



**Integridade de Dados(Data Integrity)-
Desenvolvimento de Processos
Robóticos Automatizados (RPA) sob a
solução de reconciliação de dados com o
objetivo de garantir a integridade dos
dados**

António Gomes

Mestrado Integrado em Engenharia Eletrotécnica e de Computadores

Orientador na FEUP: Pedro João

Orientador na Empresa: Sandrina Azevedo

20 de Janeiro de 2019

Resumo

Os dados e a sua integridade e consistência estão intrinsecamente ligados às organizações e ao modo como estas operam. A sua integridade é um recurso valioso para o sucesso de qualquer organização, uma vez que são tomadas decisões com base na análise dos mesmos. Na indústria do retalho tem-se vindo a verificar um aumento exponencial do volume de dados, sendo portanto de maior importância, para uma boa gestão dos negócios, manter a sua integridade.

Todos os grandes retalhistas usam o *software Oracle Retail* para a gestão dos seus negócios, existindo dentro deste várias alternativas e diferentes soluções. Para a implementação e personalização do *software*, esses retalhistas, por norma, contratam empresas externas em regime de *outsourcing*. Normalmente, além da implementação e personalização, as empresas contratadas prestam também um serviço continuado de suporte e manutenção.

Esta dissertação foi desenvolvida na *Wipro*, uma empresa dedicada à implementação do *software Oracle Retail* em grandes retalhistas e que, por norma, integra também uma equipa de suporte e manutenção nesses clientes, após a implementação. Estes serviços de suporte têm como objetivo manter a integridade e consistência dos dados utilizados pelo cliente, conseguindo estes serviços detetar e corrigir falhas e problemas que possam surgir.

O principal objetivo desta dissertação é automatizar alguns serviços de suporte que a empresa dispõe, através do desenvolvimento de um RPA que aumente a eficácia e reduza o tempo de execução, os recursos necessários e consequentemente os custos do processo.

Para tal foi feita uma análise dos processos de suporte com a finalidade de entender o seu funcionamento e identificar as oportunidades de melhoria, o que definiu os objetivos da solução a desenvolver.

Este documento descreve toda a metodologia de desenvolvimento adotada, bem como uma análise da situação e uma descrição da solução proposta que cumpra os objetivos definidos. Inclui também a descrição da prova de conceito realizada, bem como os benefícios da mesma. A abordagem proposta tem como principal objetivo automatizar as tarefas repetitivas de modo a que as equipas realizem apenas tarefas que acrescentam valor, aumentando assim a eficácia das equipas.

Abstract

The data integrity and consistency are intrinsically linked to the organizations and how they operate. Data integrity is a valuable resource for the success of any organization since decisions are made based on their analyzes. An exponential increase of data has been verified in the retail industry, and it is therefore of greater importance for a good business management to maintain its integrity.

Oracle Retail software is used by all major retailers to manage their business, having within it several alternatives and different solutions. For the implementation and customization of these software, normally those retailers hire external companies in an outsourcing mode. Usually in addition to the implementation and customization, the contracted companies also provide a continuous service of support and maintenance.

This dissertation was developed at Wipro, a company dedicated to the implementation of the Oracle Retail software in large retailers and by standard also integrates a support and maintenance team on those clients after the implementation. These support services aim to maintain the integrity and consistency of the data used by the customer, once these services can detect and correct any failures and problems that may appear.

The main objective of this dissertation is to automatize some of the company support services, through the development of an RPA that increases efficiency and reduces the execution time, resources required and consequently the process cost.

To this end and with the purpose of understanding its operation and identifying the opportunities of improvement, an analysis of the support processes was made. That analysis defined the objectives of the solution to be developed.

This document describes all the development methodology adopted, as well as an analysis of the situation and a description of the proposed solution that meets the defined objectives. It also includes the description of the proof of concept that was made as well as the benefits of it. The main objective of the proposed approach is to automate repetitive tasks so that the teams only perform tasks that add value, and by that increase the effectiveness of teams.

Agradecimentos

Durante todo o curso houve pessoas que me deram muito apoio e a elas estarei eternamente grato.

Agradeço ao meu orientador, o professor Pedro Alexandre Rodrigues João, a sua disponibilidade, acompanhamento e ajuda. A confiança e calma que me transmitiu foram essenciais para a realização desta dissertação.

Agradeço também à professora Maria Antónia Carravilla e ao professor José Fernando Oliveira por toda a preocupação em disponibilizar aos alunos do MIEEC da área de Gestão Industrial uma vasta lista de temas na escolha da dissertação.

À minha orientadora na *Wipro*, Doutora Sandrina Azevedo, um muito obrigado pela boa recepção e pelo acompanhamento e disponibilidade durante a dissertação. Agradeço também à equipa de desenvolvimento do DIaaS, André Couto e Daniel Costa, por toda a ajuda prestada. Gostaria também de agradecer à *Wipro* pela possibilidade de realizar a minha dissertação num ambiente de trabalho espetacular, com muito companheirismo e onde me fizeram sempre sentir bem vindo.

Quero também deixar um especial agradecimento à minha namorada, Joana Lourenço Ribeiro, por me tornar uma pessoa melhor e mais feliz, por ter sido o meu apoio e a minha força. Por me motivar e apoiar incondicionalmente. Quando precisei esteve sempre lá, mostrou-me que eu era capaz e reergueu-me nos momentos mais difíceis. À Joana um muito obrigado pois sem ela eu não teria conseguido.

À minha mãe, ao meu pai, ao meu irmão, à minha irmã, à minha tia e à minha avó um agradecimento muito especial pelo suporte que me deram durante este longo percurso. Com um especial agradecimento aos meus pais, por estarem sempre ao meu lado e nunca me deixarem desistir, vocês são os pilares da minha vida, sem vocês nada seria possível.

António Gomes

“The best way to predict the future is to create it.”

Peter Drucker

Conteúdo

1	Introdução	1
1.1	Apresentação da Empresa	1
1.1.1	O Espírito da <i>Wipro</i>	1
1.1.2	<i>Wipro</i> Portugal	2
1.2	Contextualização	3
1.3	Descrição e Objetivos do Projeto	4
1.4	Metodologia	5
1.5	Estrutura da Dissertação	5
2	Enquadramento Teórico	7
2.1	O Retalho	7
2.1.1	Cadeia de Logística no Retalho	7
2.1.2	A Importância dos Dados no Retalho	8
2.2	Tecnologias de Informação	9
2.2.1	Sistemas de Informação no Retalho	10
2.2.2	Características e Objetivos dos Sistemas de Informação no Retalho	11
2.3	<i>Software Oracle Retail (OR)</i>	12
2.4	Novas Tecnologias	13
2.4.1	Processos Robóticos Automatizados (RPA)	13
2.4.2	<i>Software as a Service</i>	14
2.4.3	<i>Machine Learning</i>	14
2.5	Solução DIIaaS	15
2.6	Integridade de Dados(<i>Data Integrity</i>)	16
2.6.1	A Importância da Integridade dos Dados	17
2.6.2	Manter a Integridade dos Dados	18
3	Situação Atual - Análise	21
3.1	O Suporte Prestado Atualmente pela <i>Wipro</i>	21
3.1.1	Processos de Suporte	21
3.2	Modelo Atual	22
3.3	Dimensão do Problema	25
3.4	Oportunidades de Melhoria	26
4	Solução Proposta - Descrição	27
4.1	Conceitos Necessários	27
4.2	Abordagem ao Problema	28
4.3	Metodologia Utilizada	30
4.4	Descrição da Solução	30

4.4.1	Requisitos da Solução	31
4.4.2	Limitações	31
4.5	Explicação do Funcionamento da Solução	31
4.5.1	Algoritmo	32
4.5.2	Alterações ao DIaaS	32
5	Prova de Conceito e Resultados Obtidos	35
5.1	Permissões e Campo de Ação	35
5.2	Tecnologias Utilizadas	35
5.2.1	Linguagens de Programação	36
5.2.2	Ambientes Utilizados	36
5.3	Descrição Detalhada do Funcionamento do Algoritmo	37
5.3.1	Tabelas	37
5.3.2	Identificação do Erro	38
5.3.3	RCA's	38
5.3.4	Soluções	39
5.4	Resultados Obtidos	39
5.4.1	Dados de Teste	39
5.4.2	Resultados Obtidos	40
5.4.3	Tempo de Execução	42
5.4.4	Comparação dos Resultados Obtidos com a Situação Atual	42
6	Conclusões e Trabalho Futuro	45
6.1	Conclusões	45
6.2	Trabalho Futuro	46
6.2.1	Implementação em Ambiente Real	46
6.2.2	Correção automática	46
6.2.3	Inteligência Artificial	46
A	Modelo Relacional Antigo	47
B	Modelo Relacional Novo	49
C	Fluxograma do Algoritmo	51
	Referências	53

Lista de Figuras

1.1	Logótipo da Wipro	2
1.2	Logótipo da Enabler	2
2.1	Cadeia de Logística do Retalho	8
2.2	Modulos que compõe o <i>Oracle Retail</i>	13
2.3	Benefícios dos dados com alta qualidade	15
3.1	Período de atividade dos processos num entreposto durante o dia	23
3.2	Processo de análise do suporte	23
4.1	Processo atual	28
4.2	Processo futuro	29
4.3	Metodologia utilizada	30
4.4	Tabelas de apresentação dos erros	33
5.1	Linguagens de programação utilizadas	36
5.2	Tabela de Identificação dos erros	37
5.3	Tabela com as RCA e reconciliações	37
5.4	Tabela das ações a aplicar	38
5.5	Tabela informativa das RCA efectuadas	39
5.6	Processo de identificação do erro	40
5.7	Processo de verificação se dentro das RCA tem extração ou reconciliação	41
5.8	Processo de execução das RCA	42
5.9	Resultado da reconciliação a nível dos entrepostos	43
5.10	Resultado da reconciliação a nível das ordens	44
5.11	Resultado da reconciliação a nível do artigo	44
5.12	Tabela comparison_error relativa as reconciliações acima apresentadas	44
5.13	Resultado de alguns erros para os quais se realizaram RCA	44

Lista de Tabelas

2.1	Tabela de Funcionalidades do Argws e do DIaaS	16
3.1	Número de artigos e tempo despendido na sua análise	25

Abreviaturas e Símbolos

RPA	Processos Robóticos Automatizados
TI	Tecnologia da Informação
SI	Sistema de Informação
OR	<i>Oracle Retail</i>
SaaS	<i>Software as a Service</i>
ML	<i>Machine Learning</i>
DIaaS	<i>Data Integrity as a Service</i>
SaaS	<i>Software as a Service</i>
RCA	<i>Root Cause Analysis</i>
IDE	<i>Integrated Development Environment</i>
pks	<i>Primary Keys</i>
RMS	<i>Retail Merchandising System</i>
RIM	<i>Retail Invoice Matching</i>
WMS	<i>Warehouse Management System</i>

Capítulo 1

Introdução

A presente dissertação, desenvolvida no âmbito do Mestrado Integrado em Engenharia Eletrotécnica e de Computadores, da Faculdade de Engenharia do Porto (FEUP), foi realizada na empresa *Wipro*. Neste primeiro capítulo, procederemos então a uma breve apresentação da referida organização e descreveremos, a título introdutório, o projeto desenvolvido, assim como a metodologia adotada na sua prossecução.

1.1 Apresentação da Empresa

Fundada em 1945, a *Wipro Limited* (Western India Products Limited) é uma multinacional indiana, especializada em serviços de tecnologia da informação, consultoria e processos de negócios, integrando computação cognitiva, automação, robótica, nuvem, análise e tecnologias emergentes. O eixo central da sua intervenção objetiva-se na possibilidade de fornecer aos clientes a capacidade de se adaptarem ao mundo digital, promovendo o sucesso dos seus negócios. Sediada em Bangalore, na Índia, a *Wipro* é uma empresa reconhecida pelo seu portefólio substantivo, nomeadamente no que concerne ao compromisso que estabelece com a sustentabilidade e a cidadania corporativa. O domínio da *Wipro* tem um amplo alcance em todo o mundo, operando em 62 países e integrando mais de 160.000 trabalhadores. Saliente-se que a empresa se encontra ainda em constante crescimento. Ao longo dos anos, a *Wipro* tem sido contemplada com vários prémios, vendo assim reconhecida toda a sua ação e assinaladas as suas notáveis conquistas. Neste âmbito é de registar o prémio que a consagrou como uma das empresas mais éticas do mundo, atribuído pelo Instituto Ethisphere em 2018 pela 7^a vez consecutiva [1]. Entre muitos outros, angariou igualmente o segundo lugar assinalado na Newsweek 2012 Global 500 Green Companies. Em 2017, foi apresentado o seu novo logótipo, representado na Figura 1.1 patente na página 2 deste documento.

1.1.1 O Espírito da *Wipro*

O Espírito da *Wipro* é a essência da *Wipro*. Os seus valores estão claramente definidos e assumidos e refletem-se no comportamento de todos. Quem são aqueles que integram a empresa



Figura 1.1: Logótipo da Wipro

assim como os princípios que a organização subscreve são questões centrais no quadro da *Wipro*, enraizadas na sua essência inalterável mas constituindo, também, os ingredientes essenciais para o que a empresa deve aspirar a ser no futuro. O Espírito da *Wipro* é, então, um farol, guiando e proporcionando um propósito claro para todos, unindo as pessoas para um objetivo comum. O Espírito da *Wipro* é estabelecido em quatro valores[2] que estão presentes desde a origem:

- Ter como foco principal o sucesso dos clientes;
- Ser global e responsável;
- Tratar cada pessoa com respeito;
- Ter sempre a integridade presente em tudo o que é feito.

1.1.2 *Wipro* Portugal

A *Wipro* Portugal tem a sua origem na compra da empresa *Enabler*. A *Enabler* foi criada em 1997 como resultado do *spinoff* de um departamento do Grupo *Sonae*, o de Sistemas de Informação da *Modelo Continente Hipermercados* [3].

Em 2006, devido ao seu vasto conhecimento e domínio na área do retalho e também por serem os líderes na implementação do pacote *Oracle Retail Enterprise Resource Planning(ERP)*, a *Wipro* adquiriu a *Enabler*, cujo logótipo está representado na Figura 1.2.

Em 2008, é inaugurado o escritório da Maia onde a *Wipro* Portugal está instalada desde então. Atualmente, trabalham na Maia cerca de 180 pessoas, sendo uma referência global na implementação de soluções *Oracle Retail(OR)*[4].



Figura 1.2: Logótipo da Enabler

Estas soluções desenvolvidas a partir de ferramentas e *software* da *Oracle* vêm em módulos que podem ser selecionados e personalizados conforme as necessidades do cliente. Esses módulos, a título de exemplo, permitem a gestão de inventário, entrepostos, preçários, promoções, entre outros aspetos. A implementação deste *software* tem como objetivos a redução dos custos e o aumento da satisfação dos clientes, permitindo, acima de tudo, a gestão e controlo mais eficaz de todo o modelo de negócio.

A Wipro Portugal fornece os seguintes serviços:

- **Consultoria;**
- **Implementação de produtos *Oracle Retail*;**
- **Manutenção e Suporte de aplicações;**
- **Transformação de Negócios.**

1.2 Contextualização

Ao longo dos últimos anos e principalmente na área do retalho, têm sido desenvolvidas novas ferramentas e soluções que ajudam as empresas a conduzir os seus negócios de forma controlada e eficiente. Como nesta área se trabalha cada vez mais com um elevado número de dados, estes precisam de ser coerentes, sendo necessário aperfeiçoar as ferramentas que servem esse propósito. Esta prática requer uma grande coordenação e validação dos dados utilizados. Podemos enumerar algumas das possibilidades de utilização dos dados: gestão dos níveis de inventário entre os entrepostos e as lojas; avaliação da coerência dos preços; descrição dos produtos nas diferentes lojas e nos entrepostos; verificação das faturas, avaliando a respetiva correspondência entre os valores apresentados a nível global (total de lojas) e a nível local. Neste contexto, existe uma elevada quantidade de dados que precisam de ser transferidos e atualizados, sendo necessário garantir a sua integridade.

Ao implementar nos clientes o *software Oracle Retail*, para que este cumpra o seu propósito, é necessário uma sincronização dos dados entre os diferentes módulos utilizados, de forma a que a sua integridade não seja posta em causa. É muito comum, após a implementação deste *software*, o cliente assinar um contrato de suporte e manutenção de forma a garantir o bom funcionamento do sistema. Ao assinar este contrato, o cliente irá ter uma equipa de suporte a trabalhar para a sua empresa e que fica responsável pela identificação e correção de problemas e erros que possam emergir. Para conseguir que a solução implementada cumpra o seu objetivo com rigor é necessário que a equipa responsável por estes procedimentos tenha um vasto conhecimento da arquitetura e das particularidades da solução implementada. Assim, nestas circunstâncias, é habitual a aquisição deste serviço complementar, dado que, por norma, os clientes não integram essas equipas especializadas.

Com o crescimento do mercado nesta área, há uma maior procura destes serviços e são necessárias mais equipas para cada projeto, surgindo a necessidade de reduzir essas equipas de modo a

baixar o preço do serviço de suporte prestado. Para isso é necessário reduzir ou simplificar as tarefas da equipa de suporte de modo a diminuir o trabalho realizado pela mesma, automatizando ou simplificando as tarefas a cumprir pelos membros das respetivas equipas. Registe-se que muitos destes procedimentos são repetitivos.

1.3 Descrição e Objetivos do Projeto

Para o suporte e manutenção a realizar após a implementação do software anteriormente mencionado, a *Wipro* desenvolveu uma ferramenta designada “Argws” que é capaz de monitorizar e alertar o utilizador da existência de inconsistências no sistema. Recentemente, e devido à ferramenta se encontrar desatualizada, foi desenvolvida outra aplicação, sob a denominação de “DIIaaS”, com tecnologias mais recentes e que para além de criar reconciliações (comparação de dados entre sistemas) permite agendá-las e visualizar o seu resultado através de *dashboards*.

Este processo de suporte, prestado após implementação do software, consiste em realizar reconciliações através de *queries* (comando para a consulta de dados presentes numa base de dados) entre vários ambientes e por vezes dentro do mesmo que vão retornar conforme existam ou não inconsistências nos dados. Estas reconciliações são usadas para verificar, por exemplo, se o stock total corresponde à soma do stock nas várias lojas, se o que foi solicitado numa encomenda chegou realmente, se a quantidade disponível de um artigo é a mesma em dois sistemas diferentes, de modo a ser possível garantir a integridade dos dados entre os vários sistemas.

Estas reconciliações retornam os erros encontrados e o utilizador tem de descobrir manualmente a origem dos mesmos, identificar o sistema que integra os dados válidos e o sistema onde constam os dados inválidos. Este processo passa por analisar individualmente as *queries* para perceber, numa primeira fase, quais as tabelas envolvidas e, numa segunda fase, analisá-las e perceber onde está o erro de forma a proceder à sua correção. Este processo é muitas vezes repetitivo e moroso o que o torna mais dispendioso dado que exige muitas horas de trabalho de suporte.

Assim, o objetivo deste projeto é desenvolver *RPA* a ser implementados no “DIIaaS” que permitam identificar e, tanto quanto possível, resolver os erros em causa. Na eventualidade de não ser possível a resolução dos mesmos, garantir, no mínimo, que se sugira ao utilizador os procedimentos a adotar evitando o dispêndio excessivo de tempo neste processo. Pretende-se, igualmente, identificar os erros que possam ter a mesma solução de modo a definir tipologias. Os objetivos passam então por:

- **Definir e caracterizar erros** - perceber os tipos de erros que podem acontecer e arranjar parâmetros que os caracterizem.
- **Verificar se é possível automatizar a solução existente** - perceber se a solução existente ou pelo menos parte dela, pode ser ou não automatizada.
- **Verificar se é possível estandardizar soluções**- tendo as soluções existentes perceber se erros diferentes podem ter a mesma solução.

Na essência, trata-se de simplificar o trabalho de suporte de modo a que se mantenha a Integridade dos Dados e se garanta a realização destas tarefas em menos tempo. Ou seja, permitir que a equipa de suporte use o seu tempo de forma mais eficiente, centrando a sua ação na análise de novas anomalias e respetivas soluções a integrar na aplicação, evitando-se o dispêndio de tempo na realização de procedimentos repetitivos.

1.4 Metodologia

Para a prossecução deste projeto foi inicialmente realizado um estudo da empresa e do seu modelo de negócio para uma melhor compreensão da pertinência da ferramenta a ser utilizada. Foi também elaborado um levantamento das especificações e requisitos da solução a ser desenvolvida.

Em seguida, foi feito um estudo e aperfeiçoamento da nova ferramenta desenvolvida pela *Wipro*, de forma a garantir que esta cumpria todos os requisitos necessários no quadro dos objetivos estabelecidos. Neste passo, foram acrescentados à aplicação alguns parâmetros e funcionalidades para que a solução a aplicar pudesse ser implementada.

Preparada a aplicação, foi elaborado o estudo de casos reais e desenvolvida uma solução para os problemas de forma a garantir que a mesma pudesse ser flexível e automatizada. Ou seja, garantir que se acrescentam soluções sempre que sejam encontrados novos erros que seriam invocadas por um algoritmo de modo a que a resposta seja automática. Desta forma, os utilizadores são libertados da correção sistemática dos mesmos erros.

Após ter tudo preparado, foi iniciada a prova de conceito onde foi desenvolvido um algoritmo que percorre uma tabela e verifica se pode aplicar solução ao problema. Também foi desenvolvida essa tabela já com algumas soluções e com a facilidade de serem acrescentadas novas soluções sempre que se encontrar um erro diferente.

Todos os erros e soluções apresentados foram obtidos em reuniões com funcionários relacionados com a parte de suporte da empresa de modo a perceber as tipologias existentes, a forma como lidavam com os erros referidos, garantindo-se que a solução desenvolvida pudesse cumprir todos os requisitos necessários para automatizar a correção.

1.5 Estrutura da Dissertação

Esta dissertação integra cinco capítulos, para além desta introdução. No capítulo 2, é descrito o enquadramento teórico que suporta o projeto, sustentado na pesquisa bibliográfica realizada e integrando a definição de todos os conceitos e tecnologias necessárias à compreensão do discurso patente nesta dissertação.

No capítulo 3, é realizada uma análise da situação atual com explicitação da ferramenta a utilizar e das tecnologias que a mesma mobiliza para permitir a perceção do problema e a identificação das oportunidades de melhoria.

No capítulo 4 é detalhada a solução proposta, enunciando-se as funcionalidades a aplicar na ferramenta de modo a cumprir os requisitos necessários e a tornar inteligível a solução proposta.

No capítulo 5 é apresentada a prova de conceito e os resultados obtidos assim como uma análise aos mesmos.

No capítulo 6 são apresentados os requisitos cumpridos, as dificuldades encontradas ao longo do projeto e o trabalho futuro que poderá ser desenvolvido.

Capítulo 2

Enquadramento Teórico

No mundo contemporâneo, enfrentamos um universo marcado pela grande quantidade de dados, nomeadamente no retalho em que os dados são cada vez mais reconhecidos como estratégia para aceder ao cliente e assim conseguir prever melhor as tendências de consumo[5]. Todos estes dados podem ser gerados por pessoas ou por máquinas e é previsível que o aumento continue a ser significativo nos próximos anos.[6] Estes dados são fulcrais no retalho não só para a previsão de vendas, mas também na tomada de decisões. Torna-se fundamental, por exemplo, determinar as encomendas a fazer, o que vai depender, nomeadamente, do nível de stock. É então crucial haver coerência e confiança nos dados que estão a ser utilizados, garantindo a sua integridade nas várias interfaces dos retalhistas.

Neste capítulo é feita uma breve introdução ao retalho, área de negócio sobre a qual a *Wipro* atua, de modo a permitir compreender as características e necessidades do modelo de negócio. Serão também abordadas as tecnologias utilizadas assim como os conceitos fundamentais inerentes à manutenção da Integridade dos Dados.

2.1 O Retalho

Retalho, por definição, é a venda de produtos ou serviços, em pequenas quantidades, ao consumidor final, não para revender mas para uso e consumo próprio[7].

2.1.1 Cadeia de Logística no Retalho

Uma cadeia de logística é um processo que ocorre entre empresas e fornecedores para a distribuição de produtos aos usuários finais, ou seja, é como um bem ou serviço que é entregue aos consumidores.

Os retalhistas devem entender a sua cadeia de logística de modo a garantir que recebem os produtos certos a um preço acessível e dentro de um prazo razoável [8]. Se ocorrer algum problema na cadeia de logística é provável que se venha a observar um aumento do custo do produto ou do tempo de entrega.

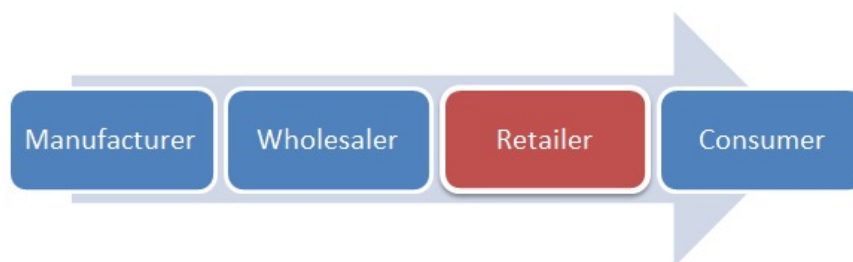


Figura 2.1: Cadeia de Logística do Retalho

Na figura 2.1, que se encontra na página 8, está representada a cadeia de logística no retalho com os principais participantes na mesma. Os papéis destes participantes são [9]:

- **Fabricante(Manufacturer)** - os fabricantes são aqueles que estão envolvidos no processo de fabrico de mercadorias com ajuda de máquinas e mão de obra, garantindo assim a transformação das matérias primas.
- **Grossista(Wholesaler)** - o grossista é aquele que compra as mercadorias aos fabricantes e vende aos retalhistas em grandes quantidades mas a um preço menor. Um grossista nunca vende diretamente ao consumidor final.
- **Retalhista(Retailer)** - um retalhista está no final da cadeia de fornecimento, compra os produtos ao fabricante ou ao grossista e vende ao consumidor final em pequenas quantidades.
- **Consumidor(Consumer)** - o consumidor final vai ao retalhista para comprar as mercadorias em pequenas quantidades para satisfazer as suas necessidades, ou seja, para uso pessoal.

No retalho, há uma grande quantidade de dados e para serem tomadas decisões sustentadas nessa informação é necessário garantir a Integridade dos mesmos. A recolha e tratamento de dados e posterior extração do seu valor é muito importante para a tomada de decisões por parte dos retalhistas, nomeadamente nas previsões de procura.

2.1.2 A Importância dos Dados no Retalho

Atualmente encontramos-nos na chamada era do *Big Data*, dado que o volume de dados existente no mundo duplica a cada 2 anos sendo, portanto, cada vez mais decisivo e fundamental, para o sucesso de qualquer empresa relacionada com o retalho, implementar culturas de decisão orientadas para a análise de dados. Todos os dias são disponibilizados nas lojas destes retalhistas uma imensidão de dados, sendo então necessária a utilização de ferramentas e dos *skills* necessários para recolher, tratar e estruturar esta informação. Conseguir estas ferramentas e *skills* constitui indubitavelmente uma vantagem competitiva para qualquer negócio[10]. Obter estes dados é, então, fundamental, mas estes têm de ser transformados em conhecimento passível de ser utilizado pelos empresários.

Devido à grande quantidade de dados disponíveis hoje em dia, o maior desafio é, acima de tudo, conseguir garantir a leitura daqueles que são mais relevantes para o negócio e que podem criar mais impacto na tomada de decisões estratégicas. A forma mais eficiente de o fazer é centrar a atenção no cliente, utilizando ferramentas que possam não só descrever, de forma assertiva, o que está a acontecer, mas também mobilizar conhecimento passado para tentar prever comportamentos. Assim, estas ferramentas permitem conhecer o negócio e perceber como este é perspectivado pelo cliente. Desta forma, este conhecimento deve ser usado para tomar decisões que permitam ganhos de eficiência e melhoria a longo prazo, tornando o negócio adequado às necessidades do presente e mais preparado para o futuro[11].

Os dados estão então a ser cada vez mais entendidos como uma forma de conhecer o cliente, permitindo, igualmente, fazer ajustamentos de preços, planear a logística ou até reajustar mapas de loja. O mundo do retalho está em constante mutação havendo cada vez mais opções para o cliente. Os retalhistas, para além de terem de competir com a concorrência que está cada vez mais habilitada e aguerrida, têm também um consumidor cada vez mais exigente resultante das variadas opções de compra que lhe são fornecidas.

Hoje em dia, os retalhistas já analisam as compras de cada cliente traçando uma espécie de perfil, daí recebermos, muitas vezes, de hipermercados, por exemplo, alertas ou cupões de desconto de produtos que costumamos comprar[12].

Para o futuro, é previsível uma continua progressão dos negócios no sentido de obterem informações que façam a diferença[13].

2.2 Tecnologias de Informação

A tecnologia da informação (TI) é o uso de qualquer computador, armazenamento, rede ou outros dispositivos físicos, infraestruturas e processos para criar, processar, armazenar e proteger e trocar todas as formas de dados eletrónicos[14]. Também se pode dizer que se refere a qualquer coisa relacionada com a tecnologia de computação, como rede, *hardware*, *software*, Internet ou às pessoas que trabalham com essas tecnologias. Cada vez mais as empresas têm departamentos de TI para administrar computadores, redes e outras áreas técnicas dos seus negócios. Os trabalhos de TI incluem programação, administração de redes, engenharia da computação, desenvolvimento da Web, suporte técnico e muitas outras tarefas relacionadas[15].

As informações críticas necessárias numa organização de retalho formam a base dos fatores críticos para o sucesso do respetivo negócio. A seguir estão representadas algumas das áreas de informação chave:

- **Informação do produto** - catálogo, disponibilidade, promoções...
- **Informação do cliente** - perfil, comportamento, atividades, preferências, distribuição...
- **Informação de operações** - alocação, logística, procura, inventário...

A necessidade de TI no retalho deve-se a uma variedade de fatores dos quais quatro são mais importantes do que nunca, sendo eles[16]:

- **Competição Intensa** - a sua adoção força a organização a ser mais eficiente e efetiva. Uma forma de obter vantagem competitiva é usar as TI, no mínimo, de quatro formas: reduzir o custo das operações através da automação; melhoria do produto ou a qualidade do serviço fornecendo garantia de qualidade; fornecimento de serviços de valor agregado que criem diferenciação; utilização da TI para inteligência competitiva.
- **Globalização das operações de negócio** - em todos os negócios, sejam eles de baixa ou alta tecnologia, de serviço ou produto, há sempre algum grau de competição global. As operações globais requerem que a barreira da distancia e do tempo sejam eliminadas ou reduzidas pelo que é essencial maximizar o uso das TI.
- **Mudanças organizacionais** - com o aumento de fusões e aquisições, a administração das organizações reconhece a necessidade de flexibilidade através de TI compatíveis.
- **Revolução tecnológica** - num curto período de tempo ocorreu a comercialização das TI, a relação entre o preço e o desempenho das TI tornou-as mais acessíveis e úteis para as organizações.

A principal motivação da utilização das TI vem da necessidade de uma melhor utilização dos recursos financeiros investidos. A sua adoção, entre outros benefícios, permite aumentar a produtividade, tornar mais fácil e rápida a comunicação entre diversas entidades e aumentar a eficácia das previsões de procura[17].

2.2.1 Sistemas de Informação no Retalho

Um sistema de Informação (SI) pode ser definido como um sistema composto por pessoas, procedimentos e equipamentos cujo objetivo central é recolher, armazenar e distribuir informação para efeitos específicos. O SI é composto por *inputs* (dados, instruções) que são processados e *outputs* (relatórios) que são produzidos por ele e disponibilizados ao utilizador e/ou a outros sistemas[18].

Nos dias de hoje, os grandes retalhistas recolhem uma grande quantidade de informação que vai ser organizada e armazenada em SI complexos obtendo com isso vantagens na sua empresa como a redução de custos, a criação de produtos e serviços únicos para o cliente e uma avaliação continua do seu desempenho através de indicadores.

Segundo um estudo, os requisitos dos SI de cadeias de lojas incluem pelo menos três aspetos[19]:

- **Serviço de Informação para a Gestão** - A troca de informação entre o retalhista e o consumidor que inclui informação sobre o negócio e sobre as reclamações e procura do cliente. A comunicação entre lojas e o serviço central que inclui informações de gestão dos recursos humanos, do plano de marketing, dados financeiros e informações sobre reclamações. A troca de informações entre o serviço central e o centro de distribuição que inclui ordens de distribuição, informações de inventário e ordens de aceitação de encomendas. E, por fim, a

troca de informação entre os serviços centrais e os fornecedores que inclui as informações de compra e fornecimento.

- **A informação crítica para a tomada de decisões** - que se refere a informações a serem usadas para a melhoria da qualidade do serviço, onde a maioria dos dados diz respeito a clientes, como vendas, resultados das análises das reclamações, capacidade dos fornecedores, entre outras.
- **Os serviços de informação de valor acrescentado** - estes têm como objetivo descobrir novas tendências, aumentar lucros e reduzir custos através da análise de estudos do mercado, inquéritos aos clientes e muitos outros.

2.2.2 Características e Objetivos dos Sistemas de Informação no Retalho

Um SI de retalho deve ter as seguintes características:

- Deve conectar todas as lojas à empresa.
- Permitir a troca de informações instantâneas entre as lojas e o departamento de gestão.
- Lidar com os vários aspetos da gestão dos produtos.
- Tratar da análise do cliente.
- Permitir que um gerente de loja defina preços ao longo de um ano financeiro.

Os objetivos de um SI no retalho são os seguintes:

- Deve fornecer informações relevantes ao gerente do retalho regularmente.
- Antecipar as necessidades e exigências do gerente.
- Ser suficientemente flexível para incorporar as necessidades que estão em constante evolução no mercado.
- Deve ser capaz de captar, armazenar e organizar todos os dados relevantes de maneira regular e contínua.
- Deve estar alinhado com os planos estratégicos e de negócio da organização, devendo ser capaz de fornecer informações que suportem e guiem este objetivo.

Os SI no retalho devem então apoiar funções básicas de retalho como a aquisição de material, o armazenamento, o envio, entre outras. Devem permitir que o gerente monitore as vendas dos produtos, o volume de vendas diárias e deve ainda ajudar na gestão do inventário.[20].

2.3 *Software Oracle Retail (OR)*

O *Oracle* é um Sistema de Gestão de Base de Dados que surgiu em 1977 quando Larry Ellison vislumbrou uma oportunidade ao encontrar uma descrição de um protótipo funcional de uma base de dados relacional e ao verificar que nenhuma empresa se tinha empenhado em comercializar essa tecnologia. Ellison e os co-fundadores da *Oracle Corporation*, Bob Miner e Ed Oates, perceberam que havia um imenso potencial de negócios no modelo de base de dados relacional, tornando-se assim a maior empresa de *software* empresarial do mundo[21].

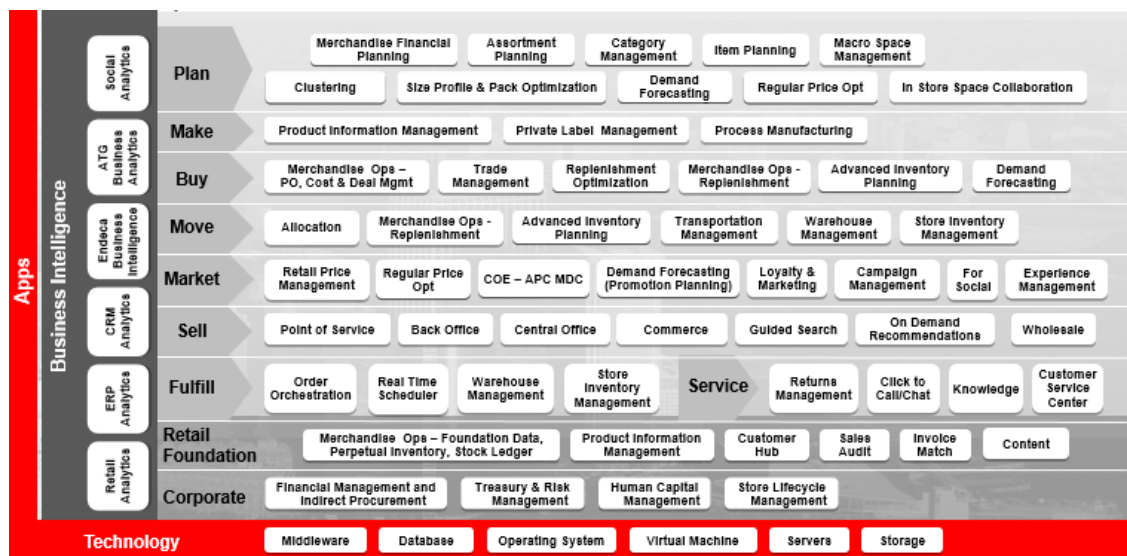
Além da base de dados, a *Oracle* criou um conjunto de ferramentas de desenvolvimento com o nome de *Oracle Developer Suite*, utilizada na construção de programas de computador que interagem com a sua base de dados. A *Oracle* criou, igualmente, a linguagem de programação PL/SQL, utilizada no processamento de transações.

As suas soluções para retalho compõem uma subdivisão do grupo denominada por *Oracle Retail*, tendo esta sido iniciada no ano de 2005 com a aquisição de uma empresa de software para retalho de nome *Retek*. O OR é a mais completa e integrada suite (conjunto de programas disponibilizados como um produto único) de aplicações para retalho do mercado. Garante ao retalhista a visibilidade das informações que ele necessita para tomar as melhores decisões. Atende tanto ao pequeno como ao grande retalho em todas as áreas, incluindo moda, produtos alimentares, entre outros.

Das principais funcionalidades do OR destacam-se [22]:

- Suite completa de otimização de preços, margens de lucro e preços de venda centrados no cliente.
- Recursos de autenticação unificados, painéis e geração de relatórios em todas as soluções.
- Funcionalidade de controlo de finanças e *stock* geral para auxiliar os retalhistas nos processos de expansão dos seus negócios para outros países.
- Através de recursos avançados de segurança, protege as informações dos clientes, além de auxiliar nas iniciativas de cumprimento da norma *Security Standards Council (PCI)* (PCI é o conjunto de regras que protegem estabelecimentos e consumidores de fraudes de cartão de crédito).

As soluções *Oracle* para o retalho são compostas por vários módulos tal como pode ser observado na figura 2.2 que se encontra na página 13. Na referida figura estão também representadas as fases em que são utilizados. Os módulos podem ser escolhidos e implementados de forma independente, conforme as necessidades e escolhas do cliente. Cada módulo possui igualmente funções e características próprias.

Figura 2.2: Módulos que compõe o *Oracle Retail*

2.4 Novas Tecnologias

Para o estudo da solução proposta e no sentido de garantir a familiarização com as ferramentas adiante mencionadas, procedeu-se a uma pesquisa de tecnologias que serão de seguida apresentadas de forma simplificada permitindo, assim, uma melhor compreensão dos capítulos seguintes.

2.4.1 Processos Robóticos Automatizados (RPA)

RPA é uma tecnologia inovadora que automatiza processos estruturados de negócio, permitindo a qualquer um configurar software de computador, ou a um “robô” imitar e integrar as ações de um ser humano interagindo com sistemas digitais para executar processos de negócio. Os robôs RPA utilizam a interface do utilizador para captar dados e manipular aplicações tal como o ser humano faz. Estes interpretam, acionam respostas e comunicam com outros sistemas para executar uma vasta variedade de tarefas repetitivas.

Em contraste com outras soluções TI tradicionais, o RPA permite às organizações automatizar uma fração do custo e do tempo encontrados previamente. São também de natureza não intrusiva e aproveitam a infraestrutura existente sem causar a interrupção dos sistemas subjacentes, difíceis e dispendiosos de substituir. Com RPA a eficiência de custos e a conformidade não são mais um custo operacional mas sim um subproduto da automação[23].

O RPA oferece às organizações a capacidade de reduzir custos de pessoal e reduzir o erro humano pois "ele nunca dorme e não comete erros". Proporciona então um aumento da conformidade e qualidade de execução dos processos e também uma melhoria na sua monitorização. Como exemplo da aplicação de RPA temos a experiência apresentada pelo diretor da Deloitte LP, David Schatsky. A situação reporta-se à implementação de RPA num banco. A instituição em causa redesenhou o seu processo de reclamações implantando 85 *bots* para executar 13 processos, lidando

com 1,5 milhões de solicitações por ano. O resultado foi o banco conseguir acrescentar capacidade equivalente a mais de 200 funcionários em tempo integral, com aproximadamente 30% do custo que tal recrutamento implicaria[24].

2.4.2 *Software as a Service*

Software as a Service (SaaS) é um modelo de licenciamento de *software* no qual o acesso ao *software* é fornecido por assinatura, com o *software* localizado em servidores externos e não em servidores internos. Geralmente, o SaaS é acessado por meio de um navegador da Web, com os utilizadores a fazer login no sistema com nome de utilizador e palavra passe. Em vez de cada utilizador ter que instalar o *software* no seu computador, passa a poder aceder ao programa pela Internet. As empresas geralmente usam o SaaS na gestão de retenção de clientes, recursos humanos e compras. Empresas de tecnologia, empresas de serviços financeiros e empresas de serviços públicos têm liderado o mundo dos negócios na adoção da tecnologia SaaS[25]. Algumas de empresas que fornecem serviços do tipo SaaS são a *Google* através das suas aplicações na nuvem como o *Google Docs*, o *Google Drive* ou a *Dropbox* com a solução de armazenamento na nuvem.

O SaaS elimina a necessidade das organizações instalarem e executarem aplicações nos seus próprios computadores ou nos seus centros de dados, sendo também fácil de atualizar e implementar. Isto elimina as despesas de aquisição, aprovisionamento e manutenção de *hardware*, e também o custo das licenças, instalação e suporte de *software*. Outras vantagens do SaaS incluem pagamentos flexíveis e atualizações automáticas.[26]

2.4.3 *Machine Learning*

Machine Learning (ML) é a ciência que possibilita a aprendizagem dos computadores fazendo com que os mesmos se comportem como seres humanos. permite igualmente que melhorem a sua aprendizagem ao longo do tempo, autonomamente, através de dados e informações na forma de observações e de interações do mundo real. É então um método de análise que automatiza a construção de modelos analíticos, sendo um ramo da Inteligência Artificial baseado na ideia de que os sistemas podem aprender com dados, identificar padrões e tomar decisões com o mínimo de intervenção humana[27].

Existem várias abordagens para fazer com que as máquinas aprendam, desde o uso de árvores de decisão até ao agrupamento de camadas de redes neuronais. Mediante o que se pretende fazer e o tipo e quantidade de dados disponíveis define-se qual o algoritmo a usar. Para treinar estes algoritmos existem três formas diferentes de o fazer:

- **Aprendizagem com supervisão** - quando são conhecidos e classificados os dados de teste.
- **Aprendizagem sem supervisão** - quando o algoritmo não tem informações sobre os dados que utiliza.

- **Aprendizagem com *reinforcements*** - que funciona em base da tentativa/erro, existindo um *feedback* externo.

Para ser utilizado o ML é então necessário recolher dados, armazená-los e escolher um algoritmo que se adapte às necessidades previstas, aplicar esse algoritmo, treina-lo e testa-lo. São exemplos práticos de aplicação do ML o reconhecimento de padrões, algoritmos para videogames e até a condução autónoma que é desenvolvida atualmente[28].

2.5 Solução DIaaS

Data Integrity as a Service (DIaaS) é uma nova solução desenvolvida pela *Wipro* com o objetivo de substituir uma solução desatualizada e na qual se vai trabalhar. A aplicação ainda está em fase de desenvolvimento. Podemos então começar por explicar a aplicação *Argws*, que é usada atualmente e que será substituída pela DIaaS.

O *Argws* é uma ferramenta de controlo da qualidade de dados criada pela *Wipro* em 2004, capaz de monitorizar e alertar o utilizador para a existência de inconsistências no seu sistema, com o objetivo de garantir a qualidade e consistência de dados entre vários sistemas. Esta realiza comparações periódicas entre dados provenientes de diferentes sistemas e alerta então o utilizador se houver inconsistências. É muito importante garantir a consistência dos dados pois os benefícios de uma grande qualidade nos dados são muitos e estão representados na figura 2.3. Esta ferramenta encontra-se desatualizada pois usa tecnologias antigas e uma interface pouco eficiente e pouco apelativa começando a não responder às necessidades do mercado atual.

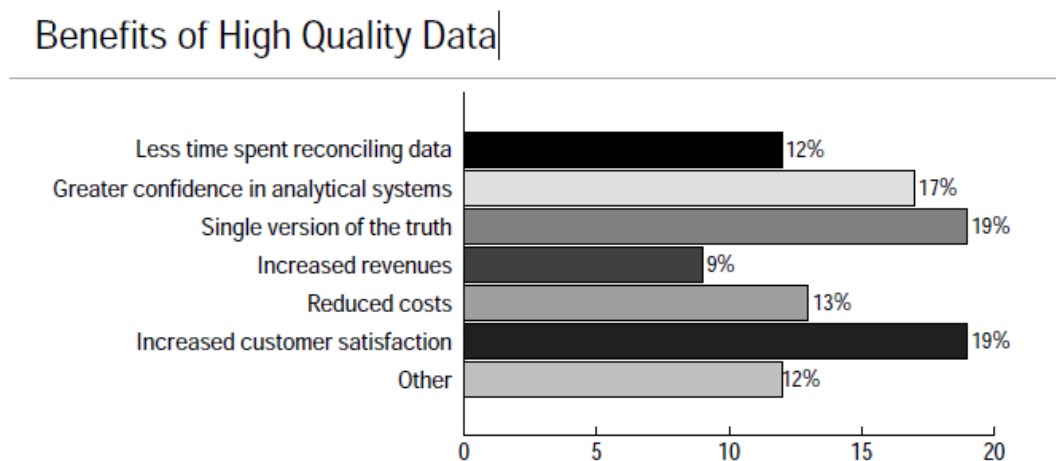


Figura 2.3: Benefícios dos dados com alta qualidade

O DIaaS é criado com o objetivo principal de substituir o *Argws* e, para além da criação das reconciliações, a aplicação permite agendá-las e visualizar os seus resultados através de *dashboards*, seguindo as melhores práticas de desenvolvimento de *software* e utilizando as tecnologias

mais recentes e apropriadas. Tem também maior conectividade pois permite a ligação a diferentes tipos de bases de dados. As tecnologias utilizadas por esta nova solução são:

- **Java Spring** - é uma *framework* criada para tornar o desenvolvimento de aplicações mais fácil, suportando também um conjunto muito variado de arquitetura e simplificando o processo de desenvolvimento de aplicações Java.
- **Angular** - é uma *framework* web desenvolvida pela Google que foi utilizada para a Interface. A utilização desta traz muitos benefícios como produtividade (templates intuitivos), velocidade e desempenho e o facto das aplicações que são construídas em Angular funcionarem em qualquer ambiente de implantação.
- **MySQL** - é um sistema de gestão de bases de dados relacionais que utiliza a linguagem *Structured Query Language* e que se destaca pela sua facilidade de uso, segurança e rápida recuperação de falhas e bloqueios.

O DIaaS é então basicamente uma ferramenta com maior compatibilidade, mais funcionalidades e mais intuitiva do que o Argws. Na tabela 2.1 estão representadas algumas funcionalidades das duas ferramentas de modo a facilitar a perceção das diferenças.

Tabela 2.1: Tabela de Funcionalidades do Argws e do DIaaS

	Argws	DIaaS
Utilizadores	Sistema básico de gestão de utilizadores (acessos baseados em permissões)	Conceito de cliente(SaaS) Gestão de utilizadores com suporte para papéis(e permissões) por cliente
Extração	Necessidade de especificar tipos de dados a receber nas <i>queries</i> Extração única por reconciliação Suporte apenas de bases de dados <i>Oracle</i>	Extração simplificada (apenas através da <i>query</i>) Adaptadores para mais bases de dados
Reconciliação	Apenas suporta a comparação de dados entre duas extrações	Multi-comparação Reutilização de extrações Suporte para diferentes <i>primary keys</i> e comparações entre extrações
Implantação	Necessidade de <i>deployment</i> no cliente	SaaS Mesmo servidor pode servir vários clientes
Tecnologias	<i>OracleDB, Java Server Pages, Pro*C, Kornshell</i>	<i>Java Spring, Angular, MySQL</i>

2.6 Integridade de Dados(*Data Integrity*)

A integridade dos dados é a manutenção e a garantia de exatidão e consistência dos dados durante todo o seu ciclo de vida, e é um aspeto crítico para o desenho, implementação e utilização de qualquer sistema que armazene, processe ou retorne dados. Pode descrever o estado dos dados

(válido ou inválido) ou o processo de garantir e preservar a validade e precisão dos mesmos. A intenção geral de qualquer técnica de integração de dados é garantir que os dados são guardados exatamente como pretendido e, numa futura utilização desses dados, garantir que são os mesmos que foram originalmente gravados. Ou seja, a integridade dos dados visa impedir alterações não pretendidas à informação[29].

2.6.1 A Importância da Integridade dos Dados

Hoje em dia as informações são movidas online ou digitalizadas, resultando numa quantidade crescente e em tipos variados de dados a ser recolhidos. Assim, os dados são essenciais para a prossecução dos objetivos das organizações. Contudo, sem integridade, os dados não são passíveis de serem utilizados convenientemente, principalmente quando nos referimos à área do retalho onde quase todas as decisões tomadas são sustentadas nestas informações.

Os dados não são estáticos, existem inúmeras ocorrências desde o momento em que são criados no sistema e durante todo o seu ciclo de vida. Podem ser transferidos para outros sistemas, podem ser alterados e atualizados várias vezes. Durante este ciclo é de extrema importância manter a integridade dos mesmos, tarefa essa que é complexa. [30].

A integridade dos dados garante a capacidade de recuperação e a procura dos dados, a rastreabilidade e a conectividade. Assim, a proteção da validade e precisão dos dados aumenta a estabilidade e o desempenho dos sistemas. Na maioria das empresas e a vários níveis, os funcionários tomam decisões de negócios que contribuem para o sucesso ou insucesso da empresa, decisões essas que são baseadas nestas informações. Portanto, qualquer empresa é capaz de garantir a integridade dos seus dados se as suas operações vitais, as decisões tomadas ou os serviços prestados estiverem a ser executados corretamente. Tudo isto afeta diretamente o futuro da empresa pois pode conduzir a erros, oportunidades perdidas, perdas monetárias, entre outras.[31]

Um novo relatório da KPMG (uma das maiores empresas de prestação de serviços profissionais, que incluem auditoria, impostos e consultoria de gestão e estratégia [32]) revela que a grande maioria dos executivos seniores não tem um nível de confiança alto na forma como a sua organização utiliza dados, análises ou Inteligência artificial. O estudo revela que apenas 35% tem um alto nível de confiança na forma de como a sua organização utiliza os dados e análises e 92% estão preocupados com o impacto negativo dos dados e análises na reputação da organização[33].

A título de exemplo prático, podemos aludir ao sistema de um banco. Estes armazenam informações relativas aos clientes nas suas bases de dados, informações essas que são utilizadas, por exemplo, quando se efetua um levantamento. Se essas informações estiverem erradas podemos não conseguir levantar (por exemplo o *pin* associado ao cartão ser um errado) ou até levantar de uma outra conta, entre outros possíveis problemas. Numa empresa de retalho temos, por exemplo, a questão das decisões tomadas com base na quantidade de um artigo para encomenda. Nestas circunstâncias, se os dados relativos ao stock desse artigo estiverem incorretos, a decisão tomada terá um impacto negativo na empresa dado que o stock realizado passa a ser excessivo ou insuficiente. Em qualquer um dos casos, pode resultar em perdas a nível financeiro.

2.6.2 Manter a Integridade dos Dados

Como referimos anteriormente, os dados são muitas vezes alterados, acidentalmente ou não, e muitas destas alterações ocorrem durante a sua transferência de um dispositivo para outro. Entre muitas ameaças à integridade dos dados as mais comuns são [33]:

- **Erro humano**
- **Erros de transferências não intencionais**
- **Configurações incorretas e erros de segurança**
- **Malware, ameaças internas e ataques cibernéticos**
- **Hardware comprometido**

Quando se vai utilizar os dados é necessário ter confiança nessas informações. Para saber se estes dados têm integridade, analisam-se as seguintes características [33]:

- **Acessibilidade e recuperabilidade** - é importante ter dados precisos nos locais adequados e no momento certo, ou seja, quando alguém está a trabalhar e necessita de mobilizar essa informação, os dados devem estar disponíveis. Sem o acesso e recuperação adequados e facilitados o negócio pode sair prejudicado abrindo caminho para que a concorrência ganhe vantagem.
- **Rastreabilidade** - pode-se rastrear todos os pontos de contacto feitos com um cliente através de um ponto de dados. Os dados podem informar quem vai tomar as decisões sustentadas na sua informação, destacar bandeiras vermelhas, deficiências ou limitações.
- **Confiabilidade** - ter métricas de negócio confiáveis e consistentes em relação aos objetivos da empresa e à concorrência é essencial para uma organização de topo.

Para preservar a integridade dos dados existem vários métodos, nomeadamente os constantes da lista de verificação que se segue [33]:

- **Validar entradas** - quando um conjunto de dados é fornecido por uma fonte conhecida ou desconhecida deve-se exigir a validação da entrada. Esses dados devem ser verificados e validados para garantir que a entrada seja precisa.
- **Validar dados** - é fundamental verificar que os dados do processo não foram corrompidos. Deve-se identificar especificações e atributos chave que são importantes para a organização, antes de validar os dados.
- **Remover dados multiplicados** - dados confidenciais de uma base de dados segura podem ser facilmente encontradas num email ou em pastas compartilhadas, tornando-os acessíveis a lementos da organização que não deveriam ter essa possibilidade. É prudente limpar sempre os dados perdidos e remover os duplicados.

- **Backup de dados** - além de remover os duplicados para garantir a segurança dos dados, é fundamental considerar que estes são uma parte do processo, são necessários e ajudam a evitar a perda permanente de dados.
- **Controlo no acesso** - tanto a nível online como a nível físico os acessos devem ser controlados.
- **Manter sempre um rasto de auditoria** - sempre que ocorre uma falha é essencial para a integridade dos dados conseguir localizar a origem. Isto fornece à organização o necessário para identificar com precisão a origem do problema. Por norma, um rasto de auditoria necessita de ser gerado automaticamente, não deve haver acesso de utilizadores para realizarem alterações, todos os eventos devem ser registados e a cada evento deve ser associado um utilizador e hora.

Capítulo 3

Situação Atual - Análise

Atendendo a que o objetivo deste projeto é o da melhoria da forma de como o suporte dado aos clientes pela *Wipro* é realizado, aumentando a sua eficiência e diminuindo o tempo disponibilizado, neste capítulo é feita uma análise detalhada desse processo no sentido de garantir a sua automação e integração na aplicação anteriormente mencionada, o DIaaS.

3.1 O Suporte Prestado Atualmente pela *Wipro*

Quando é implementada numa empresa uma solução OR pela *Wipro*, normalmente segue-se um processo de suporte para resolver incidentes e problemas relacionados com as soluções implementadas ou para a manutenção dos processos necessários e das bases de dados utilizadas.

Por norma, a *Wipro* só não fornece suporte das soluções implementadas no cliente quando este tem na sua organização um departamento de TI habilitado a executar essas funções. Pode também ocorrer situações em que o processo de suporte é dividido pela *Wipro* e pelo cliente.

É também importante referir que, muitas vezes, para se implementar uma solução é necessário a personalização da mesma. A *Wipro* procede no sentido de garantir a devida adaptação às necessidades do modelo de negócio do cliente.

Estes serviços de suporte prestados pela *Wipro* estão sempre em constante negociação com o cliente de modo a permitir que o serviço prestado satisfaça todas as suas necessidades.

3.1.1 Processos de Suporte

A *Wipro* fornece diferentes tipos de suporte aos seus clientes, suporte esse que pode ser de primeira ou segunda linha. No suporte de segunda linha não são executadas operações diretamente no cliente, são antes enviadas soluções ou sugestões ao respetivo departamento TI. Normalmente, quando é o cliente a efetuar o suporte de primeira linha, este resolve apenas os casos mais simples e para os outros espera pelas soluções enviadas pela *Wipro*.

O tipo de suporte prestado pode também ser classificado como funcional ou técnico. Funcional quando é explicado o modo de funcionamento e o conjunto de passos a seguir para resolver casos

específicos. O técnico refere-se ao serviço de suporte técnico que normalmente é a alteração do código, de modo a implementar melhorias.

De seguida, são apresentados os processos de suporte relevantes para a compreensão deste projeto:

- **Serviços *on call*** - no mundo do retalho é comum que grande quantidade de dados seja processada durante a noite, pois é neste período que a atividade se encontra parada (como por exemplo para comparar o nível de stock entre dois sistemas, durante o dia a comparação não pode ser feita pois estão constantemente a ocorrer alterações e ao fazer a extração dos dados não vamos ter um valor coerente nos dois sistemas). Para realizar este tipo de serviço é necessário que haja alguém (24 horas por dia, 7 dias por semana) disponível para responder a solicitações e pedidos ligados à monitorização dos processos e ocorrência de falhas.
- **Monitorização e Análise das Diferenças** - neste tipo de processo são detetadas, analisadas e corrigidas anomalias que possam surgir e que podem vir a ter um grande impacto na gestão do cliente, como por exemplo a análise das diferenças de inventário entre dois sistemas cujo impacto na gestão da empresa é significativo.
- **Resolução de Incidente e Problemas** - este é o tipo de suporte que representa a principal função de uma equipa de suporte durante o seu período de atividade. Neste, o cliente reporta funcionalidades que não estão a funcionar conforme especificado e a equipa de suporte analisa a causa dessa anomalia para, de seguida, proceder à sua correção.

É importante que todo o suporte dado aos clientes seja rápido e eficaz pois caso contrário ocorrerá um impacto negativo no negócio do cliente e consequentemente a insatisfação do mesmo para com os serviços prestados.

3.2 Modelo Atual

O suporte prestado atualmente pela *Wipro* utiliza uma ferramenta de controlo de qualidade designada *Argws*, já descrita anteriormente no capítulo 2 e que é capaz de monitorizar e alertar o utilizador para a existência de inconsistências no seu sistema. O *Argws* compara então dados entre dois módulos pertencentes ao *software OR* e emite alerta caso sejam encontradas inconsistências nesses dados. Ou seja, através do *Argws* os utilizadores conseguem observar a existência ou não de inconsistências entre os dados.

De modo a facilitar a compreensão do problema, consideramos o caso das inconsistências corresponderem a diferenças de valor (por exemplo: diferenças no valor do inventário, de faturas...). Quando se presta suporte a um cliente da área do retalho é importante que os dados estejam coerentes entre todos os módulos. Nesse sentido, o cliente e a *Wipro* estão em constante comunicação. Nada neste processo de suporte é automatizado e a comunicação entre a equipa de suporte e o cliente é feita através do envio de ficheiros excel.

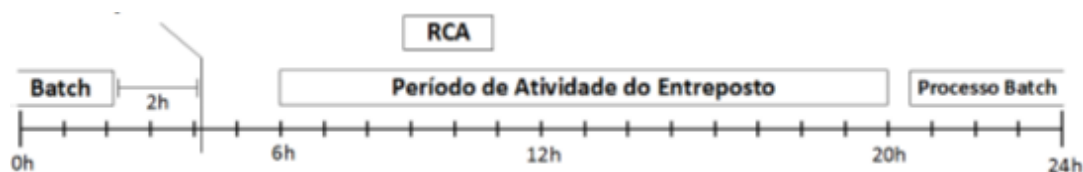


Figura 3.1: Período de atividade dos processos num entreposto durante o dia

A extração dos dados normalmente é feita após o processo *batch* ter terminado e não se encontrando dentro do período de atividade do entreposto, como representado na figura 3.1.

Após a extração dos dados dá-se início ao processo de suporte efectuado pela *Wipro*. Este começa de madrugada com a criação de um ficheiro excel que contém os dados comparados nas reconciliações e os seus valores nos respetivos módulos. De seguida, é realizado o tratamento das diferenças e enviado um relatório final para o cliente. Este tratamento dos dados extraídos é feito manualmente e são executados os seguintes passos antes da sua concretização de modo a analisar as causas das diferenças, representados na figura 3.2 e explicados a seguir:

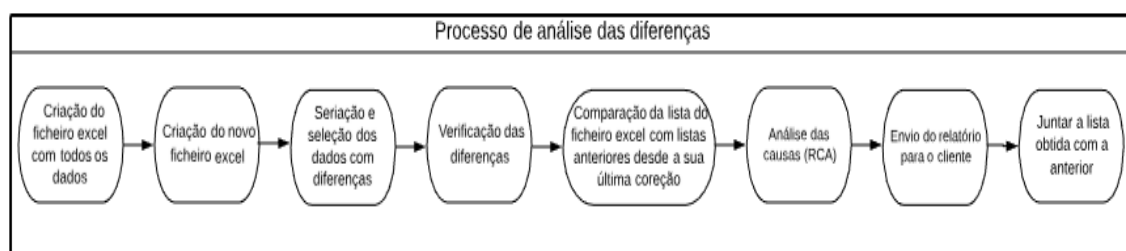


Figura 3.2: Processo de análise do suporte

- **Criação do ficheiro excel referido anteriormente** - este vai conter os dados a comparar e os valores dos mesmos nos respetivos módulos.
- **Criação de um novo ficheiro excel** - este vai servir de orientação para todo o processo, contendo todas as colunas necessárias ao relatório final a ser enviado para o cliente.
- **Sieriação e seleção dos dados com diferenças** - neste passo é feita uma seleção dos dados que contêm diferenças a partir do ficheiro excel criado no primeiro passo que depois vão ser copiados para o ficheiro criado no segundo passo.
- **Verificação das diferenças** - este processo é uma segunda verificação dos dados com diferenças e serve para descartar à priori falsas diferenças uma vez que a eventual causa da diferença já possa ter sido corrigida e o valor ajustado. Para isso é feita uma *query* à base de dados do cliente.
- **Comparação da lista do ficheiro excel com listas anteriores desde a sua última correção** - este tem como objetivo retirar dados com diferenças dos quais já foram analisadas e

identificadas as causas das mesmas. Isto deve-se ao facto de as ações de correção serem semanais, o que faz com que os mesmos erros apareçam até que a solução seja aplicada.

- **Análise das causas (RCA)** - este passo tem por base o conhecimento adquirido pela equipa de suporte e é executado segundo o método de tentativa e erro, através de *queries* que permitem identificar causas já conhecidas. Caso a causa não seja descoberta é feita uma análise individual e manual que pode ser muito morosa e inconclusiva.
- **Envio do relatório para o cliente** - é enviado por email o ficheiro excel criado no segundo ponto contendo o resultado das análises e as ações a tomar.
- **Juntar a lista obtida com a anterior** - o que permite ter um relatório geral com todos os dados que apresentam diferenças.
- **Executar as ações de correção** - quando chega a altura prevista, são executadas, na base de dados do cliente, de forma a reestabelecer os valores reais (corretos). Este passo normalmente é executado no final da semana, quando os serviços logísticos do cliente estão fechados.

Como já foi referido anteriormente, para um bom funcionamento do sistema é necessária uma sincronização dos dados entre todos os módulos do OR, e para isso a RCA é essencial pois permite a descoberta da solução a aplicar. Este processo designado RCA pode ser dividido nas seguintes tarefas:

- **Extração e comparação dos dados** - aqui é feita a extração dos dados das bases de dados e comparados os mesmos.
- **Análise das causas das diferenças** - são analisados os dados com diferenças.
- **Envio ao cliente** - após a análise é elaborado e enviado para o cliente um relatório que contém os dados com diferenças e as suas causas.
- **Correção** - de acordo com a resposta do cliente, ao relatório enviado são executadas as ações de correção de forma a garantir que as diferenças são corrigidas e que os valores presentes nos sistemas são reais.

Todos os dias na *Wipro* há elementos dedicados ao suporte do cliente que despendem muito tempo a analisar e corrigir erros dos quais, na sua maioria, são já conhecidas as causas e soluções. O trabalho destas pessoas é assim repetitivo, dispendioso e, de certa forma, pouco produtivo, dado que em vez de se proceder exclusivamente à análise de novos casos, são também trabalhadas situações já conhecidas. É também importante referir que nenhum dos passos mencionados neste capítulo é automatizado.

3.3 Dimensão do Problema

Com o objetivo de ajudar à compreensão da dimensão e do volume de trabalho associados ao suporte dado atualmente pela *Wipro* irão ser apresentados alguns dados sobre os processos de suporte realizados. Como não há dados concretos do tempo despendido por uma equipa de suporte realizou-se, com base em reuniões informais, uma estimativa. Podemos afirmar que o suporte está dividido em várias áreas e que para cada uma delas temos uma equipa que normalmente é constituída por cerca de 6 elementos. Estes elementos perdem, em média, 2 horas por dia só no processo de garantir que os dados estão síncronos entre os módulos OR. Existindo dados com discrepâncias será despendido mais tempo para proceder à sua análise e correção, o que permite concluir que por mês sejam despendidas, no mínimo, cerca de 60 horas só para garantir a sincronia dos dados.

Foi feito recentemente um estudo interno na *Wipro* relativo ao tempo de análise perdido na identificação da causa do problema (discrepância entre os módulos) referente à análise das diferenças de stock para um cliente durante um período de 4 meses. No estudo, verificou-se que o tempo gasto na análise dessas diferenças foi em média de 49,1 horas por mês, como se pode verificar na tabela 3.1. A tabela em causa integra, para além do tempo despendido, o número de artigos analisados.

Tabela 3.1: Número de artigos e tempo despendido na sua análise

Mês	Número de artigos	Tempo de Análise
Janeiro	1600	26,7 h
Fevereiro	4574	57,9 h
Março	7844	41,4 h
Abril	4081	44,1h
Maio	6186	75,4h
Média	4857	49,1 h

Como se pode verificar pelos dados da tabela acima apresentada, a análise dos artigos que contêm diferenças de *stock* gasta em média 49,1 horas por mês o que corresponde a mais de 6 dias de trabalho de um trabalhador, sendo portanto um processo demorado e, por isso mesmo, muito dispendioso.

No sentido de diminuir o tempo gasto com esta análise e com a verificação da sincronia dos dados é necessário automatizar os processos ou simplificar as tarefas dos utilizadores de modo a que o seu tempo seja gasto em atividades mais produtivas.

Como já explanado no capítulo 2, a ferramenta utilizada atualmente, o *Argws*, está desatualizada para estes processos e não permite a automatização dos mesmos. Daí ter sido criada a *DIaaS* com melhores tecnologias e mais funcionalidades e que suporta alterações de modo a melhorar as suas funcionalidades e até criar novas.

O DIaaS permite acrescentar todas as funcionalidades necessárias devido às tecnologias com que foi desenvolvido. Contudo, nunca foi utilizado nem testado num ambiente real, o que faz com que sejam necessárias algumas alterações base, para além do acréscimo de funcionalidades que suportem as referidas adições.

3.4 Oportunidades de Melhoria

A análise de todo o processo descrito na secção anterior 3.2, foi essencial para detetar os pontos críticos do processo realizado atualmente e permitiu também destacar algumas oportunidades de melhoria. De seguida, são explicadas as oportunidades de melhoria que serviram como base para o desenvolvimento da solução proposta descrita no capítulo seguinte.

Dentro dos pontos mais críticos deste processo de suporte destaca-se a total ausência de automatismo nas tarefas realizadas, que são tarefas repetitivas, demoradas e que têm de ser executadas diariamente.

A inexistência desses automatismos faz recair todos os procedimentos nos meios humanos, tornando as tarefas dispendiosas monetariamente e com uma maior tendência à ocorrência de erros e falhas nomeadamente de duplicação de dados, de erros de digitação que podem levar a erros de comunicação ou de dados, ou até mesmo do esquecimento ou falha no processo de copiar os artigos com erros para as folhas de Excel. Além destes inconvenientes temos ainda o facto de, por vezes, ser apenas um elemento a realizar a tarefa, o que pode levar a atrasos caso este não consiga responder a todos os pedidos na sequência de uma sobrecarga de tarefas ou até mesmo por incumprimento resultante da falta ao serviço.

Além das vantagens acima mencionadas e relativas ao automatismo, temos também o facto de estar a ser ineficaz a utilização de profissionais altamente qualificados, com muitos conhecimentos numa área tão específica como é o Retalho, que acabam por despende o seu tempo em tarefas repetitivas de suporte e que não acrescentam valor ao sistema. Registe-se ainda a falta de uma junção dos resultados obtidos nas análises realizadas e dos conhecimentos obtidos. Uma vez que a *Wipro* funciona a nível mundial com vários clientes, a centralização dos resultados obtidos permitiria a partilha de informação entre diferentes equipas de suporte, o que iria facilitar muito o trabalho dessas equipas dado que muitas das análises efetuadas, apesar de estarem relacionadas com clientes diferentes, podem ser iguais.

Destacam-se também oportunidades de melhoria na ferramenta a ser utilizada, o DIaaS, pois como este nunca foi usado em ambiente real e nunca foi criado um algoritmo para analisar os erros, serão necessárias algumas alterações de modo a que a ferramenta esteja pronta para a implementação do referido algoritmo.

Capítulo 4

Solução Proposta - Descrição

Na sequência da análise descrita no capítulo anterior, deu-se início à construção de uma solução capaz de auxiliar a equipa de suporte na realização do seu trabalho. Esta deve também ser dinâmica de modo a ser possível adaptá-la para ser usada por várias equipas, trazendo assim benefícios para a empresa, tanto a nível da inovação como a nível económico.

Foi então pensado um algoritmo que pudesse ser flexível e adicionado à ferramenta usada pelo suporte de modo a automatizar parte do seu trabalho e adaptado para que possa ser utilizado por todas as equipas nos respetivos clientes.

De seguida, são apresentados alguns conceitos importantes que serão mobilizados frequentemente, sendo os mesmos indispensáveis para a compreensão do projeto. Será também realizada uma abordagem ao problema, apresentada a metodologia adotada e descrita a solução.

4.1 Conceitos Necessários

Dado que a solução aqui apresentada trata de um tipo de suporte em particular, há terminologia utilizada que só faz sentido no contexto do suporte referido, não sendo usual a sua mobilização. Procederemos então a uma breve explicação de todos os conceitos utilizados no sentido de garantir um melhor entendimento dos assuntos abordados.

- **Extração** - refere-se, como o nome indica, à extração de dados de um sistema.
- **Reconciliação** - é a criação de extrações de dados de vários sistemas do cliente e a comparação desses mesmos dados.
- **Erro** - refere-se ao problema ou falha que está na origem do processo de análise e que se manifesta como uma diferença entre as informações desse sistema. Quando se diz que uma reconciliação deu erro, esse erro refere-se à diferença.
- **Causa** ou Causa Raiz (*Root Cause*) - refere-se ao motivo ou justificação de um erro. Para um erro pode haver mais do que uma causa.

- **Análise da Causa Raiz(RCA)** - é o processo de procurar e identificar a causa da ocorrência daquele erro.
- **Solução** - a solução indica o que deve ser feito para se corrigir um erro, ou seja, discrimina as ações que se devem realizar. Uma solução pode ter uma ou várias ações.
- **Ação** - é o procedimento a executar para corrigir os erros. A ação a aplicar deve estar de acordo com o tipo de erro e a causa do mesmo. Um exemplo de ação é o ajuste manual de inventário no RMS (módulo do *Software OR*) para um valor igual ao do WMS (módulo do *Software OR*)

4.2 Abordagem ao Problema

Numa primeira abordagem, mais geral, a solução pode ser separada em duas partes: a automatização das tarefas e a dinâmica da solução, que vai ser feito através da criação de um algoritmo que deverá identificar e corrigir os erros.

Numa abordagem mais específica, a solução desenvolvida pode-se separar em três partes: a caracterização do erro, as suas RCA e as soluções.

O objetivo da solução proposta é a automatização das tarefas realizadas que são repetitivas e não acrescentam valor ao processo permitindo assim que a equipa de suporte apenas se dedique à realização das tarefas que não podem ser automatizadas, conseguindo-se assim aproveitar mais eficientemente a experiência e o conhecimento dos profissionais que constituem as equipas.

Para decidir o que poderia ou não ser automatizado e perceber o atual funcionamento do processo efetuado pelo suporte, foram realizadas várias reuniões internas nas quais foi possível retirar o modelo atual (modelo *As Is*) do processo realizado e que se encontra representado na figura 4.1, onde se pode verificar que o mesmo é manual com exceção da execução das reconciliações que é um processo automático.

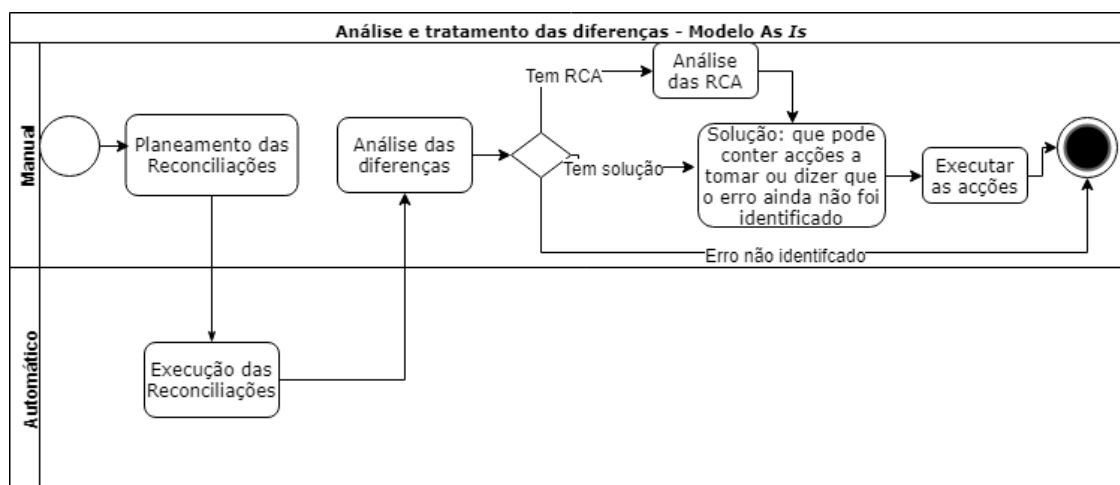


Figura 4.1: Processo atual

Após o estudo deste modelo, foram definidas as tarefas que poderiam ou não ser automatizadas e com base nessas decisões foi então desenvolvido um novo modelo, o modelo *To Be*, representado na figura 4.2. Neste, podemos verificar que quase todos os passos do processo foram automatizados, com exceção da correção. Para que este passo possa ser automatizado será necessário, antes de mais, que o cliente ganhe alguma confiança na solução.

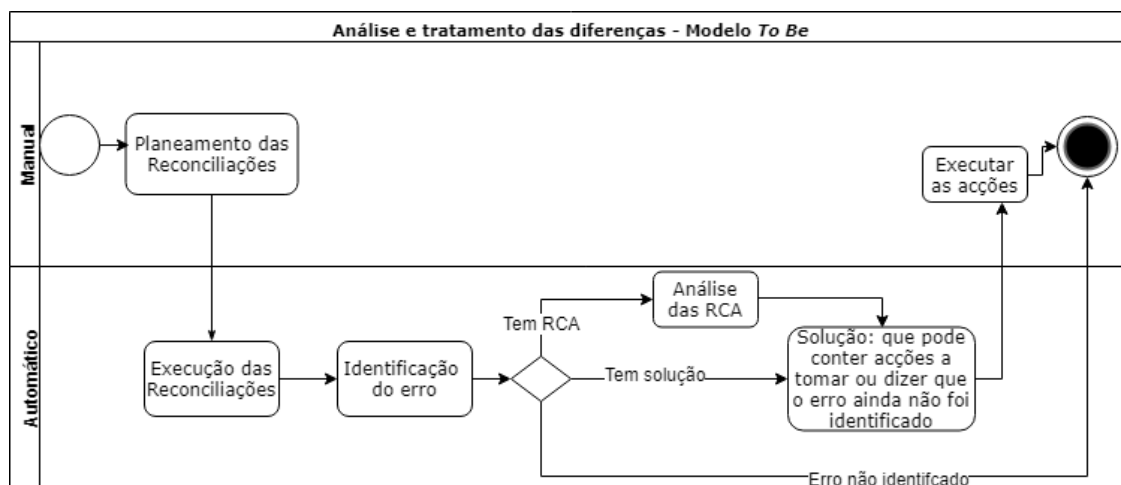


Figura 4.2: Processo futuro

No novo modelo, a equipa de suporte dá início ao processo, isto é, define as reconciliações a serem realizadas, determinando se as vai realizar no momento ou se deverão ser agendadas. Após a definição, o programa vai correr a reconciliação e verificar se tem erros. Tendo erros vai identificá-los. Este processo de identificação é fundamental, tal como explicaremos mais à frente. Após a identificação vai-se verificar se têm solução ou se é necessário correr as causas (RCA) para encontrar uma solução e, por fim, enviar as ações necessárias a realizar pela equipa de suporte de forma a aplicar a respetiva solução. A equipa de suporte passa assim a desempenhar apenas as seguintes tarefas: iniciar o processo, aplicar as ações e analisar novos erros, na eventualidade destes surgirem e de forma a acrescentá-los ao algoritmo.

Após redefinir o processo foi introduzida uma componente, Identificação do erro, e retirada a análise das diferenças, como é possível verificar a figura 4.2. Este processo de identificação do erro é muito importante pois tem de ser particular e específico para cada um destes. Após identificado o erro, vamos ver as soluções ou RCA correspondentes. Se este não for corretamente identificado podemos estar a adotar procedimentos incorretos. A solução proposta deverá então ter um conjunto de dados e informações que permita identificar os erros e as soluções ou causas correspondentes e sugerir as ações a tomar para os corrigir.

É de maior importância que esta solução seja dinâmica de modo a que seja possível ser adaptada a vários clientes e dessa maneira ser utilizada por várias equipas de suporte.

Para além dos aspetos anteriormente referidos, foi também necessário analisar a ferramenta a ser utilizada de modo a verificar se esta estava pronta para a implementação do algoritmo. Foi então verificado que era necessário alterar o modo de execução das reconciliações e o resultado das

mesmas de forma a que o resultado da reconciliação fosse mais complexo, com mais parâmetros, para uma melhor identificação do erro.

4.3 Metodologia Utilizada

A metodologia que foi seguida no desenvolvimento desta solução passou por quatro fases que estão representadas na figura 4.3.

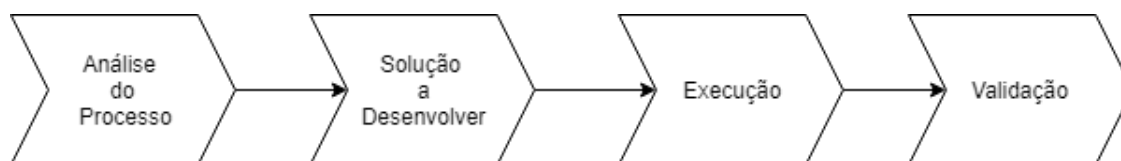


Figura 4.3: Metodologia utilizada

Como se pode observar na figura, a primeira fase passou por uma análise do processo que resultou na elaboração do Modelo *As Is* representado na figura 4.1 que se encontra no capítulo 4.2. Este tem como objetivo a compreensão do processo atual de modo a que possam ser identificados os pontos críticos e as oportunidades de melhoria do mesmo.

Na segunda fase, foi pensada a solução e como resultado elaborou-se o Modelo *To Be* que está representado na figura 4.2. Este modelo tem como objetivo explorar as oportunidades de melhoria verificadas no passo anterior e perceber os benefícios que daí irão surgir.

Posteriormente, seguiu-se a terceira fase, a da execução, ou seja, o desenvolvimento da prova de conceito. Para a sua realização foram necessárias alterações no DIaaS para que o algoritmo a desenvolver pudesse ser aplicado. A prova de conceito é então neste caso muito importante na demonstração de como vai identificar os erros e aplicar as soluções para que no futuro os clientes onde se vai aplicar a solução ganhem confiança e permitam que o processo seja totalmente automático. A complexidade da prova de conceito está de acordo com o tempo e quantidade de dados disponibilizados, pois quanto maior for a quantidade de dados, mais erros serão aplicados.

Numa última fase, foi feita uma validação da solução, ou seja, foi verificado se a solução funciona para vários tipos de erros e se pode ser integrada em vários ambientes.

4.4 Descrição da Solução

Tendo em conta que a solução é para ser integrada na aplicação DIaaS e que vai lidar com vastas quantidades de dados em ambientes reais e em diferentes clientes, foram efetuadas reuniões com elementos de equipas de suporte onde foi elaborada uma lista de requisitos e limitações da solução, para que esta correspondesse ao pretendido.

4.4.1 Requisitos da Solução

- **Flexibilidade** - A solução será usada em ambientes e clientes diferentes, logo deve estar preparada para ser utilizada em diversas situações.
- **Redução do esforço humano** - A intervenção humana deverá ser bastante reduzida com a aplicação da solução pois serão automatizados processos que anteriormente eram manuais. Esses processos automatizados também conduzirão a uma redução do tempo de operação.
- **Extensibilidade** - A solução deverá ser extensível, ou seja, deverá ser possível acrescentar novos erros à solução sempre que surgirem e desse modo torná-la mais complexa ao longo do tempo. Ao acrescentar tipos de erros também se acrescentam as soluções de modo a que o algoritmo ao percorrer os erros os identifique e aplique a solução correspondente.
- **Assertividade** - A solução deverá ser o mais assertiva possível pois identificar mal um erro ou aplicar soluções que não correspondem ao erro podem trazer consequências devastadoras para o cliente. Portanto, o processo de identificação deve garantir que não identifica erros diferentes como sendo do mesmo tipo.
- **Informativa** - Deverá ser possível aos utilizadores visualizarem todos os passos realizados pelo algoritmo, em que estado estão e qual as soluções que foram aplicadas. E quando realizar RCA deve apresentar o resultado retornado por elas.

4.4.2 Limitações

Devido á especificidade da área de negócio que é o Retalho existem algumas limitações que foram consideradas e resolvidas.

Como a solução é para ser utilizada por equipas de suporte em clientes retalhistas, vai lidar com informações críticas do cliente e este pode não permitir que aplique soluções diretamente nas suas bases de dados, pelo que as ações a aplicar não serão automatizadas numa fase inicial.

Os ambientes reais onde a solução será aplicada pertencem a clientes e como tal não foi possível o acesso a esses ambientes, sendo por isso necessário criar um ambiente de teste que, embora não possua tantos dados, integra os dados necessários para que a solução fosse o mais realista possível.

4.5 Explicação do Funcionamento da Solução

Neste subcapítulo será apresentado o funcionamento da solução e as alterações efetuadas ao programa. O funcionamento vai ser brevemente explicado com o objetivo de se compreender a realização de alterações ao programa. As alterações efetuadas ao programa, no sentido de poder aplicar a solução, serão explicadas mais detalhadamente logo de seguida.

4.5.1 Algoritmo

De um modo mais geral, o funcionamento do algoritmo desenvolvido pode ser dividido em três fases. Numa primeira fase o algoritmo vai receber o resultado da reconciliação e vai verificar se para o tipo de erro encontrado já existe.

Existindo o erro, passa para a segunda fase em que é verificado se o erro identificado tem associada uma solução ou RCA. Quando já integra a solução esta é aplicada diretamente e termina.

Tendo RCA passa para uma terceira fase em que pode ser corrida uma nova reconciliação ou realizadas as RCA para descobrir a causa. Realizando as RCA's retorna os seus resultados e aplica a solução correspondente.

4.5.2 Alterações ao DIaaS

Para que o algoritmo pudesse funcionar perfeitamente, identificar os erros e aplicar as soluções corretamente, foi pedido à equipa de desenvolvimento do DIaaS que realizasse algumas alterações ao programa. De entre todas as alterações realizadas destacam-se:

- Na reconciliação o acréscimo de um parâmetro, *schedule*, que permite uma reconciliação correr quando o algoritmo dá ordem para tal sendo automaticamente agendada como sendo a próxima a realizar. No modelo antigo não era possível correr a reconciliação automaticamente sem ter sido manualmente agendada.
- O modo como os erros eram apresentados, tal como constatamos na figura 4.4 que se encontra na página 33. No modelo antigo, os erros eram apresentados apenas com os campos visíveis em *recon_report_errors* onde apenas era possível visualizar os valores comparados e a reconciliação a que correspondiam. No modelo novo, temos o *comparison_error* onde para além dos valores e da reconciliação a que correspondem apresenta também as colunas a que pertencem e as *primary keys* correspondentes, assim como o tipo de erro (este é um tipo como por exemplo *value dif* quando os valores comparados são diferentes e *record not exists* quando para as mesmas *primary keys* tem valor num sistema e no outro não existe).

Além das alterações realizadas pela equipa de desenvolvimento do DIaaS, foram também desenvolvidas algumas tabelas que vão ser percorridas ou preenchidas pelo algoritmo. Estas serão descritas mais à frente e podem ser vistas no Anexo B, sendo as mais relevantes:

- **recon_solution_config**
- **recon_rca**
- **recon_report_rca**
- **recon_solution_action**

Todas as alterações ao DIaaS acima mencionadas podem ser vistas nos modelos relacionais, antigo e novo, que estão representados respetivamente nos Anexos A e B.

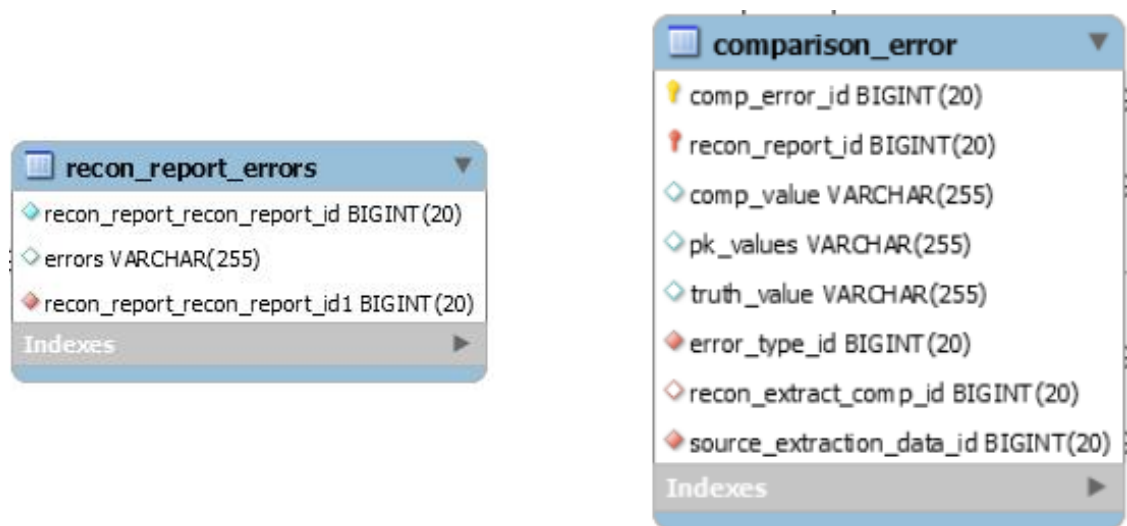


Figura 4.4: Tabelas de apresentação dos erros

Capítulo 5

Prova de Conceito e Resultados Obtidos

A prova de conceito realizada teve como objetivo automatizar ao máximo o processo levado a cabo pelas equipas de suporte de um modo estandardizado/uniforme/padronizado, de forma a que seja utilizada por diferentes equipas de suporte e permita comprovar que a utilização da solução vai simplificar e reduzir o tempo gasto na realização das tarefas a cumprir.

5.1 Permissões e Campo de Ação

Durante a realização da prova de conceito, uma vez que a solução é para ser estandardizada/uniforme/padronizada, e devido ao tempo disponível, definiram-se alguns erros mais comuns aos quais o algoritmo deverá responder numa fase inicial e mais tarde acrescentar todos os possíveis.

Atendendo à confidencialidade dos clientes não foi possível aceder a um ambiente real e foi então necessário popular um ambiente de teste com tabelas e dados de modo a que se enunciassem os erros pretendidos.

Através de um conjunto de reuniões com elementos das equipas de suporte, foi possível definir alguns erros e o modo como esses deveriam ser caracterizados e corrigidos.

A prova de conceito vai então receber como parâmetro de entrada os dados comparados que contêm erros apresentados pelos DIaaS, com as respetivas informações associadas a esses dados como os módulos a que pertencem as pks. Após a receção vai analisar e dar uma sugestão de resolução. A saída do algoritmo não será a correção do erro mas sim as ações necessárias para a sua correção, pois numa primeira fase enquanto o cliente não ganha confiança na solução, as ações não serão automáticas.

5.2 Tecnologias Utilizadas

As tecnologias utilizadas no desenvolvimento da solução são as mesmas que o DIaaS utiliza, uma vez que o algoritmo é integrado no DIaaS. Foi decidido que deveria ser implementado no

DIaaS pois este utiliza as tecnologias mais recentes e apropriadas e será a ferramenta a ser utilizada pelas equipas de suporte num futuro próximo.

5.2.1 Linguagens de Programação

As linguagens de programação utilizadas foram então as mesmas que o DIaaS utiliza e estão representadas na figura 5.1.

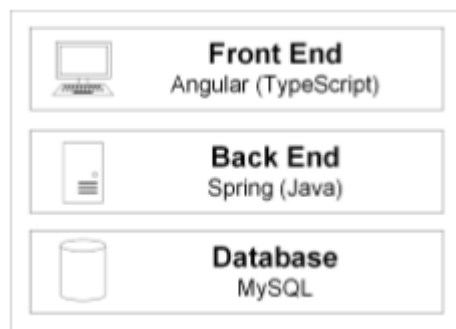


Figura 5.1: Linguagens de programação utilizadas

Na *Interface*, vulgarmente chamada de *front end*, foi usado o Angular que é uma *framework* que pode ser transversal a todas as plataformas e utiliza como linguagens o HTML, o CSS e o *TypeScript*.

No *back end* foi utilizado o java na *framework* Spring, esta *framework* torna o desenvolvimento de aplicações mais simples devido a todas as funções e bibliotecas que integra.

Quanto à base de dados, esta foi desenvolvida em *MySQL* que utiliza linguagem SQL e se destaca por ser estável e robusta.

5.2.2 Ambientes Utilizados

Um *Integrated Development Environment* (IDE), é uma ferramenta utilizada por programadores para o desenvolvimento de *software* pois além de ter uma *Interface* gráfica agradável contém um conjunto de funcionalidades que facilitam o desenvolvimento e teste de aplicações.

Foi então utilizada para o *back end* a ferramenta *Spring Tool Suite* que está adaptada para o desenvolvimento de aplicações em *Spring*(*framework web*, desenvolvida na linguagem de programação Java).

No *front end* foi utilizado o *Visual Studio Code* que é um IDE com um vasto leque de extensões que ajudam a testar código em diferentes plataformas.

Para a base de dados, foi utilizado o *MySQL Workbench* que é um software desenvolvido pela *Oracle* e que se destaca pela sua facilidade de uso, segurança e rápida recuperação de falhas e bloqueios.

5.3 Descrição Detalhada do Funcionamento do Algoritmo

O funcionamento do algoritmo passa por percorrer e preencher tabelas de modo a identificar o erro, mostrar as ações realizadas e dar uma sugestão de solução para o erro. Para uma melhor compreensão do funcionamento do algoritmo apresentamos o fluxograma do mesmo representado no Anexo C.

5.3.1 Tabelas

As tabelas criadas foram então as seguintes:

- **recon_solution_config**, esta está representada na figura 5.2 e possui todos os campos necessários para a identificação do erro, nomeadamente as colunas que são comparadas (column_a e column_b) a que sistemas pertencem (system_a_id e system_b_id), as *primary keys* que tem de ser iguais e garantem que se está a comparar os mesmos dados (pks_a e pks_b) e o tipo de erro (error_type). Além destes campos para a identificação do erro possui também outros dois que direcionam o algoritmo para outras tabelas, o rca_id e o solution_id, quando o erro é identificado.

recon_solution_config_id	column_a	column_b	error_type	pks_a	pks_b	rca_id	solution_id	system_a_id	system_b_id
1	value	value	VALUE DIF	RECEIVE DATE.LOCATION	RECEIVE DATE.LOCATION	1	NULL	2	3
2	NULL	value	RECORD NOT FOUND	NULL	RECEIVE DATE.LOCATION	NULL	2	2	3
3	value	NULL	RECORD NOT FOUND	RECEIVE DATE.LOCATION	NULL	NULL	3	2	3
4	SOH	SOH	VALUE DIF	ITEM	ITEM	2	NULL	2	4
5	value	value	VALUE DIF	RECEIVE DATE.LOCATION.ORDER NO	RECEIVE DATE.LOCATION.ORDER NO	3	NULL	2	3
6	value	value	VALUE DIF	RECEIVE DATE.LOCATION.ORDER NO.ITEM	RECEIVE DATE.LOCATION.ORDER NO.ITEM	NULL	4	2	3

Figura 5.2: Tabela de Identificação dos erros

- **recon_rca** esta está representada na figura 5.3 e contém as RCA ou a nova reconciliação a ser feita quando certo tipo de erro é identificado.

recon_rca_id	rca	rca_id	rule_validation	solution_id	extraction_id	recon_id
1	go to next level	1	>0	NULL	NULL	2
2	inv adi rms	2	>0	1	7	NULL
3	customer orders missina RMS	2	>0	1	8	NULL
4	customer orders missina WMS	2	>0	1	9	NULL
5	go to item level	3	>0	NULL	NULL	3

Figura 5.3: Tabela com as RCA e reconciliações

- **recon_solution_action** esta contém as ações a realizar para a correção dos erros, o sistema nas quais essas ações devem ser executadas e está representada na figura 5.4 apresentada na página 38.
- **recon_report_rca** esta é auto-preenchida sempre que o algoritmo realiza uma RCA de modo a informar o utilizador quais as RCA que foram corridas e qual o seu resultado, tal como pode ser visto na figura 5.5 apresentada na página 39.

action_id	solution_id	action	client_system_id
1	1	CREATE ADJUSTMENT	2
2	2	INSERT DATA ON RMS	2
3	3	INSERT DATA ON RIM	3
4	4	ADJUST OTY MATCHED(received)	3

Figura 5.4: Tabela das ações a aplicar

5.3.2 Identificação do Erro

O primeiro passo do algoritmo é a identificação do erro. Para uma melhor compreensão do processo temos a parte do fluxograma encarregue dessa identificação representada na figura 5.6 que se encontra na página 40. O erro é caracterizados pelas suas pks, sistemas comparados, o tipo de dados comparados e o tipo de erro apresentado.

Quando uma reconciliação é realizada, o algoritmo recebe o seu resultado que inclui os sistemas que foram comparados, as colunas comparadas as pks e o tipo de erro e compara esses dados com os campos da tabela recon_solution_config. Se esses campos coincidirem com alguma linha da tabela o erro é identificado. O erro também é identificado se coincidirem todos os campos com exceção das pks e colunas de apenas um sistema.

Quando o erro não é identificado, o utilizador vê uma mensagem a sugerir que adicione esse novo erro à tabela correspondente. Sendo identificado o erro, o algoritmo verifica se o mesmo tem uma solução ou uma RCA associada.

5.3.3 RCA's

Após a identificação do erro, numa primeira fase é verificado se esse erro tem uma RCA ou solução associada. Não tendo nenhuma das duas é apresentado um *print* ao utilizador informando-o de que aquele erro não tem solução nem RCA associada.

Se tiver associada uma RCA, vai então verificar se essa RCA tem associada uma reconciliação ou uma extração. A solução será explicada no subcapítulo seguinte. Como se pode verificar na figura 5.7, se estiver associada uma reconciliação, o algoritmo vai planeá-la e executá-la. Após a execução, o algoritmo vai então receber novas informações voltando ao seu início.

Tendo associada uma extração vai também planeá-la e executá-la verificando ao fim de cada execução o resultado da RCA, como pode ser visto na figura 5.8, na página 38. O resultado pode ser negativo ou positivo. Se for negativo não é aquela a causa, se for positivo é. Se o resultado for positivo, verifica a solução associada e após confirmar se todas as RCA associadas já foram executadas, se for falso passa diretamente a verificar se já foram todas executadas. Após executar as RCA, verifica sempre se já todas foram corridas. Se já foram todas corridas e nenhuma deu positiva, o algoritmo vai apresentar ao utilizador um *print* a informar que existe um problema nas RCA. Se uma ou mais deram positivo, o algoritmo apresenta as soluções associadas às RCA chegando assim ao seu fim.

recon_report_rca_id	column_values	pk_values	rca_id	rca_result	status	update_date	comp_error_id	recon_report_id	extraction_id	recon_solution_config_id
1	260	70281835	2	FALSE	DONE	2018-12-27 16:05:21	2	8	7	4
2	260	70281835	2	FALSE	DONE	2018-12-27 16:05:22	2	8	8	4
3	260	70281835	2	FALSE	DONE	2018-12-27 16:05:22	2	8	9	4
4	303	70281859	2	TRUE	DONE	2018-12-27 16:05:22	3	8	7	4
5	303	70281859	2	FALSE	DONE	2018-12-27 16:05:22	3	8	8	4
6	303	70281859	2	FALSE	DONE	2018-12-27 16:05:23	3	8	9	4

Figura 5.5: Tabela informativa das RCA efectuadas

Sempre que uma RCA é corrida a tabela recon_report_rca é preenchida informando o utilizador de todas as que foram corridas e o seu resultado.

5.3.4 Soluções

Durante o processo, sempre que o Algoritmo invoca uma solução são verificadas as ações associadas a ela e que se têm de realizar e apresentadas ao utilizador. Estas ações não são automáticas devido ao fatores já referidos anteriormente.

Podem verificar todas as situações em que o Algoritmo invoca uma solução no fluxograma representado no Anexo C. Estas situações ocorrem quando uma RCA dá resultado positivo e quando o erro tem logo uma solução associada levando sempre ao fim do processo.

5.4 Resultados Obtidos

A análise dos resultados obtidos a nível temporal não será totalmente viável pois o tempo do processo é apenas uma estimativa, visto que nunca foram retirados tempos pelas equipas de suporte relativos à duração das tarefas.

5.4.1 Dados de Teste

Para ser possível realizar testes, utilizou-se o ambiente de teste disponibilizado pela *Wipro* no qual foram necessárias criar algumas tabelas e popular outras de modo a ser possível realizar dois tipos de testes diferentes.

Num tipo, que daqui em diante vai ser designado por tipo1, perante o erro apresentado é necessário realizar novas reconciliações de modo a que se consiga descobrir a raiz do problema. E o outro tipo que vai ser designado de tipo2 em que perante o erro apresentado é necessário realizar RCA para descobrir a causa do problema.

No teste de tipo1, é comparado o número de artigos a nível do entreposto (ou seja, se o número de artigos para um determinado entreposto é igual nos dois módulos comparados) entre o RMS e o RIM onde se vão apresentar quantidades diferentes. Após verificar que a nível do entreposto estão diferentes, vai-se verificar a nível da ordem e após para as ordens que tiverem quantidades diferentes verifica-se a nível do item. Neste teste provocou-se que a quantidade fosse diferente para um artigo de modo a verificar se o algoritmo identifica qual é, realizando novas reconciliações de modo a ir mais ao detalhe do problema

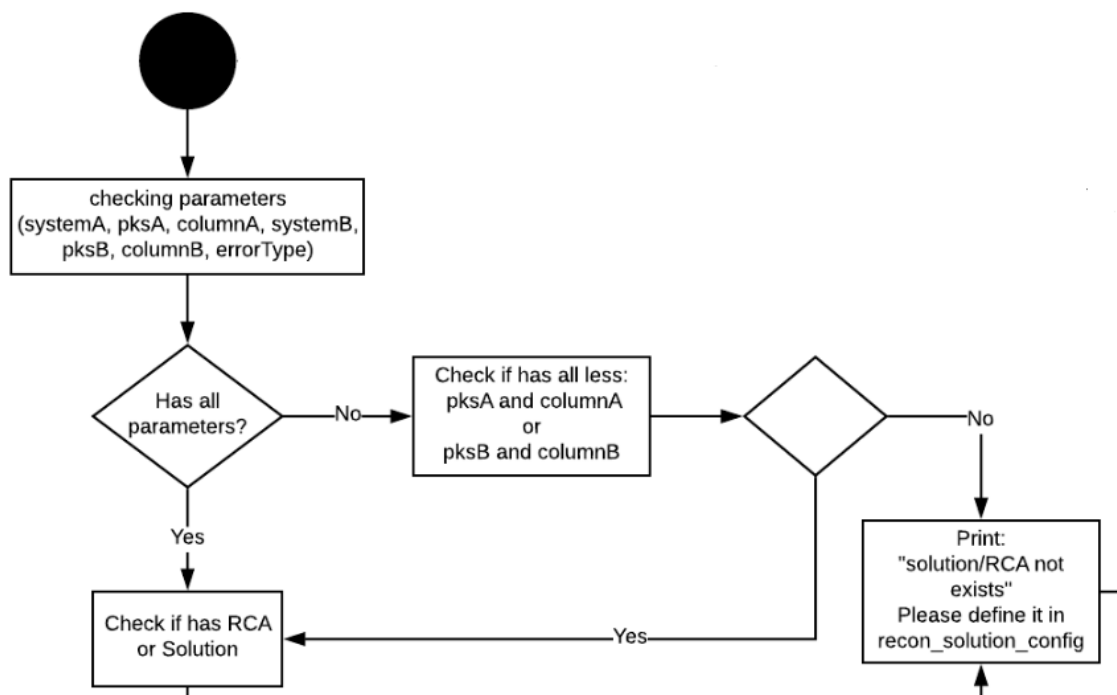


Figura 5.6: Processo de identificação do erro

No teste de tipo2, são comparados os níveis de stock entre o RMS e o WMS onde obtendo erros é necessário correr as RCA correspondentes para cada um de modo a descobrir a causa. Neste teste já se tem 649 erros para os quais são necessárias realizar três RCA.

5.4.2 Resultados Obtidos

No teste de tipo1, o algoritmo recebe o resultado da reconciliação em que foi comparado o número de artigos dos entrepostos entre o RMS e o RIM. Este apresenta erro para um dos entrepostos como se pode verificar na figura 5.9 onde também se apresenta a nova ação que o algoritmo vai realizar.

O algoritmo vai então executar uma nova reconciliação de modo a ir "mais a fundo" no problema e descobrir para que ordem ou ordens é que naquele entreposto se apresenta o erro, como se pode verificar na figura 5.10.

Após descobrir em que ordem se encontra o erro, o algoritmo vai então executar nova reconciliação para verificar em que artigos se encontra o erro e indicar ao utilizador a ação ou ações necessárias para a sua correção, como se pode visualizar na figura 5.11.

Como pode ser observado nas imagens anteriores, o *front end* apresenta os valores mas não as pks a que correspondem, essas podem ser vistas na tabela *comparison_error* representada na figura 5.12.

No teste de tipo2 é realizada uma reconciliação que retorna 14152 artigos com níveis de *stock* diferentes de onde se encontram 649 para os quais é necessário realizar RCA. O algoritmo vai

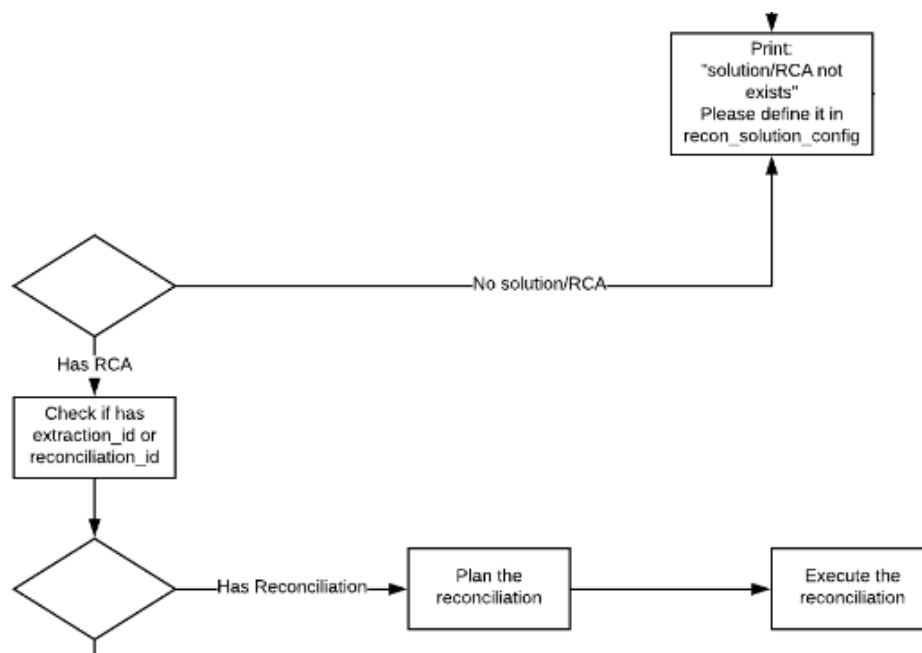


Figura 5.7: Processo de verificação se dentro das RCA tem extração ou reconciliação

então analisar cada erro individualmente sendo que para 13503 deles vai apresentar que não tem solução pois não existem num módulo e existem no outro e para os restantes 649 vai realizar para cada três RCA e apresentar o resultado obtido. Para cada RCA realizada, como pode ser visto na figura 5.5, na página 36, apresenta:

- o seu resultado que pode ser TRUE ou FALSE dependendo se retornou >0 ou <0 . Sendo TRUE foi essa a causa do problema. Se nenhuma der true o programa informa o utilizador de um problema nas RCA.
- o seu estado que pode ser DONE(feita), WAITING(a espera) ou FAILED(falhada).

No *front end*, o utilizador consegue visualizar por comparação se alguma das RCA deu verdadeira tal como pode ser visto na figura 5.13, da página 41, onde se verifica o resultado retornado das 3 RCA executadas para cada um dos 3 erros apresentados. No primeiro exemplo, temos um caso em que o algoritmo não identifica solução nem RCA associada ao erro e apresenta então ao utilizador a informação de que deve acrescentar na tabela. No segundo exemplo temos o resultado de 3 RCA realizadas para o erro em que todas retornaram falso, tendo portanto no final um print informando de que o problemas não foi encontrado nas RCA. No terceiro exemplo temos o resultado das 3 RCA realizadas em que uma deu verdadeira tendo a ação correspondente e as outras deram falsas.

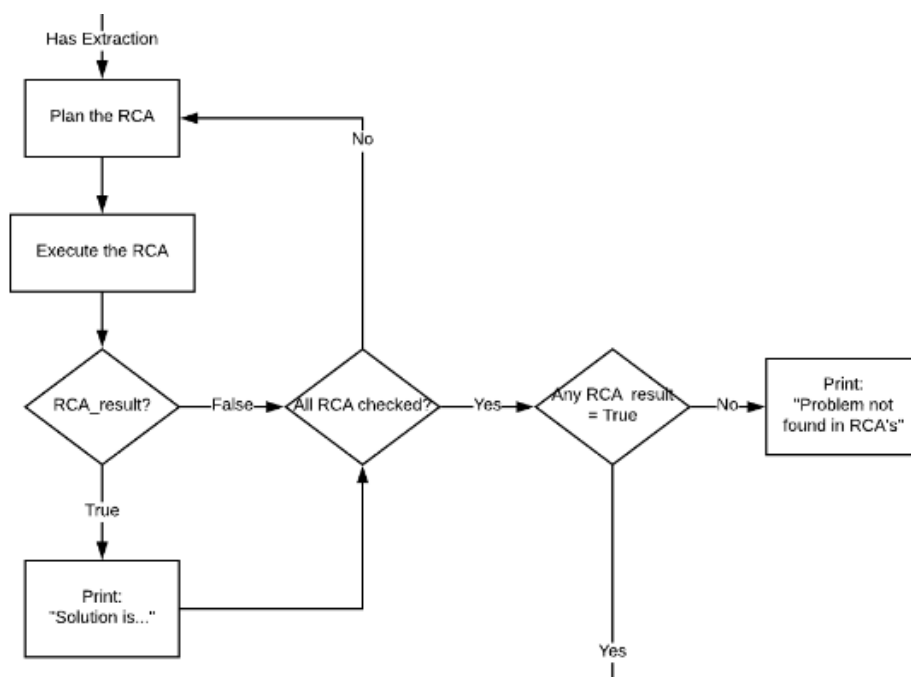


Figura 5.8: Processo de execução das RCA

5.4.3 Tempo de Execução

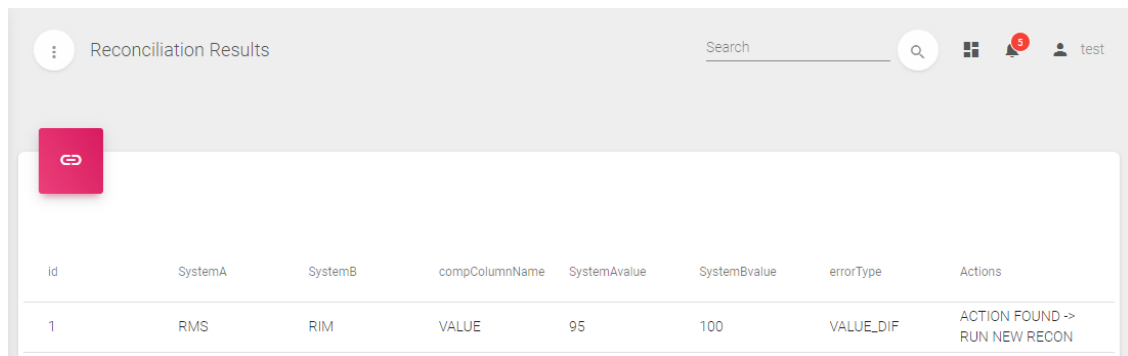
Para o teste de tipo1, o tempo de execução desde que recebe o resultado até descobrir o item e sugerir a ação a realizar, corresponde a cerca de 8 segundos. Não existe informação exata relativamente ao tempo despendido na anterior realização desta tarefa mas pode-se estimar que seria de cerca de 7 a 10 minutos pois o utilizador teria de receber o resultado da reconciliação, verificar qual o entreposto e fazer a segunda reconciliação inserindo o entreposto e o mesmo para a terceira de modo a descobrir qual o item, tudo isto efetuado manualmente.

Para o teste de tipo2, apenas se irá considerar o tempo despendido nos artigos para os quais são necessárias de realizar RCA. O algoritmo demora cerca de 13 minutos a fazer 1947 RCA, 3 para cada um dos 649 artigos. Esta tarefa se for realizada manualmente e contando que quem a efetua é como um "robô" e não perde tempo a pensar, demoraria cerca de 20 segundos para cada RCA o que daria um total de 649 minutos.

5.4.4 Comparação dos Resultados Obtidos com a Situação Atual

Comparar diretamente os resultados obtidos a nível temporal com a situação atual pode induzir a conclusões incorretas pois o tempo despendido atualmente varia consideravelmente com o tipo de erros, as suas causas, a equipa que o realiza e o número de erros com causas desconhecidas que necessitam de ser analisadas.

Pode-se afirmar então que o tempo gasto com erros que contêm causas já conhecidas passa a ser nulo, pois após inserido nas tabelas percorridas pelo algoritmo um erro e a sua correção,



id	SystemA	SystemB	compColumnName	SystemAvalue	SystemBvalue	errorType	Actions
1	RMS	RIM	VALUE	95	100	VALUE_DIF	ACTION FOUND -> RUN NEW RECON

Figura 5.9: Resultado da reconciliação a nível dos entrepostos

sempre que esse erro aparecer será automaticamente corrigido.

Com a implementação do algoritmo, o tempo despendido pela equipa de suporte seria exclusivamente dedicado à análise de novos erros de modo a acrescenta-los às tabelas para tornar a sua correção automática. Anulando assim completamente o tempo em tarefas repetitivas.

id	SystemA	SystemB	compColumnName	SystemAvalue	SystemBvalue	errorType	Actions
1	RMS	RIM	VALUE	45	50	VALUE_DIF	ACTION FOUND -> RUN NEW RECON

Figura 5.10: Resultado da reconciliação a nível das ordens

id	SystemA	SystemB	compColumnName	SystemAvalue	SystemBvalue	errorType	Actions
1	RMS	RIM	VALUE	45	50	VALUE_DIF	Action= ADJUST QTY_MATCHED(rece ived) in system= RIM

Figura 5.11: Resultado da reconciliação a nível do artigo

	comp_error_id	recon_report_id	comp_value	pk_values	truth_value	error_type_id	recon_extract_comp_id	source_extraction_data_id
1	1	1	100	2017-11-15. 2	95	2	1	2
1	2	2	50	2017-11-15. 2. 130072	45	2	2	4
1	3	3	50	2017-11-15. 2. 130072. 100065002	45	2	3	6

Figura 5.12: Tabela comparison_error relativa as reconciliações acima apresentadas

id	SystemA	SystemB	compColumnName	SystemAvalue	SystemBvalue	errorType	Actions
1	RMS	WMS				RECORD_NOT_FOUN D	Solution/RCA not exists, please define it in recon_solution_con fig
2	RMS	WMS	SOH	637	260	VALUE_DIF	<u>Problem not in this RCA Problem not in this RCA Problem not in this RCA</u> Problem not found in RCA's
3	RMS	WMS	SOH	1159	303	VALUE_DIF	Action= CREATE ADJUSTMENT in system= RMS <u>Problem not in this RCA Problem not in this RCA</u>

Figura 5.13: Resultado de alguns erros para os quais se realizaram RCA

Capítulo 6

Conclusões e Trabalho Futuro

Neste capítulo serão apresentadas as conclusões relativas a todo o trabalho efetuado nesta dissertação, a satisfação dos objetivos propostos e o trabalho a realizar futuramente.

6.1 Conclusões

O principal objetivo desta dissertação era o desenvolvimento de uma abordagem automatizada da forma como são executados atualmente alguns processos de suporte realizados pela *Wipro*. Esta abordagem parte da utilização de um algoritmo acrescentado ao programa DIaaS.

Através de uma análise detalhada dos processos de suporte efetuados pela *Wipro* foi possível identificar os seus pontos críticos e as respetivas oportunidades de melhoria. Foi então desenvolvida uma nova abordagem que explora todas as oportunidades de melhoria encontradas, cumprindo os requisitos previamente discutidos e definidos.

A solução proposta faz um levantamento dos requisitos do algoritmo de modo a ser possível a automatização das tarefas e a identificação dos erros de forma standardizada para que seja viável a sua utilização por todas as equipas de suporte.

A prova de conceito desenvolvida permitiu comprovar que esta nova abordagem desenvolvida é uma mais valia pois automatiza o processo e permite também verificar que os objetivos do projeto foram cumpridos.

O objetivo central do estudo não é de substituir as equipas de suporte, mas sim de retirar às mesmas as tarefas repetitivas que não acrescentam valor, garantindo que estes elementos podem dedicar o seu tempo à análise de novos erros. O objetivo é então que estas equipas possam inserir o conhecimento obtido nas análises na solução, evitando que realizem a mesma tarefa mais do que uma vez. Por outro lado, permite uma substantiva redução do tempo atualmente gasto na realização destas tarefas.

A implementação da solução proposta na empresa apenas trará benefícios para a organização. A mesma tarefa passará a ser realizada com maior assertividade, em menos tempo e com custos menores. Num mercado tão competitivo como é o do Retalho, é essencial tirar proveito das tecnologias emergentes para estar sempre um passo à frente da concorrência.

6.2 Trabalho Futuro

Considerando o ponto em que ficou esta dissertação e de modo a dar continuidade ao trabalho desenvolvido existem alguns aspetos que podem ser aprofundados e trabalhados e que serão, de seguida, enunciados.

6.2.1 Implementação em Ambiente Real

A solução não chegou a ser testada em ambiente real devido às questões de confidencialidade dos clientes. É então importante garantir esse procedimento em ambiente real de modo a avaliar devidamente os benefícios que esta mobilizará para as organizações. Seria igualmente importante realizar uma análise temporal e financeira do antes e depois relativamente ao processo de suporte realizado.

6.2.2 Correção automática

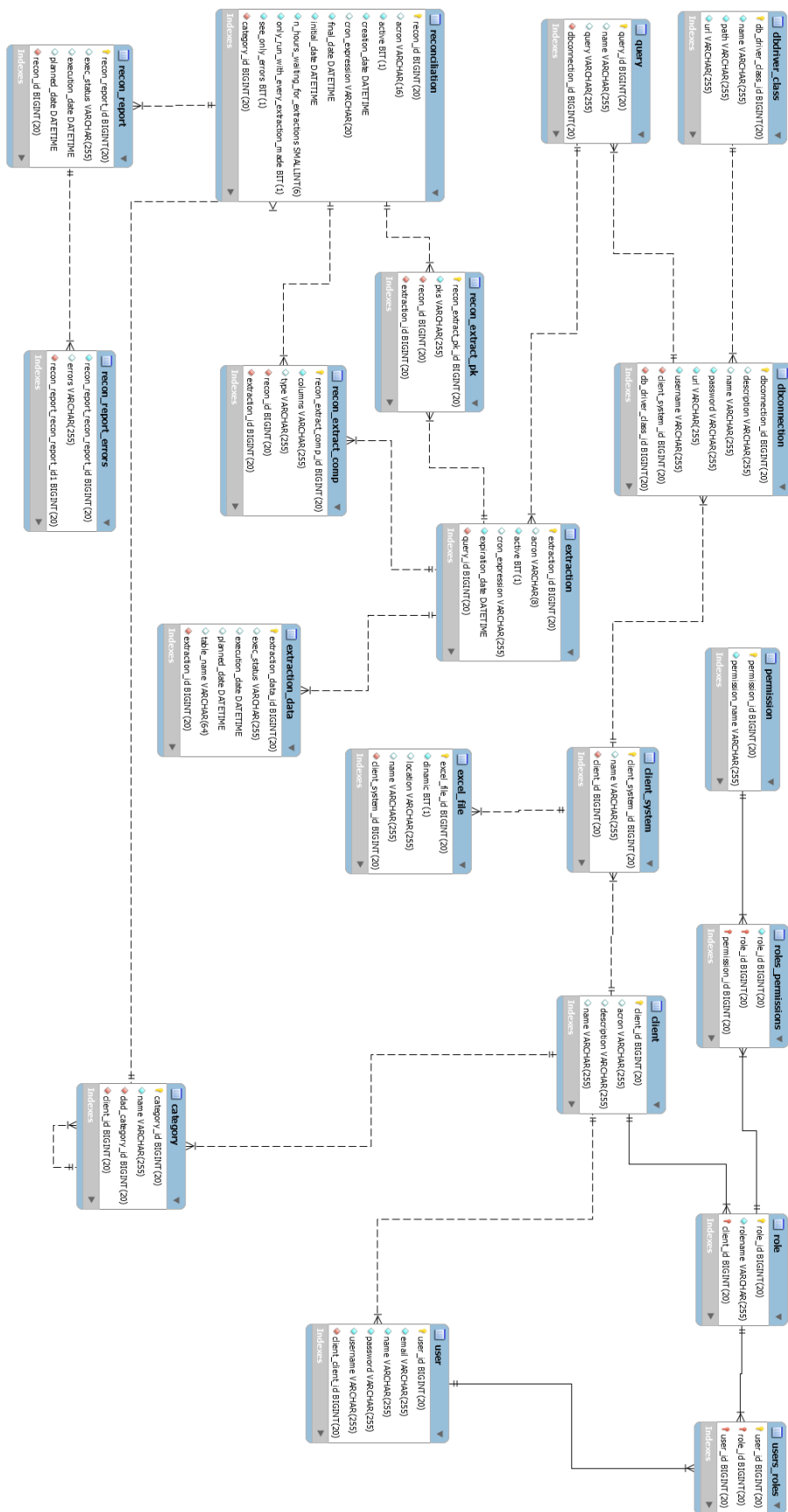
Após ser testada em ambiente real, colocar as ações a serem realizadas automaticamente será uma mais valia significativa para a organização. Assim, garantir-se-á que todos os procedimentos serão totalmente automatizados.

6.2.3 Inteligência Artificial

Este trabalho não se dá aqui por concluído uma vez que a solução aqui proposta carece ainda da introdução de inteligência artificial para que atinja o desiderato que desejamos. Fica portanto em aberto, a elaboração de um trabalho futuro que, no complemento deste, encerre o objetivo a que nos propusemos. Assim, dever-se-á procurar no futuro uma solução no âmbito da mencionada inteligência artificial que permita identificar e corrigir novos erros, tornando a aplicação totalmente automatizada.

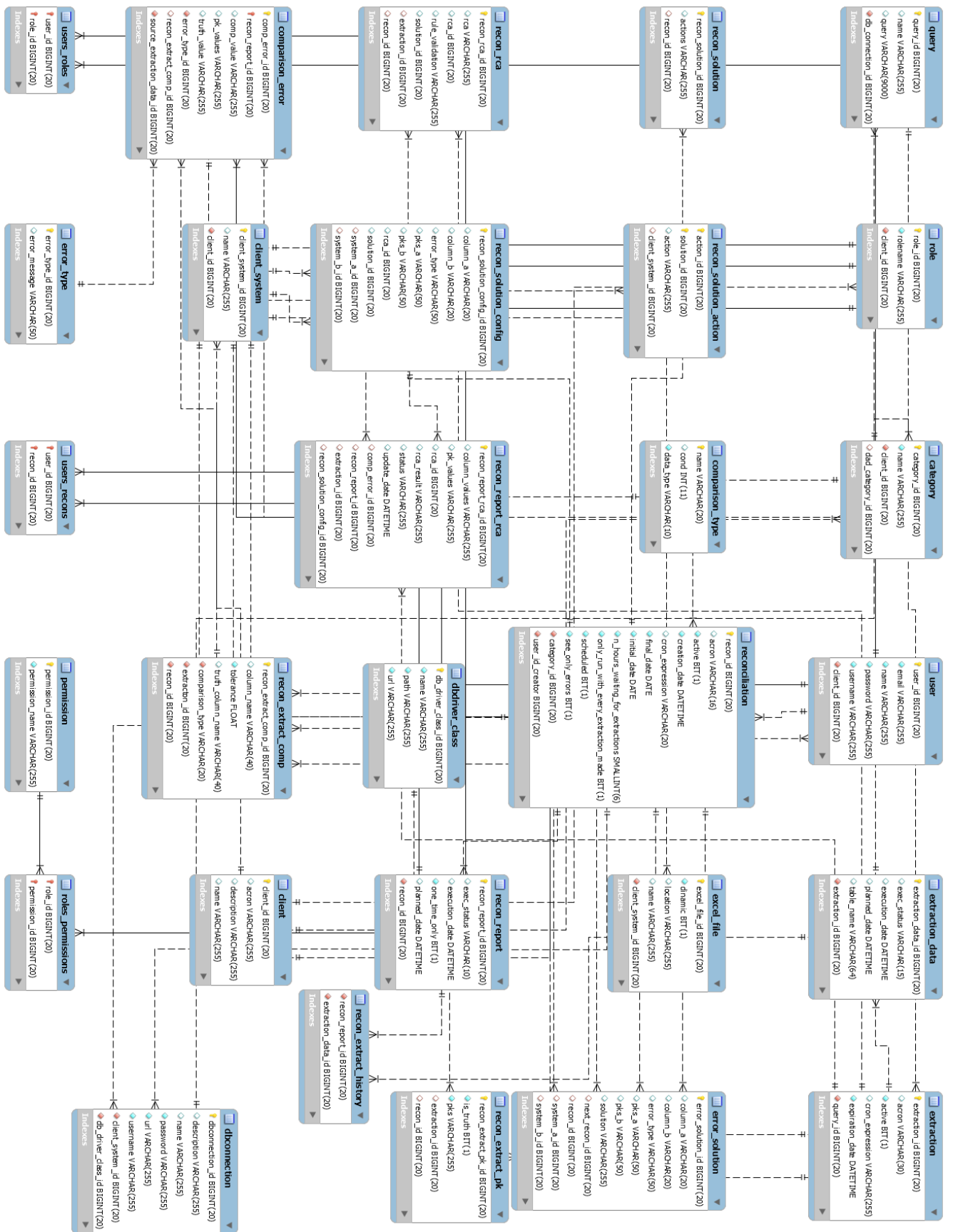
Anexo A

Modelo Relacional Antigo



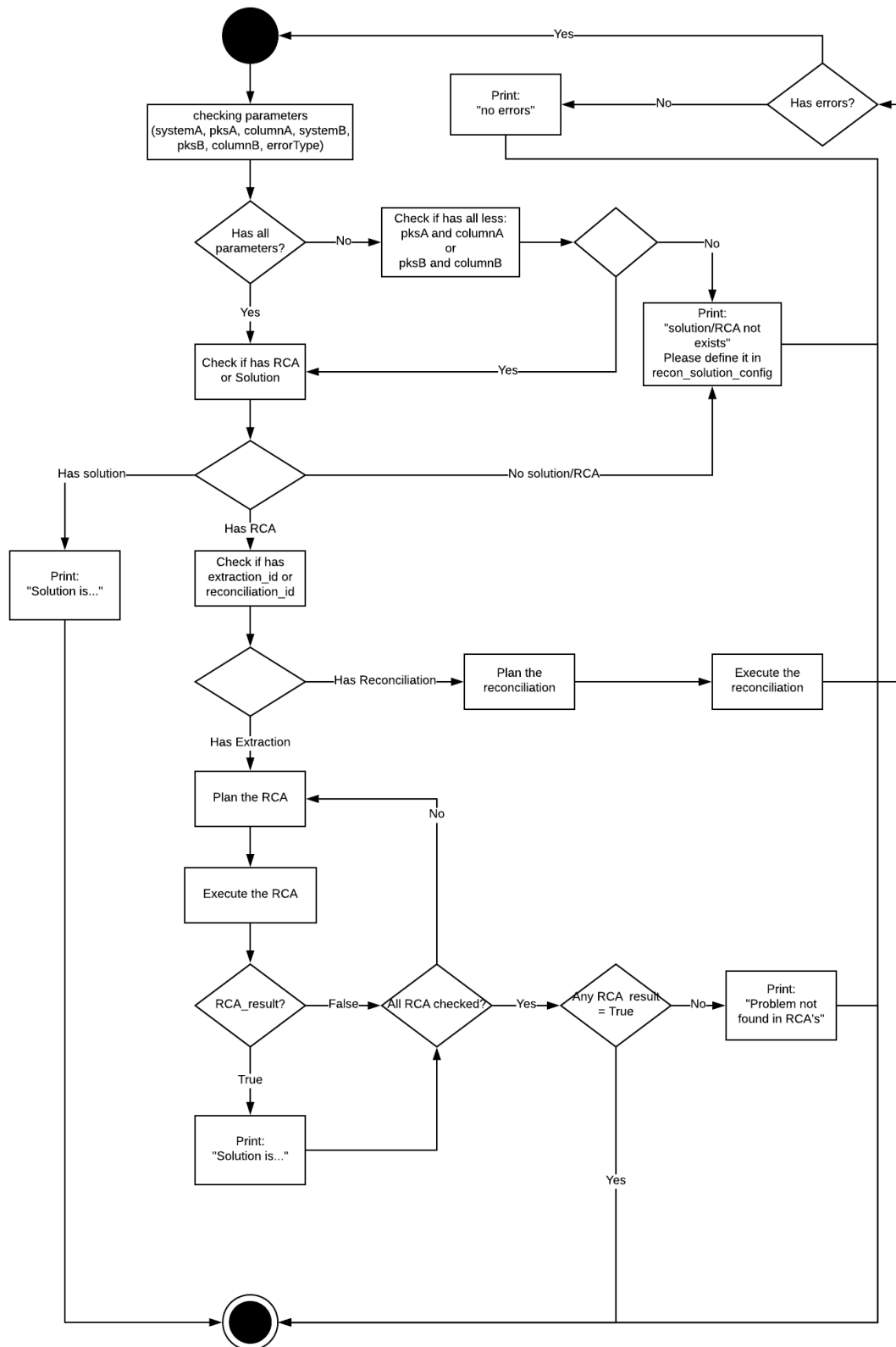
Anexo B

Modelo Relacional Novo



Anexo C

Fluxograma do Algoritmo



Referências

- [1] Wipro Named by Ethisphere Institute as One of the World's Most Ethical Companies. URL: <https://www.wipro.com/newsroom/press-releases/archives/wipro-named-by-ethisphere-institute-as-one-of-the-worlds-most-et/> [último acesso em 01-10-2018].
- [2] About Us - Wipro. URL: <https://www.wipro.com/about-us/> [último acesso em 01-10-2018].
- [3] Enabler Wipro – Wikipédia, a enciclopédia livre. URL: https://pt.wikipedia.org/wiki/Enabler_Wipro [último acesso em 02-10-2018].
- [4] Revismarket/Electro. URL: <http://rmelectro.com/cgi/detn?id=9072> [último acesso em 03-10-2018].
- [5] Preveja tendências de mercado antes da concorrência com Big Data. URL: <http://www.bigdatabusiness.com.br/preveja-tendencias-de-mercado-antes-da-concorrencia-por-meio-de-big-data/> [último acesso em 05-10-2018].
- [6] The Retail Data Overload – Sifting Through the Clutter - Wipro. URL: <https://www.wipro.com/analytics/the-retail-data-overload/> [último acesso em 05-10-2018].
- [7] Definition, Types, and Examples of Retailing. URL: <https://www.thebalancesmb.com/what-is-retail-2892238> [último acesso em 08-10-2018].
- [8] What Is Retail? URL: <https://erply.com/what-is-retail/> [último acesso em 09-10-2018].
- [9] Retail – Definition and Meaning | relivingmbadays. URL: <https://relivingmbadays.wordpress.com/2012/10/16/retail-definition-and-meaning/> [último acesso em 11-10-2018].
- [10] Decisão orientada para os dados é fundamental para o retalho. URL: <https://www.distribuicao hoje.com/insights/decisao-orientada-para-os-dados-e-fundamental-para-o-retalho/> [último acesso em 15-10-2018].
- [11] A revolução dos dados no retalho - Business-IT. URL: <https://business-it.pt/2018/03/03/revolucao-dos-dados-no-retalho/> [último acesso em 15-10-2018].

- [12] Method and apparatus for profiling clients of a database-driven transaction system and marketing according to profile data. 12 2000. URL: <https://patents.google.com/patent/US20020095328A1/en>.
- [13] O grande volume de dados no retalho - ShopperTrak. URL: <https://pt.shoppertrak.com/article/o-grande-volume-de-dados-no-retalho/> [último acesso em 18-10-2018].
- [14] What is Information Technology (IT)? - Definition from WhatIs.com. URL: <https://searchdatacenter.techtarget.com/definition/IT> [último acesso em 19-10-2018].
- [15] IT (Information Technology) Definition. URL: <https://techterms.com/definition/it> [último acesso em 19-10-2018].
- [16] Saroja S. Information Technology - Key Success Factor in Retail. URL: <https://www.gjimt.ac.in/web/wp-content/uploads/2017/10/N22.pdf> [último acesso em 22-10-2018].
- [17] Christiane Amanda Lima Alves. A IMPORTÂNCIA DA TECNOLOGIA DA INFORMAÇÃO NAS EMPRESAS. Relatório técnico. URL: <https://semanaacademica.org.br/system/files/artigos/artigosistemadeinformacaonasesempresas.pdf> [último acesso em 22-10-2018].
- [18] Ana Lucas, Cristiane Pedron, Fernando Naves, José Camacho, Luis Vaz Henriques, e Filipa Pires da Silva. Conceitos fundamentais de sistemas e tecnologias de informação e de gestão do conhecimento. Cadeira de Tecnologias de Informação. Relatório técnico, 2008. URL: <https://www.iseg.ulisboa.pt/aquila/getFile.do?fileId=8003&method=getFile> [último acesso em 23-10-2018].
- [19] Haibin Zhang. General Framework of Chain Store Information System base on Supply Chain Management Theory. Relatório técnico 4, 2008. URL: <https://pdfs.semanticscholar.org/b67f/be4c3815a1b97ac8d9d7b22ac3887a1ed348.pdf>.
- [20] Prachi Juneja. Information System in Retail Sector. URL: <https://www.managementstudyguide.com/information-system-in-retail-sector.htm> [último acesso em 26-10-2018].
- [21] What is Oracle? - Computer Business Review. URL: <https://www.cbronline.com/what-is/what-is-oracle-4931117/> [último acesso em 26-10-2018].
- [22] Oracle Retail - 3CON. URL: <http://www.trescon.com.br/oracle-retail/> [último acesso em 29-10-2018].
- [23] What is Robotic Process Automation? - RPA Software | UiPath®. URL: <https://www.uipath.com/rpa/robotic-process-automation> [último acesso em 30-10-2018].
- [24] Clint Boulton. What is RPA? A revolution in business process automation | CIO. 2018. URL: <https://www.cio.com/article/3236451/business-process-management/what-is-rpa-robotic-process-automation-explained.html>.
- [25] Software As A Service (SaaS). URL: <https://www.investopedia.com/terms/s/software-as-a-service-saas.asp> [último acesso em 02-11-2018].

- [26] Margaret Rouse. What is Software as a Service (SaaS)? - Definition from WhatIs.com. URL: <https://searchcloudcomputing.techtarget.com/definition/Software-as-a-Service> [último acesso em 05-11-2018].
- [27] Daniel Fagella. What is Machine Learning? - An Informed Definition. 2018. URL: <https://www.techemergence.com/what-is-machine-learning/> [último acesso em 05-11-2018].
- [28] Machine Learning. URL: https://www.sas.com/en_us/insights/analytics/machine-learning.html [último acesso em 05-11-2018].
- [29] What is Data Integrity in Databases? - Definition from Techopedia. URL: <https://www.techopedia.com/definition/811/data-integrity-databases> [último acesso em 05-11-2018].
- [30] Henning Lund. Data integrity: what is it and why is it important?, 2 2017. URL: <https://www.rapidionline.com/blog/data-integrity-what-and-why> [último acesso em 07-11-2018].
- [31] Marvin Magusara. The Importance of Data Integrity for Your Business - TweakYourBiz, 2018. URL: <https://tweakyourbiz.com/technology/the-importance-of-data-integrity-for-your-business> [último acesso em 08-11-2018].
- [32] Sasha Chavkin e Bart Mos Michael Hudson. *Big 4 Audit Firms Play Big Role in Offshore Murk*. The International Consortium of Investigative Journalists, 11 2014. URL: <http://www.icij.org/project/luxembourg-leaks/big-4-audit-firms-play-big-role-offshore-murk> [último acesso em 12-11-2018].
- [33] Cindy Ng. What is Data Integrity and How Can You Maintain it?, 2018. URL: <https://blog.varonis.com/data-integrity/> [último acesso em 12-11-2018].