



INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Gestão de Dados Não Estruturados na Era da Transformação Digital: Práticas e Impactos na Eficiência Organizacional

Rui Filipe Alves dos Santos Pires

Mestrado em Gestão

Orientador:

Professor Doutor Carlos Miguel Correia Hernandez Jerónimo,
Professor Auxiliar Convidado,
ISCTE - Instituto Universitário de Lisboa

Outubro, 2024



BUSINESS
SCHOOL

Departamento de Marketing, Operações e Gestão Geral

**Gestão de Dados Não Estruturados na Era da
Transformação Digital: Práticas e Impactos na Eficiência
Organizacional**

Rui Filipe Alves dos Santos Pires

Mestrado em Gestão

Orientador:

Professor Doutor Carlos Miguel Correia Hernandez Jerónimo,
Professor Auxiliar Convidado,
ISCTE - Instituto Universitário de Lisboa

Outubro, 2024

"There is nothing outside the text."

Jacques Derrida,

Filósofo, 1930-2004

*À Ana, a minha Voz Decidida,
à Francisca e à Carolina, as minhas maiores realizações!*

Agradecimentos

Gostaria de expressar minha profunda gratidão a todos que me apoiaram durante a jornada desta tese de mestrado. Primeiramente, à minha família, cujo apoio inabalável foi essencial para que eu pudesse dedicar-me plenamente. Em especial, à Ana, Francisca e Carolina, que são o que tenho de mais precioso. O amor e encorajamento que recebi tornaram este processo não apenas mais leve, mas também mais significativo.

Agradeço também ao *ISCTE Executive Education* pelo contributo do mestrado executivo em *Gestão de Programas e Projetos*, que me proporcionou as ferramentas e conhecimentos indispensáveis para dar este importante passo em direção ao mestrado em Gestão. A experiência e as lições aprendidas neste programa foram verdadeiros alicerces que me impulsionaram neste novo desafio.

Sou imensamente grato a todos que contribuíram para a recolha de dados, sem o qual seria impossível realizar esta pesquisa. O caminho foi enriquecido pelas interações e lições compartilhadas.

Por fim, mas não menos importante, meu sincero agradecimento ao Prof. Carlos Jerónimo, cuja orientação foi crucial em momentos de incerteza. A sua habilidade para me manter "acima da linha de água" e as suas instruções precisas sobre como "nadar até à margem" foram essenciais para conduzir este trabalho até à sua conclusão.

A todos, o meu sincero agradecimento!

Resumo

Os dados não estruturados são uma fonte rica de conhecimentos que, se forem bem captados e utilizados, podem ser de utilidade relevante tanto para as empresas como para a sociedade em geral. Mas a ausência de medidas de controlo sobre estes dados pode resultar em incumprimento ou na perda de oportunidades. À medida que a tecnologia evolui, as organizações devem valorizar o papel que os dados não estruturados desempenham e adotar técnicas que encorajem a recolha e a utilização destes dados, adquirindo uma vantagem competitiva e promovendo o estado de consciência da sociedade. Este estudo contribui para um entendimento mais claro da gestão dos dados não estruturados, centrando-se nas barreiras e nos fatores que permitem a sua integração nos sistemas de gestão existentes, acompanhados por processos de regulamentação adequada.

As novas tecnologias, como o Processamento de Linguagem Natural (PNL) e a Aprendizagem Automática (AM), devem ser adotadas, mas em ambientes que valorizem a formação e a criatividade, permitindo que fatores como a cultura de uma organização e os seus recursos possam acelerar a adoção dessas tecnologias. Para beneficiar plenamente dos dados não estruturados, as organizações devem dispor de um quadro de governação claramente definido e incorporar adequadamente novos processos. Por conseguinte, a presente investigação salienta que a conversão de dados não estruturados em recursos com base nas TI permite às empresas não só aumentar a sua eficácia operacional, mas também a capacidade da organização para reagir rapidamente às exigências de um ambiente dinâmico.

Palavras-chave: Transformação Digital, Gestão de Informação, Dados Não Estruturados, Inteligência Artificial, Governação de Dados, Eficiência Organizacional, Desafios de Integração, Segurança e Conformidade, Análise Avançada de Dados.

Códigos de classificação JEL: O33 - Mudança Tecnológica: Escolhas e Consequências; Processos de Difusão e C81 - Metodologia para Recolha, Estimção e Organização de Dados Microeconómicos; Acesso a Dados.

Abstract

Unstructured data is a rich source of knowledge which, if well captured and used, can be of significant use to both companies and society in general. But the absence of control measures over this data can result in non-compliance or lost opportunities. As technology evolves, organizations must value the role that unstructured data plays and adopt techniques that encourage the collection and use of this data, gaining a competitive advantage and promoting society's state of awareness. This study contributes to a clearer understanding of the management of unstructured data, focusing on the barriers and factors that enable its integration into existing management systems, accompanied by appropriate regulatory processes.

New technologies, such as Natural Language Processing (NLP) and Machine Learning (ML), should be adopted but in environments that value training and creativity, allowing factors such as an organization's culture and resources to accelerate the adoption of these technologies. To fully benefit from unstructured data, organizations must have a clearly defined governance framework and properly incorporate new processes. This research therefore highlights that converting unstructured data into IT-based resources enables companies not only to increase their operational efficiency, but also the organization's ability to react quickly to the demands of a dynamic environment.

Keywords: Digital Transformation, Information Management, Unstructured Data, Artificial Intelligence, Data Governance, Organizational Efficiency, Integration Challenges, Security and Compliance, Advanced Data Analytics.

JEL Classification Codes: O33 - Technological Change: Choices and Consequences; Diffusion Processes: and C81 - Methodology for Collecting, Estimating, and Organizing Microeconomic Data; Data Access.

Índice

1.	INTRODUÇÃO	1
1.1	GESTÃO DE “TODOS” OS DADOS	1
1.2	PROBLEMA DE INVESTIGAÇÃO	1
1.3	OBJETIVOS	2
2.	REVISÃO DA LITERATURA	3
2.1	TRANSFORMAÇÃO DIGITAL NA DEFINIÇÃO DA GESTÃO DA INFORMAÇÃO	3
2.2	DADOS NÃO ESTRUTURADOS COMO FOCO ESTRATÉGICO NAS ORGANIZAÇÕES	4
2.3	DATA GOVERNANCE, SEGURANÇA E COMPLIANCE DA INFORMAÇÃO	6
2.4	EFICIÊNCIA ORGANIZACIONAL, DESAFIOS DE INTEGRAÇÃO E ESCALABILIDADE.	8
2.5	ANÁLISE AVANÇADA DE DADOS NÃO ESTRUTURADOS NA ERA DA TRANSFORMAÇÃO DIGITAL	9
3.	METODOLOGIA	13
3.1	MÉTODO	13
3.2	AMOSTRA DO ESTUDO	13
3.3	PROCEDIMENTO DE RECOLHA DE DADOS.....	14
3.4	TRATAMENTO DOS DADOS	15
4.	ANÁLISE DE DADOS	17
4.1	ANÁLISE QUALITATIVA DOS RESULTADOS	17
4.2	ANÁLISE QUANTITATIVA DOS RESULTADOS	18
4.2.1	CORRELAÇÕES ENTRE TECNOLOGIAS	19
4.2.2	MODELAGEM DE REGRESSÃO.....	19
4.2.3	ANÁLISE DE CONSISTÊNCIA E OUTRAS ANÁLISES.....	20
5.	DISCUSSÃO	23
5.1	TRANSFORMAÇÃO DIGITAL E GESTÃO DA INFORMAÇÃO	23
5.2	DADOS NÃO ESTRUTURADOS COMO FOCO ESTRATÉGICO NAS ORGANIZAÇÕES	24
5.3	DESAFIOS DE GOVERNAÇÃO E CONFORMIDADE	25
5.4	EFICIÊNCIA, INTEGRAÇÃO E ESCALABILIDADE	26
5.5	ANÁLISE AVANÇADA DE DADOS NÃO ESTRUTURADOS	27
6.	CONCLUSÕES, PROPOSTAS, LIMITAÇÕES E OUTROS FOCOS DE INVESTIGAÇÃO	29
6.1	CONCLUSÕES.....	29
6.2	PROPOSTA DE PROJETO DE DEFINIÇÃO DE MODELO DE GESTÃO DE DADOS NÃO ESTRUTURADOS	30
6.3	SOLUÇÕES TECNOLÓGICAS PARA A GESTÃO DE DADOS NÃO ESTRUTURADOS.....	32
6.4	LIMITAÇÕES.....	33
6.5	CONTRIBUTO PARA FUTUROS FOCOS DE INVESTIGAÇÃO	34
7.	REFERÊNCIAS BIBLIOGRÁFICAS	37

Índice de figuras, tabelas e anexos

Figura 1 – Demografia profissional	14
Tabela 1 – Resultados da análise quantitativa.....	21
Anexo A - Questionário.....	41

Glossário

AGILE Methodology | Metodologia AGILE - Abordagem de gestão de projetos centrada na flexibilidade e entregas iterativas.

Artificial Intelligence (AI) | Inteligência Artificial (IA) - Simulação de processos de inteligência humana por máquinas, especialmente sistemas computacionais.

Advanced Unstructured Data Repository (AUDR) | Análise de Dados em Tempo Real - Processo de monitorização contínua e análise de dados atuais para insights imediatos.

Big Data | Grandes Volumes de Dados - Conjuntos de dados massivos e complexos que exigem tecnologias e ferramentas especializadas para processamento e análise.

Big Data Analytics (BDA) | Análise de Grandes Volumes de Dados - Processo de recolha, análise e interpretação de conjuntos massivos de dados para extrair informações valiosas.

Blockchain - Tecnologia de registo descentralizado que garante segurança e transparência nas transações, sendo utilizada em diversas aplicações, incluindo a autenticação de dados.

Cloud Service Provider (CSP) | Fornecedor de Serviços em Nuvem - Empresa que oferece serviços de computação em nuvem para indivíduos e organizações.

Cloud Based | Soluções residentes na internet - Refere-se a aplicações ou serviços acessíveis pela internet, permitindo o armazenamento e processamento de dados remotamente.

Compliance | Conformidade - Aderência a regulamentos e políticas relativas à gestão e utilização de dados.

Data Analysis | Análise de Dados - Processo de inspecionar, limpar e modelar dados para descobrir informações úteis e suportar a tomada de decisões. Envolve técnicas estatísticas e ferramentas de visualização, sendo comum em áreas como negócios, ciência e investigação social.

Data Compliance | Conformidade de Dados - Aderência a regulamentos de gestão e proteção de dados.

Data Curation | Curadoria de Dados - Processo de organização e integração de dados de diferentes fontes para os tornar úteis.

Data Governance | Governança de Dados - Definição e implementação de políticas, processos e normas para a gestão de dados.

Data Lake | Lago de Dados - Repositório centralizado que armazena dados brutos no seu formato nativo até que sejam necessários.

Data Warehousing | Armazenamento de Dados - Processo de recolha, integração e gestão de dados de várias fontes num repositório central para análise.

Digital Transformation | Transformação Digital - Processo de utilização de tecnologias digitais para mudar fundamentalmente a forma como uma empresa opera.

Extract-Transform-Load (ETL) | Extração-Transformação-Carga - Processo usado em data warehousing para recolher, transformar e carregar dados de várias fontes num banco de dados de destino.

Metadata Management | Gestão de Metadados - Organização e gestão de metadados para facilitar a descoberta e o uso de dados.

Inteligência Competitiva (IC) - Processo de recolha e análise de informações estratégicas sobre o mercado e concorrentes para apoiar a tomada de decisões.

Insight - Compreensão profunda de um problema ou situação, resultante de análise ou reflexão, que revela padrões ou relações não evidentes, levando a novas ideias ou soluções.

Interoperability | Interoperabilidade - Capacidade de sistemas de computação para trocar e usar informações de forma eficaz.

Internet of Things (IoT) | *Internet das Coisas* - Rede de objetos físicos equipados com sensores, software e outras tecnologias que se conectam e trocam dados.

Machine Learning (ML) | Aprendizagem Automática - Tipo de inteligência artificial que permite que sistemas computacionais aprendam com dados sem programação explícita.

Named Entity Recognition (NER) | Reconhecimento de Entidades Nomeadas – Sub-tarefa da extração de informações que procura localizar e classificar entidades nomeadas mencionadas em texto não estruturado em categorias predefinidas, como nomes de pessoas, organizações, locais, códigos médicos, expressões temporais, quantidades, valores monetários, percentagens, etc.

Natural Language Processing (NLP) | PLN (Processamento de Linguagem Natural): Ramo da inteligência artificial que trata das interações entre computadores e a linguagem humana.

Pequenas e Médias Empresas (PME) - Empresas que normalmente empregam menos de 250 pessoas.

Regulamento Geral sobre a Proteção de Dados (RGPD) - Regulamento da União Europeia sobre proteção de dados e privacidade.

Text Mining | Extração de Texto - Extração de informações relevantes de grandes volumes de texto.

Tecnologias de Informação (TI) - Tecnologia utilizada na criação, armazenamento, recuperação, troca e manipulação de dados.

Structured Data | Dados Estruturados - Dados organizados num formato predefinido, como bancos de dados relacionais.

Unstructured Data | Dados Não Estruturados – Dados que não têm um formato ou organização predefinidos, como documentos de texto, imagens, áudio e vídeo

1. Introdução

1.1 Gestão de “todos” os dados

There is nothing outside the text! (ou "Não há nada fora do texto") de Derrida, J. (1967) é uma frase frequentemente citada que provoca diversas interpretações no campo da crítica literária e filosófica. Uma das interpretações centrais dessa frase sugere que o significado de um texto não pode ser compreendido sem considerar o contexto e as relações ao seu redor; em outras palavras, o texto deve ser analisado dentro da sua própria estrutura e nas múltiplas interações que ele estabelece. Deste modo, o desenvolvimento deste estudo pretende realçar a relevância de todos os dados contextuais da intervenção de uma organização, quer sejam estruturados ou não estruturados. A diferença marcante entre eles é que os dados estruturados estão organizados em sistemas, como bases de dados, enquanto os dados não estruturados, incluindo emails, documentos de texto, imagens, som e vídeos, pela sua heterogeneidade possuem distintas formas de recolha, organização e arquivo. A correta gestão dos dados não estruturados é relevante para um entendimento mais claro do ambiente onde a organização se insere, quer interna, quer externamente.

Estamos hoje perante uma era de digitalização em que as empresas estão a incorporar tecnologias digitais nas suas operações, transformando-as nos seus processos, cultura e tecnologia e gerando um volume cada vez maior de dados: estruturados e não estruturados. A gestão eficaz dos dados não estruturados pode ser vista como um fator crucial para a criação de uma vantagem competitiva nas organizações. Conforme destacado por Ma *et al.* (2022), a melhoria da qualidade na partilha de informações e nos processos de rastreabilidade é essencial para a sustentabilidade e desempenho das operações. Ao integrar dados contextuais da intervenção de uma organização, tanto em níveis internos quanto externos, é possível não apenas otimizar a gestão da informação, mas também alinhar as práticas operacionais às necessidades do mercado em tempo real.

1.2 Problema de Investigação

É preciso lidar com muitas barreiras na gestão de dados não estruturados. Vărzaru *et al.* (2022) referem que o aspeto mais significativo do aumento da eficácia operacional da organização é a integração de novas tecnologias que garantam o fornecimento de informações vitais para a tomada de decisões de gestão atempadas e sólidas. A ausência de políticas adequadas ou a implementação ineficaz da governação dos dados tende a criar um ambiente propício a um maior risco com a segurança e privacidade dos dados, para além de impactar com a gestão da própria empresa. Além disso, Beheshti

et al. (2016) também salientam que os dados e as informações que não são exatos e honestos distorcem a qualidade das decisões tomadas. Um bom exemplo é o caso das organizações financeiras. O facto de não conseguirem obter informações exatas desses conjuntos de dados leva a conclusões erradas baseadas em avaliações erróneas que não apoiam os objetivos organizacionais (Ma *et al.* 2022).

Além disso, a integração de dados não estruturados noutros sistemas de informação previamente existentes, na sua maioria estruturados, torna-se malsucedida ou bastante complicada, o que, tal como referido por Anthony *et al.* (2021), pode criar impacto negativo nas operações e privar a organização de boas oportunidades de negócio. A crescente procura de desmaterialização de processos conduziu ao crescimento de espaço das bases de dados que armazenam um volume de dados cada vez maior. No entanto, na ausência do emprego de estratégias que envolvam a integração da informação, como observam Alabdali & Salam (2022), resultará na incapacidade de melhorar os indicadores de desempenho da empresa, causando atrasos na perfeita adaptação às mudanças do mercado.

Além disso, não ter em consideração as questões relacionadas com a escalabilidade dos sistemas de gestão de dados pode resultar numa situação em que a organização não tem outra opção senão ficar limitada a crescer para além das barreiras estabelecidas pelas tecnologias da informação (TI). Isto está de acordo com Vimal (2021), que salienta a necessidade de uma abordagem mais agressiva da governação dos dados para minimizar essas ameaças.

1.3 Objetivos

A investigação procura analisar e destacar os métodos e as tecnologias de gestão de dados não estruturados no contexto da mudança digital em curso, centrando-se em como contribui para o desempenho da organização. O objetivo central é avaliar as oportunidades oferecidas pela transformação digital em termos de melhoria da gestão de dados, de modo a criar informação que otimize a competitividade das empresas respondendo à questão: como assegurar a gestão de dados não estruturados tendo em vista melhorar o desempenho da organização e que práticas podem ser utilizadas?

Um objetivo subjacente, é avaliar como as melhores práticas de governação de dados devem integrar e equilibrar a proteção e a gestão da informação como resposta à evolução dos desafios em matéria de privacidade e segurança. Além disso, este estudo procura fornecer orientações e recomendações práticas que ajudem as organizações a gerir os dados não estruturados de uma forma mais coordenada que incorporem não só as tecnologias emergentes, incluindo a Inteligência Artificial, mas também a implementação estratégica de dados no contexto organizacional no ambiente atual.

2. Revisão da literatura

2.1 Transformação Digital na definição da gestão da informação

O processo de transformação digital nesta nova era pode ser descrito como a adoção em massa da globalização e da internet, que não só se torna parte integrante de todos os negócios, mas também gera profundas transformações na estrutura do próprio negócio. Para ajudar, o impacto da transformação precisa também ser comentado em diversas esferas da sociedade, que incluem o aprimoramento da cultura da gestão do conhecimento, mudança de estruturas sociais e interações construídas através de tecnologias de gestão mais eficazes.

Um dos impactos mais relevantes da digitalização é, sem dúvida, a maneira como as informações são geridas e compartilhadas nas organizações. Autores sugerem que, através da transformação digital, as capacidades de partilha, que são altamente essenciais para o desempenho da cadeia de abastecimento, bem como a eficácia operacional, são melhoradas. Por exemplo, Ma *et al.* (2022) salientam que o processo de transformação digital provocou uma melhoria na partilha de informações e nos processos de rastreabilidade na cadeia de abastecimento farmacêutica, melhorando assim a sustentabilidade na cadeia de abastecimento da empresa. Tal afirmação é apoiada por Alabdali & Salam (2022) que destacam que o processo de transformação digital aumenta em velocidade a procura pelos mecanismos de compartilhamento eficientes, que são vitais na aquisição de vantagem competitiva na cadeia de abastecimento.

Além disso, Vărzaru *et al.* (2022) advogam que o emprego de tecnologias digitais para a realização de funções rotineiras não apenas agiliza o resultado, mas também aumenta a qualidade e a disponibilidade de informações para a tomada de decisões de gestão eficazes. Por sua vez, a transformação digital influencia diretamente as evoluções nas estruturas organizacionais, o que é de extrema relevância quando se trata da gestão da informação. Bonanomi *et al.* (2019) propõe a análise das relações entre diferentes estruturas organizacionais formais e informais e a informatização nos meios de ensino de arquitetura e engenharia. A partir desse estudo, eles constataram que novas funções e relações que foram criadas com o uso de tecnologias digitais passaram a promover uma melhor informação e interação entre os membros da equipa.

Para tal, é fundamentada a análise proposta por Zhang *et al.* (2023). De acordo com a linguagem, a integração de tecnologias digitais implica obrigatoriamente numa reconstrução dos modelos e estruturas organizacionais dos processos para melhorar a economia da informação. Os resultados sugerem que as organizações devem modificar suas estruturas organizacionais a partir da compreensão dos impactos da transformação digital para melhorar suas práticas de gestão da informação.

A digitalização trouxe inúmeras mudanças na forma como os processos de gestão da informação funcionam. Introduziu novas práticas e tecnologias de gestão, como refere Mokhtar *et al.* (2020). Estes autores sugerem que o processo de transformação digital é tal que existem produtos, serviços e processos novos e radicais que fazem a organização repensar algumas das suas estratégias de informação. Feroz *et al.* (2021) argumentam que essa abordagem permite que as empresas ofereçam ao mercado modelos de negócios mais amigáveis ao meio ambiente, o que altera as próprias abordagens de sustentabilidade e de comunicação com os *stakeholders* sobre essas informações. Além disso, Mu (2023) esclarece que a esfera de influência da transformação digital é ampla, pois também melhora a qualidade da informação contabilística. Ou seja, o uso de ferramentas digitais melhora a qualidade da informação financeira, que é fundamental para a tomada de decisão.

Além disso, pode-se dizer que a liderança desempenha um papel importante para facilitar a transformação da área de gestão da informação. Como defendem Mukhtamar B *et al.* (2023), as organizações devem possuir uma liderança capaz de lidar com as especificidades da transformação digital, bem como abranger as novas expectativas impostas à gestão da informação através da nova lente digital. Essa liderança é importante para nutrir uma mentalidade de cultura organizacional e inventividade que é muito crítica na adoção efetiva de tecnologias digitais destinadas a aumentar a eficácia da gestão da informação. Os princípios de investigação de Warner & Wager (2019) destacam ainda a importância deste desenvolvimento no contexto da promoção do desenvolvimento de capacidades dinâmicas para efeitos de transformação digital, que se centra na transformação contínua das estratégias e práticas de gestão da informação, de modo a incorporar eficazmente as tecnologias digitais contemporâneas nos processos da empresa.

No que diz respeito à gestão da organização e dos seus colaboradores, tem-se assistido a uma mudança de paradigma na forma como as pessoas e as organizações do sector gerem e armazenam a informação. Trata-se da representação da gestão na estrutura das organizações, e isto tem a ver com o facto de a tecnologia estar a mudar a gestão enquanto tal. Como nos dizem Zhang, L. & Chen, Z. (2023), as abordagens de gestão dos trabalhadores em termos de processos de recursos humanos em relação à sua informatização também mudam, e esta prática de gestão específica é melhorada com o contributo dos trabalhadores. Esta mudança na economia política exige a modificação de práticas de gestão de recursos humanos mais sensíveis à mudança, uma vez que as ferramentas digitais melhoram a gestão e a análise da informação, o que torna todo o processo de tomada de decisão mais eficiente.

2.2 Dados não estruturados como foco estratégico nas organizações

No atual panorama da estratégia organizacional, surgiu uma vasta área de enfoque nos dados não estruturados, que se integram no processo central de tomada de decisões estratégicas de uma

organização e ajudam a criar uma vantagem competitiva. As organizações passaram a apreciar os dados não estruturados, que podem ser derivados de texto, imagens, áudio e vídeo. Este tipo de dados é fundamental para revelar informações que podem ser utilizadas para orientar estratégias de resolução de problemas, bem como para melhorar a eficácia operacional. Onwujekwe *et al.* (2020) sublinham que os dados não estruturados devem ser incluídos nos sistemas de gestão do conhecimento para se poder compreender a estratégia relacionada com a tomada de decisões numa organização, uma vez que esses dados são altamente diversificados e podem ser capturados sistematicamente. Além disso, as conclusões de Beheshti *et al.* (2016) amplificam a necessidade de refinamento de dados, porque os dados brutos não estruturados têm de ser transformados em inteligência útil para o apoio na tomada de decisões. A utilização efetiva de dados não estruturados para obter vantagens competitivas pode ser observada na inteligência competitiva (IC). Lutz & Bodendorf (2020) afirmam que a gestão organizada de dados qualitativos não estruturados tem o potencial de auxiliar não apenas o nível estratégico, mas também o nível operacional, disponibilizando assim informações críticas a diferentes níveis da organização. Isto não só ajuda nos processos de tomada de decisão, como também promove uma mentalidade orientada para os respetivos dados em toda a organização.

Além disso, Rave *et al.* (2021) sublinham a importância da análise de texto no processo de tomada de decisão multicritério e o papel que os dados não estruturados podem desempenhar nos debates posteriores sobre as questões estratégicas complexas. Além disso, as abordagens automatizadas, como a análise de ML e IA de dados não estruturados, alteraram a forma como as organizações extraem informações. Por exemplo, Wang, J. (2023) aborda a utilização destas tecnologias na otimização da estratégia de investimento através da análise de dados não estruturados, o que expande as fronteiras da análise quantitativa em finanças. No mesmo contexto, Kudalkar (2023) ilustra o poder preditivo da análise de sentimentos na utilização de dados dos meios de comunicação social, que é um tipo de dados não estruturados, e na promoção da interatividade e do marketing estratégico de produtos.

Mas a gestão de dados não estruturados também abre espaço a desafios. Como salientaram Wei & Liu (2010), os sistemas convencionais de bases de dados relacionais revelam-se frequentemente inadequados quando confrontados com a magnitude do volume e da complexidade dos dados não estruturados. Esta limitação exige um conjunto totalmente novo de estruturas e sistemas, como o *Advanced Unstructured Data Repository* (AUDR), que são eficientes na acomodação e no processamento de uma variedade de dados não estruturados (Liu *et al.*, 2011). Além disso, Adnan & Akbar (2019) observam que a diversidade e a riqueza dos grandes dados não estruturados requerem melhores métodos de extração de informação e as organizações têm de ser inteligentes na utilização destes dados se quiserem colher os seus benefícios.

2.3 Data Governance, Segurança e Compliance da Informação

Nesta era de transformação digital, os dados não estruturados surgem como um ativo relevante para as organizações. O desenvolvimento exponencial de tecnologias digitais conduz à criação e à recolha de grandes quantidades de dados não estruturados que apresentam desafios quanto à sua governação, segurança e conformidade.

A governação de dados é fundamental para garantir que as organizações podem gerir eficazmente os seus dados estruturados e não estruturados, enquanto cumprem os requisitos regulamentares e as políticas internas. O caminho para a transformação digital necessita de uma estrutura robusta de governação de dados que englobe a sua qualidade, origem e administração. Huang, Z. (2023) sublinha a importância de uma abordagem equilibrada à transformação digital nas instituições de ensino, destacando a necessidade de uma proteção vigilante dos ecossistemas digitais contra ameaças de cibersegurança e violações da privacidade dos dados. Este sentimento é partilhado por Ullah *et al.* (2021), que salientam a importância da agregação e transmissão seguras de dados no contexto da *internet of things* (IoT), sublinhando a necessidade de quadros de governação abrangentes para proteger informações sensíveis.

Além disso, a integração dos sistemas *cloud based* nas operações organizacionais veio enfatizar a necessidade de implementar uma governação dos dados. Vimal (2021) salienta que o acesso não autorizado e a perda de dados em ambientes de nuvem podem levar a graves repercussões, incluindo perdas financeiras e danos à reputação. Portanto, as organizações devem cultivar uma compreensão completa dos conceitos de segurança de dados e das práticas recomendadas para mitigar esses riscos. O estabelecimento de políticas claras de governação de dados é essencial para garantir que os dados não estruturados são geridos de forma eficaz, com controlos de acesso adequados e medidas de conformidade em vigor.

Como as organizações dependem cada vez mais dos fornecedores de serviços *Cloud* (CSP) para o armazenamento e processamento de dados, a segurança dos dados torna-se uma preocupação premente. A implementação de medidas de segurança robustas é fundamental para proteger contra o acesso não autorizado e as violações de dados. Wang *et al.* (2019) discutem vários métodos de proteção de dados, incluindo técnicas criptográficas e marcas de água, que podem ser empregues para salvaguardar informações privadas durante a transmissão. Além disso, a adoção da tecnologia de *blockchain* tem sido proposta como um meio de melhorar a segurança dos dados em ambientes *Cloud*, como salientam Benavides *et al.* (2020), que exploram o potencial da *blockchain* para a partilha segura de dados em instituições de ensino superior.

Para além da governação e da segurança, a conformidade com os quadros jurídicos e regulamentares é um aspeto vital da gestão de dados não estruturados. O aumento dos regulamentos

de proteção de dados, como o Regulamento Geral sobre a Proteção de Dados (RGPD), exige que as organizações implementem medidas rigorosas para garantir a privacidade e a segurança dos dados. Olaniyi (2023) salienta o papel das agências reguladoras na supervisão das práticas de gestão de dados, observando que as organizações devem tratar os dados dos utilizadores com cuidado e a devida diligência para minimizar os riscos de exploração. Este cenário regulamentar obriga as organizações a adotarem estratégias de conformidade proactivas que englobem a classificação dos dados, os controlos de acesso e os protocolos de resposta a incidentes.

Os desafios associados aos dados não estruturados vão para além da governação, da segurança e da conformidade; abrangem também a necessidade de práticas eficazes de gestão de dados. O grande volume e a complexidade dos dados não estruturados exigem a utilização de tecnologias e metodologias avançadas para o processamento e análise de dados. Tu, Z. (2023) destaca as vulnerabilidades inerentes à tecnologia da informação informática, salientando a necessidade de técnicas de encriptação inovadoras para proteger os dados sensíveis. Além disso, a aplicação da aprendizagem automática e da inteligência artificial na análise de dados pode melhorar a capacidade de extrair conhecimentos significativos de dados não estruturados, impulsionando assim a inovação e a tomada de decisões informadas.

Além disso, a integração da análise de grandes volumes de dados nos processos organizacionais pode facilitar a identificação de padrões e tendências nos dados não estruturados. Yang *et al.* (2019) defendem a utilização de métodos de classificação hierárquica para proteger a segurança da informação na era dos megadados, sugerindo que tais abordagens podem alterar qualitativamente as práticas de gestão e governação de dados. Ao tirar partido da análise avançada, as organizações podem obter uma vantagem competitiva e, ao mesmo tempo, garantir que as suas estruturas de governação de dados permaneçam robustas e em conformidade com as normas regulamentares.

À medida que as organizações continuam a adotar a transformação digital, a importância de promover uma cultura de segurança e conformidade dos dados não pode ser exagerada. O rápido avanço da tecnologia exige uma educação e formação contínuas dos funcionários para garantir que estão equipados para lidar com dados não estruturados de forma responsável. Spivakovskyy *et al.* (2021) salientam a necessidade de melhorar as medidas de segurança da informação em resposta aos desafios colocados pela digitalização, defendendo uma abordagem proactiva à gestão de dados que dê prioridade à segurança e à conformidade. Ao cultivar uma cultura de consciencialização e responsabilização, as organizações podem navegar melhor pelas complexidades da gestão de dados não estruturados na era digital.

2.4 Eficiência Organizacional, Desafios de Integração e Escalabilidade.

Como mencionado, os dados não estruturados são constituídos por diferentes tipologias de textos, imagens e sons, mas o principal problema deste tipo de dados é o seu formato, ou seja, a ausência de uma estrutura rígida que dificulta a sua integração nos sistemas existentes e, mais ainda, a sua análise. Em contextos semelhantes, o problema é ainda mais acentuado em áreas como o marketing, as finanças e o retalho, em que os modelos de negócio da organização se baseiam em dados para otimizar o processo de tomada de decisões e melhorar as interações com os clientes (Miklošík & Evans, 2020; Xie, 2024; Zou, 2024).

A inclusão de dados não estruturados nos processos de integração nas organizações é difícil, sobretudo devido aos diferentes bancos de dados e sistemas existentes em muitas organizações. Como observam Anthony *et al.* (2021), é a integração da informação das várias unidades a partir dos sistemas remotos que dificulta a concretização dos objetivos pretendidos com a digitalização. Esta fragmentação não só tem o efeito de tornar obsoleta a disponibilidade de dados, como também cria uma amálgama de processos analíticos que precisam ser empregues para derivar decisões-chave. Além disso, a falta de formatos comumente aceites para os dados agrava os problemas de integração desses dados e permite que as organizações que pretendem tirar partido dos seus recursos de dados o façam de uma forma ineficaz e inadequada (Jiang, 2023).

Atualmente, as organizações procuram cada vez mais empregar tecnologias modernas como *big data analytics* (BDA) e machine learning (ML). Estas últimas ajudam no desenvolvimento de processos de integração que geram grandes quantidades de dados não estruturados, permitindo que as organizações reconheçam e analisem tendências e padrões que estavam irremediavelmente perdidos (Miklošík & Evans, 2020; Chan J., 2020). Por exemplo, a utilização de tecnologias de aprendizagem automática pode ser empregue para prescrever a identificação e a estruturação dos dados não estruturados, o que consequentemente poupa tempo durante o processo de integração e aumenta a eficácia geral das estratégias de gestão de dados das organizações (Zhu, 2023; Tripathi, S., 2023). Além disso, a integração de soluções *cloud based* permite que as organizações lidem com volumes de dados dinâmicos de uma forma que lhes permita responder mais rapidamente ao seu mercado interno e externo (Xie, 2024; Chan J., 2020).

A escalabilidade é um aspeto que tem de ser tido em conta quando se empreende uma transformação digital, que é direcionada para satisfazer a crescente procura de dados dentro da organização. É vital que a organização implemente eficazmente soluções de gestão de dados lineares para obter uma produtividade ótima nos mercados em constante evolução e competitivos. No entanto, Zou (2024), indica que a adaptação dos sistemas de gestão pelas empresas com a incorporação de tecnologias digitais está a conduzi-las para o desenvolvimento de estruturas

operacionais empresariais mais escaláveis e sustentáveis. Esta transformação exige a redefinição das atuais estruturas de dados existentes e o desenvolvimento de quadros resilientes que possam ser modificados eficazmente de acordo com a evolução dos requisitos de dados (Xie, 2024).

Esta abordagem só é possível se existir uma forte liderança na organização, que é fundamental para defender uma cultura de dados. À medida que as organizações se transformam digitalmente, os conjuntos de competências dos líderes em toda a hierarquia mudam, o que exige que se concentrem na inteligência estratégica e na comunicação eficaz (Wang, M. 2023). Os líderes devem encorajar a incorporação de dados não estruturados nos processos das organizações, de modo que os dados sejam considerados como recursos empresariais em vez de serem vistos como resultados de processos. Embora seja uma alteração cultural que tem de ser feita para vencer o conflito da mudança, bem como para garantir que a utilização de dados é uma estratégia adotada a todos os níveis da organização (Yuanming, 2023; Lu & Zhou, 2022).

Trazer novos talentos digitais também é crucial para enfrentar os desafios de competências pessoais que normalmente fazem parte dos processos de transformação digital. As organizações precisam de se concentrar na contratação de profissionais com competências em *data analysis*, *machine learning* e *cloud based technologies* para poderem aproveitar plenamente o potencial dos dados não estruturados (Gilch & Sieweke, 2020; Jiang, 2023). Esta estratégia de capital humano não só melhora as capacidades da organização, como também ajuda a fomentar a criatividade e a capacidade de resposta às necessidades do mercado.

2.5 Análise avançada de dados não estruturados na era da transformação digital.

O termo *big data* refere-se ao grande volume de dados que é gerado a partir de diversas fontes e que, devido à sua complexidade e variedade, requer tecnologias e métodos avançados para armazenamento, processamento e análise.

Este termo tornou-se onnipresente no mundo empresarial devido às possibilidades económicas associadas e aos benefícios sociais na resolução de problemas do mundo real. As organizações adotaram uma abordagem tecnológica para responder a esta necessidade, incluindo a análise de grandes volumes de dados, a IA e a soluções *cloud based*. É bastante claro que a maioria das organizações têm uma tendência contínua para adotar estas tecnologias e, assim, gerar uma grande quantidade de dados não estruturados que têm o potencial de revelar conhecimentos valiosos quando utilizados corretamente (Motjoloopane, 2023; Feroz *et al.*, 2021).

O conceito de parte dos dados não estruturados tem como um dos seus significados voz, vídeos e textos. As abordagens clássicas à análise de dados e à sua eficiente tomada de decisões e gestão

empresarial parecem falhar mais frequentemente nesta área, o que não é surpreendente dada a diversidade e variabilidade que estes tipos de dados possuem. A extração de texto e a análise de sentimentos, por outro lado, apresentam resultados promissores no domínio da análise de informações não estruturadas quando combinadas com técnicas de ML e de PLN (El-Taliawi *et al.*, 2021). Por exemplo, a extração de texto e a análise de sentimentos permitem o envolvimento do cliente, recuperando dados valiosos da secção de comentários, dos contactos nas redes sociais e de outros recursos baseados em texto (El-Taliawi *et al.*, 2021; Vasilopoulou *et al.*, 2023).

Os dados não devem ser analisados apenas nos processos das grandes empresas. Da mesma forma, a análise de dados não estruturados é absolutamente crítica, para dizer o mínimo, quase sempre quando se trata de pequenas e médias empresas (PME), uma vez que estas raramente dispõem dos recursos para uma análise de dados automatizada e eficaz. No entanto, é precisamente nestas PME que a implementação da utilização de tecnologias de digitalização, com o objetivo de aumentar a sua eficiência e competitividade, deve ser melhor implementada. Há indicações de que se as PMEs passarem por processos de transformação digital, estarão em melhor posição para responder às mudanças do mercado e dos consumidores (Parra-Sanchez, 2023; Pan & Lee, 2020). Assim, estes indivíduos ou empresas veem os processos de negócio onde os dados são tratados como uma ferramenta para obter mais negócios, bem como o desenvolvimento desses negócios (Król & Zdonek, 2020).

Além disso, à medida que análises mais sofisticadas são incorporadas nos processos de negócio, torna-se mais fácil compreender as condições de mercado num contexto mais amplo. Neste sentido, no setor da indústria, por exemplo, tem sido demonstrado que a aplicação de modelos de análise de dados é benéfica para a vantagem competitiva através de um controlo mais robusto das cadeias de abastecimento ou mesmo de processos de eficiência produtiva (Song *et al.*, 2022; Xu *et al.*, 2020). Tudo isto acaba por resultar em avanços em larga escala, como maior eficiência operacional e redução de custos através da utilização eficiente de dados não estruturados obtidos a partir de redes IoT, bem como de consumidores finais (Zeng & Wang, 2023).

A integração da análise de dados não estruturados não só é valiosa para melhorar o desempenho operacional, como também é essencial para melhorar a experiência do cliente. As organizações que se envolvem ativamente com a informação recolhida dos clientes podem personalizar os seus produtos e serviços de acordo com os clientes, alcançando assim taxas de retenção mais elevadas (Halim, 2023; Sánchez *et al.*, 2022). Com a ajuda de ferramentas sofisticadas de análise de dados, as organizações conseguem identificar e analisar a composição da sua base de clientes, o seu comportamento de consumo e até o seu comportamento futuro, o que é vital para o planeamento de estratégias de marketing e de produto (Hussain *et al.*, 2023).

Outra forma de olhar para este dilema é observar a maturidade digital da organização. Estudos realizados por Aagaard *et al.* (2021) e Santiso (2022) sugerem que as empresas ou organizações que se promovem em relação às outras no ranking de maturidade digital são geralmente mais esclarecidas sobre o que lhes falta e como chegar lá. Esta análise passa normalmente por uma análise bastante profunda das interfaces tecnológicas previstas para a organização, da gestão e desenho do seu espaço de dados e da estrutura e competências dos seus recursos humanos. Desta forma, juntamente com a compreensão do problema, as organizações podem planejar e atribuir recursos à tecnologia apropriada e a cursos adequados para melhorar os seus processos de análise de dados. (Król & Zdonek, 2020).

Outra área crítica em que a análise de dados não estruturados pode ajudar é a interação entre a transformação digital e a sustentabilidade ambiental. Os recursos são desperdiçados de várias formas, incluindo o consumo, a produção de resíduos e as emissões, à medida que as organizações continuam a trabalhar para atingir um objetivo de sustentabilidade (Feroz *et al.*, 2021). Estes autores defendem que as organizações podem tirar partido destas tecnologias avançadas para melhorar a eficiência e reduzir os impactos ambientais, a fim de conciliar as estratégias empresariais com os objetivos de sustentabilidade (Vasilopoulou *et al.*, 2023; Qian, 2023).

A entidades governamentais que utilizam dados não estruturados têm o potencial de aumentar o nível de envolvimento dos cidadãos e a eficácia dos mecanismos de prestação de serviços (Vasilopoulou *et al.* 2023; Santiso, 2022). Através da utilização de dados provenientes de diferentes fontes, as agências governamentais são capazes de compreender as necessidades e os desejos dos cidadãos e de adaptar os serviços para serem mais direcionados na futura atribuição de recursos (Mendonça & Dantas, 2020). Este modelo promove a responsabilização, bem como a confiança entre os cidadãos e as instituições governamentais.

No ambiente atual, em que as organizações estão a passar por uma transformação digital, é evidente que a governação de dados, as políticas e os processos têm de ser implementados. Quando os dados são devidamente governados, apresentam-se como um ativo que pode ser aproveitado pelas organizações para melhorar a qualidade dos dados, a segurança e os requisitos de conformidade (Król & Zdonek, 2020; Gaffley & Pelser, 2021). Isto é especialmente importante no caso de dados não estruturados, em que a ausência de formatos universais pode criar discrepâncias e falhas no processo de avaliação. Ao dispor de regras e instruções bem definidas para a realização de processos de gestão de dados, as organizações são muito provavelmente capazes de melhorar a sua capacidade em relação aos dados não estruturados (El-Taliawi *et al.*, 2021).

3. Metodologia

Neste capítulo, é apresentada a abordagem metodológica que foi utilizada neste estudo que teve como objetivo obter as perspetivas dos profissionais em relação à sua experiência no tratamento de dados não estruturados. O questionário (anexo A) construído teve por base a revisão de literatura e foi concebido para responder à questão principal da investigação: como assegurar a gestão de dados não estruturados tendo em vista melhorar o desempenho da organização e que práticas podem ser utilizadas?

3.1 Método

Foi utilizada uma metodologia quantitativa para recolher dados primários. O questionário foi o instrumento de recolha desses dados, tendo sido devidamente organizado de modo a que permitisse a colocação de diferentes tipos de perguntas destinadas a recolher diferentes tipologias de dados. O questionário foi construído utilizando uma plataforma online (www.typeform.com) disponibilizando-se um link para acesso.

3.2 Amostra do estudo

A amostra do estudo foi constituída por profissionais que trabalham em diversas áreas, tendo sido enviada para diferentes entidades através de redes sociais e outros meios de comunicação profissionais. O critério utilizado para a seleção foi assegurar que os grupos que receberiam o convite para participar no questionário teriam diferentes níveis de experiência, a fim de obter diferentes pontos de vista e conhecimento que refinassem qualitativamente os resultados.

Para efeitos desta investigação, junto com o envio do link foi fornecida informação suficiente sobre o objetivo do estudo e a forma como os seus dados estariam protegidos de serem utilizados. Foram tomadas medidas adequadas para garantir a confidencialidade e a segurança dos dados dos participantes.

O questionário foi acedido por 65 pessoas tendo participado 43 (66%). 24 concluíram o questionário, o que corresponde a 37% da amostra.

A nível da demografia dos questionados, destacamos os seguintes dados:

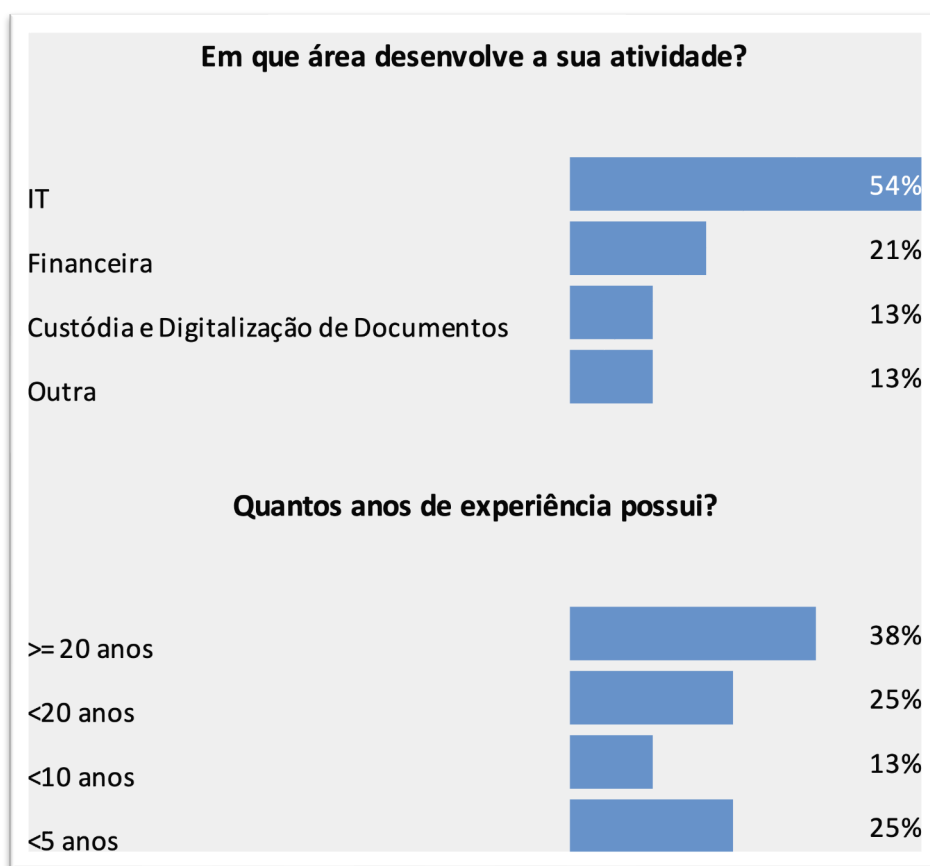


Figura 1 – Demografia profissional

Os dados mostram que, quanto à experiência profissional mais de 63% possuem acima de 10 anos, e 37,5% possuem mais de 20 anos. Quanto à área de atividade o maior peso é atribuído a áreas de IT (54,2%) e financeira (20,8%).

3.3 Procedimento de recolha de dados

A recolha de dados foi efetuada através de um questionário acessível online através de um link enviado na mensagem de convite para a participação. Este link permitia o acesso exclusivo ao questionário, sem acesso a respostas de outros questionados, durante um período limitado. O questionário foi elaborado de forma a ser breve (tempo médio 16 min).

Incluímos no questionário:

- Perguntas de escolha múltipla: perguntas com várias respostas preparadas e em que se esperava que os inquiridos escolhessem as mais adequadas para eles. Estes poderiam ainda excluir as propostas e indicar outra resposta numa pergunta posterior.

- Escala de *Likert*: estas perguntas visavam obter o grau de opinião (entre 0 e 10) dos questionados acerca da não utilização de técnicas e ferramentas de gestão.
- Perguntas abertas: que permitiam aos inquiridos dar respostas com maior detalhe, de modo a obter uma apreciação qualitativa.

Esta estrutura de questionário tinha como objetivo recolher informações qualitativas e quantitativas que enriquecessem os dados para posterior análise e extração de *insights*.

3.4 Tratamento dos dados

Os dados recolhidos foram processados utilizando estratégias de análise estatística descritiva e inferencial. As respostas ao questionário foram introduzidas em ferramentas de apoio como o *MS Excel* e ferramentas de IA tendo como objetivo:

- *Análise descritiva*: ou seja, apreciar as características gerais da população em estudo.
- *Análise de correlação*: determinar a relação entre as variáveis em questão.
- *Modelação de Regressão*: identificar os fatores determinantes que explicam a utilização de tecnologias de gestão de dados não estruturados.

A análise qualitativa das respostas abertas foi efetuada através do processo de codificação temática, em que os dados ajudaram a compreender a gestão de dados não estruturados, bem como os desafios enfrentados pelos profissionais. Para ajudar na compreensão dos resultados qualitativo foram utilizadas ferramentas de IA.

4. Análise de dados

Este capítulo apresenta uma análise dos dados obtidos através de resposta ao questionário (anexo A). A análise aplica um método misto, tanto quantitativo como qualitativo, de modo a dar uma visão global dos padrões e tendências verificadas. A análise quantitativa, que se baseia em estatísticas descritivas e inferenciais, examina a relação entre as diferentes tecnologias utilizadas no tratamento de dados não estruturados (PLN, aprendizagem automática, extração de texto e similares) e os efeitos de fatores contextuais, como a área de atividade e experiência, na utilização dos processos mencionados. Algumas das técnicas analíticas aplicadas nesta parte incluem análise de correlação, regressão linear simples e múltipla e testes de normalidade para avaliar a validade dos pressupostos dos modelos estatísticos.

A componente qualitativa da análise utiliza o conteúdo das perguntas abertas, explorando os dados textuais dos participantes na escolha de determinadas tecnologias e as suas opiniões sobre a gestão de dados não estruturados, a governação dos dados e o efeito das tecnologias emergentes.

4.1 Análise qualitativa dos resultados

As opiniões dos entrevistados em relação à forma não estruturada dos dados indicam que há uma mudança no sentido de apreciar o valor desses dados no contexto organizacional quando se tomam decisões e se melhora a prestação de serviços. A análise qualitativa revela vários ângulos que os profissionais consideram fulcrais para o tratamento eficaz destes dados.

- *Perspetiva das tecnologias emergentes:* para os inquiridos, a IA e a aprendizagem automática, entre outras tecnologias emergentes, são recursos importantes que podem mudar a forma como os dados não estruturados são tratados. Este tipo de entusiasmo está presente em todas os grupos, independentemente da sua experiência profissional, sugerindo que se trata de uma convicção mais alargada de que os desenvolvimentos nestas tecnologias irão automatizar processos e garantir novas perspetivas, conhecimentos, segurança e privacidade. Por outro lado, a informação de que a tecnologia é, por si só, altamente complexa e de difícil utilização foi também outra reflexão comum, apontando para a necessidade de mecanismos que apoiem a supervisão humana da tecnologia.
- *Barreiras de escalabilidade e desempenho:* no que diz respeito à sua perspetiva sobre como lidar com as barreiras causadas por problemas de escalabilidade e desempenho, a maioria dos inquiridos indica que a implementação de soluções como repositórios e tecnologias de armazenamento distribuído, como os *data lake*, é benéfica. A maioria entende que estas soluções não só são vantajosas para aumentar a eficiência durante o processamento de uma

grande quantidade de dados, como também a integração de diferentes fontes de dados é crucial. Mas os obstáculos à inovação e à implantação de novas tecnologias continuam a incluir processos ineficientes e a ausência de escolhas tecnológicas mais coerentes, o que indica algum grau de dinamismo nas suas abordagens e uma procura persistente de eficiência.

- *A governação dos dados e a sua relevância para os dados não estruturados:* a governação dos dados é considerada uma das áreas mais importantes, uma vez que assegura a definição de regras e normas sobre os dados não estruturados. O maior número dos inquiridos especifica a gestão de metadados e as políticas de acesso e retenção como as principais práticas de aplicação da qualidade e conformidade dos dados. O reconhecimento dessa situação enfatiza a compreensão da complexidade relacionada à governança e à necessidade de mecanismos estruturados que evitem desordens e inconsistências nos dados.
- *Efeito das Políticas na Gestão de Dados:* uma questão central que todos concordam em relação aos efeitos da falta de políticas eficazes de gestão de dados é a probabilidade de efeitos tangenciais e críticos, por exemplo, aumento de custos e falta de competitividade. A maioria dos entrevistados insiste que também são necessárias políticas adequadas e fortes para garantir que os dados provêm da imagem de uma organização. Isto faz sentido porque os inquiridos têm uma boa percepção da governação de dados como sendo uma solução empresarial dentro de um contexto maior.
- *Recolha e proteção de dados:* a recolha de dados não estruturados, tais como críticas de consumidores e processos de verificação relacionais prolongados, é certamente uma grande preocupação. As sugestões para implementar a análise de sentimentos e a modelação de tópicos mostram que há um esforço para se orientar para a extração de valor dos dados disponíveis. Em relação à segurança, o emprego de práticas otimizadas, como a encriptação, e de tecnologias, como a cadeia de blocos, mostra uma grande preocupação com os riscos enfrentados no tratamento de dados sensíveis.

4.2 Análise quantitativa dos resultados

A informação e os padrões relativos à recolha de dados foram intrigantes, especialmente no que diz respeito à utilização do Processamento de Linguagem Natural (PLN) e de outras tecnologias. Foi através da análise exploratória que surgiram correlações bastante fortes, enquanto a modelação de regressão linear tratou das relações que envolvem a experiência e outras variáveis relativas à utilização do PLN. Além disso, foram efetuados testes de normalidade, análise fatorial e ANOVA para melhorar os resultados da análise.

4.2.1 Correlações entre tecnologias

A análise de correlação produziu as seguintes relações:

- *PLN e algoritmos de aprendizagem automática (ML)*: parece haver uma correlação fraca (0,258) que sugere que estas tecnologias são suscetíveis de serem utilizadas em conjunto, presumivelmente devido à sua complementaridade. Os resultados de PLN são melhorados pela aprendizagem automática.
- *PLN e ferramentas de reconhecimento de entidades nomeadas (NER)*: existe uma correlação negativa moderada de (-0,467). Isto significa que, no caso desta amostra, a utilização de PLN é suscetível de estar associada a uma menor utilização de NER. Esta tendência pode representar a escolha de diferentes mecanismos concorrentes para utilização em determinados contextos, provavelmente devido à necessidade de economizar recursos ou a um maior apelo a técnicas mais simples.
- *Algoritmos de aprendizagem automática e extração de texto*: verificou-se uma forte correlação de (-0,775), o que sugere que a maioria dos utilizadores que trabalham com dados não estruturados adota provavelmente uma destas abordagens em detrimento da outra. Esta relação pode ser explicada pela sobreposição de funcionalidades ou pela falta de necessidade.
- *Análise de sentimentos e outras técnicas*: a padronização da análise de sentimentos e do NER/Text Mining oferece uma possibilidade de escolha entre eles, o que reforça ainda mais a observação feita sobre o predomínio da análise de sentimentos, sugerindo uma possível segmentação ou especialização nas estratégias tecnológicas.

4.2.2 Modelagem de Regressão

Foram efetuadas análises de regressão linear simples e múltipla para verificar o efeito da experiência profissional na adoção do uso da PLN.

- *Métodos diretos de análise de regressão*: com base nos anos de experiência como variável explicativa que prevê a aplicação da PLN, foi excluído um valor R-quadrado de 0,049 e um valor p de 0,600, incluindo a experiência como variável relevante aceitável para o nível de utilização da PLN. Provavelmente, outras variáveis como o ambiente de trabalho, o tipo de projeto ou a formação são mais importantes.
- *Métodos indiretos de análise de regressão, incluindo anos de experiência*: verificou-se um pequeno aumento do valor R ao quadrado, que foi registado em 0,150, o que não é apreciável.

Nenhuma das variáveis incluídas foi significativa com valores de $p > 0,05$, o que sugere o impacto de outras variáveis não consideradas no modelo.

4.2.3 Análise de Consistência e outras análises

Foram utilizadas seis análises diferentes para analisar a relação entre a adoção de diferentes estratégias integradas na sua atividade:

- *Consistência interna*: o alfa de Cronbach é de 0,99, pelo que se pode concluir que o questionário apresenta elevados níveis de consistência interna.
- *Matriz de correlação*: a partir da matriz de correlação, verificou-se que a “Definição de objetivos claros” e o “Capacitar a equipa” estão correlacionados com um coeficiente de 0,98; a “Seleção de tecnologia” e o “Capacitar a equipa” têm um coeficiente de 0,93; a “Garantia de escalabilidade” e o “Capacitar a equipa” estão correlacionados com um coeficiente de 0,93. Isto indica a existência de fortes correlações entre estas práticas.
- *Testes de normalidade*: a maioria das variáveis não obedeceu à curva de distribuição normal, o que deve ser considerado na análise paramétrica dos dados.
- *Análise de Regressão*: o valor de R^2 ajustado foi de 0,76 numa regressão simples dos parâmetros “Definição de Objetivos Claros” e “Monitorização e Avaliação de Resultados”, indicando que existe uma forte relação.
- *ANOVA*: não foram observadas diferenças estatisticamente significativas entre as áreas de atividades no que diz respeito à “Monitorização e Avaliação de Resultados”.
- *Análise Fatorial*: três fatores principais emergiram da análise, sendo que o primeiro destes fatores é responsável por uma grande parte da variância e está altamente correlacionado com a maioria das variáveis.

Em resumo, apresentamos a Tabela 1 que esquematiza o detalhe e os resultados das diferentes técnicas estatísticas utilizadas:

Tabela 1 – Resultados da análise quantitativa

Técnica	Detalhes e Resultados
Correlações entre Tecnologias	PLN e ML: 0,258 (Frac) PLN e NER: -0,467 (Moderada Negativa) ML e Extração de Texto: -0,775 (Forte Negativa)
Análise de Regressão (Métodos Diretos)	R ² : 0,049, p: 0,600 (não relevante)
Análise de Regressão (Métodos Indiretos)	R ² : 0,150 (não significativo, p > 0,05)
Consistência Interna	Alfa de Cronbach: 0,99 (elevada consistência)
Matriz de Correlação	Objetivos claros e Capacitar a Equipa: 0,98 Seleção de Tecnologia e Capacitar a Equipa: 0,93 Escalabilidade e Capacitar a Equipa: 0,93
Testes de Normalidade	A maioria das variáveis não distribuídas normalmente
R ² ajustado (Análise de Regressão)	0,76 entre “Definição de Objetivos Claros” e “Monitorização e Avaliação de Resultados”
ANOVA	Sem diferenças estatisticamente significativas
Análise Fatorial	Emergiram 3 fatores principais, com o primeiro fator correlacionado à maioria das variáveis

5. Discussão

Nesta parte do estudo, são feitas reflexões sobre os dados recolhidos e estudos anteriores sobre os temas da transformação digital, da gestão da informação, bem como da integração de dados estruturados e não estruturados.

5.1 Transformação Digital e Gestão da Informação

Os dados analisados a partir do questionário revelam várias perspectivas sobre o tratamento de dados não estruturados, refletindo algumas tendências emergentes na literatura. Os entrevistados destacam a crescente valorização dos dados não estruturados no contexto organizacional, uma opinião que está alinhada com a transformação digital discutida por autores como Vărzaru *et al.* (2022) e Zhang *et al.* (2023). Estes autores enfatizam que a digitalização reforça a partilha de informações e a eficácia organizacional, promovendo melhorias substanciais na gestão da informação. A análise qualitativa mostra que as tecnologias emergentes, como a IA e a aprendizagem automática, são vistas como elementos transformadores, um ponto de vista que ressoa com as mudanças tecnológicas descritas na literatura e que sublinha a necessidade de uma cultura organizacional adaptativa.

Paralelamente, a análise quantitativa fornece insights sobre a integração de tecnologias digitais, revelando desafios na implementação em ambientes organizacionais dinâmicos. As correlações entre tecnologias, como PLN e algoritmos de aprendizagem automática, sugerem uma utilização complementar, mas também destacam a complexidade na escolha e na adoção de novas ferramentas. Isso reflete o argumento de autores como Mokhtar *et al.* (2020) de que a digitalização introduz novas práticas que podem exigir soluções inovadoras e enfrentar barreiras de escalabilidade. A utilização de data lakes e outras tecnologias de armazenamento é mencionada como uma abordagem eficaz para mitigar dificuldades de desempenho e escalabilidade, alinhando-se com o contexto de transformação digital contínua.

Os entrevistados também abordaram a importância da governança de dados, um tema crucial na literatura sobre transformação digital. A necessidade de políticas claras e estruturas bem definidas para garantir a qualidade e conformidade dos dados ecoa as conclusões de Zhang *et al.* (2023) sobre como as organizações devem adaptar suas estruturas para integrar efetivamente as novas tecnologias digitais. A análise de regressão quantitativa, que examinou o papel da experiência profissional na adoção de PLN, sugere que fatores como o ambiente de trabalho ou tipo de projeto são vitais,

reforçando a visão de Mukhtar B *et al.* (2023) sobre a importância da liderança e da inovação cultural na implementação de tecnologias.

Finalmente, a análise de consistência interna dos dados indica que práticas como empowerment e definição de objetivos claros estão fortemente correlacionadas com o sucesso organizacional. Este resultado está em sintonia com a adaptação estrutural mencionada por Zhang, L. & Chen, Z. (2023), que é necessária para acomodar mudanças induzidas pela transformação digital. Tanto os dados qualitativos quanto quantitativos reforçam a percepção de que, embora a digitalização ofereça um potencial significativo para transformar a gestão da informação, ela também traz desafios significativos que precisam ser enfrentados com estratégia e liderança eficazes. A interação entre a inovação tecnológica e as práticas organizacionais existentes é, portanto, crítica para maximizar os benefícios da transformação digital.

5.2 Dados não estruturados como foco estratégico nas organizações

No atual panorama organizacional, a análise de dados não estruturados emergiu como um elemento fundamental para a tomada de decisões estratégicas. Esse tipo de dado, conforme mencionado por Onwujekwe *et al.* (2020), deve ser integrado nos sistemas de gestão do conhecimento para uma compreensão mais abrangente da estratégia relacionada à tomada de decisões. Na pesquisa, observou-se que as organizações que adotam tecnologias como o Processamento de Linguagem Natural (PLN) e a aprendizagem automática (ML) não apenas melhoram a eficácia operacional, mas também revelam informações valiosas para guiá-las em situações problemáticas. Assim, a abordagem de Beheshti *et al.* (2016) sobre a necessidade de refinar dados resulta em ações que transformam dados brutos em inteligência acionável, essencial para a tomada de decisões.

A vida organizacional é impactada por desafios relacionados à gestão de dados não estruturados, conforme destacado por Lutz & Bodendorf (2020), que enfatizam a importância da gestão organizada de dados qualitativos. A pesquisa indica que a falta de políticas claras de tratamento de dados leva a erros e redundâncias, comprometendo a eficiência operacional. As organizações que não estabelecem diretrizes adequadas para a gestão de dados podem enfrentar operações diárias ineficazes e desafios legais, refletindo a crítica de Wei & Liu (2010) sobre a inadequação de sistemas tradicionais em lidar com volumes crescentes de dados não estruturados.

A análise quantitativa dos dados destaca correlações que conectam diferentes tecnologias e suas aplicações nas decisões estratégicas. Por exemplo, houve uma correlação fraca entre PLN e algoritmos de aprendizagem automática, o que sugere um uso conjunto, conforme esperado na literatura. Entretanto, a correlação negativa observada entre PLN e ferramentas de reconhecimento de entidades

nomeadas (NER) sugere que as organizações podem optar por diferentes abordagens com base nas circunstâncias, algo que Kudalkar (2023) reforça ao discutir a interatividade e marketing estratégico.

Finalmente, as análises quantitativas, incluindo a correlação entre "Definição de Objetivos Claros" e "Capacitar a Equipa", evidenciam a necessidade de estratégias bem definidas que promovem a inovação. Os altos níveis de consistência interna do questionário reforçam a validade das conclusões. Assim, os dados não estruturados representam uma área de crescente interesse para as organizações que desejam melhorar a qualidade da tomada de decisões e, conseqüentemente, aumentar sua vantagem competitiva no mercado. A integração eficaz desses dados em sistemas existentes, apoiada por metodologias analíticas avançadas, pode permitir que as organizações transformem desafios em oportunidades.

5.3 Desafios de Governança e Conformidade

As conclusões das respostas ao questionário mostram que muitas das organizações inquiridas ainda não conseguem implementar quadros de governança de dados eficazes. Embora compreendam a importância deste aspeto, os participantes referem que políticas mal concebidas impossibilitam a proteção e a segurança dos dados. Esta situação está em consonância com os pontos de vista da literatura, em que Huang, Z. (2023) propôs que o ciberespaço nas instituições de ensino deve ser protegido de forma equilibrada contra as ameaças digitais e o risco para a privacidade. De ambas as perspectivas - os dados recolhidos e a literatura - a ausência de uma governança adequada coloca as empresas no caminho dos riscos legais e de marca.

As respostas ao questionário também revelam questões relacionadas com a segurança dos dados nos ambientes de computação em nuvem. Observa-se que os participantes têm preocupações em relação à fuga e perda de dados que, entre outras literaturas, são exploradas por Vimal (2021) que considera esta questão como um problema grave, dados os efeitos prejudiciais destas falhas a nível financeiro e de reputação. Esta exigência de avaliação de riscos e a necessidade de valorizar soluções de gestão de segurança que beneficiem as práticas de governança estão em concordância com o que os inquiridos e a literatura académica enfatizaram.

Outro problema que foi frequentemente encontrado nas respostas ao questionário foi a conformidade com a regulamentação. Tal como referido por vários participantes, é um desafio aderir a regulamentos rigorosos e efetuar ajustamentos sempre que necessário. Olaniyi (2023) salienta a importância dos organismos reguladores na implementação de práticas de gestão de dados, uma vez que as regras não reduzem apenas os riscos, mas também protegem a confiança dos clientes. O facto de estes desafios estarem subjacentes aos dados recolhidos realça a necessidade de uma gestão preventiva dos riscos regulamentares relacionados com os dados.

A pesquisa aponta para a maior atenção ao uso de tecnologias, como blockchain e criptografia avançada, para a proteção de dados, o que Wang et al (2019) e Benavides et al (2020) argumentam. Os questionários revelaram que, embora algumas organizações sejam líderes na aplicação dessa tecnologia, grande parte delas ainda está em processo de avaliação. Isto indica uma inadequação nas lacunas de governação que devem ser resolvidas de modo a aumentar os níveis de segurança e eficiência na gestão dos dados.

A digitalização dos processos empresariais coloca complexidades adicionais na governação dos dados não estruturados, o que é evidente nos relatos de gestores que têm dificuldade em lidar com a sobrecarga de dados. De acordo com Spivakovskyy *et al.* (2021), sugere-se que se reforcem constantemente as medidas de proteção da informação, o que segue a estratégia de gestão ativa dos dados. Os participantes e a literatura sublinham conjuntamente esta necessidade de mudança incessante para práticas organizacionais que valorizem a importância da proteção dos dados.

5.4 Eficiência, Integração e Escalabilidade

Os resultados do questionário ilustram que os exercícios de transformação digital têm a capacidade de aumentar a eficiência operacional das organizações. No entanto, muitos dos inquiridos admitem que existem resistências internas na primeira fase de tentativa de implementação, uma constatação que vai ao encontro da literatura de Miklošík e Evans (2020), onde se constata, à partida, que existem complexidades associadas à implementação de tecnologias digitais e que a otimização vem depois. As empresas bem-sucedidas na digitalização relataram uma redução do tempo nos processos e a eliminação de redundâncias, o que prova que, uma vez que as atitudes são positivas, o quadro de gestão da mudança é a única dificuldade associada que é gerida de forma eficaz para revelar resultados.

A fusão de todos estes tipos de dados não estruturados recolhidos de diferentes fontes foi considerada um dos atributos preocupantes pelos inquiridos no questionário. Alguns dos inquiridos abordam a questão da integração de dados provenientes de diferentes sistemas, o que reflete algumas das preocupações de Anthony *et al.* (2021). A ausência de um formato comum acordado para os dados ou de uma estrutura empresarial universal impede uma integração harmoniosa e, por conseguinte, os dados e a análise não podem ser consolidados para ajudar a tomada de decisões baseadas em factos. Não obstante, a literatura identifica também a utilização de sistemas de informação avançados e de aplicações como a aprendizagem automática e a análise de grandes volumes de dados, que têm a capacidade de melhorar esta integração, tal como referido por Chan, J. (2020).

Como se pode ver nos dados do questionário, a maioria das organizações tem como objetivo expandir continuamente e ajustar rapidamente as soluções de dados para satisfazer as exigências do mercado em constante mudança. Este facto faz sentido à luz de Zou (2024), que afirma que, à medida que as tecnologias digitais das empresas evoluem, há mais requisitos para desenvolver modelos operacionais escaláveis e sustentáveis. A investigação mostra que as tecnologias de computação em nuvem são frequentemente utilizadas para este fim, permitindo às empresas escalar os seus cativos de forma eficiente sem uma forte dependência de cativos fixos.

As dificuldades de eficiência, integração e escalabilidade são também desafios que dependem da liderança organizacional para promover uma cultura centrada nos dados. Os inquiridos sublinharam a importância da liderança adaptativa na gestão das complexidades da transformação digital, o que está de acordo com o trabalho de Yuanming (2023) e Lu & Zhou (2022) que estudaram este fenómeno no contexto chinês. A gestão de liderança é capaz de mudar a cultura organizacional na medida em que os dados passam a ser valorizados como um recurso estratégico e todos os componentes da organização são orientados para a utilização otimizada dos dados disponíveis.

Por último, os resultados do questionário destacam o papel da formação constante e do desenvolvimento de talentos para facilitar a transformação digital nas organizações. Nesse sentido, trazer novas pessoas com competências capazes de preencher as lacunas na análise de dados e nas novas tecnologias é considerada uma prática importante defendida por Gilch & Sieweke (2020). Esta abordagem não só melhora as competências analíticas da organização, como também garante que a empresa está bem equipada para lidar com a dinâmica do mercado e as mudanças tecnológicas.

5.5 Análise avançada de dados não estruturados

Com base nos dados do questionário, verificou-se que a maioria das organizações reconhece a existência de dados não estruturados, mas não consegue utilizá-los para a maioria dos fins. As várias formas de dados, que incluem texto, imagens, vídeos, etc., são difíceis de combinar e utilizar com sistemas de análise estruturados, uma vez que a sua composição é complexa. No entanto, a literatura, incluindo El-Taliawi *et al.* (2021), afirma que métodos avançados como a Aprendizagem Automática e o Processamento de Linguagem Natural (PLN) parecem atenuar estes desafios e permitir que as organizações recolham informações úteis a partir de cativos ricos em informação, mas terrivelmente complexos.

A partir dos resultados do questionário, pode deduzir-se que a análise de dados não estruturados é bastante popular em tarefas como o feedback dos clientes e as redes sociais. Os participantes registam a incorporação bem-sucedida da análise de sentimentos e da extração de texto na prática

como ferramentas eficazes para interpretar as preferências e percepções dos clientes. Esta utilização é também bem apoiada pela literatura académica disponível, que salienta a capacidade derivada destas técnicas para informar a tomada de decisões relativas ao marketing e ao desenvolvimento de produtos. (Vasilopoulou *et al.*, 2023).

Os resultados do questionário destacam que, especificamente, as PME têm desafios na utilização da análise de dados não estruturados devido a recursos limitados. No entanto, o questionário sugere que as PME com esse potencial podem ser capazes de competir com empresas maiores depois de adotarem estas tecnologias e melhorarem a sua eficácia e competitividade no mercado. Parra-Sánchez (2023) afirma que a transformação digital pode ser um nivelador para estas empresas, uma vez que podem alavancar tecnologias atualizadas nas suas estratégias.

Os participantes no questionário observaram que, com uma implementação adequada, a análise avançada de dados não estruturados tem a capacidade de melhorar a cadeia de abastecimento e o desempenho operacional em vários sectores, incluindo a indústria transformadora e os cuidados de saúde. Esta opinião é partilhada em estudos como o de Song *et al.* (2022), que afirmam que a presença combinada com o espaço para analisar esses dados, como os dispositivos IoT, permite uma melhor tomada de decisões e a utilização de recursos.

Uma questão diferente levantada nos resultados é a forma como os dados não estruturados também podem ser utilizados para melhorar a experiência do cliente. As empresas que tiram partido destas informações também tendem a conseguir uma melhor personalização da interação e clientes mais fiéis. Além disso, Harrsain *et al.* (2023) argumentam que a rotatividade de clientes com base na análise de dados ajuda a antecipar o comportamento dos clientes e, por conseguinte, a aprofundar o marketing.

6. Conclusões, propostas, limitações e outros focos de investigação

Este estudo investigou como assegurar a gestão de dados não estruturados tendo em vista melhorar o desempenho da organização e que práticas podem ser utilizadas. Apresentamos neste capítulo as conclusões, as limitações e as recomendações para investigação futura.

6.1 Conclusões

Resulta desta investigação que a governação da informação não estruturada, tema central desta tese, é frequentemente dificultada pela falta de regulamentação adequada, bem como pela incorporação desses dados no modelo de gestão de dados existente. As barreiras identificadas, a procura de novos recursos e a procura de regulamentação, são essencialmente a ilustração do problema que está a ser investigado: como é que estas organizações adaptam a gestão destes dados não estruturados às circunstâncias da digitalização. A compreensão dos investigadores do papel da governação de dados para as organizações, juntamente com a integração de dados provenientes de muitas fontes, cria uma forte necessidade de avançar com novos desenvolvimentos que acompanhem o mercado contemporâneo, ligando assim problemas reais de gestão de dados aos problemas teóricos abordados anteriormente.

O processo de gestão de dados não estruturados apresenta muitos desafios. Embora o Processamento de Linguagem Natural (PLN) seja uma tecnologia útil, está interligado com outras tecnologias e não pode ser efetuado isoladamente de qualquer prática profissional. A ausência de uma relação positiva entre a aprendizagem automática e a extração de texto indica que os utilizadores tendem a adotar um método de âmbito bastante vasto e específico para as funções da organização.

Os resultados também mostram que fatores organizacionais como a cultura organizacional, os recursos e a formação têm um impacto mais forte do que a experiência anterior quando se trata da adoção do ML. A correlação entre “Definir Objetivos Claros”, “Capacitar a Equipa” e “Assegurar a Escalabilidade” confirma a importância de adotar uma boa estratégia para garantir um sucesso eficiente nas tecnologias de análise de dados.

Para completar a transformação digital contemporânea, é essencial que as organizações adotem novas tecnologias, bem como as suas práticas organizacionais. A receita para o sucesso deve incluir tecnologia, estruturas organizacionais ágeis e uma liderança convincente que dê às empresas uma vantagem competitiva através da utilização de informações digitais. É necessário efetuar mais

investigação sobre a forma como as diferentes indústrias abordam questões semelhantes de transformação digital neste contexto.

Para garantir que os dados não são desperdiçados, as organizações precisam de adotar tecnologias emergentes e estabelecer mecanismos eficazes de governação e escalabilidade. No atual panorama organizacional, os líderes são cruciais para promover uma cultura de criatividade e agilidade. As melhores práticas discutidas podem ser implementadas na prática com a ajuda de tabelas e gráficos.

Os dados digitais estão sujeitos a governação e conformidade, pelo que os dois mecanismos são básicos na gestão dos dados digitais. É necessário um investimento em tecnologia avançada e uma força de trabalho qualificada para eliminar os riscos de segurança e cultivar um ambiente que cumpra todos os princípios de governação de dados.

É certo que existem desafios associados à eficiência, integração e escalabilidade, mas estes podem ser resolvidos com a ajuda de tecnologia adequada, liderança e desenvolvimento de recursos humanos. As empresas que implementam a digitalização nos processos e se concentram na melhoria do pessoal terão mais hipóteses num ambiente em constante mudança.

Em última instância, embora se saiba que os dados não estruturados representam algumas ameaças, oferecem grandes potencialidades para as empresas que estão dispostas a incorporar técnicas analíticas avançadas nos seus processos empresariais. É possível transformar esses dados num ativo com a aplicação eficaz da aprendizagem automática e de outros algoritmos como a PLN, melhorando assim a eficiência das empresas e promovendo o desenvolvimento estratégico das mesmas. Para tal, será necessário um investimento sustentado no desenvolvimento de aptidões e na criatividade de competências, a fim de adaptar as organizações às exigências do mercado em mutação para uma competitividade a longo prazo.

Como resultado deste estudo, propomos um plano de projeto, alto nível, para a implementação de um modelo de gestão de dados não estruturados assim como uma lista de seis ferramentas tecnológicas que podem ser consideradas no projeto.

6.2 Proposta de projeto de definição de modelo de gestão de dados não estruturados

Face ao que foi identificado nos dados recolhidos, tendo em consideração a discussão dos mesmos com a revisão de literatura, consideramos que o desenvolvimento de um projeto para um modelo de gestão de dados não estruturados poderia passar por:

1. *Avaliação Inicial/Diagnóstico*
 - a. Identificação dos Dados Não Estruturados

- i. Tipos de dados presentes (textos, imagens, vídeos).
 - b. Análise do Contexto Organizacional
 - i. Avaliar a cultura organizacional; disposição para a mudança.
- 2. *Definição de Objetivos*
 - a. Objetivos Estratégicos
 - i. Melhorar a tomada de decisões.
 - ii. Aumentar a eficiência operacional.
 - iii. Promover inovação e competitividade.
 - iv. Outros. Quais?
 - b. Alinhamento com Processos de Negócio
 - i. Integrar dados não estruturados nas estratégias existentes.
- 3. *Desenvolvimento de Políticas de Data Governance*
 - a. Estabelecimento de normas
 - i. Criar/rever políticas claras para a gestão dos dados.
 - b. Práticas de Segurança
 - i. Implementar medidas de segurança para proteção de dados.
 - ii. Assegurar conformidade com normas regulamentares aplicáveis.
- 4. *Seleção de Tecnologias e Ferramentas*
 - a. Definição de Tecnologias a Utilizar
 - i. Processamento de Linguagem Natural (PLN), Aprendizagem Automática (ML).
 - ii. Outras. Quais?
 - b. Integração de Sistemas
 - i. Garantir a interoperabilidade das ferramentas selecionadas
- 5. *Formação e desenvolvimento de equipas*
 - a. Desenvolvimento de Talentos
 - i. Promover formação sobre tecnologias adotadas e prática de gestão de dados.
 - b. Promover uma Cultura de Dados
 - i. Incentivar a utilização de dados na tomada de decisões em todos os níveis da organização.
- 6. *Execução e Monitorização*
 - a. Fase de Execução
 - i. Iniciar projeto-piloto na abordagem adaptativa (ex. AGILE) dada a incerteza relacionada com o produto final e os seus requisitos, e à relevância de obter entregáveis intermédios, antes da implementação total.
 - b. Monitorização e Avaliação

- i. Avaliar desempenho e eficácia das soluções implementadas, avaliando qual foi a evolução face à situação inicial.

7. *Melhoria Contínua*

- a. Feedback e Iteração
 - i. Recolher feedback dos utilizadores e ajustar processos conforme necessário, documentando lições aprendidas para apoio aos projetos seguintes, numa fase mais madura.
- b. Inovação Contínua
 - i. Identificar novas tecnologias ou novas metodologias que tenham entretanto atingido maior maturidade e possam acrescentar valor para a gestão dos recursos da organização.

Cada um destes pontos deve ser enquadrado num plano de projeto, considerando que:

1. A *Avaliação Inicial/Diagnóstico* é o ponto de partida, pois fornece a base para a Definição de Objetivos.
2. Os *Objetivos Estratégicos* e *Governança de Dados* devem ser conduzidos em paralelo porque permite criar a base para alcançar os objetivos definidos.
3. A *Seleção de Tecnologias e Ferramentas* devem respeitar o que foi definido nos objetivos estratégicos e regras de governança de dados.
4. A *Formação e Desenvolvimento* deve acompanhar o processo de implementação para que as equipas impactadas com os novos desenvolvimentos saibam como utilizar as ferramentas e executar as políticas.
5. A *Melhoria Contínua* deve acompanhar o pós-entrada em produção para garantir que o modelo de gestão de dados evolua junto com as necessidades da empresa e as mudanças no ambiente tecnológico.

6.3 Soluções tecnológicas para a gestão de dados não estruturados

Conforme previsto no ponto 4 da *Proposta de projeto de definição de modelo de gestão de dados não estruturados*, algumas das soluções tecnológicas que podem ser utilizadas, são:

- *Implementar um Repositório Centralizado de Dados (Data Lake)*: consolidar dados não estruturados de várias fontes num local central. Isto permite o acesso, análise e gestão mais fácil em comparação com bancos de dados dispersos e distintos.

- *Utilizar a Gestão de Metadados*: implementar um sistema para etiquetar, classificar e descrever dados não estruturados usando metadados. Este passo crucial torna os dados localizáveis, compreensíveis e utilizáveis para análise.
- *Utilizar Avaliação e Limpeza da Qualidade dos Dados*: estabelecer processos para identificar, corrigir e remover erros e inconsistências nos dados não estruturados. Isto garante a qualidade dos dados que alimentam a análise e a tomada de decisões a jusante.
- *Integrar Dados Estruturados e Não Estruturados*: projetar sistemas para combinar dados estruturados e não estruturados para obter uma visão mais holística da informação empresarial. Vincular dados não estruturados (ex.: feedback do cliente) com dados estruturados (ex.: dados de vendas) pode revelar correlações valiosas.
- *Utilizar Técnicas de Processamento de Linguagem Natural (PLN)*: utilizar PLN para extrair informações significativas de dados textuais (ex.: e-mails, documentos, redes sociais). Isto permite a análise automatizada de sentimentos, identificação de tópicos e outros insights valiosos.
- *Aplicar Algoritmos de Aprendizagem Automática (ML)*: utilizar algoritmos de ML (ex.: para classificação, previsão ou agrupamento) para analisar e interpretar a informação extraída de dados não estruturados através de PLN ou outros métodos. Isto pode automatizar a análise, identificar padrões e melhorar a tomada de decisões.

6.4 Limitações

Para a interpretação dos resultados deste estudo, há que ter em conta algumas limitações. Para começar, a dimensão da amostra, embora razoável, reduz possivelmente a validade da aplicação dos resultados a populações mais alargadas. Uma vez que o estudo utilizou um desenho transversal, só foi capaz de reportar associações e não relações causais entre as variáveis. O facto de os dados terem sido recolhidos através de um questionário em linha como única fonte de recolha de dados pode criar um enviesamento na amostra do estudo, excluindo os profissionais que têm uma base tecnológica baixa ou aqueles que não estão familiarizados com esta abordagem de recolha de dados. A utilização de um instrumento em vez de meras enumerações para a recolha de informações pode ter afetado os resultados. É pouco provável que a avaliação auto-relatada da utilização das diferentes tecnologias seja um indicador do grau real de utilização dessas tecnologias na prática.

6.5 Contributo para futuros focos de investigação

Ao analisar os dados dos questionários em relação à literatura anterior, surgem algumas áreas de semelhanças e diferenças sobre o tema da gestão de dados não estruturados e a forma como esta afeta o desempenho organizacional. A primeira área de convergência está relacionada com a aceitação de que tecnologias avançadas como a Inteligência Artificial (IA) e a Aprendizagem Automática (AM) podem revolucionar a utilização de dados não estruturados. Tanto a revisão da literatura como os dados recolhidos sugerem que estas tecnologias ajudam a automatizar o processo, a gerir a inteligência dos dados e a melhorar a tomada de decisões globais nas organizações. Esta correlação torna imperativo que as organizações procurem novas abordagens para lidar com o problema que surge devido ao aumento do volume e da variabilidade dos dados não estruturados.

No entanto, uma diferença digna de nota é a prática efetiva da estrutura de governação de dados. Embora os autores sublinhem a necessidade de uma forte governação dos dados para facilitar a conformidade, aumentar a segurança e gerir eficazmente os dados não estruturados, os resultados dos questionários mostram que muitas organizações continuam a ter problemas em estabelecer políticas de governação eficazes. Esta lacuna mostra que, apesar de ambos deduzirem o mesmo resultado, existe um acordo sobre a necessidade de uma boa governação dos dados a nível teórico, a aplicação efetiva na prática não está em conformidade com tal, pelo que é necessário analisar as dificuldades que as organizações enfrentam durante a aplicação dos quadros.

Para aprofundar a pesquisa sobre a governação de dados, é crucial identificar barreiras que dificultam a implementação dessas políticas nas organizações, explorando casos de empresas que superaram esses desafios para encontrar soluções práticas. É igualmente importante entender a relação entre as práticas de gestão de dados e as novas tecnologias, pois isso permitirá que as organizações aproveitem plenamente a transformação digital e se tornem mais flexíveis num ambiente em constante mudança.

Tendo em conta as limitações do presente estudo, considero que podem ser ponderadas as seguintes ações, em estudos posteriores, com vista à redução das limitações verificadas e à ampliação do âmbito do estudo:

1. *Aumentar o tamanho da amostra*: no futuro, devem ser incluídos participantes de uma maior heterogeneidade de sectores e de dimensões organizacionais, para que se possam obter resultados de investigação mais sólidos e generalizáveis.
2. *Adotar uma conceção longitudinal*: este tipo de investigação ajudaria a identificar quem adaptou as tecnologias ao longo do tempo, permitindo assim examinar a mudança na utilização e, por conseguinte, a causa, caso exista.

3. *Adotar técnicas mistas de recolha de dados:* a utilização de questionários, de perspetivas observacionais e de entrevistas qualitativas pode ajudar a compreender em pormenor os fatores que influenciam a adoção de tecnologias de tratamento de dados não estruturados de forma abrangente.

Creio que este trabalho subsequente poderá ajudar a refinar os resultados que permitam perceber a realidade por detrás da utilização dos dados não estruturados e o seu impacto nas organizações.

7. Referências bibliográficas

- Aagaard, A., Presser, M., Collins, T., Beliatas, M., Skou, A., & Jakobsen, E. (2021). The role of digital maturity assessment in technology interventions with industrial internet playground. *Electronics*, 10(10), 1134. <https://doi.org/10.3390/electronics10101134>
- Adnan, K., and Akbar, R. (2019). Limitations of information extraction methods and techniques for heterogeneous unstructured big data. *International Journal of Engineering Business Management*, 11. <https://doi.org/10.1177/1847979019890771>
- Alabdali, N. S., and Salam, A. A. (2022). The impact of digital transformation on supply chain procurement for creating competitive advantage: An empirical study. *Sustainability*, 14(19), 12269. <https://doi.org/10.3390/su141912269>
- Anthony, B., Petersen, S., Helfert, M., & Guo, H. (2021). Digital transformation with enterprise architecture for smarter cities: A qualitative research approach. *Digital Policy Regulation and Governance*, 23(4), 355-376. <https://doi.org/10.1108/dprg-04-2020-0044>
- Beheshti, S., Tabebordbar, A., Benatallah, B., and Nouri, R. (2016). Data curation APIs. <https://doi.org/10.48550/arxiv.1612.03277>
- Benavides, L. M. C., Arias, J. A. T., Serna, M. D. A., Bedoya, J. W. B., and Burgos, D. (2020). Digital transformation in higher education institutions: A systematic literature review. *Sensors*, 20(11), 3291. <https://doi.org/10.3390/s20113291>
- Bonomi, A., *et al.* (2019). The impact of digital transformation on formal and informal organizational structures of large architecture and engineering firms. *Engineering Construction and Architectural Management*, 27(4), 872-892. <https://doi.org/10.1108/ecam-03-2019-0119>
- Chan, J. (2020). Digital transformation in the era of big data and cloud computing. *International Journal of Intelligent Information Systems*, 9(3), 16. <https://doi.org/10.11648/j.ijis.20200903.11>
- Derrida, J. (1967). *Writing and difference* (A. Bass, Trans.). University of Chicago Press.
- El-Taliawi, O., Goyal, N., & Howlett, M. (2021). Holding out the promise of lasswell's dream: Big data analytics in public policy research and teaching. *Review of Policy Research*, 38(6), 640-660. <https://doi.org/10.1111/ropr.12448>
- Feroz, E. H., Zo, H., and Lee, H. (2021). Digital transformation and environmental sustainability: A review and research agenda. *Sustainability*, 13(3), 1530. <https://doi.org/10.3390/su13031530>
- Gaffley, G. and Pelser, T. (2021). Developing a digital transformation model to enhance the strategy development process for leadership in the south african manufacturing sector. *South African Journal of Business Management*, 52(1). <https://doi.org/10.4102/sajbm.v52i1.2357>
- Gilch, H., and Sieweke, J. (2020). Recruiting digital talent: The strategic role of recruitment in organizations' digital transformation. *German Journal of Human Resource Management*, 35(1), 53-82. <https://doi.org/10.1177/2397002220952734>
- Halim, H. (2023). Digital transformation strategy to optimize company performance. *Jurnal Manajemen Bisnis Akuntansi Dan Keuangan*, 2(2), 189-200. <https://doi.org/10.55927/jambak.v2i2.7022>
- Huang, Z. (2023). Modeling and constructing a network security system for colleges and universities in the context of digital transformation. *Advances in Education Humanities and Social Science Research*, 7(1), 392. <https://doi.org/10.56028/aehtsr.7.1.392.2023>
- Hussain, H., Alabdullah, T., Ahmed, E., and Jamal, K. (2023). Implementing technology for competitive advantage in digital marketing. *International Journal of Scientific and Management Research*, 06(06), 95-114. <https://doi.org/10.37502/ijsmr.2023.6607>
- Jiang, X. (2023). Research on the impact and countermeasures of the digital economy era on the development of real enterprises. *Frontiers in Business Economics and Management*, 9(1), 43-46. <https://doi.org/10.54097/fbem.v9i1.8286>
- Król, K. and Zdonek, D. (2020). Analytics maturity models: An overview. *Information*, 11(3), 142. <https://doi.org/10.3390/info11030142>

- Kudalkar, D. (2023). Social media sentiment analysis. *International Journal for Research in Applied Science and Engineering Technology*, 11(10), 1696-1708. <https://doi.org/10.22214/ijraset.2023.56290>
- Liu, X., Lang, B., Yu, W., Luo, J., & Huang, L. (2011). AUDR: An advanced unstructured data repository. In 2011 6th International Conference on Pervasive Computing and Applications (pp. 462-469). Port Elizabeth. <https://doi.org/10.1109/ICPCA.2011.6106548>
- Lu, X. and Zhou, Z. (2022). Digital transformation, business risk and firm value. *Frontiers in Sustainable Development*, 2(5), 25-32. <https://doi.org/10.54691/fsd.v2i5.695>
- Lutz, C., and Bodendorf, F. (2020). Analyzing industry stakeholders using open-source competitive intelligence – A case study in the automotive supply industry. *Journal of Enterprise Information Management*, 33(3), 579-599. <https://doi.org/10.1108/jeim-08-2019-0234>
- Ma, X., et al. (2022). The effect of digital transformation on the pharmaceutical sustainable supply chain performance: The mediating role of information sharing and traceability using structural equation modeling. *Sustainability*, 15(1), 649. <https://doi.org/10.3390/su15010649>
- Miklošik, A. and Evans, N. (2020). Impact of big data and machine learning on digital transformation in marketing: A literature review. *IEEE Access*, 8, 101284-101292. <https://doi.org/10.1109/access.2020.2998754>
- Mokhtar, S., Hussin, N., Tokiran, N., Wahab, H., and Ibrahim, A. (2020). Digital transformation in information management. *International Journal of Academic Research in Business and Social Sciences*, 10(11). <https://doi.org/10.6007/ijarbss/v10-i11/9071>
- Mendonça, F. and Dantas, M. (2020). Covid-19: where is the digital transformation, big data, artificial intelligence and data analytics? *Revista Do Serviço Público*, 71, 212-234. <https://doi.org/10.21874/rsp.v71i0.4770>
- Motjolopane, I. (2023). Digital transformation dimensions for evaluating smes' readiness for big data analytics and artificial intelligence: A review. *International Journal of Research in Business and Social Science* (2147-4478), 12(7), 583-595. <https://doi.org/10.20525/ijrbs.v12i7.2837>
- Mu, B. (2023). Does enterprise digital transformation affect quality of accounting information? *BCP Business and Management*, 50, 198-210. <https://doi.org/10.54691/bcpbm.v50i.5609>
- Muktamar B, A., Bachtar, A., Guntoro, G., Riyantie, M., and Ridwan, N. (2023). The role of leadership in digital transformation management in organizations. *Jurnal Minfo Polgan*, 12(1), 1306-1314. <https://doi.org/10.33395/jmp.v12i1.12731>
- Olaniyi, O. (2023). Whatsapp data policy, data security and users' vulnerability. *International Journal of Innovative Research and Development*. <https://doi.org/10.24940/ijird/2023/v12/i4/apr23021>
- Onwujekwe, G., Osei-Bryson, K. M., & Ngwum, N. (2020). A framework for capturing and analyzing unstructured and semi-structured data for a knowledge management system. *arXiv preprint arXiv:2007.07102*. <https://doi.org/10.48550/arXiv.2007.07102>
- Parra-Sánchez, I. (2023). Digital transformation in small and medium enterprises: A scientometric analysis. *Digital Transformation and Society*, 3(3), 257-276. <https://doi.org/10.1108/dts-06-2023-0048>
- Pan, G. and Lee, B. (2020). Leveraging digital technology to transform accounting function: Case study of a sme. *International Journal of Accounting and Financial Reporting*, 10(2), 24. <https://doi.org/10.5296/ijafr.v10i2.17052>
- Qian, J. (2023). The significance of financial accounting transformation in the context of big data. *Frontiers in Sustainable Development*, 3(4), 60-66. <https://doi.org/10.54691/fsd.v3i4.4761>
- Rave, J., Álvarez, G., and Morales, J. (2021). Multi-criteria decision-making leveraged by text analytics and interviews with strategists. *Journal of Marketing Analytics*, 10(1), 30-49. <https://doi.org/10.1057/s41270-021-00125-8>
- Sánchez, D., Sarmiento, L., Cuadros, J., & Guerrero, C. (2022). Chief information officer's role for iot-based digital transformation in colombian smes. *Revista Colombiana De Computación*, 23(2), 43-54. <https://doi.org/10.29375/25392115.4607>
- Santiso, C. (2022). Govtech against corruption: what are the integrity dividends of government digitalization? *Data & Policy*, 4. <https://doi.org/10.1017/dap.2022.31>

- Sapach, A., & Perkins, R. (2021). Ethical data governance: The competitive advantage. *International Journal of Information Management*, 57, 102-110. <https://doi.org/10.1016/j.ijinfomgt.2020.102110>
- Song, J., Wang, A., Liu, P., Li, D., Han, X., and Yan, Y. (2022). A hybrid model for value-added process analysis of manufacturing value chains. *IET Collaborative Intelligent Manufacturing*, 5(1). <https://doi.org/10.1049/cim2.12071>
- Spivakovskyy, S., Kochubei, O., Shebanina, O. B., Sokhatska, O., Yaroshenko, I., and Nych, T. (2021). The impact of digital transformation on the economic security of Ukraine. *Studies of Applied Economics*, 39(5). <https://doi.org/10.25115/eea.v39i5.5040>
- Tripathi, S. (2023). Digital transformation and cloud computing: System dynamics modeling approach. *International Journal for Multidisciplinary Research*, 5(6). <https://doi.org/10.36948/ijfmr.2023.v05i06.11417>
- Tu, Z. (2023). Research on vulnerabilities in computer information technology and encryption techniques. *Frontiers in Computing and Intelligent Systems*, 6(1), 42-45. <https://doi.org/10.54097/fcis.v6i1.08>
- Ullah, A., Azeem, M., Ashraf, H., Alaboudi, A., Humayun, M., and Jhanjhi, N. (2021). Secure healthcare data aggregation and transmission in iot—A survey. *IEEE Access*, 9, 16849-16865. <https://doi.org/10.1109/access.2021.3052850>
- Vărzaru, A., Bocean, C., Mangra, M., & Simion, D. (2022). Assessing users' behavior on the adoption of digital technologies in management and accounting information systems. *Electronics*, 11(21), 3613. <https://doi.org/10.3390/electronics11213613>
- Vasilopoulou, C., Theodorakopoulos, L., and Giotopoulos, K. (2023). Big data analytics: A catalyst for digital transformation in e-government. *Technium Social Sciences Journal*, 45, 449-459. <https://doi.org/10.47577/tssj.v45i1.9134>
- Vimal, B. (2021). Data security in cloud computing. *Mathematical Statistician and Engineering Applications*, 70(2), 1716-1724. <https://doi.org/10.17762/msea.v70i2.2462>
- Wang, J. (2023). Optimization of quantitative investment strategies in the financial big data environment. *Frontiers in Business Economics and Management*, 12(2), 52-54. <https://doi.org/10.54097/fbem.v12i2.14595>
- Wang, M., Chen, C., Chen, K., Lo, H. (2023). Leadership competencies in the financial industry during digital transformation: An evaluation framework using the Z-DEMATEL technique. *Axioms*, 12(9), 855. <https://doi.org/10.3390/axioms12090855>
- Wang, H., Lin, H., Gao, X., Cheng, W., and Chen, Y. (2019). Reversible ambtc-based data hiding with security improvement by chaotic encryption. *IEEE Access*, 7, 38337-38347. <https://doi.org/10.1109/access.2019.2906500>
- Warner, K. S. R., and Wäger, M. (2019). Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal. *Long Range Planning*, 52(3), 326-349. <https://doi.org/10.1016/j.lrp.2018.12.001>
- Wei, L., and Liu, B. (2010). A tetrahedral data model for unstructured data management. *Science China Information Sciences*, 53(8), 1497-1510. <https://doi.org/10.1007/s11432-010-4030-9>
- Xie, M. (2024). Exploring improvement of business performance of transaction processing systems in the retail sector. *Advances in Economics Management and Political Sciences*, 87(1), 77-83. <https://doi.org/10.54254/2754-1169/87/20240972>
- Xu, K., Li, Y., Liu, C., Xu, L., Hao, X., Gao, J., ... and Maropoulos, P. (2020). Advanced data collection and analysis in data-driven manufacturing process. *Chinese Journal of Mechanical Engineering*, 33(1). <https://doi.org/10.1186/s10033-020-00459-x>
- Yang, G., Yang, M., Salam, S., & Zeng, J. (2019). Research on protecting information security based on the method of hierarchical classification in the era of big data. *Computers, Materials & Continua*, 61(3), 19-26. <https://doi.org/10.32604/jcs.2019.05947>
- Yuanming, X. (2023). Traditional publishing enterprises shift to digitalization: Basis for strategic plan. *The Quest Journal of Multidisciplinary Research and Development*, 2(3). <https://doi.org/10.60008/thequest.v2i3.147>

- Zeng, Y. and Wang, L. (2023). Regional fintech development and enterprise digital transformation: Theoretical and mechanism analysis. *BCP Business & Management*, 47, 168-184. <https://doi.org/10.54691/bcpbm.v47i.5188>
- Zhang, J., and Chen, Z. (2023). Exploring human resource management digital transformation in the digital age. *Journal of the Knowledge Economy*, 15(1), 1482-1498. <https://doi.org/10.1007/s13132-023-01214-y>
- Zhang, X., Xu, Y., & Ma, L. (2023). Information technology investment and digital transformation: the roles of digital transformation strategy and top management. *Business Process Management Journal*, 29(2), 528-549. <https://doi.org/10.1108/bpmj-06-2022-0254>
- Zhu, Y. (2023). The digital transformation of enterprise accounting: Big data, IA and financial sharing. *Advances in Economics Management and Political Sciences*, 31(1), 187-192. <https://doi.org/10.54254/2754-1169/31/20231536>
- Zou, L. (2024). Measuring corporate digital transformation: Methodology, indicators and applications. *Sustainability*, 16(10), 4087. <https://doi.org/10.3390/su16104087>

Anexo A - Questionário

O questionário é composto por 9 partes:

A. *Parte 1: Área de atividade*

Pergunta: Em que área desenvolve a sua atividade?

Objetivo: Classificar os inquiridos em categorias específicas do sector, tais como (TI, Gestão Financeira, Legalização e Digitalização de Documentos, Outros, e Certificação e Formação) e permitir que as respostas fossem comparadas com algumas respostas específicas do sector.

B. *Parte 2: Experiência*

Pergunta: Quantos anos de experiência possui?

Objetivo: Situar os inquiridos numa escala de experiência (≥ 20 anos, < 5 anos, < 20 e < 10 , respetivamente) e perceber se o tempo de experiência profissional tem impacto na resposta fornecida.

C. *Parte 3: Identificar tipologia de resposta a situações prováveis*

Questão 1: A sua empresa está a tentar consolidar dados não estruturados de feedback dos clientes provenientes de várias fontes para melhorar a oferta de serviços. Qual é a melhor abordagem para facilitar esta integração?

Objetivo: Avaliar as perspetivas dos inquiridos relativamente à integração de dados não estruturados, tais como investigação primária ou estudos periféricos, nos dados existentes (repositório centralizado, abordagem ETL, plataforma de análise de texto, outros).

Questão 2: A sua entidade pretende utilizar dados de mercado não estruturados de várias fontes para tomar decisões estratégicas. Que abordagem analítica recomendaria para obter informações financeiras?

Objetivo: Avaliar a aplicação, pelos inquiridos, de métodos analíticos para dados de mercado não estruturados (modelação de tópicos, reconhecimento de entidades, análise de sentimentos, outros).

Questão 3: Com o aumento dos dados não estruturados que contêm informações financeiras sensíveis, como daria prioridade às medidas de segurança?

Objetivo: Avaliar o nível de melhores práticas que os inquiridos demonstraram em relação à segurança dos dados (encriptação, monitorização, controlos de acesso, outros).

Questão 4: A que práticas-chave daria prioridade no âmbito da estrutura de governação de dados para garantir a qualidade dos dados e a conformidade regulamentar dos dados não estruturados?

Objetivo: Avaliar o nível de reconhecimento pelos inquiridos dos principais aspetos da governação de dados (gestão de metadados, gestão do ciclo de vida dos dados, análise da qualidade dos dados, outros).

Questão 5: Tendo em conta o panorama dinâmico do sector financeiro, que tecnologia emergente prevê como mais benéfica para melhorar as capacidades de gestão de dados não estruturados e o serviço ao cliente?

Objetivo: Avaliar o nível de reconhecimento pelos inquiridos dos principais aspetos da governação de dados (gestão de metadados, gestão do ciclo de vida dos dados, análise da qualidade dos dados, outros).

D. Parte 4: Evolução futura da gestão de dados

Questão: Como se posiciona perante estas técnicas de tratamento de dados?

Objetivo: Conhecer a utilização por parte das organizações dos inquiridos de novas tecnologias relacionadas com o tratamento de dados:

- *Processamento de linguagem natural (NLP)*
- *Algoritmos de machine learning*
- *Análise de Sentimento*
- *Ferramentas de NER (Named Entity Recognition)*
- *Mineração de Texto*
- *Processamento de Imagens*
- *Análise de Voz*
- *Aprendizagem de Máquina Não Supervisionado*
- *Indexação e Recuperação de Informação*
- *Integração de Dados*

E. Parte 5: Capacidade de tratamento de dados não estruturados

Questão: Como avaliaria a capacidade de uma empresa tratar dados estruturados sem ...

Objetivo: Classifique numa escala de 1 a 10, em que 1 (um) refere-se a uma capacidade nula e 10 (dez) refere-se a uma capacidade plena para o sucesso do processo de tratamento de dados não estruturados, por não aplicar as seguintes medidas

- *definir objetivos claros*
- *capacitar a sua equipe*
- *escolher a tecnologia certa*
- *garantir escalabilidade da solução*
- *utilizar metadados*
- *integração dos dados estruturados*

- *limpeza e pré-processamento de dados*
- *implementar segurança de dados*
- *monitorizar e avaliar resultados*
- *um plano de melhoria contínua*

F. Perguntas abertas

Perguntas:

- Como é que vê as tecnologias emergentes (ex.: IA, Aprendizagem Automática) a influenciar a gestão de dados não estruturados no futuro?
- Como é que lida com os desafios de escalabilidade e desempenho associados ao processamento de grandes volumes de dados não estruturados?
- Que papel acredita que a governação de dados desempenha na gestão eficaz de dados não estruturados nas organizações?
- Qual o impacto no seu negócio ou organização de não existirem políticas de tratamento dados eficazes?

Objetivo: abrir espaço para que os inquiridos possam transmitir a sua opinião de forma aberta e conseguir retirar das suas respostas padrões que possam depois ser avaliados com o restante questionário.