



Instituto Superior de Economia e Gestão

UNIVERSIDADE TÉCNICA DE LISBOA

DESDE 1911

MESTRADO
GESTÃO DE SISTEMAS DE INFORMAÇÃO

TRABALHO FINAL DE MESTRADO
DISSERTAÇÃO

*FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS
DE
MASTER DATA MANAGEMENT*

RUI JORGE RAPOSO E SILVA

SETEMBRO - 2012



Instituto Superior de Economia e Gestão

UNIVERSIDADE TÉCNICA DE LISBOA

DESDE 1911

MESTRADO EM GESTÃO DE SISTEMAS DE INFORMAÇÃO

TRABALHO FINAL DE MESTRADO DISSERTAÇÃO

*FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS
DE
MASTER DATA MANAGEMENT*

RUI JORGE RAPOSO E SILVA

ORIENTAÇÃO:

PROF. ENG^a. ANA MARIA MARQUES RIBEIRO DOS SANTOS LUCAS

SETEMBRO - 2012

Agradecimentos

O presente trabalho de dissertação não é apenas fruto de um esforço individual, resultando da combinação de diversos contributos de várias pessoas. Por essa razão, gostaria de aqui deixar um sincero agradecimento a todas as pessoas que, direta ou indiretamente, contribuíram para a realização deste trabalho.

Um agradecimento especial à professora Ana Lucas por ter aceitado a orientação do meu Trabalho Final de Mestrado, através do qual espero conseguir retribuir a confiança depositada. Reconheço especialmente a sua inesgotável paciência e capacidade para gerir as dúvidas, apontar o melhor caminho em cada situação e manter elevada a motivação.

Ao Eng.º Manuel Dias, da Microsoft Portugal, pelo tempo que disponibilizou e por todos os contributos e ideias que propôs para o desenvolvimento e consolidação do trabalho.

Ao Professor Doutor Palma dos Reis pela valiosa contribuição que forneceu na fase de avaliação e validação do modelo proposto para a *framework* e pela sua perspicaz visão analítica.

Ao Eng.º Bruno Marques, da Portugal Telecom, por ter disponibilizado toda a sua experiência e visão pragmática sobre as temáticas do *Master Data Management* e todos os seus comentários sobre a *framework* proposta no trabalho de dissertação.

Aos colegas do curso de mestrado por todas as dicas e sugestões e pela solidariedade demonstrada ao longo desta nossa etapa final do curso.

Símbolos, Abreviaturas, Siglas e Convenções

B2B	<i>Business to Business</i>
B2C	<i>Business to Consumer</i>
BPM	<i>Business Process Management</i>
CDI	<i>Customer Data Integration</i>
CDC	<i>Change Data Capture</i>
CPD	<i>Centro de Processamento de Dados</i>
CRM	<i>Customer Relationship Management</i>
DG	<i>Data Governance</i>
DQ	<i>Data Quality</i>
DS	<i>Design Science</i>
EAI	<i>Enterprise Application Integration</i>
ECM	<i>Enterprise Content Management</i>
EDR	<i>Enterprise Data Replication</i>
EII	<i>Enterprise Information Integration</i>
EIM	<i>Employee Information Management</i>
ERP	<i>Enterprise Resource Planning</i>
ETL	<i>Extract, Transform and Load</i>
FTP	<i>File Transfer Protocol</i>
HTML	<i>Hypertext Markup Language</i>
HTTP	<i>Hypertext Transfer Protocol</i>
IEEE-LOM	<i>Learning Object Metadata</i>
JEE	<i>Java Enterprise Edition</i>
JMS	<i>Java Message Service</i>
KPI	<i>Key Performance Indicator</i>
MARC21	<i>Machine Readable Cataloging</i>
MDM	<i>Master Data Management</i>
METS	<i>Metadata Encoding & Transmission Standard</i>
MODS	<i>Metadata Object Description Schema</i>
ONIX	<i>Online Information Exchange</i>
PIM	<i>Product Information Management</i>
RDF/XML	<i>Resource Description Framework</i>

ROI	<i>Return on Investment</i>
SaaS	<i>Software as a Service</i>
SGBD	Sistema de Gestão de Bases de Dados
SLA	<i>Service Level Agreement</i>
SO	Sistema Operativo
SOA	<i>Service-oriented Architecture</i>
TCO	<i>Total Cost of Ownership</i>
TI	Tecnologias de Informação
VIM	<i>Vendor Information Management</i>
XHTML	<i>Extensible Hypertext Markup Language</i>
XML	<i>Extensible Markup Language</i>

Resumo

Existe atualmente no mercado de *software* uma grande variedade de aplicações de *Master Data Management* e tem vindo a ser conduzido um processo de consolidação entre os fabricantes daquelas ferramentas. Um dos grandes desafios, no que diz respeito à tecnologia para o *Master Data Management*, está relacionado com o seu grande espectro de aplicabilidade e utilização. Esta situação configura um problema para as equipas de Tecnologias de Informação na avaliação e pesquisa da ferramenta de *Master Data Management* mais adequada para a sua organização.

O presente trabalho de dissertação tem como objetivo determinar quais as características mais significativas a considerar na avaliação de uma ferramenta de *Master Data Management*, tendo em vista a construção de uma *framework* que permita a avaliação qualitativa daquele tipo de *software*. Para a concretização desse objetivo, foi realizado um estudo orientado por uma metodologia baseada no paradigma do *Design Science*, tendo sido executada apenas uma iteração do processo.

Não obstante ter sido possível elencar um conjunto de vinte e duas características fundamentais das aplicações de *Master Data Management* com base na literatura disponível sobre o tema e, a partir delas, conceptualizar uma *framework* de avaliação, o estudo realizado aponta para a necessidade de introdução de melhorias no modelo construído, quer ao nível do detalhe dos parâmetros de avaliação e respetivos valores possíveis, quer ao nível do âmbito de análise da própria *framework*. O processo de validação da *framework* proposta sugere ainda a construção de uma matriz que permita uma aplicação mais facilitada do modelo e o aumento da abrangência da *framework* de forma a contemplar outros fatores de análise, tais como a vertente financeira, TCO, ROI, perfil do fabricante, portfólio de serviços e a análise de risco.

PALAVRAS-CHAVE: *Framework, Master Data Management, MDM, Master Data, Software MDM.*

Abstract

Currently, it exists in the software market a wide variety of Master Data Management applications and it has been conducted a process of consolidation among the manufacturers of those tools. One of the major challenges with regard to technology for Master Data Management is related to its wide range of applicability and use. This sets up a problem for the IT teams in evaluation and research of most appropriate Master Data Management tool for the organization.

This dissertation work aims to determine the most important features to consider in evaluating a Master Data Management tool and build a framework with these features that allow the qualitative assessment of that kind of software. To achieve this goal, a study guided by a methodology based on the paradigm of Design Science was performed, considering only a single iteration of the process.

It was possible to list a set of twenty-two fundamental characteristics of Master Data Management applications, based on the available literature on the subject and, with them, was conceptualized an evaluation framework. However, the study points to the need to introduce improvements in the model built, both at the level of detail of the assessment parameters and their possible values, and in the scope of analysis of the framework itself. The validation process of the proposed framework also suggests building a matrix that allows an easier application of the model and increasing the scope of the framework considering other analysis factors such as the financial slope, TCO, ROI, manufacturer's profile, services portfolio and risk analysis.

KEYWORDS: *Framework, Master Data Management, MDM, Master Data, MDM Software.*

Índice Geral

1. Introdução	1
1.1. <i>Objetivos do Estudo</i>	2
1.2. <i>Contexto do Estudo</i>	2
1.3. <i>Questão de Investigação</i>	3
1.4. <i>Relevância do Estudo</i>	3
2. Revisão da Literatura	4
2.1. <i>Master Data Management</i>	4
2.1.1. <i>Estilos de Implementação MDM</i>	6
2.2. <i>Abordagens de MDM</i>	8
2.3. <i>Ferramentas de MDM</i>	9
2.3.1. <i>O que procurar numa Plataforma MDM</i>	11
2.3.2. <i>O Mercado de Ferramentas MDM</i>	12
3. Metodologia de Investigação	15
3.1. <i>A Metodologia</i>	15
3.2. <i>Desenho da Investigação</i>	17
4. A Framework	19
4.1. <i>Características Desejáveis do Software de MDM</i>	19
4.2. <i>Construção da Framework</i>	26
4.3. <i>Avaliação de Software de MDM</i>	32
4.4. <i>Avaliação da Framework</i>	33
5. Conclusões e Investigação Futura	35
5.1. <i>Conclusões</i>	35
5.2. <i>Investigação Futura</i>	35
6. Referências Bibliográficas	36
ANEXO A - Framework: os valores possíveis	40
ANEXO B - Avaliação de Ferramentas de MDM	44

Índice de Tabelas

Tabela I - Conceito de <i>Master Data</i>	5
Tabela II - Conceito de <i>Master Data Management</i>	6
Tabela III - Estilos de Implementação MDM.	8
Tabela IV - Características das Ferramentas de MDM.	21
Tabela V - <i>Framework</i> para a Avaliação de Ferramentas de MDM.	27
Tabela VI - Referências utilizadas na avaliação das ferramentas MDM.	32
<i>Tabela VII - Especialistas convidados.</i>	33
<i>Tabela VIII - Avaliação da Framework.</i>	34
Tabela IX - <i>Framework</i> : os valores possíveis.	40

Índice de Figuras

Figura 1 - Capacidades Técnicas do MDM.	10
Figura 2 - Evolução das Soluções de <i>Master Data Management</i>	13
Figura 3 - <i>Framework</i> para a Investigação em Sistemas de Informação.	16

1 Introdução

Nos dias de hoje as organizações lidam com informação heterogénea que necessita de ser gerida e mantida. Dada a estrutura departamental de muitas organizações, cada departamento gere os seus próprios dados. As organizações possuem os seus dados distribuídos em diversos sistemas, pelo que, esta situação origina frequentemente dados inconsistentes, falta de informação, dados duplicados e não normalizados. A existência destes problemas de qualidade dos dados pode conduzir a tomadas de decisão pouco eficientes e, conseqüentemente, a organização pode consumir recursos desnecessariamente ou até mesmo ver reduzidos os seus rendimentos (Galhardas, Torres, & Damásio, 2010).

Desta forma, no caso particular das ferramentas de *Master Data Management* (MDM), o objetivo é, claramente, gerir a informação fundamental para o negócio da organização, garantindo a qualidade dos dados, a sua integridade e consistência (Otto & Ebner, 2010). O MDM representa um dos temas mais atuais na área dos Sistemas de Informação (Moss, 2007). A gestão de *Master Data*, tal como foi brevemente introduzida, é muito promissora para ultrapassar todos os obstáculos referidos (Cleven & Wortmann, 2010).

O entendimento do âmbito da justificação, desenho, planeamento, implementação e manutenção de uma aplicação de MDM é, em si mesmo, um investimento significativo (Loshin, 2009). O investimento em *software* de MDM tem de, pelas razões apresentadas, ser devidamente ponderado, respeitando desejavelmente os fatores críticos de sucesso para o MDM.

Atualmente, o estado da arte em soluções MDM é reclamado por empresas de *software* como a IBM e a Oracle, que prometem dados de qualidade para a tomada de decisão. A *Master Data* de uma organização constitui o seu ativo mais crítico e cai geralmente em quatro grupos: pessoas, coisas, locais e conceitos. No sistema de informação da organização, a *Master Data* está armazenada em diferentes sistemas como o CRM e ERP (Wang, Ming, & You, 2009).

Tendo em vista o investimento adequado em *software* MDM para uma determinada realidade de uma organização, é necessário, desde logo, o recurso a instrumentos de avaliação de ferramentas MDM que permitam a seleção da aplicação que melhor se adequa a essa realidade, através de uma comparação razoável entre as diversas ferramentas disponíveis no mercado. Para que esses instrumentos de avaliação sejam disponibilizados é desejável e útil dar resposta às questões: (1) Quais as características

mais significativas de uma ferramenta MDM? e (2) De que forma deverá ser operacionalizada a análise e avaliação de ferramentas MDM?

O presente trabalho de investigação propõe-se precisamente criar uma *framework* para a comparação de ferramentas MDM, que permita avaliar o funcionamento e arquitetura de uma solução face às características desejáveis de uma ferramenta dessa natureza e face à realidade da organização onde será eventualmente implementada.

1.1. Objetivos do Estudo

Os objetivos desta investigação consistem na construção de uma *framework* que permita a avaliação de ferramentas de *Master Data Management* (MDM). A investigação deverá reunir um conjunto de características e fatores para uma avaliação eficaz de ferramentas MDM, bem como a sistematização da forma como deverá ser feita essa avaliação, delineando claramente uma ferramenta metodológica. O modelo a construir será testado em campo através da análise e avaliação de duas ferramentas de MDM atualmente existentes no mercado.

Assim, o estudo pode ser resumido em quatro grandes fases:

1. Construção de um referencial teórico através da revisão da literatura existente sobre o objeto do estudo;
2. Definição de uma *framework*, caracterizando, justificadamente, todas as suas componentes;
3. Avaliação de duas ferramentas de MDM disponíveis no mercado, utilizando a *framework* construída na fase anterior;
4. Avaliação da *framework* construída através de entrevistas efetuadas a especialistas nesta matéria.

1.2. Contexto do Estudo

“As organizações, principalmente nas economias mais evoluídas, investem elevados e crescentes recursos financeiros em Sistemas e Tecnologias de Informação na perspetiva de melhorar a sua eficiência e eficácia e fazer face a um contexto em permanente evolução, onde a competição é crescente e os objetivos organizacionais cada vez mais exigentes” (Serrano & Caldeira, 2001, p.99).

Num ambiente empresarial competitivo, os gestores e decisores precisam de tomar decisões informadas e devem possuir métodos e ferramentas que lhes permitam

realizar uma efetiva análise e gestão de benefícios, no que diz respeito a investimentos em Sistemas de Informação.

No caso particular das ferramentas de MDM, qualquer investimento nesta área exige uma análise cuidada e ponderada, uma vez que se trata de sistemas que vão gerir e manipular a informação mais sensível e importante da organização, que está intimamente associada a decisões estratégicas.

Uma aplicação MDM é concebida para suportar as necessidades de negócio de uma organização, providenciando acesso a perspetivas consistentes das entidades de *Master Data* univocamente identificáveis na infraestrutura da aplicação operacional (Loshin, 2009).

1.3. Questão de Investigação

O presente trabalho de investigação tem como objetivo dar resposta à seguinte questão de investigação:

Quais as características mais significativas a considerar na avaliação de uma ferramenta de *Master Data Management*?

1.4. Relevância do Estudo

Nas diversas pesquisas efetuadas à literatura científica existente sobre o *Master Data Management*, não foi possível até ao momento encontrar publicações que se dedicassem à presente questão de investigação.

Desta forma, o estudo pretende:

- Preencher o *gap* identificado na literatura;
- Apresentar uma solução para o problema de investigação;
- Contribuir para o aumento do conhecimento na área em estudo;
- Fornecer uma ferramenta conceptual que auxilie gestores, profissionais de TI e fabricantes na avaliação de ferramentas de MDM.

2 Revisão da Literatura

A revisão da literatura é fundamental para qualquer projeto académico (Webster & Watson, 2002). Nesta secção pretende-se apresentar o “estado da arte” relativamente à temática do MDM e elencar de forma sistematizada os fatores críticos de sucesso para uma implementação de uma ferramenta de MDM. Pretende-se igualmente apresentar os principais conceitos na perspetiva dos autores mais relevantes, relativamente ao tema em investigação.

2.1. *Master Data Management*

O ambiente empresarial atual obriga as organizações a encontrarem formas de se diferenciarem em relação à concorrência e prosperar face à pressão de alcançarem o sucesso (Cervo & Allen, 2011). O forte ambiente competitivo constitui um *driver* para a reorganização de processos, pessoas e tecnologias de informação para que as empresas consigam atingir níveis de competitividade e performance que lhes permitam sobreviver e prosperar.

A transição entre aplicações desenhadas para atender a necessidades de negócio das empresas deve assentar numa visão consistente de dados, construída a partir de várias fontes (Loshin, 2009).

Tendo em vista a compreensão do significado de *Master Data Management* e da sua dinâmica, é necessário antes de mais introduzir o conceito de *master data*. Encontram-se sistematizadas na Tabela I quatro definições propostas para este conceito.

Durante muitos anos ocorreu uma grande proliferação de dados devido ao armazenamento relativamente barato e aumento de digitalização de informação. Por outro lado, as soluções compartimentalizadas contribuíram para esta fragmentação da informação de forma considerável, incrementando a duplicação de dados e ausência de uma identificação de uma identidade comum.

As organizações chegaram à conclusão que a forma mais eficaz de resolver este problema crescente é a criação de uma abordagem de fonte única para a gestão da *master data* baseada em altos padrões de qualidade e governança que sirva completamente o negócio (Cervo & Allen, 2011).

Em cada instância, o principal objetivo é a criação de uma visão sincronizada e consistente das entidades do *core business* da organização (Loshin, 2009).

Tabela I – Conceito de *Master Data*.

Definição de <i>Master Data</i>	Fonte
Dados que representam a informação mais valiosa que uma organização possui. Representam a informação fundamental acerca de um determinado negócio tal como clientes, fornecedores, produtos e contas, bem como os relacionamentos entre si. Cada um destes domínios de <i>Master Data</i> representa informação que é necessária em diversos processos de negócio, em diversas unidades organizacionais e entre sistemas operacionais e sistemas de suporte à decisão. A <i>Master Data</i> define uma organização.	(Dreibelbis <i>et al.</i> , 2008)
Informação crítica para as operações e suporte à decisão de uma empresa e é normalmente categorizada em áreas de entidades de <i>master data</i> (designadas frequentemente por domínios de dados) tais como clientes, produtos, fornecedores, parceiros, empregados, materiais, entre outros.	(Cervo & Allen, 2011)
Objetos fundamentais do negócio utilizados em diferentes aplicações na organização, em conjunto com a <i>metadata</i> associada, atributos, definições, papéis, ligações e taxonomias. Os objetos de <i>master data</i> são aqueles elementos chave que mais importam – são os elementos que são registados nos sistemas transacionais, medidos e reportados nos sistemas de <i>reporting</i> e analisados nos sistemas analíticos.	(Loshin, 2009)
Conjunto de identificadores consistentes e uniformes e respetivos atributos que descrevem as entidades fundamentais da empresa, sendo utilizados nos diversos processos de negócio.	(Radcliffe, 2010)

As práticas de *Master Data Management* surgiram inicialmente para endereçar estas questões de qualidade dos dados e de fragmentação (Cervo & Allen, 2011). Num mundo ideal, existe um único lugar onde toda a *master data* de uma organização está armazenada e é gerida. Os dados serão precisos, consistentes e são mantidos de uma forma coerente e segura. Todas as atualizações serão executadas nesta cópia única de *master data* e todos os utilizadores destes dados interagirão com esta fonte única e independente de informação (Dreibelbis *et al.*, 2008).

Para a manipulação da *master data* de uma organização, é necessária a existência de mecanismos bem definidos que permitam sistematizar os fluxos de informação em conformidade com as regras do negócio. Isto leva-nos ao conceito de *Master Data Management*. Na Tabela II encontram-se sistematizadas várias definições para o conceito de *Master Data Management*.

A gestão de *master data* governa os métodos, as ferramentas, a informação e os serviços (Loshin, 2009). Os sistemas MDM que se focam exclusivamente na gestão de informação acerca de clientes são frequentemente designados por *Customer Data Integration* (CDI). Os sistemas MDM que se focam exclusivamente na gestão de informação de produtos são designados por *Product Information Management* (PIM).

Tabela II – Conceito de *Master Data Management*.

Definição de <i>Master Data Management</i>	Fonte
Coleção das melhores práticas da gestão de dados que gere <i>stakeholders</i> chave, utilizadores e clientes de negócio. Incorpora as aplicações de negócio, métodos de gestão de informação para implementar políticas, procedimentos, serviços e infraestrutura, tendo em vista o suporte da captura, integração e subsequente utilização partilhada de <i>Master Data</i> consistente e precisa.	(Loshin, 2009)
Disciplina de negócio baseada em tecnologia que auxilia as organizações a alcançarem uma “versão única da verdade” em áreas tão importantes como clientes, produtos e contas. Em MDM, o negócio e as TI trabalham em conjunto para assegurar a uniformidade, precisão, persistência semântica, administração e responsabilização da <i>master data</i> partilhada.	(White & Radcliff, 2010)
Conjunto de políticas, procedimentos, aplicações e tecnologias utilizadas para harmonizar e gerir o sistema de registo e sistemas de lançamento dos dados e <i>metadata</i> associados às entidades chave de uma organização.	(White C. , 2007)

Os sistemas de MDM que manipulam múltiplos domínios de *master data* e que suportam diversos estilos de implementação são designados por *multi-form* MDM (Dreibelbis *et al.*, 2008).

2.1.1. Estilos de Implementação MDM

Os sistemas de MDM são implementados para aumentar a qualidade da *master data* e fornecer uma utilização consistente desta informação naquele que é frequentemente um ambiente complexo. Diferentes combinações de implementação e requisitos de utilização levaram à evolução de vários estilos de implementação MDM. São comuns as implementações híbridas que combinam múltiplos estilos de implementação. Os quatro estilos de implementação mais comuns são: Consolidação, *Registry*, Coexistência e *Hub* Transacional. À medida que os estilos progredem do primeiro para o último, é possível constatar um aumento de funcionalidade e a necessidade de implementações mais sofisticadas (Dreibelbis *et al.*, 2008).

Estilo de Consolidação

Este estilo de implementação reúne *master data* de vários sistemas existentes, quer de bases de dados quer ficheiros, num único *hub* MDM. Ao longo do processo, os dados são transformados, depurados, verificados e integrados com o objetivo de fornecer um “*golden record*” para um ou mais domínios de *master data*. Este registo otimizado serve de fonte confiável de sistemas a jusante para a produção de análises e relatórios, ou como um sistema de referência para outras aplicações. Alterações aos dados são originadas em sistemas que o alimentam - é o chamado sistema *read-only*.

A implementação do estilo de consolidação é uma fase inicial e natural de um *roll-out* de MDM com várias fases. Um sistema MDM com o estilo de consolidação fornece uma base para os estilos de implementação em *hub* transacional e de coexistência. Porque o estilo de consolidação é alimentado por sistemas a montante, nem sempre contém a informação mais atualizada (Dreibelbis *et al.*, 2008).

Estilo Registry (Registo)

O estilo de implementação em *registry* pode ser útil para fornecer uma fonte só de leitura de *master data* como uma referência para os sistemas a jusante, com um mínimo de redundância de dados. O sistema de MDM possui uma quantidade mínima de informação necessária para identificar univocamente um registo de *master data*. Fornece também referências cruzadas para informação detalhada que é gerida noutros sistemas e bases de dados. O estilo *registry* de implementação MDM serve como um sistema *read-only* de referência para outras aplicações. As solicitações ao sistema MDM em *registry* reúnem a informação necessária em dois passos: primeiro, a informação identificadora é pesquisada no sistema MDM; depois, utilizando essa informação de identidade e referência cruzada, é obtida informação relevante a partir de outros sistemas. No caso de um dos sistemas fonte falhar, o sistema MDM será diretamente afetado (Dreibelbis *et al.*, 2008).

Estilo de Coexistência

O estilo de implementação MDM de coexistência envolve *master data* que pode ser armazenada em várias localizações e que envolve um “*golden record*” instanciado fisicamente no sistema de MDM, que é sincronizado com os sistemas fonte. O “*golden record*” é construído da mesma forma que no estilo de consolidação, tipicamente através de importação em *batch* e pode ser atualizado no sistema MDM. Se o sistema MDM fizer uma sincronização bidirecional com os sistemas fonte, é necessário evitar ciclos de atualização onde as alterações de um sistema entrem em conflito com as alterações de outro sistema. A vantagem do estilo de coexistência é que pode fornecer um conjunto completo de capacidades MDM sem causar alterações significativas no ambiente existente. A desvantagem é que não é o único local onde a *master data* pode ser alterada, nem sempre estando atualizado.

Estilo de Hub Transacional

O estilo de implementação de *hub* transacional consiste num conjunto centralizado e completo de *master data* para um ou mais domínios. É um sistema de registo que serve como uma única versão para a *master data* que gere. Este estilo de

implementação é com frequência uma evolução das implementações de consolidação e coexistência. Dado que se trata de um sistema de registo, as atualizações à *master data* acontecem diretamente neste sistema, utilizando os serviços fornecidos pelo *hub*. À medida que as transações acontecem, a *master data* é verificada e mantida de forma a preservar a sua qualidade. Depois de as atualizações serem aceites, o sistema distribui estas alterações aos utilizadores e aplicações indicadas para o efeito. Sendo a implementação em *hub* transacional um sistema de registo, a segurança e *governance* desempenham um papel especialmente crítico na manutenção da integridade da *master data*. As principais desvantagens do estilo de *hub* transacional são os custos e a complexidade. Este tipo de implementação significa frequentemente que os sistemas existentes e os processos de negócio têm de ser alterados quando o *hub* transacional passa a ser o ponto único de atualização no ambiente (Dreibelbis *et al.*, 2008).

Na Tabela III estão sistematizadas as vantagens e desvantagens dos estilos descritos.

Tabela III - Estilos de Implementação MDM.

Estilo	CONSOLIDAÇÃO	REGISTRY	COEXISTÊNCIA	HUB TRASACIONAL
Descrição	Agrega <i>master data</i> num repositório comum para <i>reporting</i> .	Mantém um sistema de registo com <i>links</i> para dados existentes nos sistemas; útil para referência em tempo real.	Gere uma visão única da <i>master data</i> , sincronizando alterações com outros sistemas.	Gere uma vista única de <i>master data</i> , fornecendo acesso através de serviços.
Benefícios	Ótimo para a preparação dos dados para alimentação de sistemas a jusante.	A visão completa dos dados é montada à medida das necessidades; de rápida construção.	Pressupõe os sistemas existentes inalterados, no entanto permite a gestão de leitura e de escrita.	Suporta aplicações transacionais novas e já existentes.
Desvantagens	Só de leitura; nem sempre existente em sistemas operacionais.	Principalmente de leitura; poderá ser de gestão mais complexa.	Nem sempre compatível com outros sistemas.	Pode exigir alterações aos sistemas existentes em exploração.

Fonte: adaptado de Dreibelbis *et al.* (2008), p. 32

2.2. Abordagens de MDM

Para os diferentes domínios de *master data*, os dados de uma empresa podem ser classificados como operacionais ou não-operacionais. Os dados operacionais consistem na coleção em tempo real de dados que suportam as necessidades da empresa nas suas atividades diárias e os dados não-operacionais são capturados num *data warehouse* e são utilizados para suporte à decisão. As implementações MDM

podem ser categorizadas em três grandes tipos de iniciativas: MDM Analítico (endereço o BI); o MDM Operacional (endereço as operações de negócio) e o MDM Empresarial (endereço o BI e as operações de negócio). Cada um possui um objetivo algo diferente e envolve diferentes níveis de complexidade, risco e impacto. (Cervo & Allen, 2011).

MDM Analítico

Em termos históricos, o MDM analítico tem sido a abordagem mais adotada. Isto deve-se em grande parte à relativa simplicidade em alavancar projetos de *data warehouse*. Uma desvantagem desta abordagem é a possibilidade de acrescentar um ou mais sistemas de dados fragmentados e incompletos. Algumas áreas de negócio são especialmente resistentes a esta abordagem, sendo necessário um forte patrocínio da gestão de topo (Cervo & Allen, 2011).

MDM Operacional

O *MDM* operacional tem como alvo sistemas e dados. Fornece a oportunidade de consolidar os sistemas operacionais de dados na empresa e tornar-se um verdadeiro sistema de referência. Dependendo das dimensões da empresa, o *operacional MDM* deverá ser desejavelmente desenvolvido por fases. Um método para uma implementação faseada é a migração gradual de cada sistema de dados para um único repositório MDM até que todos os sistemas envolvidos tenham atingido o fim-de-vida. Outro método para dividir as fases é através da migração gradual de porções de dados de um único sistema. A combinação dos dois métodos é também comum (Cervo & Allen, 2011).

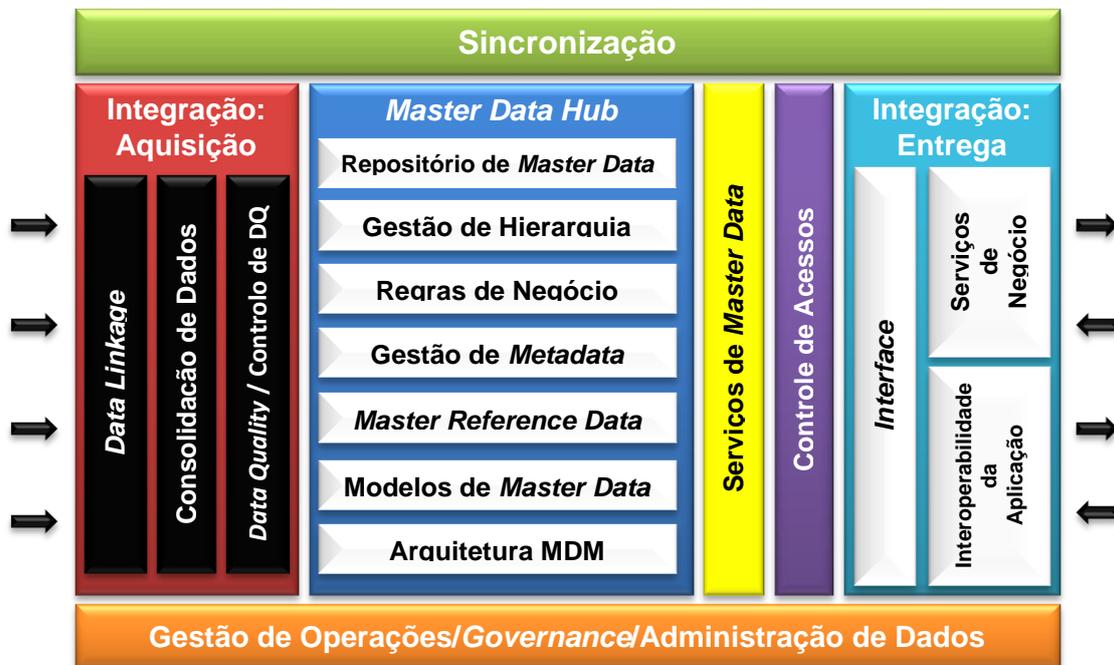
MDM Empresarial

O *MDM* empresarial é uma combinação do MDM operacional e do MDM analítico. Desta forma, pode ser implementado através da combinação das arquiteturas anteriormente apresentadas. O número de combinações de abordagens MDM e arquiteturas é muito grande. As categorias anteriores são linhas de orientação geral e as implementações mais comuns. É importante atender ao domínio de dados considerado, ao objetivo da gestão dos dados (operacional ou analítico) e à arquitetura adotada (Cervo & Allen, 2011).

2.3. Ferramentas de MDM

O lançamento de um extenso projeto ou programa de MDM, em virtude da sua posição como uma aplicação empresarial, requer coordenação e planeamento para além da

aquisição de tecnologia. Quando a organização está no ponto em que todos os *stakeholders* estão alinhados para a implementação do MDM, é útil avaliar as necessidades organizacionais do negócio e avaliar apropriadamente as ofertas dos fornecedores. Podemos agrupar as capacidades técnicas do MDM em sete grandes áreas de infraestrutura e serviços, como se mostra na Figura 1 (Loshin, 2012).



Fonte: adaptado de Loshin (2012), p. 4 - Knowledge Integrity, Inc.

Figura 1 - Capacidades Técnicas do MDM.

Segundo Loshin (2012), as sete áreas de infraestrutura e serviços são:

1. **Hub Central de Master Data:** O *hub* de *master data* combina a infraestrutura, a gestão de dados, a gestão de dados de referência e a gestão de regras de negócio. Alguns produtos de fornecedores possuem pacotes com modelos de dados que podem ser desenvolvidos; outros fornecedores esperam que os implementadores integrem os seus próprios modelos na arquitetura;

2. **Integração de Dados e Consolidação:** Dois aspetos críticos do MDM, relativamente à gestão dos dados no *hub* de *master data* são: obter os dados para o *hub* e disponibilizar os dados a partir do *hub*. O primeiro aspeto requer a integração dos dados e componentes de qualidade dos dados, tais como: conectores para os sistemas de dados fonte, transformação dos dados, melhoria da qualidade dos dados, normalização e carregamento;

3. **Serviços de Master Data:** A disponibilização dos dados a partir do *hub* consiste em vários serviços de entrega de dados. Os serviços básicos de *master data*

podem ser configurados para criar, ler, atualizar e apagar, quer o objeto em si, quer os atributos associados;

4. **Entrega (*Delivery*):** A entrega de informação tem o objetivo de adicionar valor aos serviços de negócio, permitindo a redução da aplicação de funções comuns implementadas diversas vezes em diferentes aplicações. A criação de uma *interface* padrão ajudará também na transição de aplicações existentes, facilitando a interoperabilidade da aplicação;

5. **Controle de Acessos:** Trata-se de um componente de valor acrescentado para uma ferramenta de MDM, dada a grande exposição da infraestrutura MDM a outras aplicações e utilizadores;

6. **Sincronização:** Em função do estilo de implementação adotado, existem diferentes requisitos para sincronizar os dados provenientes dos diferentes sistemas de dados. Todos os fornecedores de MDM devem ser capazes de descrever de que forma e com que frequência o seu produto sincroniza os dados e como garante a consistência dos mesmos;

7. **Operações e Governança dos Dados:** Para além da monitorização contínua da qualidade dos dados, deverá existir governança operacional e serviços de administração de forma a suportar a análise e melhoria da qualidade dos dados, bem como o controle da sua governança no ambiente de *master data*.

2.3.1. O que procurar numa Plataforma MDM

Apesar do MDM não dever ser considerado um projeto de tecnologia, é claro que o desenvolvimento de um programa de MDM não pode ser realizado sem a alavancagem de aplicações e tecnologia. Os aspetos técnicos da gestão de *master data* assentam em ferramentas que suportam o *data profiling*, análise complexa de dados, gestão de *metadata*, criação de modelos de dados, integração de dados, normalização, *linkage* e *matching* de registos, SOA, controle de acessos, *data federation*, entrega dos dados, entre outros. O MDM é uma composição de pessoas, ferramentas, métodos e políticas que irão moldar a forma como será explorado o valor da informação de uma organização (Loshin, 2009).

Segundo Loshin (2012), no processo de avaliação e seleção de *software* MDM devem ser observadas as seguintes funcionalidades:

Resolução de Identidades (*Identity Resolution*): Uma vez que o objetivo de uma aplicação de MDM é fornecer uma visão única de *master data*, a ferramenta deve oferecer funcionalidades de resolução de identidades (processo operacional para a

ligação de diferentes fontes de dados, que aplica valores de probabilidade para determinar se as entidades são coincidentes e que possíveis relacionamentos existem entre elas);

Modelos Principais de Dados: Algumas características e tipos de dados podem ser encapsulados num modelo base que pode ser ajustado para satisfazer necessidades do cliente. O mesmo se passa noutros conceitos de dados normalizados como os de cliente, produto, fornecedor, entre outros;

Serviços de Dados: A ferramenta de MDM deve sugerir ou providenciar uma biblioteca de serviços utilizada para criar, modificar e substituir instâncias de objetos de *master data*;

Funcionalidade de Gestão em Hub: Se a ferramenta de MDM é utilizada para reunir e consolidar uma cópia única de entidades de *master data*, deve existir uma capacidade sistematizada de instanciar e gerir esse repositório principal de dados;

Aplicações de Data Governance: Quando são identificados casos de falsos positivos ou falsos negativos na integração dos dados, deve ser garantida uma *interface* para a remediação manual dos problemas identificados, assegurando a administração e governo dos dados;

Federação de Dados (Data Federation): A aplicação deve contemplar funcionalidades de *data federation* (técnica que consiste na virtualização de dados, utilizando, por exemplo, CDC) para suportar a sincronização e atualização dos dados;

Integração de Dados: O produto deve incluir ferramentas com capacidades de extração, transformação e carregamento (ETL);

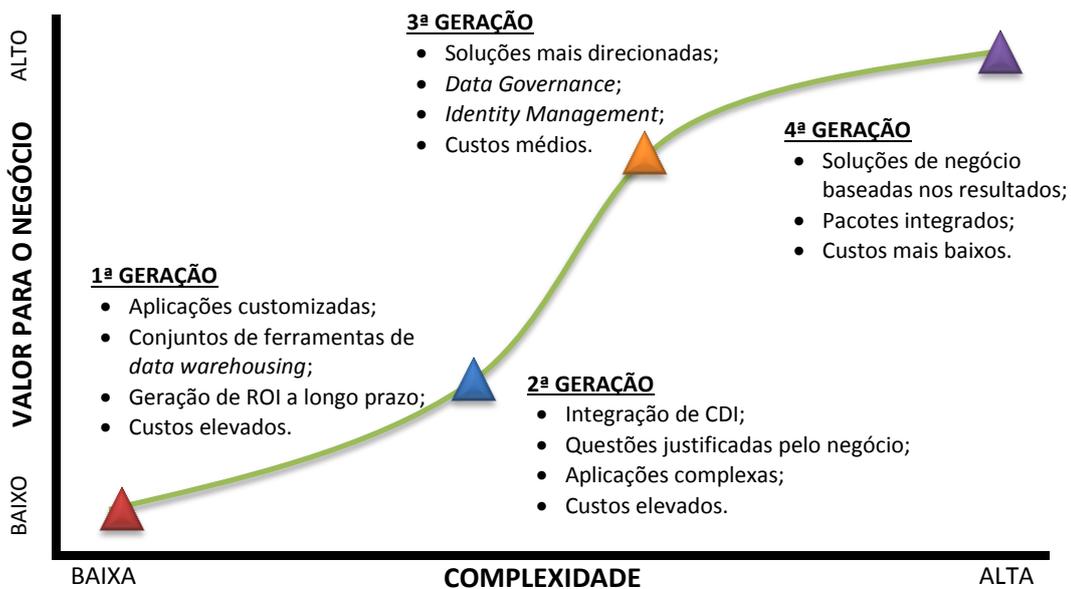
Normalização e Enriquecimento (Standardization e Enhancement): Funcionalidades integradas em regra num pacote de *data quality* que reduzem a incoerência dos dados e melhoram o processo de *matching* na agregação de registos;

Perfil dos Dados (Data Profiling): Ferramenta indispensável para a identificação da *metadata* e avaliação da qualidade dos dados. Tecnicamente é normalmente um componente isolado, embora alguns fornecedores o agreguem a outras funcionalidades.

2.3.2. O Mercado de Ferramentas MDM

Atualmente, o mercado de produtos de MDM é bastante extenso e diversificado. O *site The MDM Institute* faz referência a 55 produtos de *Master Data Management* (CDI, PIM e *Data Governance*) (The MDM Institute, 2012). Desde pacotes comerciais até ferramentas *open source*, a oferta de aplicações de MDM está também associada a

muitos fabricantes de *software* marginal e fornecedores de serviços de *marketing* que estão na periferia do mercado de MDM. Segundo a Gartner (2012), além dos produtos líderes de mercado, existem também fabricantes cujas receitas são muito pequenas ou que se focam numa área geográfica muito limitada, outros existem que se focam unicamente num único mercado de indústria vertical como as ciências da vida ou a área financeira. Para além destas classes de ferramentas, existem também *brokers* de bases de dados que desenvolvem e mantêm um conjunto de dados de domínios de dados como cliente, fornecedor e veículos. São vistos como complementares às soluções de MDM, dado que podem fornecer, por exemplo, alterações ou atualizações de endereços e, desta forma, enriquecer e acrescentar valor à *master data* existente (Radcliffe, 2010). Em termos de evolução, Srinivasan (2010) considera que as soluções de MDM se classificam em 4 gerações. A Figura 2 ilustra e caracteriza essas gerações.



Fonte: adaptado de Srinivasan (2010), p. 7

Figura 2 - Evolução das Soluções de *Master Data Management*.

Os maiores domínios do mercado de *software* de MDM são os de dados de cliente e de produto, prevendo-se que cada um desses domínios duplique em dimensão nos próximos 4 anos (Gartner, 2012). A recessão mundial provocou consequências no mercado de *software* em 2009, com alguns setores a verem diminuídas as suas receitas. No entanto, de acordo com estimativas anteriores, o mercado de aplicações de MDM para os dados de cliente resistiu bem à recessão e cresceu cerca de 11% em 2009. O primeiro trimestre de 2010 foi um período de grandes transformações no

mercado de MDM de dados de cliente, com dois dos melhores fabricantes, Iniate Systems e Siperian, a serem adquiridos pela IBM e Informática, respectivamente. Foi ainda possível observar a *Microsoft* a entrar no mercado no segundo trimestre de 2010 (Radcliffe, 2010). De acordo com a Gartner, as receitas de *software* de MDM atingirão 1,9 mil milhões de dólares em 2012, um aumento de 21% em relação a 2011. Prevê-se que o mercado atinja 3,2 mil milhões de dólares em 2015 (Gartner, 2012).

À medida que o mercado de MDM cresce e evolui, novas tendências ajudarão a marcar a sua evolução. Segundo John Radcliffe (Gartner, 2011), as três principais tendências que terão um impacto significativo no mercado nos próximos anos, são:

Procura crescente de *software* MDM multi-domínio: Prevê-se que, em 2014, 66% das organizações da Fortune 1000 tenham implementado duas ou mais soluções MDM para suportar as suas estratégias empresariais. Em termos históricos, a maioria das iniciativas MDM foram orientadas para um único domínio de *master data*. No entanto, muitas organizações estão agora a esforçar-se para alcançar uma visão mais alargada de como conseguir capacidades em diversos domínios de *master data*.

Adoção crescente de MDM na *cloud*: Prevê-se que, em 2015, 10% das implementações de MDM sejam entregues como SaaS em *cloud* pública. Atualmente, o MDM é implementado tipicamente em soluções *in-house*. Muitas indústrias (ex.: mercado financeiro) estão relutantes em colocar informação tão importante como a *master data* fora do seu CPD. Questões de privacidade também têm impedido a adoção de MDM como SaaS em algumas regiões, como é o caso da Europa. No entanto, algumas soluções de MDM têm vindo a ser progressivamente integradas com aplicações SaaS e existem vários serviços de dados na *cloud* (ex.: *data integration as a service*) que complementam as implementações de MDM.

Aumento de *links* entre o MDM e as redes sociais: Prevê-se que, em 2015, 15% das organizações terão adicionado *social media data* dos seus clientes aos atributos da *master data* de clientes nos sistemas de MDM. A análise da *social media* pode ser muito útil para compreender o que é que o mercado pensa sobre uma empresa, os seus produtos e serviços. No futuro, o *software* MDM para os dados de cliente das organizações deverá ser orientado para as redes sociais. Esta integração de fontes de dados provenientes de redes sociais no MDM é frequentemente designada por *Social MDM*.

3 Metodologia de Investigação

3.1. A Metodologia

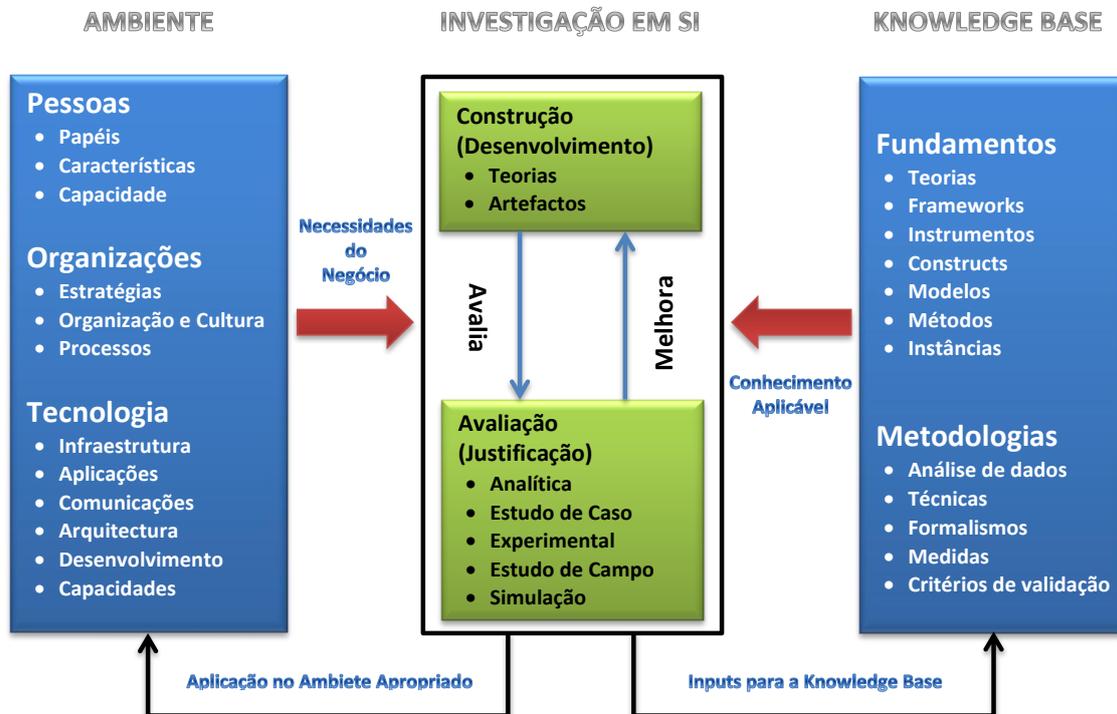
Os Sistemas de Informação são implementados no seio de uma organização com o objetivo de melhorar a eficácia e eficiência dessa mesma organização. O objetivo da investigação em Sistemas de Informação é o de adquirir o conhecimento e o entendimento que permita o desenvolvimento e implementação de soluções baseadas em tecnologia que enderecem importantes problemas de negócio que estejam ainda por resolver (Hevner *et al.*, 2004). A investigação na área dos Sistemas de Informação deve promover o uso de paradigmas de investigação que se dediquem quase exclusivamente à solução do problema de investigação, contribuindo para a solução prática do problema identificado numa determinada organização. Neste sentido, é importante seguir uma metodologia que permita a cabal resposta a todas as questões de investigação em estudo, o desenho de uma solução concreta para o problema em análise e simultaneamente a avaliação e validação do modelo de resposta construído. Nos anos mais recentes, vários investigadores tiveram sucesso, nas comunidades de investigadores e Sistemas de Informação, em provar a validade e o valor do *Design Science* como um paradigma de investigação em Sistemas de Informação (Peffer *et al.*, 2007).

Tendo em vista a natureza dos resultados que se pretende alcançar, o *Design Science* parece ser a metodologia que mais se adequa ao processo de investigação aqui proposto. O *Design Science* é fundamentalmente um paradigma orientado à resolução de problemas.

A metodologia do *Design Science* tem como base a identificação e manipulação de artefactos. Os artefactos de TI são definidos geralmente como *constructs* (vocabulário e símbolos), modelos (abstrações e representações), métodos (algoritmos e práticas) e instâncias (sistemas implementados e protótipos) (Hevner *et al.*, 2004). Os artefactos são assim construídos para endereçar problemas não resolvidos e são avaliados em termos de utilidade na resolução desses mesmos problemas, permitindo aos investigadores aprender acerca do mundo real, avaliar de que forma o artefacto o afeta e como os utilizadores melhor o poderão utilizar (Hevner *et al.*, 2004).

Para a sistematização de todo o trabalho de investigação, tendo em vista o objetivo de assegurar resultados reais e confiáveis, faz todo o sentido a utilização de uma *framework* que defina claramente o âmbito de uma investigação em Sistemas de

Informação. A Figura 1 apresenta uma *framework* conceptual para a execução e avaliação de uma investigação em Sistemas de Informação.



Fonte: adaptado de Hevner *et al.* (2004), p. 80

Figura 3 - *Framework* para a Investigação em Sistemas de Informação.

Na *framework*, o Ambiente constituído por pessoas, organizações e tecnologia, cria as necessidades de negócio que estão na origem do problema a resolver no âmbito de uma investigação. Por seu lado, o processo de investigação em si constrói teorias e/ou artefactos que vão ser avaliados e validados através de processos analíticos, experimentais, por estudos de caso, estudos de campo, ou através de simulação. Este processo de avaliação pode e deve contribuir para a melhoria das teorias ou artefactos criados no processo de investigação. Todo o processo de investigação, quer seja na sua fase de desenvolvimento, quer seja na fase de validação/avaliação, vai ser suportado por instrumentos de uma *Knowledge Base* (base de dados de conhecimento), que podem ser fundamentos e/ou metodologias. Com este sistema de investigação, os resultados do estudo irão não só ter uma aplicação prática na envolvente Ambiente, resolvendo o problema por ela despoletado, como também contribuir para o enriquecimento da *Knowledge Base*, adicionando-lhe valor acrescentado.

3.2. Desenho da Investigação

A metodologia de investigação *Design Science* incorpora princípios, práticas e procedimentos necessários para a condução de uma determinada investigação e cumpre três objetivos: é consistente com a literatura existente, fornece um modelo de processo para a investigação e fornece um modelo mental para a apresentação e avaliação de uma investigação em Sistemas de Informação (Peppers *et al.*, 2007).

Segundo Peppers *et al.* (2007), o processo de *Design Science* inclui seis passos: identificação do problema e motivação do estudo, definição de objetivos para uma solução, *design* e desenvolvimento, demonstração, avaliação e comunicação. Desta forma, a metodologia de *Design Science* aplicada à questão de investigação do presente trabalho, deverá considerar os passos seguintes:

Identificação e Motivação do Problema

Nesta fase é definido o problema de investigação específico e é justificado o valor de uma solução para o mesmo. A motivação para o estudo tem como objetivos não só cativar o investigador e os *stakeholders* da investigação para a solução e aceitar os resultados, bem como compreender a envolvente e o conhecimento do investigador sobre o assunto. Esta atividade é executada através de uma revisão da literatura exaustiva e cuidada (Peppers *et al.*, 2007).

Definição de Objetivos para uma Solução

Nesta atividade são estabelecidos os objetivos de uma eventual solução a partir do problema a investigar e do conhecimento do que é possível fazer no processo de investigação. No caso proposto, os objetivos são qualitativos, devendo ser inferidos racionalmente a partir da especificação do problema.

Neste passo alguns investigadores dão relevância à natureza do processo iterativo de pesquisa (Hevner *et al.*, 2004).

Design e Desenvolvimento

Nesta fase é criado o artefacto. Esta atividade envolve a descoberta da funcionalidade do artefacto no processo de investigação.

No estudo proposto, o artefacto a gerar é a *framework* de avaliação de ferramentas. Para a sua construção é necessário encontrar previamente um conjunto de características desejáveis das ferramentas de MDM. Esta descoberta decorre da

análise da bibliografia existente sobre o tema em estudo. A sistematização da *framework* será efetuada numa única iteração da metodologia.

Demonstração

Nesta atividade, a ideia é provar que o conceito de conceção do artefacto resulta efetivamente. Deve ser demonstrada a utilização do artefacto (*framework*) em várias instâncias do problema, que no caso em estudo consistem em duas ferramentas de MDM. A demonstração será efetuada segundo o método de experimentação do modelo desenhado para a *framework*.

Avaliação

Trata-se de observar e medir se o artefacto desenhado suporta a solução para o problema a resolver. Nesta fase da metodologia serão analisados os resultados obtidos na atividade de demonstração. Pretende-se retirar conclusões acerca da eficácia da *framework* desenhada e se a mesma necessita de ajustes e melhorias. Caso sejam desejáveis melhorias no modelo da *framework* poder-se-á optar por realizar uma iteração da atividade de *Design e Desenvolvimento*, o que não aconteceu no presente trabalho devido a constrangimentos de tempo.

Comunicação

Esta atividade envolve a comunicação da questão de investigação, da proposta de solução encontrada para o problema de investigação, da eficácia da solução e das conclusões a que se chegou.

Atendendo a que a proposta de investigação avançada no presente trabalho tem como objetivo a produção de um Trabalho Final de Mestrado, este processo de comunicação consistirá na produção do presente Relatório e na apresentação e defesa do trabalho efetuado.

4 A Framework

Nesta secção será apresentado o estudo efetuado, que corresponde à fase de *design* e desenvolvimento da metodologia de investigação de *Design Science*, na sua primeira e única iteração. No trabalho apresentado neste documento, apenas será efetuada uma única iteração, dado o limitado período de tempo disponível para a sua realização. Num cenário ideal, o presente trabalho envolveria mais do que uma iteração, com o objetivo de alcançar um artefacto mais abrangente e consolidado – a nossa *framework*.

Tendo em vista a resposta à questão de investigação a que o presente trabalho se propõe, pretende-se construir uma *framework* que evidencie as características e atributos desejáveis que permitam a avaliação de uma qualquer ferramenta de MDM. Apesar disso, não existe a pretensão de elencar exaustivamente todas as características, mas antes disponibilizar uma ferramenta de trabalho para equipas de TI que constitua um referencial para aferição qualitativa deste tipo de *software*, a partir das características que têm diretamente a ver com o MDM.

O modelo construído, cuja utilidade parece não oferecer qualquer dúvida, deve ser o mais intuitivo possível, possuir alguma flexibilidade no sentido em que deverá adequar-se a qualquer *software* de MDM e seu fabricante e deverá ser de simples utilização. A construção e definição dos diversos componentes da *framework* foram efetuadas exclusivamente a partir da revisão da literatura.

O trabalho de desenvolvimento apresentado em seguida possui as fases seguintes:

1. Levantamento das características desejáveis para uma ferramenta de MDM;
2. Construção da *framework* para a avaliação de ferramentas de MDM;
3. Avaliação de 2 ferramentas de *software* MDM disponíveis no mercado, através da aplicação da *framework* concebida no ponto anterior;
4. Avaliação da *framework* proposta através de entrevistas a realizar a investigadores, fabricantes de *software* e proponentes de iniciativas de MDM.

4.1. Características Desejáveis do Software de MDM

Existe atualmente no mercado uma grande variedade de aplicações de MDM e tem vindo a ser conduzido um processo de consolidação entre os fabricantes das ferramentas de MDM. Esta situação configura um problema para as equipas de TI na avaliação e pesquisa da ferramenta de MDM mais adequada para a sua organização.

Qualquer implementação de MDM envolve, quer a equipa de TI, quer o negócio. Desta forma, durante a fase de seleção de uma ferramenta, as atividades e decisões devem envolver ambas as partes. Este processo envolve em algum ponto a definição dos requisitos do negócio e dos requisitos técnicos a cumprir pela aplicação selecionada, bem como a criação de uma matriz de seleção. Em termos de requisitos técnicos devem ser considerados os requisitos funcionais, não-funcionais e os de integração de sistemas.

É aconselhável ter em consideração diversos aspetos no processo de avaliação de *software*, tais como: os objetivos e visão do negócio, os custos da ferramenta, custos de desenvolvimento, tempo necessário e esforço envolvido, usabilidade da ferramenta, capacidade de integração com os sistemas existentes, arquitetura e sincronização dos dados, perfil do fabricante, entre outros.

Um dos grandes desafios, no que diz respeito à tecnologia para o MDM, está relacionado com o seu grande espectro de aplicabilidade e utilização. Algumas ferramentas são utilizadas pela equipa de TI, algumas pelo negócio e outras por ambos. Por outro lado, os utilizadores da equipa de TI podem ser capazes de entender mais facilmente os conceitos que envolvem ferramentas mais complexas, mas não sabem propriamente como melhor os aplicar para resolver uma questão do negócio (Cervo & Allen, 2011).

Na análise de uma qualquer ferramenta de MDM é possível sistematizar uma série de características que são desejáveis e várias outras que são aconselháveis.

Ainda assim, não obstante as perspetivas serem divergentes de autor para autor, é possível sistematizar um conjunto de características fundamentais que são reconhecidos como tal por diversos peritos e autores considerados em matéria de MDM. Cada uma das características identificadas configura um parâmetro de avaliação que compõe a *framework*.

A Tabela IV reúne as vinte e duas características identificadas na bibliografia disponível, apresentando para cada uma algumas referências retiradas da literatura consultada.

Tabela IV - Características das Ferramentas de MDM.

Característica C₁: *Identity Resolution* (Resolução de Identidades)

O *Identity Resolution* é utilizado mais recentemente de forma pró-ativa como um processo de *data governance* para reduzir a introdução de duplicados ou informação de entidades inconsistentes. Os processos de *identity resolution* fornecem valor acrescentado em aplicações analíticas (Loshin, 2010). A *Identity Resolution* refere-se à capacidade de determinar que duas ou mais representações de dados podem ser resolvidas numa única representação. Para um dado conjunto de dados, a *identity resolution* pode ser vista como um processo de duas fases: a primeira é de descoberta (combina *data profiling* com uma revisão manual dos dados); a segunda fase, alimentada pelo desenvolvimento de modelos probabilísticos simples, consiste na contagem por semelhança e *matching*, com o objetivo da ligação dos registos (Loshin, 2009).

O objetivo do parâmetro de avaliação gerado por esta característica é avaliar se a ferramenta possui um motor de *identity resolution* (ou *middleware*) que permita a conexão de diferentes fontes de dados para a identificação de identidades coincidentes e relacionamentos entre elas.

Característica C₂: *Data Profiling* (Perfil dos Dados)

O *data profiling* teve origem num conjunto de algoritmos para análise estatística e avaliação da qualidade dos valores de dados de um conjunto, bem como para explorar relacionamentos entre coleções de valores de conjuntos de dados (Loshin, 2009). Exemplos de técnicas de *data profiling* são: análise de padrões, distribuição de frequências, análise de domínios, identificação de tipos, verificação de interdependências e análise de redundâncias (Cervo & Allen, 2011). As ferramentas de *data profiling* são utilizadas para vários fins: descobrir erros nos dados, monitorizar os dados em relação a erros para suportar a governança dos dados, descobrir padrões de dados e suportar a descoberta de relacionamentos implícitos nos dados numa fonte e entre fontes de dados (Howard, 2011). As técnicas de *data profiling* podem também ser utilizadas para a avaliação da qualidade dos dados das fontes de dados de uma organização (Loshin, 2009; Cervo & Allen, 2011; Howard, 2011). Existe uma grande variedade de ferramentas de *data profiling* disponibilizadas pelos fabricantes de ferramentas de MDM, bem como opções *open source* (Cervo & Allen, 2011).

O objetivo do parâmetro de avaliação gerado por esta característica é aferir se a aplicação de MDM possui funcionalidades de *data profiling* e que técnicas utiliza para as implementar.

Característica C₃: *Parsing e Standardization* (Análise e Normalização)

As ferramentas de *parsing* permitem que o analista de dados defina padrões que possam alimentar motores de regras que são utilizados para distinguir valores de dados válidos de inválidos. Quando é encontrado um padrão válido, os componentes podem ser extraídos para uma representação padrão. Quando é reconhecido um padrão inválido, a aplicação pode tentar transformar o valor inválido num que cumpra a regra definida (Loshin, 2009). A normalização executa duas tarefas básicas: rearranja conjuntos de *tokens* reconhecidos em formato padrão e, por outro lado, identifica erros conhecidos e corrige-os (Loshin, 2010).

O objetivo do parâmetro de avaliação gerado por esta característica é averiguar se a aplicação disponibiliza ferramentas e técnicas de *parsing* e *standardization* e qual a o nível de especialização que apresentam.

Característica C₄: *Data Enhancement* (Enriquecimento dos Dados)

O *data enhancement* é executado através do mapeamento de um conjunto de dados noutros conjuntos de dados e da reunião de informação proveniente de múltiplas fontes. O objetivo do *enhancement* é identificar conhecimento utilizável de coleções de dados que são usadas nos processos de negócio. Na sua essência, o *enhancement* pode ser utilizado para aprender mais acerca das entidades, utilizando dados obtidos de um conjunto de bases de dados e aumentar a qualidade dos dados (Loshin, 2010).

O objetivo do parâmetro de avaliação gerado por esta característica é aferir se a ferramenta de MDM realiza o enriquecimento dos dados e qual a abordagem que utiliza. Este parâmetro depende dos parâmetros de *parsing* e *standardization* e *identity resolution*.

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Característica C₅: *Data Lineage* (Origem dos Dados)

A funcionalidade de *data lineage* permite a pesquisa da origem de “pedaços” ou conjuntos de *master data*. Torna-se um processo fundamental quando a *master data* de vários sistemas de informação distribuídos está consolidada (Otto, Humer, & Osterle, 2011). Em virtude de poderem existir registos de diferentes sistemas que representam uma única entidade, é necessário documentar que fontes de dados contribuem para a consolidação principal dos dados e, em certos tipos de arquiteturas MDM, que *links* existem entre o índice principal e os registos fonte originais, com o objetivo de materializar a informação solicitada. Neste sentido, a *data lineage* pode ser encarada como uma forma de mitigar a ocorrência inevitável de erros na integração de dados (Loshin, 2009).

O parâmetro de avaliação gerado por esta característica, que possui uma forte componente de *data governance*, tem o objetivo de avaliar se a aplicação de MDM efetua a monitorização das operações efetuadas sobre os dados.

Característica C₆: *Segurança e Auditoria*

O acesso à *master data* por todos os seus utilizadores é gerido pelos serviços de controlo de acessos e de gestão da segurança, Estes serviços são parte integrante da estrutura de *governance* do *software* de MDM (Loshin, 2009).

O objetivo do parâmetro de avaliação gerado por esta característica consiste em avaliar se o *software* de MDM possui mecanismos de controlo de acessos, autenticação, mecanismos de segurança e capacidades de auditoria.

Característica C₇: *Processos de Data Governance* (Governança dos Dados)

Um dos maiores valores de uma iniciativa de MDM é que, por se tratar de uma iniciativa empresarial, facilita o desenvolvimento de um programa de *data governance* na organização. Apesar do sucesso do MDM assentar na governança dos dados, um programa de *governance* pode ser aplicado em diferentes domínios operacionais, proporcionando uma implementação que envolva toda a organização (Loshin, 2009).

O parâmetro de avaliação gerado por esta característica tem o objetivo de aferir se a aplicação possui processos de *data governance* que incluam a parametrização de políticas, a criação de processos e controlos e a disponibilização de ferramentas de auditoria para o controlo de conformidades e regras de negócio.

Característica C₈: *Domínios de Master Data*

Os objetos de *master data* podem ser classificados segundo uma hierarquia. Essa classificação é efetuada em domínios de *master data* com base no papel desempenhado pela categoria dos dados (Loshin, 2009). Segundo Dreibelbis et al. (2008), a *master data* pode ser categorizada de acordo com os tipos de questões que endereçam, sendo as questões mais comuns: “quem?”; “o quê?” e “como?”. Estas questões são endereçadas pelos domínios de *master data party* (indivíduo/cliente, fornecedor, empregado, organização, parceiro, cidadão, ...), *product* (produto, pacote de produtos, peça, serviço, bem, ...) e *account* (conta financeira, *loyalty*, acordo, contrato, ...). Cada um destes domínios representa uma classe. As aplicações de MDM estão, regra geral, especializadas para endereçar um ou mais domínios de *master data*. Para a avaliação e seleção de uma aplicação de MDM, o domínio de *master data* é uma característica fundamental.

O parâmetro de avaliação gerado por esta característica tem o objetivo de estabelecer qual ou quais os domínios de *master data* tratados pela ferramenta em análise.

Característica C₉: *Estilo de Implementação*

As 3 dimensões principais das soluções de MDM são: os domínios de *master data*, os métodos de utilização da ferramenta e os estilos de implementação. (Dreibelbis et al., 2008). Todas as arquiteturas devem suportar um acesso partilhado e transparente a uma representação única da *master data*, mesmo para diferentes situações de utilização (Loshin, 2009). Diferentes combinações de implementação e requisitos de utilização levaram à evolução de vários estilos de implementação do MDM. As organizações podem começar com um estilo de implementação mais simples e mais tarde endereçar necessidades de negócio adicionais expandindo a implementação de forma a contemplar estilos adicionais (Dreibelbis et al., 2008).

O objetivo do parâmetro de avaliação gerado por esta característica consiste em avaliar que estilo ou estilos de implementação a ferramenta de MDM pode suportar.

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Característica C₁₀: Conectividade com as Fontes de Dados

Em qualquer iniciativa de MDM é necessário identificar duplicados, criar relacionamentos, criar *intelligence*, minimizar custos de manutenção, controlar acessos, tendo em vista a manutenção de um repositório de alta qualidade com dados fiáveis e governáveis. Por estas razões a integração dos dados é a fundação de tudo o resto (Cervo & Allen, 2011). A integração de um sistema MDM operacional com os sistemas já existentes necessita do suporte de uma grande variedade de estilos de comunicação e protocolos, incluindo estilos síncronos e assíncronos, transações globais e comunicações unidirecionais (Dreibelbis et al., 2008). As aplicações de MDM disponíveis atualmente no mercado possuem uma série de conectores para diversos sistemas de gestão de bases de dados, tecnologias e protocolos para a comunicação e integração dos dados no repositório de *master data*.

O objetivo do parâmetro de avaliação gerado por esta característica é aferir quais os conectores e protocolos de comunicação disponibilizados pelo *software* de MDM.

Característica C₁₁: Sincronização de Master Data

A sincronização é um desafio tanto maior quanto mais dados estão distribuídos em diversos sistemas, física ou virtualmente (Cervo & Allen, 2011). Os arquitetos de MDM deverão avaliar as aplicações e suas necessidades futuras para determinar quais são os requisitos para a sincronização da *master data*. Estas avaliações permitem aos arquitetos da solução definir qual a arquitetura que melhor satisfaz as necessidades organizacionais (Loshin, 2009). Todos os fornecedores de ferramentas MDM deverão ser capazes de descrever de que forma os seus produtos sincronizam os dados e mostrar a forma de ajustar a frequência de sincronização (Loshin, 2012).

O parâmetro de avaliação gerado por esta característica tem o objetivo de apurar que metodologias e abordagens de sincronização são disponibilizadas pelo *software*.

Característica C₁₂: Modelo de Master Data

Todos os dados em vários formatos e estruturas necessitam de ser apresentados como um recurso centralizado que contempla as diferenças entre as fontes de dados existentes. Por esta razão deverá existir um modelo consolidado para a representação da *master data*, bem como modelos para a extração e troca de dados (Loshin, 2009). O modelo de master data deve estabelecer como serão os registos mestre, ou seja, que atributos são incluídos, que dimensão e tipo de dados possuem, que valores são permitidos, entre outras características. Este passo deve também incluir o mapeamento entre o modelo de *master data* e as fontes de dados existentes (Wolter & Haselden, 2006).

O objetivo do parâmetro de avaliação gerado por esta característica consiste em avaliar se a aplicação possui ferramentas para a modelização da *master data* e como funcionam.

Característica C₁₃: Data Integration (Integração de Dados)

As implementações de MDM iniciam-se normalmente com um esforço de integração de dados, que não implica necessariamente a transferência física dos dados. Outras entidades e informação transacional têm de ser conectados ao repositório de MDM, que, por si só, constitui outro nível de integração (Cervo & Allen, 2011). Os produtos de integração de dados evoluíram ao ponto em que podem adaptar-se a praticamente qualquer *framework* de representação de dados e fornecem os meios para transformar os dados existentes numa forma que pode ser materializada, apresentada e gerida através de um sistema de *master data* (Loshin, 2009). Uma aplicação de MDM deve contemplar ferramentas de integração de dados com capacidades de ETL (Loshin, 2012).

O objetivo do parâmetro de avaliação gerado por esta característica é estabelecer se a aplicação de MDM possui ferramentas de integração de dados e quais as metodologias utilizadas.

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Característica C₁₄: *Data Federation* (Federação de Dados)

Apesar do “santo graal” do MDM ser uma fonte única completamente sincronizada com os dados que suportam todas as aplicações da organização, o MDM pode ser criado por um conjunto de apontadores para atributos de dados relevantes, associados num repositório central, com capacidade de indexação, que funciona como um *registry* para os dados distribuídos na organização. Este tipo de modelo de informação “federada” é frequentemente suportado por estilos de ferramentas de integração de dados como EAI e EII. Esta capacidade é importante em sistemas de MDM que utilizem uma qualquer *framework* que não mantenha todos os atributos no repositório (Loshin, 2009). Os sistemas de *data federation* fornecem normalmente uma infraestrutura mais robusta que uma implementação simples em *registry* (Cervo & Allen, 2011). Na avaliação de uma ferramenta de MDM deve ser observada a funcionalidade de *data federation* para suportar a sincronização ou para capturar e comunicar atualizações (Loshin, 2012).

O objetivo do parâmetro de avaliação gerado por esta característica consiste em estabelecer se a ferramenta de MDM possui capacidades de *Data Federation* e quais as técnicas disponibilizadas na integração de dados.

Característica C₁₅: *Social MDM*

O *social MDM* envolve a inclusão no sistema de MDM de fontes adicionais de dados e a integração destas fontes com as fontes do tradicional CRM e MDM. No futuro, o sistema MDM para os dados de cliente das empresas deverá ser orientado para as redes sociais (Gartner, 2011).

O objetivo do parâmetro de avaliação gerado por esta característica é avaliar se a aplicação de MDM faz integração de *social media* no repositório de *master data* e quais as fontes de dados que poderá integrar.

Característica C₁₆: *Mobile MDM*

A integração de computadores e da internet tornou-se muito comum um pouco por todo o lado. Os conteúdos estão a ser cada vez mais disponibilizados em vários formatos e a partir de qualquer localização, sendo esta tendência impulsionada pela utilização de dispositivos móveis. Um dos fatores que influenciarão o MDM do futuro é a utilização de aplicações móveis (Messerschmidt & Stuben, 2011). Alguns fabricantes de *software* de MDM disponibilizam já ferramentas que permitem a utilização e acesso da plataforma a partir de dispositivos móveis.

O objetivo do parâmetro de avaliação gerado por esta característica é aferir se a ferramenta de MDM pode ser executada a partir de dispositivos móveis.

Característica C₁₇: *Processos e Workflow*

As políticas do negócio são traduzidas em processos de *workflow* para cumprirem os objetivos traçados. Os *workflows* são configurados especificamente para integrar os eventos do ciclo de vida da *master data* nas aplicações correspondentes e nos contextos específicos do negócio (Loshin, 2009). Em regra, as ferramentas de MDM disponibilizam mecanismos de *workflow* para a gestão dos processos de negócio.

O parâmetro de avaliação gerado por esta característica tem como objetivo a aferição da existência de mecanismos de criação e gestão de processos e de *workflow*.

Característica C₁₈: *Gestão de Metadata*

A *metadata* é o DNA de todas as iniciativas de integração de dados, incluindo o MDM, dado que servem três funções essenciais: a documentação, navegação e administração dos dados. A gestão da *metadata* tem vindo a ser desenvolvida desde há muito tempo, mas as organizações têm vindo a adotá-las muito lentamente. A *metadata* é um método essencial para gerir os dados de uma organização (Moss, 2007). Segundo Otto & Huner (2009), uma ferramenta de MDM deve contemplar, no âmbito da gestão de *metadata*, a documentação de regras de negócio, um glossário/dicionário para o MDM e as funcionalidades de importação, transporte, visualização e publicação de *metadata*.

O objetivo do parâmetro de avaliação gerado por esta característica é avaliar como é efetuada a gestão da *metadata* e que *standards* são utilizados.

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Característica C₁₉: *Reporting* (Gestão de Relatórios)

Certos tipos de regras e regulamentações requerem que a organização prepare documentos e relatórios que demonstrem a conformidade. A capacidade de conduzir auditorias requer a supervisão sistemática dos processos que implementam as políticas de informação. Para assegurar a consistência do *reporting* e sua precisão são necessárias tarefas de administração e *governance* dos dados utilizados para popular esses relatórios (Loshin, 2009).

O objetivo do parâmetro de avaliação gerado por esta característica é avaliar quais as ferramentas de *reporting* disponibilizadas pelo *software* nas áreas de *data quality*, estatísticas de utilização, monitorização e suporte a auditorias.

Característica C₂₀: *Usability* (Usabilidade)

A avaliação da qualidade do *software* é muito importante, não só para determinar o nível da qualidade fornecida, mas também quando é necessário efetuar uma opção entre dois produtos de *software* similares mas concorrenciais. Questões como: O produto é de fácil aprendizagem? Responde às solicitações? As tarefas podem ser realizadas eficientemente?, determinam a sua aceitação e sucesso no mercado de *software* (Folmer & Bosch, 2002).

Na norma ISO9241-11, a *usability* é definida como o grau com que um produto pode ser usado por utilizadores definidos de forma a alcançar objetivos específicos com eficácia, eficiência e satisfação (ISO9241-11, 1998). Existem diferentes abordagens para a forma como a *usability* deve ser medida e avaliada (Folmer & Bosch, 2002).

O objetivo do parâmetro de avaliação gerado por esta característica é avaliar o nível de usabilidade da ferramenta de MDM.

Característica C₂₁: *Exception Handling* (Tratamento de Exceções)

Por muito bem planeada que esteja uma aplicação de MDM é muito provável que um grande número de falsos positivos e falsos negativos ocorra durante a migração inicial dos dados para o repositório, à medida que os dados entram no ambiente de MDM. Em qualquer aplicação de MDM, são detetadas inconsistências nas estruturas de ficheiros e valores de dados. Quando são encontrados esses erros, o *software* deve estar capacitado de mecanismos de *exception handling* que permitam colocar de parte os casos de erro, continuando o processo de importação de registos. Os casos de erro poderão ser depois tratados automática ou manualmente (Loshin, 2009).

O objetivo do parâmetro de avaliação gerado por esta característica é aferir se a aplicação de MDM possui alguma capacidade para o tratamento de exceções.

Característica C₂₂: Requisitos Tecnológicos

Para avaliar a qualidade do *software*, a forma como o mesmo funciona no seu domínio aplicacional tem de ser tomada em consideração, ao invés de avaliar o *software* fora do seu contexto (Folmer & Bosch, 2002). Em qualquer processo de avaliação e seleção de *software* de MDM é importante aferir os requisitos do sistema no que diz respeito à sua compatibilidade em termos de sistemas operativo e de gestão de bases de dados, de forma a adequar a ferramenta ao ambiente tecnológico já existente.

O objetivo do parâmetro de avaliação gerado por esta característica é aferir quais as plataformas tecnológicas compatíveis com o funcionamento da ferramenta de MDM.

4.2. Construção da Framework

Da análise da literatura existente sobre o MDM foram selecionadas vinte e duas características desejáveis e aconselháveis para a avaliação de uma ferramenta de MDM. Cada característica identificada foi transformada num parâmetro de avaliação para a construção da *framework*.

Foram consideradas seis categorias nas quais foram enquadrados os parâmetros que compõem o artefacto – a *framework*. As categorias consideradas foram: *Master Data Quality*, *Segurança e Governance*, *Arquitetura*, *Controlo e Auditoria*, *Interface e Plataforma Tecnológica*. Não foi atribuído nenhum critério de ordenação para os parâmetros de avaliação e também não foram consideradas ponderações para cada um dos parâmetros.

A Tabela V apresenta a Framework construída, onde se efetuou para cada um dos parâmetros de avaliação a identificação do âmbito em que se insere, a caracterização do seu significado/natureza, o enquadramento dos valores que pode assumir e o seu mapeamento na bibliografia relacionada. Os parâmetros refletem não só ferramentas e funcionalidades desejáveis para este tipo de *software*, mas também capacidades que resultam de tendências atuais para este tipo de aplicações, disponibilizadas já no mercado de *software* de MDM.

Na Tabela IX do Anexo A foram elencados e explicados todos os valores possíveis atribuídos a cada um dos parâmetros da *framework*.

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Tabela V - Framework para a Avaliação de Ferramentas de MDM.

ÂMBITO						CARACTERIZAÇÃO DO PARÂMETRO	VALORES POSSÍVEIS	REFERÊNCIAS BIBLIOGRÁFICAS
Master Data Quality	Segurança e Governance	Arquitetura	Controle e Auditoria	Interface	Plataforma Tecnológica			
Parâmetro P₁: Identity Resolution (Resolução de Identidades)								
X	X					<p>Processo operacional, tipicamente potenciado por um motor de <i>identity resolution</i> ou <i>middleware</i>, através do qual as organizações podem conectar diferentes fontes de dados para a compreensão de possíveis identidades coincidentes e relacionamentos pouco óbvios em diversos silos de dados. Aplica valores de probabilidade para determinar que entidades são coincidentes e que possíveis relacionamentos existem entre essas entidades.</p>	<ul style="list-style-type: none"> • Redução/eliminação de duplicados; • <i>Merge/purge</i>; • <i>Householding</i>; • Não possui. 	(Loshin, 2009); (Loshin, 2010); (Loshin, 2012)
Parâmetro P₂: Data Profiling (Perfil dos Dados)								
X						<p>Utilizado para a identificação quantitativa de questões de qualidade dos dados. É realizado com base na “<i>completeness</i>” dos campos, que determina a usabilidade do campo para <i>matching</i> (reunião e comparação de dados provenientes de diferentes fontes). Está relacionado com a identificação e modificação de padrões para <i>standardização</i> e <i>matching</i> de dados.</p>	<ul style="list-style-type: none"> • <i>Attribute Profiling</i>; • <i>Dependencies Profiling</i>; • <i>Redundancy Profiling</i>; • Não possui. 	(Loshin, 2010); (Cervo & Allen, 2011) (Howard, 2011) (Loshin, 2009); (Naumann & Sattler, 2006); (Loshin, 2012)
Parâmetro P₃: Parsing e Standardization (Análise e Normalização)								
X						<p>Ferramentas que combinam diferentes algoritmos para reconhecimento de padrões e validação dos valores de dados. A análise dos dados é efetuada por <i>tokens</i> que permitem posterior mapeamento dos dados fonte numa representação destino. Juntas formam uma combinação de técnicas utilizadas para verificar os valores de dados e compará-los com domínios de valores conhecidos, formatos de dados e padrões.</p>	<ul style="list-style-type: none"> • Nome pessoa; • Endereço e telefone; • Cliente; • Organização; • Produto; • Não possui. 	(Loshin, 2009); (Loshin, 2010); (Kumar, 2010); (Loshin, 2012)

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Parâmetro P ₄ : <i>Data Enhancement</i> (Enriquecimento dos Dados)								
x						<p>É o processo para aumentar o valor de uma instância de dados, ou conjunto de dados, adicionando conhecimento de valor acrescentado. O seu objetivo é identificar conhecimento utilizável das coleções de dados usado para melhorar os processos de negócio. Pode ser efetuado por abordagens em <i>batch</i> ou em <i>inline</i>.</p>	<ul style="list-style-type: none"> • <i>Batch</i>; • <i>Inline</i>; • Não possui. 	(Loshin, 2010); (Loshin, 2012)
Parâmetro P ₅ : <i>Data Lineage</i> (Origem dos Dados)								
x	x		x			<p>Capacidade da gestão de <i>metadata</i> com a funcionalidade de determinar de onde provêm os dados, de que forma são transformados e para onde vão. Traça o ciclo de vida da informação entre diversos sistemas, incluindo as operações que são executadas sobre os dados.</p>	<ul style="list-style-type: none"> • <i>App Browser</i>; • <i>Reporting</i>; • Não possui. 	(Loshin, 2009) (Otto, Humer, & Osterle, 2011)
Parâmetro P ₆ : Segurança e Auditoria								
	x		x			<p>Mecanismos que fazem parte da estrutura de governança dos dados e que gerem o acesso aos dados, nomeadamente através de serviços de controlo de acessos, autenticação, gestão de segurança e auditoria.</p>	<ul style="list-style-type: none"> • Controlo de Acessos; • Autenticação; • Gestão de segurança; • Auditoria. 	(Loshin, 2009)
Parâmetro P ₇ : Processos de <i>Data Governance</i> (Governança dos Dados)								
	x		x			<p>Asseguram que os dados cumprem as expectativas de todos os objetivos do negócio no contexto de administração de dados, conformidade, privacidade, segurança, sensibilidade dos dados, gestão da <i>metadata</i> e gestão da <i>master data</i>. Cada aspeto da governança dos dados está relacionado com a especificação de políticas de informação que refletem as necessidades do negócio.</p>	<ul style="list-style-type: none"> • Políticas; • Processos; • Controlo; • Auditoria. 	(Loshin, 2009); (Cervo & Allen, 2011); (Dreibelbis <i>et al.</i> , 2008); (Loshin, 2012); (Russom, 2012)
Parâmetro P ₈ : Domínios de <i>Master Data</i>								
		x				<p>Áreas compartimentadas através das quais se descrevem, caracterizam e classificam conceitos de negócio e são definidas políticas e regras para a natureza da <i>master data</i>.</p>	<ul style="list-style-type: none"> • <i>CDI (Customer Data Integration)</i>; • <i>PIM (Product Information Management)</i>; • <i>EIM (Employee Information Management)</i>; • <i>VIM (Vendor Information Management)</i>; • <i>Parts Information Management</i>; • <i>Location Information Management</i>; • <i>Distributor Information Management</i>; • <i>Multi-Domain</i>; • <i>Custom-Domain</i> 	(Dreibelbis <i>et al.</i> , 2008); (Cervo & Allen, 2011); (Loshin, 2009)

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Parâmetro P ₉ : Estilo de Implementação								
		✘				Abordagem de desenvolvimento do ambiente MDM com base em diferentes combinações de implementação e requisitos de utilização que levaram à evolução de vários estilos. São frequentes as implementações híbridas que combinam dois ou mais estilos de implementação.	<ul style="list-style-type: none"> • Consolidação; • <i>Registry</i>; • Coexistência; • <i>Hub</i> transacional; • Estilo híbrido. 	(Cervo & Allen, 2011); (Dreibelbis <i>et al.</i> , 2008) (Loshin, 2009)
Parâmetro P ₁₀ : Conectividade com as Fontes de Dados								
		✘		✘		Mecanismos e motores disponibilizados pela plataforma aplicacional para conexão da mesma a múltiplas fontes de dados e consequente geração de um repositório único de <i>master data</i> .	<ul style="list-style-type: none"> • JMS (<i>Java Message Service</i>); • <i>Web Service</i>; • FTP; • HTTP; • CDC (<i>Change Data Capture</i>); • Pacotes de conectores. 	(Cervo & Allen, 2011); (Dreibelbis <i>et al.</i> , 2008); (Otto & Hüner, 2009)
Parâmetro P ₁₁ : Sincronização de Master Data								
✘		✘				Técnica de atualização dos dados, com manutenção da sua consistência, que se encontra em diversas aplicações, bases de dados ou sistemas. A necessidade de sincronização dos dados pode ser permanente (entre sistemas operacionais) ou temporária (por exemplo numa migração). A metodologia de sincronização depende da arquitetura adotada (estilo de implementação).	<ul style="list-style-type: none"> • <i>Batch</i>; • <i>Inline</i>; 	(Loshin, 2009); (Cervo & Allen, 2011); (Loshin, 2012)
Parâmetro P ₁₂ : Modelo de Master Data								
✘		✘				Representação do nível conceptual/lógico da organização dos dados no serviço de <i>Master Data</i> , sendo constituída por objetos como entidades, atributos e grupos de atributos, hierarquias e coleções. Dependendo da filosofia do fabricante, poderão ser mais rígidos (pacote pré-definido na aplicação) ou mais flexíveis, admitindo customizações e adaptações à realidade do negócio onde a ferramenta será implementada.	<ul style="list-style-type: none"> • <i>Data Model editing</i>; • Modelização gráfica; • Suporte a <i>standards</i> de negócio; • Controlo de versões. 	(Loshin, 2009); (Dreibelbis <i>et al.</i> , 2008); (Wolter & Haselden, 2006); (Otto & Hüner, 2009); (Loshin, 2012)
Parâmetro P ₁₃ : Data Integration (Integração de Dados)								
		✘				Ferramentas e técnicas que permitem a criação de uma perspetiva geral de todos os dados da organização, independentemente do local físico em que se encontram localizados. O esforço de integração de dados pode acontecer numa única migração inicial de dados ou acontecer numa base contínua.	<ul style="list-style-type: none"> • Integração física; • Integração virtual; • Integração Inicial (única); • Integração Contínua. 	(Loshin, 2009); (Loshin, 2012); (Cervo & Allen, 2011)

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Parâmetro P ₁₄ : <i>Data Federation</i> (Federação de Dados)									
		✘	✘			<p>Técnica que consiste na virtualização de dados, em que os dados armazenados num conjunto heterogêneo de <i>data stores</i> autônomos são colocados à disposição como se tratando de um único <i>data store</i>. Esta técnica de integração de dados é normalmente implementada utilizando diversas ferramentas e combinações dessas ferramentas.</p> <ul style="list-style-type: none"> • ETL (<i>Extract, Transform, Load</i>); • EII (<i>Enterprise Information Integration</i>); • EAI (<i>Enterprise Application Integration</i>); • EDR (<i>Enterprise Data Replication</i>); • ECM (<i>Enterprise Content Management</i>); • Não possui. 	<p>(Loshin, 2009); (Loshin, 2012); (Cervo & Allen, 2011)</p>		
Parâmetro P ₁₅ : <i>Social MDM</i>									
		✘		✘		<p>Introdução no ambiente de <i>Master Data</i> de dados de <i>social media</i>, provenientes das redes sociais. O <i>social media master data</i> consiste em perfis existentes em serviços como o Facebook, principalmente em atividades B2C, o LinkedIn principalmente em atividades B2B e o Twitter em atividades da mesma natureza.</p> <ul style="list-style-type: none"> • Facebook; • LinkedIn; • Twitter; • Outros; • Não possui. 	<p>(Gartner, 2011)</p>		
Parâmetro P ₁₆ : <i>Mobile MDM</i>									
		✘		✘	✘	<p>Integração do <i>mobile computing</i> no ambiente de <i>master data</i> da organização, permitindo que utilizadores/dispositivos móveis possam aceder ao repositório de <i>master data</i> e à <i>social data</i> remotamente. Permite o acesso a informação específica para uma determinada localização, com o suporte de mecanismos de segurança e controle de acessos.</p> <ul style="list-style-type: none"> • iOS; • Android; • Outros dispositivos móveis; • Não possui. 	<p>(Messerschmidt & Stuben, 2011)</p>		
Parâmetro P ₁₇ : <i>Processos e Workflow</i>									
	✘	✘	✘			<p>Serviços e processos que suportam as políticas e tarefas/ações associadas ao utilizador ou com a utilização da aplicação da <i>master data</i>. Os processos de <i>workflow</i> autorizam tarefas através de regras ou SLA's e são especificamente configurados para integrar os eventos do ciclo de vida da <i>master data</i> nas aplicações correspondentes, no contexto específico do negócio.</p> <ul style="list-style-type: none"> • Agregação de atividades; • Modelização gráfica; • Gestão de <i>workflows</i>; • Não possui. 	<p>(Loshin, 2009); (Wolter & Haselden, 2006); (Otto & Hüner, 2009)</p>		
Parâmetro P ₁₈ : <i>Gestão de Metadata</i>									
		✘	✘			<p>Ferramentas e técnicas para a criação e gestão de uma base de dados concebida para a reunião e armazenamento de informação contextual relacionada com o negócio. Deverá executar mapeamentos e tabelas de referência entre os domínios de dados e dos valores de dados. Existem diversos <i>standards</i> para a <i>metadata</i>, cada um deles orientado para uma determinada área de atuação (natureza dos dados).</p> <table style="width: 100%; border: none;"> <tr> <td style="width: 50%; vertical-align: top;"> <p><u>Esquema:</u></p> <ul style="list-style-type: none"> • MARC21; • DublinCore; • MODS; • IEEE-LOM; • Onix; • ... </td> <td style="width: 50%; vertical-align: top;"> <p><u>Sintaxe:</u></p> <ul style="list-style-type: none"> • HTML; • XHTML; • XML; • RDF/XML; • METS; • ... </td> </tr> </table>	<p><u>Esquema:</u></p> <ul style="list-style-type: none"> • MARC21; • DublinCore; • MODS; • IEEE-LOM; • Onix; • ... 	<p><u>Sintaxe:</u></p> <ul style="list-style-type: none"> • HTML; • XHTML; • XML; • RDF/XML; • METS; • ... 	<p>(Loshin, 2009); (Dreibelbis <i>et al.</i>, 2008); (Cervo & Allen, 2011); (Otto & Hüner, 2009); (Moss, 2007)</p>
<p><u>Esquema:</u></p> <ul style="list-style-type: none"> • MARC21; • DublinCore; • MODS; • IEEE-LOM; • Onix; • ... 	<p><u>Sintaxe:</u></p> <ul style="list-style-type: none"> • HTML; • XHTML; • XML; • RDF/XML; • METS; • ... 								

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Parâmetro P ₁₉ : <i>Reporting</i> (Gestão de Relatórios)								
x			x			<p>Ferramenta de produção de relatórios e publicação de <i>master data</i> em diversos suportes. A atividade de <i>reporting</i> exige a existência de práticas estabelecidas para a validação dos dados, requerendo a supervisão sistemática dos processos que implementam as políticas da informação. O <i>reporting</i> produzido deverá cobrir as áreas de <i>data quality</i>, estatísticas de utilização, monitorização e suporte às auditorias.</p>	<ul style="list-style-type: none"> • Relatórios de <i>Data Quality</i>; • Estatísticas de utilização; • Monitorização de tarefas; • Suporte a auditorias. 	(Loshin, 2009); (Otto & Hüner, 2009)
Parâmetro P ₂₀ : <i>Usability</i> (Usabilidade)								
				x		<p>Facilidade de utilização da ferramenta de MDM. Inclui métodos de medição (<i>needs analysis</i>) e o estudo de princípios e regras a observar na manipulação da <i>master data</i> tendo em vista a eficiência e <i>design</i> da solução. Deve ser aferida a facilidade de utilização e fácil acesso a funcionalidades da <i>aplicação</i> de MDM, a sua performance, o tempo necessário para a aprendizagem de uma determinada função e eventuais erros na execução de funcionalidades.</p>	<ul style="list-style-type: none"> • <i>Completeness</i> (integridade); • Tempo de aprendizagem; • Erros; • <i>Performance</i>; • Satisfação; • Atitude. 	(Folmer & Bosch, 2002)
Parâmetro P ₂₁ : <i>Exception Handling</i> (Tratamento de Exceções)								
x				x		<p>Processo de resposta a ocorrências, durante o processamento de <i>master data</i>, de exceções (situações anómalas ou excepcionais que necessitam de processamento especial) que alteram tipicamente o fluxo normal da plataforma de MDM.</p>	<ul style="list-style-type: none"> • <i>Flagging</i>; • Correção automática; • Correção manual; • <i>Send back to provider</i>; • <i>Obliteration</i>; • Não possui. 	(Dreibelbis <i>et al.</i> , 2008); (Loshin, 2009)
Parâmetro P ₂₂ : Requisitos Tecnológicos do Sistema (S.O. / SGBD)								
		x		x	x	<p>Ambiente de <i>software</i> de base compatível, necessário e suficiente para o funcionamento da aplicação de MDM, nomeadamente ao nível do Sistema Operativo e do Sistema de Gestão de Base de Dados.</p>	<p>Sistema Operativo:</p> <ul style="list-style-type: none"> • AIX; • <i>Enterprise SuSE Linux</i>; • <i>Red Hat Enterprise Linux</i>; • Solaris; • <i>Windows Server 2008 R2</i>. • ... <p>SGBD:</p> <ul style="list-style-type: none"> • DB2; • Oracle; • SQL Server; • ... 	(Folmer & Bosch, 2002)

4.3. Avaliação de Software de MDM

Com o objetivo de dar seguimento à metodologia do *Design Science*, adotada no desenvolvimento do trabalho aqui apresentado, foi descrita nesta secção a fase de demonstração da metodologia, na qual se pretende provar que o conceito de conceção do artefacto resulta efetivamente. Neste sentido, tentou-se demonstrar a utilização da *framework* apresentada na secção anterior em duas instâncias do problema, através da aplicação do modelo desenhado a duas ferramentas de MDM disponíveis no mercado.

Para a aplicação da *framework*, foram consideradas duas soluções MDM que estão entre os principais produtos para a gestão de *master data* de cliente e de produto, são aplicações de utilização genérica, embora não sejam específicas ou especialmente desenhadas para uma determinada indústria (Zornes, 2012). O *The MDM Institute* referencia-as no “*top 50*” de ferramentas de MDM (The MDM Institute, 2012).

As ferramentas selecionadas foram: o *IBM InfoSphere Master Data Management Server* (versão 10.0) e o *Microsoft SQL Server 2012 Master Data Services*.

Na Tabela VI foram sistematizadas todas as fontes e referências bibliográficas utilizadas na avaliação técnica das aplicações referidas.

Tabela VI - Referências utilizadas na avaliação das ferramentas MDM.

FERRAMENTA	FONTES
IBM InfoSphere MDM Server	(Ballard, et al., 2011); (IBM, 2008); (InformationCenter, 2012); (InfoSphere, 2011); (Borean, 2011)
Ms SQL Server 2012 Master Data Services	(Graham & Selhorn, 2011); (Technet, 2012); (MSDN, 2012); (Dicken, 2010); (Microsoft-SQLServer, 2012)

O relatório *Enterprise MDM Market Review & Forecast for 2008-12* (MDM Institute) apresenta algumas considerações em relação às referidas ferramentas, destacando-se, para cada uma delas, as seguintes:

IBM InfoSphere MDM Server

Devido à sua origem e pontos fortes no mercado de serviços financeiros, este produto teve uma performance global muito boa no suporte a grandes bancos e companhias de seguros. Este conjunto multi-nível de serviços empresariais fornece *frameworks* aceleradoras mas muitas vezes à custa de um conjunto considerável de *software* da IBM. Ainda assim, esta é a solução MDM líder para os dados de cliente na banca e

seguros com adaptação e sucesso em outros mercados como o do retalho e comunicações. Em número de instalações e valor monetário associado às vendas (software e serviços), a IBM foi o líder destacado para o *Enterprise MDM*, durante o ano de 2007. Os desafios recentes de marketing da aplicação foram rapidamente ultrapassados com o lançamento de iniciativas e pacotes para reduzir a complexidade de instalação e customização (Zornes, 2008).

Ms SQL Server Master Data Services

A *Microsoft* teve um começo lento no mercado de MDM que está em rápido movimento, mas espera-se que o acompanhe em breve para fornecer capacidades de MDM robustas para o mercado das pequenas e médias empresas. Parte da sua estratégia para acompanhar o mercado passa pela aquisição de algumas das tecnologias mais complexas. Recentemente, a *Microsoft* adquiriu a *Stratature* e a *Zoomix* – duas empresas cujos produtos tiveram um sucesso modesto e poderão agora proliferar com a distribuição da *Microsoft* como parte da família formal de produtos e suas soluções integradas. Apesar de ter tido uma penetração tardia no mercado, espera-se que domine o mercado das PME nos próximos dois a cinco anos, através de um esquema de licenciamento e *marketing* apropriado para aquele mercado (Zornes, 2008).

No Anexo B foi sistematizada a avaliação realizada às duas ferramentas selecionadas.

4.4. Avaliação da Framework

Nesta secção está descrita a fase de avaliação da *framework* de DS utilizada no presente trabalho. Esta fase consiste em observar e medir se o artefacto desenhado suporta a solução para o problema a resolver (questão de investigação). Serão analisados os resultados obtidos na atividade de demonstração (avaliação de *software* MDM). Pretende-se retirar conclusões concretas acerca da eficácia da *framework* desenhada e se a mesma necessita de ajustes e melhorias. Para o efeito, três especialistas foram convidados a comentar a *framework* construída, relativamente à sua conceptualização, conteúdo e aplicabilidade prática. Os especialistas estão identificados na Tabela VII.

Tabela VII - Especialistas entrevistados.

Nome	Função	Organização	Data
Eng.º Manuel Dias (MD)	<i>Solution Specialist</i>	Microsoft	04.09.2012
Eng.º Bruno Marques (BM)	Consultor	Portugal Telecom	11.09.2012
Prof. Palma dos Reis (PR)	Professor/Investigador	UTL/ISEG	13.09.2012

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Através da avaliação de ferramentas executada neste trabalho foi possível tirar algumas conclusões acerca da *framework* desenhada, nomeadamente:

- A necessidade de uma matriz detalhada para uma aplicação mais prática;
- O parâmetro “**usability**” foi impossível de avaliar com a informação disponível;
- É necessário maior detalhe em relação aos valores possíveis da *framework*.

Na Tabela VIII estão sistematizados os comentários referidos na análise efetuada pelos especialistas convidados.

Tabela VIII – Avaliação da Framework.

Comentário	Referido por:
1. A <i>framework</i> é útil e poderá ter aplicabilidade prática;	(MD); (BM); (PR)
2. Em termos de características de MDM, a <i>framework</i> reúne todas aquelas que são fundamentais;	(MD)
3. A <i>framework</i> está genericamente bem estruturada;	(PR)
4. A <i>framework</i> possui parâmetros muito genéricos, mas está perceptível e objetiva;	(MD)
5. Faz sentido o enquadramento dos parâmetros da <i>framework</i> em áreas ou domínios de atuação;	(PR)
6. É necessário o conhecimento da forma como foi concebida a <i>framework</i> para a aplicar facilmente;	(BM)
7. As diferenças entre duas ferramentas de MDM só poderão ser apuradas através de um referencial mais detalhado;	(MD)
8. Os valores possíveis da <i>framework</i> não devem ser muito abrangentes sob pena de serem subjetivos e por isso estarem dependentes de quem avalia;	(PR)
9. O impacto de um determinado parâmetro pode ser maior ou menor em função do fim a que se destina a aplicação;	(PR)
10. Seria interessante desdobrar a <i>framework</i> numa <i>checklist</i> , em que cada parâmetro daria origem a um conjunto de funcionalidades;	(MD); (BM)
11. É importante atribuir pesos aos parâmetros de forma a poder aferir o grau de realização de cada um pela ferramenta de MDM;	(MD)
12. Seria interessante alinhar os parâmetros com o ciclo de vida dos dados numa fase inicial, numa fase de monitorização e numa fase de “ <i>discovery</i> ”;	(MD)
13. A <i>framework</i> deverá contextualizar um guião de análise que oriente o decisor a avaliar a solução de <i>software</i> em função do contexto em causa;	(PR)
14. A <i>framework</i> deveria possibilitar uma avaliação quantitativa das ferramentas;	(MD)
15. Na análise de um <i>software</i> de MDM poderá interessar avaliar as características do fabricante e do esquema de serviços;	(PR)
16. Seria interessante que a <i>framework</i> contemplasse uma análise que incluísse preços, TCO, ROI e recursos (equipa, formação,...);	(MD)
17. A <i>framework</i> deve ser complementada com outros vetores de análise relacionados com a organização e gestão do risco;	(MD)

5 Conclusões e Investigação Futura

5.1. Conclusões

O presente trabalho, através da sua abordagem metodológica, reuniu vinte e duas características significativas a considerar na avaliação de uma ferramenta de MDM. As referidas características deram origem a outros tantos parâmetros de avaliação, com os quais foi construída uma *framework* para a avaliação qualitativa de *software* de MDM. Através da aplicação da *framework* a duas ferramentas disponíveis no mercado e da sua análise realizada por especialistas na matéria, foi possível concluir que o modelo construído necessita de melhorias, como sejam a criação de um referencial mais detalhado e a diminuição da abrangência dos valores possíveis para os seus parâmetros. A validação da *framework* sugere ainda a inclusão na sua estrutura de outros vetores de análise como sejam: a vertente financeira, TCO, ROI, perfil do fabricante, portfólio de serviços e análise de risco.

A *framework* avançada pelo estudo apresentado poderá constituir uma ferramenta de suporte para gestores e equipas de TI, contextualizando um guião de análise que oriente os decisores na credibilização de investimentos em *software* na área do MDM.

5.2. Investigação Futura

O estudo apresentado perspetiva a realização de outros trabalhos de investigação. Algumas sugestões para trabalhos futuros são:

- A realização de uma ou mais iterações da metodologia baseada em DS que contemplem os contributos de um painel de peritos em MDM, fabricantes de *software*, implementadores de soluções MDM e organizações que implementaram soluções de MDM, tendo em vista o aperfeiçoamento da *framework*;
- A construção de uma *framework* de âmbito mais alargado que considere também outros vetores de análise como sejam a vertente financeira, TCO, ROI, perfil do fabricante, portfólio de serviços e análise de risco;
- Criação de uma matriz de análise, a partir do modelo da *framework*, para uma avaliação facilitada do *software* de *Master Data Management*.

6 Referências Bibliográficas

- The MDM Institute (2012). Obtido em 05 de junho de 2012, de The MDM Institute: <http://www.tcdii.com>.
- Ballard, C., Deswal, P., Flores, P., Guitard, P., Jia, C. & Pittman, M. (2011). *Master Data Management: IBM InfoSphere Rapid Deployment Package (Second Edition)*. NY - USA: Redbooks - IBM International Technical Support Organization.
- Borean, D. (2011). *Implementing InfoSphere MDM Server for High Performance: Best Practices*. (I. C. Limited, Ed.) Obtido em 30 de julho de 2012, de IBM: <http://www-01.ibm.com/software/data/master-data-management/>.
- Cervo, D. & Allen, M. (2011). *Master Data Management in Practice: Achieving True Customer MDM*. New Jersey: John Wiley & Sons.
- Cleven, A., & Wortmann, F. (2010). *Uncovering four strategies to approach Master Data Management*. 43rd Hawaii International Conference on System Sciences. 978-0-7695-3869-3/10. Hawaii: IEEE Computer Society.
- Dicken, D. (2010). *An Overview of SQL Server Master Data Services*. Obtido em 30 de julho de 2012, de *Database Journal - The Knowledge Center for Database Professionals*: <http://www.databasejournal.com/features/mssql/article.php/3898626/An-Overview-of-SQL-Server-Master-Data-Services.htm>.
- Dreibelbis, A., Hechler, E., Milman, I., Oberhofer, M. & Run, P. (2008). *Enterprise Master Data Management: An SOA Approach to Managing Core Information* (1^a ed.). New Jersey: IBM Press.
- Folmer, E. & Bosch, J. (2002). *Architecting for usability: a survey*. *The Journal of Systems and Software*, 70 (2004), 61–78.
- Galhardas, H., Torres, L. & Damásio, J. (2010). *Master Data Management: a proof of concept*. *International Conference on Information Quality*. USA: INESC-ID.
- Gartner. (2011). *Press releases: Gartner Highlights Three Trends That Will Shape the Master Data Management Market*. Obtido em 05 de junho de 2012, de Gartner website: <http://www.gartner.com/it/page.jsp?id=1666414>.
- Gartner. (2012). *Press Releases. Gartner Says Master Data Management Software Revenue to Grow 21% in 2012*. Obtido em 05 de junho de 2012, de Gartner website: http://www.gartner.com/it/page.jsp?id=1886314&goback=%2Eanp_2258008_13

- 37202500874_1%2Egmp_2258008%2Egde_2258008_member_87496081[16-05-2012 22:16:20].
- Graham, T. & Selhorn, S. (2011). *Microsoft SQL Server 2008R2 - Master Data Services: Implementation & Administration*. McGraw-Hill.
- Hevner, A., March, S., Park, J. & Ram, S. (2004). *Design Science in Information Systems Research*. MIS Quarterly. 28(1), 75-105.
- Howard, P. (2011). *Data Quality: more than just master data*. London, UK: Bloor Research.
- IBM. (2008). *InfoSphere Master Data Management Server: Overview*. (I. S. Group, Ed.) Obtido em 30 de julho de 2012, de IBM Corporation: http://www-947.ibm.com/support/entry/portal/documentation/software/information_management/infosphere_master_data_management_server.
- InformationCenter. (2012). *InfoSphere MDM Server v10*. Obtido em 30 de julho de 2012, de IBM: <http://publib.boulder.ibm.com/infocenter/mdm/v10r0m0/index.jsp>.
- InfoSphere. (2011). *Software product documentation*. Obtido em 30 de julho de 2012, de IBM: <http://www-01.ibm.com/support/docview.wss?uid=swg27013919>.
- ISO9241-11. (1998). *Guidance on usability (First ed.)*. International Organization for Standardization.
- Kumar, J. (2010). *Entity Standardization: An MDM Challenge*. Obtido em 14 de maio de 2012, de Infosys White Paper: <http://www.infosys.com/CRM/idea-center/Documents/entity-standardization-MDM-challenge.pdf>.
- Loshin, D. (2009). *Master Data Management (1ª ed.)*. USA: Morgan Kauffmann OMG Press. Elsevier.
- Loshin, D. (2010). *The Practitioner's Guide to Data Quality Improvement (1 ed.)*. USA: Morgan Kauffmann OMG Press. Elsevier.
- Loshin, D. (2012). *Choosing and understanding MDM software*. SearchDataManagement.com. Obtido em 10 de junho de 2012, de E-Guide: *Tips for Selecting MDM Software Tools*: http://searchdatamanagement.rl.techtarget.co.uk/detail/RES/1331224656_24.html.
- Messerschmidt, M. & Stuben, J. (2011). *Hidden Treasure: A global study on master data management*. Obtido em 20 de junho de 2012, de PricewaterhouseCoopers AG: www.pwc.com.
- Microsoft-SQLServer. (2012). *SQL Server - Enterprise Information Management*. Obtido em 30 de julho de 2012, de Microsoft Corporation:

- <http://www.microsoft.com/sqlserver/en/us/solutions-technologies/enterprise-information-management/master-data-management.aspx>.
- Moss, L. (2007). Critical Success Factors for *Master Data Management*. *Cutter IT Journal - The Journal of Information Technology Management*, 20, 7-12.
- MSDN. (2012). *Master Data Services*. Obtido em 30 de julho de 2012, de Microsoft Corporation: <http://msdn.microsoft.com/en-us/library/ee633763.aspx>.
- Naumann, F. & Sattler, K. (2006). *Information Quality: Fundamentals, Techniques, and Use*. Obtido em 20 de junho de 2012, de Hasso-Platter Institut website: http://www.hpi.uni-potsdam.de/fileadmin/hpi/FG_Naumann/publications/EDBT06Tutorial_IQ.pdf.
- Otto, B. & Ebner, V. (2010). Measuring Master Data Quality: Findings from an Expert Survey. *MKWI 2010 – Business Intelligence*, (pp. 1101-1112). Göttingen.
- Otto, B. & Hüner, K. (2009). *Functional Reference Architecture for Corporate Master Data Management*. University of St. Gallen - for Business Administration, Economics, Law and Social Sciences (HSG). Institute of Information Management, St. Gallen - Switzerland.
- Otto, B., Humer, K. & Osterle, H. (2011). Toward a functional reference model for master data quality management. *Information Systems and E-Business Management*.
- Peppers, K., Tuunanen, T., Rotherberger, M. & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3), 45-77.
- Radcliffe, J. (2010). *Magic Quadrant for Master Data Management of Customer Data*. RAS Core Research Note G00206031. Gartner Inc.
- Russom, P. (2012). *Seven Reasons Why Master Data Management needs Data Governance*. *TDWI Checklist Report*. TDWI Reseach.
- Serrano, A. & Caldeira, M. (2001). *Gestão de investimentos em Sistemas e Tecnologias de Informação – uma revisão crítica*. *Revista Sistemas de Informação*, 15, 99-207.
- Srinivasan, D. (2010). *Using Lessons learnt to create a robust master data program: MDM in a box*. *MDM & Data Governance Summit*. Singapore: Cognizant Technology Solutions.
- Technet. (2012). *Master Data Services*. Obtido em 30 de julho de 2012, de Microsoft Corporation: <http://lab.technet.microsoft.com/pt-br/library/hh231022%28v=sql.110%29.aspx>.

- Wang, L., Ming, X. & You, J. (2009). *The steps and methodology of identifying master data from business processes*. *World Conference on Software Engineering*. 978-0-7695-3570-8/09. IEEE Computer Society.
- Webster, J. & Watson, R. (2002). *Analyzing the past To prepare the future: Writing a Literature Review*. *Webster & Watson/Guest Editorial. MIS Quarterly*, 26(2), XIII-XXIII.
- White, C. (2007). *Using Master Data in Business Intelligence*. (T. Bishop, Ed.) *BI Research*.
- White, A. & Radcliff, J. (2010). *Hype Cycle for Master Data Management, 2010*. ID Number: G00206123. Gartner Inc.
- Wolter, R. & Haselden, K. (2006). *The What, Why, and How of Master Data Management*. Obtido em 10 de março de 2012, de Microsoft: <http://msdn.microsoft.com/en-us/library/bb190163.aspx>.
- Zornes, A. (2008). *Enterprise Master Data Management: Market Review & Forecast for 2008-12*. Obtido em 17 de julho de 2012, de MDM Institute MarketPulse™: <http://www.tcdii.com/mdmresearch/industrywhitepapers.html>.
- Zornes, A. (2012). *MDM & Data Quality Vendor Products*. Obtido em 05 de junho de 2012, de <http://www.tcdii.com/links/vendorproducts.html>.

ANEXO A

VALORES POSSÍVEIS PARA OS PARÂMETROS DA FRAMEWORK

Tabela IX - Framework: os valores possíveis.

<p>Parâmetro P₁: Identity Resolution (Resolução de Identidades)</p> <p>Segundo Loshin (2010), são três as aplicações da <i>identity resolution</i>:</p> <ul style="list-style-type: none"> • Redução/eliminação de duplicados - consiste na identificação de múltiplas representações da mesma entidade num conjunto de dados e eliminação de todas essas, exceto uma; • Merge/purge - processo similar ao anterior mas que envolve a agregação de vários conjuntos de dados, seguida de eliminação de duplicados; • Householding - resolução de um conjunto de características associadas a conjuntos de dados que mapeiam com uma entidade de ordem mais elevada.
<p>Parâmetro P₂: Data Profiling (Perfil dos Dados)</p> <p>Os valores possíveis para este parâmetro correspondem aos três principais tipos de <i>data profiling</i>, que são:</p> <ul style="list-style-type: none"> • Attribute Profiling – processo de <i>data profiling</i> ao nível do atributo; • Dependencies Profiling – processo de <i>data profiling</i> ao nível das dependências existentes nos dados; • Redundancy Profiling – Processo de <i>data profiling</i> ao nível da redundância e duplicação de dados.
<p>Parâmetro P₃: Parsing e Standardization (Análise e Normalização)</p> <p>A combinação destas duas técnicas é em regra executada em função do tipo de dados em análise ou especializada em termos de um determinado domínio de dados, ao qual é aplicado o processo de análise e normalização. Os valores possíveis para este parâmetro são alguns exemplos de tipologias ou domínios de dados (nome, endereço, telefone, cliente, organização, produto, ...).</p>
<p>Parâmetro P₄: Data Enhancement (Enriquecimento dos Dados)</p> <p>O <i>data enhancement</i> possui essencialmente dois tipos de abordagem:</p> <ul style="list-style-type: none"> • Batch enhancement - extração de coleções de dados e execução de análise (<i>parsing</i>), normalização (<i>standardization</i>), consolidação e agregação; • Inlined enhancement - melhoria incremental de instâncias de dados.
<p>Parâmetro P₅: Data Lineage (Origem dos Dados)</p> <p>A aplicação deverá possuir capacidades de:</p> <ul style="list-style-type: none"> • App Browser – módulo aplicacional para consulta e edição; • Reporting – geração de relatórios e listagens.
<p>Parâmetro P₆: Segurança e Auditoria</p> <p>Segundo Loshin (2009), os principais serviços são:</p> <ul style="list-style-type: none"> • Controlo de acessos - Atores em diferentes papéis possuem diferentes níveis de acesso a objetos de <i>master data</i> específicos. O controlo de acessos é gerido com base na classificação dos papéis assignados a cada ator; • Autenticação - Qualificam os solicitantes para garantir que são quem dizem ser; • Gestão da segurança - Supervisionam a gestão do controlo de acessos e informação de autenticação, incluindo a gestão dos dados para os atores (pessoas ou aplicações) autorizados a aceder ao repositório de <i>master data</i>; • Auditoria - Gerem o registo de acessos (<i>logging</i>) bem como a análise e apresentação do histórico de acessos.

Parâmetro P₇: Processos de *Data Governance* (Governança dos Dados)

Os valores possíveis para este parâmetro têm a ver com elementos fundamentais para a implementação de uma iniciativa de *master data governance*, são eles:

- **Políticas** - normas e regras que estabelecem os responsáveis pelo acesso e gestão da *master data*;
- **Processos** – para garantir a forma como os dados são armazenados, arquivados, salvaguardados e protegidos de ataques e furto;
- **Controles** – indicadores para a aferição de regras de conformidade;
- **Auditoria** – procedimentos de verificação da conformidade de regras e normas configuradas no sistema.

Parâmetro P₈: Domínios de *Master Data*

As aplicações de MDM estão, regra geral, especializadas para endereçar um ou mais domínios de *master data*. O domínio de *master data* é um parâmetro obrigatório, tendo em vista assegurar a adequação da ferramenta à realidade da organização. Os tipos mais comuns são:

- **CDI** – Para dados de cliente;
- **PIM** – Para dados de produto;
- **EIM** – Para dados de empregado;
- **VIM** – Para dados de fornecedor;
- **Parts Information Management** – Para dados de peças e sobressalentes;
- **Location Information Management** – Para dados relativos a localizações;
- **Distributor Information Management** – Para dados relativos a distribuidor;
- **Multi-Domain** – Para diversos domínios em simultâneo;
- **Custom-Domain** – Flexibilização para criação de domínio(s) de *master data* à medida.

Parâmetro P₉: Estilo de Implementação

Segundo Dreibelbis *et al.* (2008), os estilos de implementação existentes são:

- **Consolidação;**
- **Registry;**
- **Coexistência;**
- **Hub transacional;**
- **Estilo híbrido.**

Nota: Os estilos de implementação foram apresentados na secção 2.1.1 do presente trabalho.

Parâmetro P₁₀: Conectividade com as Fontes de Dados

Os valores possíveis para este parâmetro correspondem a alguns mecanismos e protocolos de ligação do repositório de dados a diversas fontes de dados:

- **JMS** – *Standard de messaging* que permite que componentes aplicativos baseados em JEE criem, enviem e leiam mensagens;
- **Web Service** – Plataforma aplicacional que suporta interoperabilidade entre diferentes sistemas através do envio e receção de dados, tipicamente em formato XML;
- **FTP** – Protocolo utilizado para a transferência de ficheiros numa rede;
- **HTTP** – Protocolo de comunicação utilizado para sistemas de informação distribuídos e colaborativos;
- **CDC** – Abordagem de integração de dados baseada na identificação, captura e alterações efetuadas nas fontes de dados;
- **Pacotes de conectores** – *Bundle* de adaptadores para diferentes SGBDs.

Parâmetro P₁₁: Sincronização de *Master Data*

Os valores possíveis para este parâmetro são as duas grandes abordagens utilizadas para a atualização da *master data* no repositório de dados. A metodologia utilizada na sincronização dos dados depende do estilo de implementação adotado (parâmetro P₉). Os valores possíveis são:

- **Batch;**
- **Inline.**

Parâmetro P₁₂: Modelo de *Master Data*

Os valores possíveis para este parâmetro dizem respeito às características que facilitam a criação e gestão do modelo de *master data*, são elas:

- **Data Model editing** – capacidade de edição e manipulação do modelo;
- **Modelização gráfica** - interface gráfica par a gestão do modelo;
- **Suporte a standards de negócio** (regras e normas internacionais);
- **Controlo de versões.**

Parâmetro P₁₃: *Data Integration* (Integração de Dados)

Os valores possíveis para este parâmetro têm a ver com duas abordagens diferentes para o esforço de integração de dados numa iniciativa de MDM.

Numa perspetiva da localização dos dados:

- **Integração física** – existe transporte físico dos dados para uma localização comum;
- **Integração virtual** – são apenas mantidos apontadores para a localização física dos dados (que neste caso permanece inalterada);

Numa perspetiva de frequência da integração dos dados:

- **Integração Inicial (única)** – acontece numa única migração inicial dos dados;
- **Integração Contínua** – acontece com uma determinada frequência, numa base contínua.

Parâmetro P₁₄: *Data Federation* (Federação de Dados)

Os valores possíveis para este parâmetro são ferramentas que poderão suportar esta técnica de integração de dados. Os valores são:

- **ETL (*Extract, Transform, Load*)** – ferramenta de *software* cuja função é extração de dados de diversos sistemas, transformação desses dados e a carga dos mesmos num repositório;
- **EII (*Enterprise Information Integration*)** – técnica de integração de dados cujo objetivo é fazer com que um conjunto de fontes de dados heterogéneas surjam para o utilizador ou sistema como uma única fonte homogénea de dados;
- **EAI (*Enterprise Application Integration*)** – ferramentas que viabilizam a interação entre fontes de dados de sistemas heterogéneos, através de serviços;
- **EDR (*Enterprise Data Replication*)** – ferramentas para a replicação dos objetos de dados a partir dos diferentes sistemas existentes no ambiente organizacional;
- **ECM (*Enterprise Content Management*)** – ferramentas que permitem capturar, gerir, armazenar, preservar e entregar informação crucial para a operação do negócio.

Parâmetro P₁₅: *Social MDM*

Os valores possíveis para este parâmetro correspondem às redes sociais mais utilizadas e que serão potenciais fontes de *master data* para o repositório de MDM:

- **Facebook;**
- **LinkedIn;**
- **Twitter;**
- **Outros.**

Parâmetro P₁₆: *Mobile MDM*

Os valores possíveis para este parâmetro correspondem às principais plataformas de dispositivos móveis como iPads, tablets e SmartPhones:

- **iPad (iOS);**
- **Android;**
- **Outros dispositivos móveis.**

Parâmetro P₁₇: Processos e *Workflow*

Os valores possíveis para este parâmetro consistem nas características principais que são executadas pelas ferramentas de gestão de processos e *workflow* (fluxo de trabalho):

- **Agregação de atividades** – *bundle* de processos que correspondem a atividades do negócio;
- **Modelização gráfica** – capacidade de criação e gestão através de um modo gráfico;
- **Gestão de workflows** – capacidade de modelização e gestão de um conjunto de tarefas a realizar por diferentes atores, no âmbito de um processo de negócio.

Parâmetro P₁₈: Gestão de Metadata

Os valores possíveis para este parâmetro são *standards* existentes para a *metadata*, quer ao nível do esquema, quer ao nível da sintaxe utilizada. Por uma questão de espaço, apenas se indicam alguns *standards*:

Esquema:

- **MARC21** – *standard* para a representação e comunicação de informação bibliográfica;
- **DublinCore** – *standard* que inclui um grande conjunto de vocabulários de *metadata* e especificações técnicas;
- **MODS** – esquema para um conjunto de elementos bibliográficos a manipular numa plataforma aplicacional;
- **IEEE-LOM** – norma que especifica a sintaxe e semântica da *Learning Object Metadata*;
- **Onix** – *standard* internacional para a representação e comunicação de informação.

Sintaxe:

- HTML;
- XHTML;
- XML;
- RDF/XML;
- METS;

Parâmetro P₁₉: Reporting (Gestão de Relatórios)

Segundo Otto & Hüner (2009), o *reporting* de uma ferramenta de MDM deverá cobrir quatro grandes áreas. Essas áreas correspondem aos valores possíveis para este parâmetro, são elas:

- **Relatórios de Data Quality** - permitem ilustrar os resultados da análise de dados através de *templates* (modelos) ou diagramas a serem utilizados em *dashboards*;
- **Estatísticas de utilização** - permitem registar em tempo real quem está a utilizar ou a requisitar dados;
- **Monitorização de tarefas** - permite monitorizar funções automáticas e avaliá-las segundo diversos indicadores;
- **Suporte a auditorias** – criação de relatórios para verificação de conformidade de *standards* ou SLA's.

Parâmetro P₂₀: Usability (Usabilidade)

Segundo a *framework* construída por Folmer & Bosch (2002), os indicadores para a medida da usabilidade dividem-se em duas categorias: indicadores de performance e indicadores da perspectiva do utilizador. Os valores possíveis para este parâmetro correspondem aos indicadores de avaliação, que são:

- **Completeness** – capacidade para atingir objetivos definidos;
- **Tempo de aprendizagem** – necessário para a execução de tarefas;
- **Erros** – o sistema deve permitir uma baixa taxa de erros de utilização;
- **Performance** – medida da eficiência na execução das tarefas;
- **Satisfação** – conforto e aceitabilidade da utilização do *software* (indicador de carácter subjetivo);
- **Atitude** – satisfação do utilizador com o sistema (indicador de carácter subjetivo).

Parâmetro P₂₁: Exception Handling (Tratamento de Exceções)

Os valores possíveis para este parâmetro correspondem a abordagens de tratamento de exceções:

- **Flagging** – processo de marcação do conjunto de dados para posterior tratamento de erros;
- **Correção automática** – tratamento de exceções de forma automatizada;
- **Correção manual** – tratamento manual de exceções (não existem automatismos configurados);
- **Send back to provider** – recusa dos conjuntos de *master data* que originaram o erro/exceção;
- **Obliteration** – não é tomada qualquer ação sobre o erro detetado (a exceção é ignorada);

Parâmetro P₂₂: Requisitos Tecnológicos do Sistema (S.O. / SGBD)

Os valores possíveis para este parâmetro correspondem a sistemas operativos e sistemas de gestão de bases de dados possíveis. Por razões meramente ilustrativas só se indicam alguns exemplos:

Sistema Operativo:

- AIX;
- Enterprise SuSE Linux;
- Red Hat Enterprise Linux;
- Solaris;
- Windows Server 2008 R2.

SGBD:

- DB2;
- Oracle;
- SQL Server;

ANEXO B

AVALIAÇÃO DE FERRAMENTAS DE MDM

GRELHA DE AVALIAÇÃO DE SOFTWARE

FABRICANTE:	Microsoft Corporation	DATA:	29-07-2012
PRODUTO:	SQL Server – Master Data Services (MDS)		
VERSÃO:	2012		

VALORES POSSÍVEIS	OBSERVAÇÕES
Parâmetro P₁: Identity Resolution (Resolução de Identidades)	
<ul style="list-style-type: none"> Redução/eliminação de duplicados. 	<ul style="list-style-type: none"> A funcionalidade de <i>Data Quality Services</i> foi integrada no <i>Master Data Services</i> (MDS) para que seja possível executar tarefas de <i>identity resolution</i> nos dados fonte a integrar no repositório de <i>master data</i>; Os componentes do <i>Data Quality Server</i> têm de ser instalados com o MDS.
Parâmetro P₂: Data Profiling (Perfil dos Dados)	
<ul style="list-style-type: none"> <i>Attribute profiling</i>; <i>Dependencies profiling</i>. 	<ul style="list-style-type: none"> Esta tarefa de <i>data quality</i> é efetuada pela <i>feature</i> do SQL Server - <i>Data Quality Services</i> (DQS); O <i>data profiling</i> efetuado pelos DQS é o processo de analisar os dados numa fonte de dados existente e de apresentar estatísticas acerca dos dados nas atividades dos DQS. Apresenta medidas automáticas de <i>data quality</i> e é dinâmico e ajustável; Nesta ferramenta o <i>profiling</i> tem dois grandes objetivos: guiar o utilizador nos processos de <i>data quality</i> e suportar as suas decisões e, por outro lado, avaliar a eficácia do processo.
Parâmetro P₃: Parsing e Standardization (Análise e Normalização)	
<ul style="list-style-type: none"> Não especificado nas fontes disponíveis. 	<ul style="list-style-type: none"> Não existem <i>features</i> no MDS para suportar a normalização de dados; Este processo de <i>data quality</i> tem de ser implementado utilizando as funções disponíveis do <i>Server Integration Services</i> (SSIS).
Parâmetro P₄: Data Enhancement (Enriquecimento dos Dados)	
<ul style="list-style-type: none"> <i>Batch</i>. 	<ul style="list-style-type: none"> Não existe capacidade da <i>feature</i> do MDS para suportar o <i>data enhancement</i>; Este processo de <i>data quality</i> tem de ser implementado utilizando as funções disponíveis do SQL Server <i>Integration Services</i> (SSIS).

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Parâmetro P ₅ : <i>Data Lineage</i> (Origem dos Dados)	
<ul style="list-style-type: none"> • App Browser; • Reporting. 	<ul style="list-style-type: none"> • Ferramenta integrada no <i>SQL Server Integration Services</i> (SSIS) e assegurada por aquela funcionalidade; • Os processos de <i>data lineage</i> podem ser manipulados de diversas formas no SSIS. A forma mais comum é adicionar um número de incrementação automática ao fluxo de dados que executa o mapeamento dos objetos de dados relativamente às suas fontes.
Parâmetro P ₆ : Segurança e Auditoria	
<ul style="list-style-type: none"> • Controlo de acessos; • Autenticação; • Gestão de segurança; • Auditoria. 	<ul style="list-style-type: none"> • O MDS tem como base a <i>Active Directory</i> para a autenticação de utilizadores e grupos; • Uma vez que todas as permissões de segurança são armazenadas na base de dados do MDS, as <i>passwords</i> e os grupos não são geridos no MDS; • A segurança no MDS está orientada em 3 áreas: segurança funcional, segurança de objetos e segurança de membros numa hierarquia.
Parâmetro P ₇ : Processos de <i>Data Governance</i> (Governança dos Dados)	
<ul style="list-style-type: none"> • Processos; • Controlos. 	<ul style="list-style-type: none"> • As capacidades de <i>data governance</i> no <i>Master Data Services</i> estão limitadas à execução de controlo de acessos para os diferentes elementos de dados no repositório; • Não possui capacidades próprias de <i>workflow</i>, dependendo fortemente de motores de <i>workflow</i> externos para suportar os requisitos de gestão de dados.
Parâmetro P ₈ : Domínios de <i>Master Data</i>	
<ul style="list-style-type: none"> • <i>Custom-Domain</i>. 	<ul style="list-style-type: none"> • Aplicação "<i>domain agnostic</i>"; • O <i>Master Data Services</i> não foi concebido para um domínio específico; • Qualquer tipo de dados e virtualmente qualquer esquema de dados pode ser suportado por um sistema MDS.
Parâmetro P ₉ : Estilo de Implementação	
<ul style="list-style-type: none"> • Consolidação; • <i>Hub</i> transacional. 	<ul style="list-style-type: none"> • O <i>Master Data Services</i> está mais habilitado para os estilos de implementação de consolidação e <i>hub</i> transacional; • O estilo em <i>registry</i> não é tão bem endereçado pelo MDS, não havendo suporte para este estilo de implementação (tem de ser construído para a instalação de MDS ao nível do modelo de dados).

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Parâmetro P ₁₀ : Conectividade com as Fontes de Dados	
<ul style="list-style-type: none"> • <i>Web Service</i>; • FTP; • HTTP; • CDC; • Pacotes de conectores. 	<ul style="list-style-type: none"> • A integração de dados é assegurada pelo <i>SQL Server Integration Services (SSIS)</i> que possui uma série de gestores de ligação, incluindo gestores de ODBC e OLEDB utilizados para a ligação de um grande número de fontes de dados (SQL Server, Oracle, DB2, MySQL, ...); • Estão disponíveis gestores de ligação para aplicações como SAP e Teradata; • O SSIS também possui gestores para a ligação de ficheiros em formato Excel, XML e ficheiros de texto; • É possível a gestão de ligações FTP e <i>Web services</i>; • Extração e identificação fácil dos dados modificados com suporte para <i>Change Data Capture (CDC)</i>; • A documentação consultada refere a existência de pacotes de conectores para diferentes bases de dados, não especificando exatamente quais.
Parâmetro P ₁₁ : Sincronização de Master Data	
<ul style="list-style-type: none"> • <i>Batch</i>; • <i>Inline</i>. 	<ul style="list-style-type: none"> • O <i>Master Data Services</i> possui ferramentas de integração e de sincronização de dados que funcionam através de tabelas de <i>staging</i> aquando da importação de dados para o repositório e para atualização dos dados em aplicações a montante; • A aplicação admite as duas abordagens como metodologia de sincronização de dados.
Parâmetro P ₁₂ : Modelo de Master Data	
<ul style="list-style-type: none"> • <i>Data model editing</i>; • Modelização gráfica; • Controlo de versões. 	<ul style="list-style-type: none"> • A estrutura do modelo é criada na área funcional da administração do sistema do <i>Master Data Manager web application</i>, ou através da utilização dos <i>web services</i>; • Para novos utilizadores do MDS existem duas formas de criação de um novo modelo de dados: <ul style="list-style-type: none"> ▪ Carregamento de um pacote de <i>deployment</i> do modelo na base de dados e modificar o modelo; ▪ Construção de um modelo de raiz; • A utilização de um modelo pré-definido pode ser uma forma mais fácil de abordar o <i>Master Data Services</i> do <i>SQL Server</i>.
Parâmetro P ₁₃ : Data Integration (Integração de Dados)	
<ul style="list-style-type: none"> • Integração física; • Integração contínua. 	<ul style="list-style-type: none"> • É possível importar a <i>master data</i> das fontes de dados existentes através de um processo de <i>staging</i> dos dados na base de dados do MDS; • O processo de <i>staging</i> pode ser efetuado utilizando o <i>SQL Server Import wizard</i> ou o <i>SQL Server Integration Services</i>; • Após o processo de <i>staging</i> é utilizado o <i>Master Data Manager</i> para processar os dados em <i>staging</i> em <i>batch</i>. O MDS move depois os dados válidos das tabelas de <i>staging</i> para as tabelas da <i>master data</i> na BD do MDS.

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Parâmetro P ₁₄ : <i>Data Federation</i> (Federação de Dados)	
<ul style="list-style-type: none"> • ETL; • EII. 	<ul style="list-style-type: none"> • O <i>Master Data Services</i> não fornece nenhuma funcionalidade de <i>extract, transform, and load</i> (ETL); • O MDS está desenhado para suportar o <i>SQL Server Integration Services</i>, embora também funcione bem com outras soluções de ETL (Informatica, InfoSphere DataStage), ou qualquer outro produto da área; • A solução não consiste realmente numa ferramenta de <i>data federation</i> pura, trabalhando sobre uma arquitetura de <i>staging</i> de dados.
Parâmetro P ₁₅ : <i>Social MDM</i>	
<ul style="list-style-type: none"> • Não possui. 	<ul style="list-style-type: none"> • Não foram encontradas referências nas fontes disponíveis sobre a possibilidade de integração de <i>social data</i> no repositório de <i>master data</i>; • É possível a conexão a fontes de dados localizadas na <i>cloud</i>.
Parâmetro P ₁₆ : <i>Mobile MDM</i>	
<ul style="list-style-type: none"> • Não possui. 	<ul style="list-style-type: none"> • Não há indicações na documentação do fabricante de existir uma versão do <i>software</i> especialmente concebida para as plataformas iOS ou Android.
Parâmetro P ₁₇ : <i>Processos e Workflow</i>	
<ul style="list-style-type: none"> • Gestão de <i>workflows</i>. 	<ul style="list-style-type: none"> • O MDS permite a utilização de regras de negócio para criar soluções básicas de <i>workflow</i>; • Os dados podem ser automaticamente atualizados e validados e é possível ter notificações de <i>e-mail</i> enviadas com base em condições especificadas pelo utilizador; • Caso as necessidades de <i>workflow</i> envolvam de um processamento de eventos mais complexo, é possível a configuração do MDS para enviar dados para o <i>SharePoint</i> e iniciar um <i>workflow</i>.
Parâmetro P ₁₈ : <i>Gestão de Metadata</i>	
<ul style="list-style-type: none"> • XML. 	<ul style="list-style-type: none"> • A funcionalidade <i>MDS metadata</i> contém um modelo de sistema para permitir que os administradores adicionem atributos aos objetos do MDS; • Este modelo de sistema é o único modelo existente na plataforma; • As entidades no modelo de <i>metadata</i> não podem ser apagadas por nenhum utilizador, seja qual for o seu nível de permissões; • Não foi possível aferir na documentação existente qual o <i>standard</i> utilizado para o esquema de <i>metadata</i>.

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Parâmetro P ₁₉ : <i>Reporting</i> (Gestão de Relatórios)	
<ul style="list-style-type: none"> • Relatórios de <i>Data Quality</i>; • Estatísticas de utilização; • Suporte a auditorias. 	<ul style="list-style-type: none"> • O <i>SQL Server Reporting Services (SSRS)</i> é uma plataforma de <i>reporting</i> de servidor que fornece uma funcionalidade abrangente de <i>reporting</i> para uma grande variedade de fontes de dados; • As ferramentas de <i>reporting</i> funcionam no ambiente do <i>Ms Visual Studio</i> e estão completamente integradas com as ferramentas e componentes do <i>SQL Server</i>.
Parâmetro P ₂₀ : <i>Usability</i> (Usabilidade)	
N/A	<ul style="list-style-type: none"> • Não foi possível encontrar informação disponibilizada pelo fabricante que permita a avaliação e caracterização deste parâmetro.
Parâmetro P ₂₁ : <i>Exception Handling</i> (Tratamento de Exceções)	
<ul style="list-style-type: none"> • <i>Flagging</i>; • Correção manual. 	<ul style="list-style-type: none"> • O MDS move os dados válidos no processo de integração de dados, a partir das tabelas de <i>staging</i>, para as tabelas de <i>master data</i> na base de dados do MDS e assinala (<i>flagging</i>) os registos inválidos para serem corrigidos na fonte de onde provêm. São depois alvo de um processo de <i>restaging</i>.
Parâmetro P ₂₂ : Requisitos Tecnológicos do Sistema (S.O. / SGBD)	
<p>Sistema Operativo:</p> <ul style="list-style-type: none"> • <i>Windows Vista Business</i>; • <i>Windows Vista Enterprise</i>; • <i>Windows Vista Ultimate</i>; • <i>Windows 7 Business</i>; • <i>Windows 7 Enterprise</i>; • <i>Windows 7 Ultimate</i>; • <i>Windows Server 2008 R2 SP1</i>; • <i>Windows Server 2008 SP2</i> <p>SGBD:</p> <ul style="list-style-type: none"> • <i>SQL Server 2008 R2 Enterprise (64-bit)</i>; • <i>SQL Server 2008 R2 Developer (64-bit)</i>; • <i>SQL Server 2012 Business Intelligence (64-bit)</i>; • <i>SQL Server 2012 Business Enterprise (64-bit)</i>; • <i>SQL Server 2012 Business Developer (64-bit)</i>; 	<ul style="list-style-type: none"> • O <i>Master Data Services</i> é uma <i>feature</i> das versões e edições do <i>SQL Server</i>; • O computador onde a Instalação do <i>SQL Server</i> é executada deve atender aos requisitos mínimos da instalação do <i>SQL Server</i>, do aplicativo <i>Web</i> do <i>Master Data Manager</i>, dos serviços <i>Web</i> e do banco de dados do <i>Master Data Services</i> (se for hospedado no mesmo computador do aplicativo <i>Web</i>); • O <i>Master Data Services</i> não tem suporte como parte de uma instalação da instância em cluster do <i>SQL Server</i> (não tem suporte em <i>cluster</i>).

GRELHA DE AVALIAÇÃO DE SOFTWARE

FABRICANTE:	IBM	DATA:	29-07-2012
PRODUTO:	InfoSphere Master Data Management Server		
VERSÃO:	v10 – Enterprise Edition		

VALORES POSSÍVEIS	OBSERVAÇÕES
Parâmetro P₁: Identity Resolution (Resolução de Identidades)	
<ul style="list-style-type: none"> • Redução/eliminação de duplicados; • Merge/purge. 	<ul style="list-style-type: none"> • Esta característica é assegurada no modulo do <i>InfoSphere MDM Server</i>, que consiste nos componentes dos controladores Txn ou Finder; • Esta funcionalidade de <i>Identity Resolution</i> pode também ser efetuada através do <i>InfoSphere Identity Insite</i> (não pertence ao software em avaliação), integrado com o <i>InfoSphere MDM Server</i>, que através de algoritmos avançados reconhece identidades similares, as suas relações e atividades.
Parâmetro P₂: Data Profiling (Perfil dos Dados)	
<ul style="list-style-type: none"> • Não possui. 	<ul style="list-style-type: none"> • Esta funcionalidade é assegurada pelo <i>IBM InfoSphere Information Server</i> que poderá ser integrado no <i>InfoSphere MDM Server</i> e desta forma realizar tarefas de <i>data profiling</i>; • O <i>IBM InfoSphere Information Analyzer</i> fornece amplas capacidades para o <i>profiling</i> dos dados fonte; • As principais funções de <i>data profiling</i> são análise de colunas, análise de chaves primárias, análise de chaves estrangeiras e análise interdomínios.
Parâmetro P₃: Parsing e Standardization (Análise e Normalização)	
<ul style="list-style-type: none"> • Nome pessoa; • Endereço e telefone; • Cliente; • Organização; • Produto; • ... 	<ul style="list-style-type: none"> • A aplicação possui uma <i>framework</i> de <i>Data standardization</i>; • Na perspetiva desta ferramenta, a normalização dos dados assegura que o <i>InfoSphere MDM Server</i> armazene dados que são, não só de alta qualidade, mas que obedecem também a uma representação <i>standard</i> definida; • A normalização é invocada de forma condicional com base em regras ou condições definidas no repositório de <i>metadata</i>; • A <i>framework</i> de normalização de entidades utiliza restrições associadas a “normalizadores” que determinam se os mesmos devem ser aplicados ou não durante uma transação.

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Parâmetro P ₄ : <i>Data Enhancement</i> (Enriquecimento dos Dados)	
<ul style="list-style-type: none"> • <i>Batch</i>. 	<ul style="list-style-type: none"> • O <i>InfoSphere MDM Server</i> possui um processo a que chama “<i>batch enrichment</i>” que utiliza a <i>feature InfoSphere MDM Server Batch Processor</i>; • O fabricante pré-configurou a conexão com bases de dados da <i>Dun & Bradstreet</i>; • Aquela ligação lê o ficheiro de <i>batch enrichment</i> de entrada (ficheiro de resposta) da <i>Dun & Bradstreet</i>, realiza <i>parsing</i> aos dados e executa a transação <i>enrichOrgWithDnBData</i> para efetuar a atualização ou acrescentar informação organizacional e de hierarquia ao repositório de dados.
Parâmetro P ₅ : <i>Data Lineage</i> (Origem dos Dados)	
<ul style="list-style-type: none"> • <i>App Browser</i>. 	<ul style="list-style-type: none"> • O <i>InfoSphere MDM Server</i> inclui, na sua versão 10, melhorias na interface do utilizador da funcionalidade <i>Data Stewardship</i> que permitem aos utilizadores ver a “<i>lineage</i>” de um dado registo de <i>master data</i>; • O <i>display</i> desta informação mostra todo o histórico de operações e a origem, transformação e destino dos dados.
Parâmetro P ₆ : Segurança e Auditoria	
<ul style="list-style-type: none"> • Controlo de acessos; • Autenticação; • Gestão de segurança; • Auditoria. 	<ul style="list-style-type: none"> • A aplicação endereça múltiplos aspetos de segurança relevantes para uma implementação de MDM; • As atividades incluem o <i>logging</i>, autenticação, controle de acessos, gestão de privacidade, bem como o suporte a atividades de auditoria.
Parâmetro P ₇ : Processos de <i>Data Governance</i> (Governança dos Dados)	
<ul style="list-style-type: none"> • Políticas; • Processos; • Controles. 	<ul style="list-style-type: none"> • Os serviços <i>MDM Data Governance</i> permitem a autorização de transações e de atributos de dados; • Os administradores podem impor regras e limitar a capacidade dos utilizadores de visualizar e efetuar a atualização da <i>master data</i> ao nível dos atributos de dados e da transação, tendo como base o papel desses mesmos utilizadores na aplicação.
Parâmetro P ₈ : Domínios de <i>Master Data</i>	
<ul style="list-style-type: none"> • <i>Multi-Domain</i>; • <i>Custom-Domain</i>. 	<ul style="list-style-type: none"> • A plataforma do <i>IBM InfoSphere</i> suporta todos os domínios – cliente, cidadão, paciente, fornecedor, capital humano, <i>provider</i>, <i>member</i>, produto, serviço, <i>account</i>, financeiro, localização e domínios customizáveis; • Suporta todas as indústrias – banca, seguros, retalho, telecomunicações, governo, saúde, entre outras...

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Parâmetro P ₉ : Estilo de Implementação	
<ul style="list-style-type: none"> • Consolidação; • Registry; • Coexistência; • Hub transacional; • Estilo híbrido. 	<ul style="list-style-type: none"> • O <i>software</i> avaliado suporta todos os estilos de implementação.
Parâmetro P ₁₀ : Conectividade com as Fontes de Dados	
<ul style="list-style-type: none"> • JMS; • Web Service; • FTP; • CDC; • Pacotes de conectores. 	<ul style="list-style-type: none"> • A conectividade é concretizada através de <i>interfaces</i> e pontos de integração que compõem pacotes de adaptadores na aplicação; • As interfaces existentes para funções e componentes permitem que os clientes se conectem a ferramentas de outros fabricantes; • A documentação consultada refere a existência de pacotes de conectores para diferentes bases de dados, não especificando exatamente quais.
Parâmetro P ₁₁ : Sincronização de Master Data	
<ul style="list-style-type: none"> • Batch; • Inline. 	<ul style="list-style-type: none"> • Executa tarefas de importação e exportação através da utilização de deltas e de <i>versioning</i>; • Admite ambas as abordagens de sincronização de <i>master data</i>.
Parâmetro P ₁₂ : Modelo de Master Data	
<ul style="list-style-type: none"> • Data model editing; • Modelização gráfica; • Suporte a standards de negócio; • Controlo de versões. 	<ul style="list-style-type: none"> • O <i>InfoSphere MDM Server</i> vem com um modelo de dados pré-definido que tem incorporadas as melhores práticas de muitas implementações de MDM; • Executa o mapeamento dos modelos dos sistemas fonte para o modelo do <i>InfoSphere MDM Server</i>; • O mapeamento suporta as atividades subsequentes (desenvolvimento de ETL e a configuração de extensões no modelo da aplicação, se necessário).
Parâmetro P ₁₃ : Data Integration (Integração de Dados)	
<ul style="list-style-type: none"> • Integração física; • Integração contínua. 	<ul style="list-style-type: none"> • A ferramenta possui motores e ferramentas para a integração dos dados de diferentes fontes para um repositório comum; • O <i>InfoSphere MDM Server</i> integra dados de diferentes domínios utilizando serviços de negócio que interagem com todas as aplicações e processos de negócio que lidam com <i>master data</i>; • A aplicação está desenhada para correr num ambiente heterogéneo, mas também está pré-integrada com produtos IBM.

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Parâmetro P ₁₄ : <i>Data Federation</i> (Federação de Dados)	
<ul style="list-style-type: none"> • ETL; • EII; • EAI. 	<ul style="list-style-type: none"> • A capacidade de <i>data federation</i> do MDM Server permite que as organizações possam executar a federação de múltiplas instâncias da aplicação, tendo em vista um entendimento único da <i>master data</i>.
Parâmetro P ₁₅ : <i>Social MDM</i>	
<ul style="list-style-type: none"> • Não possui. 	<ul style="list-style-type: none"> • Não foram encontradas referências nas fontes disponíveis sobre a possibilidade de integração de <i>social data</i> no repositório de <i>master data</i>.
Parâmetro P ₁₆ : <i>Mobile MDM</i>	
<ul style="list-style-type: none"> • Não possui. 	<ul style="list-style-type: none"> • Não foram encontradas referências nas fontes disponíveis sobre a existência de uma versão do <i>software</i> para funcionamento em plataformas de dispositivos móveis.
Parâmetro P ₁₇ : <i>Processos e Workflow</i>	
<ul style="list-style-type: none"> • Agregação de atividades; • Modelização gráfica; • Gestão de <i>workflows</i>. 	<ul style="list-style-type: none"> • A plataforma permite fluxos de trabalho colaborativos, sendo assim possível à organização definir fluxos de trabalho que reflitam processos de negócio existentes e regras de negócio.
Parâmetro P ₁₈ : <i>Gestão de Metadata</i>	
<ul style="list-style-type: none"> • HTML; • XML; • XSD (<i>XML Schema Definition</i>). 	<ul style="list-style-type: none"> • O <i>InfoSphere MDM</i> possui ferramentas para a gestão eficaz da <i>metadata</i>, permitindo a importação e exportação dos ficheiros de <i>metadata</i> e mapeamento da informação contextual com os dados do repositório.
Parâmetro P ₁₉ : <i>Reporting</i> (Gestão de Relatórios)	
<ul style="list-style-type: none"> • Relatórios de <i>Data Quality</i>; • Estatísticas de utilização; • Monitorização de tarefas; • Suporte a Auditorias. 	<ul style="list-style-type: none"> • O <i>InfoSphere MDM Server</i> permite que alguma análise e <i>reporting</i> sejam efetuados em relação aos dados armazenados no sistema; • O <i>reporting</i> pode ser efetuado com o <i>Cognos® reports</i>. Uma versão desta ferramenta está disponibilizada na distribuição da ferramenta em avaliação; • O <i>reporting</i> fornecido abrange as áreas de <i>data quality</i>, utilização, monitorização de processo e tarefas e providencia suporte a auditorias.

FRAMEWORK PARA A AVALIAÇÃO DE FERRAMENTAS MDM

Parâmetro P ₂₀ : <i>Usability</i> (Usabilidade)	
N/A	<ul style="list-style-type: none"> • Não foi possível encontrar informação disponibilizada pelo fabricante que permita a avaliação e caracterização deste parâmetro; • O centro de suporte do <i>software</i> disponibiliza uma série de parâmetros de usabilidade que controlam as capacidades e a aparência de vários módulos e consolas de interface com o utilizador.
Parâmetro P ₂₁ : <i>Exception Handling</i> (Tratamento de Exceções)	
<ul style="list-style-type: none"> • <i>Flagging</i>; • Correção automática; • Correção manual. 	<ul style="list-style-type: none"> • A <i>framework</i> para a manipulação de exceções foi melhorada relativamente às <i>releases</i> anteriores no sentido de expor a causa raiz dos erros. • O <i>exception handling</i> é suportado por uma série de <i>scripts</i> para controlo e resposta de erros e exceções detetadas pelo sistema nos mais diversos tipos de transações.
Parâmetro P ₂₂ : Requisitos Tecnológicos do Sistema (S.O. / SGBD)	
<p>Sistema Operativo:</p> <ul style="list-style-type: none"> • <i>AIX 6.1 TL07</i> • <i>AIX 7.1 SP03</i> • <i>Solaris 10 SPARC</i> • <i>HP-UX 11i v3 IA64</i> • <i>Red Hat Enterprise Linux (RHEL) 5.0</i> • <i>Red Hat Enterprise Linux 6</i> • <i>Red Hat Enterprise Linux (RHEL) 6.0</i> • <i>SUSE Linux Enterprise Server (SLES) 11.0</i> <p>SGBD:</p> <ul style="list-style-type: none"> • <i>DB2 Enterprise Server Edition v9.7</i> • <i>DB2 for z/OS v10.1</i> • <i>DB2 for z/OS v9.1</i> • <i>Oracle Database 11g Enterprise Edition Release 2</i> 	<ul style="list-style-type: none"> • Com o <i>IBM InfoSphere Master Data Management V10</i>, a IBM unificou 3 ofertas MDM de mercado: o <i>InfoSphere MDM Server</i>, o <i>Initiate Master Data Service</i> e o <i>InfoSphere MDM Server for PIM</i>. Estas 3 plataformas fazem parte das edições seguintes: <ul style="list-style-type: none"> • <i>IBM InfoSphere MDM Collaborative Edition</i>; • <i>IBM InfoSphere MDM Standard Edition</i>; • <i>IBM InfoSphere MDM Advanced Edition</i>; • <i>IBM InfoSphere MDM Enterprise Edition</i>. <p>IBM InfoSphere Master Data Management Server Incluído no <i>IBM InfoSphere Master Data Management - Enterprise Edition</i> e <i>Advanced Edition</i>.</p> <p>IBM Initiate Master Data Service Incluído no <i>IBM InfoSphere Master Data Management - Enterprise Edition</i>; <i>Advanced Edition</i> e <i>Standard Edition</i>.</p> <p>IBM InfoSphere Master Data Management Collaboration Server Incluído no <i>IBM InfoSphere Master Data Management - Enterprise Edition</i> e <i>Collaborative Edition</i>.</p>