controlo e partilha de *Master Data* (*Master Data Management*), e enquadrar a sua necessidade no mercado segurador e a aplicação prática numa organização particular através da metodologia de *Action Research*.

Palavras-Chave: *Master Data, Master Data Management, Data Governance, Data Quality, Seguros, SOA*

Abstract

In financial companies like insurance and banks their main asset is their data. Consistent, secure, accurate, accessible and timely data lead to knowledge and enabling the business to make the most correct decisions.

Any organization has data about their clients, suppliers, localizations, accounts, etc also known as Master Data, which are critical to their business process.

When an insurance company has his data spread out over multiple systems, having multiple different copies of the same data, without quality, of the same business object, then the organization will not be able to compete effectively in such a mature and regulated market.

This work intends to detail the basic of concepts of the process, tools, method and technology involved in the acquisition, correction, integration, maintenance, control and sharing of Master Data (Master Data Management). Is also an objective of the paper to detail the need of Master Data Management in the insurance industry and the practical implementation in a particular company through an Action Research.

Keywords: Master Data, Master Data Management, Data Governance, Data Quality, Insurance, SOA

“Action is the key to all success”

Pablo Picasso

Índice

1	Introdução	1
1.1	Questão de Investigação e Estrutura do relatório	1
2	Revisão literatura	3
2.1	Dados Informação e Conhecimento	3
2.2	O que é a <i>Master Data</i>	4
2.3	<i>Master Data Management</i>	5
2.4	Dimensões do sistema de MDM	9
2.4.1	Formas de Uso	9
2.4.2	Estilos de implementação	11
2.4.3	Domínio de aplicação	14
2.5	Arquitectura de referência	15
2.6	Fases de implementação	16
3	Os seguros	18
3.1	Definição	18
3.2	Supervisão na actividade seguradora	18
3.3	Processos de negócio	19
3.3.1	Cadeia de valor	19
3.4	Nível de adopção de Sistemas de informação	21
4	<i>Action Research</i>	22
5	A Investigação	26
5.1	A Companhia de Seguros	26
5.2	Definição do problema	27
5.2.1	Enquadramento global	27
5.2.2	Enquadramento específico	28
5.3	Plano de acção	31

5.4	Acção	31
5.5	Avaliação	32
5.6	Aprendizagem	33
6	Conclusões (contributos, limitações e investigações futuras)	35
6.1.1	Contributo	35
6.1.2	Limitações.....	37
6.1.3	Investigações futuras	37
7	Bibliografia	38
8	Anexos	41
8.1	Elementos chave no MDM	41
8.1.1	<i>Data Governance</i>	41
8.1.2	<i>Data Quality</i>	42
8.1.3	<i>Data Stewardship e Ownership</i>	43
8.2	Regras de Integração de dados.....	43
8.2.1	Atributos de dados de identificação de Entidades	43
8.2.2	Atributos destinados ao armazenamento de Contactos	48
8.2.3	Atributos de Morada associados à Entidade	48

Lista de Figuras

Figura 1- Dados, Informação e Conhecimento adaptado de (Thomas H. Davenport & Prusak, 1998)	4
Figura 2 - Dimensões do MDM (Dreibelbis et al., 2008).....	9
Figura 3 - Estilo de implementação – Consolidação (Dreibelbis et al., 2008)	12
Figura 4 - Estilo de implementação – Registo (Dreibelbis et al., 2008).....	12
Figura 5 - Estilo de implementação <i>Coexistência</i> (Dreibelbis et al., 2008)	13
Figura 6 - Estilo de implementação - Transaccional (Dreibelbis et al., 2008)	14
Figura 7 – Arquitectura de referência de um <i>Hub</i> (Bersov & Dubov, 2007)	15
Figura 9 – Macro funções do MDM na arquitectura funcional proposta por Otto et al., (2009).....	16
Figura 10 - Possível cadeia de valor interna da indústria seguradora	19
Figura 11 - Possível cadeia de valor externa na indústria seguradora	20
Figura 12 - Overall Insurance IT spend, 2004-2009 (..... Erro! Marcador não definido.	
Figura 13 – Ciclo de <i>Action Research</i> adaptado de (Baskerville & Wood-Harper, 1996)	23
Figura 14 - Evolução da produção Vida e Não Vida em Portugal (APS, 2009)	28
Figura 15 - Ligação entre os vários sistemas da empresa (AS-IS).....	29
Figura 16 – Ligações dos vários sistemas (To-Be).....	30
Figura 17 - Os passos a efectuar para resolver o problema.	31
Figura 18 - Principais razões de implementação de <i>data governance</i> (Teradata, 2010)	42

Lista de Tabelas

Tabela 2 - Mapeamento das dimensões de MDM com a solução definida	30
Tabela 2 - Categorias e características da qualidade dos dados (Wang et al., 1997)	43
Tabela 3 - Regra N° RI001 - Normalização do Nome da Entidade	44
Tabela 4- Regra N° RI002 - Correção de Erros Ortográficos no nome de Entidade	44
Tabela 5 - Regra N° RI003 - Determinação do valor do atributo 'Sexo' para não preenchidos	45
Tabela 6 - Regra N° RI004- SEXO – Identificação de valores fora do domínio esperado	45
Tabela 7 – Regra N° RI005- Tratamento de inconsistências entre o género e o 1º Nome de Entidade	46
Tabela 8 - Regra N° RI006 - Validação e correção de NIF's Portugueses	46
Tabela 9 - Regra N° RI008 - Data de Nascimento – Limpeza de inválidos	48
Tabela 10 - Regra N° RC001- Email – Identificação de Inválidos.....	48
Tabela 11 - Regra N° RM001- Estrangeiras – Correção do Código de País para moradas Portuguesas	49
Tabela 12 Regra N° RM002- Moradas Portuguesas - validação do Código Postal.....	50

1 Introdução

A *Master Data* ou dados mestre são a matéria-prima essencial para o desempenho de qualquer processo de negócio de uma organização. Estes dados descrevem os objectos de negócio principais, como por exemplo clientes, fornecedores, produtos, contactos que estão armazenados num sistema de informação. Como cada processo de negócio necessita de informação distinta, tipicamente vamos encontrar estes dados dispersos, inconsistentes e dessincronizados em diferentes sistemas de informação que muitas vezes não “falam” entre si.

Sendo o mercado segurador extremamente competitivo e regulado, e deparando-se actualmente com um panorama de crescimento lento ou até decréscimo do seu volume de negócios onde os clientes estão cada vez mais informados e possuem um conhecimento tecnológico que lhes permite facilmente procurar o melhor negócio, é essencial possuir uma visão harmonizada dos dados mestre. Uma companhia de seguros ágil, eficiente e focada no futuro está obrigada a gerir, distribuir e sincronizar os seus dados mestres dispersos por múltiplos sistemas. A criação de uma “única visão da verdade” destes dados é sem dúvida um imperativo para alcançar o sucesso, quer através de ganhos operacionais, melhorando os relatórios financeiros ou na retenção e fidelização do cliente.

1.1 Questão de Investigação e Estrutura do relatório

Esta investigação procura usando a metodologia *action-research* responder à questão de “como implementar um conceito de repositório de *Master Data* numa Seguradora.”

Este relatório pretende evidenciar a estratégia, métodos e acções realizadas na criação de um repositório de dados mestre. O relatório apresenta uma visão global de alguns conceitos relacionados com o *Master Data Management*, e através de uma experiência real numa organização concreta procura analisar os motivos e estratégia usada para realizar a integração de dados num único repositório de dados, que constituirão padrões comuns em outras organizações com o mesmo objectivo.

Este relatório encontra-se organizado em 6 partes. Logo após a introdução é apresentada a revisão de literatura sobre a *Master Data Management*. De seguida é apresentada sucintamente uma descrição do negócio segurador. No capítulo seguinte, são

apresentadas as características metodológicas da investigação são apresentadas. Em sequência é exposto o estudo realizado sobre uma companhia de seguros.

Finalmente, apresentam-se as contribuições do estudo e respectivas considerações finais.

2 Revisão literatura

Este capítulo pretende apresentar conceitos base que suportam a definição de *Master Data Management* (MDM) e o que envolve a sua implementação numa organização de forma a compreender e acompanhar de forma activa o processo de implementação.

2.1 Dados Informação e Conhecimento

Os três conceitos estão relacionados entre eles, mas têm significados diferentes. O sucesso ou insucesso de uma organização depende em grande parte de saber distingui-los, sabendo o que existe o que se pode ou não fazer com cada um deles. (Thomas H. Davenport & Prusak, 1998). Os dados são representações de factos como texto, números, imagens, gráficos som ou vídeo. Estes factos são capturados e guardados e transformados em dados, tipicamente, numa organização, através de uma infra-estrutura tecnológica (Thomas H. Davenport & Prusak, 1998). Os dados em si não permitem tomar decisão nem interpretação mas são a matéria-prima para criação de informação e por isso são de elevada importância numa organização (Thomas H. Davenport & Prusak, 1998). A informação surge quando é atribuído um significado, utilidade e contexto aos dados. (Thomas H. Davenport & Prusak, 1998).

Ao serem transformados em informação os dados podem ser colocados em categorias, analisados de forma estatística ou matemática, corrigidos ou condensados por máquinas ou humanos. (Thomas H. Davenport & Prusak, 1998). A informação permite responder a questões como “quem?”, o “quê?”, “onde?” e “quando?” (Bellinger, Castro, & Mills, 2004). Assim como a informação deriva dos dados também o conhecimento deriva da informação. O conhecimento pode ser definido como “informação que muda qualquer coisa ou alguém, servindo de bases para uma acção ou tornando um indivíduo (ou organização) capaz de acções diferentes ou mais eficazes”.

A aquisição de conhecimento é uma actividade humana e entre humanos (Thomas H. Davenport & Prusak, 1998) que reside no utilizador e na possibilidade de tomar acções sustentadas na informação disponível. (Lucas, 2010) permitindo responder à questão “como?” (Bellinger et al., 2004).



Figura 1- Dados, Informação e Conhecimento adaptado de (Thomas H. Davenport & Prusak, 1998)

2.2 O que é a *Master Data*

Os dados existentes numa organização são categorizados de forma ligeiramente diferente por vários autores. Analisando as diferenças e semelhanças podemos identificar com facilidade seis tipos de dados numa organização:

1. Não estruturados - aqui podemos encontrar informação contida em emails, artigos de revista, intranet, portais, fax, *pdfs*, etc.(Bai, Li, Li, & Song, 2010) (Wolter & Haselden, 2006)
2. *Master Data* (dados mestre) - Os dados mestre representam os principais objectos utilizados na organização sobre os quais são executadas as transacções pelas aplicações e também representam as dimensões chave dos dados analíticos. Normalmente encaixam em quatro grupos: entidades, produtos, localizações e conceitos. (Wolter & Haselden 2006, Dreibelbis et al. 2008, (Butler, 2011)
3. Transaccionais – as operações de qualquer organização é suportada por aplicações que ajudam a automatizar os seus processos negócio como por exemplo, vendas, compras, produção, facturação, contabilidade, etc. Os dados produzidos, que correspondem aos objectos envolvidos e sobre as próprias transacções realizadas pelas aplicações são considerados dados transaccionais (Wolter & Haselden 2006; Dreibelbis et al. 2008).
4. Analíticos ou históricos – representam os dados transaccionais e dados mestre acumulados ao longo do tempo numa organização. (Dreibelbis et al., 2008) Os dados analíticos são utilizados para suportar o processo de tomada de decisão de uma organização (Butler, 2011). Este tipo de informação tipicamente é armazenada em grandes tabelas de factos e dimensões como cliente, produto, fornecedor, localização em *Data Warehouses* e possivelmente em *data marts* mais pequenos desenhados para que suportam agregação, *data mining* e *ad hoc queries*.
5. *Metadata* – assim que os dados são movimentados e manipulados dentro da organização surge a informação sobre donde vieram, quem alterou, o que foi

alterado, etc. A *metadata* é assim informação estruturada que descreve, explica e localiza outra informação, permitindo-a obter, usar e gerir. Esta informação sobre informação pode residir num repositório ou em outro formato como documentos XML, descrições de colunas, ficheiros de *log*, ficheiros de configuração, etc (Wolter & Haselden 2006; Dreibelbis et al. 2008).

6. *Reference data* – enquanto a *metadata* descreve a estrutura, origem e significado a *reference data* procura definir e distribuir um conjunto de valores comuns. A existência de *reference data* potencia a eficiência operacional e analítica ao definir uma definição comum de valores para abreviações, códigos e validações comuns (Dreibelbis et al., 2008).

Embora pareça fácil identificar quais os dados mestre de uma organização, nem todos os dados que à partida encaixam na definição, podem ser tratados como tal. Cada sector de actividade deve identificar o seu próprio domínio de dados mestre. Por exemplo os requisitos, importância, ciclo de vida e as actividades exercidas sobre dados de um cliente serão distintos numa seguradora de uma empresa do sector têxtil. Podem, no entanto, ser identificados os principais domínios de dados encontrados em qualquer organização: entidades (ex: clientes, vendedores, empregados), produtos (ex: produtos, lojas), localizações (dados geográficos) e conceitos (ex: apólices, contratos, garantias) (Cleven & Wortmann, 2010).

Dreibelbis et al. (2008) identifica três grandes categorias: produto, entidades e conceitos que permitem responder às questões Quem? Como? e o Quê?. O domínio de localização é utilizado de forma transversal a todos e por isso é colocado num subdomínio. Estes diferentes domínios de dados necessitam de ter processos de consolidação, standardização, limpeza e distribuição para que se possa garantir a utilidade dos dados dentro da organização.

2.3 Master Data Management

O *Master Data Management* (MDM) é consistentemente definido na vária literatura consultada como a combinação de aplicações, tecnologias, ferramentas e processos necessários para criar e manter consistente e correcta os diferentes dados mestre de uma organização. O *Master Data Management* consolida, limpa e aumenta os dados mestre corporativos, e sincroniza-os com todas as aplicações, processos de negócios e

ferramentas analíticas. (Kumar, 2011), (Gartner, 2010),(Seiler et al., 2011)(Cleven & Wortmann, 2010) (Dreibelbis et al., 2008)

O *Master Data Management* (MDM) é definido Bersov & Dubov (2007) como “uma *framework* ” de processo e tecnologias destinadas a criar e manter uma ambiente de dados autorizado, de confiança, sustentável e fidedigno que representam a “única versão da verdade” ou “a cópia dourada” usado de forma transversal dentro e da organização”.

Embora que as definições encontradas na bibliografia sejam consistentes vamos considerar para este relatório a definição de *Master Data Management* de Loshin,(2009) o “incorpora aplicações de negócio, métodos de gestão de informação, e ferramentas de gestão de dados para implementar os procedimentos e infraestruturas que suportam a captura, integração e partilha de dados mestre fidedignos, consistentes, completos e atempadamente”.

Uma iniciativa de *Master Data Management* é aplicável a todos os sectores actividade e abrange um conjunto alargado de informação. Organizações como Bancos, Seguradoras, Telecomunicações, distribuição, manufacturação, fornecedores de energia, todas partilham a necessidade de ter acesso a informação completa, fidedigna, em tempo e segura sobre o seu negócio, performance financeira e operacional, produtos, clientes, fornecedores e competidores e outra informação importante e relevante. (Bersov & Dubov, 2007) .

Embora a necessidade de um sistema de *Master Data Management* não seja de todo recente, o interesse em iniciativas de MDM têm crescido em quase todos os sectores. Segundo a *Gartner* até 2014 66% das empresas do *Fortune* 1000 já terão implementado 2 ou mais soluções de MDM (Gartner, 2010).

Este interesse não é ocasional e está relacionado com alguns factores chave que levam as organizações a olharem com mais atenção a este tipo de soluções:

- *Regulamentação e legislação*

Uma série de escândalos empresariais levaram à criação de nova regulamentação e legislação como o *Sarbanes-Oxley Act*, e o *Basel II Capital Accord*. Estas novas directivas estão focadas na necessidade de fornecer, usar e reportar informação financeira de confiança, verificável e relevante sobre a performance financeira e eventos

que possam causar impacto financeiro. Para além das regulamentações internacionais as empresas têm que cumprir com a legislação e regulamentação nacional, que engloba a protecção de dados.

Esta nova legislação e regulamentação obriga as organizações mantenham a integridade, segurança, integridade, e acutilância dos seus dados e que ponham em prática controlos eficazes e verificáveis necessários sobre os dados organizacionais.

Ter e manter uma “única versão da verdade” permite perceber melhor os factores e tendências que podem afectar o negócio. (Bersov & Dubov, 2007)

- *Reduzir custos*

Ao ter um único ponto de controlo de dados mestre permite a redução de custos através da eliminação de aplicações que criam e usam versões locais dos dados. (Bersov & Dubov, 2007)

- *Ganhos Operacionais*

Ter informação correcta e completa sobre um cliente e as suas interacções com a organização permite compreender melhor o cliente, os seus objectivos, necessidades, capacidades, e a sua propensão a adquirir novos produtos ou serviços. Isto vai possibilitar oportunidades de *cross-sell* e *up-sell*.

- *Relacionamento com os Clientes*

Uma visão global do cliente permite a uma empresa oferecer serviços personalizados e apropriados melhorando o seu relacionamento com o cliente. As aplicações de *Customer Relationship Management* (CRM) e os benefícios associados serão potenciados com uma visão consolidada do cliente e da sua relação com a organização.

- *Arquitecturas orientadas a serviços*

A forte aposta nos últimos anos em sistemas orientados a serviços (SOA) e *Software as Service* (SaaS) torna o *Master Data Management* um tema actual e critico. Não será possível criar um serviço que disponibilize os detalhes de um cliente se este estiver disperso em várias bases de dados possuindo várias moradas e números de telefone.

A criação de um *Hub* central de dados mestre parece ser um desafio de enorme dificuldade, no entanto para além dos benefícios já referidos outros também serão evidentes:

- Criação de uma visão completa e estandardizada do cliente e das suas relações,
- Implementa uma plataforma de integração de dados que permite obter um único ponto de gestão,
- Normaliza, limpa e consolida dados mestre,
- Cria uma definição comum a toda a organização de uma determinada entidade,
- Define o formato e *syntax* de dados que deve ser utilizada na organização
- Permite a integração de várias aplicações até agora dispersas,
- Potencia o foco nos dados mestre da organização ao invés de uma visão departamental,
- Potencia a qualidade dos dados e a criação de procedimentos para a manter,
- Facilita a tomada de decisão,
- Reduz custos de manutenção e de novas implementações.,
- Melhor relacionamento com entidades e aumento de satisfação do cliente,
- Facilita o *compliance* com normas e legislações,
- Facilita a criação de relatórios financeiros e torna-os mais exactos,
- Processos aquisições e fusões deixam de ser tão problemáticos na integração de dados,
- Leva à implementação de processos de *data governance* ,
- Permite a criação de uma verdadeira arquitectura orientada a serviços,

A introdução de um sistema de MDM numa organização, coloca desafios à estratégia de gestão de dados organizacional, *data governance*, *data quality*, a um conjunto alargado de tecnologias, e aos perfis e responsabilidades que permitem a integração e interoperabilidade entre o *Hub* de Dados e as suas fontes e consumidores (pessoas e aplicações). Como tal é importante enquadrar a estratégia de gestão de dados mestre dentro da visão holística e multi-dimensional de uma *framework* de arquitectura empresarial e dentro da sua dimensão de gestão de dados.

2.4 Dimensões do sistema de MDM

O *Master Data Management* é um tema complexo e abrange vários aspectos e que podem ser abordados de diferentes maneiras. Foi escolhida a abordagem pela classificação pelas dimensões. Estas dimensões são características comuns identificadas em qualquer solução de MDM.



Figura 2 - Dimensões do MDM (Dreibelbis et al., 2008)

2.4.1 Formas de Uso

Ao analisar a forma como os dados mestre são criados, mantidos e utilizados dentro da organização, são identificados três formas de uso: Colaborativo, Operacional e Analítico. (Dreibelbis et al., 2008) Oberhofer & Dreibelbis 2008, MDM-Institute n.d.)

2.4.1.1 Colaborativo

Segundo o *MDM Institute* o modelo colaborativo é “a definição, criação e sincronização de dados mestre através de *workflow* e serviços de entrada-saida como por exemplo no *Product Information Management* (PIM) “. Esta forma de uso significa segundo Dreibelbis et al., (2008) a participação de diferentes utilizadores com diferentes perfis e múltiplas aplicações no mesmo processo de dados mestre. Esta abordagem é mais utilizada na gestão de dados mestre de produtos, um exemplo típico é o processo de criação de um produto onde desde o desenho, produção, definição do preço e marketing vários especialistas são chamados a colaborar no preenchimento da informação sobre o produto. A utilização deste método implica alguns requisitos chave para permitir uma coordenação eficiente e garantir a integridade dos dados: (Dreibelbis et al., 2008)

- Utilização de *workflows* com mecanismos de check-in/out
- Gestão de tarefas
- Monitorização do estado dos dados mestre.
- Gestão de relacionamentos e hierarquias

Esta abordagem colaborativa provoca a necessidade da existência de *data stewards* que possam garantir a qualidade de dados, bem como a existência de serviços que garantam a segurança e integridade e que suportem a administração, validação e importação/exportação de dados mestre. No entanto tudo isto é comum ao MDM independentemente da forma como é utilizado.

2.4.1.2 Operacional

O *MDM Institute* define o modelo Operacional como “a definição, criação e sincronização de dados mestre necessário para sistemas transaccionais e entregues através de arquitecturas orientadas a serviços (SOA).” Esta forma de uso é importante quando o sistema de MDM tem de funcionar como *Online Transaction Processing* (OLTP). Tipicamente um número alargado de utilizadores e aplicações requerem acesso rápido a dados mestre para retirar ou alterar dados através de serviços MDM invocados pelos processos de negócio como “Abertura de Conta”. Estes serviços são normalmente utilizados dentro de um contexto de arquitecturas orientadas a serviços (SOA – *Service Oriented Architecture*). Este tipo de arquitectura necessita de suportar vários milhares de transacções por segundo em milhões de registos. (Dreibelbis et al., 2008) À semelhança da forma Colaborativa também aqui são necessários serviços para limpeza, validação, detecção e correcção de duplicados e segurança e privacidade. (Dreibelbis et al., 2008)

2.4.1.3 Analítico

O *MDM Institute* define o método de uso analítico como “a definição, criação e análise de dados mestre.” O modelo analítico é usado quando o MDM e o *Business Intelligence* (BI) se tocam. O BI é um tema vasto que engloba o *data warehouse*, reports, *data mining*, *data marts* e outros. Seja qual o for a área de BI em causa, irá sempre necessitar de dados correctos e relevantes de forma a ser útil. Segundo Dreibelbis, Hechler, Milman, Oberhofer, Run, & Wolfson, (2008a) existem três pontos de contacto distintos entre o MDM e o BI:

- MDM como fonte de dados de confiança – A primeira função de um MDM é fornecer dados mestre ao *data warehouse*. Numa segunda fase o próprio *data warehouse* poderá completar os dados mestre. Por exemplo ao colocar no cliente métricas calculadas.
- Análise sobre os dados mestre – Aqui o próprio sistema de MDM possui mecanismo de análise sobre os seus dados que o permitam compreender melhor. Por exemplo necessita de responder a questões como “Quantos novos clientes recebi nos últimos dias?”
- Análise como função dentro do MDM – em alguns sistemas existem necessidades específicas de análise que permitam por exemplo verificar a identidade de um cliente ou descobrir relações desconhecidas.

2.4.2 Estilos de implementação

Os dados mestre tipicamente estão dispersos dentro de uma organização numa série de sistemas o que a torna inconsistente e de fraca qualidade. Um sistema MDM é implementado para resolver estes problemas para além da habilidade de distribuir dados mestre às várias aplicações em ambientes diversos. A implementação de um sistema de MDM para além de resolver os problemas relacionados com os dados mestre, como a qualidade a consistência e gestão, necessita também de responder a diversos requisitos técnicos e de negócio. Estes requisitos variam com o tempo o que torna as iniciativas de MDM dentro de uma organização um processo longo em vez de um projecto único e isolado. Para corresponder a estes requisitos são identificados por Dreibelbis, Hechler, Milman, Oberhofer, Run, & Wolfson, (2008a) diferentes estilos de implementação padrão utilizados na criação do *Hub* de dados mestre. Uma organização pode optar por utilizar uma implementação híbrida combinando vários destes estilos consoante os seus requisitos.

2.4.2.1 Consolidação

Este estilo de implementação agrupa e consolida a informação de várias fontes e sistemas num único *Hub* de dados mestre. Durante este processo os dados são transformados, limpos, agregados, e integrados de forma a obtermos o “registo dourado” de um ou mais domínios de dados mestre. A Figura 3 ilustra o estilo de Consolidação, onde um hub central de dados recebe dos vários sistemas heterogéneos com escrita e leitura e depois disponibiliza uma versão consolidada dos dados agora só

em modo de leitura, por exemplo a *data warehouses*.

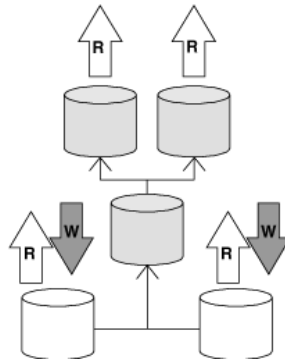


Figura 3 - Estilo de implementação – Consolidação (Dreibelbis et al., 2008)

2.4.2.2 Registo

Neste estilo também são disponibilizados dados mestre somente em modo de leitura. O sistema de MDM, como ilustrado na Figura 4, desempenha um papel diferente. O sistema de MDM não consolida nem realizada qualquer operação de *data quality* sobre a informação dos sistemas de origem. O sistema de MDM possui apenas uma parte dos dados mestre necessários para garantir a unicidade e referências para a aplicação que detêm o resto dos dados mestre. Este estilo de implementação é útil para remover duplicados e fornecer uma forma fácil e segura de aceder aos dados. A grande vantagem de implementação é o baixo custo de implementação em comparação com os outros, para além das poucas alterações realizadas nas actuais aplicações. Este método ao contrário dos outros não possibilita a existência de um verdadeiro registo de dados mestre, único e de alta qualidade, mas sim um registo com a referência para outros dados não geridos.

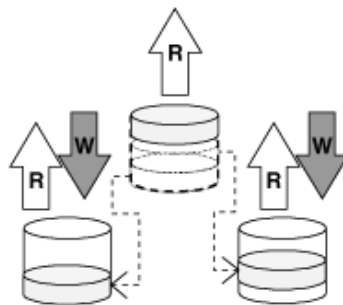


Figura 4 - Estilo de implementação – Registo (Dreibelbis et al., 2008)

2.4.2.3 Coexistência

Este estilo materializa todos os atributos dos dados mestre no sistema de MDM. É

possível alterar dados através do sistema de MDM como através das aplicações de origem. Neste estilo de implementação os dados obtidos do sistema de MDM não estão consistentes uma vez que podem existir alterações por sincronizar. O custo associado a este estilo de implementação é elevado. Os dados presentes no sistema de MDM devem ser tratados antes de inseridos e a implementação do mecanismo de sincronização também tem um elevado custo associado. Em relação ao modelo de *Registry* os benefícios obtidos são relevantes uma vez que a qualidade dos dados é superior e o acesso aos dados é mais rápido uma vez que não existe a necessidade de obter informações de outros sistemas. A utilização de *workflows* para edição e criação de dados mestre é facilitada. Os relatórios são mais fáceis de construir uma vez que agora os dados mestres estão presentes num único local

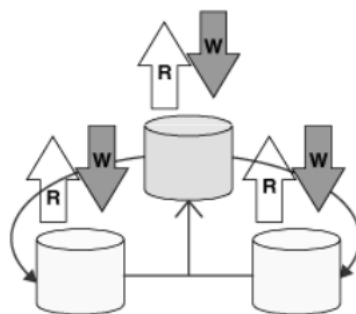


Figura 5 - Estilo de implementação *Coexistência* (Dreibelbis et al., 2008)

2.4.2.4 Transaccional

Com este estilo de implementação os dados mestres estão sempre consistentes, fidedignos e completos. A principal diferença para o modelo de coexistência é que ambos podem ler e escrever dados mestre mas agora isso é feito através do sistema de MDM. Isto significa que agora todas as aplicações que necessitem de modificar dados mestre, fazem-no através de serviços especificamente disponibilizados pelo sistema de MDM. Com isto é possível obter consistência total dos dados mestre. A utilização deste estilo de implementação implica uma alteração profunda da forma como as aplicações lidam com transacções de negócio sobre os dados mestre e a implementação de mecanismo de gestão de transacções.

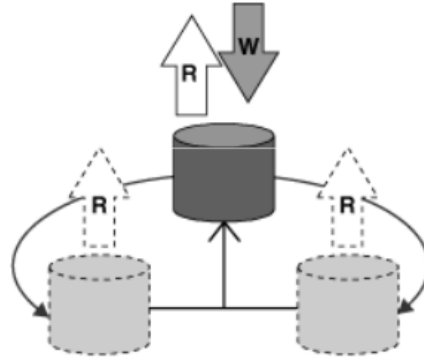


Figura 6 - Estilo de implementação - Transaccional (Dreibelbis et al., 2008)

2.4.3 Domínio de aplicação

Como referenciado no ponto 2.2 quando foi definido o conceito de *Master Data* esta pode ser pertencer a vários domínios. Definição de MDM não se restringe ao domínio de actuação dando a entender que este deverá ser desenhado para gerir todo o tipo de entidades identificadas como *Master Data*. O conceito de MDM até muito recentemente misturava-se com a sua aplicação a domínios mais restritos como o cliente no *Customer Data Integration (CDI)* e o produto no *Product Integration Management (PIM)*. A *Gartner* por exemplo, publica o quadrante mágico para cada subdomínio (Cliente e Produto) não existindo um para as soluções globais de MDM abrangentes Multi-domínio. (Gartner n.d. , Greenberg 2010) O conceito de *Master Data Management* multi-domínio é recente e surge á medida que as vertentes foram evoluindo e tornou-se evidente que embora o *CDI* tenha o seu foco em dados de clientes por vezes é conveniente ter os dados dos produtos ou das contas que o cliente tem, tal como no *PIM* é relevante ter dados dos fornecedores (Dreibelbis et al., 2008). A localização é referida também como um dos domínios importantes de *Master Data*, um cliente ou fornecedor terão uma ou mais moradas e um produto é vendido e ou enviado para um destino. Normalmente uma localização está sempre ligada a um cliente ou produto e poderá ser tratado de como subdomínio em alguns casos. (Dreibelbis et al., 2008)

A informação que deve ser considerada como *Master Data* varia consoante o negócio e o sector de actividade. Uma seguradora terá em atenção apólices, clientes e contas enquanto uma empresa de distribuição preocupa-se com produtos, localização e fornecedores. Este tipo de especificidade levou a que existam soluções específicas para determinado sector de actividade, adaptadas aos seus processos negócio.

2.5 Arquitectura de referência

Uma arquitectura de referência é uma abstracção ou generalização de uma solução técnica. Estas arquitecturas são o resultado do conhecimento adquirido em soluções implementadas com sucesso e incorporam padrões e melhores práticas identificadas. Uma arquitectura de referência para um sistema de MDM deverá ser uma proposta agnóstica e independente em termos de indústria onde aplicar e produto a utilizar que permita a implementação segundo as várias formas de uso e estilos de implementação.

Sendo várias as arquitecturas de referência encontradas na literatura consultada, estas pela natureza da sua definição tendem a ser semelhantes. Mesmo sendo o estudo das várias arquitecturas um assunto extenso e fora do escopo de análise é apresentado aqui a proposta efectuada por Bersov & Dubov, (2007) e a de Otto, Back, & Brenner, (2009). Estas duas arquitecturas de referência identificam claramente quais os componentes e serviços necessários para a implementação de um sistema de MDM. Estas arquitecturas de referência são especialmente úteis para uma organização definir os requisitos quando se decide por adquirir um produto.

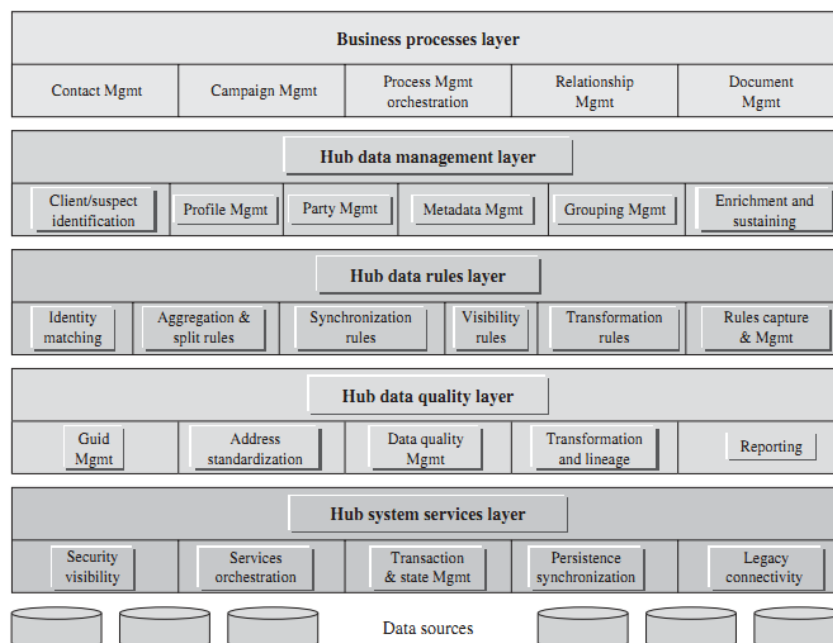


Figura 7 – Arquitectura de referência de um *Hub* (Bersov & Dubov, 2007)

Bersov & Dubov, (2007) define primeiro uma arquitectura de referência centrado para o desenho do *Hub* de dados (Figura 7) organizando os seus componentes e serviços distribuindo-os em diversas camadas consoante as suas funções. Depois apresenta uma

possível arquitectura enquadrando o *Hub* de dados numa visão da sua integração no fluxo de informação de uma organização.

Otto et al., (2009) realiza um estudo com o propósito de apresentar uma arquitectura funcional de referência focada nas aplicações que identifique as principais funções que um sistema de MDM deve realizar. Otto et al., (2009) identifica nesta arquitectura seis categorias funcionais dividido em 19 subcategorias (Figura 8) - que por sua vez se dividem em 72 funções precisas, que podem ser utilizadas para analisar a oferta de sistemas de MDM.

A Master Data Lifecycle Management	1 Data Creation	2 Data Maintenance	3 Data Deactivation	4 Data Archiving
B Metadata Management and Master Data Modeling	1 Data Modeling	2 Model Analysis	3 Metadata Management	
C Master Data Quality Management	1 Data Analysis	2 Data Enrichment	3 Data Cleansing	
D Master Data Integration	1 Data Import	2 Data Transformation	3 Data Export	
E Cross Functions	1 Automation	2 Reports	3 Search	4 Workflow Management
F Administration	1 Data History Management	2 User Management		

Figura 8 – Macro funções do MDM na arquitectura funcional proposta por Otto et al., (2009)

A descrição destas arquitecturas não faz parte do âmbito deste estudo, no entanto foi considerado relevante apresentá-las e aconselha-se o seu estudo aprofundado de estas e outras consultadas como a apresentada por Oberhofer & Dreibelbis (2008) ou por Bhagi (2011).

2.6 Fases de implementação

Uma iniciativa de MDM como qualquer outra deve ser implementada por fases dentro de uma organização. Da literatura consultada (Wolter & Haselden 2006, Radcliffe 2009, Butler 2011, Moreno & Mancuso 2007) podemos aferir uma sequência de passos habitualmente necessários:

- Análise da situação actual
 - Identificar e perceber quais as fontes de dados e a qualidade dos mesmos

- Identificar as aplicações e consumidores dos dados mestre identificados
- Recolher e analisar os dados mestre e a sua *metadata* (como por exemplo tipos de dados, valores por defeito, dependências e os donos da informação)
- Consolidar os dados num repositório central (*hub* de dados)
 - Desenvolver o modelo de dados
 - Implementar a infra-estrutura de suporte e escolher as ferramentas adequadas para limpar, transformar, consolidar e enriquecer os dados mestre consoante as regras previamente definidas.
 - Criar o *hub* de dados utilizando as ferramentas desenvolvidas ou adquiridas para integrar e consolidar os dados mestre.
- Implementar controlos
 - Identificar *data stewards*
 - Implementar um processo de *data governance*
- Partilha
 - Identificar as alterações necessárias a fazer nas aplicações.
 - A constituição do *hub* central de dados só será útil se existirem mecanismos de partilha com as aplicações, idealmente através de uma arquitectura *SOA*.
- Manutenção
 - Implementar as ferramentas, processos e perfis adequados à manutenção da qualidade dos dados.

Os seguros

Este capítulo tem como objectivo explicar os conceitos gerais relacionados com a actividade seguradora.

2.7 Definição

São várias as definições de seguro que se encontram na literatura. Praticamente todas as definições têm em comum a partilha do risco entre duas entidades mediante o pagamento de uma quantia como forma de premiar a entidade que aceita o risco (incerteza associada a um acontecimento futuro, seja quanto à sua realização, ao momento em que ocorre e aos danos dele decorrentes).

De uma forma mais simples, um seguro é um contrato entre uma entidade (companhia de seguros) que assume a cobertura de determinados riscos, mediante o recebimento de determinada quantia (prémio ou prestação), comprometendo-se a satisfazer as indemnizações, ou a pagar o capital seguro em caso de ocorrência de sinistro, nos termos acordados (ISP, 2010). A prestação do que ficou acordado no contrato pode ser efectuada à pessoa ou entidade do qual o seguro é celebrado (o segurado) ou de terceiro designado pelo tomador de seguro (o beneficiário) ou ainda uma terceira pessoa ou entidade que tenha sofrido prejuízos que o segurado deva indemnizar – o terceiro lesado (ISP, 2010). O negócio dos seguros está, tecnicamente, dividido pela natureza dos bens a segurar. Esta divisão é conhecida como o ramo do seguro. Os seguros tipicamente são classificados em seguros do ramo vida e não-vida. O primeiro inclui os seguros temporários de vida risco (genericamente cobram situações de morte e invalidez), os de capitalização, os seguros mistos e os de poupança (como por exemplo Planos Poupança Reforma- PPR). O ramo não-vida inclui o seguro automóvel, o seguro de incêndio, o de responsabilidade civil, o multiriscos, o seguro de saúde, o de acidentes de trabalho e o de acidentes pessoais (Silva & Dias 2002; Gilberto 2008).

2.8 Supervisão na actividade seguradora

A natureza da actividade seguradora onde quer as empresas de seguros quer os fundos de pensões assumem responsabilidades pelo pagamento de indemnizações e prestações que, por vezes, só se concretizam no futuro e na ocorrência de determinado evento, leva

a que praticamente em todos os países do mundo, existam autoridades de supervisão de seguros. Estas autoridades vigiam a capacidade das seguradoras e das entidades gestoras de fundos de pensões para assumirem e respeitarem os compromissos celebrados. Em Portugal, compete ao Instituto de Seguros de Portugal (ISP) supervisionar a actividade seguradora e resseguradora, para além dos fundos de pensões e a actividade de mediação de seguros (Gilberto, 2008).

2.9 Processos de negócio

2.9.1 Cadeia de valor

Para compreensão da actividade seguradora e para contextualizar a análise foi feito um estudo da actividade das seguradoras segundo o modelo da cadeia de valor. (Porter, 1998)

Este modelo analisa a empresa nas suas actividades de relevância estratégica, no sentido de se compreender o comportamento dos custos e as fontes existentes e potenciais de diferenciação. Segundo (Porter, 1998) uma empresa ganha vantagem competitiva executando estas actividades estrategicamente importantes de uma forma barata ou melhor que a concorrência.



Figura 9 - Possível cadeia de valor interna da indústria seguradora

A gestão de sinistros continua a ser a área mais ineficiente e com custos elevados dentro

Para uma companhia de seguros, a cadeia de valor é composta pelos processos, actividades e operações que suportam a actividade da gestão do risco e serviços associados. A Figura 9 ilustra uma possível cadeia de valor dos seguros.

de uma seguradora, correspondendo a sensivelmente 80% dos custos de operação de uma seguradora. (Europe et al., 2009) O desenho de produtos é um factor diferenciador entre os concorrentes, caso permita um correcto *time-to-market* juntamente com a definição de preço e coberturas.

A cadeia de valor de uma empresa, por sua vez, encaixa-se numa corrente maior de actividades, denominada de cadeia de valor externa, onde se situam os fornecedores e os canais de vendas (Porter, 1998).



Figura 10 - Possível cadeia de valor externa na indústria seguradora

Tradicionalmente uma empresa seguradora ainda está dependente dos mediadores de seguros. Eles têm uma relação próxima com o cliente final e acabam por possuir dados que acabam por não ser integrados nos sistemas das seguradoras, ora por questões tecnológicas ou culturais e de protecção do negócio. A cadeia de valor interna, evidencia com clareza que os dados são a matéria-prima dos seguros. Sem estes e sem um sistema de informação adequado à respectiva gestão, as companhias de seguros não conseguem exercer a sua actividade ou exercem-na com grandes limitações.

Com base nesta análise à cadeia de valor de uma seguradora, conclui-se que as actividades de gestão de apólices, gestão de sinistros e vendas assumem particular importância no desenvolvimento do negócio, uma vez que estão directamente relacionadas com os clientes.

2.10 Nível de adoção de Sistemas de informação

Os sistemas de informação sempre tiveram uma grande utilização no sector financeiro como banca e seguros. O uso de Sistemas de Informação é essencial para suportar o negócio em toda a cadeia de valor, reduzindo custos, fazendo mais com menos e possibilitando vantagens competitivas. Num mercado altamente competitivo e com pouca diferenciação no produto e no preço, grande parte do orçamento disponível é canalizado para o investimento em SI/TI na busca de otimização de processos, redução de custos operacionais e vantagens competitivas.

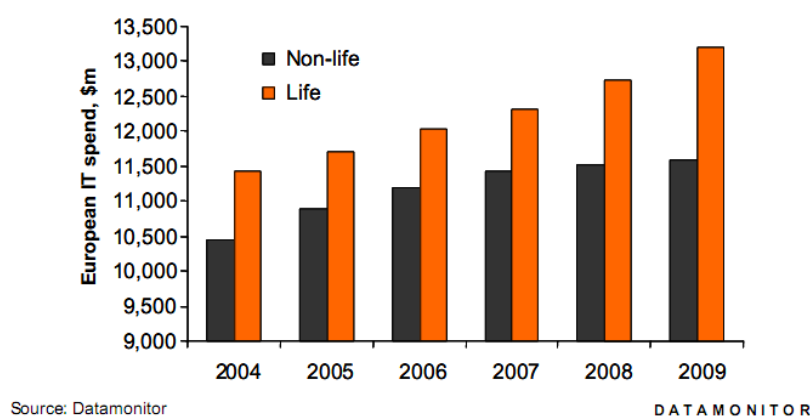


Figura 11- Gastos com IT entre 2004 e 2009

3 *Action Research*

Segundo vários autores consultados o método de *Action Research* baseia-se nos estudos e trabalhos realizados pelo psicólogo Kurt Lewin, nos Estados Unidos em meados de 1940, num contexto de pós-guerra (Baskerville and Wood-Harper 1996; Baskerville and Myers 2004).

Embora existam várias formas de *Action Research* (Baskerville, 1999; Baskerville and Wood-Harper, 1998), todas elas pressupõem o envolvimento e a colaboração estreita entre o investigador e os investigados (teoria e prática) através de acção e reflexão na abordagem de um problema concreto, sempre dentro de uma *framework* ética. O *action research* em geral coloca o investigador numa função em que pode ajudar a organização estudada e pode produzir importantes descobertas. (Baskerville, 1999a) Esta metodologia contribui para a compreensão de preocupações práticas e problemas imediatos, enquanto aumenta o conhecimento científico (Avison e tal, 1999). Ao passo que outras metodologias procuram estudar um processo mas não influenciá-lo ou alterá-lo, a aplicação do *Action Research* pode influenciar e alterar o processo e no limite até estudar o processo de alteração. Com a evolução dos SI, que se posicionam cada vez mais numa área de desenvolvimento e estruturação organizacional, esta metodologia ganha cada vez mais adeptos. Avison et al. (1999) refere que a *Action Research* é uma metodologia qualitativa única na forma como associa simultaneamente a investigação e a prática resultando em sinergias.

O método *Action Research*, é um processo sistemático e iterativo, onde é suposto existir a colaboração entre os participantes. O processo é descrito (Baskerville, 1999b; Baskerville & Wood-Harper, 1996), em cinco fases percorridas ciclicamente. Este processo ideal requer que numa primeira instância seja estabelecida uma infra-estrutura cliente-sistema ou ambiente de pesquisa (Baskerville & Wood-Harper, 1996). Esta infra-estrutura estabelece as regras e o acordo estabelecido entre investigador e investigados para a realização da pesquisa.

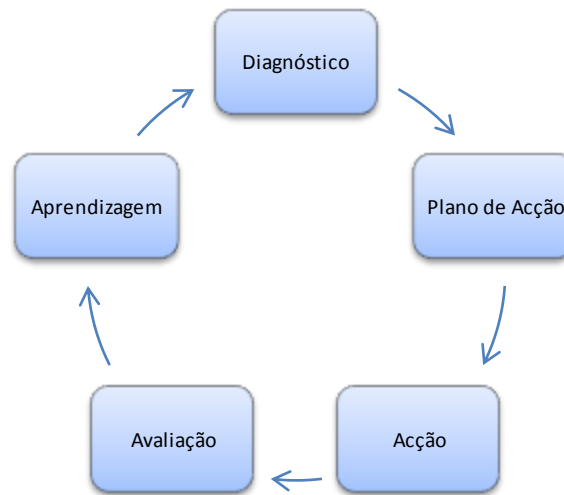


Figura 12 – Ciclo de *Action Research* adaptado de (Baskerville & Wood-Harper, 1996)

1. Diagnóstico

Segundo Baskerville (1999), a fase de diagnóstico corresponde à identificação dos problemas iniciais que são as causas subjacentes e que geram a necessidade de mudança. O diagnosticar envolve a auto interpretação do problema organizacional de uma forma holística, consequentemente irá ser criada um base teórica sobre o problema e eventuais soluções. Esta fase corresponde à identificação do problema.

2. Plano de Acção

Baskerville (1999) defende que nesta fase devem ser definidas as acções que melhorarão os problemas inicialmente identificados. As acções planeadas são guiadas pelo referencial teórico que substanciam a alterações a efectuar de modo a resolver o problema. O plano deve estabelecer a meta para a mudança e a forma de realizar. Neste ponto assiste-se à passagem da teoria à prática.

3. Acção

A acção é implementada tendo em conta o seu planeamento definido previamente Baskerville (1999). Os profissionais e os investigadores colaboram de forma activa na intervindo na organização, realizando as alterações identificadas. Existem diversas formas de estratégia de intervenção que podem ser adoptadas, nomeadamente: a intervenção directiva – a investigação conduz à mudança; ou não directiva – em que a mudança é solicitada indirectamente. Nesta fase é iniciada a transformação.

4. Avaliação

Após a conclusão das acções os profissionais e investigadores cooperam de forma a avaliarem os resultados (Baskerville 1999). A avaliação inclui verificar se os efeitos teóricos da acção foram alcançados, e se estes efeitos atenuam os problemas encontrados. Quando a mudança é bem-sucedida, a avaliação deve analisar de modo a saber se as acções realizadas foram o único motivo da resolução do problema. Quando a alteração não é bem sucedida, deve-se criar um novo *framework* para a próxima iteração do ciclo. Neste passo é realizada a reflexão e a tomada de consciência dos resultados obtidos.

5. Especificação da Aprendizagem

Embora seja o último passo esta actividade é um processo contínuo e recorrente (Baskerville 1999). O conhecimento adquirido com a condução de um *Action Research* independentemente de a acção ser bem ou mal sucedida, pode ser dirigida a três audiências:

1. Reorganização da organização de modo a reflectir o conhecimento adquirido durante a investigação.
2. Onde a mudança não for bem sucedida o conhecimento adicional poderá servir de base para o diagnóstico e como base para novas acções de *Action Research*
3. O sucesso ou fracasso do *framework* teórico fornece dados importantes para a comunidade científica em futuras de investigações.

O ciclo da *Action Research* pode continuar mesmo que a acção não seja bem sucedida, de modo a se desenvolver futuros conhecimentos sobre a organização (Baskerville 1999).

Na realização de uma *Action Research* o investigador depara-se com a dificuldade de obter o rigor desejado e daí ser importante ter em conta alguns aspectos como:

- Mudança de paradigma – a *Action Research* não se insere na tradicional perspectiva positivista e deve ser assegurado que existe consenso na utilização da metodologia.
- Estabelecer um acordo formal de investigação - o investigador deve obter o consentimento formal ou pelo menos o consentimento informado dos envolvidos.

- Estabelecer um enunciado teórico do problema - as acções desenvolvidas pelo investigador devem ter uma fundamentação teórica.
- Método de recolha de dados – a natureza empírica da *Action Research* leva a que o métodos qualitativos sejam mais utilizados, provocando um problema de validade dos dados devido à perspectiva interpretativista. Daí ser sugerida a recolha sistemática dos dados por exemplo através de um diário escrito. (Jepsen e tal., 1989)
- Colaboração e aprendizagem subjectiva – A colaboração é essencial, qualquer um dos intervenientes pode ter informação pertinente para a resolução do problema.
- Generalização – A possibilidade de generalização é um dos principais factores de investigação, alguns autores consideram tão importante como o rigor na investigação em Sistemas de Informação (Keen, 1991)., Sendo uma intervenção directa no problema de uma organização, dificilmente o resultado de uma *Action Research* pode vir a ser generalizada. No entanto é legítimo generalizar a partir de uma ou poucas observações desde que tal seja evidenciado e pela representatividade da amostra.

A realização deste estudo procurou seguir os elementos teóricos expostos sobre o método de investigação utilizado, a *Action Research*. Como colaborador da organização estudada, o investigador teve a oportunidade de intervir activamente nas rotinas da empresa, interagindo e colaborando activamente com outros colaboradores. Este facto facilitou o papel do investigador enquanto observador e interveniente directo, permitindo uma monitorização atempada das acções implementadas e consequente aprendizagem e adaptação.

Posteriormente é discutido o conjunto de passos metodológicos seguidos na organização em questão.

4 A Investigação

Neste capítulo pretende apresentar a aplicação prática numa organização do conhecimento teórico adquirido. Este capítulo será dedicado a apresentar a organização estudada, o contexto geral e específico do problema e qual o plano e acções realizadas.

A investigação ocorreu entre 15 de Outubro a 30 de Novembro e nela identificaram-se os serviços que a Informática do oferece à organização. A identificação foi feita inicialmente através dos incidentes já registados. Posteriormente realizaram-se reuniões one-to-one e reuniões conjuntas para discutir os serviços identificados. Os dados que se analisam ao longo deste capítulo foram obtidos de duas fontes:

entrevistas e documentos de controlo de gestão dos últimos três anos da organização em estudo. As entrevistas foram efectuadas a alguns elementos chave na implementação de um projecto de Balanced Scorecard e, para o efeito consideram-se dois momentos. Foi feita uma análise das ferramentas existentes no mercado e através da sua utilização, que

era muitas vezes limitada pelo facto de ser necessário licenças e de se ter acesso apenas a uma

demonstração, o investigador fez uma comparação entre elas, escolheu a que lhe pareceu

mais adequada e elaborou um relatório sobre as duas ferramentas que eram melhores. Está

representado no Anexo 2 a tabela comparativa de funcionalidades que foi utilizada no

relatório para justificar a escolha.

A informação foi recolhida através de entrevistas e documentos aos quais o investigador teve acesso.

4.1 A Companhia de Seguros

A companhia de seguros estudada é a filial de uma multinacional, com sede nos Estados Unidos da América, e é constituída por um conjunto de companhias internacionais de serviços financeiros diversificados e é um dos maiores grupos seguradores dos Estados Unidos da América. Com mais de 45.000 Colaboradores distribuídos por mais de 900 escritórios em todo o mundo, oferecendo uma ampla gama de produtos e serviços de elevada qualidade para particulares e empresas. Neste momento esta companhia de seguros está presente em 15 países. Esta companhia de seguros conta com oito décadas de experiência e está presente em Portugal desde 2003 através da aquisição de outra companhia seguros que já operava no mercado nacional. Já em 2010 existiu a incorporação de outra companhia de seguros que operava essencialmente no mercado de seguros ligados ao crédito habitação. Actualmente, conta com a dedicação de cada um dos seus cerca de 500 colaboradores em Portugal, na procura de melhores soluções de protecção para as famílias portuguesas, os particulares e para as micro, pequenas e médias empresas. Em todo o território nacional esta companhia de seguros conta com uma vasta rede comercial de escritórios e agentes de seguros, aliados estrategicamente à companhia, através dos quais se oferece um leque variado de produtos e serviços nos ramos vida e não vida que permitem aos clientes usufruir de uma vida mais segura e protegida. No ano de 2010, de acordo com o relatório de contas, o resultado atingiu o valor de 9,4 milhões de euros e em termos globais, esta companhia de seguros alcançou um volume de prémios brutos emitidos de 200,1 milhões de euros. Em termos do negócio não-vida, segundo o seu relatório de contas, o volume de prémios brutos emitidos atingiu o montante de 174,9 milhões de euros, e em termos do negócio vida, o volume de prémios brutos emitidos atingiu o montante de 25,2 milhões de euros.

4.2 Definição do problema

4.2.1 Enquadramento global

A história e natureza do negócio não deixam adivinhar grandes alterações na actividade seguradora. As companhias de seguros actuam num mercado maduro, saturado e fortemente regulado, onde o preço é cada vez mais um factor determinante. O espaço para inovar ou crescer é diminuto. Isto é comprovado pela Figura 13 onde podemos observar que o volume de negócio entre 2005 e 2009 no ramo não-vida se manteve praticamente idêntico assistindo-se até a uma quebra.

> PRODUÇÃO VIDA E NÃO VIDA										
	2005	2006	2007	2008	2009	+05/04	+06/05	+07/06	+08/07	+09/08
TOTAL VIDA	9.136	8.762	9.369	11.005	10.427	46,2%	-4,1%	6,9%	17,5%	-5,2%
Seguros de Vida	5.418	4.685	5.247	6.121	7.241	30,7%	-13,5%	12,0%	16,7%	18,3%
Seg. Ligados a Fund. de Invest.	3.161	3.376	3.245	3.994	3.161	78,1%	6,8%	-3,9%	23,1%	-20,9%
Operações de Capitalização	558	700	878	890	25	69,2%	25,5%	25,3%	1,4%	-97,2%
TOTAL NÃO VIDA	4.308	4.361	4.382	4.324	4.132	2,0%	1,2%	0,5%	-1,4%	-4,4%
Acidentes e Doença	1.311	1.350	1.372	1.396	1.353	3,1%	3,0%	1,6%	1,7%	-3,1%
Acidentes de Trabalho	779	774	763	741	674	0,7%	-0,7%	-1,4%	-2,8%	-9,1%
Doença	372	408	440	483	500	7,7%	9,7%	7,8%	9,6%	3,5%
Incêndios e Outros Danos	684	689	706	732	744	1,3%	0,6%	2,5%	3,7%	1,7%
Automóvel	1.997	2.004	1.944	1.810	1.666	1,6%	0,3%	-3,0%	-6,9%	-8,0%
Transportes, Rc Geral e Div.	315	318	360	383	368	2,4%	1,1%	13,1%	6,5%	-3,8%

U: milhões de euros Quadro 2

Figura 13 - Evolução da produção Vida e Não Vida em Portugal (APS, 2009)

Apesar desta aparente estagnação, as companhias de necessitam de se adaptar constantemente às necessidades dos clientes, alterações legislativas, exigências do mercado e desenvolvimentos tecnológicos. As companhias de seguros para além de estarem atentas à sua envolvente, desenvolvendo novos produtos e soluções financeiras actuam preferencialmente internamente, nos seus processos de negócio, na modernização de serviços, na maior rentabilidade e optimização de recursos, na criação de novas parcerias, na satisfação e retenção de clientes e no *up* e *cross selling*. Todos os processos de uma seguradora são completamente dependentes de dados para que possam realizar a sua função de forma eficiente e eficaz. A emissão de uma apólice e a regularização de um sinistro são fluxos intensos de informação entre a seguradora e as entidades com que se relaciona na sua cadeia de valor como clientes, prestadores de serviços, agentes, etc.

As companhias de seguros procuram nas soluções de *Master Data Management* uma forma quer de aumentar os lucros, melhorar a gestão de risco, *compliance*, optimização de processos, satisfação do cliente ou diferenciação e para obter vantagens competitivas.

4.2.2 Enquadramento específico

Como foi referido anteriormente a empresa de seguros em causa iniciou a sua actividade no mercado nacional através da aquisição de uma companhia de seguros e em 2010 incorporou outra. Para além deste facto a companhia de seguros adquirida já tinha passado por um processo anterior de fusão. Estando o sector de seguros tradicionalmente dividido em ramos vida e não-vida nesta organização por razões

históricas existem dois sistemas *Core* desagregados e desenvolvidos à medida para cada ramo. Estas duas situações aliadas a dados históricos acumulados sem grandes cuidados na sua gestão e qualidade, e novos requisitos legais e regulamentares, impostos quer por normas internacionais como o *Sarbanes Oxley Act*, quer nacionais pelo ISP contribuíram um problema evidente de *Master Data Management*. Mesmo tendo em conta esta realidade a organização conseguiu encontrar soluções alternativas internamente, através de aplicações que agrupavam, quando possível, os dados dos dois sistemas dando a ilusão de uma “única visão do cliente”. No entanto estas aplicações começaram a criar e a manter alguns dados mestre para responder aos requisitos cada vez mais exigentes do negócio.

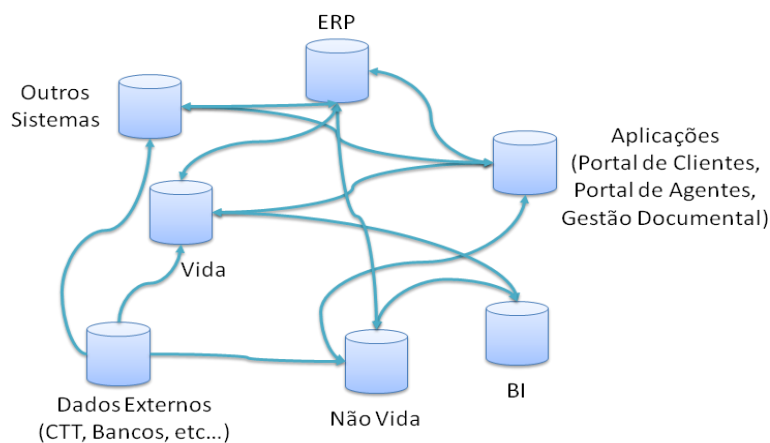


Figura 14 - Ligação entre os vários sistemas da empresa (AS-IS)

Em 2009 a companhia de seguros decidiu substituir os dois sistemas existentes através da aquisição de um pacote único para ambos os ramos vida e não-vida, integrando também o sistema incorporado em 2010.

Esta nova solução vai servir como *Hub* central de *master data* e *reference data* da seguradora disponibilizando também uma camada de *WebServices*. Esta camada de serviços permitirá que as aplicações satélite possam efectuar operações *CRUD* (*Create Read Update and Delete*) sobre esses mesmos dados garantindo simultaneamente a segurança, qualidade e autorizações necessárias e potenciando uma arquitectura orientada a serviços (*SOA*). Esta solução vai corresponder a um método de uso operacional e a um estilo de implementação transaccional.

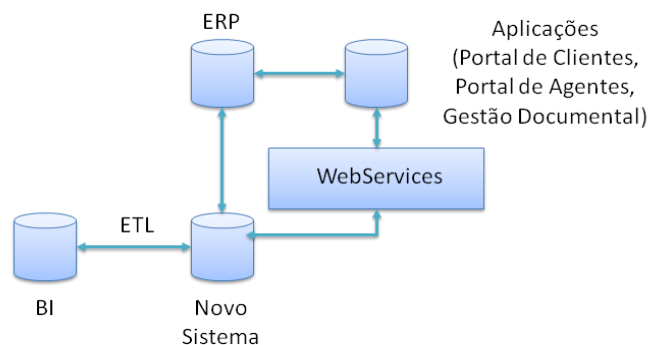


Figura 15 – Ligações dos vários sistemas (To-Be)

Dentro das tarefas críticas para que este projecto seja um sucesso é a migração e integração dos dados. Tendo em conta que os produtos vida e não-vida têm especificações próprias e devem ser tratados de forma isolada, o mesmo não se passa com as várias entidades relacionadas com a actividade da organização.

Um tomador de apólice vida, por exemplo, pode também existir em não-vida mas poderá ter diferentes moradas ou atributos diferentes ou complementares nos dois sistemas. Para além disso o paradigma de gestão de entidades é alterado no sistema adquirido. Os sistemas actuais fazem a gestão e manutenção de cada tipo de entidade de forma independente, ao contrário o novo sistema generaliza e cria os conceitos de *Party* (entidade) e *Stake* (papel). As entidades quer sejam individuais ou organizações passam a ser consolidadas nos seus atributos comuns como nome, nif, sexo, data de nascimento e depois é lhes atribuído um papel (tomador, terceiro, empregado, beneficiário, etc.) que define os seus atributos específicos. Os dados referentes a localização e contactos são tratados como dados mestre isolados. Cada entidade para além de um papel terá um ou mais contactos (morada, telefone, email, etc.). Desta forma podemos estabelecer neste momento como evidenciado na Tabela 1 - Mapeamento das dimensões de MDM com a solução definida o mapeamento entre a solução a implementar e as dimensões de *Master Data Management* já referenciadas.

Forma de uso	Operacional, Analítica
Implementação	Transaccional, e Consolidação
Domínio	Entidades, Conceitos (apólices), Localização

Tabela 1 - Mapeamento das dimensões de MDM com a solução definida

4.3 Plano de acção

Tipicamente a arquitectura de um *Master Data Management*, como analisado referido anteriormente na análise das arquitecturas de referência, é composta por quatro grandes componentes: uma repositório de dados mestre, um repositório de metadados, um componente de *data quality* e um componente de integração de dados.

Tendo presente estes componentes essenciais a um *Master Data Management* e seguindo a sequência típica apresentada anteriormente o plano acção identificado foi o seguinte:

1. Identificar e perceber quais as fontes de dados e a qualidade dos mesmos;
2. Recolher e analisar os dados mestre e a sua *metadata*;
3. Consolidar os dados no repositório central
 - a. Implementar a infra-estrutura de suporte e escolher as ferramentas adequadas para limpar, transformar, consolidar e enriquecer os dados mestre consoante as regras previamente definidas;
 - b. Criar o *Hub de Dados* utilizando as ferramentas desenvolvidas ou adquiridas para integrar e consolidar os dados mestre.

Cada um destes itens corresponde a um ciclo de *Action Research*.

4.4 Acção

No momento em que é efectuado este relatório só foi executada na totalidade o primeiro ciclo. Para tal foram identificadas a um nível macro, 4 acções principais consoante a figura apresentada:

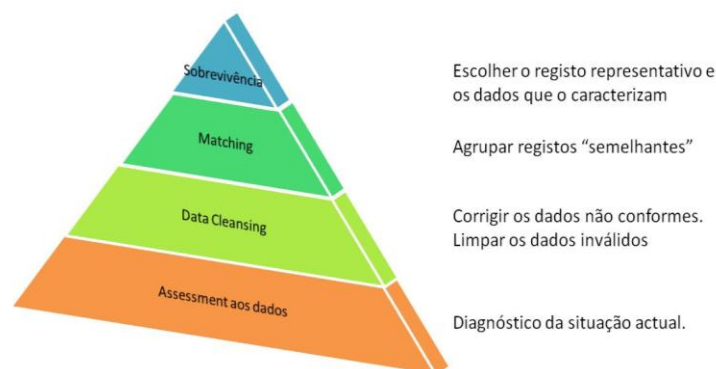


Figura 16 - Os passos a efectuar para resolver o problema.

1. Identificar as várias fontes de dados e a qualidade dos mesmos;
2. A implementação de um conjunto de processos de *cleansing* (limpeza de dados) aos dados de entidades identificadas com base nas várias regras definidas e que se encontram em anexo;
3. A implementação de um processo de *matching* de Entidades, utilizando critérios de comparação aproximada e que visa a identificação de registos que representam a mesma entidade, registos não relacionados e dispersos pelos vários sistemas;
4. A implementação de um processo de consolidação desses mesmos dados de entidades (após passo anterior), que visa a construção de uma visão única dos dados da entidade, através da implementação de regras de escolha automática desses mesmos dados representativos.

Estes 4 eixos de actuação vão dotar a organização dos mecanismos de limpeza futuros e proporcionam-lhe a preparação prévia dos dados para satisfazerem os requisitos de informação (e de modelo de dados) para o próximo ciclo de integração dos dados presentes nos actuais sistemas para um novo sistema futuro.

4.5 Avaliação

Foram analisados aproximadamente 1,5 milhões de registos, dos quais 26,5% provém de 7 fontes de informação do sistema vida e os restantes 73,5% provenientes de 2 fontes não-vida. A análise a atributos de identificação de entidade (seu nome, género, NIF, N° de BI e data de nascimento) permitiu identificar;

- Presença de erros ortográficos no Nome de Entidade (< 1% do universo). O campo género (Sexo) apresenta-se em aproximadamente 18.395 casos não preenchido. Destes, em 8.626 casos, pela análise do 1º nome, seria possível preencher correctamente o atributo. Foram registadas inconsistências entre Nome e Género (Nomes femininos classificados como masculino e vice-versa) numa fatia inferior a 1% do universo.
- O número de identificação Fiscal está vazio em quase 10% do universo (aproximadamente 155.000 casos). Nos casos preenchidos, registámos 18,8% de inválidos (296.501 casos) na regra do *checkdigit*. Registámos também inconsistências entre a natureza jurídica da entidade e o 1º dígito do NIF. No

universo de empresas, 2,4% deste possui NIF's começados pelo dígito 1 e 2. De forma idêntica, no universo de particulares, 1,15% deste possui NIF começado por 5 ou 6.

- 14.482 casos de NIF's começados por 8 que se encontram desactualizados (pertenciam aos antigos empresários em Nome Individual. Estes casos foram descontinuados pelas finanças). Recomendamos a sua limpeza.
- O número de Bilhete de Identidade está preenchido apenas em 45% do universo. Este é um campo tradicionalmente “difícil”, pois não permite a adopção de mecanismos de validação mais rigorosos. Analisámos o formato, valores “suspeitos” e valores de grande frequência. Dos 55% preenchidos, 24,4% deles parecem-nos “suspeitos”, devendo ser alvo de investigação mais profunda na fase de *cleansing*.
- Na data de nascimento, registamos 4,75% de vazios. Dos preenchidos, registamos 333.000 datas inválidas. Suspeitamos, também, da existência de uma franja de datas válidas mas não verdadeiras (datas no futuro ou datas muito antigas, por exemplo).
- Para todos os casos recomendamos a implementação de processos automáticos de limpeza/correção (limpeza de NIF's inválidos, por exemplo. Para o bilhete de identidade, as eventuais regras de limpeza a definir devem ser cuidadosamente analisadas.) Já em relação às datas de nascimento vazias, não haverá muito a fazer.

Da análise ao universo fornecido para caracterização de Entidades Únicas (por oposição, estudo por critérios de comparação aproximada que caracteriza os casos que representam a mesma entidade), revelou que à partida cerca de 86,4% do universo é constituído por entidades únicas. Por oposição, para 13,6% dos registos, encontrou-se pelo menos 1 entidade com semelhanças.

4.6 Aprendizagem

Este estudo permitiu balizar, de uma forma geral, a percentagem do universo total em que se esperam operações de consolidação de dados. De facto, será em aproximadamente 13% do universo que a organização terá de intervir com regras de consolidação de informação já definidas e apresentadas em anexo para o próximo passo

do plano e criar mecanismos activos de controlo de qualidade e de gestão para evitar que situações idênticas venham a ocorrer.

5 Conclusões (contributos, limitações e investigações futuras)

5.1.1 Contributo

Este trabalho de fim de mestrado levado a cabo segundo a metodologia de *Action Research* teria obrigatoriamente de resultar numa proposta de acções práticas a serem efectuadas pela organização.

Identificar as mesmas entidades, actualmente dispersas em vários sistemas, e fornecer uma “visão” consolidada e com qualidade da informação das mesmas e a criação das primeiras regras de limpeza, consolidação, normalização de dados, controlos e métodos de qualidade, são reconhecidamente as pedras basilares para uma verdadeira estratégia de *Master Data Management* numa Companhia de Seguros.

Porém, como referido sem uma estratégia de *data governance* e metodologias adequadas, este passo pode ser insuficiente ou mesmo ineficaz.

Grande parte do sucesso dessa estratégia vai bem além do plano tecnológico necessário: passa por um processo de mudança organizacional, de mudança de hábitos e costumes na forma e processos de trabalho da organização. Gerir esta mudança é um processo complexo e, por vezes, penoso. Porém, superar este desafio torna qualquer organização mais ágil, mais apta para enfrentar as inevitáveis e rápidas mudanças na conjectura dos mercados.

Abraçar esta estratégia passa, por endereçar eficazmente os seguintes vectores basilares:

- Definir claramente quem são os *owners* da informação – O *ownership* dos dados está a transitar do departamento de TI (tradicionalmente apontado como "dono" dos dados ou responsabilizado pelas suas inconsistências) para as unidades de negócio. São elas os "consumidores" da informação, quem beneficia (ou sofre) com a boa (ou fraca) qualidade dos dados que utiliza e dos quais depende. Entendemos que, se são as áreas de negócio o público-alvo dos dados, deverão, também, ser estas a assumir para si o papel de donos da informação. Ser "dono" da informação não implica assumir (mais uma) responsabilidade. Implica, também, ter os meios (orçamento, calendário e recursos) para gerir as acções sobre os mesmos, garantindo que se gere a qualidade dos dados exactamente da

mesma forma como se gere qualquer projecto: Definindo prioridades de actuação, possuindo um orçamento, calendário e recursos afectos e funcionando coordenado com os outros departamentos da empresa;

- Definir os segmentos de informação a controlar e implementar os mecanismos de monitorização periódica da qualidade dessa mesma informação - Sabemos que nenhum sector da empresa é levado a sério quando aponta inconsistências ou explica o quanto os seus processos de negócio "sofrem" com essas inconsistências. É certo que, de uma forma qualitativa, devemos entender as causas e os impactos que um menor nível de qualidade dos dados nos impõe. Porém, esta nossa análise deve estar fundamentada pelo aspecto quantitativo: Sermos capazes de "medir" eficazmente a extensão dos problemas. Seleccionar a informação a controlar e implementar os processos de medição periódica dos níveis de qualidade desses mesmos dados é fundamental. Superado este objectivo, o passo lógico seguinte será a definição de objectivos a atingir e a classificação da qualidade da informação como "suficiente" ou "insuficiente" para utilização no processo de negócio A ou B da empresa.
- Gerir a qualidade dos dados de forma activa - Uma significativa parte das organizações é reactiva no que toca a assuntos relacionados com a qualidade dos seus dados. Quando o fraco nível de qualidade dos mesmos já afectou (em grande parte dos casos, grandemente) as operações da empresa, quando nos vimos perante impactos consumados, reagimos com a tentativa de correcção desses mesmos dados. É necessário alterar este paradigma e de mudar da capítulo reactivo para o capítulo activo: Aprender que a qualidade dos dados deve passar a ser um processo gerido onde, o mais possível, é a empresa que tem o controlo e que é capaz de determinar e priorizar o que deve ser corrigido e quando se devem levar a cabo tais intervenções.
- Criar um centro de competências na área da qualidade dos dados - Criar, internamente, as competências de "*know how*" técnico e metodológico reunidas em torno de uma equipa de profissionais que assume a responsabilidade de liderar a implementação dos mais variados processos de auditoria e limpeza dos dados da empresa. Essa equipa deverá misturar as variadas competências técnicas necessárias para estes trabalhos e actuar como veículo de divulgação metodológica, definindo "Boas práticas" de desenvolvimento, formando outras áreas da empresa, etc.

- Implementar e fomentar uma cultura de responsabilização - Os "produtores" de informação, devem ser premiados se gerirem a informação com qualidade e fomentando um maior nível exigência nos restantes. Este é, será um dos aspectos mais difíceis de atingir mas, a longo prazo, dos objectivos mais críticos: Sensibilizar quem gera informação para deixar de ser "parte do problema" para passar a ser "parte importante da solução". Globalmente, todos os intervenientes beneficiam com a existência de informação de qualidade. Todos deverão ser parte activa na melhoria e protecção de um dos mais preciosos activos da empresa: Os seus dados.

5.1.2 Limitações

O tema de *Master Data Management* é complexo e vasto tendo ligações a toda a arquitectura da organização. O âmbito deste trabalho é restrito e foca-se numa pequena parte do *Master Data Management*. Este trabalho é também limitado nas interpretações do investigador enquanto colaborador.

5.1.3 Investigações futuras

No seguimento do que foi feito até ao momento na organização, propõe-se a realização de novos ciclos de *Action Research* contextualizados nos problemas identificados e conclusões iniciais. No decorrer deste trabalho vários outros temas mereciam um estudo mais aprofundado:

- Qual o impacto de fusões e aquisições na qualidade de dados de uma Seguradora?
- Como integrar e enriquecer os dados de entidades com dados não estruturados, de redes sociais, ou georeferenciação num *Customer Data Integration*?
- Qual o impacto do *Cloud Computing* no *Master Data Management*?

6 Bibliografia

- APS. (2009). PANORAMA DO MERCADO.
- Ana Lucas. (2011). *Corporate Data Quality Management From Theory to Practice. Information Systems and Technologies* (pp. 1-7). Lisboa. doi:10.1007/s00038-011-0292-2
- Bai, X., Li, P., Li, H., & Song, X. (2010). Enterprise Mater Data Manage Project Practice. *System*, (Iccasm), 220-223.
- Baskerville, R. L. (1999a). INVESTIGATING INFORMATION SYSTEMS WITH ACTION TUTORIAL INVESTIGATING INFORMATION SYSTEMS WITH ACTION. *October*, 2(October).
- Baskerville, R. L. (1999b). A critical perspective on action research as a method for is.pdf. communications of the Association for Information Systems.
- Baskerville, R. L. (1999c). Investigating Information Systems with Action Research. *Communications of the AIS*, 2(October). communications of the Association for Information Systems. Retrieved from http://www.cis.gsu.edu/~rbaskerv/CAIS_2_19/CAIS_2_19.html
- Baskerville, R. L., & Wood-Harper, A. T. (1996). A Critical Perspective on Action Research as a Method for Information Systems Research. *Journal of Information Technology*, 11, 235-246.
- Bellinger, G., Castro, D., & Mills, A. (2004). Data, information, knowledge, and wisdom. URL: <http://www.systems-thinking.org/dikw/dikw.htm>. Retrieved from http://www.dsd.go.th/itrain/km/kboc/DIKUW_KM.pdf
- Bersov, A., & Dubov, L. (2007). *Master Data Management and Customer Integration for the Goblal Enterprise*. (The McGraw-Hill Companies, Ed.)*Building*. The McGraw-Hill Companies.
- Bhagi, C. S. (2011). Practical Approach for Master Data Management. *Information Technology Journal*, 1(5), 213-216.
- Butler, D. (2011, June 1). Master Data Management. *Combinatorial chemistry & high throughput screening*. Oracle.
- Cleven, A., & Wortmann, F. (2010). Uncovering four strategies to approach master data management, 1-10.
- Data Governance Institute. (n.d.). Data Governance. Retrieved September 12, 2011, from http://www.datagovernance.com/adg_data_governance_definition.html
- Dreibelbis, A., Hechler, E., Milman, I., Oberhofer, M., Run, P. V., & Wolfson, D.

- (2008). *Enterprise Master Data Management - An SOA Approach to Managing Core Information. Data Management*. IBM Press. Retrieved from <http://media.techtarget.com/searchDataManagement/downloads/MasterDataManagementSOA1.pdf>
- Europe, D., House, C., Kingdom, U., Germany, D., Pacific, D. A., & Park, R. D. (2009). *European Insurance Technology Strategies. Strategies*.
- Gartner. (2010). Master Data Management Summit 2010 Summit Overview Who should attend. In Gartner (Ed.), *Data Management*. Las Vegas: Gartner.
- Gartner. (n.d.). Gartner Magic Quadrants. Retrieved September 5, 2011, from <http://www.gartner.com/technology/research/methodologies/magicQuadrants.jsp#m>
- Gilberto, F. (2008). *Manual Prático dos Seguros*. (Lidel, Ed.).
- Greenberg, C. (2010). The Great MDM Debate: Specialist or Generalist, Depth or Breadth? *Sandhill*. Retrieved from http://sandhill.com/opinion/daily_blog.php?id=75&post=661.
- ISP. (2010). *Guia de seguros e fundos de pensões*. (I. S. Portugal, Ed.). Instituto de Seguros de Portugal.
- Kumar, M. (2011). Master Data Management and It's Implementation Challenges. Retrieved from <http://www.ibmsoftwareservicesindia.com/master-data-management-and-its-implementation-challenges/>
- Loshin, D. (2009). *Master Data Management*. Elsevier.
- Lucas, A. (2010). Corporate data quality management: From theory to practice. *Information Systems and Technologies CISTI 2010 5th Iberian Conference on*. Retrieved from http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5556606
- MDM-Institute. (n.d.). MDM Institute. Retrieved September 24, 2011, from <http://www.the-cdi-institute.com/>
- Moreno, A., & Mancuso, G. (2007). Master Data Management: Transcending Technology, Solving Business Problems. *cutter IT Journal*, 20(9). Retrieved from http://www.johnbryce.co.il/owpdf/STKI_MDM_cutter_sept2007.pdf#page=13
- Oberhofer, M., & Dreibelbis, A. (2008). An introduction to the Master Data Management Reference Architecture. *Data Management*, (Mdm), 1-26. Retrieved from <http://download.boulder.ibm.com/ibmdl/pub/software/dw/dm/db2/dm-0804oberhofer/dm-0804oberhofer-pdf.pdf>
- Oracle. (2009). *How Technology Enables Data Governance*. San Francisco: Oracle.
- Otto, B., Back, A., & Brenner, W. (2009). Functional reference architecture for corporate master data management. *Institute of Information Management*.

Retrieved from
[http://web.iwi.unisg.ch/org/iwi/iwi_pub.nsf/wwwPublYearGer/F429893B96A8328CC125762D002DD840/\\$file/MdmFunctionalReferenceArchitecture.pdf](http://web.iwi.unisg.ch/org/iwi/iwi_pub.nsf/wwwPublYearGer/F429893B96A8328CC125762D002DD840/$file/MdmFunctionalReferenceArchitecture.pdf)

Porter, M. E. (1998). *Competitive Advantage, Creating and sustaining superior performance*. (F. Press, Ed.). New York.

Radcliffe, J. (2009). The Seven Building Blocks of MDM : A Framework for Success. *Gartner*, (May).

Seiler, K. P., Bodycombe, N. E., Hawkins, T., Shell, R., Lemon, A., De Souza, A., & Clemons, P. A. (2011). Master Data Management: Getting your House in Order. *Combinatorial chemistry high throughput screening*, (June). Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/21631416>

Silva, P., & Dias, M. (2002). *À descoberta dos seguros*. EDIDECO.

Teradata. (2010). How Data Governance Links Master Data Management e Data Quality. Retrieved from <http://www.teradata.com/assets/0/206/308/111d116f-fd36-401f-8c0e-73f46b4b8ad1.pdf>

Thomas H. Davenport, & Prusak, L. (1998). *Working knowledge: how organizations manage what they know*. (H. B. S. Press, Ed.) (pp. 1-6). Retrieved from <http://www.google.com/books?hl=pt-PT&lr=&id=QIyIWVhdYoYC&oi=fnd&pg=PR5&dq=to+m+davenport+knowledge&ots=Z2YJeoMWNS&sig=ECwzbAfI5axb5VQNfOHuPGbN8TY#v=onepage&q=to+m+davenport+knowledge&f=false>

Wang, R. Y., Strong, D. M., & Lee, Y. W. (1997). Data Quality In Context. *Communications of the ACM*, 40(5), 103-110.

Wolter, R., & Haselden, K. (2006). The What, Why, and How of Master Data Management. Retrieved September 1, 2011, from <http://msdn.microsoft.com/en-us/architecture/bb190163.aspx>

7 Anexos

7.1 Elementos chave no MDM

Para melhor compreensão de alguns elementos mencionados são descritos sucintamente alguns elementos chave de um *Master Data Management*.

7.1.1 Data Governance

Se uma organização vai dedicar tempo, esforço, recursos e dinheiro a implementar um *Master Data Management*, então faz todo o sentido que ponha em prática os processos necessários e adequados para gerir e proteger este investimento. A organização que não compreenda e não ponha em prática um sistema de *Data Governance*, não irá ter sucesso.

A definição de *data governance* não é clara existindo várias definições da mesma. Segundo o *Data Governance Institute* (Data Governance Institute, n.d.) o *data governance* é: “um sistema de decisões e direitos e responsabilidades para processos relacionados com informação, executada de acordo a modelos previamente acordados que descrevem quem pode determinar decisões com certa informação, e quando, e sobre quais circunstâncias, utilizando que métodos.” O próprio *Data Governance Institute* alerta para o facto de que *governance* pode ter significados distintos dependendo do contexto.

O conceito de *data governance* é definido pela *Data Management Association* (DAMA) como: “ O exercício de autoridade, controlo e partilha de decisões (planeamento, monitorização e aplicação) sobre a gestão dos dados. O *data governance* é o planeamento e controlo de alto nível sobre a gestão de dados”

O *data governance* para (Bersov & Dubov, 2007) é “um processo focado na gestão da qualidade, consistência, segurança e disponibilidade da informação. O *data governance* está assim intimamente ligado aos conceitos de *data ownership* e *stewardship*.”

Uma boa estratégia de *data governance* para além de alinhar o negócio com o IT na resolução de problema relacionados com dados, também define *data ownership* e outras políticas e processos de qualidade de dados, direitos de decisão e procedimentos. (Oracle, 2009)

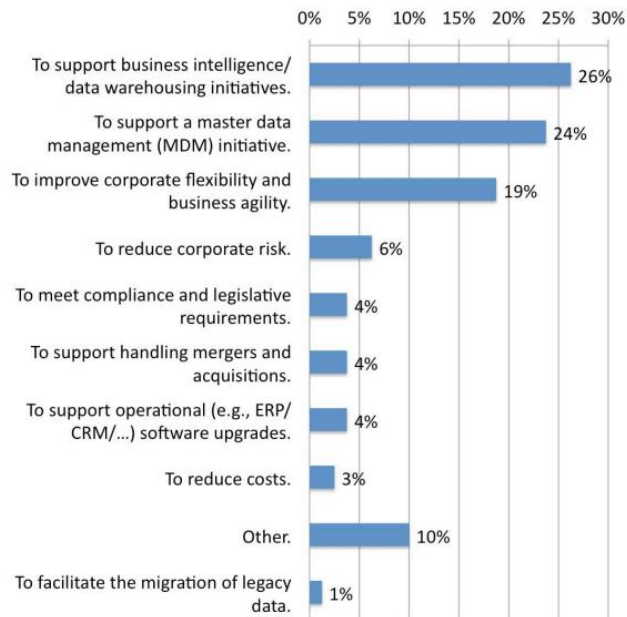


Figura 17 - Principais razões de implementação de *data governance* (Teradata, 2010)

A partir das várias definições fica claro que o *data governance* é uma parte essencial para qualquer iniciativa de criação de um *Hub* de dados. Um estudo da Teradata evidencia claramente esta relação directa entre o MDM e o *data governance*. Uma boa estratégia de *data governance* permite entregar informação apropriada, com a devida qualidade a utilizadores ou aplicações com as autorizações necessárias. O *Data Governance* abrange todo o ciclo de vida dos dados mestre desde a captura até à sua eliminação, incluindo a qualidade, segurança e acessibilidade. O *Data Governance* introduz um processo formal denominado "*stewardship*". (Radcliffe, 2009)

7.1.2 Data Quality

A constituição de um *Hub* central de dados mestre que vai ser utilizado por múltiplas aplicações, provoca uma dependência da qualidade destes dados. A fraca qualidade dos dados leva a informação imprecisa e incorrecta que pode prejudicar a organização nos seus processos negócio, na tomada de decisão e nas suas relações com os seus clientes e outras entidades. Por exemplo a morada incorrecta de um cliente pode provocar o envio errado da factura e do produto ou por exemplo o preço errado num produto pode revelar-se um desastre em termos de marketing. Para além disto por exemplo a regulamentação de alguns sectores como o *Sarbanes-Oxley* ou o *Basileia II* obrigam a certos parâmetros de qualidade nos dados.

A qualidade dos dados é consensualmente definida como "*fitness for use*"(Wang,

Strong, & Lee, 1997), (Ana Lucas, 2011). A qualidade não é uma característica intrínseca aos dados, como a qualquer outro objecto, só a poderemos aferir quando é utilizada ou aplicada e se revela adequada ao propósito de quem a solicitou.

Os dados de qualidade possuem certas características que determinam o seu “*fitness for use*”. Wang et al., 1997 agruparam estas características em determinadas categorias consoante a tabela 1.

Categoria	Características
Intrínseca	Correcção, objectividade, reputação e veracidade
Acessibilidade	Acessibilidade e segurança
Contextual	De valor acrescentado, relevante, disponível, complete, no volume apropriado, intemporal
Representacional	Interpretabilidade, compreensão, consistência e concisa

Tabela 2 - Categorias e características da qualidade dos dados (Wang et al., 1997)

7.1.3 Data Stewardship e Ownership

Os *Data Owners* são aqueles que estão na posição de obter, criar e possuem algum controlo sobre a informação. Os *Data Stewards* desempenham um papel diferente uma vez que não possuem informação nem têm controlo sobre ela. O seu papel é de assegurar que as normas, procedimentos e métricas de qualidade de dados acordadas são correctamente mantidas

7.2 Regras de Integração de dados

Para a criação de um único repositório de dados é necessário identificar regras a aplicar às várias fontes de dados de forma a normalizar, completar e agrupar os dados mestre.

7.2.1 Atributos de dados de identificação de Entidades

7.2.1.1 Nome

Regra Nº RI001	
Atributos Envolvidos	<p>NOME</p> <p>Fontes Aplicáveis</p> <ul style="list-style-type: none"> Clientes Agentes Aderentes - Vida Grupo Centros de Facturação Pessoas Seguras – Vida Individual Procuradores e Terceiros Beneficiários

Universo Aplicável	Todos os registos de cada uma das fontes acima indicadas
Descritivo das Funcionalidades	<p>O processo actuará nas seguintes vertentes:</p> <ul style="list-style-type: none"> - Conversão de letras minúsculas para maiúsculas; - Conversão de letras acentuadas pelas suas correspondentes não acentuadas. - Remoção dos caracteres identificados no ponto - Remoção de partículas de Ligação correspondentes às palavras DE, DA, DAS e DOS. Remoção da palavra DO em todos os casos excepto quando estamos na presença das palavras “DO Ó” ou “DO O” (Correspondente não acentuado). - Remoção dos Pronomes de Tratamento / Axiónimo, títulos Académicos e Honoríficos identificados
Precedências	

Tabela 3 - Regra Nº RI001 - Normalização do Nome da Entidade

Regra Nº RI002			
Atributos Envolvidos	NOME	Fontes Aplicáveis	Clientes Agentes Aderentes - Vida Grupo Centros de Facturação Pessoas Seguras – Vida Individual Procuradores e Terceiros Beneficiários
Universo Aplicável	Todos os registos de cada uma das fontes acima indicadas		
Descritivo de funcionalidades	<p>Analisar o nome para detectar a presença de erros ortográficos de escrita. A detecção será feita com recurso a um dicionário de erros ortográficos comuns para nomes Portugueses disponibilizado pela ferramenta utilizada.</p> <p>Se for detectada uma palavra que apresente erros ortográficos, será feita a correcção automática da palavra errada no nome da entidade, com base no Dicionário indicado.</p> <p>A regra <u>não considerará os caracteres latinos</u>, de acordo com a estratégia de migração de dados para o. Como tal, não serão considerados como erros ortográficos de escrita os casos em que o(s) nome(s) se apresentem substituídos dos referidos caracteres (JOSE em vez de JOSÉ, GRACA em vez de GRAÇA, etc.). A mesma premissa será aplicada ao resultado da regra, ou seja, as eventuais correcções ao nome serão disponibilizadas com as respectivas substituições, De acordo com a regra RI0001 acima.</p>		
Precedências:	RI0001		

Tabela 4- Regra Nº RI002 - Correcção de Erros Ortográficos no nome de Entidade

7.2.1.2 Sexo

Regra Nº RI003			
Atributos Envolvidos	SEXO – Género	Fontes Aplicáveis	Pessoas Seguras – Vida Individual Procuradores e Terceiros
Universo Aplicável	Todos os registos das fontes acima indicadas, em que o valor do campo Sexo seja igual a '0' (Não preenchido / Não especificado)		
Descritivo das Funcionalidades	Perante a presença de valor não preenchido (valor 'zero') neste atributo, o processo executará as seguintes funções:		

	<p>- Caso o nome de entidade possua palavras e padrão típicos de um nome de empresa, será atribuído o valor 3 (três) para este campo. A detecção de palavras típicas de uma entidade colectiva será feita com recurso a um dicionário correspondente existente na ferramenta utilizada;</p> <p>- Caso a entidade não tenha padrão típico de empresa, será analisado o 1º nome da entidade, tentando determinar-se o género da mesma, com base em dicionário existente na ferramenta. O processo atribuirá, sempre que o determinar, os seguintes valores:</p> <ul style="list-style-type: none"> • Se a entidade tiver 1º nome Feminino, será atribuído valor 'F'; • De forma semelhante, quando o processo conduzir que a entidade é Masculina, o campo deverá conter o valor 'M'. • Caso o processo não consiga apurar o género, o campo permanecerá com o valor 0 (Não preenchido / Não especificado).
Precedências:	RI004

Tabela 5 - Regra Nº RI003 - Determinação do valor do atributo 'Sexo' para não preenchidos

Regra Nº RI004				
Atributos Envolvidos	<table border="1"> <tr> <td>SEXO – Género</td> <td>Fontes Aplicáveis</td> <td>Pessoas seguras - Vida individual Procuradores e terceiros Aderentes – Vida Grupo Clientes Agentes</td> </tr> </table>	SEXO – Género	Fontes Aplicáveis	Pessoas seguras - Vida individual Procuradores e terceiros Aderentes – Vida Grupo Clientes Agentes
SEXO – Género	Fontes Aplicáveis	Pessoas seguras - Vida individual Procuradores e terceiros Aderentes – Vida Grupo Clientes Agentes		
Universo Aplicável	Todos os registos das fontes acima indicadas.			
Descritivo das Funcionalidades	<p>O processo identificará o lote de registos com valor inválido neste atributo, de acordo com a seguinte regra:</p> <ul style="list-style-type: none"> • Para as fontes 'Pessoas seguras - Vida individual', 'Procuradores e terceiros' e 'Aderentes – Vida Grupo' marcará os registos com valor diferente de 0, 1 e 2; • Para as fontes 'Clientes' e 'Agentes', marcará os registos com valor diferente de E, F ou M. <p>Não será feito qualquer tratamento adicional a estes casos, apenas serão identificados, devendo ser fornecidas listagens, se solidadas.</p>			
Precedências:				

Tabela 6 - Regra Nº RI004- SEXO – Identificação de valores fora do domínio esperado

Regra Nº RI005				
Atributos Envolvidos	<table border="1"> <tr> <td>SEXO - Género</td> <td>Fontes Aplicáveis</td> <td>Pessoas seguras - Vida individual Procuradores e terceiros Clientes Agentes</td> </tr> </table>	SEXO - Género	Fontes Aplicáveis	Pessoas seguras - Vida individual Procuradores e terceiros Clientes Agentes
SEXO - Género	Fontes Aplicáveis	Pessoas seguras - Vida individual Procuradores e terceiros Clientes Agentes		
Universo Aplicável	<p>- Para as fontes 'Pessoas seguras - Vida individual' e 'Procuradores e terceiros', todos os registos contendo os valores 1 e 2 neste campo;</p> <p>Para as fontes 'Clientes' e 'Agentes', todos os registos contendo os valores M e F neste campo</p>			
Descritivo das Funcionalidades	<p>O processo da analisará a correlação entre o 1º Nome da entidade e o valor presente no campo Sexo, detectando inconsistências entre os mesmos. Será analisado o 1º nome da entidade, tentando determinar-se o género da mesma, com base em</p>			

	dicionário existente na ferramenta. O processo decidirá com base na seguinte tabela:									
	<table border="1"> <thead> <tr> <th>Valor encontrado no campo</th> <th>Valor determinado pela análise do 1º nome</th> <th>Valor a corrigir</th> </tr> </thead> <tbody> <tr> <td>1 ou M</td> <td>F</td> <td>F</td> </tr> <tr> <td>2 ou F</td> <td>M</td> <td>M</td> </tr> </tbody> </table>	Valor encontrado no campo	Valor determinado pela análise do 1º nome	Valor a corrigir	1 ou M	F	F	2 ou F	M	M
Valor encontrado no campo	Valor determinado pela análise do 1º nome	Valor a corrigir								
1 ou M	F	F								
2 ou F	M	M								
Observações	Ficará excluído desta regra o valor 'E' de <u>Empresa</u> , que será alvo de um tratamento posterior; - Será também excluda a fonte <u>Aderentes - Vida Grupo</u> , por a eventual alteração do campo Sexo poder ter implicações no Prémio dos produtos Vida.									
Precedências:	RI004, RI003									

Tabela 7 – Regra Nº RI005- Tratamento de inconsistências entre o género e o 1º Nome de Entidade

7.2.1.3 Número de Identificação Fiscal

Regra Nº RI006													
Atributos Envolvidos	Número de Identificação Fiscal	Fontes Aplicáveis	Clientes Agentes Aderentes - Vida Grupo Centros de Facturação Pessoas Seguras – Vida Individual Procuradores e Terceiros Beneficiários										
Universo Aplicável	Todos os registos, das fontes indicadas.												
Descritivo das Funcionalidades	<p>O Processo deverá validar o número de identificação Fiscal, pela aplicação do algoritmo público de <i>checkdigit</i> para o NIF Português. O processo agirá com base nas seguintes regras:</p> <table border="1"> <thead> <tr> <th>Situação</th> <th>NIF Corrigido</th> </tr> </thead> <tbody> <tr> <td>NIF original vazio ou 0000000</td> <td rowspan="5">111111111</td> </tr> <tr> <td>NIF original igual a 123456789 ou 5000000000</td> </tr> <tr> <td>NIF Dummie, conforme lista presente no Anexo 2 - Números de Identificação Fiscal "Dummies"</td> </tr> <tr> <td>NIF original com checkdigit inválido</td> </tr> <tr> <td>NIF original cujo 1º dígito seja 8</td> </tr> <tr> <td>NIF com checkdigit válido ou NIF igual a 111111111</td> <td>Não efectua qualquer correcção</td> </tr> </tbody> </table>			Situação	NIF Corrigido	NIF original vazio ou 0000000	111111111	NIF original igual a 123456789 ou 5000000000	NIF Dummie , conforme lista presente no Anexo 2 - Números de Identificação Fiscal " Dummies "	NIF original com checkdigit inválido	NIF original cujo 1º dígito seja 8	NIF com checkdigit válido ou NIF igual a 111111111	Não efectua qualquer correcção
Situação	NIF Corrigido												
NIF original vazio ou 0000000	111111111												
NIF original igual a 123456789 ou 5000000000													
NIF Dummie , conforme lista presente no Anexo 2 - Números de Identificação Fiscal " Dummies "													
NIF original com checkdigit inválido													
NIF original cujo 1º dígito seja 8													
NIF com checkdigit válido ou NIF igual a 111111111	Não efectua qualquer correcção												
Obs.:	Os valores válidos de alta frequência (>1) devem ser analisados e corrigidos manualmente pois são também potencialmente inválidos.												
Precedências:													

Tabela 8 - Regra Nº RI006 - Validação e correcção de NIF's Portugueses

Regra Nº RI007			
Atributos Envolvidos	Número de Identificação Fiscal	Fontes Aplicáveis	Clientes Agentes
Universo	Todos os registos, das fontes indicadas que se encontrem classificados		

Aplicável	como Empresa (atributo SEXO com o valor E), cujos NIF's possuam <i>checkdigit válido</i> e começado pelo dígito 1 ou 2 . Excluem-se os casos cujo NIF tenha o valor 11111111 .									
Descritivo das Funcionalidades	<p>Será implementado um processo com o objectivo de validar a consistência entre a natureza das entidades classificadas como colectivas e o 1º dígito do seu NIF. Sabemos que as entidades colectivas deverão possuir NIF começado por 5 ou 6. O processo actua sobre os casos classificados como Empresa (SEXO='E') e cujo NIF seja válido e comece por 1 ou 2. Nestes casos, executam-se a seguinte validação:</p> <ul style="list-style-type: none"> • Validação 1 - Valida-se se o nome de entidade possuir palavras e padrão típicos de um nome de empresa, de forma a verificar a referida coerência entre campos. A detecção de palavras típicas de uma entidade colectiva será feita com recurso a um dicionário correspondente existente na ferramenta da utilizada no projecto. • Validação 2 - Se o nome não possuir palavras típicas de empresa, tenta identificar o género com base no 1º nome da Entidade, utilizando o dicionário existente na ferramenta. <p>Mediante o resultado da validação, será aplicada a seguinte regra de correcção:</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr style="background-color: #4F81BD; color: white;"> <th>Resultado da validação</th> <th>NIF corrigido</th> <th>SEXO corrigido</th> </tr> </thead> <tbody> <tr style="background-color: #D3D3D3;"> <td>A entidade possui palavras típicas de entidade colectiva</td> <td style="text-align: center;">11111111</td> <td>Não altera o valor do atributo</td> </tr> <tr> <td>A entidade não possui palavras típicas de entidade colectiva</td> <td style="text-align: center;">Não altera o valor ao atributo</td> <td>F ou M ou O (Feminino, Masculino, Não apurado, respectivamente), mediante o resultado da validação 2 descrita acima.</td> </tr> </tbody> </table>	Resultado da validação	NIF corrigido	SEXO corrigido	A entidade possui palavras típicas de entidade colectiva	11111111	Não altera o valor do atributo	A entidade não possui palavras típicas de entidade colectiva	Não altera o valor ao atributo	F ou M ou O (Feminino, Masculino, Não apurado, respectivamente), mediante o resultado da validação 2 descrita acima.
Resultado da validação	NIF corrigido	SEXO corrigido								
A entidade possui palavras típicas de entidade colectiva	11111111	Não altera o valor do atributo								
A entidade não possui palavras típicas de entidade colectiva	Não altera o valor ao atributo	F ou M ou O (Feminino, Masculino, Não apurado, respectivamente), mediante o resultado da validação 2 descrita acima.								
Precedências:	RI006									

Tabela 1 – Regra Nº RI007 - Empresas – Validação da coerência com o 1º dígito do NIF

7.2.1.4 Data de Nascimento

Regra Nº RI008			
Atributos Envolvidos	Data de Nascimento	Fontes Aplicáveis	Pessoas Seguras - Vida individual Beneficiários Procuradores e terceiros Aderentes – Vida Grupo Clientes Agentes
Universo Aplicável	Todos os registos das fontes acima indicadas cujo atributo se encontre preenchido.		
Descritivo das Funcionalidades	O processo analisará o conteúdo do campo no formato AAAAMMDD, e aplicará as seguintes duas regras de limpeza:		

	<ul style="list-style-type: none"> • Situação 1- O conteúdo do campo não representa uma data válida, isto é, o ano, mês e dia não existem como data de calendário. Neste caso, o processo substituirá o conteúdo pelo valor 00000000 (zero); • Situação 2 – O conteúdo do campo é válido (representa uma data válida) mas o seu valor é superior à data de execução do processo (datas no futuro). Neste caso, o processo substituirá o conteúdo pelo valor 00000000 (zero).
Precedências	

Tabela 9 - Regra Nº RI008 - Data de Nascimento – Limpeza de inválidos

7.2.2 Atributos destinados ao armazenamento de Contactos

7.2.2.1 Email

Regra Nº RC001			
Atributos Envolvidos	EMAIL	Universo Aplicável	Aderentes – Vida Grupo Clientes
Universo Aplicável	Todos os registos das fontes indicadas com Email preenchido.		
Descritivo das Funcionalidades	<p>O processo deverá detectar e isolar as situações cujo conteúdo presente no campo não representa um email válido, aplicando as regras internacionais de constituição de um endereço de Email.</p> <p>Todos os registos inválidos deverão ser identificados. Não ficam, porém, identificadas quaisquer operações de limpeza sobre estes casos.</p> <p>Os casos poderão ser fornecidos para análise e identificação dos procedimentos de correcção.</p>		
Precedências:			

Tabela 10 - Regra Nº RC001- Email – Identificação de Inválidos

7.2.3 Atributos de Morada associados à Entidade

7.2.3.1 Moradas

Regra Nº RM001			
Atributos Envolvidos	Vários campos destinados ao armazenamento de morada e Código de País	Universo Aplicável	Pessoas Seguras (Vida Individual) Moradas de Correspondência (Vida Individual) Moradas de Correspondência (Vida Grupo) Beneficiários Procuradores e Terceiros Aderentes (Vida Grupo) Centros de Facturação Clientes Pessoas (Agentes)
Universo	Todos os registos de morada das 9 fontes indicadas cujo atributo 'código		

Aplicável	de País' possua valor diferente de 0 (Portugal).						
Descritivo	<p>O processo analisará os campos destinados ao armazenamento da morada, procurando a presença de palavras típicas de arruamentos Portugueses. Esta detecção utilizará um dicionário de palavras a construir durante a implementação desta regra. Será aplicada a seguinte regra de limpeza:</p> <table border="1"> <thead> <tr> <th>Resultado da validação</th> <th>Código de País corrigido</th> </tr> </thead> <tbody> <tr> <td>A morada possui palavras típicas de arruamentos Portugueses.</td> <td>0 (Portugal)</td> </tr> <tr> <td>A morada não possui palavras típicas de arruamentos Portugueses.</td> <td>Não altera o valor ao atributo</td> </tr> </tbody> </table>	Resultado da validação	Código de País corrigido	A morada possui palavras típicas de arruamentos Portugueses.	0 (Portugal)	A morada não possui palavras típicas de arruamentos Portugueses.	Não altera o valor ao atributo
Resultado da validação	Código de País corrigido						
A morada possui palavras típicas de arruamentos Portugueses.	0 (Portugal)						
A morada não possui palavras típicas de arruamentos Portugueses.	Não altera o valor ao atributo						
Precedências:							

Tabela 11 - Regra Nº RM001- Estrangeiras – Correção do Código de País para moradas Portuguesas

Regra Nº RM002	
Atributos Envolvidos	<p>Vários campos destinados ao armazenamento de morada e Código de País</p> <p>Universo Aplicável</p> <p>Pessoas Seguras (Vida Individual) Moradas de Correspondência (Vida Individual) Moradas de Correspondência (Vida Grupo) Beneficiários Procuradores e Terceiros Aderentes (Vida Grupo) Centros de Facturação Clientes Pessoas (Agentes)</p>
Universo Aplicável	Todos os registos das fontes indicadas cujo atributo ' código de País' possua o valor 0 (Portugal).
Descritivo	<p>O processo fará uso de um componente específico para análise de moradas Portuguesas (Address Validator) da ferramenta. Será submetido ao processo o universo acima identificado e tipificadas, como resultado, as seguintes 4 situações:</p> <ul style="list-style-type: none"> • Situação 1 - Morada, CP e localidade dos dados preenchidos. O Código Postal obtido é igual ao Código Postal presente nos dados, sendo elevado o grau de certeza sobre a validade dos elementos da morada. • Situação 2 - Morada, Código Postal e localidade preenchidos. O Código Postal obtido é diferente do Código Postal presente nos dados, pelo que o CP presente nos dados deverá estar desactualizado. • Situação 3 - Morada, Código Postal e localidade preenchidos. O processo não validou todos os componentes da morada contra a fonte de referência. O Código Postal obtido é igual ao Código Postal presente nos dados. O grau de certeza da comparação é menor, mas, provavelmente, o código postal presente nos dados é

	<p>o correcto.</p> <ul style="list-style-type: none"> • Situação 4 – Elementos da Morada preenchidos. Porém, o processo não encontrou a respectiva morada na fonte de referência, não conseguindo atribuir Código Postal e, conseqüentemente, validar a morada. <p>O processo fará as seguintes correcções ao código postal e respectivo descritivo postal, com base nas situações acima descritas:</p>										
	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr style="background-color: #333; color: white;"> <th>Código Postal corrigido</th> <th>Localidade Postal corrigida</th> </tr> </thead> <tbody> <tr style="background-color: #ccc;"> <td>Não efectua alterações</td> <td>Não efectua alterações</td> </tr> <tr> <td>Substitui o CP pelo valor apurado pelo processo</td> <td>Substitui a localidade Postal pelo valor apurado pelo processo</td> </tr> <tr style="background-color: #ccc;"> <td>Não efectua alterações</td> <td>Não efectua alterações</td> </tr> <tr style="background-color: #ccc;"> <td>Não efectua alterações</td> <td>Não efectua alterações</td> </tr> </tbody> </table>	Código Postal corrigido	Localidade Postal corrigida	Não efectua alterações	Não efectua alterações	Substitui o CP pelo valor apurado pelo processo	Substitui a localidade Postal pelo valor apurado pelo processo	Não efectua alterações	Não efectua alterações	Não efectua alterações	Não efectua alterações
	Código Postal corrigido	Localidade Postal corrigida									
	Não efectua alterações	Não efectua alterações									
	Substitui o CP pelo valor apurado pelo processo	Substitui a localidade Postal pelo valor apurado pelo processo									
Não efectua alterações	Não efectua alterações										
Não efectua alterações	Não efectua alterações										
<p>No processo a desenvolver para o efeito deverá ser melhorado o <i>parsing</i> de componentes da morada, na tentativa de minimizar o número de casos que, após validação, cairão na situação 4 acima descrita.</p>											
<p>Precedências: RM001</p>											

Tabela 12 Regra Nº RM002- Moradas Portuguesas - validação do Código Postal