INSTITUTO
SUPERIOR
DE CONTABILIDADE
E ADMINISTRAÇÃO
DO PORTO
POLITÉCNICO
DO PORTO

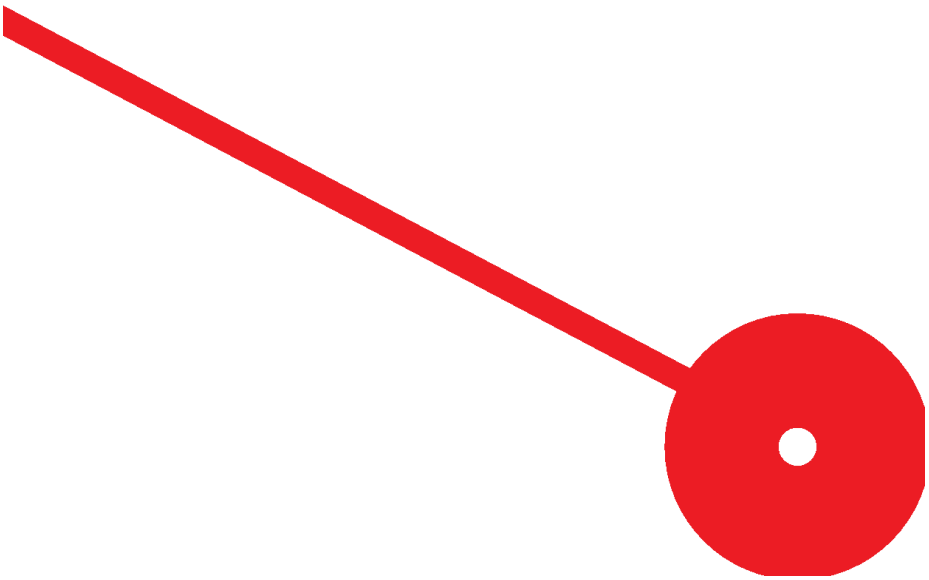
M MESTRADO
NEGÓCIO ELECTRÓNICO

Estudo de Caso - Migração de Sistemas de Gestão de Conteúdo numa empresa Multinacional

Daniela Alexandra Coelho Silva

10/2022

Esta versão contém as críticas e sugestões dos elementos do júri



Daniela Alexandra Coelho Silva. Estudo de Caso: Migração de Sistemas de Gestão de Conteúdo numa Empresa Multinacional 10/2022

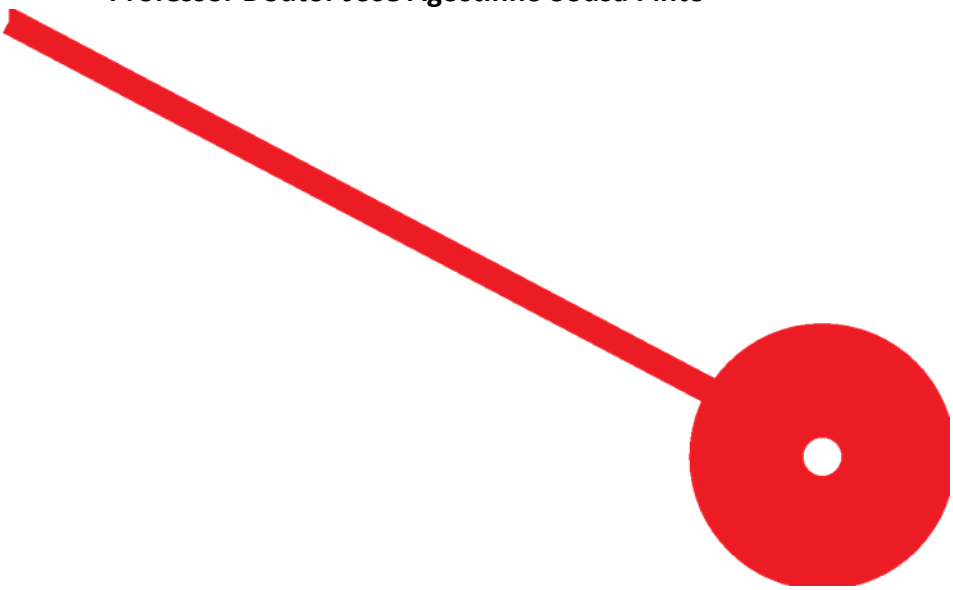
INSTITUTO
SUPERIOR
DE CONTABILIDADE
E ADMINISTRAÇÃO
DO PORTO
POLITÉCNICO
DO PORTO

M MESTRADO
NEGÓCIO ELETRÓNICO

Estudo de Caso - Migração de Sistemas de Gestão de Conteúdo numa empresa Multinacional

Daniela Alexandra Coelho Silva

Dissertação de Mestrado apresentado ao Instituto Superior de Contabilidade e Administração do Porto para a obtenção do grau de Mestre em Negócio Eletrónico sob orientação de Professor Doutor Rui Humberto Ribeiro Pereira e Professor Doutor José Agostinho Sousa Pinto



Daniela Alexandra Coelho Silva. Estudo de Caso: Migração de Sistemas de Gestão de Conteúdo numa Empresa Multinacional
10/2022

Agradecimentos

Em primeiro lugar, gostaria de agradecer à minha família, pelo apoio e encorajamento durante todo o projeto.

Em segundo lugar, agradecer aos meus orientadores, Professor Doutor Agostinho Sousa Pinto e Professor Doutor Rui Humberto Ribeiro Pereira, por todo o apoio prestado e por colaborar neste estudo.

De seguida, gostaria de agradecer à empresa que se mostrou disponível em colaborar neste estudo, em especial ao André Pereira pelo tempo disponibilizado.

À Letícia Azevedo e à Catarina Vieira, que me apoiaram e incentivaram durante todo o projeto.

Resumo:

Com a contínua evolução dos sistemas e softwares de gestão de conteúdo, as organizações empresariais são obrigadas a trocar os seus sistemas obsoletos ou no final do seu ciclo de vida e atualizar por novas aplicações. Com esta constante necessidade de atualização é necessário realizar a migração do conteúdo e dados desses sistemas. Este estudo aborda a temática da migração de softwares de gestão de conteúdo e plataformas de e-commerce numa empresa multinacional. De modo a contextualizar o tópico, é realizada uma primeira abordagem com a fundamentação teórica que introduz os temas, o que são plataformas de e-commerce, o que são sistemas de gestão de conteúdo e como eles podem dar suporte às vendas online. Como a migração de dados é realizada e quais os diferentes métodos que podem ser adotados para os processos de migração. Como resultado do estudo apresentam-se os sistemas de origem e a nova aplicação, bem como a técnica utilizada pela organização para o seu processo de migração e realiza-se uma descrição do processo em termos temporais e tecnológicos. Apontam-se quais as limitações encontradas no estudo e os pontos positivos.

Palavras-chave: e-commerce; migração de dados; sistemas de gestão de conteúdo, CMS.

Abstract:

With the continuous evolution of content management systems (CMS), the organizations are forced to replace obsolete or end-of-life systems and upgrade to new ones. With this constant need to upgrade, it is necessary to migrate the content and data of these systems. This study addresses the topic of migration of CMS software and e-commerce platforms in a multinational company. To put the topic in context, the first approach made is the theoretical basis that introduces the topics, what are e-commerce platforms, what are content management systems and how they can support online sales. How data migration is performed and what are the different methods that can be adopted for a migration process. As a result of the study, the source systems and the new application are presented, as well as the technique used by the organization for its migration process and a description in temporal and technological terms. The limitations found in the study and the positive points are pointed out.

Key words: e-commerce; data migration; content management systems, CMS.

Índice geral

Capítulo - Introdução 1

Contexto 2

Problema a Resolver 3

Questões de investigação 4

Metodologia 4

Estrutura do trabalho 5

Capítulo I – Fundamentação Teórica 7

1.1 e-commerce 8

1.2 Diferentes modalidades de e-commerce 12

1.3 Plataformas de e-commerce 14

1.4 Content Management System (CMS) 16

1.5 Migração 18

1.5.1 Migração de dados 18

1.5.2 Estratégias de migração de dados 19

1.5.3 ETL como processo de migração de dados 20

1.5.4 Ferramentas para migração de dados 27

1.6 Trabalhos Relacionados 29

1.6.1 Migração de dados em Sistemas de Gestão de Conteúdo Web 29

1.6.2 Desenvolvimento de uma ferramenta de migração de banco de dados 30

1.6.3 Migração de bases de dados relacionais para NoSQL - Métodos de Análise 30

1.6.4 Metodologia de migração de dados em contexto de migração de sistemas .. 31

1.6.5 Processos para uma migração de dados 32

Capítulo II – Processo de migração 34

2.1 Sistemas origem 35

TYPO3 35

Sitecore 37

Joomla 39

SAP Web Channel 41

2.2 Sistema destino dos dados - Adobe Experience Manager (AEM) 43

2.3 Estratégia adotada para migração 45

2.4 Processo de Migração 52

2.4.1 Início 53

2.4.2 Instalação e configuração 55

2.4.3 Extração 56

2.4.4 Análise 59

2.4.5 Seleção e Mapeamento 61

2.4.6 Desenho da Migração de Conteúdo 62

2.4.7 Transformação 63

2.4.8 Teste de importação 64

2.4.9 Importação para produção 65

2.4.10. Encerramento 66

Capítulo III – Discussão de Resultados 67

4.1 Análise do processo de migração 70

Capítulo IV – Conclusões 75

4.1 Considerações finais 76

4.2 Limitações 77

4.3 Recomendações para futuras investigações 77

Referências bibliográficas 79

Apêndices 84

Índice de Figuras

Figura 1: Proporção de pessoas dos 16 aos 74 anos que utilizaram comércio eletrónico nos últimos 12 meses anteriores à entrevista 10
Figura 2: Número de compradores digitais em todo mundo de 2014 até 2021 (em biliões) 11
Figura 3: Vendas de comércio eletrónico no retalho a nível mundial de 2014 a 2021, com previsões de 2022 a 2025 12
Figura 4: Arquitetura do Método ETL 22
Figura 5: Interface de uma ferramenta de processamento em lotes 24
Figura 6: Interface de uma ferramenta com base em código 24
Figura 7: Interface de uma ferramenta com base em nuvem 24
Figura 8: Interface de uma ferramenta com base em Open Source 25
Figura 9: Interface de uma ferramenta com base GUI 25
Figura 10: Interface de uma ferramenta com base em tempo real 26
Figura 11: Interface de uma ferramenta com base NoSQL 26
Figura 12: Diagrama do processo ETL adaptado ao projeto do Estudo de Caso 46
Figura 13: Ciclo da migração de dados 51
Figura 14: Cronograma da primeira etapa – Início 54
Figura 15: Script exemplo para configuração no sistema Sitecore 55
Figura 16: Cronograma da segunda etapa – Instalação e Configuração 55
Figura 17: Script exemplo utilizado na extração de dados 56
Figura 18: Extração realizada ao sistema TYPO3 57
Figura 19: Extração realizada ao sistema Sitecore 58
Figura 20: Extração realizada ao sistema Joomla 58
Figura 21: Cronograma da terceira etapa – Extração 58
Figura 22: Interface de Análise providenciado pela empresa contratada 60
Figura 23: Insights com diferentes setores de análise 60
Figura 24: Cronograma da quarta etapa – Análise 60
Figura 25: Mapa de componentes 61
Figura 26: Cronograma da quinta etapa – Seleção e mapeamento 62
Figura 27: Cronograma da sexta etapa – Desenho da migração de conteúdo 63
Figura 28: Cronograma da sétima etapa – Transformação 64
Figura 29: Cronograma da oitava etapa – Teste de importação 65

Figura 30: Cronograma da nona etapa – Importação para produção 65
Figura 31: Cronograma da décima etapa – Encerramento 66

Índice de Tabelas

Tabela 1 - Desvantagens e Vantagens do Sistema TYPO3 37
Tabela 2 - Desvantagens e Vantagens do Sistema Sitecore 39
Tabela 3 - Desvantagens Vantagens do Sistema Joomla..... 41
Tabela 4 - Desvantagens e Vantagens do Sistema SAP Web Channel 42
Tabela 5 - Desvantagens e Vantagens do Sistema AEM..... 45

Lista de abreviaturas

AEM – Adobe Experience Manager

B2B – Business to Business

B2C – Business to Consumer

B2G – Business to Government

C2C – Consumer to Consumer

CMS - Content Management System

CRM - Customer Relationship Management

DMS - Document Management System

DNA - Deoxyribonucleic Acid

DW - Data Warehouse

E2E – End to End

ECM - Enterprise content management

ERP - Planeamento de Recursos Empresarias

ETL - Extract, Transform and Load

EUA – Estados Unidos da América

GUI – Graphical User Interface

HTML - Hypertext Markup Language

PHP - Hypertext Preprocessor

SDLC - Software Development Life Cycle

SEO - Search Engine Optimization

SGBD - Sistema de Gestão de Banco de Dados

SQL - Structured Query Language

UX – User Experience

WCEM - Web Channel Experience Manager

WCMS - Web Content Management Systems

WYSIWYG - What You See Is What You Get

WWW – World Wide Web

XML - Extensible Markup Language

CAPÍTULO - INTRODUÇÃO

Neste capítulo é apresentado o contexto do problema, uma breve descrição, identificação das questões de investigação, qual a metodologia adotada para aplicar na investigação e a estrutura do trabalho.

Contexto

A internet começou a ganhar popularidade para o público em geral apenas em 1994. Por volta dos anos 2000 algumas empresas na sua maioria residentes nos Estados Unidos e na Europa Ocidental começaram a promover os seus serviços através da World Wide Web. A internet sofreu uma evolução significativa nas últimas décadas tornando-se, progressivamente parte da vida diária da nossa sociedade (Mohapatra, 2013).

A internet foi o impulso para a origem do comércio eletrónico como uma das mais importantes conquistas científicas. A tecnologia entranhada no e-commerce (comércio eletrónico) dá origem a uma revolução na economia circular. O e-commerce permitiu que o comércio tradicional obtivesse resultados económicos de forma mais rápida, maiores e em geral melhores. Estes resultados derivam do facto de que as tecnologias associadas ao e-commerce quebram os limites de tempo e espaço, o padrão de comércio altera-se, assim como a circulação de mercadoria se intensifica (Qin et al., 2014).

Segundo Qin et al. (2014), o comércio eletrónico é uma revolução industrial e a mais significativa desde a Revolução industrial em 1760, para além de ser a revolução com mais influência na humanidade que as duas revoluções industriais anteriores. O e-commerce tem uma mudança significativa devido à produtividade ter um aumento considerável, as operações económicas tornaram-se mais eficientes e foi possível diminuir os custos, toda a revolução é acompanhada de evolução e com isso muitos factos que originalmente seriam impossíveis atualmente são realizáveis, além destas alterações o e-commerce também influencia os estilos de vida das pessoas em todos os aspetos sociais.

Com todas estas alterações as empresas começaram a perceber que a internet pode trazer valor ao seu negócio e compreenderam que para além de ser uma grande oportunidade é uma forma rápida de alcançar novos clientes com um simples investimento. A sociedade começou a confiar cada vez mais no e-commerce e desta forma alterar os seus hábitos de compra do offline para o online. Em 2021, os negócios de e-commerce atingiram o valor de 4,9 triliões de dólares em todo o mundo e a

previsão para os seguintes anos é de evolução tendo como expectativa no ano 2025 atingir o valor de 7.4 triliões de dólares (Statista, 2022)[1] .

No entanto, existem empresas que abriram os seus negócios online e com o passar dos anos as plataformas que usam deixam de conseguir responder às necessidades dos utilizadores e da empresa. Os sistemas começam a ficar obsoletos e com isso as empresas procuram novas soluções (Shoaib et al., 2017).

Problema a Resolver

Com a evolução da internet e com o aumento de consumidores, os websites necessitam de alguns ajustes, precisam de mais funcionalidades, aumentar a capacidade para armazenar conteúdo, serem mais acessíveis e de fácil navegação para os seus utilizadores. Com o crescimento de utilizadores e evolução das tecnologias estes procuram sempre mais funcionalidades, um design mais apelativo e uma experiência rápida. As novas plataformas precisam de ser rápidas o suficiente de forma a corresponder às expectativas de todos os utilizadores, independentemente das suas condições de acesso e ao mesmo tempo entregar-lhes uma experiência única e agradável.

O problema surge para empresas que adquiriram os seus sistemas de CMS e plataformas de e-commerce há vários anos, estas apercebem-se que as tecnologias que utilizam já não correspondem ao que os utilizadores procuram por diversas razões. Muitos destes sistemas estão a chegar ao fim do seu ciclo de vida, para estas empresas conseguirem acompanhar a evolução de tecnologias muitas vezes tem de procurar por atualizações, mas que por vezes só essas alterações não são suficientes. Começam a procurar por novos sistemas que vão responder às suas expectativas e às expectativas dos utilizadores, implicando que as empresas comecem a efetuar migração dos seus sistemas para os mais recentes. O que pode ser reconhecido como uma limitação pois a maioria das empresas não tem equipas com conhecimentos técnicos sobre as plataformas e projetos de migração.

[1] Chevalier, S. (2021, Outubro 28) Global Retail E-Commerce Sales from 2014 to 2026)https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/

Para desenvolver um processo de migração de sistemas não é simples, e muitas empresas encontram dificuldades nestes processos por diversas razões, regra geral técnicas.

A investigação para o presente estudo de caso recaí na necessidade de uma organização, com atividade global, executar um processo de migração de dados dos seus sistemas de origem, os sistemas descontinuados, para um único sistema destino. No ano de 2019, esta organização adquiriu uma empresa dinamarquesa de grande escala resultando na fusão das duas, desde então a organização começou a desenvolver um projeto para a atualização das suas tecnologias. Durante estes três anos, a organização conta com três websites diferentes (dois sistemas da organização inicial mais um sistema da empresa dinamarquesa) e uma loja online, os respetivos CMS's serão identificados posteriormente. Estes sistemas continuam a ser mantidos para que não percam a sua posição nos motores de busca. A organização compreendeu que existia uma necessidade de unir, alinhar a sua imagem e atividade digital através de uma ferramenta mais desenvolvida. Com a intenção de expandir a plataforma online, de forma a promover a sua identidade, os seus valores e a sua essência.

Questões de investigação

O objetivo principal desta dissertação foi desenvolver uma investigação sobre o processo de migração de sistemas de gestão de conteúdo em uma empresa multinacional. Para ajudar a atingir o objetivo principal do estudo, foram delineados objetivos específicos:

- Identificar características dos sistemas de origem e sistema destino;
- Definir vantagens e desvantagens dos sistemas;
- Identificar a estratégia adotada para o processo de migração de dados;
- Analisar as etapas percorridas durante o projeto e a sua ordem cronológica;

Por conseguinte, existe uma questão de investigação que está associada ao objetivo principal da dissertação e objetivos específicos com o intuito de ajudar a responder à questão principal. A questão de investigação considerada como o ponto de partida para este estudo é: Como uma organização internacional executou um processo de migração de dados entre diferentes sistemas de gestão de conteúdo?

Metodologia

A migração de dados entre sistemas de gestão de conteúdo não é um processo que ocorra diariamente nas organizações, depois de pesquisar e investigar as diversas metodologias de estudo a abordagem escolhida para esta investigação incidiu sobre a metodologia estudo de caso.

Keeves (1997) define estudo de caso como o recurso a um "termo geral" para a investigação de um indivíduo, grupo ou fenômeno. Esta metodologia é classificada como uma forma de investigação com fundamento no estudo de um fenômeno em específico, em um contexto real. Devido às características próprias de um Estudo de Caso e a sua capacidade de estudo e compreensão de um determinado contexto, é extremamente benéfico quando é aplicado em uma circunstância especial e se pretende ter uma compreensão detalhada uma vez que existe muita informação.

Para Bogdan e Knopp Biklen (1994) o estudo de caso pode recorrer a um único caso ou recorrer a múltiplos casos. Neste caso o estudo será aplicado a uma única organização, sendo apenas um único caso.

Yin (1984) afirma que um estudo de caso poderá ter um foco mais preciso e objetivo em um grupo, organização e ainda especificamente num indivíduo. O autor diferencia que o estudo de caso pode ter também um foco abstrato dando o exemplo de mudanças na organização, implementação de processos entre outros.

A presente investigação classifica-se com um estudo de caso, com uma abordagem do tipo descritivo. A recolha de dados foi por meio de observação e registo detalhado do projeto de migração.

Estrutura do trabalho

O documento é estruturado em 5 capítulos. A parte introdutória é apresentado o âmbito do trabalho, a origem do problema e as questões de investigação. Na seção seguinte (capítulo I) é apresentado uma fundamentação teórica do tópico, bem como uma análise de trabalhos relacionados ao tema em estudo.

O capítulo seguinte (capítulo II) é dedicado à apresentação do processo de migração, é realizada uma descrição dos sistemas que foram descontinuados e no novo sistema adotado, é apresentada a técnica utilizada para a migração de dados e por fim o processo de migração é descrito em termos temporais e tecnológicos.

O capítulo III visa a discussão de resultados que ocorre com a análise dos dados recolhidos no capítulo anterior. No último capítulo (capítulo IV) é um espaço reservado para as conclusões retiradas do estudo, as limitações encontradas na investigação e recomendações para futuros estudos.

CAPÍTULO I – FUNDAMENTAÇÃO TEÓRICA

Neste capítulo aborda-se a fundamentação teórica que visa explorar o tema e proporcionar uma base de conhecimento sobre o tema de migração de dados de sistemas de gestão de conteúdo e plataformas de comércio eletrónico (e-commerce).

1.1 e-commerce

Para Laudon & Traver (2016) a Internet é uma rede de computadores utilizada a nível mundial e a World Wide Web (WWW) é um serviço disponibilizado por esta rede que disponibiliza biliões de páginas web. Por vezes existe confusão por uma certa parte de utilizadores, que pensam que estes dois termos (internet e web) têm o mesmo significado, no entanto, eles estão interligados, mas com diferentes características. A internet é uma rede global de computadores interligados e a WWW permite a interligação de documentos e outros recursos. O acesso aos recursos disponíveis na WWW é realizado com o recurso ao navegador (browser). Este software, permite visualizar e interagir com os múltiplos tipos de formatos, hiperligações e aplicações que constituem a WWW. O e-commerce usa a internet, e a WWW para realizar transações comerciais entre organizações e indivíduos.

O e-commerce é uma curta versão do termo comércio eletrónico. Atualmente, designa-se por comércio eletrónico todas as transações eletrónicas entre organizações e outra terceira parte, todas as transações financeiras ou não financeiras como pedido de informações e apoio ao cliente são também considerados parte do comércio eletrónico (Bălan, 2014).

Qin et al., (2014) apontam que a tecnologia do e-commerce impulsionou uma revolução no sistema de circulação, ao romper com a fronteira de tempo e espaço, uma vez que altera o padrão do comércio ao aperfeiçoar a circulação de mercadoria, informação e capital, com todas características positivas que a tecnologia traz para o e-commerce faz com que as empresas tenham mais vantagens sobre outras empresas que não adotem esta forma de negócio.

Qin et al., (2014) afirmam que o desenvolvimento do e-commerce terá influência e um grande impacto em todos os aspetos da sociedade atual. E que o e-commerce tem um grande peso na manufatura, contribuindo para a quarta revolução industrial, com mais influência sobre a humanidade do que a segunda e terceira revolução industrial. Para além do e-commerce, aumentar a produtividade, operações económicas e diminuir custos de operação, o comércio eletrónico influencia os estilos de vida e todos os

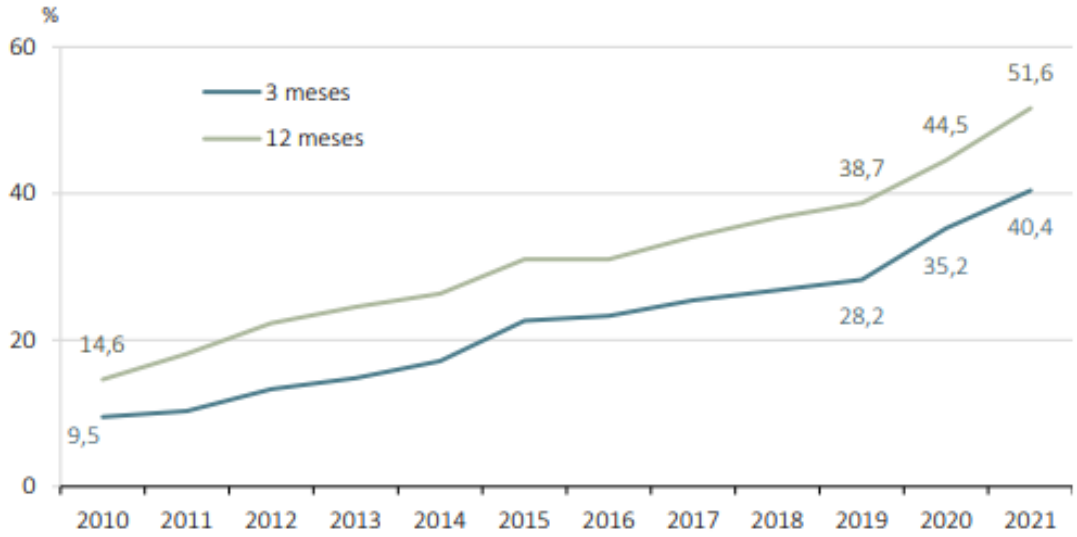
aspetos sociais da humanidade atual e com isso muda a sua visão do mundo e a sua metodologia.

Pearson (2017) defende que o e-commerce é essencial para a existência de empresas com marcas estabelecidas. As capacidades de e-commerce que estas empresas podem desenvolver são um fator muito importante para o sucesso das mesmas.

Com o aparecimento da doença coronavírus (Covid-19) e as medidas restritivas, forçou diversas empresas a migrar para o comércio eletrónico, uma alternativa para que empresas não vissem o seu negócio fracassar e continuassem a obter os seus lucros. Com a adesão de muitos negócios ao e-commerce existiu um grande desenvolvimento num curto espaço de tempo.

Num estudo realizado pelo Instituto Nacional de Estatística (INE) em 2020, 44,5% das pessoas com idades compreendidas dos 16 aos 74 anos fizeram encomendas pela internet nos 12 meses anteriores ao estudo e 35,2% nos 3 meses anteriores. A percentagem de utilizadores de comércio eletrónico registou em 2020 o maior aumento da série iniciada em 2002. Noutro estudo do INE, o crescimento da proporção de utilizadores do comércio eletrónico manteve-se (**Figura 1:** Proporção de pessoas dos 16 aos 74 anos que utilizaram comércio eletrónico nos últimos 12 meses anteriores à entrevista). Cerca de 82,3% da população residente em Portugal, com idade compreendida entre os 16 aos 74 anos, utiliza a internet. Estes resultados sustentam o reforço do crescimento verificado no ano anterior.

Figura 1: *Proporção de pessoas dos 16 aos 74 anos que utilizaram comércio eletrónico nos últimos 12 meses anteriores à entrevista*



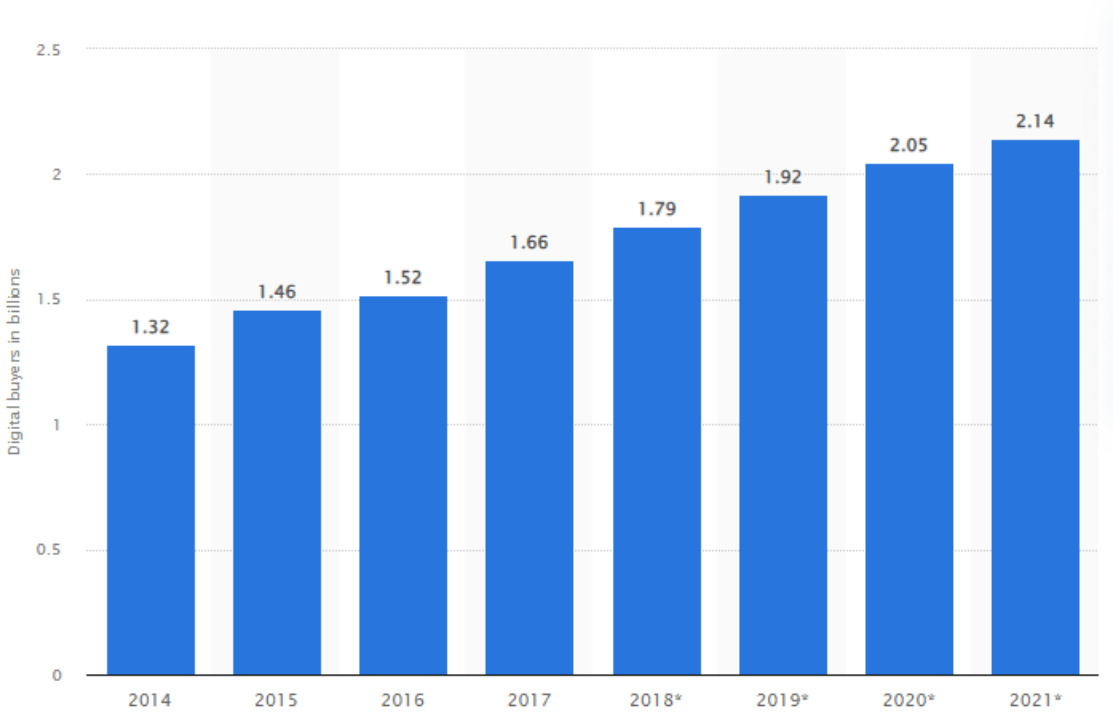
Fonte: INE, Inquérito à Utilização de Tecnologias da Informação e da Comunicação

No estudo anterior apenas são referidos dados sobre o comportamento da população portuguesa.

Num estudo publicado por Daniela Coppola (2021) no website Statista[2], é possível observar que o número de compradores online aumenta de ano para ano. A autora explica ainda que as previsões para o futuro são que as compras online continuem a ser populares nos EUA. Na **Figura 2:** Número de compradores digitais em todo mundo de 2014 até 2021 (em biliões) é possível observar um aumento contínuo no número de compradores digitais de 2014 a 2021.

[2] Coppola, D. (2021, Outubro 28). Number of Digital Buyers Worldwide from 2014 to 2021. https://www.statista.com/statistics/251666/number-of-digital-buyers-worldwide/

Figura 2: *Número de compradores digitais em todo mundo de 2014 até 2021 (em biliões)*

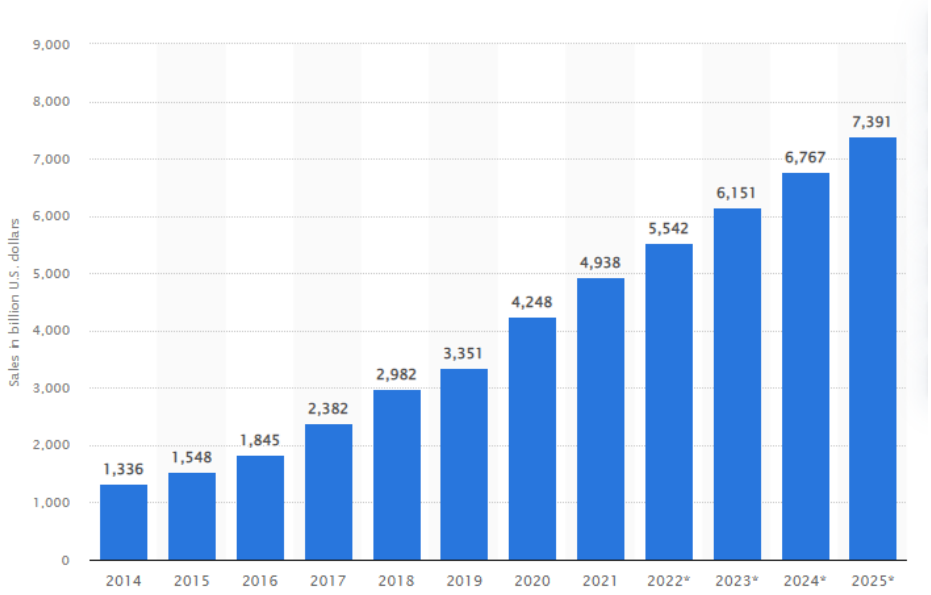


Fonte: Statista, https://www.statista.com/statistics/251666/number-of-digital-buyers-worldwide/

Na **Figura 3:** Vendas de comércio eletrónico no retalho a nível mundial de 2014 a 2021, com previsões de 2022 a 2025 é possível observar que existe um crescimento acentuado entre o ano de 2019 para 2020 em biliões de dólares para vendas no retalho através do comércio eletrónico e com projeções para os anos seguintes de progresso. Um estudo publicado por Stephanie Chevalier (2022) no site Statista[3], informa que o crescimento acentuado entre 2019-2020 é explicado pelo aparecimento da Covid-19 e as respetivas restrições que levaram aos consumidores a realizarem mais compras online.

[3] Chevalier, S. (2021, Outubro 28) Global Retail E-Commerce Sales from 2014 to 2026)https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/

Figura 3: *Vendas de comércio eletrónico no retalho a nível mundial de 2014 a 2021, com previsões de 2022 a 2025*



Fonte: Statista, https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/

1.2 Diferentes modalidades de e-commerce

Mohapatra (2013)diferencia as seguintes formas negócio: business-to-business (**B2B**), business-to-consumer (**B2C**), business-to-government (**B2G**), consumer-to-consumer (**C2C**) e por fim mobile commerce (**m-commerce**).

Mohapatra (2013) define B2B como sendo o comércio eletrónico entre empresas, suportando por via eletrónica as transações comerciais, tipicamente, entre clientes e fornecedores. Estas transações podem ocorrer entre fornecedor e comprador, fabricante com retalhista, colaboração entre duas empresas.

Laudon & Traver (2016) definem B2B como negócios que focam na venda para outras empresas, representando a maior porção de comércio online. Os autores afirmam que a dimensão total do B2B é potencialmente muito maior com uma capacidade de crescimento muito significativo. Acrescentam ainda que existem dois modelos de negócio primários usados no B2B, redes privadas industriais e *net marketplaces*[4] que inclui distribuidores e empresas de aprovisionamento eletrónico, bem como bolsas e consórcios industriais.

[4] É uma plataforma de vendas online onde se concentram diversas lojas no mesmo local.

Por outro lado, o B2C é o comércio entre empresas e consumidores, permitindo que os consumidores comprem um produto físico, ou apenas obter informação sobre o produto ou serviço. Esta forma de comércio reduz algumas barreiras de entrada nos mercados, como por exemplo, dado que o custo de uma loja física ser muito superior à manutenção de uma plataforma de e-commerce.

Para os autores Laudon & Traver (2016) o comércio eletrónico de empresas para consumidores B2C, é o tipo de comércio mais discutido, onde os negócios online tentam atingir os consumidores individuais. Estão incluídos a compra de bens a retalho, viagens e conteúdo online, é a forma de comércio que os consumidores mais facilmente se deparam.

Mohapatra (2013) define B2G, como a relação entre empresas e o governo, existem algumas características deste mercado que se referem a procedimentos públicos utilizando a internet, licenciamentos, entre muitas outras operações que envolvem o governo. Muitas vezes o governo interpreta um papel de liderança em estabelecer procedimentos online, é do interesse do setor público colocar os sistemas de compra a funcionar de forma eficiente.

Laudon & Traver (2016) declararam que a categoria de comércio eletrónico B2G, negócio entre empresas e o governo, está dentro do mesmo conceito que o B2B, colocando o setor público como uma simples forma de negócio que atua como um comprador de bens e/ou serviços.

Comércio entre pessoas individuais e consumidores é definido por Mohapatra (2013) como C2C. Neste tipo de comércio verifica-se o crescimento de *marketplaces* e leilões online, este tipo de comércio é mais focado para indústrias verticais onde negócios podem realizar ofertas do valor desejados entre outros fornecedores.

Para Laudon & Traver (2016) C2C é definido como consumidores venderem a outros consumidores, com a ajuda de plataformas de mercado online, um exemplo deste tipo de plataformas é o eBay. Nesta categoria de e-commerce o consumidor faz todas as preparações que o produto necessita para entrar no mercado, coloca o produto para venda ou leilão e a plataforma de mercado online providencia os catálogos, motores de busca e ainda capacidades de transação para que os produtos sejam expostos, encontrados e pagos.

Mohapatra (2013) define Mobile Commerce (M-Commerce) como a venda e a compra de bens e serviços por tecnologia wireless, como telemóveis. A autora acredita ainda que como a entrega de conteúdo via wireless se torna mais rápida e mais segura o m-commerce irá ultrapassar o volume de comércio através da rede fixa e será o método de transações digitais mais escolhido.

Laudon & Traver (2016) o M-commerce ou comércio mobile refere-se ao uso de dispositivos móveis para realizar transações online. Que envolve o uso de redes wireless para conectar computadores, tablets, telemóveis (iPhone ou telemóveis android) à internet. Quando a ligação é estabelecida, os consumidores conseguem realizar transações de ações, comparação de preços, operações no banco, reserva de viagens, compra de conteúdo de informação entre muitas outras ações.

Laudon & Traver (2016) acrescenta duas variantes de comércio eletrónico: social e-commerce e local e-commerce.

O Social e-commerce é realizado através das redes sociais e relações sociais online, é uma forma de negócio em crescimento devido a diversos fatores como a facilidade de iniciar sessão através de redes sociais como Facebook. Esta variante permite a partilhar a aprovação e desaprovação de produtos/serviços, ferramentas de compra online colaborativas e ainda a possibilidade de realizar uma pesquisa social e conhecer as recomendações de outros amigos de confiança online.

Local e-commerce é definido por Laudon & Traver (2016) como uma forma de comércio focada em envolver o consumidor dentro da sua localização geográfica com os comerciantes locais que recorrem a diversas técnicas de marketing para que estes se dirijam às suas lojas. Esta categoria e as duas anteriores, constituem a tendência de comércio eletrónico, móvel, social e local, alimentada pela procura de serviços locais.

1.3 Plataformas de e-commerce

Com o rápido crescimento do e-commerce nos últimos anos, as organizações com pequeno porte até às de grande escala adaptaram os seus negócios ao digital. As tecnologias utilizadas no e-commerce, diferenciam-se de outras ferramentas das últimas décadas, o que precede a diferenças sem precedentes no comércio (Laudon & Traver, 2016).

As plataformas de e-commerce podem ser consideradas ferramentas de software de soluções, que representam um espaço para apresentar produtos de forma digital e que podem desempenhar um papel importante para organização logística, para o marketing e para diversos processos relacionados com a componente financeira das transações (Dragomirov, 2020).

As tecnologias do e-commerce ajudam os comerciantes a ter acesso a mais informações sobre os consumidores e como podem utilizar essas informações de forma mais eficiente (Laudon & Traver, 2016).

Atualmente existem diversas soluções disponíveis de plataformas de e-commerce. Laudon & Traver (2016) afirmam que para selecionar a ferramenta correta é necessário elaborar uma avaliação das mesmas. Realizar uma decisão sobre qual ferramenta utilizar é sem dúvida um passo importante e uma das decisões mais incertas a ser realizada aquando da construção de um website e vendas online. Além da necessidade de escolher qual plataforma será utilizada é necessário ter em conta outros fatores como os custos que podem não ser constatados inicialmente, entre eles a formação necessária para os profissionais com contacto direto com a ferramenta, quer seja como editores, administradores ou desenvolvedor web.

Um dos primeiros passos das empresas é considerar as diversas opções que existem para escolher uma plataforma de comércio eletrónico tendo em consideração categoria de negócio em que a empresa se insere (B2B, B2C, C2C).

Mathur (2016) aponta para a existência de opções como: *Hosted Shopping Cart Software* e *Open Source e-commerce Platforms*.

Hosted Shopping Cart Software: Os carrinhos de compras hospedados permitem usar o sistema de gestão de conteúdo e o alojamento que a empresa escolher, hospedam apenas o processo de carrinho de compras e o registo de saída. O mecanismo de carrinho alojado em sistema externo trata dos detalhes de segurança, pois o pagamento é processado pelo site, mas o site inteiro não é hospedado pelos fornecedores.

Open Source e-commerce Platforms: As plataformas de comércio eletrónico baseadas em código aberto (*open source*) têm muitas vantagens, pois permitem fazer o download e ser hospedadas em qualquer lugar. Para os designers e programadores, as soluções de código aberto podem ser uma boa opção, mas quando a solução é de

utilização livre não existe suporte técnico dado pelo fabricante do software, ficando a equipa de desenvolvimento muito dependente do apoio da comunidade de utilizadores.

Neste subcapítulo referimos a importância do e-commerce e dos diferentes tipos de plataformas de e-commerce. É imprescindível para o presente documento apresentar os sistemas de gestão de conteúdos (Content Management System), os quais são a base para muitas empresas apresentarem os seus produtos ou serviços, realizarem campanhas de marketing, presentear os seus clientes com informações e avaliações do seu negócio.

1.4 Content Management System (CMS)

Laudon & Traver (2016) define CMS (Sistema de Gestão de Conteúdos) como um software que permite a gestão de conteúdos através de uma interface amigável. Esses conteúdos podem ser ficheiros (dados não estruturados) e/ou registos em base de dados (dados estruturados). É ainda possível que os gestores criem, organizem, publiquem e apaguem conteúdos do site.

De forma sucinta para os autores Alrasheed & Amir D. EL-Masri (2013) a função principal dos CMS resume-se a possibilitar os utilizadores sem conhecimento técnico a editar conteúdo apresentado no website da forma mais fácil.

Para Rohilla (2017) os sistemas de gestão de conteúdo ajudam na criação, gestão e edição de conteúdo de forma fácil. A autora define conteúdo como, texto escrito, imagens, configurações de variáveis para os produtos disponíveis. Com os sistemas CMS é possível ao gestor de conteúdos da plataforma fazer alterações no conteúdo de qualquer parte do mundo com acesso à Internet e no horário que pretende de forma livre. Os gestores destes sistemas podem criar, modificar e remover conteúdo do website e não requer conhecimento de linguagem código para realizar as ações anteriormente mencionadas.

Os autores Laudon & Traver (2016) afirmam que se o objetivo do website não for orientado para vendas, que necessitem de um carrinho de compras, uma das estratégias que as empresas podem adotar é a aquisição de um CMS por não ser uma solução dispendiosa.

Para Alrasheed & Amir D. EL-Masri (2013) os CMS são uma forma de suporte para o e-commerce que providenciam a facilidade de editar conteúdo no website.

Rohilla (2017) no seu artigo afirma que com o crescimento mundial de compras e vendas online de bens e serviços, o número de sistemas CMS e plataformas de e-commerce crescem de forma paralela.

Engelfriet (2010) explica de forma muito clara o porquê da necessidade de referir os CMS's e a sua relação com o e-commerce. Segundo o autor, alguns utilizadores podem achar que construir uma plataforma de e-commerce é simples, bastando apenas escolher um sistema de carrinho de compras e seguir. Atualmente os consumidores não compram um produto apenas a olhar para as especificações e o preço do mesmo, precisam de ver avaliações do produto e a sua respetiva classificação. Existem ainda produto relacionados com a internet, e para este tipo de bem é essencial criar material que o suporta.

Engelfriet (2010) afirma ainda que as descrições dos produtos têm ficado cada vez mais complexas, tem diversas cores, variações, tamanhos, padrões entre outros aspetos. Para aumentar as vendas, é necessário fazer *up-sell* e *cross-sell* (incentivar os clientes a adquirir uma versão mais sofisticada do produto e a venda de produtos ou serviços relacionados respetivamente). É preciso também apresentar diversas imagens do produto para que os utilizadores tenham uma melhor perceção do que o produto é realmente. É essencial que os consumidores possam comparar produtos de modo a facilitar a melhor escolha e oferecer descontos. Mais importante do que vender, é cativar o cliente para que este sinta a vontade de voltar a comprar, fidelizando-o. Para isso, existem diversas formas de fazer com que o cliente não se sinta esquecido e para isso é essencial enviar newsletter, escrever um blog, ter um fórum de suporte, e enviar ofertas especiais. Para que o site obtenha tráfego de qualidade e estar bem qualificado nos motores de busca (Engelfriet, 2010).

Barker (2016) explica de forma mais detalhada, o CMS é um software que utiliza a automação para realizar tarefas úteis para a gestão de conteúdo. Os CMS têm de forma geral multiutilizadores, que interagem com o conteúdo armazenado num repositório. Esse mesmo repositório pode ter a mesma localização no mesmo servidor como pode fazer parte do mesmo software ou até mesmo armazenado separadamente. Os CMS's permitem aos editores criarem conteúdos, editarem o conteúdo que já existe, executarem processos de edição no conteúdo e também disponibilizar esse conteúdo para outras pessoas para poderem consumir esse conteúdo. Um CMS é constituído por diversos componentes, por exemplo, a interface de edição, o repositório, os mecanismos

de publicação entre outros componentes. Devido à arquitetura dos CMS baseadas em camadas, estes componentes podem ser autónomos e separados. Para um utilizador do sistema que não tenha muito conhecimento técnico, estas partes são vistas com um todo.

1.5 Migração

Para Trias et al., (2013) as migrações podem ocorrer na maioria das vezes devido à evolução da tecnologia. Atualmente, os softwares estão em constante evolução provocando migrações contínuas de componentes de software e consequentemente os dados em base de dados. Além dos motivos mencionados, também existe o caso dos sistemas descontinuados, podendo ser substituídos por outros baseados em tecnologias mais modernas.

Zarmpou et al., (2009) indica diversas razões pelas quais organizações decidem avançar com migrações entre softwares, são:

- Tecnologia ultrapassada, o sistema já não consegue responder às necessidades, como o aumento de volume de conteúdo, as funcionalidades não são suficientes, baixo desempenho, entre outros;
- Pressão dos utilizadores;
- Necessidade de criar um único sistema de informação coerente quando ocorre a fusão de empresas;
- Novos softwares no mercado com tecnologia mais avançada e menos dispendiosos;
- Busca por uma solução mais económica devido a cortes no orçamento;
- O fornecedor do sistema parou o serviço de suporte, deixando o sistema sem atualizações de segurança e sem correção de erros.

As migrações podem acontecer em pequena ou grande escala. Segundo Shoaib et al. (2017), as migrações podem ocorrer utilizando apenas um único sistema origem, que corresponde a migração de pequena escala. Se a migração contém mais do que um sistema origem é considerado uma migração de larga escala. As migrações podem ocorrer ao nível do código, plataforma, base de dados, sistema operativo, entre outros. A presente dissertação foca-se na migração de dados.

1.5.1 Migração de dados

A migração, segundo Ramalho et al., (2009) "é o processo de transferência de informação digital entre dois ou mais sistemas de informação." (p.2).

A migração de dados é o processo de mover dados de uma plataforma para outra. Envolve a transferência de dados de um sistema origem ou vários sistemas para um novo sistema destino. Este tipo de processo envolve várias etapas, desde a análise de dados antigos até ao carregamento de dados na nova aplicação (Sarmah, 2018).

Para Silva de Oliveira & Abud Marcelino (2012) a migração de dados estruturados significa mover os dados de um SGBD (Sistema de Gestão de Banco de Dados) para outro SGBD. Apesar de existirem padrões de tipos de dados, os dados de cada campo de uma tabela podem variar em tamanho/formato e cada SGBD possui características próprias. Num CMS existem dois tipos de dados, dados estruturados e não estruturados.

Caso os dados estejam disponíveis nas aplicações existentes e o seu volume seja muito grande, estes não devem ser criados de novo, devem ser migrados das aplicações existentes para a nova aplicação, segundo os autores Heleno & de Mendonça (2009). No entanto, alguns projetos de migração não consideram que todos os dados de origem são essenciais, pelo que será necessário filtrar os dados mais relevantes. Esta análise de relevância dos dados deve ser realizada com os utilizadores impactados diretamente com as mudanças. Existem projetos de migração que por exigência do negócio todos os dados devem ser migrados, e requerem um planeamento mais detalhado e mais esforço por parte da equipa que lidera o projeto.

Existem diversos fatores que se deve considerar num processo de migração para Mello (2009): a análise da motivação da migração e a tecnologia escolhida irá corresponder às novas necessidades, o autor aponta como os dois principais fatores a ter em consideração.

1.5.2 Estratégias de migração de dados

Antes de iniciar o processo de migração de dados é essencial as organizações decidirem qual a estratégia que vão adotar, para que os dados que são alterados e transferidos para um novo banco de dados, mantenham a integridade e significado. Existem duas estratégias principais para a migração de dados: *Big Bang* e *Trickle* (Heleno & de Mendonça, 2009).

Big Bang: esta estratégia é caracterizada por realizar migrações de uma vez só. Para ocorrer, os sistemas de origem devem ficar fora de serviço enquanto os dados são extraídos ou carregados no sistema destino. Contudo, muitas organizações não podem ficar com a plataforma fora de serviço. Como forma de minimizar a interrupção de serviço, frequentemente as migrações são realizadas aos fins de semana ou feriado (Soares Vian, 2020).

Trickle: esta estratégia executa as migrações de dados de forma incremental. O sistema de origem e destino são executados em paralelo e a migração ocorre em etapas. Para manter os dados atualizados em ambos o sistema é necessário um processo, que transfira as atualizações do sistema de origem para o sistema destino (Soares Vian, 2020).

1.5.3 ETL como processo de migração de dados

Em todos os projetos de migração é necessário definir qual o processo de migração será adotado. Durante a pesquisa surgiram estudos relacionados ao tópico em investigação "migração de dados", nesses estudos os autores apresentavam o processo de ETL como uma metodologia aplicada. Por conseguinte, neste estudo será discutido o processo ETL (E*xtract, Transform and Load*).

Extração, Transformação e Carregamento (ETL - Extract, Transform and Load)

ParaVassiliadis (2002), o processo ETL não é somente utilizado na migração de dados, é aplicado em projetos de *Data Warehouse* (DW), em português armazém de dados. O processo de ETL é responsável pela extração de dados de diversas fontes, subsequentemente a limpeza, tratamento e carregamento nos DW.

O processo ETL divide-se em três fases: extração, transformação e carregamento. A fase de transformação pode ser opcional, uma vez que irá depender da necessidade ou não da transformação de dados.

Na primeira fase é realizada a extração de dados, que consiste no processo de extrair os dados do sistema origem e armazená-los numa área de trabalho do processo ETL designada por DSA[5] (*Data Staging Area*). É muito frequente estes dados fonte serem provenientes de vários sistemas, como de um sistema de gestão de dados, documentos

[5] É uma área de armazenamento intermediária e combina dados de diferentes fontes.

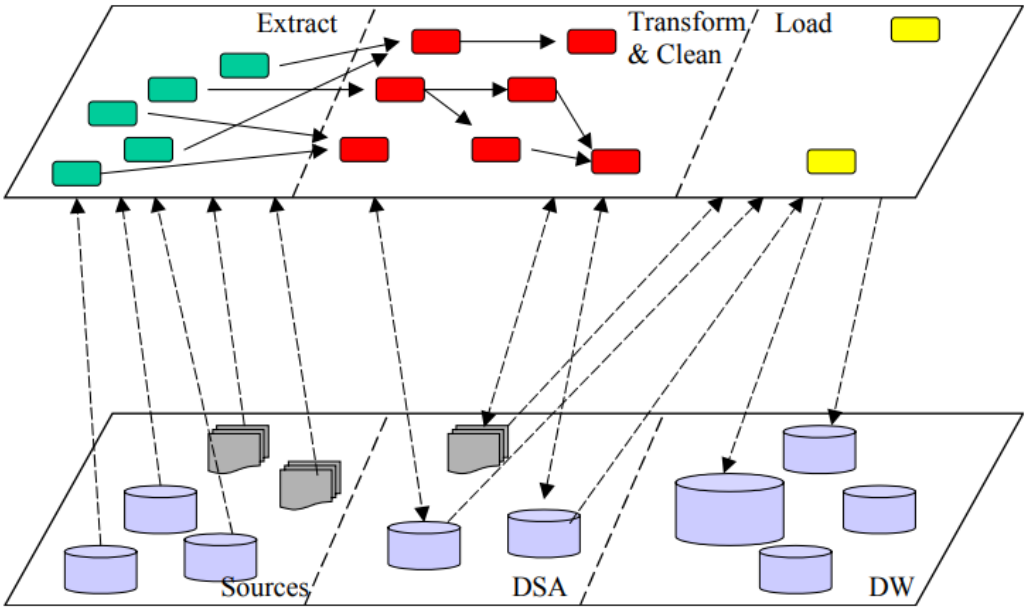
Excel, de arquivos entre muitos outros. Segundo Silva de Oliveira & Abud Marcelino (2012) é importante durante o processo de extração documentar de forma nítida como os dados estão organizados no sistema, para ajudar na fase de transformação, para que esta seja uma etapa sem perdas de dados e pormenorizada.

Para Silva de Oliveira & Abud Marcelino (2012) o sucesso na segunda fase, transformação, está relacionado com a fase anterior, pois a organização dos dados, o detalhe e a qualidade podem ajudar a tornar este processo mais fácil. Nesta fase é realizado o mapa do conjunto de relacionamentos entre os dados de origem para os da tabela destino, bem como todos os processos necessários para manipular e converter informações para se adaptar ao novo sistema. Para Tripathy & Kshirasagar (2015) nesta fase é possível descartar dados desnecessários ou realizar melhorias para que se obtenha um desempenho superior.

Por último, existe a fase de carga ou carregamento (*load*) que consiste na transferência de dados concebidos na etapa de extração para serem inseridos na base de dados no sistema destino (Silva de Oliveira & Abud Marcelino, 2012).

Na figura 4 é apresentado as etapas do processo ETL, na camada superior é possível observar os dados provenientes das fontes onde são extraídos, em seguida os dados são limpos e transformados para serem carregados para o novo armazém de dados (DW – *Data Warehouse*).

Figura 4: *Arquitetura do Método ETL*



Nota: Figura retirado do estudo de Vassiliadis, 2002

Ferramentas de ETL

Para o desenvolvimento de DW o processo ETL é utilizado globalmente. Atualmente, existem diversas ferramentas de ETL. Analisar e escolher qual a solução mais apropriada vai depender dos requisitos, tipos de dados e infraestrutura da organização (Patel & Patel, 2021).

No estudo realizado por Jacobi et al. (2012), as ferramentas de ETL são peças de software que dão suporte aos utilizadores nas fases de modelação e execução de processos ETL. Com o aumento de tamanho das empresas, o crescimento de sistemas operacionais que precisam de ser integrados e dos relatórios que devem ser criados para responder às necessidades dos utilizadores, os processos de ETL tornam-se cada vez mais complexos e maiores.

No estudo de Almeida et al. (2021) estas operações podem ser processadas por linguagem de programação. Com o crescimento da complexidade destes projetos existiu a necessidade de utilizar abordagens mais fáceis de utilizar. Estas abordagens estão focadas em documentar o processo, enquanto outras abordagens foram concebidas para uma Interface Gráfica do utilizador (GUI) para especificar os fluxos de trabalho ETL.

Biswas et al. (2019) afirma que algumas ferramentas de ETL providenciam soluções com uma interface gráfica do utilizador fácil de utilizar, para que o mapeamento de

dados entre a fonte e o sistema seja simples para o utilizador e possa ser realizado rapidamente. Para Biswas et al. (2019) é mais fácil selecionar e usar uma ferramenta de ETL comercial quando comparado com a criação e manutenção de um sistema ETL com código personalizado feito à medida.

Num estudo realizado por Tirumala et al. (2015), estes autores afirmam que os processos de ETL podem proceder recorrendo a várias ferramentas, que podem ser de código aberto ou de propriedade. Os autores afirmam ainda que nos processos de ETL, todas as linguagens de programação podem ser aplicadas e executadas. Criar um programa de raiz pode ser muito complexo e por esta razão ser necessário utilizar ferramentas.

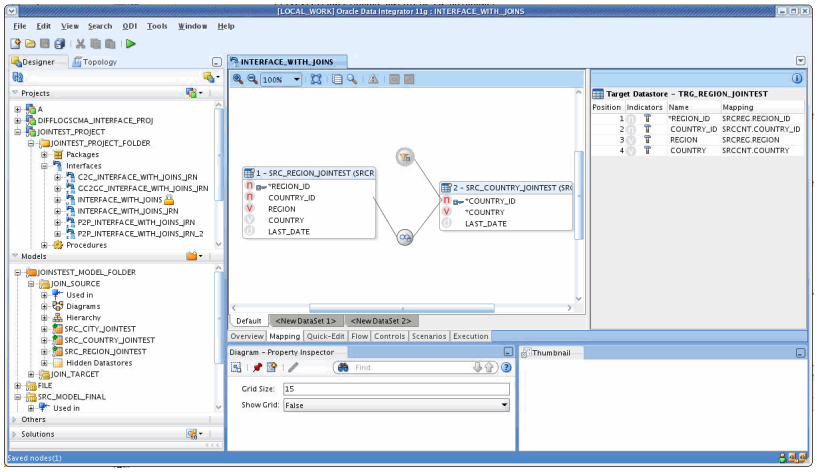
Biswas et al. (2019) afirma no seu estudo que existem desvantagens nas ferramentas de código aberto e comerciais, quando comparadas com uma ferramenta ETL personalizada. Estas ferramentas apresentam dificuldades ao serem utilizadas num cenário específico, apenas disponibilizam itens limitados. Em projetos de grande escala, como movimentar ícones, configurar propriedades e até desenhar linhas de fluxo pode ser uma tarefa complicada. No caso de projetos mais complexos, uma solução de código personalizado pode ser vantajosa.

No seu estudo, Biswas et al. (2019), afirma que a produtividade dos sistemas deve ser levada em consideração. As ferramentas GUI não asseguravam uma produtividade maior que as ferramentas de código personalizado. Para o autor, os utilizadores sem conhecimentos de programação avançada, as ferramentas de ETL de código aberto e comerciais tem uma interface gráfica mais fácil de utilizar e conseguem desenhar o fluxo de dados de forma visual recorrendo ao processo de ETL.

Atualmente, existem diversas ferramentas disponíveis, ferramentas ETL, estas podem ser comerciais ou código aberto, ambas para a gestão de dados estruturados. No estudo de Patel & Patel (2021) existem sete categorias diferentes para ferramentas de ETL, no entanto, a maioria das ferramentas não entra em apenas uma classe, mas em diversas, em cada categoria apresentamos uma imagem de uma interface exemplo.

Processamento em lotes – uma forma tradicional de processamento, primeiro junta todos os dados antes de processá-los em lotes.

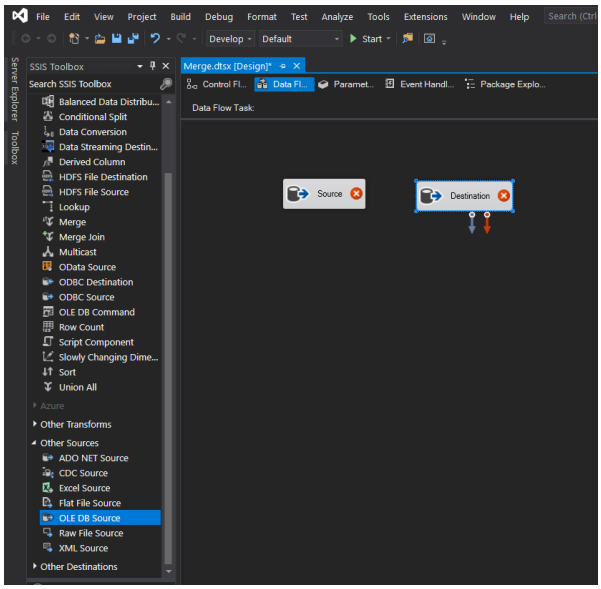
Figura 5: *Interface de uma ferramenta de processamento em lotes*



Fonte: https://docs.oracle.com/en/

Com base em código – estas ferramentas com base em código são programas compilados e não estão escritos com linguagem de programação proprietárias. Ferramentas com base em código ou com base em motor concedem um desempenho de excelência, mas é necessário ter conhecimento para personalizar código.

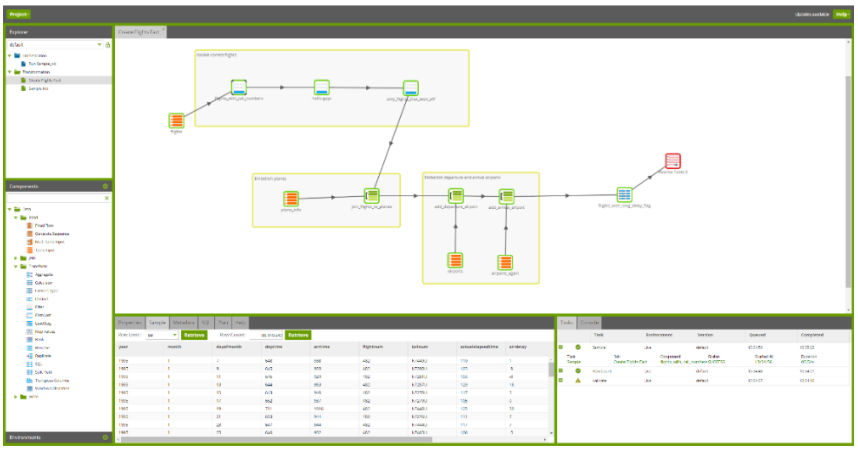
Figura 6*: Interface de uma ferramenta com base em código*



Fonte: https://learn.microsoft.com/en-us/sql/integration-services/

Com base em nuvem – estas ferramentas de ETL precisam de suportar aumentos na sua carga sem que existam problemas no seu desempenho, processamento de dados em tempo real, integrações mesmo com o aumento do número de fontes de dados.

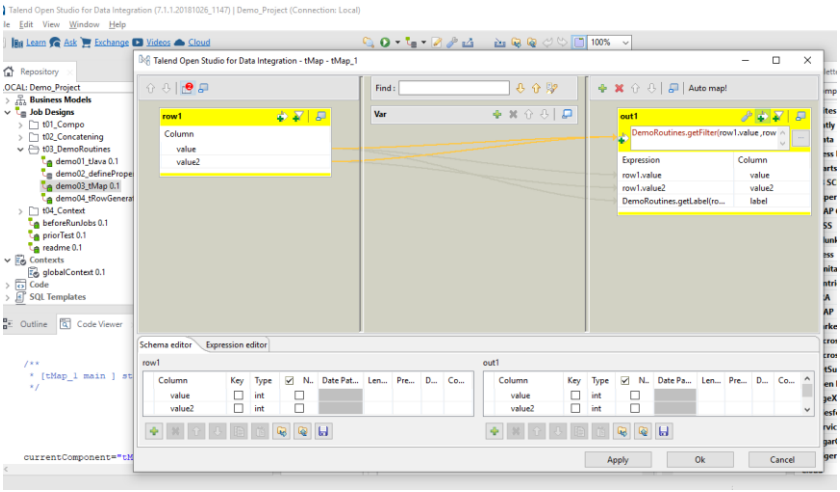
Figura 7*: Interface de uma ferramenta com base em nuvem*



Fonte: www.matillion.com

Open source – grande parte das organizações opta por soluções de código devido ao seu baixo custo comparado com soluções mais caras como soluções comerciais, são menos complexas e relatórios obrigatórios.

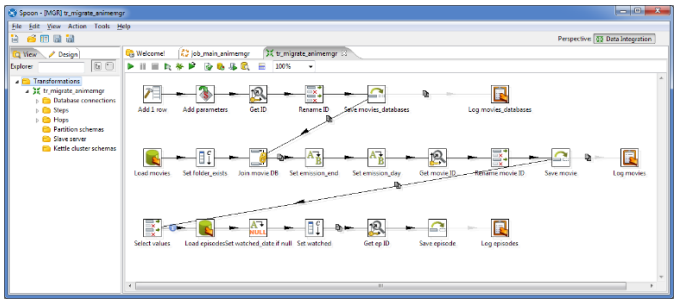
***Figura 8**: Interface de uma ferramenta com base em Open Source*



Fonte: www.talend.com/

Com base GUI – são as ferramentas mais usadas no mercado por serem de fácil utilização, este tipo de soluções concedem a funcionalidade de agarrar e mover para carregamento de dados e análise.

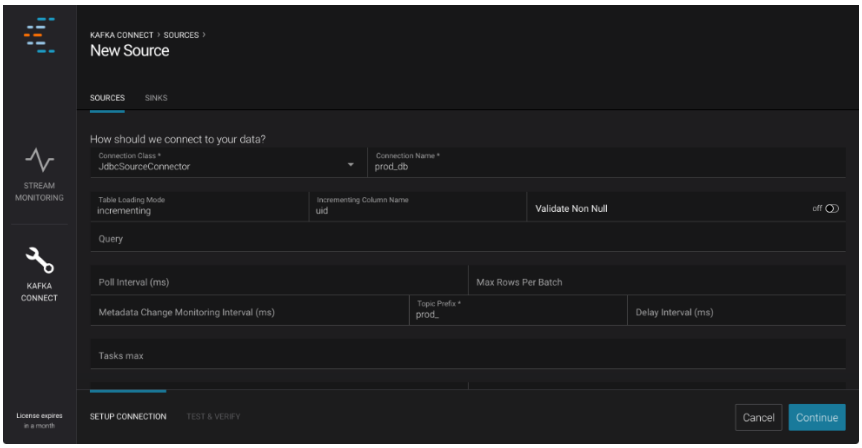
***Figura 9**: Interface de uma ferramenta com base GUI*



Fonte: www.pentaho.com

Em tempo real – a procura por dados em tempo real alterou a arquitetura do processo de ETL, com a evolução de dados e fontes, esta característica do processamento de dados em tempo real passa a ser um requerimento.

***Figura 10**: Interface de uma ferramenta com base em tempo real*

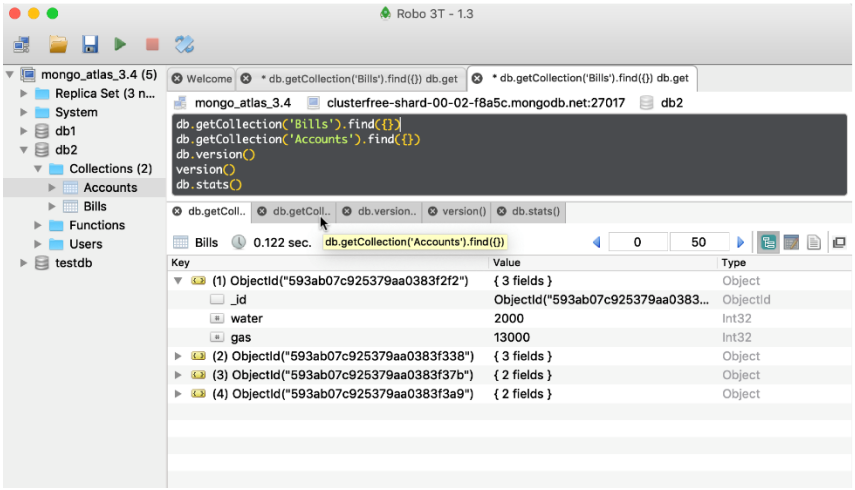


Fonte: www.confluent.io

Com base NoSQL – o processo de ETL que recolhe dados digitais e converte estas informações para uma forma percetível, como relatórios. Com o desenvolvimento de base de dados *schema-less*[6], a realização de análises torna-se complicado.

***Figura 11**: Interface de uma ferramenta com base NoSQL*

[6] É um recurso oferecido por SGBD que recorrem à abordagem NoSQL para armazenar dados, sem se importar com a sua normalização.



Fonte: www.mongodb.com

1.5.4 Ferramentas para migração de dados

Existem ferramentas que auxiliam as organizações nos processos de migração para que estas sejam mais seguras, mais práticas e ágeis. Ao longo do estudo surgiram trabalhos relacionados com a migração de dados como artigos, dissertações e *whitepapers* de algumas ferramentas utilizadas para realizar este processo. Dos documentos encontrados identificou-se três ferramentas, apresenta-se algumas das ferramentas analisadas:

Talend Open Studio

É um software de integração de dados de código aberto, é utilizado na integração de sistemas operativos e como ferramenta de ETL para armazenamento de dados, business intelligence (BI) e migração de dados. Para satisfazer as necessidades de todos os tipos de organização, a empresa fornece soluções abertas, inovadoras e poderosos softwares (Talend, 2020).

Com esta ferramenta é possível sincronizar dados de fontes heterogéneas, fácil de usar, uma interface intuitiva e com muitas ferramentas de modelação. Contem mais de 450 conectores de dados, com um ambiente de desenvolvimento integrado (IDE) amigável e compreensível, capacidade de criar código Java, este código pode ainda ser modificado para obter maior controlo e flexibilidade (Talend, 2020).

Talend utiliza Perl e Javascript como linguagem de código para definir o próprio componente, é uma ferramenta escalável e funciona em Windows e Linux. É especialmente utilizado para integração de dados, qualidade e gestão de dados (Badiuzzaman Biplob et al., 2018).

Vamosa Migration Architect

A Vamosa Technologies é uma divisão da T-Systems, especializados em migração de conteúdo em *Cloud*, *On-Premises* ou *Hybrid Cloud Systems*. *Migration Architect* é nome da ferramenta de migração utilizada por Vamosa, foi construída de raiz e com suporte interno. Este software é contruído sobre a plataforma .NET e utiliza uma base de dados de documentos NoSQL (Cockburn, 2019).

O software contém uma vasta variedade de fontes fora da caixa e conectores de destino utilizados para interações com as diversas APIs[7]. Para estender as suas capacidades é executado scripts[8] personalizados e regras, com motor em IronPython scripting incorporado como coluna vertebral.

O conteúdo é então manipulado dentro do software e não tem impactos na fonte enquanto as transformações são aplicadas.

As condutas de migração flexíveis podem ser unidas para formar fluxos de trabalho complexos. Os fluxos de trabalho podem ser executados em série ou em paralelo devido à sua capacidade avançada para agendamentos. O software contém relatórios extensos e flexíveis, esta ferramenta também tem um HTTP REST API incorporado que permite a integração com outros sistemas de software.

Xillio

Organização especialista em análise, transformação e migração de dados não estruturados, utiliza a sua própria plataforma ETL para conteúdo (Content ETL Platform). Recorre ao Modelo Universal de Dados (UDM) enraizado na linguagem script Xill IDE. O modelo de dados permite ligar diferentes fontes de conteúdo (ECM, DM, entre outros) entre si de forma padronizada, sem ter a necessidade de adaptar os modelos de dados subjacentes de cada sistema individual (Xillio, 2020a).

Xill é uma linguagem própria criada pela organização Xillio, utilizada para definir e executar migrações mais complexas. Além disso, a organização criou uma extensa biblioteca de aplicações baseadas em Xill utilizadas como conectores, para se ligar a

[7] É uma interface de programação de aplicações, é uma forma de dois ou mais programas de computadores comunicarem-se entre si.
[8] É um conjunto de instruções para que uma função seja executada numa determinada aplicação.

grandes sistemas ECM e WCM. Uma seleção destas aplicações é utilizada pela API da Xillio através do XDE Connector (é o Ambiente Distribuído Xillio, o conector XDE liga-se a este ambiente distribuído e fornece acesso a uma seleção de conectores)(Xillio, 2020b).

A Xillio afirma que a sua plataforma ETL de conteúdo consegue recuperar conteúdo não estruturado de qualquer sistema de conteúdo de forma simples, rápida e sem recorrer ao uso de programação extensiva. O conteúdo é analisado, enriquecido e por fim migrado. A plataforma contém várias funções integradas, módulos, conectores fora da caixa para os diversos sistemas de gestão de conteúdo. A utilização do Xill IDE é uma forma mais rápida, fácil e controlada do que o tradicional scripting (Xillio, 2020b).

1.6 Trabalhos Relacionados

Para apoiar a investigação, foi realizada uma pesquisa por estudos e trabalhos relacionados, foram analisados e recolhidos. Faremos uma abordagem a esses estudos para descrever semelhanças, entender o processo e identificar falhas desses processos.

1.6.1 Migração de dados em Sistemas de Gestão de Conteúdo Web

Zarmpou et al., (2009) desenvolveram um trabalho sobre migração de dados entre sistemas de gestão de conteúdo web (*Data Migration Between Web Content Management Systems*). Neste estudo foi realizado uma pesquisa através de questionários, direcionados para grandes empresas onde foram executadas migrações de dados de sistemas antigos para sistemas mais recentes.

Neste estudo os autores basearam-se em três opções de migrações: migração manual, migração semiautomática e migração automática. Durante a pesquisa, os autores procuraram saber como a migração foi realizada em cada organização, que ferramentas foram utilizadas, quanto tempo o processo de migração demorou, o volume de informação transferida e problemas encontrados com o processo adotado.

O estudo apresenta fatores que podem afetar o processo de migração de dados em empresas. Independentemente dos diferentes tipos de negócio foi possível visualizar com este estudo que existem falhas repetitivas que afetam o processo.

Os autores do estudo não apresentaram metodologias utilizadas nestes processos, pode ser visto como uma limitação, apesar de terem uma amostra de empresas considerável e

conseguirem levantar bastante informação sobre as migrações que efetuaram. A informação levantada sobre o método que utilizaram para o processo de migração ficou em falta para dar a conhecer a futuros leitores interessados no tópico.

1.6.2 Desenvolvimento de uma ferramenta de migração de banco de dados

Soares Vian (2020)apresenta um estudo sobre a relação do aumento da quantidade de dados que uma empresa produz e a necessidade de relacionar esses dados, armazenar e segmentar em diversos bancos de dados. O autor relaciona ainda o tempo necessário para investir num processo de migração, bem como a falta de experiência das empresas.

Para resolver o problema encontrado, o autor apresentou um software web, uma ferramenta para realizar migrações. Para ajudar o leitor a entender todo o processo, o autor apresentou as características do projeto, as funcionalidades e parte do código desenvolvido. Soares Vian recorreu à estratégia de migração Big Bang, utilizando como metodologia, a migração ETL.

Através da ferramenta desenvolvida, iMigration, o autor da pesquisa conseguiu comprovar que com o recurso a esta ferramenta foi possível reduzir a complexidade do processo de migração proposto, bem como a diminuição do tempo necessário para a migração ocorrer.

O estudo desenvolvido pelo autor foi aplicado a pequenas e médias empresas através de um questionário, o autor concluiu que a ferramenta de migração desenvolvida seria útil para empresas que optam por um processo de migração manual.

Algumas limitações que podem ser pontadas ao estudo do autor estão relacionadas com a quantidade de dados, existem empresas com diversos tamanhos e com isso a quantidade de dados também varia, no estudo a migração apenas foi aplicada a pequenas empresas, e a ferramenta utilizada não é comercializada.

1.6.3 Migração de bases de dados relacionais para NoSQL - Métodos de Análise

O autor Vieira De Oliveira (2017) realizou uma investigação para validar métodos de migração já existentes, utilizados e realizar uma adaptação, com a intenção de perceber qual era o método mais eficiente para proceder a uma migração de uma base de dados relacional para uma base de dados NoSQL.

O trabalho realizado pelo autor foi realizado num contexto prático, em laboratório, onde realizou uma simulação de carga real e recorrendo a ferramentas oficiais e normalizadas para o efeito.

Na investigação nos cenários testados pelo autor, este conseguiu identificar quais os requisitos necessários e identificar as várias fases de uma migração, analisar os métodos de migração disponíveis para a migração de base de dados. O autor recorreu aos métodos de migração online e offline e realizou uma avaliação desses métodos recorrendo a métricas específicas. O autor realizou ainda uma comparação dos diversos métodos e determinou qual o método de migração era mais eficiente.

O autor conclui que o método mais eficiente para realizar a migração de base de dados seria o método de migração offline automático, no entanto, levanta objeções devido à disponibilidade do ambiente, pois a janela de indisponibilidade é considerável.

A investigação realizada pelo autor é bastante completa e efetua uma abordagem ao tópico de migração de dados muito teórica e técnica, o autor realiza migração de dados, em ambiente de laboratório e por essa razão não fica sujeito a incidentes e atrasos que podem ocorrer em ambiente empresarial, o que pode se apresentar como uma limitação na investigação.

1.6.4 Metodologia de migração de dados em contexto de migração de sistemas

Heleno & de Mendonça (2009) desenvolveu o seu projeto "Metodologia de Migração de Dados num Contexto de Migração de Sistemas Legados" que consistia na integração de algumas estratégias de migração de dados existentes com a prática na gestão de projetos, para o autor conseguir definir uma metodologia genérica de migração de dados para aplicar num contexto de migração de sistemas legados (sistema de origem) recorrendo a um caso real.

Com a aplicação de metodologia composta por cinquenta e duas atividades distribuídas em seis etapas:

1. Planeamento;
2. Acompanhamento e Controlo;
3. Profiling e Auditorias;
4. Construção e Design;
5. Execução;

6. Testes e validação de dados.

O autor conseguiu melhorias significativas no projeto. Algumas das medidas passavam por considerar um projeto de migração independente das restantes atividades e um líder exclusivamente dedicado, bem como as fases do processo de migração detalhadas.

Apesar das vantagens com a metodologia apresentada o autor fez algumas sugestões de melhorias, uma vez que o estudo realizado não foi num ambiente institucional, o autor apontou alguns pontos que podem ser vistos como limitações, por exemplo, as dificuldades na realização de tarefas relacionadas com o interesse pessoal da empreitada. Contudo, o autor conseguiu realizar a análise de várias estratégias de migração de dados e contribui com a proposta de uma metodologia para migração de dados num caso de estudo real.

1.6.5 Processos para uma migração de dados

Sarmah (2018) apresentou um documento sobre o processo essencial para migração de dados. Neste documento a autora evidenciou uma lista das principais considerações para definir uma estratégia de migração.

A autora partilhou as principais fases de uma migração de dados através das cinco fases do ciclo de vida de desenvolvimento de software (SDLC - software development life cycle) que são:

1. Avaliação de dados;
2. Limpeza dos dados;
3. Teste de extração e carga;
4. Extração e carga final;
5. Validação da migração;
6. Atividades pós-migração;

A autora faz ainda uma descrição detalhada das diferentes fases da migração de dados, que grupos/pessoas são essenciais em cada fase e os resultados adquiridos com cada fase do processo. Após a partilha das fases do processo de migração de dados, a autora apresenta as melhores práticas para a realização de testes/simulações antes da conversão oficial para o ambiente de produção.

A autora realiza uma comparação de como a migração de dados pode ser alcançada utilizando duas abordagens Big Bang ou Migração por Fases. É apontado no estudo a

necessidade de identificar os riscos de gestão que podem ser enfrentados em qualquer projeto e forma de como mitigar esses problemas em cada fase da migração.

No estudo a autora faz ainda referência às melhores práticas na utilização de ferramentas e da base de dados, alerta para os possíveis custos ocultos envolvidos nas atividades de migração.

A autora concluiu que para prevenir riscos as organizações necessitam de adotar uma metodologia consistente e fiável, para que a organização consiga planear, desenhar, migrar e validar a migração, aponta a necessidade de realizar uma avaliação sobre que ferramenta efetuará a migração.

Apesar da autora realizar uma apresentação das fases de uma migração entre muitas outras características essenciais para este processo, é apenas um estudo de uma teoria relacionada ao tópico. Não existe nenhuma referência a estudos/aplicação em contexto real.

CAPÍTULO II – PROCESSO DE MIGRAÇÃO

Neste capítulo, são apresentados os sistemas *TYPO3*, *Sitecore*, *Joomla* e *SAP Web Channel*. Na descrição de cada um dos sistemas, aborda-se o seu aparecimento, as vantagens e desvantagens de cada sistema. Seguidamente, será apresentado o sistema destino *Adobe Experience Manager* (AEM), algumas das suas características positivas e negativas encontradas com o seu uso. Finalmente, é apresentado o processo de migração de dados entre os sistemas origem e destino, em termos da estratégia adotada pela organização e o processo de migração, com ênfase no planeamento temporal e nas tecnológicas utilizadas.

2.1 Sistemas origem

TYPO3

O *TYPO3*[9] é um sistema de gestão de conteúdo livre, baseado em software de código aberto. Este sistema foi criado em 1997, por Kasper Skårhøj, um programador dinamarquês, que recorreu à linguagem de programação *PHP*, utilizando *MySQL* para base de dados. O *TYPO3* apenas se tornou num sistema de código aberto em 2001, com capacidade para criar sites de médio e grande porte.

Em termos de instalação o *TYPO3* é um sistema difícil para realizar a instalação e configuração de um website. É necessário tempo para aprender a usar a ferramenta, definitivamente mais tempo de aprendizagem em comparação com, por exemplo, *Wordpress* e *Joomla*. É uma ferramenta recomendada para programadores avançados.

A nível de intuição de navegação, é uma ferramenta mais complexa de navegar, comparativamente com outros sistemas conhecidos (ex: *Wordpress*, *Wix* etc). Poderá ser necessário ter formação para conseguir navegar pelo sistema com facilidade. Depois de algum tempo de utilização é muito simples criar páginas, mover, editar entre outras funcionalidades.

É um sistema flexível, isto é, consegue-se adicionar qualquer componente nas páginas (imagens, formulários, tabelas, multimédia, texto, entre outros), não usa *templates*[10]. Tem diversas opções de escolha para o tipo de página que se pretende criar como, por exemplo, normal (standard), hiperligação para URL externo, atalho, separador de menus, entre outros. É possível mover componentes dentro da página e mover a página

[9] Typo3. (n.d). The TYPO3 Project and Community – Open-Source CMS. https://typo3.org/
[10] É um modelo a ser seguido, com uma estrutura predefinida, que ajuda o desenvolvimento e criação de conteúdo.

por toda a árvore de conteúdo. Em termos de componente de edição existem muitas opções de escolha que permite ao utilizador da ferramenta utilizar o elemento de sua preferência na página. Apesar do sistema permitir a tradução das páginas para outros idiomas, não é possível mover os elementos ou componentes nas páginas traduzidas. Uma vez que não é possível mover elementos nas páginas traduzidas, a única opção é criar um elemento igual ao da versão original.

O *TYPO3* permite a criação de componentes personalizados, como, por exemplo, tabelas com documentos específicos para download e elementos visuais do tipo CAD. É possível desenvolver estes elementos para corresponder às necessidades de cada negócio, estes requerem um elevado conhecimento em programação.

O sistema *TYPO3* oferece centenas de extensões (*plug-ins*), que podem ser instaladas de forma rápida e fácil. Apesar do sistema providenciar muitas informações sobre como instalar novos *plug-ins,* nem sempre as extensões que se pretende ter no site estão disponíveis. Por essa razão é necessário envolver programadores para ajudar a obter uma determinada extensão.

Quanto à característica da elasticidade, dos gestores no sistema, este permite aplicar configurações às contas de cada utilizador. Para os utilizadores com contas de editores a gestão de conteúdos em *TYPO3* é simples. No caso dos utilizadores com contas administrativas, estes são obrigados a ter um pouco mais de conhecimento sobre o sistema, para os programadores é um "osso duro de roer".

Se na eventualidade de se tratar de um sistema com vários gestores, é possível bloquear páginas e apenas utilizadores com a configuração de administrador consegue editar e publicar a página. As contas com a atribuição de editores, não conseguem procurar ou identificar as páginas bloqueadas e necessitam de contactar o administrador para desbloquear essas páginas. O *TYPO3* disponibiliza um módulo para desfazer mudanças, desta forma é possível restaurar versões anteriores de páginas praticamente sem qualquer limitação.

É possível encontrar imensa informação sobre este sistema no website do *TYPO3*. Para quem está a iniciar o seu projeto a organização fornece informação pormenorizada da construção de um site e quadro para discussões, pode ainda encontrar informações relativamente à modificação e utilização do *script*.

***Tabela 1** - Desvantagens e Vantagens do Sistema TYPO3*

Desvantagens	Vantagens
TYPO3 não é fácil de configurar, recomendado para programadores mais avançados.	A estrutura das páginas não tem ordem pré-fixada de blocos ou secções.
Os administradores são obrigados a ter um pouco mais conhecimento.	Na edição de páginas o conhecimento em HTML será suficiente.
Para os programadores pode ser um sistema difícil de operar.	O *TYPO3* permite a criação de elementos personalizados.
Estrutura do mapa pode ser de difícil interpretação para novos utilizadores.	Para os perfis de editores a gestão de conteúdos em *TYPO3* é simples.
Apesar de ter alguns documentos, a documentação carece de detalhes e precisa de mais exemplos.	É possível encontrar alguma informação e documentos sobre este sistema.
Muitas extensões não são atualizadas há anos.	Desfazer alterações ou restaurar a versão anterior da página.

Sitecore

Sitecore Experience Manager[11] é um sistema de gestão de conteúdo empresarial que proporciona um software para automatização de marketing multicanal. O sistema foi criado em 2001, fundado por Michael Seifert como presidente da empresa, recorre ASP.NET como linguagem de programação. *Sitecore* é uma plataforma de *Closed-Source*[12], significa que é necessário pagar pela licença tornando-se assim uma plataforma comercial.

Não é um sistema muito intuitivo de aprender inicialmente. Demora algum tempo a aprender as melhores práticas. Depois de trabalhar com a plataforma durante algum

[11] Sitecore. (n.d). About Sitecore . https://www.sitecore.com/company

[12] É um software de código-fechado, significa que o público não tem acesso ao código-fonte e os utilizadores precisam de uma licença válida e autenticada para usar o sistema.

tempo, encontrar novas características e formas de criar páginas e desenhá-las tornar-se-á intuitivo.

No requisito de flexibilidade, a plataforma *Sitecore* é flexível, os programadores podem reescrever com funcionalidades ou exportar funcionalidades personalizadas. Uma flexibilidade notável do *Sitecore* é a possibilidade de alterar itens nas versões em línguas locais. Por exemplo, a China não tem a possibilidade de utilizar YouTube, é possível inserir outro item apenas para que a página de língua chinesa possa visualizar um vídeo nessa página. No entanto, ao criar uma página predefinida, cria-se sempre em "Layout Partilhado", o que significa que esta página predefinida terá o mesmo layout em todas as línguas. Se estiver a editar e a personalizar a página chinesa por exemplo, terá de mudar de Layout Partilhado para Layout Final. Todas as versões linguísticas precisam de editar e personalizar no layout final para evitar texto chinês, de forma a aparecer na versão inglesa.

O *Sitecore* é um bom sistema quanto à personalização, pois é capaz de personalizar o layout, desenhar e configurar páginas inserindo e movendo componentes. É possível utilizar ainda uma agência digital para personalizar modelos para o *Sitecore*, uma vez que é preciso um pouco de conhecimento de codificação e tempo. Uma desvantagem do *Sitecore* na personalização é a necessidade de ter muito conhecimento em código se pretender criar modelos de anúncios.

O *Sitecore* em relação ao conhecimento de programação, se o utilizador possuir conhecimento em código HTML, poderá percorrer um longo caminho. Poderá desenhar botões e pequenas personalizações para ter um aspeto mais recatado do que o modelo oferece.

Em termos de segurança o *Sitecore* disponibiliza a opção de quem pode realizar a gestão de conteúdo e controlar quem tem os direitos de criar e editar determinadas páginas. Assim podemos ter um ambiente mais seguro e controlado dentro da plataforma e esta funcionalidade permite eliminar erros enormes. Por exemplo, apenas os utilizadores administrativos podem escrever por cima de páginas bloqueadas. No sistema, se alguém estiver a trabalhar em uma página, este deve bloqueá-la e desbloqueá-la quando o conteúdo da página estiver pronto. Por vezes, pode ser necessário criar itens em vários locais dentro do mapa de conteúdo e nem todos os utilizadores têm acesso a todo o

mapa. O utilizador é obrigado a contactar os utilizadores administrativos e este tem de criar o trabalho que o utilizador inicial pretendia fazer.

Tabela 2 - *Desvantagens e Vantagens do Sistema Sitecore*

Desvantagens	**Vantagens**
Não é intuitivo na fase de aprendizagem, demora-se algum tempo a aprender e a melhorar a prática.	Depois de algum treino com o sistema, tornar-se-á intuitivo configurar, criar páginas e desenhá-las.
Necessita de muito conhecimento em código se pretender criar *templates*.	É flexível, de uma versão é possível replicar para várias versões linguísticas.
Necessita de personalizações pesadas na sua configuração, requer conhecimentos de tecnologia de informação.	Conhecimento básico em HTML para desenhar botões para um display mais interativo.
Pouca documentação para ajudar novos utilizadores.	Interface de utilizador fácil de usar e intuitiva.
Uma licença "por utilizador" o que pode aumentar significativamente os custos.	Possibilidade de personalizar, desenhar e configurar páginas inserindo e movendo componentes.
Apenas os utilizadores administrativos podem subescrever itens bloqueados.	Controlo sobre editores, resulta em um ambiente mais seguro e controlado.

Joomla

O sistema *Joomla*[13] pertence à categoria de sistema de gestão de conteúdo de código aberto. Foi oficialmente lançado em 2005, fundado por um grupo de programadores, é desenvolvido em PHP e funciona em servidores Apache ou IIS. Conta com o suporte de diversas bases de dados como MySQL ou Azure. Este CMS é provavelmente um dos CMS's mais procurados, com a maior comunidade e recursos ((*Joomla! - Content Management System to Build Websites & Apps*, n.d.)).

[13] Joomla. (n.d).Joomla! - Content Management System to Build Websites & Apps. https://www.joomla.org/about-joomla.html

É mais complexo de navegar do que o *Wordpress*, mas é mais completo. Comparativamente ao *TYPO3* é de fácil navegação, mas tem muito menos características. Está posicionado neste lugar onde não é algo que se instala e que se possa simplesmente terminar um simples website num dia. Mesmo assim alguém com um pouco de experiência prévia em CMS pode fazer algo com relativa facilidade. Não é muito fácil de entender fora da caixa. No aspeto de navegação o *Joomla* tem um sistema bastante compreensível, consegue gerir diversas hierarquias e permite uma fácil gestão do site mesmo com centenas de subpáginas.

Joomla é um meio termo entre as ofertas mais leves como *Wordpress* e *Wix* e as opções mais pesadas como *TYPO3*. Com *plug-ins* suficientes e um pouco de experiência, pode criar algo com relativa facilidade. Fora da caixa, as suas opções são um pouco limitadas. No entanto, pode editar qualquer ficheiro no *back-end*[14] para satisfazer as suas necessidades.

O *Joomla* é uma das ferramentas CMS com mais suporte no mercado. Existem vários *plug-ins* para diversas funcionalidades. Com tantos *plug-ins* gratuitos, a qualidade e compatibilidade são as preocupações. Por vezes, não são suportadas para novas versões ou estão subdesenvolvidas. A solução é a procura de *plug-ins* pagos apenas com grandes avaliações e uma empresa bem estabelecida a apoiá-los.

Na personalização, o sistema *Joomla* contem muitos modelos que permitem editar cores e fontes com um editor visual. Dependendo dos seus *plug-ins*, poderá precisar de saber como editar JavaScript e PHP.

No painel para gestão da administração, o sistema providencia muitas funcionalidades que poderá ser intimidante no princípio, mas com o tempo é possível executar uma boa gestão e usar na sua capacidade máxima. Pode personalizar as definições de cache e outros aspetos que afetam a velocidade.

Em relação a documentos e suporte por parte da comunidade o *Joomla* tem uma base de utilizadores internacional gigante. O que significa que existem documentos oficiais de suporte, diversos artigos, vídeos e agências de programadores com muito conhecimento.

[14] É a estrutura que possibilita operar o sistema, enquanto o front-end é a parte visual, como apresentação, design, cores entre outros.

***Tabela 3** - Desvantagens Vantagens do Sistema Joomla*

Desvantagens	Vantagens
Não é fácil de entender para utilizadores iniciantes, tem muitas funcionalidades.	Navegação abrangente, de fácil gestão mesmo com centenas de subpáginas.
A compatibilidade entre *plug-ins*, necessidade de código para ter algumas funcionalidades.	Muitos *templates* permitem editar cores e fontes com o editor visual.
Gestão de Media é muito limitada, editar imagens, adicionar ficheiros, cortar e redimensionar.	Uma das ferramentas de CMS mais suportadas no mercado com plugins para tudo.
Suscetível a ataques de informática, o sistema foi atacado no mesmo dia que o site foi publicado.	Existem muitos documentos oficiais de apoio, diversos artigos, vídeos e fórum de discussão.

SAP Web Channel

O *SAP Web Channel Experience Management* (WCEM[15]) é um sistema antigo, uma das primeiras soluções que permitia às empresas transformar a internet em um canal de vendas e interações. Este sistema foi concebido e otimizado com o objetivo principal de possibilitar o e-business para processos SAP. Esta aplicação suporta a criação de ordens de venda em tempo real, preços dinâmicos, personalização, verificação de disponibilidade, restrições regionais e diferentes processos de pagamento.

O SAP WCEM está integrado com dados e processos que estão configurados no *back-end* dos sistemas de ERP[16] e CRM[17]. É uma solução que está totalmente integrada com SAP ERP/SAP CRM, todos os dados que estão disponíveis no SAP ERP estão automaticamente disponibilizados na loja web ((*Introduction and Overview to the New SAP Web Channel Experience Management | SAP Blogs*, n.d.)).

[15] SAP Web Channel. (2012, 14 Maio). (Introduction and Overview to the New SAP Web Channel Experience Management). https://blogs.sap.com/2012/05/14/introducion-and-overview-to-sap-web-channel-experience-management/

[16] Significa planeamento de recursos empresarias, estes softwares gerem as atividades de finanças, operações, relatórios, cadeia de fornecimento e recursos humanos de uma empresa.

[17] É um software de gestão de relacionamento com o Cliente, e se refere às práticas, estratégias de negócio e tecnologias que tem como foco o relacionamento com o cliente.

Não é necessário sincronizar alterações de lógica empresarial, uma vez que o SAP ERP e SAP CRM recorrem ao SAP WCEM para tal. Sempre que ocorrer a ativação de funcionalidades no SAP ERP/CRM estas aplicam-se automaticamente á solução de comércio eletrónico.

O sistema exibe todas as informações especificas do cliente, como encomendas, reclamações, pedidos de serviços, campanhas, pontos de fidelidade entre outros que são criados na SAP ECC/CRM ou na loja web.

Na apresentação de produto o sistema fornece ao cliente uma variedade de formas de apresentação do catálogo de produtos. O sistema tem capacidade de pesquisa com sugestão, proporciona ainda uma navegação intuitiva por todo o catálogo. Concede assim ao utilizador uma grande variedade de informação sobre os produtos, pode-se incluir imagens, anexos, ficheiros e textos html.

Os processos de venda, serviço e marketing que a SAP desenvolveu ficaram alinhados com os processos e dados dos sistemas de *back-end*. Tudo isto está envolvido em uma arquitetura WEB 2.0, para além da SAP ter prestado suporte para a resolução de problemas, este suporte terminou em 2021, tornando-se num sistema ultrapassado.

Tabela 4 - *Desvantagens e Vantagens do Sistema SAP Web Channel*

Desvantagens	**Vantagens**
Não é possível realizar qualquer tipo de personalização.	Tem a possibilidade de realizar as configurações do produto.
O acompanhamento da viagem do cliente a partir de A-Z não está disponível.	Gestão de cotações para confirmar quantidades de produtos, preços e entrega.
Não tem disponível nenhuma função para recomendações de produtos e serviços.	Catálogos específicos do cliente + condições.
Não contem produtos existentes dos outros dois sistemas, apenas dos produtos do sistema *TYPO3*.	Disponibilidades em tempo real dos produtos.
Dependente do legado CRM.	Contem a possibilidade de ver o

	seguimento de encomendas.
Não está ligado, nem existe forma de conexão ao Salesforce.	

2.2 Sistema destino dos dados - Adobe Experience Manager (AEM)

O *Adobe Experience Manager* (AEM[18]) pertence ao *Adobe Experience Cloud* e combina a gestão de ativos digitais (*Digital Asset Management* - DAM) com os poderes de um sistema de gestão de conteúdo. O AEM é um sistema de gestão de conteúdo alimentado por Inteligência Artificial (IA), é construído sobre uma base de agilidade, segura em uma fundação em nuvem (*cloud-native*). Para criar e gerir experiências digitais através da web, canais móveis e em outros canais de marketing independente do seu formato. O AEM possibilita aos seus utilizadores a criação de conteúdo, a gestão de atualizações com a reutilização de conteúdo e experiência de fragmentos (*Experience Fragments*).

O AEM contem um editor WYSIWYG[19], com estruturas de aplicação em páginas únicas (*Single Page Application*) que criam uma experiência mais natural, o carregamento total do resultado de páginas é executado em menos tempo, tudo disponibilizado pelo JavaScript SDK. O AEM baseia-se numa plataforma bastante consolidada, utilizando Java como linguagem de programação base, armazenamento e estrutura feita em JCR (*Java Content Repository*). Isto faz da plataforma um sistema altamente extensível ((*AEM Sites Content Management | Adobe for Business*, n.d.)).

A criação de páginas recorre ao uso de *templates* pré-definidos e uma biblioteca de componentes, é um ponto positivo do AEM. Para se obter o máximo de aptidão para os editores, é necessário um certo nível de desenvolvimento prévio para se criar os componentes e *templates*.

Com o AEM a capacidade de gestão de conteúdo é elevada, para que o utilizador não tenha muito trabalho o sistema oferece opções de criar componentes de conteúdo

[18] Adobe Experience Cloud. (n.d). (EM Sites Content Management | Adobe for Business. https://business.adobe.com/products/experience-manager/sites/aem-sites.html

[19] É uma sigla para a expressão de origem inglesa “What you see is what you get”, que em português significa “O que vê é o que obtêm”.

personalizados e *templates* para que seja mais simples. O AEM oferece infinitas opções para a criação de conteúdo, a principal vantagem desta plataforma, é a possibilidade de criar experiências de fragmentos (Experience Fragments[20]) para diversos canais de comunicação. Obrigando os programadores a criar muitas regras, *templates*, padrões e configurações para garantir que tudo funcione como o desejado.

Em termos de aprendizagem o AEM pode apresentar alguma dificuldade, especialmente se o utilizador em questão não tiver muita aptidão com tecnologias. No entanto, a plataforma oferece uma interface com um design de fácil utilização. Por ser uma plataforma com bastante flexibilidade, também é conhecida por proporcionar aos profissionais de marketing uma "dor de cabeça", devido às complexidades presentes a nível de experiência do utilizador (UX).

O AEM não é a plataforma ideal para empresas de pequena dimensão ou empresas que não tenham recursos internos para o desenvolvimento ou manutenção do código fonte. O mesmo não acontece com as plataformas mais simples, como *Wordpress*, onde a configuração, instalação de *plug-ins* é uma tarefa que um utilizador com conhecimento tecnológico consegue mesmo sem grande conhecimento técnico. No AEM para modificar, personalizar e adicionar requer alguma alteração no código ou pelo menos um conhecimento mais profundo em algumas tecnologias.

Em relação à segurança a plataforma tem o assunto com seriedade, as melhores práticas são passadas para o software para que as equipas possam prevenir, detetar e rapidamente eliminar qualquer ameaça ou desvio. No AEM é possível atribuir às contas dos seus utilizadores propriedades para a publicação de conteúdo, desta forma apenas alguns utilizadores têm disponível a propriedade para publicação, o que torna a possibilidade de erro baixa. O sistema permite dar a cada utilizador o seu próprio espaço de trabalho para os seus projetos. Esta característica permite menos painéis de controlo (dashboards) desordenados, permite ainda ver os níveis de produtividade, o que evita o risco de colisão entre equipas na interface. Todos sabem quais são as suas tarefas nos diferentes grupos de trabalho, o que reforça a estrutura organizacional, para cada membro é atribuído um papel específico.

[20] É um conjunto de componentes quando combinados criam experiências, por exemplo, um título, imagem, descrição e botão.

É possível encontrar diversos documentos de apoio sobre o AEM, vídeo explicativos, informações sobre atualizações, cursos e apoio da comunidade AEM em fórum.

Tabela 5 - *Desvantagens e Vantagens do Sistema AEM*

Desvantagens	Vantagens
Um utilizador com pouco conhecimento técnico, pode demorar muito tempo a aprender, é uma plataforma altamente flexível e pode confundir.	DAM com uma vasta variedade de funções, pesquisa visual, marcação inteligente para armazenar, gerir e encontrar rapidamente ativos.
Exige o acompanhamento de um serviço terceirizado fortemente envolvido na instalação inicial e nas operações diárias.	Integração mais fácil com outras ferramentas de marketing.
Um custo alto, devidos às funcionalidades, a grande base de dados para gerir e serviços externos.	Cada membro recebe um papel atribuído na estrutura o que facilita o trabalho entre diferentes grupos.
Manutenção por parte dos serviços terceirizados forçando a depender sempre que surgem problemas.	Controlo total sobre a forma como o conteúdo é apresentado nos diversos canais.
	Componentes *out-of-the-box* incluem etapas de fluxo de trabalho, *templates* padrão, entre muitos outros.
	Oferece uma pré-visualização do conteúdo que os utilizadores desenvolvem, possibilidade de ver as páginas como um visitante.

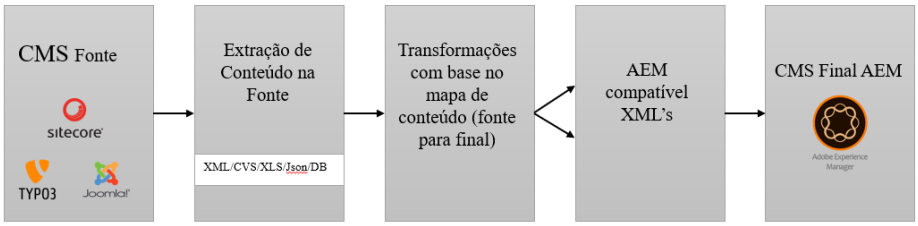
2.3 Estratégia adotada para migração

Para solucionar o Problema a Resolver apresentado no subcapítulo da Introdução, a organização desenvolveu um projeto, focado na migração de conteúdo de vários sistemas de gestão de conteúdo para apenas um único sistema. Para avançar com este projeto a organização teve de adotar uma estratégia para realizar esta migração, uma vez que os conteúdos de três sites teriam de ser consolidados em apenas uma única solução.

Depois de um estudo aprofundado da organização a estratégia escolhida para aplicar durante o processo foi ETL. A empresa contratada para realizar este processo, em conjunto com a equipa interna do projeto, utilizam uma ferramenta desenvolvida pela primeira. Esta ferramenta de migração de dados segue o modelo ETL. A plataforma consegue recuperar conteúdo não estruturado de qualquer fonte (ECM, DMS e CMS). O conteúdo ETL refere-se aos processos que são utilizados para trocar dados entre múltiplos repositórios de conteúdo. Depois de investigar a organização decidiu a era a melhor opção para resolver o problema em mãos.

Na **Erro! A origem da referência não foi encontrada.**, é possível observar o processo de migração, que conta com a identificação dos sistemas origem (*TYPO3*, *Sitecore* e *Joomla*), as fases principais do processo ETL, extração, as transformações e o carregamento de dados para o novo sistema AEM.

***Figura 12**: Diagrama do processo ETL adaptado ao projeto do Estudo de Caso*



Seguidamente, apresenta-se a planificação utilizada na realização da migração de conteúdo. Em cada fase existem marcos que necessitam de ser entregues para continuar e seguir para o próximo passo.

1. Início
2. Instalação e Configuração
3. Extração
4. Análise
5. Seleção e Mapeamento
6. Desenho da Migração de Conteúdo
7. Transformação
8. Testes de Importação
9. Importar para produção
10. Encerramento

Na fase inicial, é delineada a abordagem para realizar a análise e o plano de migração. Logo após a descrição detalhada das atividades, seguem as contribuições necessárias que cada organização precisa de executar e os resultados a apresentar. A abordagem ao processo de migração de dados consiste nas atividades de 1 a 6: 1. Início; 2. Instalação e Configuração; 3. Extração; 4. Análise; 5. Seleção e Mapeamento; 6. Desenho da Migração de Conteúdo.

Segue-se uma explicação individual de cada uma das atividades mencionadas anteriormente em detalhe.

1. **Início**

Na fase inicial é discutido e inspecionado o plano inicial com a equipa. Os elementos da equipa do projeto são essenciais a estarem presentes para que sejam introduzidos e conectados entre si. É apresentado todos os pré-requisitos necessários para a instalação e configuração da ferramenta, como partilhar acessos aos sistemas. Quando esta atividade estiver concluída, a equipa é reunida, todas as ações necessárias para a fase seguinte de instalação são definidas, esclarecidas para que o plano do projeto seja aprovado e todos terem conhecimento dos próximos passos. Durante esta fase, é criado um conjunto de regras, procedimentos que determinam como irá ocorrer a gestão do projeto e a supervisão.

2. **Instalação e configuração do Software**

Nesta fase, o software é instalado são realizadas as configurações dos conectores aos sistemas de origem e destino, a empresa externa responsabiliza-se pela instalação. As atividades a serem realizadas nesta etapa são:

- Instalação do software de migração;
- Estabelecer e configurar os conectores de extração nos sistemas de origem e testar o desempenho da extração;
- Implementar e testar o acesso remoto (de acordo com as políticas de segurança);
- Configurar os conectores no sistema de destino e testar o desempenho da importação;
- Estabelecer a estrutura e regras da migração.

Os conectores são para a extração nos sistemas de origem, no caso são Sitecore, *TYPO3* e *Joomla*. Os conectores são aprimorados para utilizar os recursos disponíveis,

maximizar o desempenho, ajudar a identificar potenciais obstáculos no início do projeto e auxiliam no processo para tomar medidas em conformidade.

3. **Extração**

Depois de os conectores estarem estabelecidos e testados, inicia-se a extração efetiva dos dados dos sistemas de origem (*Sitecore*, *TYPO3* e *Joomla*). Durante a extração, o software extrai os metadados para o modelo de dados unificado (UDM) utilizado pela empresa. O modelo de dados unificado, possibilita a criação de todos os conhecimentos necessários para a próxima fase. São criados e extraídos relatórios de forma padronizada, que são fundamentais para analisar a informação existente e chegar a um desenho de migração mais detalhado.

4. **Análise**

Na quarta etapa é analisado a estrutura, quantidade e qualidade dos dados dos sistemas de origem. Para identificar valores anómalos e padrões que requerem atenção. São definidas as regras de negócio para a migração, posteriormente registadas no desenho da migração. Durante a análise abordam-se funcionalidades específicas como formulários, tabelas, imagens e todos os dados extraídos na etapa anterior. Em resumo, nesta etapa são recolhidos dados sobre o aspeto da migração (em termos de seleção, mapeamento e enriquecimento) e estabelecemos um plano detalhado e o desenho da migração a partir daí.

5. **Seleção e mapeamento**

Nesta fase, são determinadas as regras de negócio para estabelecer o âmbito da migração e mostrar o impacto destas decisões. São estabelecidos modelos de mapas para a estrutura, dados, valores e permissões para capturar as regras, quando os dados alvo do modelo são conhecidos.

Na seleção é determinado o conteúdo com mais valor e decide-se o conteúdo que necessita de ser transferido para o novo sistema. É possível excluir conteúdos duplicados, não publicados e sem hiperligações.

Para o mapeamento são definimos modelos para facilitar as diferentes necessidades durante a próxima etapa:

- Mapeamento da estrutura – preenchimento do modelo para mapear a estrutura do sistema *Sitecore* e *TYPO3* e *Jommla*.
- Mapeamento de metadados - mapear os metadados actuais e metadados enriquecidos para o novo modelo. Os metadados associados às páginas, podem ser utilizados para decisões de migração sobre a localização do alvo e as propriedades dos metadados.
- Mapeamento de valores - mapear os valores (listas válidas) dos campos de metadados.

6. **Desenho da migração de conteúdo**:

A conceção da migração de conteúdos, inclui regras estabelecidas com a organização e registadas no desenho de migração:

- Descrição completa de todas as regras de negócio, modelos e processo de migração;
- Resumo da abordagem de migração por tipo de conteúdo *Sitecore* e tipo de conteúdo *TYPO3*;
- Plano detalhado para a implantação da migração e atividades por local;
- Guia detalhado para migração por ambiente;
- Panorama dos papéis e responsabilidades;
- Relatórios e estrutura de administração.

O conteúdo de cada atividade (7 a 10) depende das decisões realizadas e do resultado das fases de 1 a 6. Apenas é possível efetuar uma estimativa aproximada do esforço e dos prazos necessários para as próximas atividades. A abordagem de migração de dados consiste na execução das atividades: 7. Transformação; 8. Testes de Importação; 9. Importar para produção; 10. Encerramento.

7. Transformação

A etapa para transformação de dados é uma das atividades mais importantes, os dados já extraídos das fontes, são reunidos, convertidos e limpos para voltarem a ser introduzidos na nova base de dados.

São executadas uma série de funções e aplicação de diversos conjuntos de regras aos dados que foram extraídos. Para que estes sejam convertidos para o novo formato padrão de modo a se enquadrar nos requisitos da base de dados final.

Devido às necessidades do negócio, o nível de manipulação de dados na transformação ETL é bastante complexa, uma vez, que existem três base de dados fonte diferentes, cada uma com as suas características, arquitetura e organização. O passo de transformação depende dos dados extraídos, inclui a sua validação e a sua rejeição caso não sejam aceitáveis.

Como referido, existem três fontes de dados diferentes exigindo muitas transformações. No entanto, é possível aplicar várias técnicas de transformação para satisfazer os requisitos técnicos e comerciais da base de dados alvo.

Os passos necessários a seguir para que a etapa de transformação seja realizada são:

1) Converter os dados consoante as exigências do negócio;
2) Formatar os dados convertidos para o novo formato padrão;
3) Limpar dados que sejam irrelevantes dentro do conjunto de dados;
4) Ordenar e filtrar dados;
5) Limpar informações duplicadas;

8. Teste de Importação

Antes de ser realizada a importação para produção são realizados testes, este processo assegura que a transferência de dados de fontes heterogéneas para o DW, ocorre com estrita adesão às regras de transformação e está consoante todas as provas de validade.

O teste de importação pode ser denominado teste de *end-to-end* (E2E[21]), durante estes testes toda a aplicação é testada num ambiente que simula o ambiente de produção. Como a comunicação com outros sistemas, rede, base de dados, interfaces, são testadas contra os volumes de dados.

Com este teste podemos identificar problemas e assegurar que o processo está pronto para ser dimensionado com os volumes crescentes de dados. Para realizar esta validação é necessário verificar:

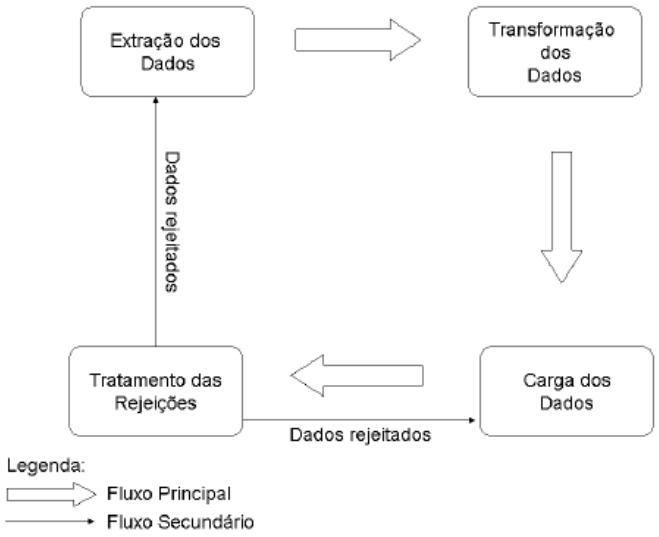
- Se as cargas e a consulta de dados são executadas dentro do período esperado;
- O volume máximo de dados foi carregado dentro do tempo estipulado;
- Qual os tempos de carga para diferentes volumes de dados.

[21] Testes E2E, ou teste de ponta a ponta, também chamados testes de integração de dados.

Mesmo com a análise e limpeza de dados nas etapas anteriores, podem ocorrer falhas, é necessário identificar a causa do erro, para que possa ser resolvido. Por exemplo, caso seja um erro técnico, a equipa externa é a responsável pela sua resolução. Caso o erro esteja relacionado com conteúdo, os indivíduos responsáveis pela gestão daquele sistema devem ser contactados para a sua resolução.

Na **Figura 13**, é possível observar como é realizado a migração de dados, quando existem erros e como são realizadas as correções, este ciclo pode se repetir inúmeras vezes.

***Figura 13**: Ciclo da migração de dados*



Nota: Representação do ciclo de migração de dados com o fluxo principal e secundário de dados by Mendonça, 2009.

9. Importar para produção

A etapa final do processo de ETL é o carregamento de dados para a base de dados final. Este passo pode ser realizado de duas formas: migrar de uma só vez todo o conteúdo, migrar passo a passo e por sincronização.

Na primeira abordagem migrar de uma só vez, todo o conteúdo é migrado apenas numa única interação. Antes da migração, vem a realização de alguns testes, assim como as avaliações, mas a migração final é realizada de uma única vez. Esta abordagem é conhecida como estratégia *Big Bang*, a organização optou por adotar esta forma de migração.

Na segunda abordagem, migração passo a passo, o sistema origem será divido e a migração será realizada por secção ou tipo de conteúdo. Esta abordagem permite tempo extra para a realização de testes de migração e proporciona mais tempo para dar formação à organização. A sincronização acontece enquanto os sistemas de origem tiverem ainda disponíveis durante um determinado período em paralelo com o novo sistema. Qualquer alteração realizada nos sistemas de origem será automaticamente migrada e publicada no novo sistema. Esta segunda abordagem é conhecida como *Trickel*.

10. Encerramento

O encerramento do projeto deu-se com a elaboração do relatório, para verificar o esquema, opções, filtros e funcionalidades da exportação. Este relatório permite às partes interessadas saber detalhes e resultados dos processos de teste, se alguma etapa estava fora do âmbito e as razões para que isso tenha acontecido.

2.4 Processo de Migração

Nesta secção são apresentadas todas as etapas percorridas durante o processo de migração de dados delineados na secção anterior, com enfase na estratégia, no planeamento e como foram executadas.

O âmbito original do projeto era a criação de um website que fornecesse, numa primeira instância, informações relacionadas com as aplicações e a indústria dos seus produtos/serviços para o público. Para servir como a face digital, unificada para o cliente e todas as marcas do grupo. Em segundo lugar, a criação de um conjunto de páginas de conteúdo focado na indústria com possibilidade para geração de leads. Por último, acrescentar informação, valor ao grupo e com possibilidades de venda cruzada aos websites existentes.

Com a mudança de gestão da organização realizaram-se alterações ao projeto, como o levantamento dos domínios existentes (três domínios diferentes), para que a marca seja alinhada e a reputação da marca existente possa estar estreitamente associada a uma única. Navegação clara, consistente, informação mais convincente sobre os produtos e sólida. Desenvolvimento do domínio para alinhar a visão com as soluções, publicar conteúdo enriquecido que reflita o "tom" da marca e o valor comercial.

O objetivo final da migração era assegurar que todo o conteúdo relevante fosse migrado para o nosso novo Sistema de Gestão de Conteúdos, *Adobe Experience Manager*. Para assegurar que a organização mantinha o nível de tráfego para as páginas, evitar a perda de toda a autoridade de domínio construída ao longo do tempo, a migração seguiu a estratégia empresarial definida pela liderança. Aplicou as melhores práticas de SEO, ao executar rastreamento pré e pós-migração, incluindo no novo ambiente para garantir que todas as páginas criadas funcionam da mesma forma que antes, sem ligações partidas e sem metadados em falta.

2.4.1 Início

Na primeira fase do projeto existiram reuniões semanais para determinar as atividades da migração de dados. No entanto, estas reuniões tornaram-se longas, demoradas, sem a finalização de algumas discussões, devido ao envolvimento de elementos de outro departamento, o que não se mostrou eficiente. Realizou-se uma separação, criou-se uma equipa exclusivamente dedicada ao projeto de migração de dados, quando necessário o conhecimento de outros indivíduos estes iriam ajudar nas atividades que lhe cabiam.

Assim, definiram-se reuniões semanais com a exclusiva finalidade para a migração de dados, geridas pelo líder do projeto. Com esta organização, o grupo ganhou tempo, o trabalho era exclusivamente focado no projeto, tornando o grupo de trabalho produtivo. O projeto teve o seu início em junho de 2021.

Uma das atividades na fase um, era o desenvolvimento de um plano da migração, uma vez que é o produto final que se pretende obter com o fim desta fase. Conforme as reuniões aconteciam o plano teve várias alterações na duração do projeto. Manter o plano de migração atualizado revelou-se ser uma tarefa por vezes esquecida. O plano de migração foi realizado para abordar todo o projeto.

Foi necessário rever o foco da migração, cujo propósito era migrar os dados dos sistemas origem (*TYPO3*, *Sitecore*, *Joomla*) para o novo sistema AEM. Os dados que estavam nos sistemas origem deveriam sofrer uma transformação com objetivo de atingir os critérios definidos. Estimou-se a migração de milhares de páginas dos sistemas de origem.

Para além de rever o foco da migração, é essencial validar a estratégia da migração. Durante esta atividade foi definido que devido ao tipo de negócio e ao volume de dados

necessários a serem migrados, os sistemas origem continuariam o seu funcionamento. Foi definido que a migração ocorreria em ciclos (ciclo 1 – sistema *Joomla* para AEM; ciclo 2 – sistema *Sitecore* para AEM; ciclo 3 – sistema *TYPO3* para AEM), no final de cada ciclo os dados seriam testados, aprovados e no caso de existirem dados rejeitados estes teriam de ser ajustados. Nesta fase, não era possível ter uma estimativa precisa do tempo para cada ciclo de migração de dados.

Como referido anteriormente, neste tópico é necessário definir a equipa que será responsável pela execução do projeto. O gestor de projeto definiu desde o início do projeto, quais os indivíduos principais do projeto e a sua dedicação exclusiva ao projeto.

A equipa ficou então definida da seguinte forma:

- Elemento 1 - Líder do projeto de migração;
- Elemento 2 - Responsável pela comunicação com a equipa contratada;
- Elemento 3 - Responsável pela migração de ficheiros de *media* e integração do novo sistema DAM;
- Elemento 4 - Programador: responsável por todo o desenvolvimento do novo sistema AEM;
- Elemento 5 - Responsável pelo mapeamento do conteúdo no novo sistema;
- Equipa Externa – empresa contratada responsável pela ferramenta e processo de migração de dados.

Como em todos os projetos foi realizado uma avaliação e identificados os riscos associados ao projeto. Entre os riscos, o principal era a subestimação do tempo necessário para a migração de dados. A qualidade dos dados dos sistemas origem apenas seria conhecida na etapa 4 de análise e não existia uma estimativa precisa do tempo necessário para a limpeza de dados. Outro risco apontado, era a possível dificuldade na sincronização da migração de dados no sistema destino, com problemas de rejeição dos dados, o que poderia ter um impacto negativo para o projeto.

Na **Figura 14** está uma representação das atividades realizadas nesta etapa com o período correspondente.

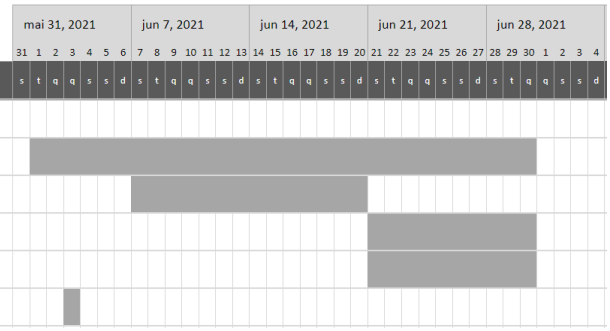
***Figura 14**: Cronograma da primeira etapa – Início*

Projeto Harmony

Projeto Inicia: ter, 6/1/2021

Por Semana: 1

Atividades	Progresso	Início	Fim
Inicio			
Desenvolver plano da migração com datas definitivas	100%	6/1/21	6/30/21
Foco da migração e validação da estratégia	100%	6/7/21	6/20/21
Definição das atividades da migração	100%	6/21/21	6/30/21
Definição dos recursos necessários	100%	6/21/21	6/30/21
Definição de comunicações entre equipas (semanal)	100%	6/3/21	6/3/21



2.4.2 Instalação e configuração

Na segunda fase do processo, apenas foi necessária a equipa Externa - responsável pela migração de dados do sistema origem para o sistema destino e o indivíduo 2, sendo o que teria o conhecimento mais abrangente de todos os sistemas.

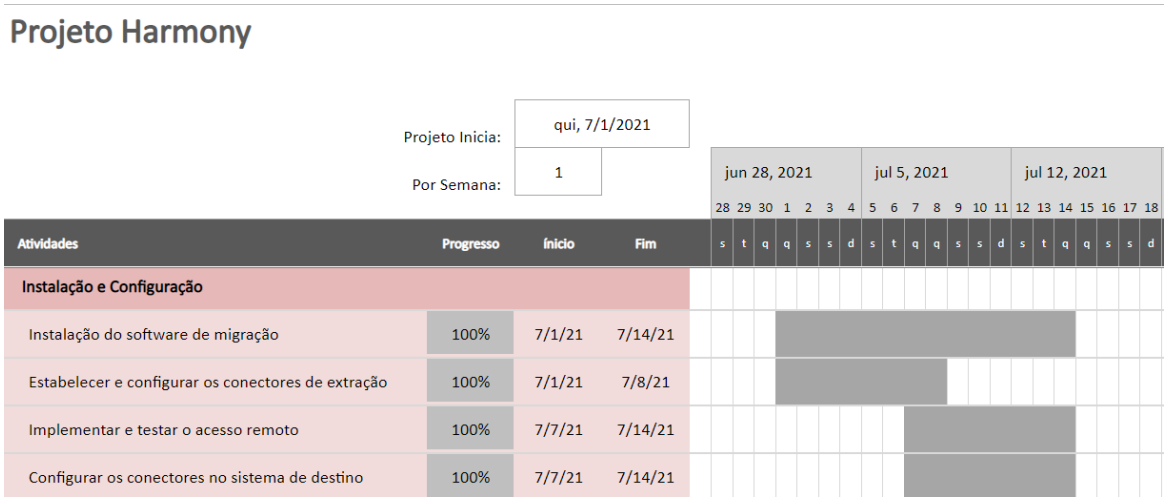
Foram dados acessos à equipa externa para que esta pudesse estabelecer, configurar os seus conectores de extração em cada um dos sistemas origem e realizar teste de desempenho da extração. Para além de serem estabelecidos os conectores no sistema origem, foram configurados conectores no sistema destino e neste caso testar qual o desempenho da importação de dados. Na **Figura 15**, é possível visualizar um exemplo usado para configurar o sistema *Sitecore*, para utilizar o conector neste sistema foi necessário instalar um plugin dentro do *Sitecore*.

***Figura 15**: Script exemplo para configuração no sistema Sitecore*

```
{
  "configurationType": "Sitecore",
  "config" : {
    "baseUrl" : "https://sitecore.mydomain.com",
    "username" : "username",
    "password" : "password",
    "database" : "master",
    "domain" : "sitecore"
  }
}
```

Na **Figura 16**, estão disponíveis as atividades necessárias para a conclusão da fase 2. Esta fase teve início no mês de julho de 2021 e necessitou de dez dias úteis para a sua conclusão.

***Figura 16**: Cronograma da segunda etapa – Instalação e Configuração*



2.4.3 Extração

Esta fase foi realizada nos sistemas de origem, cada uma das extrações foi agendada. A extração é considerada uma atividade de alguma dificuldade, uma vez que existiram alguns constrangimentos técnicos. Um dos constrangimentos está relacionado com a criação de canais de acesso, uma vez que a empresa contratada não tinha conectores específicos para o *TYPO3*. Foi necessário dar acesso à equipa para que pudesse ter acesso ao banco de dados. Outro constrangimento, estava relacionado com conteúdo, uma vez que existia conteúdo utilizado apenas durante um pequeno período, páginas que não foram publicadas e não estavam disponíveis na vista do visitante do site.

O conector utilizado está escrito numa linguagem de programação desenvolvida pela equipa externa, a **Figura 17**, é um exemplo de script utilizado pela equipa. Tem de ser preenchido com as variáveis de cada website, as informações da organização foram ocultadas.

***Figura 17**: Script exemplo utilizado na extração de dados*

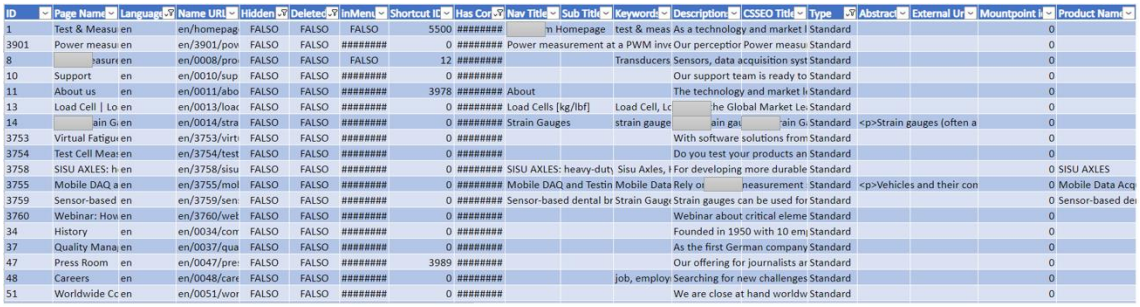
```
#udm
db.name=udm_documentum
udm.identity=documentum

#paths
project.root=D:/projects/connector_documentum/
project.exportPath=D:/projects/connector_documentum/export/
project.tempPath=D:/projects/connector_documentum/temp/
project.binaryPath=D:/projects/connector_documentum/export/binaries/
project.imagePath=D:/projects/connector_documentum/export/images/

#export

source.apiUrl1=http://localhost/dctm-rest/
source.username1=dmadmin
source.password1=password
#set how permissions will be resolved. Accepted parameters are  "DQL" and "basic"
source.permissionsType=DQL
source.extensionFile=fileextensions.sql
source.rootName1=MyRepo
source.root1=MyRepo
```

A própria ferramenta faz a extração do conteúdo dos sistemas de origem, são convertidos para um modelo unificado de dados (UDM) que pode ser considerada uma forma de transformação. Todos os sistemas de gestão de conteúdo contêm uma forma para realizar exportação do conteúdo, seja em documentos individuais, lista de documentos, seja um ficheiro Excel, ou em formatos mais utilizados XML Json.

Como os sites têm a sua percentagem de complexidade, era necessário saber a forma estruturada do conteúdo que está dentro daquelas páginas, por exemplo, que imagem é, onde está, qual o nome da imagem e onde está armazenada.

Durante a extração, todas as páginas e conteúdo foram extraídos, todos os metadados que estivessem associados (como datas, estado de publicação e língua), campos e estruturas (campos que as páginas contêm, formatos, listas de valor, dependências e restrições), relações (conteúdos interligados e a que conteúdo se refere), local (onde aquele item se encontra no site, se faz parte da navegação e se este se repete). A **Figura 18** é um exemplo de uma das extrações realizadas do *TYPO3*, alguns dados foram retirados para manter o sigilo do nome da organização.

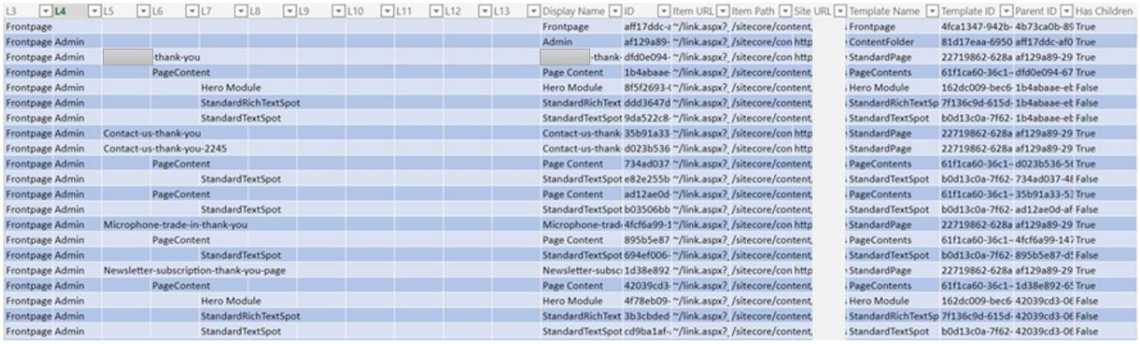
***Figura 18**: Extração realizada ao sistema TYPO3*



ID	Page Name	Language	Name URL	Hidden	Deleted	inMenu	Shortcut ID	Has Cor	Nav Title	Sub Title	Keywords	Description	CSSEO Title	Type	Abstract	External Url	Mountpoint I	Product Name
1	Test & Measu	en	en/homepag	FALSO	FALSO	FALSO	5500	########	[redacted]m Homepage		test & meas	As a technology and market l		Standard			0	
3901	Power measu	en	en/3901/pov	FALSO	FALSO	########	0	########	Power measurement at a PWM inv			Our perceptior	Power measu	Standard			0	
8	[redacted]asuri	en	en/0008/pro	FALSO	FALSO	FALSO	12	########			Transducers	Sensors, data acquisition syst		Standard			0	
10	Support	en	en/0010/sup	FALSO	FALSO	########	0	########				Our support team is ready to		Standard			0	
11	About us	en	en/0011/abo	FALSO	FALSO	########	3978	########	About			The technology and market l		Standard			0	
13	Load Cell | Lo	en	en/0013/loac	FALSO	FALSO	########	0	########	Load Cells [kg/lbf]		Load Cell, Lc	[redacted]he Global Market Le		Standard			0	
14	[redacted]ain G	en	en/0014/stra	FALSO	FALSO	########	0	########	Strain Gauges		strain gauge	[redacted]ain ga[redacted]ain G		Standard	<p>Strain gauges (often a		0	
3753	Virtual Fatigu	en	en/3753/virt	FALSO	FALSO	########	0	########				With software solutions from		Standard			0	
3754	Test Cell Mea	en	en/3754/test	FALSO	FALSO	########	0	########				Do you test your products an		Standard			0	
3758	SISU AXLES: h	en	en/3758/sisu	FALSO	FALSO	########	0	########	SISU AXLES: heavy-duty		Sisu Axles, I	For developing more durable		Standard			0	SISU AXLES
3755	Mobile DAQ a	en	en/3755/mol	FALSO	FALSO	########	0	########	Mobile DAQ and Testin		Mobile Data	Rely o[redacted]neasurement		Standard	<p>Vehicles and their con		0	Mobile Data Acq
3759	Sensor-based	en	en/3759/sen	FALSO	FALSO	########	0	########	Sensor-based dental br		Strain Gauge	Strain gauges can be used for		Standard			0	Sensor-based de
3760	Webinar: Hov	en	en/3760/wel	FALSO	FALSO	########	0	########				Webinar about critical eleme		Standard			0	
34	History	en	en/0034/con	FALSO	FALSO	########	0	########				Founded in 1950 with 10 em		Standard			0	
37	Quality Mana	en	en/0037/qua	FALSO	FALSO	########	0	########				As the first German company		Standard			0	
47	Press Room	en	en/0047/pre	FALSO	FALSO	########	3989	########				Our offering for journalists ar		Standard			0	
48	Careers	en	en/0048/car	FALSO	FALSO	########	0	########			job, employ	Searching for new challenges		Standard			0	
51	Worldwide Cc	en	en/0051/wor	FALSO	FALSO	########	0	########				We are close at hand worldw		Standard			0	

A **Figura 19** é um exemplo da extração realizada ao sistema *Sitecore*, a extração ocorreu na mesma semana em paralelo com a extração realizada ao *TYPO3*.

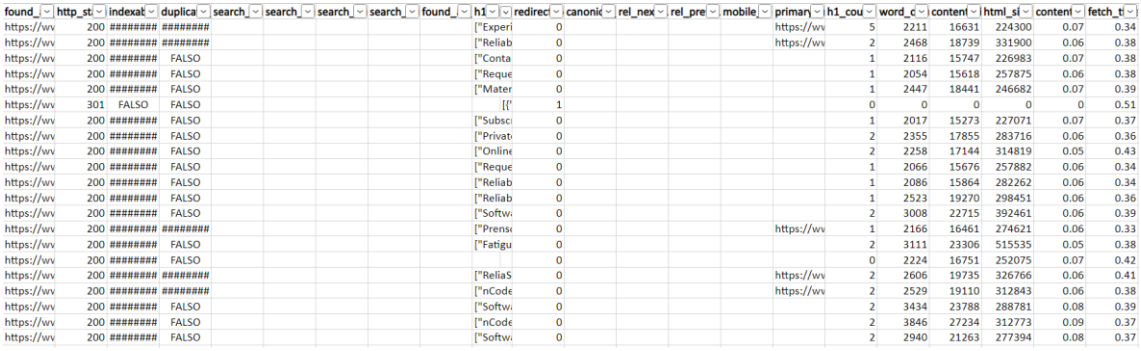
***Figura 19**: Extração realizada ao sistema Sitecore*



L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	Display Name	ID	Item URL	Item Path	Site URL	Template Name	Template ID	Parent ID	Has Children
Frontpage											Frontpage	aff17ddc-a	~/link.aspx?_	/sitecore/content		Frontpage	4fca1347-942b-	4b73ca0b-85	True
Frontpage	Admin										Admin	af129a89-	~/link.aspx?_	/sitecore/con	http	ContentFolder	81d17eaa-6950	aff17ddc-af0	True
Frontpage	Admin	[redacted]	thank-you								[redacted]-thank-	dfd0e094-	~/link.aspx?_	/sitecore/con	http	StandardPage	22719862-628a	af129a89-29	True
Frontpage	Admin		PageContent								Page Content	1b4abaae-	~/link.aspx?_	/sitecore/content		PageContents	61f1ca60-36c1-	dfd0e094-67	True
Frontpage	Admin			Hero Module							Hero Module	8f5f2693-(	~/link.aspx?_	/sitecore/content		Hero Module	162dc009-bec6	1b4abaae-e(	False
Frontpage	Admin			StandardRichText	Spot						StandardRichText	ddd3647d	~/link.aspx?_	/sitecore/content		StandardRichTextSp	7f136c9d-615d-	1b4abaae-e(	False
Frontpage	Admin			StandardTextSpot							StandardTextSpot	9da522c8-	~/link.aspx?_	/sitecore/content		StandardTextSpot	b0d13c0a-7f62-	1b4abaae-e(	False
Frontpage	Admin	Contact-us-thank-you									Contact-us-thank-	35b91a33	~/link.aspx?_	/sitecore/con	http	StandardPage	22719862-628a	af129a89-29	True
Frontpage	Admin	Contact-us-thank-you-2245									Contact-us-thank-	d023b536	~/link.aspx?_	/sitecore/con	http	StandardPage	22719862-628a	af129a89-29	True
Frontpage	Admin		PageContent								Page Content	734ad037	~/link.aspx?_	/sitecore/content		PageContents	61f1ca60-36c1-	d023b536-5(	True
Frontpage	Admin			StandardTextSpot							StandardTextSpot	e82e255b	~/link.aspx?_	/sitecore/content		StandardTextSpot	b0d13c0a-7f62-	734ad037-4(	False
Frontpage	Admin		PageContent								Page Content	ad12ae0d	~/link.aspx?_	/sitecore/content		PageContents	61f1ca60-36c1-	35b91a33-5(	True
Frontpage	Admin			StandardTextSpot							StandardTextSpot	b03506bb	~/link.aspx?_	/sitecore/content		StandardTextSpot	b0d13c0a-7f62-	ad12ae0d-af	False
Frontpage	Admin	Microphone-trade-in-thank-you									Microphone-trad	4fcf6a99-1	~/link.aspx?_	/sitecore/con	http	StandardPage	22719862-628a	af129a89-29	True
Frontpage	Admin		PageContent								Page Content	895b5e87	~/link.aspx?_	/sitecore/content		PageContents	61f1ca60-36c1-	4fcf6a99-147	True
Frontpage	Admin			StandardTextSpot							StandardTextSpot	694ef006-	~/link.aspx?_	/sitecore/content		StandardTextSpot	b0d13c0a-7f62-	895b5e87-d(	False
Frontpage	Admin	Newsletter-subscription-thank-you-page									Newsletter-subsc	1d38e892	~/link.aspx?_	/sitecore/con	http	StandardPage	22719862-628a	af129a89-29	True
Frontpage	Admin		PageContent								Page Content	42039cd3-	~/link.aspx?_	/sitecore/content		PageContents	61f1ca60-36c1-	1d38e892-6(	True
Frontpage	Admin			Hero Module							Hero Module	4f78eb09-	~/link.aspx?_	/sitecore/content		Hero Module	162dc009-bec6	42039cd3-0(	False
Frontpage	Admin			StandardRichText	Spot						StandardRichText	3b3cbded	~/link.aspx?_	/sitecore/content		StandardRichTextSp	7f136c9d-615d-	42039cd3-0(	False
Frontpage	Admin			StandardTextSpot							StandardTextSpot	cd9ba1af-	~/link.aspx?_	/sitecore/content		StandardTextSpot	b0d13c0a-7f62-	42039cd3-0(	False

E por último, a **Figura 20** é um exemplo da extração do *Joomla*, algumas informações foram retiradas para manter a confidencialidade.

***Figura 20**: Extração realizada ao sistema Joomla*



found_	http_st	indexat	duplica	search_	search_	search_	search_	found_	h1	redirec	canonic	rel_nex	rel_pre	mobile	primar	h1_cou	word_c	content	html_si	content	fetch_t
https://wv	200	########	########						["Experi	0					https://wv	5	2211	16631	224300	0.07	0.34
https://wv	200	########	########						["Reliab	0					https://wv	2	2468	18739	331900	0.06	0.38
https://wv	200	########	FALSO						["Conta	0						1	2116	15747	226983	0.07	0.38
https://wv	200	########	FALSO						["Reque	0						1	2054	15618	257875	0.06	0.38
https://wv	200	########	FALSO						["Mater	0						1	2447	18441	246682	0.07	0.39
https://wv	301	FALSO	FALSO						[{"	1						0	0	0	0	0	0.51
https://wv	200	########	FALSO						["Subsc	0						1	2017	15273	227071	0.07	0.37
https://wv	200	########	FALSO						["Privat	0						2	2355	17855	283716	0.06	0.36
https://wv	200	########	FALSO						["Online	0						2	2258	17144	314819	0.05	0.43
https://wv	200	########	FALSO						["Reque	0						1	2066	15676	257882	0.06	0.34
https://wv	200	########	FALSO						["Reliab	0						1	2086	15864	282262	0.06	0.34
https://wv	200	########	FALSO						["Reliab	0						1	2523	19270	298451	0.06	0.36
https://wv	200	########	FALSO						["Softwa	0						2	3008	22715	392461	0.06	0.39
https://wv	200	########	########						["Prens	0					https://wv	1	2166	16461	274621	0.06	0.33
https://wv	200	########	FALSO						["Fatigu	0						2	3111	23306	515535	0.05	0.38
https://wv	200	########	FALSO							0						0	2224	16751	252075	0.07	0.42
https://wv	200	########	########						["ReliaS	0					https://wv	2	2606	19735	326766	0.06	0.41
https://wv	200	########	########						["nCode	0					https://wv	2	2529	19110	312843	0.06	0.38
https://wv	200	########	FALSO						["Softwa	0						2	3434	23788	288781	0.08	0.39
https://wv	200	########	FALSO						["nCode	0						2	3846	27234	312773	0.09	0.37
https://wv	200	########	FALSO						["Softwa	0						2	2940	21263	277394	0.08	0.37

Após as extrações serem realizadas foi criado um relatório e inventário com a extração de dados. Posteriormente, foi partilhado com a equipa interna (organização), na **Figura 21** é possível visualizar as atividades associadas à Extração.

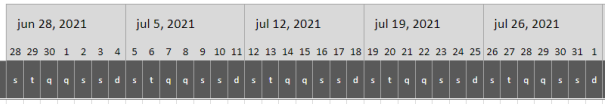
***Figura 21**: Cronograma da terceira etapa – Extração*

Projeto Harmony

Projeto Inicia: qui, 7/1/2021

Por Semana: 1

Atividades	Progresso	Início	Fim
Extração			
Extração de dados (Joomla)	100%	7/15/21	7/21/21
Extração de dados (TYPO3)	100%	7/22/21	7/29/21
Extração de dados (Sitecore)	100%	7/22/21	7/29/21
Relatório de extração	100%	7/15/21	7/29/21



2.4.4 Análise

Para a equipa interna da nossa organização, era a primeira vez que realizavam um projeto neste campo desta magnitude e por esta razão todos os passos foram uma aprendizagem. A análise de dados dos sistemas origem foi um tópico bastante debatido, a organização chegou a contratar serviços de uma outra empresa para fazer análise de conteúdo, antes de contratar a empresa externa para realizar a migração de dados. Essa empresa tinha como finalidade realizar a seleção das páginas com melhor conteúdo, o que acabou por ser uma má decisão. Nomeadamente, porque o trabalho entregue não correspondia ao que fora pedido.

Depois da extração e documentação de todo conteúdo pela equipa externa, com os dados extraídos foi possível produzir relatórios. Estes relatórios dão informação sobre que conteúdo existe, como está estruturado, metadados e hiperligações. Com uma ferramenta providenciada pela equipa externa (**Figura 22**) para a análise, esta dá uma visão de todos os dados que foram extraídos, é possível navegar por categorias, ver a data de criação e onde este conteúdo se encontra. É possível ainda utilizar o filtro para encontrar algo específico, a **Figura 23** mostra a *dashboard* com estas possibilidades. Aqui é possível verificar os dados quando não estão em concordância com as regras de negócio e ainda verificar conteúdo duplicado. Verificar quando foi a última visita que determinada página recebeu, ou ainda se recebeu visitas num determinado período.

Os sistemas origem contêm alguns anos e a equipa tinha a certeza que seria necessário realizar uma limpeza aos dados e analisar a sua qualidade. No entanto, foi possível encontrar erros com alguma especificidade e a equipa decidiu corrigir estes erros nos sistemas, para tornar o processo de migração mais fácil.

***Figura 22**: Interface de Análise providenciado pela empresa contratada*

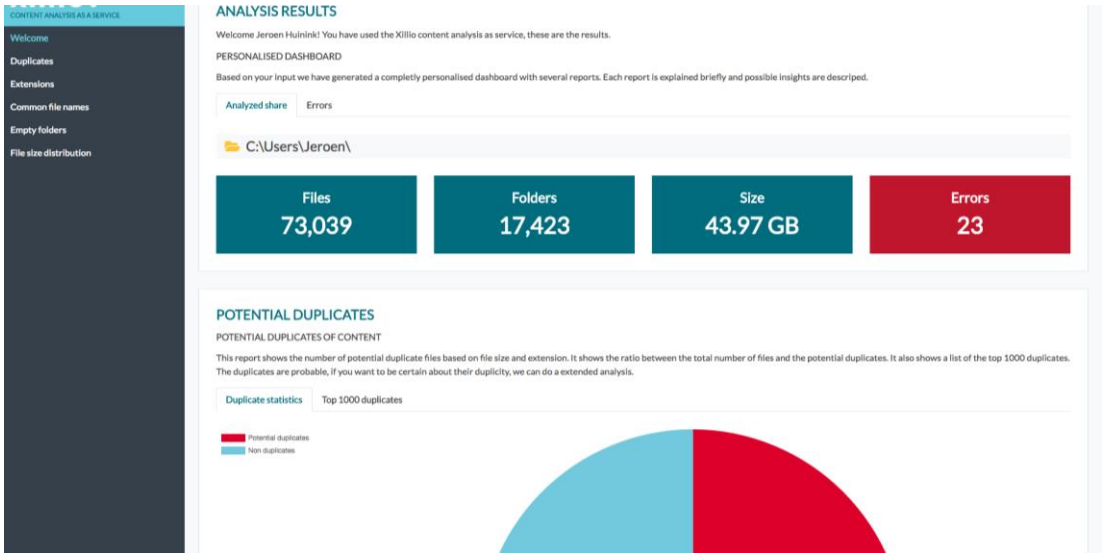
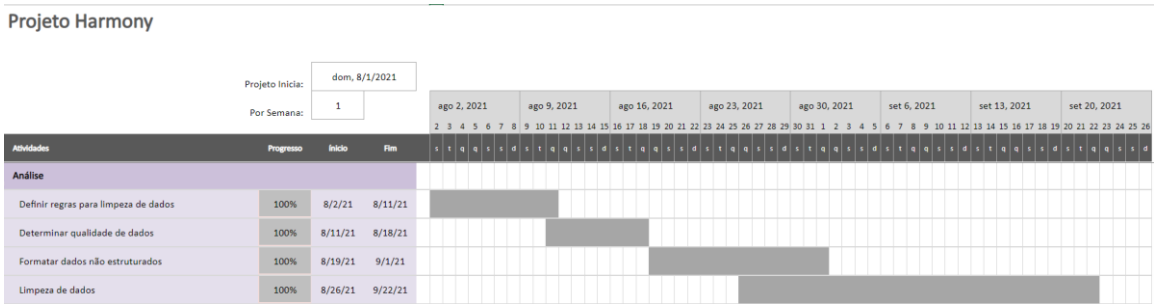


***Figura 23**: Insights com diferentes setores de análise*

Apesar de a ferramenta providenciar relatórios com base nas extrações feitas, a tarefa de realizar a limpeza de dados exigiu muito tempo e esforço da equipa. Esta atividade inicia-se a agosto de 2021 e termina a setembro de 2021, na **Figura 24** é possível visualizar as atividades e a sua duração. Com 4 semanas total para realizar a limpeza, duas semanas para o sistema *Joomla* e duas semanas para os sistemas *TYPO3* e *Sitecore*.

***Figura 24**: Cronograma da quarta etapa – Análise*



2.4.5 Seleção e Mapeamento

Cada um destes sites tem a sua estrutura de organização de dados, e para o sistema mais antigo como o *TYPO3*, existiu a criação de tipo de componentes personalizados para responder às necessidades da empresa nos seus primeiros anos. Dessa forma, foram criados mapas, uma vez que estes tipos de dados são personalizados e foi necessário realizar adaptações.

Para além da necessidade de mapear a estrutura dos sistemas origem, é essencial para a migração existir o mapa da estrutura do novo sistema, um passo essencial, sendo que é um sistema novo (AEM) existem muitos elementos em desenvolvimento, por construir ou em período de melhoria. Esta atividade contou com o acompanhamento da equipa externa de migração e o indivíduo 2, foi preenchido o mapa ao nível dos campos, para cada campo nos sistemas de origem foi necessário decidir a que campo correspondia no novo sistema, mas um processo que compensou durante a migração.

Esta etapa é dedicada para selecionar o conteúdo que se pretende mover, em que o resultado, é não migrar conteúdo sem valor para o novo ambiente. Todas as ferramentas fonte têm uma estrutura grande, para tornar o processo mais fácil e automático, gerou-se um ficheiro Excel com a estrutura completa dos sites, incluindo informações como o autor da página, data de criação e última atualização.

Criaram-se ainda regras para a migração, por exemplo, todas as páginas pertencentes a uma secção do website não precisavam ser migradas, ou todas as páginas que pertencem a uma secção de indústria precisam de ser transferidas. Outra regra que a empresa definiu estava relacionada com as visitas às páginas, todas as páginas que não recebiam visitas no último ano podiam ser arquivadas. Em todo o conteúdo extraído na fase de extração sofreram a aplicação destas regras de migração.

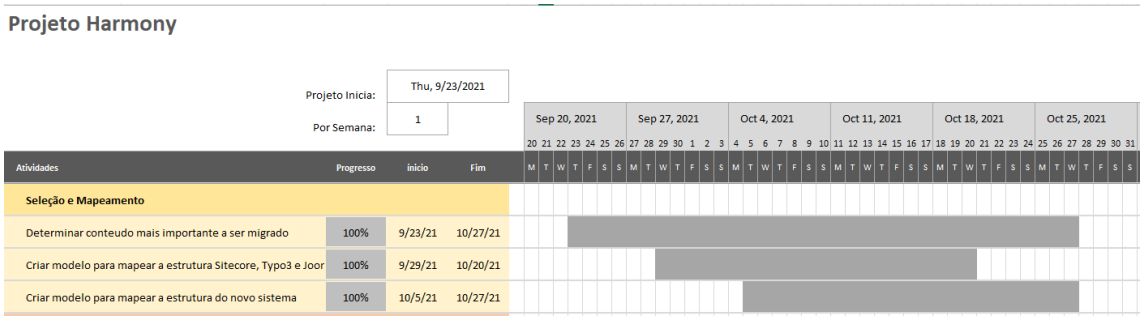
Uma pequena amostra do resultado está disponível na **Figura 25**, que deu origem a um documento com a lista de tipos de campos dos sistemas, com informações como o nome da propriedade, tipo, valor e as opções.

***Figura 25**: Mapa de componentes*

Component Name	Node Name	Property Name	Type	Value	Options	Description
Title (page title)	title	sling:resourceType	String	digitalxn/base/components/title/v1/title		
		jcr:title	String	My Test Title		if empty, page property title will be used
		type	String	h1	h1, h2, h3, h4, h5	
Text (RTE)	text	sling:resourceType	String	digitalxn/base/components/text/v1/text	<p> & Co.	
		textIsRich	String	true	true, false	
Image	image	alt	String	false	true, false	
		altValueFromDAM	String	false	true, false	
		jcr:title	String	My cool image caption		Caption T
		titleValueFromDAM	String	false	true, false	
		sling:resourceType	String	digitalxn/base/components/image/v1/image		
		linkURL	String	/content/ /global/en/demo/components/image		
Column Control	columncontrol			digitalxn/base/components/columncontrol/v1/columncontrol		
	columncontrol/col_1 -5	sling:resourceType	String	digitalxn/base/components/columncontrol/v1/columncontrol/container		
List	list	childDepth	String	1		
		downloadLabel	String	Download		
		sling:resourceType	String	hbkworld/base/components/list/v1/list		
		linkItems	String	true	true, false	
		listFrom	String	static		
		loadMoreEnabled	String	false		
		LoadMoreLinkLabel	String	Load More		
		maxItems	String	10		
		pages	String[]	/content/ /global/en/solutions/acoustics, /content/ global/en/solutions		
		parentPage	String	/content/ /global/en/solutions		
		readMoreEnabled	String	Read More		
		readMoreLinkLabel	String	true	true, false	
		showDescription	String	true	true, false	
		showDownload	String	true	true, false	
		showImage	String	true	true, false	
		showModificationDate	String	false	true, false	
		showPageCount	String	false	true, false	
		showSubtitle	String	true	true, false	
		showTagFilter	String	false	true, false	
		showTags	String	false	true, false	
		sortOrder	String	asc	asc, des, etc	
		tagsMatch	String	any		
Embed	embed	sling:resourceType	String	digitalxn/base/components/embed/v1/embed		
		embeddableResource	String	/base/components/embeddable/wistia		
		type	String	embeddable	embeddable, html, url	

A **Figura 26** é uma apresentação de todas as tarefas alocadas a esta etapa da migração de dados, teve o seu início a 23 de setembro de 2021 e finalizou-se a 27 de outubro de 2021. A equipa externa dedicou 5 semanas, para criar um modelo para os mapas, acompanhar e dar suporte à equipa interna. O mapeamento dos sistemas de origem foi realizado pelos membros da equipa interna.

***Figura 26**: Cronograma da quinta etapa – Seleção e mapeamento*



2.4.6 Desenho da Migração de Conteúdo

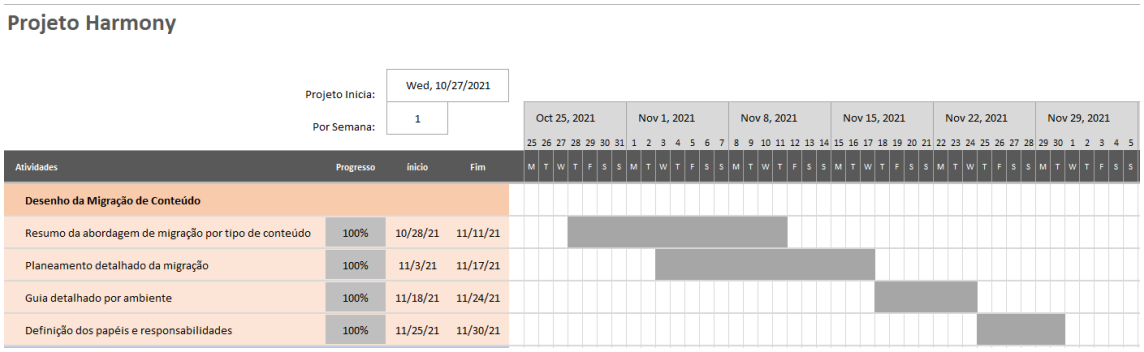
A fase de desenho da migração de conteúdo foi dedicada a planear, abordar quais os próximos passos seriam necessários para a transformação de dados, tais como as regras de negócio e o processo de migração.

Foi uma fase centrada na definição de que tipo de conteúdo seria migrado e definir qual teria prioridade na migração. Nesta atividade foram também definidos os papéis e a responsabilidade de cada indivíduo no processo de transformação. É uma fase que não necessita de muito tempo para ser definida. Na **Figura 27** é possível observar cada

atividade correspondente a esta fase no processo de migração, com o início em outubro de 28 de 2021 e a sua finalização a 30 de novembro de 2021.

O cronograma contém uma falha durante o mês de dezembro, devido ao facto de a empresa suspender as suas atividades em algumas partes do mundo e por essa razão durante esse mês não foram previstas atividades relacionadas com a migração.

***Figura 27**: Cronograma da sexta etapa – Desenho da migração de conteúdo*



2.4.7 Transformação

A etapa de transformação assinala o início da migração de dados em produção, como referido anteriormente, a existência de três sistemas diferentes contribui para o facto de as atividades de extração e transformação ocorrerem em diferentes ciclos.

A transformação foi realizada pela empresa externa contratada pela organização, das únicas fases que decorreu de forma contínua e sem interrupções, os ficheiros concebidos já estavam modificados para o formato de carga e foram depois transferidos.

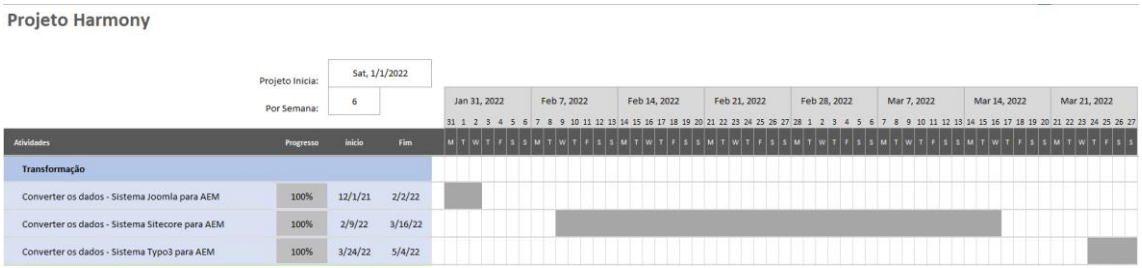
A empresa contratada criou uma ferramenta própria de migração de dados, um script, que consegue dar ordens, como pedir para ir ao banco de dados do *TYPO3* e ler todas as páginas, campo por campo, como os campos de *frontend*, lê o conteúdo da página, mas lê também os *metadata description*, *metadata title* e qualquer informação sobre a estrutura da página.

É executada uma consulta no banco de dados do sistema para construir uma página, vai na tabela e indica que para a página com determinado endereço é necessário buscar uma imagem específica, busca a imagem nesse campo da tabela, busca o título desse campo, e ainda define também a ordem dos componentes.

Na **Figura 28** estão estipuladas as atividades aplicadas a esta fase da migração de dados, com início a 1 de dezembro de 2021 e encerramento a 04 de maio de 2022.

Com o início a 01 de dezembro de 2021 a Equipa externa trabalhou na implementação da lógica da migração do Jommla durante duas semanas e parou devido às férias de Natal comunicado pelas duas empresas. A empresa retomou a sua atividade a 5 de janeiro de 2022 e continuou o processo de lógica, terminando a 02 de fevereiro de 2022. Para a implementação do processo de lógica do sistema *Sitecore* para AEM a equipa iniciou a sua atividade a 09 de fevereiro de 2022 e terminou a 16 de março de 2022. Por último, a empresa externa iniciou a atividade de transformação de dados do sistema *Typo3* para AEM, a 24 de março de 2022 terminando a 04 de maio de 2022. Entre a fase de transformação de dados de um sistema para o outro eram iniciados os testes de migração dos dados, para o qual já existia uma lógica estabelecida pela equipa externa.

***Figura 28**: Cronograma da sétima etapa – Transformação*



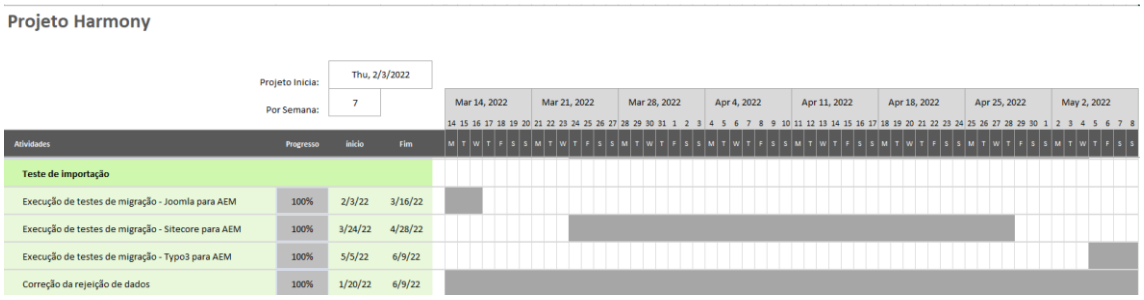
2.4.8 Teste de importação

Antes do avanço definitivo de grandes cargas de dados serem migrados, foram realizados testes de pequenas proporções de dados. Estes testes foram iniciados visando realizar a importação de dados para o novo sistema, como este reagiria e que erros eram identificados. Assim que estes erros fossem corrigidos, era realizada uma nova importação de uma nova porção de dados maior, verificava-se a existência de erros, era realizada uma nova correção dos mesmos.

Na **Figura 29** é disponibilizado as datas correspondentes, esta etapa iniciou-se em 03 de fevereiro de 2022 e terminou a 09 de junho de 2022. Para cada sistema existiram cerca de 3 testes, entre estes testes ocorriam os resultados da migração, se necessário eram realizadas atualizações aos mapeamentos. Assim que os testes eram finalizados era iniciada a migração final.

Um dos erros encontrados estava relacionado com dependências, havia algumas informações dependentes de informações oriundas de outros sistemas, que não foram carregadas e este problema comprometeu alguns testes levando a que o tempo desta atividade se estendesse mais do que o período estipulado inicialmente.

***Figura 29**: Cronograma da oitava etapa – Teste de importação*



2.4.9 Importação para produção

O último passo do processo de ETL – o carregamento, nesta fase a equipa externa contratada pela organização e o indivíduo responsável pela comunicação com a empresa, necessitam de supervisionar e cancelar cargas segundo o desempenho do servidor.

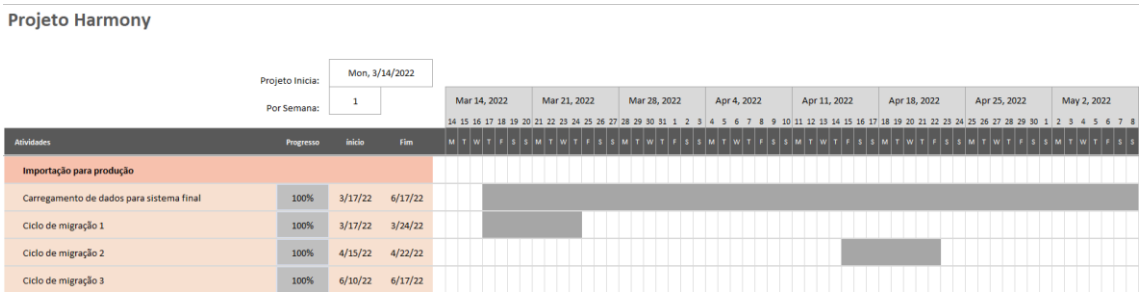
O carregamento de dados foi realizado de forma intercalada, assim que era terminado os testes de migração e a migração já não apresentava erros era realizado a migração final. O carregamento dos dados para produção, era migrado em massa todo o conteúdo selecionado. A escolha desta estratégia recaiu no facto de que a organização ter dados de três sistemas diferentes, o que levou a organização a decidir qual a migração de dados correspondente a cada sistema seria.

A primeira carga de dados foi do sistema *Joomla* para o novo sistema (AEM), a decisão de ser o primeiro sistema está relacionado com a insuficiência de funcionários para continuar a realizar a gestão do sistema, de forma a diminuir a carga de trabalho de outros funcionários.

Assim que os testes de migração terminassem entre o sistema *Sitecore* e AEM, era realizado o carregamento final, do sistema *Sitecore* para AEM.

Posteriormente, realizou-se o carregamento dos dados do *TYPO3* para o AEM um sistema com muitos dados para serem migrados, devido ao tipo de negócio e pela idade do site (onze anos). Estes ciclos de carregamento realizavam-se durante os fins de semana e períodos de baixa atividade nos sistemas. No cronograma (**Figura 30**) é visível as datas de migração que correspondem a cada sistema, com início a 17 de março de 2022 e encerramento a 17 de junho de 2022.

***Figura 30**: Cronograma da nona etapa – Importação para produção*

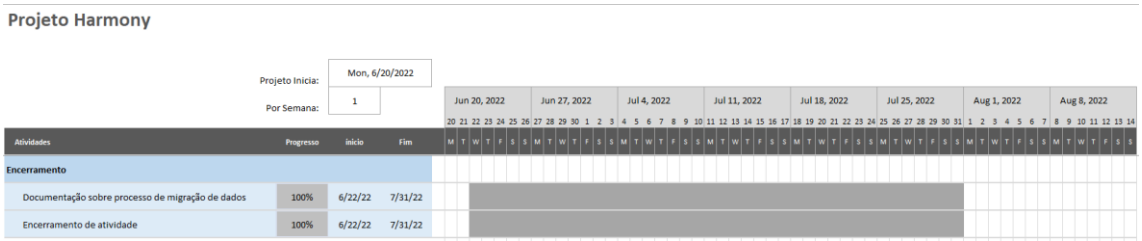


2.4.10. Encerramento

A elaboração do relatório foi a etapa final, a migração de dados já ocorreu, é a etapa mais calma de todo o processo. A equipa externa entregou à organização toda a documentação sobre a migração de dados, ficheiros, manuais, assim como um relatório sobre a qualidade de dados, a capacidade de infraestrutura, relatórios sobre o sistema final e a sua precisão.

Depois da entrega de toda a documentação sobre o projeto, encerrou-se assim toda a atividade (**Figura 31**).

***Figura 31**: Cronograma da décima etapa – Encerramento*



CAPÍTULO III – DISCUSSÃO DE RESULTADOS

O capítulo III pretende realizar uma análise crítica ao caso estudado, responder à questão de investigação colocada inicialmente e como os objetivos específicos ajudaram.

O primeiro objetivo específico tinha como meta encontrar vantagens e desvantagens que se encontravam na adoção da nova solução (AEM) em comparação com os sistemas antigos. Após realizar uma análise a todos os sistemas criou-se tabelas com as vantagens e desvantagens. Cada sistema tem as suas peculiaridades, a plataforma AEM foi adotada devido à agilidade de trabalho, possibilidade de editor visual, uma experiência mais natural, uma vez que se visualiza as páginas como o utilizador vê. Este sistema oferece ainda muitas possibilidades de criação de conteúdo e de se criar experiências para diversos canais de comunicação. A plataforma disponibiliza ainda um DAM com diversas funções, tais como: realizar pesquisas visuais, fazer marcação inteligente para encontrar mais rapidamente os ficheiros, gerir e armazenar. É uma plataforma que disponibiliza integração com diversas ferramentas de marketing através dos seus conectores AEM. Apesar de ser um sistema de fácil navegação para utilizadores com um conhecimento tecnológico considerável, para utilizadores que não tenham muita habilidade pode representar uma linha de aprendizagem longa, devido às diversas funcionalidades disponibilizadas pela plataforma. O AEM é um sistema para empresas de dimensão considerável, dispostas a gastar uma grande quantia do orçamento, porque é necessário a contratação de serviços terceirizados para instalação e manutenção, pois requer conhecimento em código para modificar, personalizar e adicionar funcionalidades. Com o novo sistema pode-se atribuir a cada conta de utilizador diferentes propriedades, com autorização para apenas editar conteúdo, autorização para publicação de conteúdo e master com acesso a tudo o que existe no sistema, desta forma a possibilidade de erro é baixa. Uma das grandes vantagens do sistema é a possibilidade de cada equipa/utilizador ter um espaço próprio para os seus projetos, para além de não existir a eventualidade de colisão com outros projetos é possível ver o nível de produtividade. Apesar de ser uma plataforma relativamente nova, é possível encontrar diversos documentos de apoio, assim como tutoriais, cursos e fóruns. Uma das desvantagens é a necessidade de um serviço terceirizado para realizar a instalação da plataforma. Dado que AEM é uma plataforma mais complexa, requer conhecimentos técnicos especializados que a empresa não dispõe, pelo que foi necessário contratar recursos externos com custos elevados. Apesar destas

desvantagens, a organização precisava de uma ferramenta onde pudesse organizar e ter o conteúdo de três sistemas diferentes, justificando o investimento.

A grande vantagem que se encontra do novo sistema em comparação aos antigos é o facto de ser um sistema mais recente, com um layout mais limpo. Apesar de a organização ter de manter os sites fonte durante mais um ano, a possibilidade de selecionar conteúdo dos outros sistemas e migrar para um único sistema era o propósito do projeto.

Numa comparação entre *TYPO3* e AEM, a diferença mais notável pelos editores é a disposição das páginas de conteúdo, por ser um sistema com alguma idade, o *TYPO3* da forma que foi construído apenas utilizadores com alguma experiência no sistema conseguiam encontrar as páginas com facilidade. Apesar de o *TYPO3* ser um sistema bastante desenvolvido, a interface do AEM é mais agradável visualmente para se trabalhar, permite visualizar a aparência da página enquanto se edita o conteúdo, no *TYPO3* isso não é possível. Para ambos os sistemas, para a configuração e manutenção é necessário recorrer a serviços terceirizados, devido à necessidade de conhecimento em programação avançada. Para editores e administradores do sistema, o *TYPO3* apresenta um nível de complexidade maior na aprendizagem. No entanto, pode existir confusão por parte de um iniciante no AEM pela sua diversidade de oferta em funcionalidades.

O *TYPO3* no que concerne a documentação sobre o funcionamento do sistema é forte em comparação ao AEM. Na rede social YouTube encontra-se mais informação sobre como trabalhar e configurar o sistema AEM com mais facilidade. Encontra-se vídeos sobre o sistema *TYPO3*, mas grande parte dos vídeos são em língua alemã, o que pode ser um problema para indivíduos que não falem o idioma. No *TYPO3*, foram criadas componentes muito personalizadas para o negócio da empresa e que não estão disponíveis nos outros sistemas.

Na comparação do *Sitecore* e AEM é possível encontrar algumas vantagens no uso do AEM. O *Sitecore* é um sistema que não é intuitivo durante a fase inicial de aprendizagem e demora algum tempo até que o utilizador se sinta confortável em navegar pelo sistema. Em comparação com o AEM, apesar de ser confuso para utilizadores iniciantes por existir uma diversidade grande de opções, consegue ser mais fácil de navegar em comparação ao *Sitecore*. Para ambos, existe pouca documentação de informação para ajudar novos utilizadores comparando com os outros sistemas

(*TYPO3* e *Joomla*). O *Sitecore* pode ser considerado um sistema com custos significativos, uma vez que é necessária uma licença por utilizador, no AEM não tem este entrave. Para o *Sitecore* é necessário algum indivíduo com conhecimento em tecnologia de informação para configurar e gerir. Comparando as interfaces de utilizador em cada um dos sistemas, o AEM apresenta maior facilidade de uso e é mais interativo quando comparado com os dois sistemas.

Por último, a comparação entre *Joomla* e AEM, ambos permitem aos utilizadores um editor visual, o que facilita a edição de conteúdo. Apesar do *Joomla* estar no mercado de CMS's há alguns anos e ter bastantes utilizadores do sistema, não é um programa fácil de entender para utilizadores iniciantes, por ser um sistema com características próprias e muitas funcionalidades. Em comparação ao sistema AEM, apesar do *Joomla* necessitar que se tenha conhecimento de JavaScript e PHP para determinadas configurações, o AEM consegue ser um sistema mais complexo no desenvolvimento do website. Como o *Joomla* é um sistema muito utilizado, existe muita documentação, fóruns onde é possível encontrar bastante informação sobre a plataforma, o que ajuda bastante na sua configuração e manutenção. Em termos de navegação, apesar de ambos terem um grau de facilidade, o sistema AEM acaba por ser mais intuitivo, por ter uma interface mais amigável.

Ao realizar a comparação entre estes sistemas, foi possível perceber que cada um deles tem as suas próprias características, apesar de todos terem a finalidade de apresentar conteúdo aos seus consumidores. A grande decisão de escolher uma nova plataforma recaiu no facto de ser necessário migrar três sistemas, cada um ter o seu próprio conteúdo e a necessidade de reordenar todo o conteúdo para ir de encontro com o que a organização pretendia apresentar aos seus clientes. A organização decidiu então avançar com a escolha de um novo sistema AEM, por ser uma plataforma recente, inovadora e que ia de encontro com o que se procurava.

4.1 Análise do processo de migração

Após decidir qual seria o novo sistema, a organização teve de decidir como seria realizada a migração de dados, a empresa optou pela contratação de um serviço externo que ficaria responsável pela execução da migração de dados dos sistemas origem para o sistema destino. No primeiro capítulo foi mencionado diversas ferramentas que realizam a migração de dados, foi selecionado um serviço terceirizado que a empresa contratou

para ajudar no projeto que recorria a uma ferramenta própria. Uma vez conhecido o sistema destino e a empresa que iria ajudar na realização do projeto de migração, foi necessário delinear a estratégia a ser adotada para a realização da migração de dados.

O processo aplicado na migração foi o ETL, suportado na ferramenta, com esta abordagem foi possível extrair, transformar e carregar os dados para o novo sistema. A organização em conjunto com a empresa que providenciou o serviço para a migração delinearam uma estratégia em que consistiu em dez atividades descritas no subcapítulo 2.3. Durante o delineamento da estratégia, a organização teve de selecionar que conteúdo seria migrado e qual o espaço de tempo necessário para ocorrer toda a migração. Como referido no subcapítulo 2.3 a abordagem pela qual a organização decidiu foi a migração única de dados. Esta abordagem permitiu que a migração de dados fosse realizada de uma única só vez, todos os dados selecionados de cada sistema foram transferidos. Esta abordagem é conhecida como a estratégia *Big Bang*, que foi de encontro com o que a organização estava à procura, assim o conteúdo foi migrado de cada sistema, dando oportunidade para que fosse realizado mais testes.

No apêndice I, está disponível o cronograma inteiro do projeto de migração, é possível visualizar as fases, atividades correspondentes e ainda o tempo gasto com cada. Não é um cronograma muito detalhado (com número de horas), a organização decidiu trabalhar com datas-limite de entrega para as atividades que teria de ser comprido, mas como em todos os projetos existiram atividades que necessitaram mais tempo do que o previsto inicialmente. No subcapítulo 2.4 tem em detalhe as atividades e ainda exemplos retirados da organização para melhorar a compreensão do que foi produzido, alguns detalhes foram cobertos para que não fosse exposto informações sobre a organização. A fase inicial do processo (1. Início) foi umas das fases com mais contratempos, por ser uma fase onde ainda faltava ter algumas decisões tomadas e ainda não existir uma equipa com foco total no projeto. Após estabelecer os indivíduos que fariam parte da equipa projeto, foi necessário identificar os riscos que estavam associados ao projeto, uma das preocupações estava relacionada com a qualidade dos dados nos sistemas de origem por terem alguns anos, apesar de a equipa ter uma perceção que existiriam alguns dados em falta ou de baixa qualidade ainda não se sabia a total extensão destes erros. Para além destes percalços durante a migração a organização sofreu ainda algumas mudanças na sua organização, o que não abalou muito o avanço do projeto de migração.

A segunda etapa (2. Instalação e configuração) foi executada sem grandes percalços e terminada em menos tempo comparada com todas as outras etapas do processo. A equipa externa conseguiu a instalação e a configuração dos conectores, é considerado um processo simples tendo em conta a complexidade de etapas do processo que necessitam de mais conhecimento e tempo.

Após a instalação e configuração dos conectores nos sistemas origem seguiu-se a etapa número 3, extração, apesar de parecer um processo simples, foi necessário explicar qual a organização e estrutura de um dos sistemas por ser complexa e com diversos níveis (*TYPO3*). Apesar de ter alguma complexidade, ocorreu de forma tranquila, é a primeira fase do processo ETL. No subcapítulo é possível visualizar amostras das extrações realizadas, foram retiradas algumas informações para que se respeite as normas de confidencialidade da organização.

Após a extração dos dados dos três sistemas, era necessário realizar uma análise à qualidade dos mesmos, a etapa 4. Apesar de a organização saber que iriam existir determinados erros nos dados, esta etapa foi fundamental para perceber a quantidade e complexidade que existia nestes erros. Para que estes erros fossem corrigidos, contactou-se os três indivíduos que realizavam a gestão destes sistemas e com a ajuda de estagiários na empresa foi possível eliminar grande parte dos erros encontrados.

Na etapa 5, seleção e mapeamento, depois da análise aos dados extraídos era necessário realizar uma seleção do conteúdo que seria migrado e quais as regras necessárias obedecer. As regras teriam de ser bem simples para que a seleção das páginas de conteúdo não se tornasse algo complexo, a regra principal era a migração de todas as páginas, menos as que não recebiam visualizações no último ano. Para além de ser necessária realizar esta seleção na etapa 5, foi realizado o mapeamento. Um passo de grande importância, o mapeamento ao nível dos campos, para cada campo nos sistemas de origem foi necessário fazer corresponder um campo no novo software. Perceber que seriam necessárias algumas adaptações, uma vez que existiam campos personalizados para o negócio da empresa nos antigos sistemas e que não existia uma componente específica no sistema destino.

Após selecionar o conteúdo e mapear os campos existentes, já existia toda a informação necessária para a próxima fase. Mas como acordado entre a organização, realizou-se o desenho da migração, como seria realizada, quem eram os indivíduos necessários para

estar envolvido no processo, definir como se iria proceder á migração e o que a empresa poderia esperar.

Após saberem o que teria de ser importado para o novo sistema, iniciou-se a etapa da transformação (etapa 7), esta etapa estava relacionada com a etapa de testes de importação (etapa 8). Apesar de as equipas terem tudo definido, os testes de importação são necessários, uma vez que o script utilizado pela equipa externa estava dividido em duas partes. Na primeira parte o script apenas ia buscar a estrutura da página, os elementos/componentes que a página necessita, por exemplo, cabeçalhos, elemento imagem, elemento texto, colunas, relações e importava essa informação. Na segunda fase do script este ia buscar o conteúdo, o texto, a imagem, links. Numa fase inicial migrou-se apenas uma página teste, a seguir três páginas, até que fosse possível realizar a migração de páginas sem percalços, durante estes testes a equipa ia a aprimorar o script.

Como mencionado anteriormente, a migração de dados neste estudo de caso seria realizada de uma única vez, devido à elevada quantidade de dados. Existia uma grande urgência de se realizar a migração do sistema *Joomla* para a nova plataforma, devido ao gestor deste sistema ter saído da organização, não existia outros funcionários com um conhecimento muito profundo de como este estava estruturado e sem mão de obra para continuar a fazer manutenção. A seguir, o sistema a realizar a migração de dados para produção (AEM) foi o *Sitecore*. Por último, realizou-se o carregamento de dados para produção do sistema *TYPO3*, foi um dos passos que apareceu mais complicações quando se realizou, uma vez que o sistema não utiliza *templates* para as páginas, por exemplo, as páginas de produto não tinham a mesma estrutura entre si. Em comparação com os outros dois sistemas onde existia essa prática. Todos estes carregamentos realizavam-se ao fim de semana.

Na estratégia de migração o relatório está como última etapa, mas é construído durante todo o processo de migração, onde são colocadas informações conforme são realizadas as etapas. Este documento serve para que a organização tenha consigo um documento de tudo o que foi executado, ter conhecimento de scripts utilizados, problemas encontrados, entre muitas outras informações.

Ao longo de 13 meses, a equipa que constituía o projeto uniram forças para que a migração de dados acontecesse, usando a estratégia e abordagem estipulada no início do

mesmo. A organização que criou o projeto de migração de dados não possuía nenhuma experiência neste campo e todos trabalhavam direta ou indiretamente com a técnica ETL. A equipa interna nunca ter realizado uma migração de dados era mais um obstáculo para o projeto, no entanto, a equipa depositou confiança na empresa externa que já tinha experiência neste campo. Apesar de encontrar alguns percalços durante o projeto, a vantagem da contratação de um serviço externo para ajudar na migração de dados foi visto como um grande avanço para que o projeto andasse com um ritmo melhor de trabalho.

CAPÍTULO IV – CONCLUSÕES

4.1 Considerações finais

Desde que Tim Berners-Lee inventou a WWW a internet tem uma dinâmica crescente ano após ano com um grande impacto na sociedade, nomeadamente, nos estilos de vida e na sua formação. Nos últimos três anos, devido à doença do coronavírus, assistiu-se à conversão de muitas empresas para o comércio eletrónico. Assim sendo, um bom conteúdo do website é essencial para qualquer campanha de marketing digital. Quando os sites são acessíveis, fáceis de navegar e credíveis as empresas conseguem tirar o maior proveito dos esforços de marketing e vendas.

Concluída a investigação que teve como propósito o estudo da migração de dados entre sistema de gestão de conteúdo, conhecer qual a estratégia adotada e os passos necessários para o projeto ser concluído dentro de uma organização internacional. A definição dos objetivos no primeiro capítulo formara o alicerce para a construção do restante estudo de caso, o que permitiu investigar e dar resposta às questões levantadas.

O primeiro objetivo específico ajudou na comparação entres os sistemas, serviu para mostrar como estes sistemas se criaram e quais são as vantagens e desvantagens de cada um. Criaram-se tabelas para ser facilmente interpretado os pontos/características positivas e negativas que se encontrou em cada um. A criação destas tabelas foi em respostas ao segundo objetivo específico estipulado, com o propósito de dar conhecimento sobre as ferramentas a pequenas empresas que não têm oportunidade de ter experiência com todas e conseguirem, pelo menos, ter uma noção do que cada sistema pode oferecer ou não.

Com a realização deste estudo e com a definição do terceiro e quarto objetivo específicos foi possível dar resposta à questão de investigação colocada inicialmente. A identificação de todas as etapas necessárias para a realização de uma migração de dados numa organização é uma situação real e que acontece a algumas empresas quando necessitam de mudar de plataformas. Para além da apresentação de todas as fases foi realizada a descrição de cada uma, das atividades que a organização teve de executar, a ordem cronológica e dos resultados obtidos. Estes resultados em algumas das etapas são possíveis de ser visualizados, assim como, os problemas enfrentados.

Concluído o estudo, é possível dizer que foi bem-sucedido, a informação compartilhada pode contribuir para outras organizações, ao leitor comum e para o campo académico sobre o tema.

Este estudo de caso contém uma limitação temporal e por essa razão não foi possível acompanhar o processo de mudança do e-commerce na nova plataforma. Devido à magnitude da empresa foi impossível investigar todos os projetos a acontecer em separado, onde o resultado é uma única solução em todos os campos relacionados com um website e loja integrado.

4.2 Limitações

As limitações encontradas durante a investigação do estudo de caso são na sua maioria referentes ao tempo. Visto que, a organização, comporta uma estrutura grande, está em período de mudança da estratégia e é inevitável que sejam feitas alterações à estrutura e que leve ao despedimento de funcionários. Neste caso, as mudanças na estrutura levaram a algumas modificações na equipa do projeto, que resultou numa paragem durante alguns dias até se encontrar qual era a melhor solução para a equipa.

Por se tratar de um projeto de grandes dimensões e devido ao limite de tempo disponível para entrega deste estudo, não foi possível colocar a integração do e-commerce no site, uma vez que ultrapassa a data-limite. Devido a algumas decisões da organização como a existência de projetos em separado, a gestão de informação sobre os clientes de três sistemas diferentes não foi possível de ser partilhado neste estudo, por se tratar de um projeto que será integrado fora da linha de tempo disponível.

Uma outra limitação, está relacionado com a falta de conhecimento da equipa interna e a escolha de serviços externos para determinadas tarefas que em nada ajudaram para o andamento do projeto. A contratação de uma empresa externa para a captação de erros e análise de páginas com maior desempenho foi uma má decisão. Inicialmente porque a informação entregue não correspondeu ao que fora pedido, em segundo, o custo do serviço e a perda de tempo que levou a que a entrega das correções necessárias no conteúdo fosse fora da data-limite estipulada.

4.3 Recomendações para futuras investigações

A conclusão deste estudo abre portas para futuras investigações e explorar vertentes que não foram abordadas durante a realização desta pesquisa. Uma recomendação é o acompanhamento do novo sistema, o AEM, e perceber o impacto que este teve na empresa. Sendo que seria interessante analisar a migração de sistema de e-commerce, para encontrar pontos comuns. Realizar uma investigação com outra

estratégia/abordagem aplicada sobre a migração de dados, o que daria outra perspetiva e possibilidade de realizar comparações entre si. Outra sugestão seria utilizar a mesma abordagem de migração, mas com uso de outra ferramenta comercial ou script criado de raiz pela própria organização. Para uma investigação mais pormenorizada, a criação de uma ferramenta para automatizar a estratégia adotada.

REFERÊNCIAS BIBLIOGRÁFICAS

About Sitecore | Sitecore. (n.d.). Retrieved October 28, 2022, from https://www.sitecore.com/company

AEM Sites Content Management | Adobe for Business. (n.d.). Retrieved October 28, 2022, from https://business.adobe.com/products/experience-manager/sites/aem-sites.html

Almeida, J. R., Coelho, L., & Oliveira, J. L. (2021). BIcenter: A collaborative Web ETL solution based on a reflective software approach. *SoftwareX*, *16*. https://doi.org/10.1016/j.softx.2021.100892

Alrasheed, A. A., & Amir D. El-Masri, Ph. D. (2013). *Comparing E-commerce Content Management*. *53*(3).

Badiuzzaman Biplob, M., Sheraji, G. A., & Khan, S. I. (2018). Comparison of Different Extraction Transformation and Loading Tools for Data Warehousing. *2018 International Conference on Innovations in Science, Engineering and Technology, ICISET 2018*, 262–267. https://doi.org/10.1109/ICISET.2018.8745574

Bălan, C. (2014). Dave CHAFFEY, digital business and E-commerce management: Strategy, implementation and practice. In *Management and Marketing* (Vol. 9, Issue 3).

Barker, D. (2016). Web Content Management Systems, Features and Best Practices. In *Journal of Chemical Information and Modeling* (Vol. 53, Issue 9). https://doi.org/10.1017/CBO9781107415324.004

Biswas, N., Sarkar, A., & Mondal, K. C. (2019). Empirical Analysis of Programmable ETL Tools. *Communications in Computer and Information Science*, *1031*, 267–277. https://doi.org/10.1007/978-981-13-8581-0_22

Bogdan, R. C., & Knopp Biklen, S. (1994). Colecção Ciências Da Educação Orientada Por Maria Teresa Estrela E Albano Estrela-Investigaçao Qualitativa Em-Educaçao Uma Introdução à Teoria E Aos Métodos.

Chevalier, S. (2022, September 21). *Global retail e-commerce sales 2026 | Statista*. https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/

Cockburn, A. (2019, January 21). *Vamosa Web Migration Considerations & Goals for Data Management*. https://vamosa.com/2019/01/migration-architect-and-its-beginnings/

Coppola, D. (2021, October 13). *Digital buyers worldwide 2021 | Statista*. https://www.statista.com/statistics/251666/number-of-digital-buyers-worldwide/

Dragomirov, N. (2020). E-Commerce Platforms and Supply Chain Management – Functionalities Study. *Economic Alternatives*, *2*, 250–261. https://doi.org/10.37075/ea.2020.2.04

Engelfriet, A. (2010). Choosing an open source license. In *IEEE Software* (Vol. 27, Issue 1). https://doi.org/10.1109/MS.2010.5

Heleno, M., & de Mendonça, R. (2009). *Metodologia de Migração de Dados em um contexto de Migração de Sistemas Legados*. www.cin.ufpe.br/~posgraduacao

Instituto Nacional de Estatística. (2020). *Inquérito à Utilização de Tecnologias da Informação e da Comunicação pelas Famílias-2020*.

Instituto Nacional de Estatística. (2021). Inquérito à Utilização de Tecnologias da Informação e da Comunicação pelas Famílias.

Introduction and Overview to the new SAP Web Channel Experience Management | SAP Blogs. (n.d.). Retrieved October 28, 2022, from https://blogs.sap.com/2012/05/14/introduction-and-overview-to-sap-web-channel-experience-management/

Jacobi, F., Krawatzeck, R., & Hofmann, M. (2012). *A Model-driven Framework for ETL Documentation Generation Meeting the Need for ETL Documentation: A Model-driven Framework for Customizable Documentation Generation* (Issue 1). http://www.ssisdocumenter.com/

Joomla! - Content Management System to build websites & apps. (n.d.). Retrieved October 28, 2022, from https://www.joomla.org/about-joomla.html

Keeves, J. P. (1997). *Educational Research, Methodology, and Measurement: An International Handbook.*

Laudon, K. C., & Traver, C. G. (2016). E-commerce 2016: business. technology. society. In *Global Edition*. www.pearsonglobaleditions.com

Mathur, A. (2016). *Open Source Platform for Ecommerce*. *4*, 173–175.

Mello, R. B., Sandro, A., de Souza, R., Analisador, P., Martins, D., & Barros, R. (2009). *Migração Entre Sistemas Gerenciadores de Banco de Dados*.

Mohapatra, S. (2013). *E-Commerce Strategy*. Springer US. https://doi.org/10.1007/978-1-4614-4142-7

Patel, M., & Patel, D. B. (2021). Progressive Growth of ETL Tools: A Literature Review of Past to Equip Future. *Advances in Intelligent Systems and Computing*, *1187*, 389–398. https://doi.org/10.1007/978-981-15-6014-9_45

Qin, Z., Chang, Y., Li, S., Li, F., Qin, Z., Chang, Y., Li, S., & Li, F. (2014). E-Commerce and E-Commerce Strategy. In *E-Commerce Strategy*. https://doi.org/10.1007/978-3-642-39414-0_1

Ramalho, J. C., Ferreira, M., Faria, L., & Castro, R. (n.d.). *Boas Práticas na Migração de Repositórios: lições aprendidas com o CALM e o ARQBASE*.

Rohilla, N. (Panjab U. (2017). *Content Management System for E-commerce Website*. *6*(7), 490–495.

Sarmah, S. S. (2018). Data Migration, Phases, ETL, Testing, Data Migration Risks and Best Practices; Data Migration, Phases, ETL, Testing, Data Migration Risks and Best Practices. *Science and Technology*, *8*(1), 1–10. https://doi.org/10.5923/j.scit.20180801.01

Shoaib, M., Ishaq, A., Ahmad, M. A., Talib, S., Mustafa, G., & Ahmed, A. (2017). Software Migration Frameworks for Software System Solutions: A Systematic Literature Review. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 8, Issue 11). www.ijacsa.thesai.org

Silva de Oliveira, C., & Abud Marcelino, M. (2012). *Metodologias E Estratégias de Migração de Dados* (Vol. 13, Issue 3).

Soares Vian, V. (2020). *Desenvolvimento de Uma Ferramenta Para Migração de Bancos de Dados*.

Talend. (2020). *Talend Open Studio for Data Integration User Guide 7.3.1*.

Tirumala, S. S., Katragadda, R., & Nandigam, D. (2015). *ETL tools for Data Warehousing: An empirical study of Open Source Talend Studio versus Microsoft SSIS*. https://www.researchgate.net/publication/271207443

Trias, F., Castro, V., López-Sanz, M., & Marcos, E. (2013). *An ADM-based Method for Migrating CMS-based Web Applications: Extracting ASTM Models from PHP Code*. 85–92. https://doi.org/10.5220/0004602900850092

Tripathy, P., & Kshirasagar, N. (2015). *Software Evolution and Maintenance Priyadarshi Tripathy Kshirasagar Naik a Practitioner's Approach*.

Vassiliadis, P., Simitsis, A., & Skiadopoulos, S. (2002). Conceptual modeling for ETL processes. *ACM International Workshop on Data Warehousing and OLAP (DOLAP)*, 14–21. https://doi.org/10.1145/583890.583893

Vieira De Oliveira, F. (2017). *Migração de bases de dados relacionais para NoSQL-Métodos de Análise*.

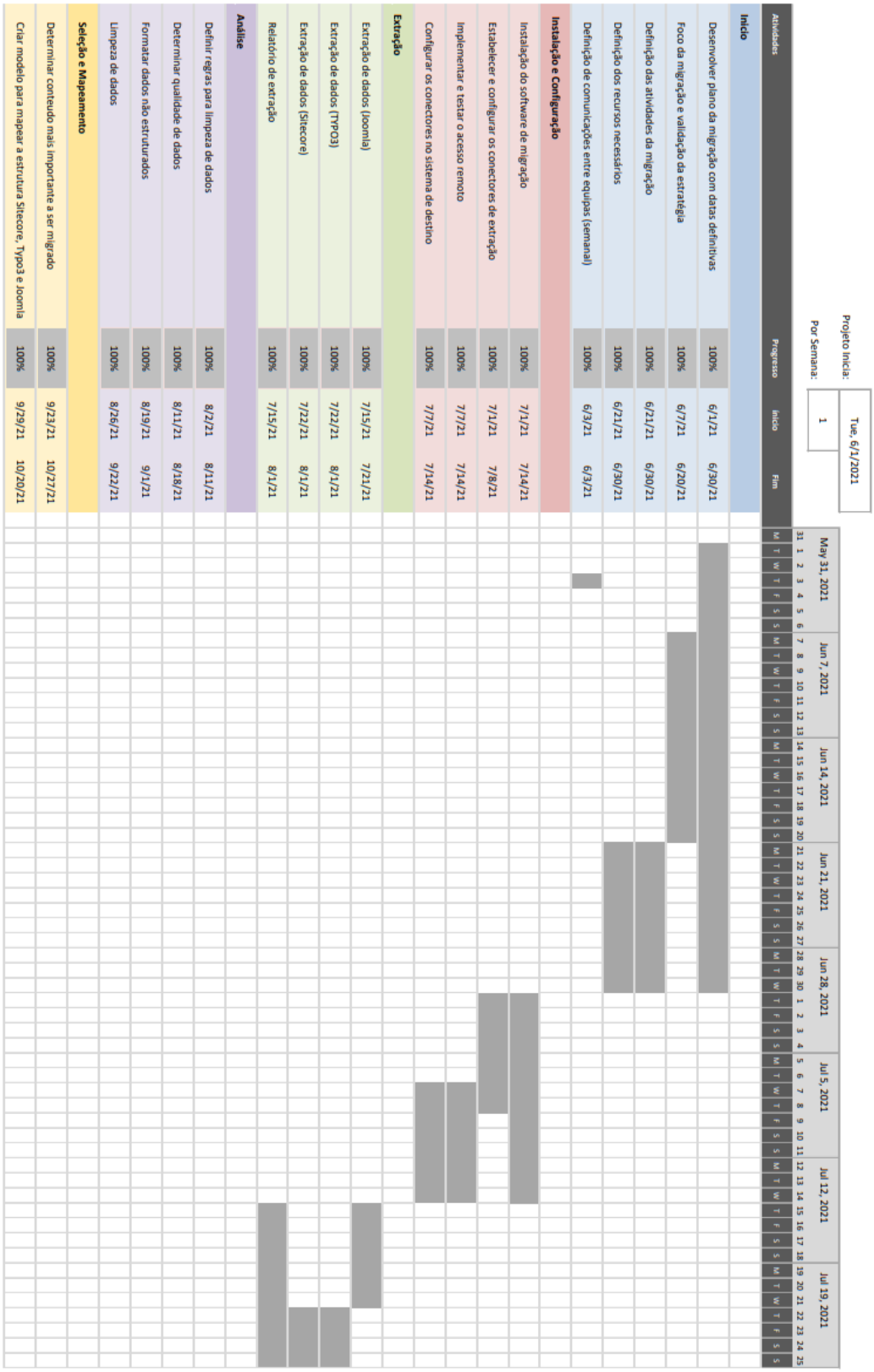
Xillio. (2020a). *Content ETL? What is it exactly?* https://www.xillio.com/what-is-content-etl

Xillio. (2020b). *Migrating Content From a Legacy ECM*. https://www.xillio.com/en-us/landing-page/whitepaper-migrating-content-from-a-legacy-ecm

Yin, R. K. (1984). Case Study Research Design and Methods. In *The Canadian Journal of Program Evaluation: Vol. 5 Edição*. https://doi.org/10.3138/cjpe.30.1.108

Zarmpou, T., Dalianis, H., & Vlachopoulou, M. (2009). *Data Migration Between Web Content Management Systems*. https://www.researchgate.net/publication/221215597

Apêndice I – Cronograma do Projeto



Por Semana: 1

Atividades	Progresso	início	Fim
Criar modelo para mapear a estrutura do novo sistema	100%	10/5/21	10/27/21
Desenho da Migração de Conteúdo			
Resumo da abordagem de migração por tipo de conteúdo	100%	10/28/21	11/11/21
Planeamento detalhado da migração	100%	11/3/21	11/17/21
Guia detalhado por ambiente	100%	11/18/21	11/24/21
Definição dos papéis e responsabilidades	100%	11/25/21	11/30/21
Transformação			
Converter os dados - Sistema Joomla para AEM	100%	12/1/21	2/2/22
Converter os dados - Sistema Sitecore para AEM	100%	2/9/22	3/16/22
Converter os dados - Sistema Typo3 para AEM	100%	3/24/22	5/4/22
Teste de importação			
Execução de testes de migração - Joomla para AEM	100%	2/3/22	3/16/22
Execução de testes de migração - Sitecore para AEM	100%	3/24/22	4/28/22
Execução de testes de migração - Typo3 para AEM	100%	5/5/22	6/9/22
Correção da rejeição de dados	100%	1/20/22	6/9/22
Importação para produção			
Carregamento de dados para sistema final	100%	3/17/22	6/17/22
Ciclo de migração 1	100%	3/17/22	3/24/22
Ciclo de migração 2	100%	4/15/22	4/22/22
Ciclo de migração 3	100%	6/10/22	6/17/22
Encerramento			
Documentação sobre processo de migração de dados	100%	6/22/22	7/31/22
Encerramento de atividade	100%	6/22/22	7/31/22