

UNIVERSIDADE FEDERAL DE SANTA CATARINA

**PROPOSTA DE GUIA PARA INICIATIVAS DE DATA ANALYTICS EM MICRO E
PEQUENAS EMPRESAS BRASILEIRAS**

Livia Silva Marques

**Florianópolis - SC
2024/2**

UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA
CURSO DE SISTEMAS DE INFORMAÇÃO

PROPOSTA DE GUIA PARA INICIATIVAS DE DATA ANALYTICS EM
MICRO E PEQUENAS EMPRESAS BRASILEIRAS

Livia Silva Marques

Trabalho de Conclusão de Curso apresentado
como parte dos requisitos para obtenção do grau
de Bacharel em Sistemas de Informação pela
Universidade Federal de Santa Catarina.

Orientador: Prof. Dr. Jean Carlo Rossa Hauck

FLORIANÓPOLIS

2024/2

LÍVIA SILVA MARQUES

PROPOSTA DE GUIA PARA INICIATIVAS DE DATA ANALYTICS EM MICRO E
PEQUENAS EMPRESAS BRASILEIRAS

Trabalho de conclusão de curso apresentado
como parte dos requisitos para obtenção do
grau de Bacharel em Sistemas de Informação.

Orientador: Prof. Dr. Jean Carlo Rossa Hauck

Banca Examinadora:

.....

.....

Ficha catalográfica gerada por meio de sistema automatizado gerenciado pela BU/UFSC. Dados inseridos pelo próprio autor.

Marques, Livia Silva

PROPOSTA DE GUIA PARA INICIATIVAS DE DATA ANALYTICS EM
MICRO E PEQUENAS EMPRESAS BRASILEIRAS / Livia Silva
Marques ; orientador, Jean Carlo Rossa Hauck, 2024.
123 p.

Trabalho de Conclusão de Curso (graduação) -
Universidade Federal de Santa Catarina, Centro
Tecnológico, Graduação em Sistemas de Informação,
Florianópolis, 2024.

Inclui referências.

1. Sistemas de Informação. 2. Data Analysis. 3. Fatores
Críticos de Sucesso. I. Hauck, Jean Carlo Rossa. II.
Universidade Federal de Santa Catarina. Graduação em
Sistemas de Informação. III. Título.

À minha mãe, Cynthia Nalila, de quem certamente herdei o interesse por processos e a disposição para melhorá-los.

AGRADECIMENTOS

Agradeço inicialmente aos meus pais que, de diferentes ângulos, me guiaram a olhar atentamente as coisas da vida e buscar agir. O tema deste trabalho surgiu de uma dessas observações.

Ao meu orientador, professor Jean, que marcou positivamente o meu período de graduação diversas vezes. Agradeço a dedicação ao ensinar e o exemplo enquanto pesquisador e profissional.

Às minhas irmãs e irmãos por estarem sempre logo ali. Quanto aos meus sobrinhos, não vou citar as vezes que deixei de estudar para brincar e vou apenas agradecer por tornarem os dias mais alegres. Às minhas amigas, especialmente Duda, Jas e Lu - que, nos difíceis e gostosos encontros de agenda, me trouxeram inúmeros momentos legais durante a graduação. Ao Doni, pela dedicação infinita, pelas ajudas com os códigos e por ter estado aqui no início e no fim dessa jornada que é a faculdade.

Por fim, agradeço às oportunidades de ensino e pesquisa que a Universidade Federal de Santa Catarina me proporcionou, especialmente devido aos excelentes professores envolvidos.

SUMÁRIO

SUMÁRIO	6
LISTA DE FIGURAS	8
LISTA DE QUADROS	9
RESUMO	10
1. INTRODUÇÃO	1
1.1 OBJETIVOS.....	2
1.1.1 Objetivo Geral.....	2
1.1.2 Objetivos Específicos.....	3
1.2 MÉTODO DE PESQUISA.....	3
2 REFERENCIAL TEÓRICO	8
2.1 DADO, INFORMAÇÃO E CONHECIMENTO: MATÉRIAS PRIMAS DA VANTAGEM COMPETITIVA.....	8
2.2 O PROCESSO DE DATA ANALYTICS.....	10
2.4 FORÇAS IMPULSIONADORAS E CONTRÁRIAS: CONHECENDO OS FATORES CRÍTICOS DE SUCESSO PARA IMPLEMENTAÇÃO DE DATA ANALYTICS.....	12
2.5 MICRO E PEQUENAS EMPRESAS DE SOFTWARE DO BRASIL.....	14
2.6 ESSENCE.....	15
3. ESTADO DA ARTE	18
3.1 DEFINIÇÃO DO PROTOCOLO DE MAPEAMENTO.....	18
3.2 EXECUÇÃO DA BUSCA.....	21
3.3 EXTRAÇÃO E ANÁLISE DOS DADOS.....	22
3.3.1 PA1 - Quais estudos identificam fatores críticos de sucesso de iniciativas de Data Analytics em micro e pequenas empresas?.....	23
3.3.2 PA2 - Quais são os fatores críticos de sucesso identificados por esses estudos?.....	26
3.3.3 PA3: Quais são as características das organizações avaliadas?.....	30
3.3.4 PA4: Quais as metodologias usadas nos processos de Data Analytics avaliados?.....	31
3.3.5 PA5: Quais resultados são categorizados como de sucesso ou de fracasso?.....	32
3.4 DISCUSSÃO.....	32
3.5 AMEAÇAS À VALIDADE.....	34
4 ESTADO DA PRÁTICA	36
4.1 PLANEJAMENTO.....	36
4.1.1 Amostragem.....	36
4.2 INSTRUMENTO DE COLETA DE DADOS.....	37
4.3 DISTRIBUIÇÃO DO QUESTIONÁRIO.....	39
4.4 ANÁLISE DOS DADOS.....	39
4.4.1 PP1 - Quantas organizações aplicam Data Analytics no seu dia a dia?...	

4.4.2 PP2 - Quais fatores são considerados críticos para o sucesso do processo de Data Analytics?.....	40
4.4.3 PP3 - Quais as características das organizações consultadas?.....	41
4.4.4 PP4 - Quais as metodologias/abordagens usadas pelas empresas para implementar Data Analytics?.....	43
4.4.5 PP5 - Quais etapas do processo de Data Analytics são efetivamente executadas pelas empresas?.....	44
4.4.6 PP6 - Quais os resultados observados com o processo de Data Analytics?.....	45
4.5 DISCUSSÃO.....	47
4.6 AMEAÇAS À VALIDADE.....	49
5. GUIA PARA IMPLEMENTAÇÃO DO PROCESSO DE DATA ANALYTICS.....	50
5.1 Estrutura do Guia e da Ferramenta de Autoavaliação.....	50
5.2 Conteúdo do Guia.....	52
5.2.1 Alpha 1: Gerenciamento do processo de dados.....	53
5.2.2 Alpha 2: Preparação de dados.....	54
5.2.3 Alpha 3: Qualificação de profissionais.....	56
5.2.4 Alpha 4: Clareza sobre objetivos.....	59
5.2.5 Alpha 5: Suporte organizacional.....	61
5.3 Apresentação final do guia e ferramenta de autoavaliação.....	63
5.4 Avaliação do Guia.....	65
5.4.1 Resultados da avaliação.....	69
5.5 Ameaças à validade do Guia.....	69
6. CONCLUSÃO.....	71
REFERÊNCIAS BIBLIOGRÁFICAS.....	73

LISTA DE FIGURAS

Figura 1 - Metodologia de pesquisa	4
Figura 2 - Visão geral do CRISP-DM	11
Figura 3 - Representação da teoria de forças de campo de Lewin.....	13
Figura 4 - Exemplos de cards Essence.....	16
Figura 5 - Fatores considerados críticos por número de ocorrências.....	41
Figura 6 - Prevalência de segmento das organizações.....	42
Figura 7 - Quantidade de profissionais dedicados a dados.....	43
Figura 8 - Análise das fases efetivamente executadas pelas organizações ...	44
Figura 9 - Menções aos impactos observados.....	46
Figura 10 - Estrutura do Guia.....	50
Figura 11 - Estrutura da ferramenta de avaliação.....	51
Figura 12 - Alpha 1: Gerenciamento do processo de dados	54
Figura 13 - Alpha 2: Preparação de dados.....	56
Figura 14 - Alpha 3: Qualificação de profissionais	58
Figura 15 - Alpha 4: Clareza sobre objetivos	60
Figura 16 - Alpha 5: Suporte organizacional	62
Figura 17 - Pré visualização do website do Guia.....	63
Figura 18 - Template da Ferramenta de Autoavaliação Guia disponibilizado	64
Figura 19 - Visão geral da abordagem GQM.....	66

LISTA DE QUADROS

Quadro 1-	Alguns elementos chave do Essence.....	16
Quadro 2-	Mapeamento de termos com PICOC adaptado	19
Quadro 3-	Expressões de busca finais.....	20
Quadro 4-	Resultado das buscas por base pesquisada	21
Quadro 5-	Informações para coleta relacionadas às perguntas de análise	22
Quadro 6-	Estudos relevantes encontrados a partir do mapeamento	23
Quadro 7-	Grupos de fatores e respectivos autores	27
Quadro 8-	Fatores por categoria e popularidade	29
Quadro 9-	Características das empresas entrevistadas por estudo.....	30
Quadro 10-	Perguntas de pesquisa do <i>survey</i>	36
Quadro 11-	Perguntas do questionário.....	38
Quadro 12-	Alphas e frequência com que aparecem nos estudos realizados.....	52
Quadro 13-	Perfil do painel de especialistas	68

RESUMO

A transformação digital e o uso estratégico de dados têm impulsionado o desempenho organizacional, especialmente em pequenas e médias empresas (PMEs) no Brasil. Este trabalho investiga os fatores críticos para a implementação de Data Analytics em micro e pequenas empresas brasileiras de software, propondo um guia baseado na abordagem Essence. A pesquisa compreendeu um mapeamento sistemático da literatura para identificar fatores críticos recorrentes e um levantamento com profissionais do mercado, resultando na priorização de cinco fatores principais: gerenciamento do processo de dados, preparação dos dados, clareza sobre a necessidade do processo, qualificação profissional e suporte organizacional. A partir dessas descobertas, foi desenvolvido um guia online que apresenta boas práticas na implementação do processo de dados e permite autoavaliação e monitoramento da evolução das organizações nesse aspecto. Os resultados indicam a relevância de iniciativas de Data Analytics e como as organizações podem buscar uma boa execução desse processo.

Palavras-chave

Data Analytics; Pequenas e médias empresas; Fatores críticos de sucesso; Essence.

ABSTRACT

Digital transformation and the strategic use of data have driven organizational performance, especially among small and medium-sized enterprises (SMEs) in Brazil. This study investigates the critical success factors for implementing Data Analytics in Brazilian micro and small software companies, proposing a guide based on the Essence approach. The research included a systematic literature review to identify recurring critical factors and a survey with market professionals, resulting in the prioritization of five main factors: data process management, data preparation, clarity regarding process needs, professional qualification, and organizational support. Based on these findings, an online guide was developed, presenting best practices for implementing data processes and enabling self-assessment and progress monitoring for organizations in this area. The results highlight the relevance of Data Analytics initiatives and how organizations can strive for successful process execution.

Keywords

Data Analytics; Small and medium-sized enterprises; Critical success factors; Essence.

1. INTRODUÇÃO

Os tempos atuais são caracterizados pela transformação digital. Se desde o início do século XXI já se pode considerar relevante o impacto da ciência e análise de dados (Cao, 2017), os últimos anos têm trazido impressionantes avanços.

Uma grande quantidade de dados é gerada por um número cada vez maior de pessoas, que utilizam smart devices e a internet (Karimi e Javdani, 2021). Com o uso de tecnologias modernas como suporte, as organizações têm o poder de aplicar análises sobre esses dados e gerar conhecimento que leve à uma tomada de decisão de sucesso e consequente melhora na performance de seu negócio (Baijens et al., 2020).

Bhatiasevi e Naglis (2018) citam a área de Business Intelligence (BI) como a principal ferramenta usada pelas empresas no que tange suas aplicações analíticas, especialmente nos últimos anos. Autores como Owusu (2017) e Anwar et.al. (2018), que estudam o impacto da adoção de BI, apontam que as empresas reportam ter percebido benefício significativo em vantagem competitiva e performance.

Em análise do cenário empresarial, Bhardwaj (2022) cita Tarek & Adel (2016) e Wang & Wang (2020), ao dizer que os pequenos e médios negócios (SME) têm crescido muito nos últimos anos, representando uma parte significativa da quantidade de empregos e receita de um país. O autor estima que esse crescimento ainda triplique até 2030, destacando seu impacto.

Ainda em relação aos pequenos e médios negócios, em contraste com sua importância - mas também conectado com sua natureza experimental - Tarek & Adel, (2016) e Wang & Wang (2020) apontam que 65% deles falham todos os anos. Dentre os diversos motivos, está a falta de uso estratégico de recursos como os dados.

Bhatiasevi e Naglis (2018) falam sobre a pressão competitiva que pequenos e médios negócios sofrem ao saber que seus concorrentes utilizam BI em suas análises e decisões. Por isso, esforços em BI que contribuam para manutenção da vantagem competitiva nas empresas devem ser estudados e empregados.

Nesse sentido, já se tem estudos que apresentam essa relação linear entre o emprego de análise de dados no cotidiano de uma empresa e a melhora em seus resultados e performance (O'Connor & Kelly, 2017). No entanto, entender essa

relação não é suficiente para garantir o sucesso de sua implantação. Ainda existe uma quantidade significativa de SMEs que não investem em análise de dados e as razões para isso podem ser muitas (Bhardwaj, 2022). Identificar os fatores críticos que inibem ou favorecem essa prática dentro das pequenas e médias empresas pode ser crucial.

Bhardwaj (2022) apresenta uma análise de pesquisas recentes relacionadas à análise de dados e SMEs. Alguns autores têm encontrado justamente a temática dos fatores determinantes para o sucesso da implementação da análise de dados nesses ambientes.

Saltz e Krasteva (2022) fizeram uma revisão sistemática da literatura para identificar abordagens de execução de projetos de big data science. Vários autores executam adaptações de metodologias já conhecidas (como CRISP-DM - *Cross-Industry Standard Process for Data Mining*) e também encontram propostas de novos métodos. A utilização de métodos bem estabelecidos visa minimizar os erros e realizar um desenvolvimento mais controlado.

Como forma de contribuir para a identificação dos fatores críticos encontrados na literatura, este estudo propõe realizar um mapeamento sistemático do estado da arte, para listar os principais fatores já identificados entre as pesquisas recentes. Depois disso, busca-se conhecer os fatores identificados por pequenas e médias empresas de software brasileiras, em suas práticas, por meio de um questionário estruturado. Finalmente, conhecendo o contexto brasileiro e o que a literatura apresenta sobre a adoção de Data Analytics, este trabalho tem como objetivo principal propor um Guia para a implementação de iniciativas de Data Analytics em micro e pequenas empresas brasileiras.

1.1 OBJETIVOS

A partir da contextualização do problema, foram definidos os objetivos do presente trabalho.

1.1.1 Objetivo Geral

O objetivo geral deste trabalho é o desenvolvimento de um Guia de boas práticas no processo de Data Analytics para micro e pequenas empresas brasileiras.

1.1.2 Objetivos Específicos

Visando o cumprimento do objetivo geral estabelecido, são definidos os seguintes objetivos específicos.

O1: Analisar e fundamentar teoricamente os conceitos relativos ao processo de Data Analytics.

O2: Identificar o estado da arte sobre os fatores que impactam o processo de Data Analytics.

O3: Identificar o estado da prática em relação aos fatores que impactam o processo de Data Analytics no contexto de micro e pequenas empresas de software brasileiras.

O4: Elaborar o Guia de boas práticas, com base nos estudos prévios.

O5: Desenvolver uma ferramenta de acesso online ao Guia.

O6: Avaliar a aplicabilidade, abrangência e cobertura do Guia desenvolvido.

1.2 MÉTODO DE PESQUISA

Esta seção descreve a metodologia de pesquisa que orientou o desenvolvimento deste trabalho, bem como o método de utilização dos dados no desenho de pesquisa.

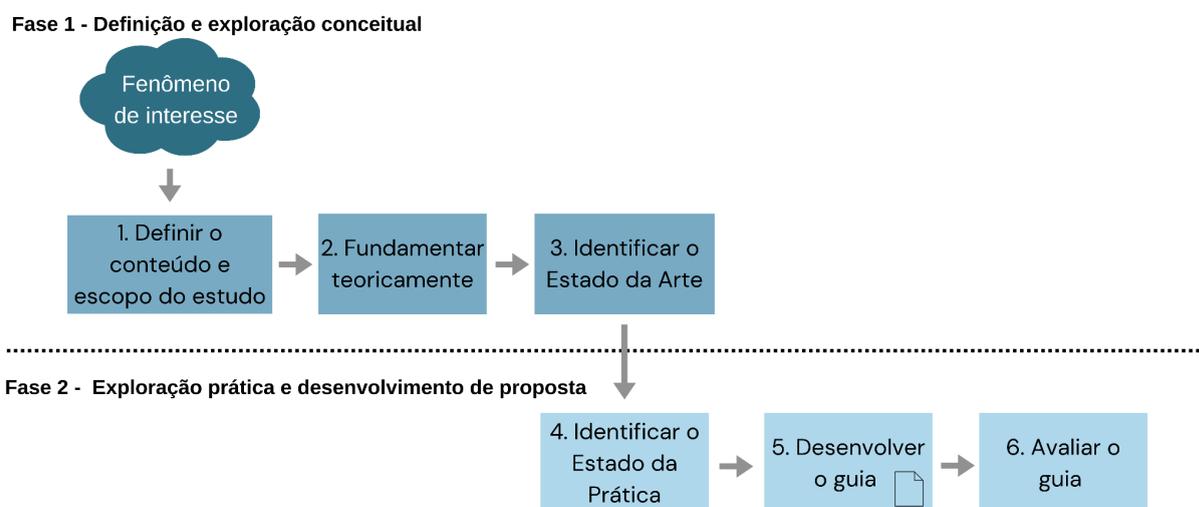
De forma geral, este trabalho foi dividido em duas fases: a primeira, inicial e teórica, traz a conceituação do problema de pesquisa e uma coleta de dados referente ao estudo do Estado da Arte (capítulo 3). A fase seguinte, de caráter mais prático, traz um *Survey* (Pinsonneault e Kraemer, 1993) descrito no capítulo 4, que identifica os principais fatores influentes no processo de Data Analytics em micro e pequenas empresas brasileiras de software, além da elaboração de um Guia (capítulo 5), que pretende auxiliar as organizações a evoluírem nesses fatores com base em práticas reconhecidas pela literatura.

Neste trabalho, o estudo com coleta de dados da primeira etapa faz o levantamento dos principais fatores que influenciam o processo de Data Analytics em micro e pequenas empresas brasileiras de software. O método de pesquisa utilizado tem como base aspectos da *Grounded Theory*, em que a teoria proposta pelo estudo se evidencia como produto da relação entre análise e coleta de dados (Goulding, 2009).

A fim de maximizar as informações que podem ser analisadas no corpo da pesquisa e aproveitar o melhor que cada método fornece (Paranhos et al., 2016), aplicou-se a metodologia multimétodo. De acordo com Creswell e Plano Clark (2011), os métodos mistos são um procedimento de coleta, análise e combinação de técnicas quantitativas e qualitativas em um mesmo desenho de pesquisa. Para os autores, a interação entre os métodos fornece melhores possibilidades analíticas.

A Figura 1, abaixo, apresenta o método de pesquisa detalhado nas atividades que o compõem.

Figura 1. Metodologia de pesquisa



Fonte: Elaborado pela autora, 2024

Atividade 1. Definir o conteúdo e escopo do estudo

A partir do fenômeno de interesse identificado, os capítulos de Introdução e Objetivos apresentaram o conteúdo e escopo do estudo.

Atividade 2. Fundamentar teoricamente

Nesta atividade do estudo, são compilados os conceitos fundamentais da literatura correlacionados com o tema do trabalho. É feita a conceitualização de termos utilizados no texto e das abordagens empregadas no desenvolvimento do trabalho.

Tarefa 2.1 – Apresentar a fundamentação dos principais conceitos relacionados ao processo de Data Analytics.

Tarefa 2.2 – Apresentar a fundamentação sobre forças impulsionadoras e restritivas.

Tarefa 2.3 – Contextualizar micro e pequenas empresas brasileiras de software.

Tarefa 2.4 – Conceituar termos e abordagens utilizadas no desenvolvimento do trabalho.

Atividade 3. Identificar o estado da arte

O processo de identificação do estado da arte, segundo os autores Vosgerau e Romanowski (2014, p. 172) “não se restringe a identificar a produção, mas analisá-la, categorizá-la e revelar os múltiplos enfoques e perspectivas”. Portanto, nesta atividade do estudo, são definidos os critérios de pesquisa e o levantamento de estudos correlacionados com processo de Data Analytics, além do desenvolvimento de análises da pesquisa realizada.

Tarefa 3.1 – Especificar o problema.

Tarefa 3.2 – Definir os critérios de pesquisa.

Tarefa 3.3 – Filtrar a busca a partir da visão de diferentes autores.

Tarefa 3.4 – Selecionar os estudos relacionados com o tema.

Tarefa 3.5 – Extrair e analisar os dados.

Atividade 4. Identificar o estado da prática

De acordo com Groves et al. (2009, p. 2), um survey é um "método sistemático para coletar informação de (uma amostra de) entidades com o propósito de construir descritores quantitativos dos atributos da população maior da qual essas entidades fazem parte". Esta atividade do trabalho é focada na construção e aplicação de um *survey* que vai complementar o estudo do estado da arte com as percepções coletadas do público alvo deste estudo.

Para esta atividade são definidas as seguintes tarefas:

Tarefa 4.1 – Planejar o survey.

Tarefa 4.2 – Aplicar o survey com o público alvo.

Tarefa 4.3 – Analisar os resultados obtidos.

Atividade 5. Desenvolver o Guia

Esta atividade do estudo tem como objetivo, a partir dos conhecimentos elencados em estudos teóricos e práticos, a elaboração do Guia e sua disponibilização.

Tarefa 5.1 – Definir o Guia.

Tarefa 5.2 – Coletar conteúdo para o Guia.

Tarefa 5.3 – Disponibilizar o Guia de forma *online*.

Atividade 6. Avaliar o Guia

Esta atividade do estudo tem como objetivo a avaliação do Guia desenvolvido, utilizando técnicas para verificar a aplicabilidade, abrangência e cobertura do conteúdo do mesmo, a fim de verificar se o que foi proposto teoricamente tem aderência prática, se o conteúdo do Guia pode ser aplicado em organizações com diferentes características e se a extensão do conteúdo fornece, em quantidade e qualidade, elementos necessários para o processo de análise de dados.

Tarefa 6.1 – Estruturar a avaliação.

Tarefa 6.2 – Desenvolver a avaliação.

Tarefa 6.3 – Realizar a avaliação, coletando dados.

Tarefa 6.4 – Analisar os dados coletados.

1.3 ESTRUTURA DO TRABALHO

O presente trabalho está organizado em 6 capítulos, que descrevem o processo elaborado a partir do método já apresentado. Esta introdução, que apresenta a motivação, objetivos e método de pesquisa é o primeiro deles.

Em seguida, o capítulo 2 apresenta o referencial teórico que fundamenta o desenvolvimento do trabalho com conceitos elementares para entendimento e contextualização do que se segue.

O capítulo 3, denominado 'Estado da Arte' apresenta o processo de levantamento do estado da arte em relação à identificação de fatores que contribuem para o processo de data analytics e os resultados desse levantamento. Na sequência, o capítulo 4 também faz um levantamento, agora do estado da prática, em que se identifica a mesma questão mas no contexto de micro e pequenas empresas brasileiras de software que participaram do survey.

O capítulo 5 apresenta o guia desenvolvido a partir dos estudos prévios, detalhando sua estrutura, conteúdo e a ferramenta de autoavaliação desenvolvida a partir dele.

O capítulo 6 é a conclusão do trabalho, descrevendo os resultados em relação aos objetivos definidos, aspectos relevantes identificados a partir do estudo e o levantamento de possíveis temas para trabalhos futuros relacionados.

2 REFERENCIAL TEÓRICO

Com o objetivo de fundamentar o desenvolvimento deste trabalho e estabelecer definições básicas que facilitem o seu entendimento, este capítulo apresenta alguns conceitos fundamentais.

É definida a diferença e a relação entre dado, informação e conhecimento e, em seguida, o processo de ciência de dados, que envolve todos eles. A teoria de Kurt Lewin (1948) apresenta o conceito de forças impulsionadoras e restritivas, para auxiliar no entendimento de como essas forças afetam a implementação de processos de Data Analytics no contexto de micro e pequenas empresas de software brasileiras - definidas na seção seguinte. Por fim, são apresentados elementos da abordagem Essence utilizados, neste estudo, no desenvolvimento do Guia.

2.1 DADO, INFORMAÇÃO E CONHECIMENTO: MATÉRIAS PRIMAS DA VANTAGEM COMPETITIVA

O avanço das tecnologias sempre representou mais facilidade para a obtenção de conhecimento. Especialmente nos últimos anos, com o lançamento de ferramentas como o ChatGPT¹ (em 2022 pela empresa OpenAI), o conceito do uso de dados para obtenção de conhecimento e aplicação em questões do cotidiano ficou mais evidente. Desde muito antes disso, no entanto, já se estuda a relação dado-conhecimento, principalmente no que diz respeito ao seu uso como fator competitivo quando relacionado ao âmbito empresarial.

Conforme Davenport e Prusak (1998), o sucesso de uma organização pode estar bastante relacionado com o entendimento dos conceitos de "dado", "informação" e "conhecimento". Confundi-los, não saber identificar quais estão disponíveis ou mesmo não saber quando e como utilizar cada um, pode levar a investimentos que não geram o resultado esperado.

Considerando a relevância desses termos na prática empresarial, a definição de cada um e o esclarecimento sobre como se relacionam se fazem necessários. As definições apresentadas por Davenport e Prusak (1988) ainda são comumente

¹ <https://chatgpt.com/>

referenciadas, com ênfase especial nos termos que diferenciam dado, informação e conhecimento. Para os autores, esses termos podem ser assim definidos:

- a) **Dado:** um fato objetivo sobre um evento; um registro que, sozinho, não apresenta significado ou relevância. Normalmente é armazenado em sistemas tecnológicos, a partir de diferentes departamentos de uma empresa, que estão constantemente gerando esses dados em suas atividades e transações.
- b) **Informação:** uma mensagem, que possui um remetente e um destinatário, e tem o objetivo de informar e possivelmente mudar - ou, ao menos, instigar a mudança - na forma como o receptor percebe algo. A informação dá significado ao dado através, por exemplo, de contexto, categorização, cálculo, correções ou agrupamentos.
- c) **Conhecimento:** uma mistura de experiência, valores, insights e informação contextualizada. O conhecimento é gerado a partir da informação, mas necessariamente da interação dessa com um (e frequentemente mais de um) componente humano. Por conta da natureza mais complexa e imprevisível do pensamento humano, o conhecimento é uma estrutura também complexa e mais difícil de representar em termos lógicos e concretos.

Alguns problemas são comuns devido à falta de conhecimento sobre os termos apresentados e suas aplicações. Crenças inválidas podem atrapalhar o processo, como a ideia de que ter maior volume de dados é sempre positivo ou que uma melhor tecnologia vai necessariamente melhorar o estado da informação (Davenport e Prusak, 1998).

A partir dessa cadeia de definições, é possível chegar mais perto da tomada de decisão. O conhecimento, ao agregar dados, informações, experiências, verdades fundamentais e "regras de ouro", julgamentos, intuição, valores e crenças, se torna um ativo corporativo sustentável.

A vantagem advinda do conhecimento é sustentável porque gera retornos crescentes e vantagens contínuas. Diferentemente dos ativos materiais, que diminuem à medida que são usados, ativos de conhecimento aumentam com o uso: ideias geram novas ideias e conhecimento compartilhado permanece com o doador enquanto enriquece o receptor. O potencial para novas ideias decorrentes do estoque de conhecimento em qualquer empresa é praticamente ilimitado - especialmente se as pessoas na empresa recebem oportunidades de pensar, aprender e conversar uns com os outros (DAVENPORT E PRUSAK, 1998, p. 13).

Dessa forma, as empresas utilizam os dados de forma a potencializar suas operações e decisões à medida que conseguem trabalhá-los para que se tornem conhecimento.

2.2 O PROCESSO DE DATA ANALYTICS

No contexto do processo de transformar dados em conhecimento útil e estratégico, Sarker (2021) menciona a variedade de termos existentes e a facilidade de confundi-los: "data analysis", "data mining", "Data Analytics", "big data", "data science", "advanced analytics", e outros, como "business intelligence".

Gökalp (2021) identifica "Data Analytics" como sendo uma área que compreende toda a pipeline do processo de dados, necessária para gerar ideias de ações a partir de dados brutos. Segundo o autor, existem frameworks conhecidos em Data Analytics como CRISP-DM, SEMMA (Amostragem, Exploração, Modificação, Modelagem e Avaliação) e KDD (*Knowledge Discovery in Databases*).

Embora diferentes, os frameworks basicamente entendem Data Analytics como o processo que engloba, de forma geral, seis processos (Gökalp, 2021):

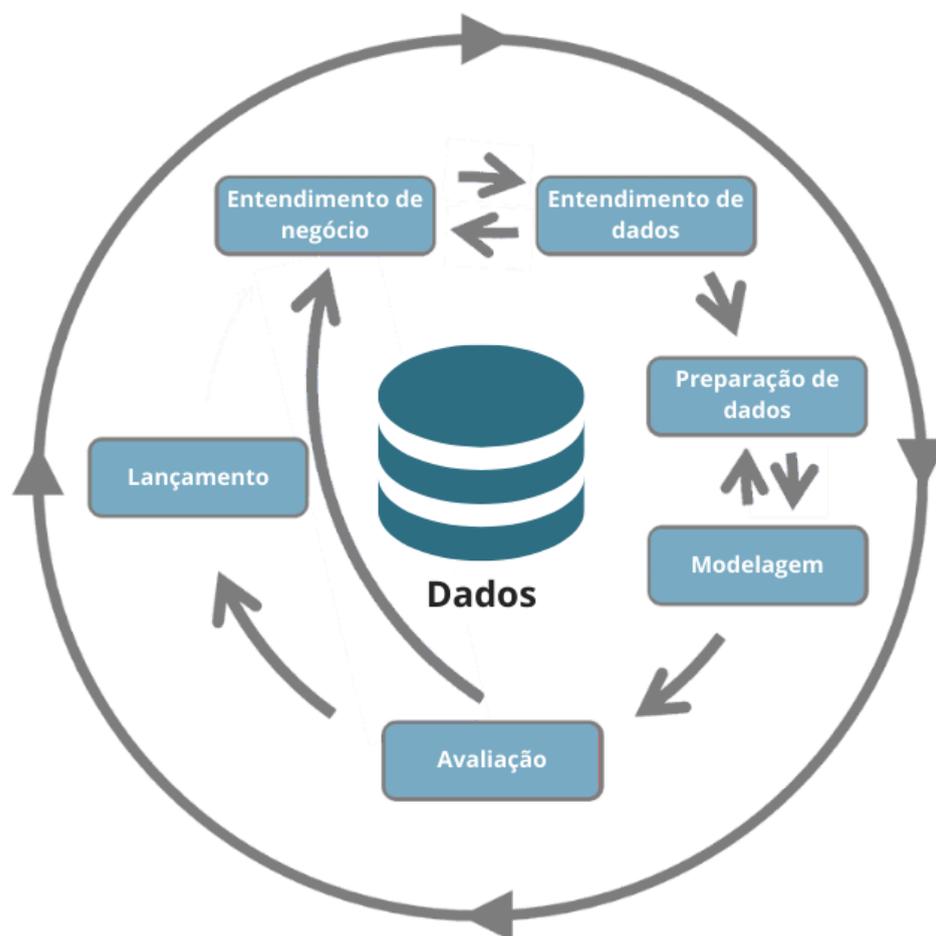
- a) Entendimento de negócio: identificação de necessidades e requisitos;
- b) Entendimento de dados: análise de necessidades e fontes de dados, qualidade de dados e entendimento de forma estatística;
- c) Preparação de dados: limpeza, manipulação, qualidade e estruturação dos dados;
- d) Construção de modelo: investigação de como o conhecimento deve ser extraído em tempo hábil e como implementar técnicas analíticas, desde análises descritivas até análises prescritivas;
- e) Avaliação: validação do sucesso dos produtos e serviços de dados desenvolvidos;
- f) Lançamento e uso: utilização dos insights gerados pelos produtos de dados.

A fim de evitar possíveis confusões em relação a nomenclatura, neste trabalho utilizaremos o termo "Data Analytics" como definido por Gökalp (2021) para

abrangência das etapas supracitadas e referência a todo processo de dados aqui estudado.

O CRISP-DM (abreviação do termo em inglês para Processo Padrão Inter-Indústrias para Mineração de Dados) é um modelo muito popular para execução dos processos de ciência de dados. Começando com o entendimento do negócio até a implementação da solução, são seis fases que compõem o modelo base - cuja execução pode não ser linear (Saltz, 2024), como ilustrado pela Figura 2.

Figura 2. Visão geral do CRISP-DM



Fonte: Elaborado pela autora, a partir de Data Science Alliance (2024)

No que diz respeito ao gerenciamento do processo de Data Analytics, Saltz et al. (2017) lembram que CRISP-DM não é um método de gerenciamento de projetos, apesar de descrever bem as etapas existentes. Utilizar apenas um modelo de ciclo

de vida como o CRISP-DM pode não ser suficiente, por não abranger aspectos de comunicação e interação entre o time.

Corroborando com essa ideia, Holtz (2024, p.02) diz que “equipes de ciência de dados que combinam uma implementação flexível do CRISP-DM com abordagens abrangentes de gerenciamento de projetos ágeis baseadas em equipe provavelmente obterão os melhores resultados”, sugerindo que uma combinação de modelos com metodologias de gerenciamento de projetos ágeis pode ser ideal, a fim de juntar aspectos próprios do processo de Data Analytics com práticas que levem a uma execução eficiente.

2.4 FORÇAS IMPULSIONADORAS E CONTRÁRIAS: CONHECENDO OS FATORES CRÍTICOS DE SUCESSO PARA IMPLEMENTAÇÃO DE DATA ANALYTICS

No contexto de introdução de um novo processo em uma organização, Montoni (2010) utiliza o termo "fator" como causa, interação e efeito que influenciam o sucesso de iniciativas de melhoria. Também cita Rockart (1979), que apresenta fatores críticos de sucesso (FCS) como o conjunto de questões importantes a que se deve estar atento para obtenção de êxito.

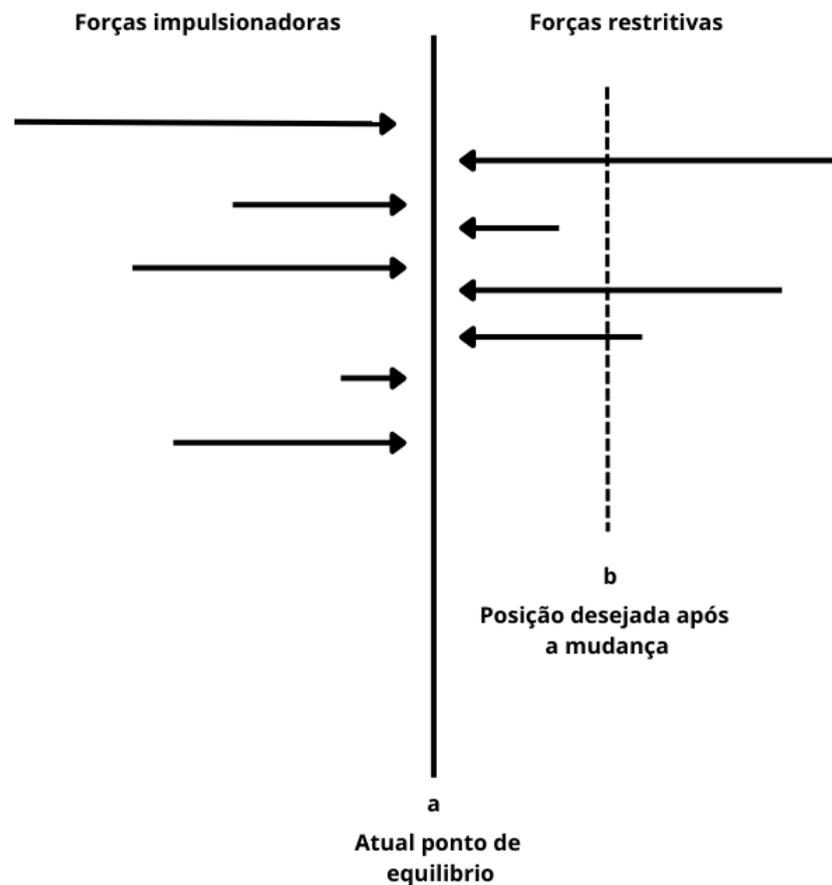
Fatores críticos de sucesso na implementação de um processo podem variar, dependendo de características da indústria em que a empresa está inserida, ou, ainda, de fatores individuais dentro de uma mesma indústria. "Algumas vezes é até mesmo necessário considerar FCS de funcionários de forma individual", afirmam Turban, McLean, & Wetherbe (2001).

Implementar um novo processo, uma nova prática ou mesmo introduzir um novo conceito em uma organização é uma tarefa que provoca atritos. Dependendo do estado atual e de fatores de impacto, pode ser mais ou menos difícil chegar ao estado desejado.

No estudo dessas mudanças de estado, Kurt Lewin (1948), em seu livro intitulado “Teoria de Campo”, propõe a teoria de que todos os seres representam um equilíbrio entre duas forças que puxam em direções opostas, levando à mudança de seus estados. A Figura 3, abaixo, apresenta uma representação visual da teoria de

Lewin.

Figura 3. Representação da teoria de forças de campo de Lewin



Fonte: Elaborado pela autora, a partir de Lewin (1948)

Dessa forma, a teoria captura a ideia de como diferentes forças se combinam para manter um estado da forma como ele é ou para alterá-lo. No geral, essa análise é feita quando se faz necessário um nível de entendimento sobre o estado atual e em quais forças mexer para alcançar o estado desejado.

Nas organizações, a análise de campo de forças de Lewin é usada para identificar as forças impulsionadoras e as forças contrárias a determinada mudança proposta. Com isso, é possível criar um plano de ação para restringir as forças contrárias e investir nas impulsionadoras.

Trazendo os conceitos de forças impulsionadoras e forças contrárias para o contexto da implementação de processos de análise de dados, alguns autores, como Bhardwaj (2022), apresentam pesquisas sobre o tema conforme revisão sistemática da literatura. Os fatores encontrados são categorizados em grandes

grupos que podem ter certa sobreposição, mas ainda assim facilitam a visualização.

2.5 MICRO E PEQUENAS EMPRESAS DE SOFTWARE DO BRASIL

A participação dos pequenos negócios na economia brasileira assume importância inquestionável. Esse segmento, liderado pelo setor de serviços, tem mostrado números expressivos, sendo responsável por 30% do PIB brasileiro e tendo acolhido 72% dos empregos criados no ano de 2022, segundo dados do SEBRAE (2023).

É nesse nicho da esfera econômica que se inserem as chamadas Micro e Pequenas Empresas de Base Tecnológica - EBT, que se dedicam, em especial, à inovação tecnológica.

Machado et al. (2001) estabelecem uma combinação entre uma proposta do congresso americano, apresentada pela OTA - *Office of Technology Assessment* e a definição adotada pelo SEBRAE para definir micro e pequenas empresas de base tecnológica. Para os autores,

Micro e pequenas empresas de base tecnológica são empresas industriais com menos de 100 empregados, ou empresas de serviço com menos de 50 empregados, que estão comprometidas com o projeto, desenvolvimento e produção de novos produtos e/ou processos, caracterizando-se, ainda, pela aplicação sistemática de conhecimento técnico-científico. Estas empresas usam tecnologias inovadoras, têm uma alta proporção de gastos com P&D, empregam uma alta proporção de pessoal técnico científico e de engenharia e servem a mercados pequenos e específicos (MACHADO ET AL., 2001, p. 07)

Portanto, o número limite de colaboradores, o alto grau de envolvimento com a busca de soluções tecnológicas para problemas e o contínuo uso do conhecimento técnico científico para chegar ao produto são prerrogativas centrais das micro e pequenas empresas de base tecnológica.

O relatório feito pela Associação Brasileira das Empresas de Software (ABES, 2023) indica que, no período da pesquisa, havia 8.478 empresas atuantes no desenvolvimento e produção de software. Dessas, cerca de 93,6% podem ser classificadas como micro e pequenas empresas, segundo análise realizada pelo critério de número de funcionários. O estudo considera microempresas aquelas com menos de 10 empregados e pequenas empresas aquelas com um número de funcionários entre 10 e 99.

2.6 ESSENCE

O Essence é um padrão para criação, uso e melhorias de práticas e métodos de engenharia de software. Criado e mantido pela *OMG International Open Standards Consortium*, o padrão se propõe a focar apenas nas informações essenciais para otimizar os processos de engenharia, fornecendo práticas leves e adaptáveis para cada contexto (OMG, 2018).

A documentação do Essence fornece um Kernel e uma Linguagem. Enquanto o Kernel descreve elementos, etapas e práticas comuns a toda execução de um processo de software, servindo de referência para times de software, a Linguagem é como se fosse o seu alfabeto.

A Linguagem apresentada pelo Essence descreve uma forma para escrita de práticas e elementos do processo de software. Por ser apresentada de maneira simples, padronizada e com elementos independentes, ela pode ser facilmente entendida e compartilhada, além de permitir diferentes combinações de seus elementos para caracterizar os processos de uma organização de forma customizada, e não como um framework fechado e inflexível.

O conjunto completo de elementos da linguagem suporta casos de uso avançados, mas o conceito de linguagem é projetado para aqueles usuários que desejam selecionar. Consequentemente, a linguagem é projetada para permitir um uso significativo mesmo com pequenos subconjuntos de elementos da linguagem. A maioria das associações e vários atributos são opcionais, de modo que os usuários não são obrigados a utilizar um grande conjunto de elementos da linguagem desde o início. Em vez disso, o conjunto completo de elementos da linguagem pode ser dividido em vários pequenos blocos que podem ser aprendidos e utilizados de forma independente e incremental (OMG, p.70).

Os elementos chave do Essence considerados relevantes para esse trabalho são apresentados no Quadro 1: alphas, estados do alpha, *patterns* e recursos.

Quadro 1. Alguns elementos chave do Essence

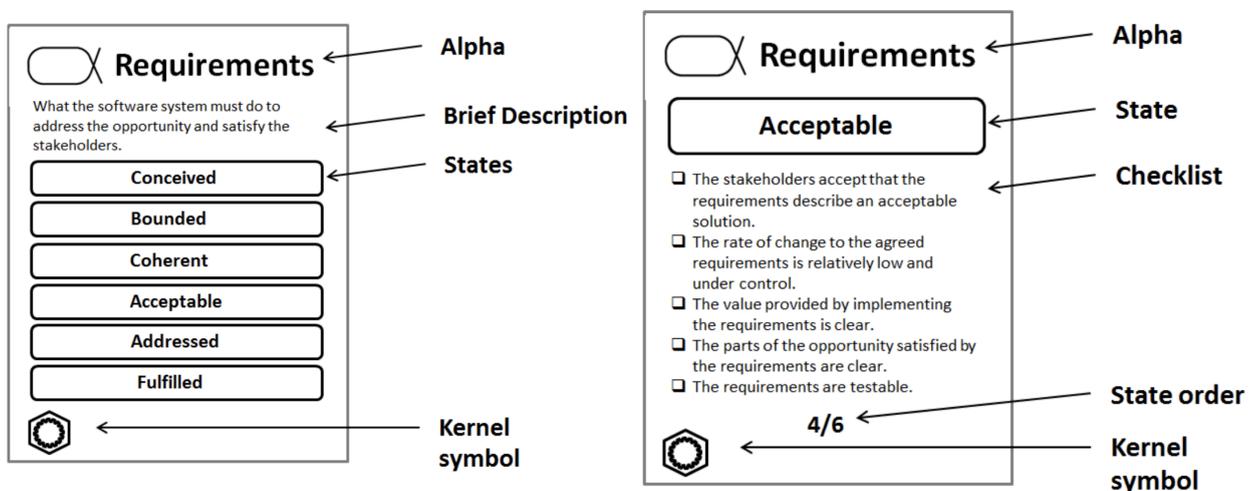
Símbolo identificador	Elemento	Descrição
	Alpha	Aspecto chave ou elemento do qual se quer monitorar o progresso; o resultado em que se quer manter o foco.
	Estados do Alpha	Especificação do estado de progresso de um Alpha. Podem ser quebrados em uma lista de itens a serem alcançados para passar ao próximo estado (ou verificar o estado atual).
	Pattern	Estrutura genérica existente para permitir descrever algum elemento adicional que se relaciona com outro elemento Essence.
	Recursos	Fonte de informação ou conteúdo (como site ou livro) que serve de apoio para aprofundamento e implementação.

Fonte: Elaborado pela autora a partir do Essence Pocket Guide v2.1, 2024

O Essence sugere a geração de cards como ferramenta de visualização dos elementos mais importantes. De acordo com a documentação do Essence, "um card apresenta um resumo sucinto das coisas mais importantes que devem ser lembradas sobre um elemento" (OMG, 2018).

A seguir, a Figura 4 traz exemplos de *cards* de um Alpha e de um dos Estados desse Alpha fornecidos na documentação.

Figura 4. Exemplos de cards Essence



Fonte: Essence 2.0 - Beta, OMG (2018)

Os cards, embora possam ser apresentados com variações (OMG, 2018), possuem os seguintes elementos em sua especificação:

- *Alpha*: nome do Alpha, já descrito anteriormente;
- *Brief description*: breve descrição do que o Alpha representa;
- *States*: estados do Alpha, conceito já descrito anteriormente;
- *Checklist*: afirmações que, quando consideradas como verdade, pela organização, marcam a evolução em direção ao próximo estado do Alpha;
- *State order*: a ordem do sub-elemento (quando há uma hierarquia);
- *Kernel symbol*: símbolo que representa o Kernel ao qual o elemento pertence. Kernel é um outro conceito da especificação Essence, não abordado neste trabalho.

De forma geral, o Essence provê elementos para a criação, uso e melhoria de métodos de engenharia de software (OMG, 2018). O uso do Essence em contextos de projetos de dados está sendo recentemente explorado, mas estudos como o de Durango Vanegas et al. (2023) já demonstraram bons resultados.

3. ESTADO DA ARTE

Neste capítulo é analisado o estado da arte quanto aos fatores que influenciam no sucesso de iniciativas de Data Analytics. Assim, para reunir esses estudos e apresentar um panorama geral do que já foi abordado pela literatura, realizou-se um Mapeamento Sistemático da Literatura (MSL). Este mapeamento, detalhado neste capítulo, segue os procedimentos propostos por Petersen et al. (2008).

3.1 DEFINIÇÃO DO PROTOCOLO DE MAPEAMENTO

A execução do MSL deve seguir um protocolo, definido previamente, para que os resultados atendam a critérios claros, permitindo uma comparação de resultados e análises consistentes.

O objetivo desse mapeamento é conhecer como fatores críticos impactam iniciativas de Data Analytics em micro e pequenas empresas de base tecnológica. Para responder a essa pergunta de pesquisa, foram definidas 6 perguntas de análise a serem respondidas a partir dos estudos encontrados no mapeamento:

PA1: Quais estudos identificam fatores críticos de sucesso de iniciativas de Data Analytics em micro e pequenas empresas?

PA2: Quais são os fatores críticos de sucesso identificados por esses estudos?

PA3: Quais são as características das organizações avaliadas?

PA4: Quais as metodologias usadas nos processos de Data Analytics avaliados?

PA5: Quais resultados são relatados nas iniciativas de Data Analytics?

Para encontrar estudos que contribuíssem na resposta dessas perguntas, foi elaborada uma expressão de busca para a pesquisa. A partir de uma busca exploratória inicial, foram encontrados os termos mais utilizados em estudos relacionados com o tema desse mapeamento.

Percebeu-se que existem variações de nomenclatura e de definição entre diferentes países, no que se refere ao porte das empresas. Por isso, trabalhos com

médias empresas foram aceitos, desde que fizessem parte da classificação "pequenas e médias", e não apenas "médias".

Além disso, apesar de ser o foco geral do presente trabalho, após experimentações iniciais de termos de busca em ferramentas de pesquisa, foi removida a limitação de estudos apenas com empresas de base tecnológica, por não serem encontrados resultados significativos com essa restrição.

Com o auxílio do framework PICOC, introduzido por Richardson et al., 1995, e adaptado para o contexto da pesquisa e da área, os termos foram classificados para melhor formulação da pesquisa e estão apresentados no Quadro 2.

Quadro 2. Mapeamento de termos com PICOC adaptado

Crítérios	Palavra chave	Sinônimos	Tradução
População	iniciativas de data Analytics	implementação de Data Analytics, processo de Data Analytics, uso de Data Analytics, big Data Analytics, business intelligence	data analysis
Intervenção	fatores críticos de sucesso	fatores de sucesso, estratégias de sucesso, processos para sucesso, desafios, barreiras	enabling/restraining force, factor, strategy, process, determinants, challenges, barriers
Comparação	-	-	-
Resultado	sucesso ou fracasso na implementação de Data Analytics	sucesso, fracasso, eficiência, desempenho, impacto	success, failure, efficiency, performance, impact, adoption
Contexto	micro e pequenas empresas de base tecnológica	MPE, micro e pequena empresa, micro empresa, pequena empresa	SME, small and medium business, small business, very small business, medium business

Fonte: Elaborado pela autora, 2024

Esses termos foram utilizados para definição de algumas expressões de busca, a fim de encontrar aquela que trouxesse melhores resultados. O critério para escolha da expressão final foi o retorno de, no mínimo, dois artigos chave previamente conhecidos pela autora (Maroufkhani et al., 2020; Ziembra e Olszak, 2012) e a apresentação de outros artigos relevantes.

A expressão final, que gerou os melhores resultados de acordo com os artigos chave, é apresentada no Quadro 3.

Quadro 3. Expressões de busca finais

Expressões de busca	
Inglês	("Data Analytics" OR "data analysis" OR "business intelligence") AND (forces OR capabilities OR factors OR success OR failure OR efficiency OR determinants OR challenges OR barriers) AND (sme OR "small and medium business" OR "small business" OR "medium business")
Português	("Data Analytics" OR "data analysis" OR "business intelligence" OR "análise de dados") AND (forças OR capacidades OR fatores OR sucesso OR fracasso OR eficiência OR determinantes OR desafios OR barreiras) AND (mpe OR "micro e pequenas empresas" OR "micro empresas" OR "pequenas empresas")

Fonte: Elaborado pela autora, 2024

Apesar da priorização da busca em inglês devido a uma maior abrangência de resultados, também foi definida uma expressão de busca em português, para minimizar o risco de exclusão, especialmente de contextos brasileiros.

Vale ressaltar que a escolha por incluir, na busca em inglês, o termo referente aos "*medium business*" (negócios de porte médio) se deu pelo fato de que não há uma padronização na classificação das empresas entre diferentes países. Conforme a Organização para a Cooperação e Desenvolvimento Econômico - OECD - empresas de médio porte possuem de 50 a 249 empregados - causando um conflito com a classificação brasileira definida na seção 2.5 deste estudo. Sendo assim, optou-se por abranger os estudos internacionais de médias empresas, para capturar o máximo de informações possíveis, já que a classificação SME (Small and Medium Business) é uma das que mais se mostrou presente nas pesquisas iniciais.

Considerando os diversos estudos que seriam resultantes da expressão de busca, foram definidos critérios para filtragem dos trabalhos que seriam considerados relevantes para inclusão neste mapeamento. estes critérios são os seguintes:

- a) Possuir evidências de que foram aplicados em micro e/ou pequenas e médias organizações.
- b) Citar fatores que, a partir da experiência em organizações, demonstraram exercer influência nas iniciativas.

- c) Ter sido realizado a partir do ano de 2010, a fim de não considerar fatores que possam já ter efeitos muito diferentes.

3.2 EXECUÇÃO DA BUSCA

Para a busca dos trabalhos, foram consideradas quatro bases de trabalhos acadêmicos: Scopus², ACM³, IEEE⁴ e Research Gate⁵. Nota-se que cada base possui um mecanismo de busca diferente, tendo sido observadas diferenças na sintaxe da busca. No entanto, a lógica e a terminologia que se utilizou em todas as bases foram as mesmas.

A busca foi realizada pela autora em junho de 2024, tendo retornado um total de 1821 resultados iniciais totais. A quantidade de trabalhos encontrados em cada base está descrita no Quadro 4, a seguir.

Quadro 4. Resultado das buscas por base pesquisada

	Scopus	ACM	IEEE	Research Gate	Total
Resultados iniciais	530	1113	78	100	1821
Resultados iniciais analisados	100	100	78	100	378
Resultados iniciais selecionados	8	1	12	15	36
Resultados potencialmente relevantes	5	0	5	8	18
Resultados relevantes	4	0	2	3	9

Fonte: Elaborado pela autora, 2024

Considerando as limitações de tempo para filtragem dos resultados iniciais, foram analisados os 100 primeiros resultados mais relevantes de cada base (quando havia mais de 100 resultados).

Os resultados iniciais selecionados foram escolhidos com base na leitura dos títulos e descrições sucintas apresentadas pelas plataformas, por meio da aplicação dos critérios de inclusão definidos. Em seguida, os resultados potencialmente relevantes foram escolhidos a partir da leitura do resumo de cada trabalho. Por fim,

² www.scopus.com

³ <https://dl.acm.org/>

⁴ <https://ieeexplore.ieee.org/>

⁵ www.researchgate.net

a leitura completa dos materiais definiu os 10 artigos relevantes finais encontrados no mapeamento.

Um outro artigo, não encontrado no mapeamento por não ter sido publicado em nenhuma das bases pesquisadas, foi incluído na análise dos dados para garantir melhor abrangência dos resultados. O artigo atende a todos os critérios de inclusão e está publicado em uma revista com critérios bem definidos de aceitação, de modo que o trabalho em questão apresenta uma contribuição relevante ao tema pesquisado.

3.3 EXTRAÇÃO E ANÁLISE DOS DADOS

Com os artigos relevantes selecionados, foram definidas quais informações seriam extraídas dos artigos, a fim de responder às perguntas de análise apresentadas na seção 3.1 deste estudo. Essas informações estão organizadas no Quadro 5.

Quadro 5. Informações para coleta relacionadas às perguntas de análise

Perguntas de análise	Item	Descrição
PA1: Quais estudos identificam fatores críticos de sucesso de iniciativas de Data Analytics em micro e pequenas empresas?	Título	Título do estudo
	Ano	Ano de publicação do estudo
	Referência	Referência do estudo encontrado
PA2: Quais são os fatores críticos de sucesso identificados por esses estudos?	Fator	Identificação do fator
	Descrição	Detalhamento (se houver)
PA3: Quais são as características das organizações avaliadas?	Setor	Setor de atuação identificado
	Localização	Localização da empresa (país)
PA4: Quais as metodologias usadas nos processos de Data Analytics avaliados?	Nome	Nome da metodologia aplicada
PA5: Quais resultados são relatados nas iniciativas de Data Analytics?	Resultado	Descrição do resultado observado
	Avaliação	Avaliação do resultado (de sucesso ou de fracasso)

Fonte: Elaborado pela autora, 2024

Dessa forma, cada item da tabela 5 foi identificado, quando presente, nos 10 estudos relevantes encontrados. A partir desses itens, foi possível responder às 6 perguntas de análise definidas.

Na sequência, cada uma das questões definidas é analisada a partir dos dados coletados.

3.3.1 PA1 - Quais estudos identificam fatores críticos de sucesso de iniciativas de Data Analytics em micro e pequenas empresas?

Foram encontrados, no total, dez estudos que elencam fatores de influência na adoção de Data Analytics por pequenas e médias empresas. A relação desses estudos pode ser vista no Quadro 6.

Quadro 6. Estudos relevantes encontrados a partir do mapeamento

Identificador	Citação	Título	Ano	Referência
1	(Scholz et al., 2010)	Benefits and Challenges of Business Intelligence Adoption in Small and Medium-Sized Enterprises	2010	Scholz, Patrick & Schieder, Christian & Kurze, Christian & Gluchowski, Peter & Böhringer, Martin. (2010). Benefits and Challenges of Business Intelligence Adoption in Small and Medium-Sized Enterprises. 18th European Conference on Information Systems, ECIS 2010.
2	(Olszak e Ziembra, 2012)	Critical success factors for implementing business intelligence systems in small and medium enterprises on the example of upper Silesia, Poland	2012	OLSZAK, Celina M.; ZIEMBA, Ewa. Critical success factors for implementing business intelligence systems in small and medium enterprises on the example of upper Silesia, Poland. <i>Interdisciplinary Journal of Information, Knowledge, and Management</i> , v. 7, p. 129-150, 2012. DOI: 10.28945/1584. Disponível em: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84862019547&doi=10.28945%2f1584&partnerID=40&md5=766ffe1471dfd5f9b6d74eb86b0ac340 .

3	(Puklavec, Oliveira e Popovič, 2014)	Unpacking Business Intelligence Systems Adoption Determinants: An Exploratory Study of Small and Medium Enterprises	2014	Puklavec, B., Oliveira, T., & Popovič, A. (2014). Unpacking Business Intelligence Systems Adoption Determinants: An Exploratory Study of Small and Medium Enterprises. <i>Economic and Business Review</i> , 16(2). https://doi.org/10.15458/2335-4216.1278
4	(Gudfinnsson e Strand, 2017)	Challenges with BI adoption in SMEs	2017	K. Gudfinnsson and M. Strand, "Challenges with BI adoption in SMEs," 2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA), Larnaca, Cyprus, 2017, pp. 1-6, doi: 10.1109/IISA.2017.8316407.
5	(Mohamed e Weber, 2020)	Trends of digitalization and adoption of big data & analytics among UK SMEs: Analysis and lessons drawn from a case study of 53 SMEs	2020	M. Mohamed and P. Weber, "Trends of digitalization and adoption of big data & analytics among UK SMEs: Analysis and lessons drawn from a case study of 53 SMEs," 2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Cardiff, UK, 2020, pp. 1-6, doi: 10.1109/ICE/ITMC49519.2020.9198545.
6	(Maroufkhani et al., 2020)	Big Data Analytics adoption: Determinants and performances among small to medium-sized enterprises	2020	Parisa Maroufkhani, Ming-Lang Tseng, Mohammad Iranmanesh, Wan Khairuzzaman Wan Ismail, Haliyana Khalid, Big Data Analytics adoption: Determinants and performances among small to medium-sized enterprises, <i>International Journal of Information Management</i> , Volume 54, 2020, 102190, ISSN 0268-4012, https://doi.org/10.1016/j.ijinfomgt.2020.102190 . (https://www.sciencedirect.com/science/article/pii/S026840122030178X).
7	(Alsyouf et al., 2022)	Factors Influencing the Adoption of Big Data Analytics in the Digital Transformation Era: Case Study of Jordnian SMEs	2022	Alsyouf, Adi & Lutfi, Abdalwali & Almaiah, Drmohammed & Alrawad, Mahmaod. (2022). Factors Influencing the Adoption of Big Data Analytics in the Digital Transformation Era: Case Study of Jordanian SMEs. <i>Sustainability</i> . 14. 10.3390/su14031802.

8	(Willetts, Atkins e Stanier, 2022)	Quantitative study on Barriers of adopting Big Data Analytics for UK and Eire SMEs	2022	Willetts, Matthew & Atkins, Anthony & Stanier, C.. (2022). Quantitative Study on Barriers of Adopting Big Data Analytics for UK and Eire SMEs. 10.1007/978-981-16-2937-2_23.
9	(Perdana et al., 2022)	Data analytics in small and mid-size enterprises: Enablers and inhibitors for business value and firm performance	2022	Arif Perdana, Hwee Hoon Lee, SzeKee Koh, Desi Arisandi, Data analytics in small and mid-size enterprises: Enablers and inhibitors for business value and firm performance, International Journal of Accounting Information Systems, Volume 44, 2022, 100547, ISSN 1467-0895, https://doi.org/10.1016/j.accinf.2021.100547 . (https://www.sciencedirect.com/science/article/pii/S146708952100049X)
10	(Justy et al., 2023)	On the edge of Big Data: Drivers and barriers to Data Analytics adoption in SMEs	2023	Théo Justy, Estelle Pellegrin-Boucher, Denis Lescop, Julien Granata, Shivam Gupta, On the edge of Big Data: Drivers and barriers to Data Analytics adoption in SMEs, Technovation, Volume 127, 2023, 102850, ISSN 0166-4972, https://doi.org/10.1016/j.technovation.2023.102850 . (https://www.sciencedirect.com/science/article/pii/S016649722300161X)

Fonte: Elaborado pela autora, 2024

Observa-se que o tema já havia demonstrado sua relevância ainda em 2010 e, com o passar dos anos, especialmente de 2020 até o presente momento, a procura por entender os fatores críticos para implementação de analytics em SME se intensificou.

Também é possível notar que alguns autores, como Olszak e Ziemba (2012), Willetts, Atkins e Stanier (2022), Alsyuf et al. (2022) e Mohamed e Weber (2020), buscaram entender esses fatores dentro do contexto de uma determinada localidade, assumindo que podem haver fatores de impacto específicos naquela região geográfica.

3.3.2 PA2 - Quais são os fatores críticos de sucesso identificados por esses estudos?

O quadro presente no Apêndice A lista os fatores pontuados por cada autor nos estudos analisados. Esses fatores foram categorizados de modo que forças impulsionadoras estão identificadas com [I], forças restritivas estão identificadas com [R] e os fatores sem identificação não têm a natureza do impacto definida pelo autor. Também foi adicionado um destaque em cores (verde para [I] e vermelho para [R]) para facilitar a identificação.

Nota-se que cada autor apresenta os fatores de impacto de diferentes formas: Puklavec, Oliveira e Popovič (2014) trazem os fatores sem detalhar se são forças impulsionadoras ou restritivas; Gudfinnsson e Strand (2017), Mohamed e Weber (2020), Willetts, Atkins e Stanier (2022) e Justy et al. (2023) apresentam apenas forças restritivas e Scholz et al. (2010), Olszak e Ziemba (2012), Maroufkhani et al. (2020), Alsyouf et al. (2022) e Perdana et al., (2022) apresentam ambas.

É possível observar que existe uma diferença na quantidade de fatores identificados pelos autores, variando de quatro a trinta e seis fatores. Alguns estudos, como em Justy et al. (2023), apresentam agrupamentos de diferentes fatores por proximidade de contexto, mas a tabela apresentada no apêndice A traz os fatores no nível mais granular em que são apresentados.

No total, foram extraídos 128 fatores, sem considerar a possível ocorrência de repetições. Foi realizada uma análise inicial, a fim de verificar a popularidade de um fator, ou seja, quantas vezes ele foi citado e, com isso, chegar a uma lista de fatores distintos.

O Quadro 7, abaixo, apresenta, de forma resumida, como esse agrupamento semântico foi realizado, considerando o fator genérico e quais autores citam fatores que o compõem. O detalhamento do agrupamento pode ser verificado no Apêndice B.

Quadro 7. Grupos de fatores e respectivos autores

Fator	Autores que citaram o fator									
	1	2	3	4	5	6	7	8	9	10
Qualidade de software	X	X				X			X	X
Segurança de software	X							X	X	X
Adequação de software		X	X					X		
Gerenciamento de tecnologia								X		
Cultura organizacional			X							
Orçamento/Recursos disponíveis	X	X	X		X			X		X
Entendimento dos benefícios		X	X	X	X	X	X			
Suporte para desenvolvimento	X	X	X	X		X	X	X		X
Interesse da gerência				X				X		X
Gestão de projetos de sucesso		X		X				X		X
Existência de referências		X	X							
Existência de suporte externo			X			X				
Existência de pressão externa							X	X		
Aptidão organizacional			X			X	X			
Existência de uma área específica				X						
Preparação de dados		X	X	X				X		

Fator	Autores que citaram o fator									
	1	2	3	4	5	6	7	8	9	10
Abertura para adoção de novas estruturas	X									
Observação da necessidade	X							X		X
Tamanho da empresa	X		X							
Serviço da empresa	X									
Tipo de gerência (familiar ou não)				X						
Qualificação dos funcionários	X	X		X				X	X	X
Existência de profissionais já qualificados					X			X		X
Resistência natural		X		X	X			X		X
Confiança em dados	X			X		X	X	X	X	
Crença de que o tópico é complexo	X	X				X	X			

Fonte: Elaborado pela autora, 2024

Considerando os agrupamentos demonstrados no Quadro 7, a lista de fatores distintos possui 26 itens. Embora alguns fatores ainda pudessem ser agrupados (por exemplo, fatores como "Crença de que o tópico é complicado" poderia se encaixar no conceito de "Cultura da organização" ou em "Preparo organizacional"), optou-se por deixar ainda separados os fatores que foram tratados de forma distinta pelos autores, para que pudessem ser analisados com maior detalhamento.

Com o objetivo de entender as raízes dos fatores encontrados, os 26 itens foram organizados em 3 categorias: Tecnológicos, Organizacionais e Pessoais.

- a. Tecnológicos: fatores que estão relacionados à tecnologia disponível para o processo de analytics.

- b. Organizacionais: fatores relacionados à organização da empresa, ou seja, como a empresa se prepara e atua para suportar o processo de analytics.
- c. Individuais: fatores relacionados a comportamentos e formas de ser dos envolvidos no processo de analytics.

O Quadro 8, abaixo, apresenta os fatores organizados por categoria e popularidade (quantidade de menções).

Quadro 8. Fatores por categoria e popularidade

Tecnológicos	Organizacionais	Individuais
Qualidade de software (11)	Suporte para implantação de sistema (13)	Qualificação de funcionários (12)
Segurança de software (4)	Gerenciamento do projeto de dados (10)	Confiança em dados (6)
Adequação de software (4)	Orçamento/Recursos disponíveis (9)	Profissionais já qualificados em dados (5)
Gerenciamento de tecnologia (1)	Conhecimento dos benefícios (9)	Resistência natural (5)
	Preparação de dados (8)	Crença de que o tópico é complexo (5)
	Observação de necessidade (5)	
	Interesse da gerência (3)	
	Preparo organizacional (3)	
	Existência de referências para auxílio (2)	
	Existência de suporte externo (2)	
	Existência de pressão externa (2)	
	Abertura para adotar novas estruturas (2)	
	Tamanho da empresa (2)	
	Serviço prestado pela empresa (2)	
	Tipo de gerência da empresa (1)	
	Cultura organizacional (1)	
	Existência de área dedicada (1)	
Total: 20	Total: 75	Total: 33

Fonte: Elaborado pela autora, 2024

Observa-se uma prevalência de fatores de origem organizacional, com 75 menções. A tabela 8 evidencia os fatores que apresentaram cinco ou mais ocorrências, o que significa um total de 98 ocorrências distribuídas em 12 fatores principais.

3.3.3 PA3: Quais são as características das organizações avaliadas?

No total de 931 empresas avaliadas pelos 10 estudos, são observadas algumas características que podem auxiliar na contextualização das descobertas do mapeamento.

No entanto, é importante notar que os estudos devem ser analisados de forma individual, uma vez que nem todas as características foram informadas para todas as empresas e, portanto, não podem gerar estatísticas ou serem generalizadas.

Quadro 9. Características das empresas entrevistadas por estudo

Estudo	Número de empresas	Áreas de atuação	Cargo dos respondentes	Localização
1	214	Serviço, manufatura, IT, outros	Diretores ou executivos sênior	Saxony, Alemanha
2	20	Comércio, serviços e consultoria	Gerentes, analistas de negócios e profissionais de TI	Alta Silésia, Polônia
3	10	Serviço, vendas, publicidade, educação, engenharia e produção, IT	Gerentes, diretores, consultores de desenvolvimento	União Europeia
4	4	Manufatura	CEO, gerentes, líderes de grupo e gerentes de projeto	Global
5	53	Consultoria, IT, educação, marketing, manufatura, turismo, alimentação e outros	-	West Midlands, Reino Unido
6	171	Manufatura	CEO, diretores ou executivos	Tehran, Iran

			sênior e proprietários	
7	116	Manufatura	Gerentes e proprietários	Jordânia
8	137	Manufatura, IT, serviços empresariais e financeiros, saúde e outros	Diretores, gerentes proprietários e profissionais de TI	Reino Unido e Irlanda
9	174	Serviços, varejo, manufatura, IT, finanças	Unidade de negócio ou gerente de departamento ou gerente executivo ou vice presidente	Singapura
10	32	Manufatura e agricultura	Gerentes	França

Fonte: Elaborado pela autora, 2024

O único critério que relacionava as organizações no MSL era a constatação de que se tratava de uma organização de porte Micro, Pequeno ou Médio. No entanto, apesar de essa constatação ter sido feita em 100% das vezes através da descrição textual, nem todos os estudos informaram qual regra de classificação utilizavam para definir o porte da organização.

Portanto, não é possível afirmar a relação entre as empresas pela quantidade de funcionários, embora a classificação geral e exclusão de empresas de grande porte minimizem esse prejuízo.

3.3.4 PA4: Quais as metodologias usadas nos processos de Data Analytics avaliados?

Nenhum dos artigos estudados informa se foi aplicada uma metodologia para implementação do processo de Data Analytics. Dessa forma, não se sabe se o uso de uma metodologia (ou a falta dela) teve impacto nos resultados dos processos de Data Analytics.

3.3.5 PA5: Quais resultados são categorizados como de sucesso ou de fracasso?

De forma similar à questão PA4, não se tem registros dos resultados ou indicadores que evidenciam sucesso ou fracasso na implementação de Data Analytics.

Isso também foi observado por Llave (2017), em seu estudo '*Business Intelligence and Analytics in Small and Medium-sized Enterprises: A Systematic Literature Review*':

Muitos frameworks e modelos emergiram para servir de Guia na identificação de fatores que dão suporte a uma implementação de BI&A de sucesso. No entanto, não há uma definição de sucesso clara. Além disso, o modo como as soluções de BI&A são usadas não foi extensivamente estudado em relação aos benefícios percebidos. Por fim, a necessidade de entregar um retorno sobre investimento (ROI) significativo não foi completamente discutido [...] (Llave, 2017, p. 202).

Considera-se, portanto, que os fatores foram elencados, nas pesquisas, com base na percepção individual de cada participante em relação ao quão bem ou mal o processo de Data Analytics está implementado em sua empresa.

3.4 DISCUSSÃO

A partir dos resultados encontrados com os estudos, percebe-se que não há uma classificação padrão ou nomenclatura padronizada em relação aos fatores. Além disso, os próprios fatores encontrados por cada estudo diferem, sendo que alguns são mais detalhados, como Olszak e Ziembra, 2012 e Justy et al., 2023, e outros mais generalistas, como Maroufkhani et al., 2020, Alsyouf et al., 2022 e Perdana et al., 2022.

A partir da classificação feita pela autora, buscou-se agrupar os fatores encontrados para entender se os principais fatores percebidos estão relacionados às tecnologias disponíveis, à preparação das empresas ou ao aprimoramento individual dos funcionários. Com os resultados dessa classificação, identificou-se uma grande quantidade de diferentes fatores no âmbito organizacional.

A literatura apresenta alguns fatores que podem ser discutidos em relação à origem da sua problemática. A exemplo, quando se fala em "Falta de treinamento do usuário e suporte" (Olszak e Ziembra, 2012) ou "Pessoas utilizando a solução de BI

não são suficientemente qualificadas” (Scholz et al., 2010) não fica claro se os fatores fazem referência a uma responsabilidade da empresa em prover esse treinamento e apoio ao desenvolvimento da habilidade no seu funcionário, ou se o fator se refere à falta de capacitação individual e de disponibilidade de profissionais capacitados no mercado.

Igualmente, fatores como ‘Complexidade’ e ‘Preocupações éticas’ não são auto explicativos. O primeiro pode ser gerado pela falta de habilidades da equipe responsável ou a utilização de um software complexo e o segundo, por exemplo, pode ser proveniente de uma falta de segurança no software ou na garantia de acesso para consumo de dados.

Outros fatores são bastante abrangentes e simplesmente mencioná-los pode não ser suficiente para explicar exatamente o que acontece nas empresas a ponto de gerar impacto no processo de Data Analytics. Como exemplo, o estudo de Puklavec, Oliveira e Popovič (2014) apresenta o fator “Cultura organizacional”, mas podem ser encontrados diversos outros fatores, na literatura, relacionados com a cultura organizacional, como ‘Falta de confiança’, em Justy et al. (2023); ‘Reação flexível às mudanças de mercado’, em Scholz et al. (2010) e ‘Suporte da gerência’, em Olszak e Ziemba (2012), Puklavec, Oliveira e Popovič (2014), Maroufkhani et al. (2020), Alsyouf et al. (2022), Willetts, Atkins e Stanier(2022) e Justy et al.(2023).

Da mesma forma, o fator “Preparação organizacional”, visto em Gudfinnsson e Strand (2017) e Maroufkhani et al(2020), pode ser detalhado em outros fatores menores, como os que compõem ‘Preparação de dados’ ou ‘Empregados qualificados’, para explicar o que caracteriza uma empresa que está preparada para trabalhar com o processo de Data Analytics.

Apenas um artigo (Justy et al., 2023) apresenta um componente de intensidade na influência dos fatores elencados. Os fatores são divididos entre os que possuem relação de influência e os que possuem forte relação de influência. Conforme o artigo, barreiras internas têm um impacto mais negativo na adoção de Data Analytics do que barreiras externas.

Os estudos de Maroufkhani et al. (2020) e Alsyouf et al. (2022) apontaram fatores irrelevantes, de acordo com suas pesquisas. “Compatibilidade”, por exemplo, não foi validado por Alsyouf et al. (2022) como um fator importante, mas aparece

para Olszak e Ziemba (2012). Da mesma forma, 'Pressão competitiva' se mostrou um fator irrelevante.

Em modo geral, os fatores relacionados à qualidade de software, à existência de suporte para implantação do sistema de analytics, à qualificação de funcionários e à qualidade do gerenciamento do projeto de dados demonstraram ser os mais relevantes. Por outro lado, o tipo de serviço prestado pela empresa, o tipo de gerência (empresa familiar ou não) e a existência de uma área dedicada a dados foram os fatores menos citados.

Esperava-se encontrar informações de contexto referentes às metodologias utilizadas para implementação do processo de Data Analytics, de modo a poder validar sua influência nos resultados do processo. No entanto, nenhuma citação foi encontrada. Esse fator parece ser relevante, pois pode indicar o desconhecimento, por parte das organizações, de metodologias já bem estabelecidas na literatura, como, por exemplo, CRISP-DM e outras apresentadas por Saltz e Krasteva (2022).

Similarmente, o fato de não haver registro de quais evidências demonstram sucesso para confirmar os fatores levantados prejudica a validação e reprodução dos experimentos, pois não há um conceito do que seria o sucesso ou fracasso produzidos pelos fatores. Este não demonstrou ser um fator de estudo, uma vez que não foram encontrados artigos que citassem especificamente métricas de monitoramento do sucesso das iniciativas de Data Analytics.

3.5 AMEAÇAS À VALIDADE

Como em outras revisões sistemáticas da literatura, o mapeamento apresentado neste trabalho possui limitações que podem influenciar os resultados apresentados.

Uma limitação que potencialmente prejudica a generalização das descobertas é a diferença de contexto geográfico e funcional das organizações estudadas. A depender da cultura de uma certa localidade, os efeitos de um fator podem ser mais ou menos relevantes e, ao compará-los todos sem essa distinção, se entende que todos têm um mesmo 'peso'.

Outra ameaça é o fato de os fatores terem sido agrupados para análise com base na interpretação da autora, possivelmente causando perda de alguma informação ou má interpretação. Para minimizar esse risco e permitir uma reavaliação dos leitores interessados, todo o processo de agrupamento foi apresentado ao longo do texto e em anexo.

Além disso, por se tratar de uma investigação manual, sempre há a possibilidade de não cobrir todos os estudos existentes nesse contexto, apesar de que o protocolo definido e seguido deve minimizar esse risco.

4 ESTADO DA PRÁTICA

Para entender como as empresas de desenvolvimento de software, no Brasil, experienciam o processo de Data Analytics e a fim de corroborar com as descobertas do estado da arte, foi realizado um *survey* com profissionais de pequenas e médias empresas de desenvolvimento de software, seguindo atividades definidas por Pinsonneault e Kraemer (1993). O resultado do *survey* deve identificar os fatores mais relevantes para o contexto de interesse que serão posteriormente utilizados como elementos chave do guia.

Neste capítulo, é apresentado o levantamento do estado da prática, realizado desde a sua concepção e planejamento até a análise dos resultados obtidos.

4.1 PLANEJAMENTO

O alvo da pesquisa são pessoas que trabalham na área de dados de pequenas e médias empresas de software do Brasil, conforme definido no capítulo 2.5. O *survey* tem como objetivo responder às perguntas de pesquisa apresentadas no Quadro 10.

Quadro 10. Perguntas de pesquisa do *survey*

Pergunta	Descrição
PP1	Quantas organizações aplicam Data Analytics no seu dia a dia?
PP2	Quais fatores são considerados críticos para o sucesso do processo de Data Analytics?
PP3	Quais as características das organizações consultadas?
PP4	Quais as metodologias/abordagens usadas pelas empresas para implementar Data Analytics?
PP5	Quais etapas do processo de Data Analytics são efetivamente executadas pelas empresas?
PP6	Quais os resultados observados com o processo de Data Analytics?

Fonte: Elaborado pela autora, 2024

4.1.1 Amostragem

Para determinar o tamanho da amostra, foi escolhido o método de amostragem probabilística (Berndt, 2020), utilizando o relatório da Associação Brasileira das Empresas de Software (2021) para definição da população de interesse.

A fim de filtrar os participantes, os critérios de inclusão na amostra são:

- a) Ter atuação direta ou indireta na análise de dados da empresa pela qual responde;
- b) Ser de uma micro ou pequena empresa de software brasileira.

Considerando os 93,6% das empresas de desenvolvimento e produção de software que correspondem às micro e pequenas empresas (seção 2.5), identificou-se que o tamanho da população de interesse nesse estudo é de aproximadamente 7.936 empresas.

A partir disso, para calcular o tamanho de uma amostra estatisticamente significativa, foi definida uma margem de erro de 5% e um grau de confiança de 90%. Assim, determinou-se que uma amostra estatisticamente significativa, para esse caso, deve ser composta de 264 respondentes.

O convite para participação na pesquisa foi feito de forma aleatória, por meio da plataforma LinkedIn⁶. Foram filtrados os profissionais de empresas de software brasileiras e, para cada profissional encontrado na busca, foi enviada uma solicitação de conexão e mensagem direta com o convite para a pesquisa.

No total, 573 convites de conexão foram aceitos, de modo que se sabe que pelo menos 573 mensagens foram recebidas. Não há controle sobre as mensagens enviadas com convites não aceitos. Considerando as limitações de tempo da pesquisa, não foi possível alcançar o tamanho da amostra prevista. Ao total, foram coletadas 133 respostas, resultando em uma taxa de resposta de 23%.

⁶ <https://www.linkedin.com/>

4.2 INSTRUMENTO DE COLETA DE DADOS

Para responder às perguntas de pesquisa, um questionário com 9 perguntas fechadas e 1 pergunta aberta foi elaborado como instrumento de coleta de dados, utilizando o Google Forms⁷ como ferramenta (Apêndice C).

O Quadro 11 apresenta as perguntas do questionário derivadas das perguntas de pesquisa.

Quadro 11. Perguntas do questionário

Pergunta de pesquisa	Pergunta do questionário
PP1 e PP5	Quais fases do processo de Data Analytics são efetivamente executadas pela sua organização?
PP2	Marque todos os fatores que, de acordo com a sua experiência na organização, você considera críticos para o sucesso do processo de Data Analytics
PP3	Qual a quantidade de funcionários da organização?
PP3	Qual o segmento de atuação da organização?
PP3	Em que região do Brasil está situada a organização?
PP3	Qual seu cargo na organização?
PP3	Qual é a sua experiência com Data Analytics?
PP3	Quantos anos de atividade a organização tem?
PP3	A organização possui pelo menos um responsável por análise de dados?
PP4	Qual metodologia tem sido utilizada pela organização para implementar o processo de Data Analytics?
PP6	Quais impactos (positivos ou negativos) foram observados na organização com a implementação do processo de Data Analytics?

Fonte: Elaborado pela autora, 2024

⁷ <https://www.google.com/forms/about/>

O questionário foi avaliado em relação a sua qualidade e confiabilidade por meio de uma aplicação piloto, com dez profissionais de empresas de desenvolvimento de software, selecionados por proximidade.

Os dez convidados responderam ao questionário e tiveram suas impressões coletadas por meio de entrevista. Foram identificadas algumas melhorias em relação à escrita e ordenação das opções de resposta do formulário.

Por fim, o coeficiente alfa de Cronbach (Cronbach, 1951) foi calculado para avaliar a confiabilidade do questionário. O cálculo representa a correlação média entre as perguntas, avaliando a consistência das respostas a partir da variância dos itens individuais e da variância da soma dos itens de cada respondente. O valor do alfa de Cronbach obtido para as 10 respostas foi 0.91, o que, segundo Nunnally (1978), indica um excelente nível de confiabilidade do instrumento de coleta de dados.

4.3 DISTRIBUIÇÃO DO QUESTIONÁRIO

O objetivo da pesquisa foi apresentado aos participantes na mensagem de convite para participação. Além disso, na página inicial do formulário elaborado na ferramenta Google Forms, foram apresentados os seus direitos e solicitado o seu consentimento por meio de um Termo de Consentimento Livre e Esclarecido (TCLE).

Todos os dados foram coletados sem identificação do participante, para assegurar a confidencialidade e a privacidade. Os participantes responderam ao questionário de forma remota, por meio da plataforma Google Forms, pela qual tiveram acesso ao TCLE de forma eletrônica. O questionário pôde ser preenchido de forma assíncrona, sem a necessidade do acompanhamento de algum pesquisador.

Os dados brutos coletados estão disponíveis em *Marques, Livia (2024), "Impressões de profissionais de software sobre Fatores Críticos de Sucesso para Data Analytics.", Mendeley Data, V1, doi: 10.17632/8y5wnspvx5.1.*

4.4 ANÁLISE DOS DADOS

A partir dos dados coletados no questionário de levantamento do estado da prática, foi possível fazer algumas análises e descobertas. Após coleta dos dados via Google Forms, foi feita uma exportação para uma planilha do Google Sheets, para tratamento e análise facilitada. Por fim, os dados foram enviados ao Looker Studio⁸ para geração dos gráficos e relatórios.

Nesta seção, são apresentadas as principais observações e análises elaboradas a partir das respostas e considerando cada pergunta de pesquisa com que se relacionam.

4.4.1 PP1 - Quantas organizações aplicam Data Analytics no seu dia a dia?

Para responder a PP1, buscou-se entender a aplicação de Data Analytics por etapas. Uma vez que o processo em questão é composto de múltiplas etapas, é válido assumir que uma organização pode executar apenas um subconjunto de etapas, ou até mesmo executar algumas delas parcialmente.

A partir das respostas da pergunta “Quais fases do processo de Data Analytics são efetivamente executadas pela sua organização?”, foi possível identificar que apenas 4 organizações consideram que não executam nenhuma parte do processo, enquanto 45 declaram que executam o processo completo. Isso leva a 84 empresas restantes que, embora sigam alguma etapa do processo de dados, realizam isso sem compromisso com a totalidade. Dessa forma, são 129 as empresas que aplicam pelo menos uma etapa de Data Analytics.

4.4.2 PP2 - Quais fatores são considerados críticos para o sucesso do processo de Data Analytics?

Para identificar os fatores considerados críticos pelos respondentes, foi utilizada a lista de fatores identificados no mapeamento do estado da arte (seção 3.3.2). Com isso, esperou-se verificar a concordância com os fatores levantados pelo estado da arte, ainda permitindo que novos fossem adicionados por meio da opção “Outros”.

⁸ <https://lookerstudio.google.com/>

A Figura 5, apresentada a seguir, indica a quantidade de ocorrências de cada fator, ou seja, quantas vezes um fator foi selecionado como relevante. Os fatores estão dispostos em ordem de relevância por número de ocorrências (do mais frequente ao menos frequente) e a cor mais forte indica os fatores convergidos por mais de 50% dos respondentes. Deve-se considerar que um respondente poderia selecionar mais de um item da lista.

Figura 5. Fatores considerados críticos por número de ocorrências

Um bom gerenciamento do processo de dados 95	Uma boa etapa de preparação de dados 94	Clareza sobre a necessidade do processo 87	Existência de profissionais qualificados (com experiência em dados) 83	Suporte organizacional (da gerência) para implementação 82
Acesso à software de qualidade 64	Orçamento adequado disponível 60	Impactados conhecerem os benefícios 57	Equipe confiante no resultado apresentados por dados 55	Suporte técnico para implementação 45
Estrutura para qualificação dos funcionários 44	Inexistência de resistência psicológica quanto ao uso de dados 28	Inexistência da crença que o tópico é complexo demais 22	Outros 5	"outros", "evangelização e localização de embaixadores de data quality", "clareza do problema a ser resolvido", "saber transformar dados em informação útil e alinhada com a estratégia", "dados de boa qualidade"

Fonte: Elaborado pela autora, 2024

Os resultados dessa pergunta evidenciaram os fatores mais populares entre os respondentes. Os fatores que mais de 50% da amostra considerou como mais importante foram cinco: “Um bom gerenciamento do processo de dados”, “Uma boa etapa de preparação de dados”, “Clareza sobre a necessidade do processo”, “Existência de profissionais qualificados (com experiência em dados)”, “Suporte organizacional (da gerência) para implementação”.

O complemento das respostas na categoria “Outros”, embora escritas de diferentes formas, podem ser conectadas, em sua maioria, com algum dos 5 fatores principais, reforçando sua relevância.

4.4.3 PP3 - Quais as características das organizações consultadas?

A fim de contextualizar as análises apresentadas, algumas perguntas foram feitas para caracterizar as organizações participantes da pesquisa.

Inicialmente, foi perguntado qual a experiência do respondente com Data Analytics, em que 57% revelou ter entre um e cinco anos de experiência. Outros

22,6% declararam um período entre 5 e 10 anos; 12% com menos de 1 ano e 8,3% com mais de 10 anos. Observa-se, então, que na amostra estudada, prevalecem profissionais com até 10 anos de experiência.

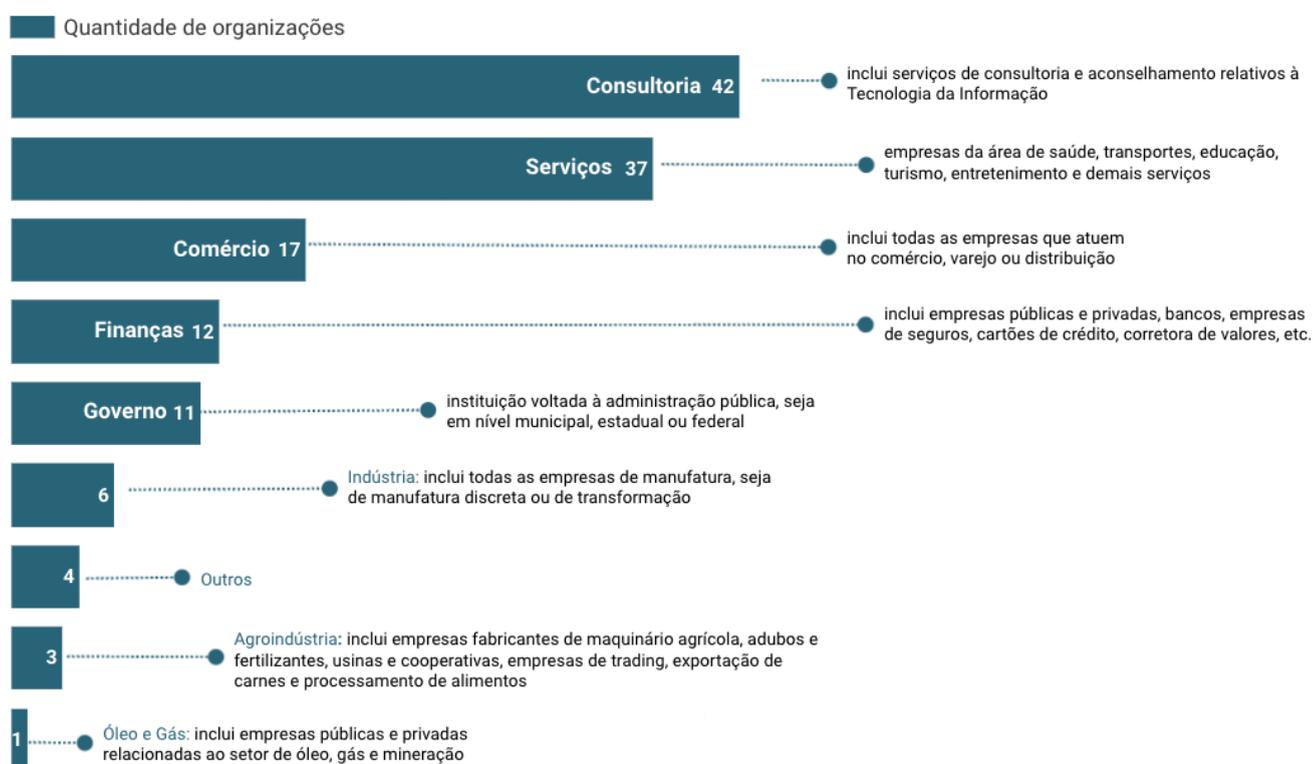
Identificou-se que, no que se refere aos cargos desempenhados, 48,8% dos participantes ocupam cargos relacionados a dados (entre cargos de liderança, gerência e engenharia), enquanto 21,8% ocupam cargos de liderança na organização. Os demais se dividem em cargos de gerência de projetos ou desenvolvimento de software (18,05%) e outros (11,28%).

Em relação à organização, buscou-se identificar a qual região brasileira ela pertence. As respostas indicam que 54% dos respondentes estão na região sul do país e 35% estão na região sudeste. Os outros 11% estão distribuídos entre regiões norte, nordeste e centro-oeste.

Também buscou-se conhecer quantos anos de atividade cada organização possuía no momento em que implementaram o processo de Data Analytics. Neste quesito, 45% das organizações pesquisadas possuíam mais de 10 anos de atividades, enquanto 23,3% possuíam entre 5 e 10 anos, 21,8% entre 2 e 5 anos e 9,8% entre 0 e 2 anos.

Em relação ao segmento aos quais as organizações pesquisadas pertencem, elas declararam que, dentro da categoria software, pertencem aos segmentos apresentados na Figura 6.

Figura 6. Prevalência de segmento das organizações

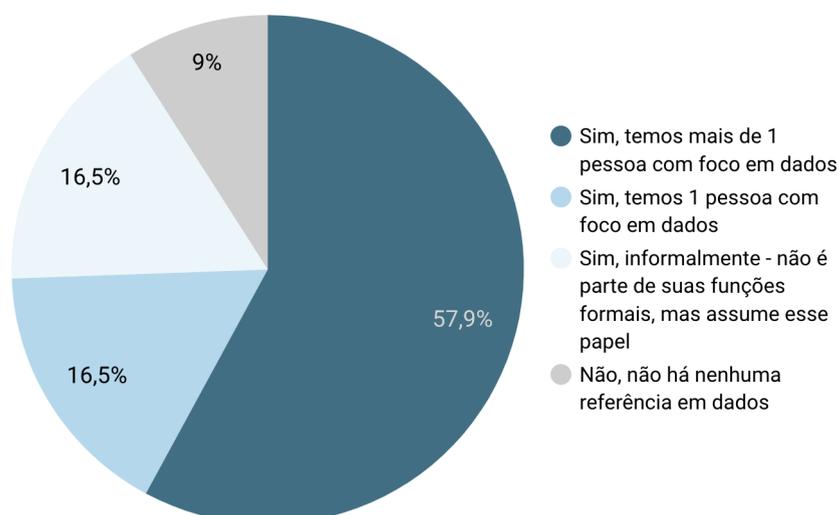


Fonte: Elaborado pela autora, 2024

A maioria das organizações lida com software para áreas de consultoria, ou serviços.

Para conhecer a estrutura organizacional em relação à área de dados, foi perguntado se a organização possui pelo menos um responsável por análise de dados. A Figura 7 evidencia que a maioria possui mais de um profissional que se dedica a essa área.

Figura 7. Quantidade de profissionais dedicados a dados



Fonte: Elaborado pela autora, 2024.

Observa-se que é baixo o percentual de organizações que não possuem profissionais atuando em atividades relacionadas a dados (9%), o que pode indicar o interesse pela área.

4.4.4 PP4 - Quais as metodologias/abordagens usadas pelas empresas para implementar Data Analytics?

A fim de conhecer as abordagens utilizadas pelas empresas no processo de dados, o *Survey* apresentava a pergunta "Qual metodologia tem sido utilizada pela organização para implementar o processo de Data Analytics?".

As respostas evidenciaram que a maioria das organizações (77,4%) não utilizam uma metodologia específica como base. Já as demais, em percentuais menores, declararam que utilizam CRISP-DM (9,8%), KDD (6%) e SEMMA (2,3%). Aproximadamente 3% das respostas indicaram que houve uma mescla entre metodologias conhecidas e desenvolvimento de metodologia própria.

Nota-se, então, que a maioria das organizações não vê um framework já estabelecido como solução completa para a implementação e gerenciamento do processo de dados.

4.4.5 PP5 - Quais etapas do processo de Data Analytics são efetivamente executadas pelas empresas?

Para conhecer quais fases do processo de Data Analytics são efetivamente executadas pelas organizações, utilizamos a mesma pergunta analisada na seção 4.4.1: “Quais fases do processo de Data Analytics são efetivamente executadas pela sua organização?”. A Figura 8 apresenta os resultados encontrados com a resposta.

Figura 8. Análise das fases efetivamente executadas pelas organizações



Fonte: Elaborado pela autora, 2024

Foi possível observar que a maioria das organizações executa pelo menos uma etapa do processo, sendo a etapa de 'Entendimento de negócio' a executada com mais frequência (89 dos 133 respondentes).

Em seguida, a "preparação dos dados", etapa que ganhou bastante destaque como um fator relevante para o processo de análise de dados, é a segunda mais executada. Mesmo assim, 32% das organizações declaram executar apenas parcialmente e quase 10% declaram não executar.

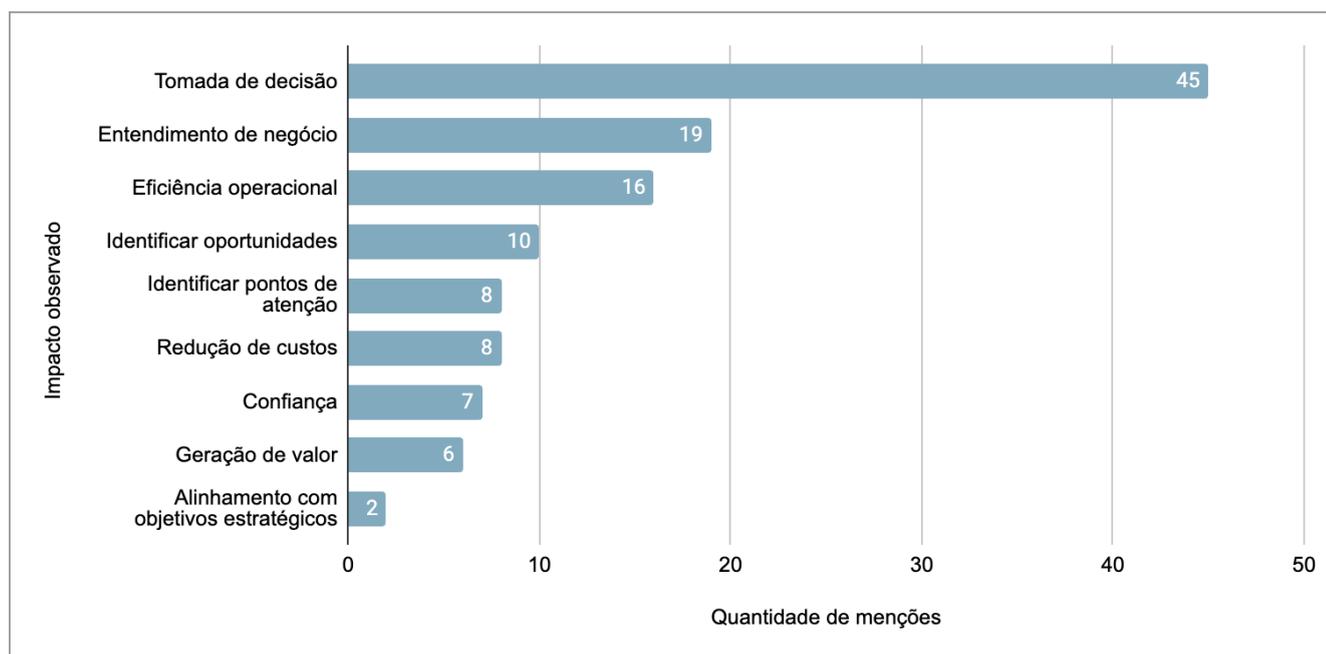
Além disso, a etapa de Avaliação dos dados modelados é a que mais comumente não se executa (15,8%).

4.4.6 PP6 - Quais os resultados observados com o processo de Data Analytics?

A pergunta "Quais impactos (positivos ou negativos) foram observados na organização com a implementação do processo de Data Analytics?" , por ser uma pergunta aberta, permitia que cada resposta pudesse conter uma, mais de uma ou até nenhuma observação. Para análise, foram desconsideradas respostas vazias ou N/A (15), respostas com textos que não agregam informações referentes à pergunta (7) e respostas que não remetem à experiência do respondente e sim à benefícios genéricos que se assume ser verdade (3). Os impactos mencionados em cada resposta foram avaliados individualmente e agrupados conforme sua natureza. A seleção e agrupamento das respostas pode ser verificada no Apêndice D.

A Figura 9 apresenta a contagem de menções a cada tipo de impacto observado.

Figura 9. Menções aos impactos observados



Fonte: Elaborado pela autora, 2024

O impacto mais observado é na melhoria do processo de tomada de decisão, mencionado 45 vezes. Os comentários incluíam aspectos como direcionamento fundamentado para evolução dos produtos, possibilidade de avaliar com base em dados históricos, entendimento das áreas em que não se deve aplicar esforços, maior velocidade e precisão, além de transparência para o processo.

Observa-se também uma grande ocorrência da percepção de impacto no entendimento do negócio. Além da possibilidade de ver com mais clareza e confiança o estado atual da organização, o uso de dados proporcionou, a partir do entendimento de negócio, mais autonomia para os colaboradores, possibilidade de avaliação de iniciativas e mais clareza sobre o perfil de clientes.

Outros impactos positivos mencionados foram a melhoria na eficiência operacional, possibilidade de identificar pontos de atenção e oportunidades, redução de custos, maior confiança estabelecida, geração de valor de modo geral e maior alinhamento com os objetivos estratégicos.

Em relação aos impactos negativos observados, pode-se considerar que as respostas apresentam aspectos negativos observados a partir da má implementação do processo ou da falta dele. Algumas respostas, que representam também outras de mesma natureza, são apresentadas a seguir:

- a) *“Aumento da dependência dos dados para tomada de decisões, implicando em excesso de análises. Com muitas análises para fazer, o time de Data analytics não conseguia se dedicar para tradução e comunicação dos dados, reduzindo o impacto das conclusões obtidas”*
- b) *“Como ponto negativo, a falta de processos rotineiros e de entendimento da importância de se seguir este trabalho fez com que a análise de dados não fizesse parte do dia a dia das áreas.”*
- c) *“O custo de implementação em casos nos quais a empresa não está preparada para usufruir da forma correta dos dados.”*
- d) *“As reuniões sem o DA parece não ter eficácia.”*
- e) *“Questões relacionadas à privacidade de dados.”*
- f) *“O desuso de Dashboards ou do controle de dados, visto que não houve um correto alinhamento entre a expectativa do cliente e o produto gerado.”*
- g) *“Falta de conhecimento técnico da equipe e os vieses de gestores.”*

Os impactos negativos identificados são, em sua maioria, relacionados ao mau gerenciamento do processo de dados, falta de alinhamento da equipe e incentivo dos gestores, falta de profissionais qualificados e custo de implementação.

4.5 DISCUSSÃO

As análises das respostas do *survey* demonstram alguns aspectos em que o Estado da Prática converge em relação ao que foi observado no Estado da Arte, apresentado no capítulo 3.

A principal observação aqui, considerando o objetivo deste trabalho, é a identificação dos fatores considerados mais relevantes para uma implementação satisfatória de Data Analytics. Os fatores relevantes do Estado da Arte que também foram considerados relevantes no Estado da Prática são:

- a) Um bom gerenciamento do processo de dados;
- b) Uma boa etapa de preparação de dados;
- c) Clareza sobre a necessidade do processo;
- d) Existência de profissionais qualificados (com experiência) em dados;
- e) Suporte organizacional (da gerência) para implementação.

Esses fatores também foram destacados a partir dos comentários sobre os impactos negativos do processo de Data Analytics. Ao discorrer sobre os problemas na implementação e no sucesso da prática, pode-se perceber que os fatores citados anteriormente, quando em falta, aparecem com frequência como limitadores da

eficiência do processo: falta de um bom gerenciamento do processo de dados, falta de apoio da gerência e baixa qualificação de profissionais.

Os fatores mais citados no Estado da Arte - 'Qualidade de software' e 'Suporte técnico para implantação de sistema' - não estiveram entre os mais votados (48% e 34% de consideração, respectivamente). Isso pode indicar que a natureza dos fatores críticos no contexto de micro e pequenas empresas brasileiras de software não é tão técnica quanto organizacional.

O fator 'Conhecimento dos benefícios', apesar de não ter sido selecionado por mais de 50% dos participantes, também foi um fator relevante, com quase 43% de considerações, contribuindo para o que foi mostrado no Estado da Arte (em que estava entre os 5 principais fatores).

De modo geral, os resultados do mapeamento dos fatores relevantes no Estado da Arte e no Estado da Prática convergiram. Especialmente se considerarmos que os fatores são provenientes de agrupamentos - que podem carregar aspectos relacionados paralelos - pode-se dizer que a maioria dos 5 fatores identificados pelo estado da prática estão entre os 5 principais fatores mapeados pelo estado da arte. Os dois fatores que não correspondem ('Uma boa etapa de preparação de dados' e 'Clareza sobre a necessidade do processo'), podem estar relacionados à características específicas do contexto investigado no Survey, em contraste com o contexto mais genérico encontrado no MSL.

Outro aspecto indicado pelos resultados do survey é a predominância do uso de uma abordagem *ad-hoc* quanto às metodologias/abordagens utilizadas para implementação do processo da Data Analytics. Esse fato segue a tendência identificada nos estudos do Estado da Arte e indica que o processo de Data Analytics ainda não foi contemplado com uma abordagem que o atenda completamente - e que, talvez, sua natureza não o permita ser. Isso evidencia a necessidade da utilização de abordagens flexíveis no contexto de dados.

A falta de uma método para gerenciamento desses projetos e o seu aparente caráter dinâmico também leva a uma variação nas etapas do processo que as organizações se comprometem a realizar. Apenas 4 das 133 organizações consultadas realizam o processo completo (considerado pelo CRISP-DM). De modo geral, as organizações fazem adaptações de implementação que se encaixam num contexto já existente.

4.6 AMEAÇAS À VALIDADE

Os resultados do *Survey* devem ser analisados levando em consideração alguns fatores que podem ameaçar a sua validade.

Em relação à generalização dos resultados para todas as empresas brasileiras de software, deve-se considerar que, apesar do estudo ter tido o objetivo de alcançar todas as regiões do país - e, portanto, não ter adicionado nenhuma restrição nesse sentido - as respostas obtidas estão majoritariamente concentradas nas regiões Sul e Sudeste. Isso pode levar a uma má representatividade das demais regiões.

Ao coletar os fatores considerados mais importantes para o processo de dados, ter disponibilizado a lista com alguns fatores pode ter adicionado um viés à resposta dos participantes, o que não aconteceria se a questão fosse aberta. A opção 'Outros' busca minimizar o impacto deste risco.

A análise do *Survey*, especialmente na questão aberta e outras que em que foi realizado agrupamento de resposta, foi sujeita a interpretações subjetivas da autora. Para suavizar o risco de interpretações incorretas, todos os agrupamentos realizados foram disponibilizados no trabalho a fim de permitir que o leitor faça suas considerações.

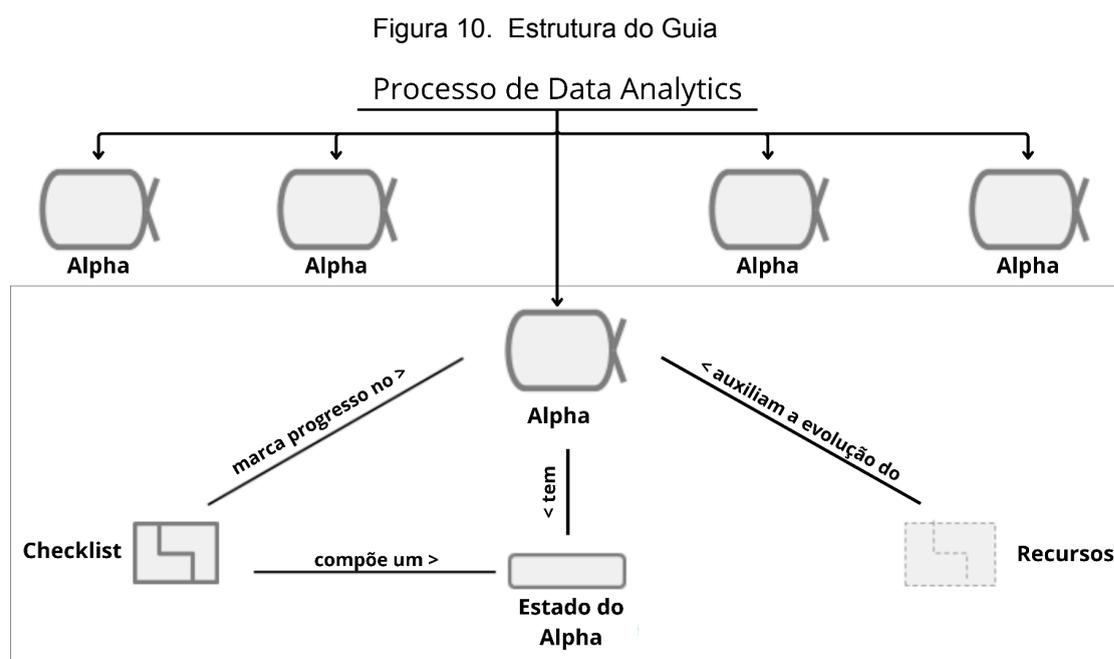
5. GUIA PARA IMPLEMENTAÇÃO DO PROCESSO DE DATA ANALYTICS

Este capítulo apresenta o Guia desenvolvido para a Implementação de Data Analytics em micro e pequenas empresas de software. Inicialmente, a estrutura do Guia é apresentada, bem como sua construção com elementos da abordagem Essence. Em seguida, o conteúdo do Guia é descrito, citando as referências teóricas que deram embasamento para cada parte. Depois, é apresentada a ferramenta de autoavaliação desenvolvida para organizações que desejam seguir as práticas sugeridas no Guia. Por fim, uma seção com os comentários dos avaliadores do Guia e as ameaças à sua validade que foram identificadas.

5.1 Estrutura do Guia e da Ferramenta de Autoavaliação

A construção do Guia é baseada na abordagem Essence (OMG, 2018), apresentada no capítulo 2.6. O Essence foi escolhido por oferecer uma notação atualizada, que abraça o dinamismo encontrado nos contextos de processos de Data Analytics e, não menos importante, de micro e pequenas empresas.

O Guia é composto por Alphas, Estados dos Alphas, Checklists que marcam a evolução nos estados de um Alpha e Recursos que contribuem para tal evolução. A Figura 10, abaixo, descreve a relação entre cada um desses elementos para um determinado Alpha, explicando a estrutura do Guia.



Fonte: Elaborado pela autora, 2024

De forma resumida, cada um dos cinco fatores mais relevantes identificados no levantamento do estado da prática (seção 4.4.2) é definido como um *Alpha* do Essence, ou seja, um fator essencial cuja evolução deve ser monitorada pela organização. Os estados de cada alpha foram extraídos da literatura, bem como práticas que podem guiar a evolução desses estados (identificadas nos checklists). Também são adicionados Recursos para servir de apoio à execução das práticas identificadas na literatura.

Com essa estrutura base e o entendimento de como cada elemento se relaciona e orienta as organizações, são criados cards do Essence com mais detalhes sobre cada elemento. Os cards serão apresentados na seção 5.2, com o conteúdo do guia.

A partir do guia, foi desenvolvida uma ferramenta de autoavaliação para que as organizações possam, dinamicamente, identificar seu status atual em relação aos 5 fatores críticos e monitorar seu progresso.

A ferramenta é apresentada como um template na ferramenta Google Spreadsheets⁹ para que as organizações tenham sua própria cópia. A Figura 11, abaixo, apresenta a estrutura dessa ferramenta.

Figura 11. Estrutura da ferramenta de avaliação

GUIA PARA IMPLEMENTAÇÃO DO PROCESSO DE DATA ANALYTICS

Este guia foi desenvolvido para apoiar micro e pequenas empresas brasileiras de software na **identificação e aprimoramento das áreas essenciais para o sucesso na implementação e uso de Data Analytics**. Ele é baseado no conceito de **Alphas**, proveniente da metodologia Essence. Alphas representam aspectos críticos que precisam ser trabalhados para alcançar um objetivo. Cada Alpha é composto por estados de evolução, que indicam o progresso de sua organização em relação a esses aspectos.

COMO UTILIZAR O GUIA:

- Identifique o estado atual de sua organização:** Para cada Alpha, revise os itens listados e marque aqueles que sua organização já realiza com sucesso. Os estados variam de 1 a 5, sendo 1 o menos evoluído e 5 o mais avançado.
- Visualize seu progresso:** Após preencher os checkboxes, utilize o gráfico gerado para obter uma visão clara das áreas mais desenvolvidas (próximas de 5) e daquelas que necessitam de maior atenção (próximas de 0).
- Planeje as próximas ações:** Com base nos estados identificados, verifique o que seriam os próximos passos. Além disso, explore os **recursos** recomendados para obter materiais de apoio e implementar as práticas necessárias.

Alpha 1	Alpha 2	Alpha 3	Alpha 4	Alpha 5
Estado 1 do Alpha				
<input checked="" type="checkbox"/> Item 1.1				
<input checked="" type="checkbox"/> Item 1.2				
<input checked="" type="checkbox"/> Item 1.3				
Estado 2 do Alpha				
<input checked="" type="checkbox"/> Item 2.1	<input checked="" type="checkbox"/> Item 2.1	<input type="checkbox"/> Item 2.1	<input checked="" type="checkbox"/> Item 2.1	<input checked="" type="checkbox"/> Item 2.1
<input checked="" type="checkbox"/> Item 2.2	<input checked="" type="checkbox"/> Item 2.2	<input type="checkbox"/> Item 2.2	<input checked="" type="checkbox"/> Item 2.2	<input checked="" type="checkbox"/> Item 2.2
<input checked="" type="checkbox"/> Item 2.3	<input checked="" type="checkbox"/> Item 2.3	<input type="checkbox"/> Item 2.3	<input checked="" type="checkbox"/> Item 2.3	<input checked="" type="checkbox"/> Item 2.3
Estado 3 do Alpha				
<input checked="" type="checkbox"/> Item 3.1	<input checked="" type="checkbox"/> Item 3.1	<input type="checkbox"/> Item 3.1	<input checked="" type="checkbox"/> Item 3.1	<input type="checkbox"/> Item 3.1
<input checked="" type="checkbox"/> Item 3.2	<input checked="" type="checkbox"/> Item 3.2	<input type="checkbox"/> Item 3.2	<input checked="" type="checkbox"/> Item 3.2	<input type="checkbox"/> Item 3.2
<input checked="" type="checkbox"/> Item 3.3	<input checked="" type="checkbox"/> Item 3.3	<input type="checkbox"/> Item 3.3	<input checked="" type="checkbox"/> Item 3.3	<input type="checkbox"/> Item 3.3
Estado 4 do Alpha				
<input checked="" type="checkbox"/> Item 3.1	<input checked="" type="checkbox"/> Item 3.1	<input type="checkbox"/> Item 3.1	<input type="checkbox"/> Item 3.1	<input type="checkbox"/> Item 3.1
<input checked="" type="checkbox"/> Item 3.2	<input checked="" type="checkbox"/> Item 3.2	<input type="checkbox"/> Item 3.2	<input type="checkbox"/> Item 3.2	<input type="checkbox"/> Item 3.2
<input checked="" type="checkbox"/> Item 3.3	<input checked="" type="checkbox"/> Item 3.3	<input type="checkbox"/> Item 3.3	<input type="checkbox"/> Item 3.3	<input type="checkbox"/> Item 3.3
Estado 5 do Alpha				
<input type="checkbox"/> Item 3.1	<input checked="" type="checkbox"/> Item 3.1	<input type="checkbox"/> Item 3.1	<input type="checkbox"/> Item 3.1	<input type="checkbox"/> Item 3.1
<input type="checkbox"/> Item 3.2	<input checked="" type="checkbox"/> Item 3.2	<input type="checkbox"/> Item 3.2	<input type="checkbox"/> Item 3.2	<input type="checkbox"/> Item 3.2
<input type="checkbox"/> Item 3.3	<input checked="" type="checkbox"/> Item 3.3	<input type="checkbox"/> Item 3.3	<input type="checkbox"/> Item 3.3	<input type="checkbox"/> Item 3.3
Recursos	Recursos	Recursos	Recursos	Recursos

Mapeamento do progresso em relação a cada Alpha

Mais informações sobre a criação do guia [aqui](#).

Fonte: Elaborado pela autora, 2024

⁹ <https://workspace.google.com/products/sheets/>

Apresentada a estrutura desenhada para o Guia e para a Ferramenta de Autoavaliação, a seção seguinte (5.2) apresenta o seu conteúdo a partir de referências da literatura.

5.2 Conteúdo do Guia

Nos capítulos anteriores deste estudo são apresentados recortes da literatura (a) e das impressões de profissionais (b) sobre a implementação de um processo de Data Analytics:

- a) o estado da arte sobre quais fatores que mais influenciam o sucesso desse processo, conforme relatado na literatura (capítulo 3, mapeamento do estado da arte);
- b) e desses, o estado da prática sobre quais os fatores mais influenciam o sucesso desse processo conforme relato de profissionais do mercado de software brasileiro (capítulo 4, estado da prática).

A partir dos resultados do estado da prática, foram selecionados os fatores que mais de 50% dos respondentes consideraram relevantes. Isso resultou em 5 fatores considerados de maior impacto (vide capítulo 4.5):

- Um bom gerenciamento do processo de dados
- Uma boa etapa de preparação dos dados
- Clareza sobre a necessidade do processo
- Existência de profissionais qualificados (com experiência) em dados
- Suporte organizacional (da gerência) para implementação

Sendo assim, propõe-se que esses 5 fatores sejam os Alphas do Guia, sendo elementos norteadores para que micro e pequenas empresas de software brasileiras possam construir e evoluir seu processo de Data Analytics.

Quadro 12. Alphas e frequência com que aparecem nos estudos realizados

Alpha	Frequência no Estado da Arte	Frequência no Estado da prática
Gerenciamento do processo de dados	10 estudos	95 respondentes
Preparação de dados	8 estudos	94 respondentes
Clareza sobre objetivos	5 estudos	87 respondentes
Qualificação de profissionais	5 estudos	83 respondentes
Suporte organizacional	9 estudos	82 respondentes

Cada um desses cinco fatores é definido como um *Alpha* do modelo Essence, conforme apresentado a seguir.

5.2.1 Alpha 1: Gerenciamento do processo de dados

O fator mais indicado pelas organizações respondentes foi 'um bom gerenciamento do processo de dados', o que deu origem ao primeiro Alpha deste modelo. Estudos compartilhados pela Data Science Process Alliance (Holtz, 2024) sugerem que, para alcançar um bom gerenciamento do processo de dados, deve-se utilizar uma composição entre "o ciclo de vida natural de data science com um framework de colaboração ágil", chamado *Agile Data Science*.

A ideia do *Agile Data Science* é a junção de princípios ágeis com práticas de Data Analytics. O processo de Data Analytics e suas etapas são respeitados e nele é incorporada a filosofia ágil.

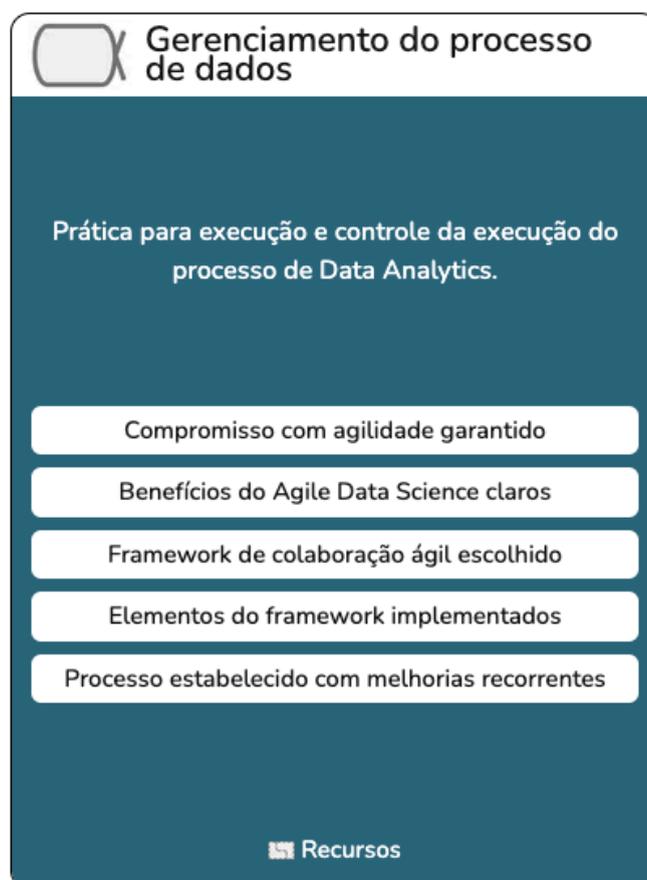
A fim de auxiliar as empresas a implementar o *Agile Data Science*, Holtz (2024) sugere que os seguintes elementos sejam considerados como marcadores de progresso em direção a uma implementação bem sucedida:

1. Garantia de compromisso com Agilidade;
2. Clareza, entre o time, sobre os benefícios do Agile Data Science;
3. Desenho ou escolha de um framework de colaboração ágil, que funcione para seu contexto;
4. Implementação do framework escolhido, iniciando pelos elementos mais fáceis ou mais importantes, de acordo com o contexto;

5. Ajuste do processo

Assim, para o Guia, as cinco práticas são organizadas como estados, pois entende-se que é importante evoluir uma prática em sequência da outra, para um bom resultado.

Figura 12. Alpha 1: Gerenciamento do processo de dados



Fonte: Elaborado pela Autora, 2024

Dessa forma, o Alpha “Gerenciamento do processo de dados” é composto por 5 estados que marcam o seu progresso, sendo 1 o inicial e 5 o seu estado final:

1. Compromisso com agilidade garantido
2. Benefícios do Agile Data Science claros
3. Framework de colaboração ágil escolhido
4. Elementos do framework implementados

5. Processo estabelecido com melhorias recorrentes

5.2.2 Alpha 2: Preparação de dados

Apesar de ser apenas uma das seis grandes etapas de um processo de Data Analytics, ter uma boa etapa de preparação dos dados foi um dos fatores com maior frequência no estado da arte e no estado da prática, sendo considerado, portanto, um dos mais importantes para o sucesso do processo.

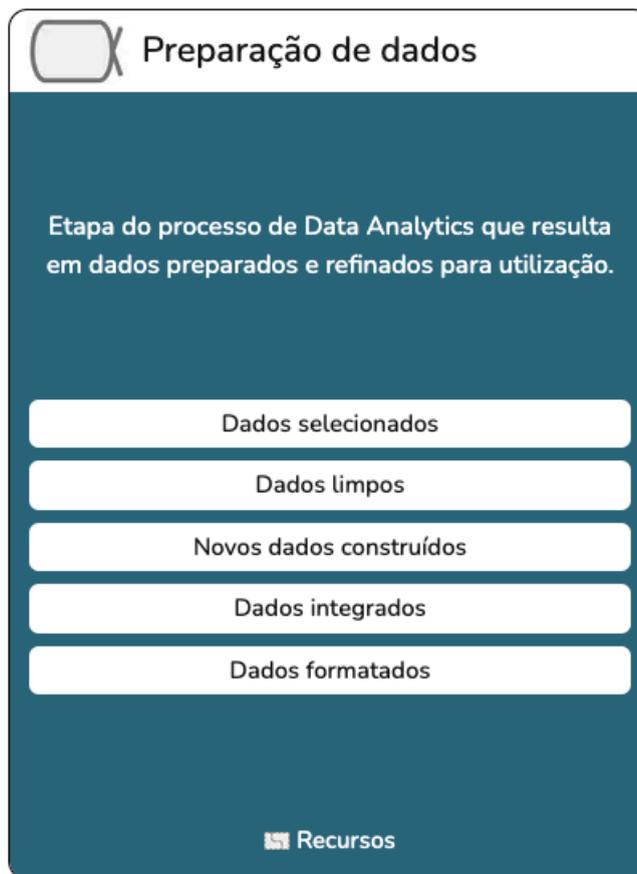
Holtz (2024) reafirma essa constatação quando diz que uma regra comum em processos de dados é que 80% do projeto consiste na etapa de preparação. Dessa forma, esse foi escolhido como segundo Alpha do modelo, para que nele seja empregado um foco maior.

A partir do modelo CRISP-DM, abordagem mais comum para projetos de dados (Holtz, 2024), são identificadas cinco tarefas: selecionar dados, limpar dados, construir dados, integrar dados e formatar dados.

Essas tarefas se traduzem nos estados do segundo Alpha do modelo, de modo que as organizações possam verificar se estão tirando o melhor de cada etapa até o final do processo de limpeza de dados.

1. Dados selecionados
 - a. Listar os dados a serem usados/excluídos
 - b. Documentar as razões para decisões de uso/exclusão
2. Dados limpos
 - a. Documentar as decisões e ações que foram tomadas para resolver problemas de qualidade de dados.
3. Novos dados construídos
 - a. Gerar novos atributos a partir de outros existentes no mesmo registro. Exemplo: $\text{area} = \text{altura} * \text{largura}$.
 - b. Gerar novos registros que adicionam conhecimento ou representam dados que não eram representados antes.
4. Dados integrados
 - a. Unir dados de diferentes tabelas sobre um mesmo objeto.
5. Dados formatados
 - a. Formatar dados para adequação a requisitos de negócios ou de ferramentas.

Figura 13. Alpha 2: Preparação de dados



Fonte: Elaborado pela Autora, 2024

O segundo Alpha do modelo, chamado “Preparação de dados”, pode ser avaliado a cada execução de uma instância do processo de dados. São 5 estados que o compõem:

1. Dados selecionados
2. Dados limpos
3. Novos dados construídos
4. Dados integrados
5. Dados formatados

5.2.3 Alpha 3: Qualificação de profissionais

O terceiro Alpha do modelo foi escolhido a partir da relevância atribuída ao fator "Existência de profissionais qualificados" nas pesquisas do estado da arte e do estado da prática.

Também nesse tema, a norma ISO 15504 (2012) (atualmente renumerada para série ISO 33000) possui um grupo de processos chamado Recurso e Infraestrutura (RIN), formado por processos focados em fornecer recursos humanos adequados para qualquer outro processo executado pela organização.

Um desses processos, o RIN.1, tem foco na garantia de profissionais qualificados. Como definido,

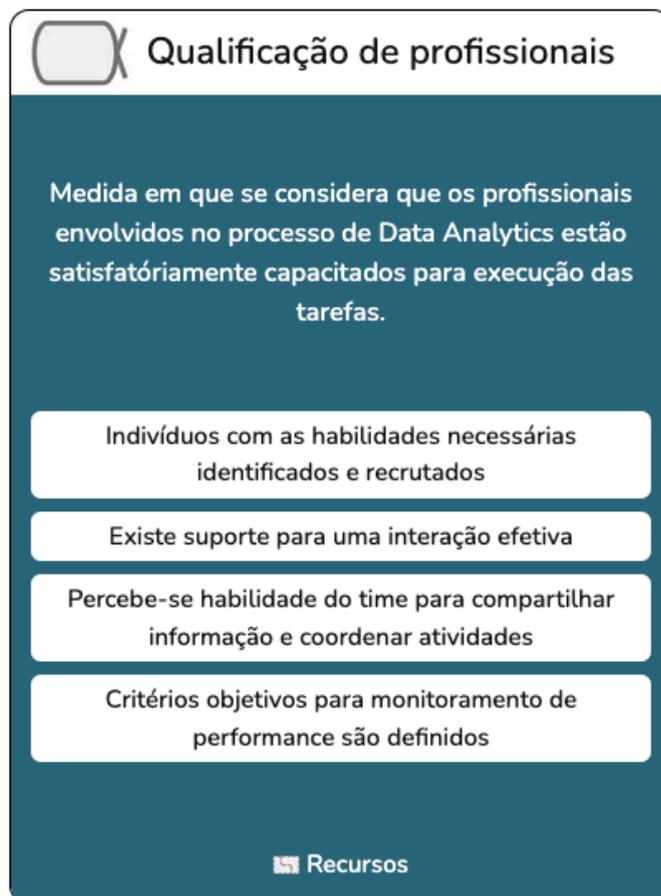
O propósito do processo de gerenciamento de recursos humanos é fornecer indivíduos à organização e aos projetos que possuam as habilidades e conhecimentos necessários para desempenhar seus papéis de forma eficaz e para trabalhar juntos como um grupo coeso (ISO15504, 2012).

A seguir são listados os resultados esperados do processo, de acordo com a RIN.1, bem como as práticas que auxiliam a alcançá-los.

1. Indivíduos com as habilidades e competências necessárias são identificados e recrutados
 - a. Identificar as habilidades e competências necessárias
 - b. Avaliar as habilidades e competências necessárias
 - c. Estabelecer um programa sistemático para recrutamento que atenda as necessidades.
2. Existe suporte para interação efetiva entre indivíduos e grupos;
 - a. Definir times para projetos e tarefas com uma estrutura e regras operacionais
 - b. Empoderar os times, garantindo que possuam um entendimento do seu trabalho, visão compartilhada do interesse comum, mecanismos apropriados para comunicação e suporte da gerência.
3. A equipe possui as habilidades para compartilhar informação e coordenar as atividades eficientemente;
 - a. Definir e oferecer oportunidades para desenvolvimento de competências.
4. Critérios objetivos são definidos para um grupo e indivíduos para que a performance seja monitorada e feedback seja recebido.
 - a. Avaliar a performance dos colaboradores considerando as suas contribuições para os objetivos definidos.

- b. Fornecer feedback para os colaboradores de acordo com avaliação de performance.
- c. Manter registros dos detalhes relacionados às habilidades, treinamentos e avaliações de performance dos colaboradores.

Figura 14. Alpha 3: Qualificação de profissionais



Fonte: Elaborado pela Autora, 2024

Conforme a Figura 14, acima, ilustra, o terceiro Alpha do modelo, “Qualificação de profissionais” é composto por 4 estados:

1. Indivíduos com as habilidades necessárias identificados e recrutados
2. Existe suporte para uma interação efetiva
3. Percebe-se habilidade do time para compartilhar informação e coordenar atividades
4. Critérios objetivos para monitoramento de performance são definidos

Embora possa ser relacionado com outras áreas, esse Alpha pode ser avaliado especificamente no contexto dos profissionais que trabalham com dados, contribuindo para a qualidade desse processo.

5.2.4 Alpha 4: Clareza sobre objetivos

Historicamente, a falta de alinhamento entre os objetivos de um processo e os objetivos da organização são responsáveis por boa parte do fracasso em esforços para melhoria de processos ou da falta de confiança para realizá-los. Essa constatação é evidenciada pelo estudo de Lepmets et al. (2012) e também pelos resultados da pesquisa do presente trabalho, que demonstram como a clareza sobre a necessidade do processo ainda é uma preocupação nos dias atuais. Ter clareza sobre isso foi um dos fatores considerados relevantes para o sucesso da iniciativa. Assim, deste fator surgiu o quarto Alpha: clareza sobre objetivos. Esse alpha evidencia a necessidade de ter clareza sobre os porquês da aplicação de Data Analytics, mostrando quais perguntas relevantes se quer responder e quais métricas poderiam compor a resposta.

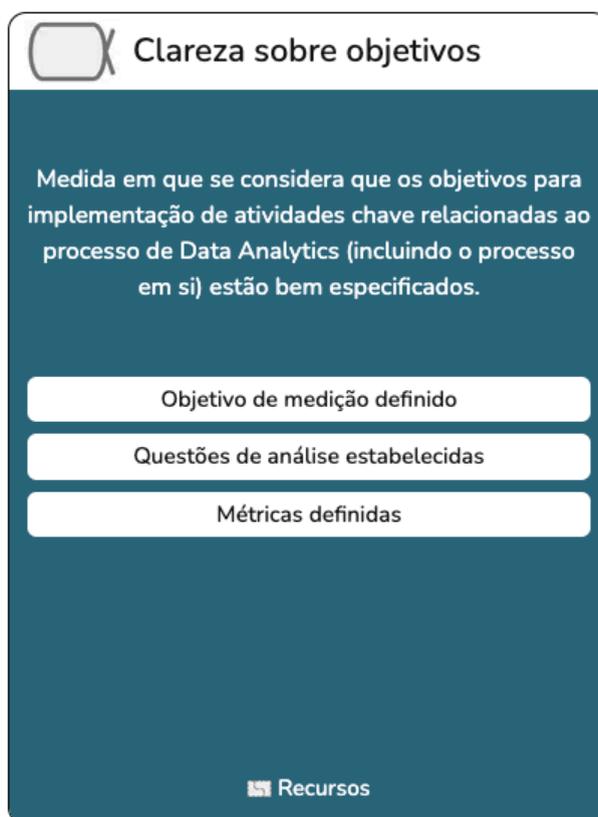
A abordagem de medição *Goal Question Metric* (GQM) apresenta um modelo sequencial de etapas que evidencia um objetivo (Goal), o qual é quebrado em questões (Question) a serem respondidas cada uma pela interpretação de métricas (Metric) (Basili et al., 1994). Assim, a implementação de um processo de Data Analytics nas organizações pode ter seu objetivo claramente definido e métricas que acompanhem a sua efetividade.

O Guia para evolução no quarto Alpha do modelo segue a abordagem apresentada por Basili et al. (1994), ainda muito utilizada atualmente para garantir que as organizações tenham uma estrutura capaz de entender os porquês do desenvolvimento das métricas com o processo de Data Analytics e, com isso, ter mais segurança no processo.

1. Objetivo de medição definido
 - a. Especificar o propósito de cada medição
 - b. Identificar qual é o objeto a ser medido
 - c. Identificar qual o ponto de vista pelo qual a medida é feita
2. Questões de análise estabelecidas

- a. Definir um conjunto de questões que caracterizam o objeto a ser medido;
 - b. Definir um aspecto de qualidade que se deseja medir sobre o objeto;
 - c. Especificar as questões com base no ponto de vista almejado.
3. Métricas definidas
- a. Definir dados objetivos e/ou subjetivos para associar às métricas e respondê-las em uma forma quantitativa.

Figura 15. Alpha 4: Clareza sobre objetivos



Fonte: Elaborado pela Autora, 2024

O quarto Alpha do modelo, “Clareza sobre objetivos”, é composto por 3 estados (listados a seguir) e sua evolução garante que as organizações tenham clareza sobre as motivações e resultados de cada iniciativa, utilizando o processo de dados.

1. Objetivo de medição definido
2. Questões de análise estabelecidas
3. Métricas definidas

5.2.5 Alpha 5: Suporte organizacional

O último Alpha do modelo se relaciona com a existência de um suporte organizacional, por parte da gerência, para implementação do processo de Data Analytics.

Para que o processo seja efetivo, os resultados do mapeamento do estado da arte e do estado da prática indicaram que a gerência deve dar suporte adequado, garantindo que o processo seja não só executado mas efetivamente institucionalizado.

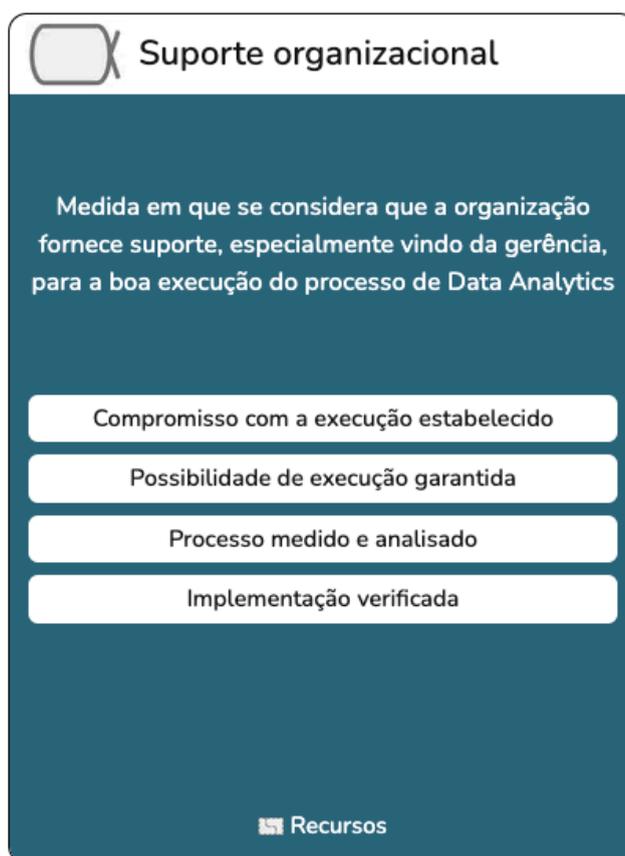
O People Capability Maturity Model apresenta, também, práticas que auxiliam na “institucionalização de práticas na cultura de uma organização de modo que sejam efetivas, repetíveis e duradouras” (PCMM, 2001, p. 54).

As quatro etapas propostas pelo People CMM e utilizadas no desenvolvimento do Guia são:

- a) Compromisso com a execução estabelecido
 - i) Criar diretrizes e regras formais que definem como os processos devem ser realizados;
 - ii) Garantir que práticas tenham visibilidade pelo apoio ativo dos líderes executivos;
 - iii) Atribuir responsabilidades específicas em toda a organização.
- b) Possibilidade de execução garantida
 - i) Disponibilizar pessoas, ferramentas, materiais, orçamento e outros recursos necessários;
 - ii) Definir claramente as funções e responsabilidades dentro da organização ou unidade;
 - iii) Capacitar a equipe para garantir que ela e a organização estejam prontas para implementar as práticas.
- c) Processo medido e analisado
 - i) Definir indicadores ou medições específicas para avaliar a implementação das práticas;
 - ii) Revisar sistematicamente os dados coletados para interpretar o desempenho das práticas;
- d) Implementação verificada

- i) Garantir que as atividades realizadas seguem as diretrizes definidas;
- ii) Avaliar e inspecionar, de maneira imparcial e com base em evidências.

Figura 16. Alpha 5: Suporte organizacional



Fonte: Elaborado pela Autora, 2024

O quinto e último Alpha do modelo, “Identificação de oportunidades” é composto pelos seguintes 4 estados:

1. Compromisso com a execução estabelecido
2. Possibilidade de execução garantida
3. Processo medido e analisado
4. Implementação verificada

5.3 Apresentação final do guia e ferramenta de autoavaliação

Como produto final da identificação dos Alphas, seus estados, itens que marcam a evolução de um Alpha e recursos que auxiliam nesse processo, temos o Guia.

A estrutura e elementos do Guia são apresentados em um website, desenvolvido utilizando HTML¹⁰ e JavaScript¹¹. O website é ilustrado na Figura 17 e pode ser acessado pelo link <https://guia-data-analytics.vercel.app/>.

Figura 17. Pré visualização do website do Guia

Fonte: Elaborado pela autora, 2024

No website, o usuário pode ler uma breve introdução ao contexto do Guia e sua estrutura, assim como navegar pelos cards criados.

A ferramenta de autoavaliação é disponibilizada como um template de arquivo Spreadsheet e permite que a organização trabalhe em sua própria cópia do template. Na ferramenta são apresentados os quatro elementos (Alphas, Estados dos Alphas, itens de checklist e Recursos) de forma interativa, auxiliando a organização a mapear sua posição atual e ter um direcionamento para evoluir.

¹⁰ <https://developer.mozilla.org/pt-BR/docs/Web/HTML>

¹¹ <https://developer.mozilla.org/pt-BR/docs/Web/JavaScript>

A partir da estrutura criada e apresentada na Figura 11, o template a ser disponibilizado foi finalizado com a substituição dos valores dos Alphas, Estados e itens por aqueles encontrados na pesquisa. Pequenas alterações de layout foram necessárias para melhor comportar o conteúdo real, como pode ser observado na Figura 18.

Figura 18. Template da Ferramenta de Autoavaliação do Guia

GUIA PARA IMPLEMENTAÇÃO DO PROCESSO DE DATA ANALYTICS				
<p>Este guia foi desenvolvido para apoiar micro e pequenas empresas brasileiras de software na identificação e aprimoramento das áreas essenciais para o sucesso na implementação e uso de Data Analytics. Ele é baseado no conceito de Alphas, proveniente da metodologia Essence. Alphas representam aspectos críticos que precisam ser trabalhados para alcançar um objetivo. Cada Alpha é composto por estados de evolução, que indicam o progresso de sua organização em relação a esses aspectos. Mais informações sobre a criação do guia nesse link.</p> <p>COMO UTILIZAR O GUIA:</p> <p>1. Identifique o estado atual de sua organização: Para cada Alpha, revise os itens listados e marque aqueles que sua organização já realiza com sucesso. Os estados variam de 1 a 5, sendo 1 o menos evoluído e 5 o mais avançado.</p> <p>2. Visualize seu progresso: Após preencher os checkboxes, utilize o gráfico gerado para obter uma visão clara das áreas mais desenvolvidas (próximas de 5) e daquelas que necessitam de maior atenção (próximas de 0).</p> <p>3. Planeje as próximas ações: Com base nos estados identificados, verifique o que seriam os próximos passos. Além disso, explore os recursos recomendados para obter materiais de apoio e implementar as práticas necessárias.</p>				
Gerenciamento do processo de dados	Preparação de dados	Qualificação de profissionais	Identificação de oportunidades	Suporte organizacional
<p>Compromisso com agilidade garantido</p> <p><input type="checkbox"/> Time entende princípios de agilidade e os aplica</p> <p>Benefícios do Agile Data Science claros</p> <p><input type="checkbox"/> Time compreende necessidade de uma abordagem ágil no processo de dados</p> <p>Framework de colaboração ágil escolhido</p> <p><input type="checkbox"/> Há concordância em relação ao framework sendo/a ser utilizado</p> <p>Elementos do framework implementados</p> <p><input type="checkbox"/> Os elementos que definem o framework escolhido (e fazem sentido para a organização) são efetivamente implementados na prática</p>	<p>Dados selecionados</p> <p><input type="checkbox"/> Dados a serem usados/excluídos são documentados</p> <p><input type="checkbox"/> Razões para decisões de uso/exclusão são documentadas</p> <p>Dados limpos</p> <p><input type="checkbox"/> Decisões e ações tomadas para resolver problemas de qualidade de dados são documentadas</p> <p>Novos dados construídos</p> <p><input type="checkbox"/> Novos atributos são gerados a partir de outros existentes no mesmo registro.</p> <p><input type="checkbox"/> Novos registros que adicionam conhecimento ou representam dados que não eram representados antes são gerados.</p>	<p>Indivíduos com as habilidades necessárias são identificados e recrutados</p> <p><input type="checkbox"/> Habilidades e competências necessárias são identificadas</p> <p><input type="checkbox"/> Habilidades e competências necessárias são avaliadas</p> <p><input type="checkbox"/> É estabelecido um programa sistemático para recrutamento que atenda as necessidades.</p> <p>Existe suporte para uma interação efetiva</p> <p><input type="checkbox"/> Times para projetos são definidos com uma estrutura e regras operacionais</p> <p><input type="checkbox"/> Times são empoderados garantindo que possuem um entendimento do seu trabalho, visão compartilhada do interesse comum, mecanismos apropriados para comunicação e suporte da gerência.</p> <p>Percebe-se habilidade do time para compartilhar informação e coordenar atividades</p>	<p>Objetivo de medição definido</p> <p><input type="checkbox"/> Propósito de cada medição especificado</p> <p><input type="checkbox"/> Objeto a ser medido identificado</p> <p><input type="checkbox"/> Ponto de vista pelo qual a medida é feita identificado</p> <p>Questões de análise estabelecidas</p> <p><input type="checkbox"/> Conjunto de questões que caracterizam o objeto a ser medido é definido</p> <p><input type="checkbox"/> Aspectos de qualidade que se deseja medir sobre o objeto são definidos</p> <p><input type="checkbox"/> Questões com base no ponto de vista almejado são especificadas</p>	<p>Compromisso com a execução estabelecido</p> <p><input type="checkbox"/> Diretrizes e regras formais são criadas para definir como os processos devem ser realizados</p> <p><input type="checkbox"/> Práticas possuem visibilidade pelo apoio ativo dos líderes executivos</p> <p><input type="checkbox"/> Responsabilidades específicas são atribuídas em toda a organização</p> <p>Possibilidade de execução garantida</p> <p><input type="checkbox"/> Recursos necessários (pessoas, ferramentas, materiais, orçamento e outros) são disponibilizados.</p> <p><input type="checkbox"/> Funções e responsabilidades são claramente definidas dentro da organização ou unidade</p> <p><input type="checkbox"/> Equipes são capacitadas para garantir que estejam, com a organização, prontas para implementar as práticas.</p>

Fonte: Elaborado pela autora, 2024

A ferramenta é composta por um cabeçalho com descrição do trabalho e modo de uso; os Alphas, seus estados e itens que contribuem para evolução; e um gráfico de radar com indicação visual da condição atual da organização em relação aos Alphas.

O modo de utilização da ferramenta é resumida em três etapas:

1. Identifique o estado atual de sua organização: Para cada Alpha, revise os itens listados e marque aqueles que sua organização já realiza com sucesso. Os estados variam de 1 a 5, sendo 1 o menos evoluído e 5 o mais avançado.

2. Visualize seu progresso: Após preencher os checkboxes, utilize o gráfico gerado para obter uma visão clara das áreas mais desenvolvidas (próximas de 5) e daquelas que necessitam de maior atenção (próximas de 0).

3. Planeje as próximas ações: Com base nos estados identificados, verifique o que seriam os próximos passos. Além disso, explore os recursos recomendados para obter materiais de apoio e implementar as práticas necessárias.

Dado que a ferramenta é disponibilizada em formato de template, as organizações devem criar uma cópia própria do arquivo para conseguir fazer o próprio mapeamento. Por isso, o link de compartilhamento do template já leva o usuário à criação da sua cópia pessoal:

 [Guia para implementação de Data Analytics em micro e pequenas empresas b...](#)

A partir do gráfico gerado, a organização tem uma representação visual do estágio em que se encontra em cada área. Também foram adicionados links para os recursos externos de apoio à evolução de cada Alpha (todos usados como referência para a construção do Guia), para que as organizações possam conhecer a teoria, visando a evolução de suas práticas.

5.4 Avaliação do Guia

Com o objetivo de avaliar o estudo, um Painel de Especialistas foi realizado. A estratégia de utilizar um painel de especialistas para avaliação consiste em compilar opiniões individuais de especialistas sobre um determinado tema, para analisar e avaliar uma proposição (HELMER, 1967).

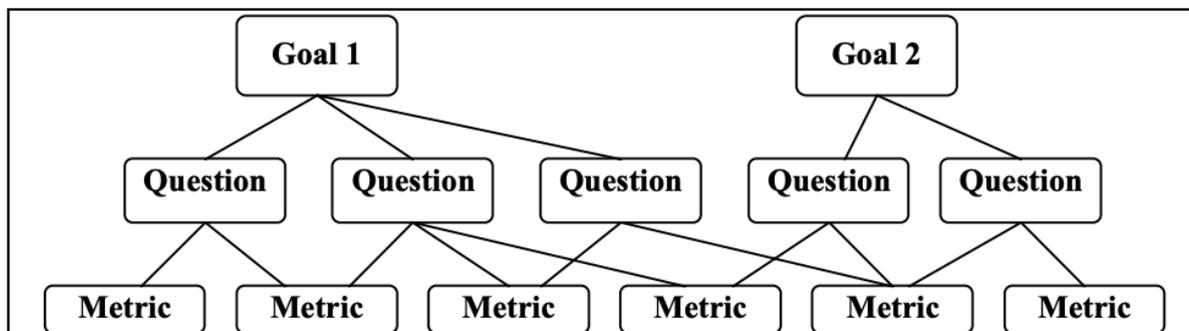
5.4.1 Planejamento da Avaliação

Buscou-se avaliar o Guia em relação à sua abrangência (utilidade do conteúdo proposto em múltiplos contextos), aplicabilidade (qualidade e utilidade prática do conteúdo do Guia) e cobertura (extensão do conteúdo em relação ao tema).

Para uma avaliação coesa e objetiva, foi utilizada a abordagem *Goal/Question Metric* (GQM), proposta por Basili, (1994). Conforme já citado, na

abordagem GQM realiza-se a definição de metas que geram perguntas e, essas, derivam métricas para análise, conforme Figura 19.

Figura 19. Visão geral da abordagem GQM



Fonte: Basili (1994)

Para avaliação do Guia proposto neste trabalho, foram definidas as seguintes metas, perguntas e métricas:

- a) **Meta 1:** Avaliar a abrangência do Guia, a partir do ponto de vista de um painel de especialistas.
 - i) **Pergunta 1.1:** O Guia pode ser aplicado em organizações de diferentes setores (desde que tenha uma área de software)?
 - 1) **Métrica 1.1.1:** Impressão subjetiva da adaptabilidade do Guia às múltiplas áreas
 - ii) **Pergunta 1.2:** O Guia pode ser adaptável a diferentes métodos ágeis escolhidos para o processo de Data Analytics?
 - 1) **Métrica 1.2.1:** Impressão subjetiva da adaptabilidade do Guia aos múltiplos modelos ágeis, como Kanban, Scrum, etc.
 - iii) **Pergunta 1.3:** O Guia pode ser adaptado a organizações de diferentes tamanhos que trabalham com desenvolvimento de software?
 - 1) **Métrica 1.3.1:** Impressão subjetiva da adaptabilidade do Guia a organizações, que desenvolvem software, de diferentes tamanhos.
- b) **Meta 2:** Avaliar a aplicabilidade do Guia, a partir do ponto de vista de um painel de especialistas.

- i) **Pergunta 2.1:** O Guia aborda práticas que efetivamente podem ser observadas?
 - 1) **Métrica 2.1.1:** Impressão subjetiva do quanto o Guia se aplica à realidade das organizações, considerando a experiência prática dos participantes.
 - ii) **Pergunta 2.2:** O Guia aborda orientações e conteúdos que podem efetivamente ser aplicados?
 - 1) **Métrica 2.2.1:** Impressão subjetiva da aplicabilidade das orientações propostas no Guia.
- c) **Meta 3** Avaliar a cobertura do conteúdo do Guia, a partir do ponto de vista de um painel de especialistas.
- i) **Pergunta 3.1:** A seção de descrição fornece embasamento necessário para entendimento do conteúdo do Guia?
 - 1) **Métrica 3.1.1:** Impressão pessoal dos participantes em relação à descrição do Guia.
 - ii) **Pergunta 3.2:** O conteúdo do Guia proposto para gestão de riscos está alinhado aos princípios e valores ágeis?
 - 1) **Métrica 3.2.1:** Impressão pessoal dos participantes em relação ao alinhamento do Guia com os princípios e valores do Manifesto ágil (BECK et al., 2001).
 - iii) **Pergunta 3.3:** O Guia apresenta atividades, técnicas, ferramentas e demais componentes suficientes para sua aplicação prática?
 - 1) **Métrica 3.3.1:** Impressão subjetiva dos participantes em relação a quantidade suficiente de recursos para aplicação do Guia.

As respostas para cada uma das perguntas segue a escala Likert de 4 pontos, variando de “Discordo totalmente” até “Concordo totalmente” (LIKERT, 1932). Não foi utilizado o ponto central de avaliação, equivalente a resposta neutra, pois esta prática evita anulação da questão (GARLAND, 1991).

Também foram adicionadas no questionário duas perguntas abertas, a fim de coletar mais impressões e possíveis sugestões que contribuam para o ajuste do

Guia. A primeira pergunta aberta é referente a quais elementos (se algum) parecem faltantes ou errôneos no Guia e a segunda abre espaço para comentários gerais.

5.4.2 Seleção dos especialistas

Foram convidados profissionais da área de dados - tanto acadêmicos quanto profissionais do mercado - para realizar a avaliação a partir das perguntas definidas no modelo GQM. Para contextualizar o painel de especialistas foram coletadas algumas informações sobre os avaliadores, apresentadas no Quadro 13:

Quadro 13. Perfil do painel de especialistas

Formação	Cargo atual	Tempo de experiência na área de dados
Bacharel em Ciência da Computação, Pós Doutora em Ciências da Educação	Docente na área de dados	30 anos
Pós graduação em tecnologia da informação	Analista de dados sênior	15 anos
Mestre em Ciência da Computação, Bacharel em Análise de Sistemas de Computação	Engenheiro de dados Sênior	8 anos
Bacharel em Ciências Contábeis	Gerente de dados	3 anos
Mestre em Administração de empresas, Bacharel em Engenharia de Controle e Automação	CEO em empresa de dados	1 ano ¹²

Fonte: Elaborado pela autora, 2024

Os especialistas foram convidados a fazer a avaliação do guia a partir da ferramenta de autoavaliação proposta. Para consolidação da avaliação, preencheram um questionário na ferramenta Google Forms, com as perguntas definidas anteriormente. O formulário pode ser visto no apêndice E.

¹² Apesar do pouco tempo de experiência em dados, o especialista foi considerado para o painel pois possui mais de 8 anos de experiência em gerenciamento no mercado de software, e isso, agregado a experiência atual em dados, foi considerado interessante para análise do guia dado o contexto em que são aplicados.

5.4.1 Resultados da avaliação

Por meio da ferramenta de autoavaliação, o Guia foi avaliado, pelos especialistas, em três aspectos principais: abrangência, aplicabilidade e cobertura. Considerando as respostas obtidas, observa-se que, em geral, a avaliação do Guia foi positiva, conforme revelam os gráficos do Apêndice F.

Em relação à abrangência, apenas 1 especialista discordou de que o Guia poderá ser aplicado em organizações de diferentes setores. Os demais concordaram com a afirmativa. Quando perguntado se o Guia poderia ser aplicado em organizações de diferentes tamanhos e adaptável a diferentes métodos ágeis, a maioria dos especialistas concordou totalmente, revelando uma abrangência significativa.

No que se refere à aplicabilidade, todos os especialistas concordaram, embora que alguns não totalmente, que o Guia aborda práticas que podem ser efetivamente observadas e conteúdos que podem ser aplicados.

Por fim, predomina a opinião de que há uma cobertura satisfatória, embora 01 especialista tenha considerado que o Guia pode ser melhor descrito.

Nos comentários das questões abertas, aparecem sugestões para o aprimoramento do Guia, tais como aperfeiçoar a descrição para aplicação do Guia e ampliar a conceituação de termos utilizados, indicações consideradas pertinentes para o objetivo deste estudo.

Além dessas sugestões, os especialistas manifestaram posicionamento favorável ao resultado do Guia, como o fato de possuir linguagem e metodologia acessíveis e apresentar uma visualização gráfica que agiliza a tomada de decisão em relação aos próximos passos para a evolução dos processos de Data Analytics.

5.5 Ameaças à validade do Guia

A utilização do Guia e da Ferramenta de autoavaliação deve levar em consideração alguns aspectos que podem comprometer sua efetividade.

Não foi executada uma avaliação prática do guia que pudesse validar os resultados positivos que se espera de sua aplicação. Para mitigar esse risco, foi realizado um painel de especialistas que contribuiu com considerações a partir da sua experiência.

Os conteúdos que compõem os Estados e Itens de Checklist do guia são recortes de outros materiais mais completos e conhecidos da literatura, mas não necessariamente foram validados no mesmo contexto em que o Guia propõe. Por isso, todos os materiais originais foram referenciados tanto neste documento quanto nos cards de Referência do Guia, a fim de permitir um aprofundamento nos frameworks em sua íntegra.

6. CONCLUSÃO

Este trabalho apresenta uma pesquisa sobre os fatores mais influentes na implementação de um processo de Data Analytics em micro e pequenas empresas brasileiras de software, bem como uma proposta de guia para essa implementação. Foram realizados um mapeamento do Estado da Arte e um mapeamento do Estado da Prática, para encontrar quais fatores são indicados pela literatura e confirmados por profissionais do mercado alvo da pesquisa. Ao final, foi desenvolvida uma proposta de guia que auxilia as organizações a evoluírem na implementação de seus processos de Data Analytics, a partir dos fatores identificados nos mapeamentos.

Inicialmente foi realizada uma pesquisa e descrição dos principais conceitos relativos ao processo de Data Analytics (O1) bem como outros elementos importantes para o desenvolvimento do trabalho, como forças impulsionadoras, a contextualização de micro e pequenas empresas e a abordagem Essence.

Na sequência foi realizado um levantamento do Estado da Arte por meio de um Mapeamento Sistemático da Literatura (O2), levando a estudos que identificam fatores de influência. Como principais resultados, esses estudos identificaram 26 fatores de influência (já agrupados), sendo 12 os mais recorrentes: 'Suporte técnico para implantação de sistema', 'Qualificação de funcionários', 'Qualidade de Software', 'Gerenciamento do projeto de dados', 'Orçamento/Recursos disponíveis', 'Conhecimento dos benefícios', 'Preparação dos dados', 'Confiança em dados', 'Profissionais já qualificados em dados', 'Observação de necessidade', 'Resistência natural' e 'Crença de que o tópico é complexo'.

A partir dos 12 fatores mais recorrentes encontrados na literatura, buscou-se conhecer a mesma questão no contexto de micro e pequenas empresas brasileiras de software. Assim, foi identificado o estado da prática por meio de um *survey* com profissionais inseridos no contexto de interesse (O3). Considerando a quantidade de participantes que selecionou cada fator como crítico, foi possível elencar 5 fatores mais relevantes: 'Um bom gerenciamento do processo de dados', 'Uma boa etapa de preparação dos dados', 'Clareza sobre a necessidade do processo', 'Existência de profissionais qualificados (com experiência) em dados' e 'Suporte organizacional (da gerência) para implementação'.

Para consolidar as descobertas dos dois mapeamentos, foi elaborado um guia (O4) com base na abordagem Essence. O guia apresenta não somente os fatores identificados como relevantes, para servir de norte às organizações, como também boas práticas encontradas na literatura que, quando bem estabelecidas, identificam o status da organização na evolução desses fatores.

O guia foi disponibilizado numa ferramenta de acesso online (O5), em que os usuários podem conhecer o objetivo, estrutura e elementos do guia, além de utilizar uma ferramenta de autoavaliação para situar seu estado atual em relação aos fatores críticos e monitorar sua evolução.

Por fim, foram avaliadas a aplicabilidade, a abrangência e a cobertura do guia, (O6) por meio de um painel de especialistas que considerou estes aspectos favoráveis e fez apontamentos relevantes para o aprimoramento do material.

Dessa forma, considera-se que o objetivo geral deste trabalho foi alcançado, gerando uma contribuição tanto para o público alvo do estudo - micro e pequenas organizações brasileiras de software - quanto para a área de pesquisa em que se encontra.

As análises realizadas neste trabalho indicaram algumas ideias sobre o processo de Data Analytics que evidenciam a necessidade de continuar explorando aspectos da área em novas pesquisas. Por exemplo, o caráter dinâmico da execução do processo gera interessantes investigações sobre o uso de frameworks para gerenciamento. Além disso, se nota uma dificuldade em mensurar resultados do processo, o que dificulta ainda mais a avaliação das práticas.

Como trabalhos futuros a partir dessa pesquisa, sugere-se a ampliação da amostra do survey, a avaliação prática do Guia gerado, a evolução do Guia com mais ferramentas fornecidas pela abordagem Essence, um recorte do estudo para considerar aspectos geográficos e uma análise temporal que identifique a mudança dos fatores ao longo do tempo.

Outro contexto para o qual se pode expandir esta pesquisa é o das organizações fora do contexto de software. Ao considerarmos os intensos avanços da tecnologia e, principalmente, da temática 'análise de dados', é importante que se analise a evolução desses assuntos em contextos mais distantes da tecnologia, especialmente dada a relevância que micro e pequenas empresas - dos mais diversos segmentos - têm no Brasil.

Similarmente, espera-se que esse trabalho contribua com a aproximação das descobertas acadêmicas com os contextos de organizações para que, juntos, tragam benefícios para a sociedade como um todo.

REFERÊNCIAS BIBLIOGRÁFICAS

ANWAR, Muhammad; KHAN, Sher Zaman; SHAH, Syed Zulfiqar Ali. Big data capabilities and firm's performance: a mediating role of competitive advantage. *Journal of Information & Knowledge Management*, v. 17, n. 04, p. 1850045, 2018.

ABES - Associação Brasileira das Empresas de Software. Mercado Brasileiro de Software: panorama e tendências, 2023 = Brazilian Software Market: scenario and trends, 2023 [versão para o inglês: Anselmo Gentile] — 1ª. ed. São Paulo: ABES - Associação Brasileira das Empresas de Software, 2023. Disponível em file:///Users/livia/Downloads/ABES_Relatorio_Estudo_de_Mercado_2023_Digital.pdf

BAIJENS, Jeroen; HELMS, Remko; KUSTERS, Rob. Data Analytics Project Methodologies: Which One to Choose?. In: *Proceedings of the 2020 International Conference on Big Data in Management*. 2020. p. 41-47. Disponível em <https://dl.acm.org/doi/10.1145/3437075.3437087>. Acesso em 5 de novembro de 2023.

BHARDWAJ, Shikha. Data analytics in small and medium enterprises (SME): a systematic review and future research directions. *Information Resources Management Journal (IRMJ)*, v. 35, n. 2, p. 1-18, 2022. 10.4018/IRMJ.291691.

BHATIASEVI, Veera; NAGLIS, Michael. Elucidating the determinants of business intelligence adoption and organizational performance. *Information development*, v. 36, n. 1, p. 78-96, 2020. Disponível em <https://doi.org/10.1177/0266666918811394>. Acesso em 9 de novembro de 2023.

BASILI, Victor R.; CALDIERA, Gianluigi H.; ROMBACH, Dieter. The Goal Question Metric Approach. Disponível em <https://www.semanticscholar.org/paper/The-Goal-Question-Metric-Approach-Basili-Caldiera/02e65151786574852007ecd007ee270c50470af0>.

CAO, Longbing. Data science: a comprehensive overview. *ACM Computing Surveys (CSUR)*, v. 50, n. 3, p. 1-42, 2017. Disponível em <https://doi.org/10.1145/3076253>. Acesso em 9 de novembro de 2023.

CARVALHO, Marly Monteiro de et al. MPEs de base tecnológica: conceituação, formas de financiamento e análise de casos brasileiros. . São Paulo: IPT/SEBRAE. Disponível em: <https://repositorio.usp.br/directbitstream/d694df4c-2505-43e7-9244-08483bab9c07/Carvalho-2001-MPEs.pdf>. Acesso em: 01 jul. 2024. , 2001

CRESWELL, J. W.; PLANO CLARK, V. L. Designing and conducting mixed methods research. 2nd. Los Angeles: SAGE Publications, 2011.

DAVENPORT, Thomas H.; PRUSAK, Laurence. Working knowledge: How organizations manage what they know. Harvard Business Press, 1998.

DASTGERDI, Amin Karimi; GANDOMANI, Taghi Javdani. On the appropriate methodologies for data science projects. In: 2021 International Conference on Information Technology (ICIT). IEEE, 2021. p. 667-673. 10.1109/ICIT52682.2021.9491712. Disponível em https://www.researchgate.net/publication/353484542_On_the_Appropriate_Methodologies_for_Data_Science_Projects. Acesso em 5 de novembro de 2023.

DURANGO VANEGAS, Claudia Elena et al . A Representation Based on Essence for the CRISP-DM Methodology. **Comp. y Sist.**, Ciudad de México , v. 27, n. 3, p. 675-689, sept. 2023 . Disponível em <http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-55462023000300675&lng=es&nrm=iso>. acesso em 16 nov. 2024. Epub 17-Nov-2023. <https://doi.org/10.13053/cys-27-3-3446>.

FAYYAD, Usama M. et al. Knowledge Discovery and Data Mining: Towards a Unifying Framework. In: KDD. 1996. p. 82-88.

GÖKALP, Mert Onuralp. Data Science Capability Maturity Model. 2021.

GOULDING, C. (2009). Grounded theory perspectives in organizational research. In D. A. Buchanan & A. Bryman (Eds.), The Sage handbook of organizational research methods. London: Sage.

GROVES, Robert M.; FOWLER, Floyd J.; COUPER, Mick P.; LEPKOWSKI, James M.; SINGER, Eleanor; TOURANGEAU, Roger. Survey Methodology. 2 ed. Willey. Ebook. Disponível em https://books.google.com.br/books?id=ctow8zWdyFgC&printsec=frontcover&hl=pt-BR&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

HELMER-HIRSCHBERG, O. Systematic Use of Expert Opinions. Report No. P-3721. Santa Monica, CA: The RAND Corporation, 1967.

HOLTZ, Nick. What is CRISP DM? 28 abr. 2024. Disponível em: <https://www.datascience-pm.com/crisp-dm-2/>. Acesso em: 2 jul. 2024.

LLAVE, Marilex Rea. Business Intelligence and Analytics in Small and Medium-sized Enterprises: A Systematic Literature Review, *Procedia Computer Science*, Volume 121, 2017, Pages 194-205, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2017.11.027>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1877050917322184>. Acesso em 23/08/2024.

LEPMETS, Marion; McBRIDE, Tom; RAS, Eric. Goal alignment in process improvement. *Journal of Systems and Software*. Volume 85, Issue 6, June 2012, Pages 1440-1452. Disponível em <https://www.sciencedirect.com/science/article/abs/pii/S016412121200026X>

LEWIN, Kurt. Teoria de campo em ciência social. [S. l.: s. n.]. 387 p.

MONTONI, Mariano Angel; DA ROCHA, Ana Regina Cavalcanti. Uma investigação sobre os fatores críticos de sucesso em iniciativas de melhoria de processos de software. In: Anais do X Simpósio Brasileiro de Qualidade de Software. SBC, 2011. p. 151-165.

MAROUFKHANI, Parisa et al. Big Data Analytics adoption: Determinants and performances among small to medium-sized enterprises, *International Journal of Information Management*, Volume 54, 2020, 102190, ISSN 0268-4012, <https://doi.org/10.1016/j.ijinfomgt.2020.102190>. Disponível em <https://www.sciencedirect.com/science/article/pii/S026840122030178X>.

NUNNALLY, J.C., 1978, *Psychometric Theory*, 2nd ed. New York, McGraw-Hill.

O'CONNOR, Christina; KELLY, Stephen. Facilitating knowledge management through filtered big data: SME competitiveness in an agri-food sector. *Journal of Knowledge Management*, v. 21, n. 1, p. 156-179, 2017. doi:10.1108/JKM-08-2016-0357.

OECD - Organisation for Economic Co-operation and Development. Enterprise by business size. Disponível em <https://www.oecd.org/en/data/indicators/enterprises-by-business-size.html#:~:text=In%20small%20and%20medium%2D-sized,employ%20250%20or%20more%20people>

OLSZAK, Celina M.; ZIEMBA, Ewa. Critical success factors for implementing business intelligence systems in small and medium enterprises on the example of

upper Silesia, Poland. *Interdisciplinary Journal of Information, Knowledge, and Management*, v. 7, p. 129-150, 2012. DOI: 10.28945/1584. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84862019547&doi=10.28945%2f1584&partnerID=40&md5=766ffe1471dfd5f9b6d74eb86b0ac340>.

OMG - OBJECT MANAGEMENT GROUP. *Essence- Kernel and Language for Software Engineering Methods*. Version 1.2, outubro 2018. Disponível em <https://www.omg.org/spec/Essence/1.2/About-Essence>.

OWUSU, Acheampong. Business intelligence systems and bank performance in Ghana: The balanced scorecard approach. *Cogent Business & Management*, v. 4, n. 1, p. 1364056, 2017. <https://doi.org/10.1080/23311975.2017.1364056>.

PARANHOS, Ranulfo; FIGUEIREDO FILHO, Dalson Britto; ROCHA, Enivaldo Carvalho da; SILVA JUNIOR, José Alexandre da; FREITAS, Diego. Uma introdução aos métodos mistos. **INTERFACES**. *Sociologias* 18 (42) • May-Aug 2016. Disponível em <https://www.scielo.br/j/soc/a/WtDMmCV3jQB8mT6tmpnzKc/#>.

PETERSEN, Kai et al. Systematic mapping studies in software engineering. In: 12th international conference on evaluation and assessment in software engineering (EASE). BCS Learning & Development, 2008.

PINSONNEAULT, Alain and KRAEMER, Kenneth. 1993. Survey Research Methodology in Management Information Systems: An Assessment. *Journal of Management Information Systems* 10, 2 (1993), 75–105.

RICHARDSON, W. Scott et al. The well-built clinical question: a key to evidence-based decisions. *ACP journal club*, v. 123, n. 3, p. A12-3, 1995.

ROCKART, J.F., 1979, "Chief executives define their own data needs", *Harvard Business Review*, v. 2, pp. 81-93.

SALTZ, Jeff. The data science maturity model. 10 abr. 2024. Disponível em: <https://www.datascience-pm.com/data-science-maturity-model/>. Acesso em: 2 jul. 2024.

SALTZ, Jeffrey S.; KRASTEVA, Iva. Current approaches for executing big data science projects—a systematic literature review. *PeerJ Computer Science*, v. 8, p. e862, 2022. doi: 10.7717/peerj-cs.862. PMID: 35494858; PMCID: PMC9044260. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9044260/> .

SALTZ, Jeffrey; SHAMSHURIN, Ivan; CONNORS, Colin. Predicting Data Science Sociotechnical Execution Challenges by Categorizing Data Science Projects. in: *Journal of the Association For Information Science and Technology*, 68(12):2720–2728, 2017. Disponível em <https://asistdl.onlinelibrary.wiley.com/doi/epdf/10.1002/asi.23873>

SARKER, Iqbal H. Data science and analytics: an overview from data-driven smart computing, decision-making and applications perspective. *SN Computer Science*, v. 2, n. 5, p. 377, 2021.

SAS. SEMMA. 2017. Disponível em: <https://documentation.sas.com/?docsetId=emref&docsetTarget=n061bzurmej4j3n1jnj8bbj1a2.htm&docsetVersion=14.3&locale=en>. Acesso em: 25 nov. 2020.

SEBRAE. Conheça as empresas de base tecnológica. 27 abr. 2023. Disponível em: <https://sebrae.com.br/sites/PortalSebrae/artigos/conheca-as-empresas-de-base-tecnologica,2b639482b04a4810VgnVCM100000d701210aRCRD>. Acesso em: 2 jul. 2024.

STRAUSS, A., CORBIN, J.M., 1998, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2nd ed., Sage Publications.

TAREK, Bel Hadj; ADEL, Ghodbane. Business intelligence versus entrepreneurial competitive intelligence and international competitiveness of North African SMEs. *Journal of International Entrepreneurship*, v. 14, p. 539-561, 2016. doi:10.1007/s10843-016-0194-8.

TURBAN, Efraim; WETHERBE, James; MCLEAN, Ephraim R. *Information technology for management: improving quality and productivity*. John Wiley & Sons, Inc., 1996.

WANG, Shouhong; WANG, Hai. Big data for small and medium-sized enterprises (SME): a knowledge management model. *Journal of Knowledge Management*, v. 24, n. 4, p. 881-897, 2020. doi:10.1108/JKM-02-2020-0081.

WIRTH, Rüdiger; HIPPE, Jochen. CRISP-DM: Towards a standard process model for data mining. In: *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*. 2000. p. 29-39.

WOHLIN, Claes et al. *Experimentation in software engineering*. Berlin: Springer, 2012.

APÊNDICE A - FATORES CRÍTICOS DE SUCESSO POR ESTUDO

1	<p>[R] Handling is too complicated [R] Processes of the BI report building are too complicated [R] Personnel using the BI solution are not qualified enough [R] Software errors [R] Inadequate security function [R] Contradictory data [R] Low speed of the product [R] Insufficient support [R] Limited data export functionalities [R] Problematic conflation of data [I]Structural fixing [I] Level of observance of budgets [I]Employee training [I]Innovativeness [I]Flexible reaction to changes in the market [I]Frequency of making operative decisions [I] Degree of time pressure [I]Kind of contact with suppliers/customers/public [I]Relative company growth [I]Service orientation [I]B2B orientation</p>
2	<p>[R] Lack of well defined business problem [R] Lack of manager's supporting [R] Lack of knowledge about BI system and its capabilities [R] No determining the expectation of BI user [R] Exceeded BI implementation budget [R] Ineffective BI project management [R] Lack of appropriate data for BI system [R] Complicated BI project [R] Lack of user training and support [R] Resistance of the "human factor" [R] Insufficient responsiveness on changing users' requirements [R] Lack of relations between business vision and BI system [R] Lack of BI system flexibility [R] Uncoordinated BI project activities [R] Misunderstanding of knowledge provided by BI [R] Different viewpoints on BI system implementation [I]The price of BI system and its implementation [I]Suitability BI for users' business needs [I]Integration BI system with operational systems [I] Well defined business problem and processes [I]Changeability and developing BI system [I]"User friendly" BI system [I]The reference list of BI supplier [I]Past cooperation with a BI supplier [I]The kind of BI technology and tools</p>
3	<p>Expected benefits Perception of strategic value Cost BIS is a part of ERP Management support Organizational culture Project champion Organizational data environment Organizational readiness Size External support</p>
4	<p>[R] Current IT support did not fully support the processes of the area [R] Limited interest from executives and owners on how to use BI as a decision support [R] When IT is outsourced, the companies have difficulties getting the support needed [R] Limited knowledge on how BI analytics can support production goals [R] Difficulties getting data that requires manual input [R] Lack of skills to use BI analytics in general [R] Lack of KPIs, with focus on more company specific KPIs [R] Lack of skills to see how BI can help increase revenue [R] Type of ownership, family owned companies seemed less interested in using BI for decision support [R] Risk of data overflow [R] Having correct data [R] Rely on gut-feeling rather than data</p>
5	<p>[R] Perceived ease of conventional business styles [R] Low awareness of open data benefits [R] Inadequate knowledge of data value and its potential [R] Limited financial resources [R] Shortage of available domain-specific data analysts [R] Inadequate up-to-date technical knowledge [R] Poor knowledge of available funds and data tools</p>
6	<p>[R] Complexity [R] Uncertainty and insecurity [I]Trialability [I]Observability [I]Top management support [I]Organizational readiness [I]External support</p>
7	<p>[R] High level of BD complexity [R] High level of BD insecurity [I] Relative advantage of BD [I]Top management support [I]Organizational readiness [I]Government regulation</p>
8	<p>[R] Complexity of data [R] Data scalability [R] Data silos [R] Infrastructure readiness [R] Lack of suitable software [R] Poor data quality [R] Ethical concerns in data use [R] Financial barriers [R] Inability to assess and address digital risks [R] Regulatory issues [R] Lack of common standards [R] Change management [R] Cultural barriers [R] Insufficient volumes of data to be analysed [R] Lack of business cases [R] Lack of in-house Data Analytics expertise [R] Lack of managerial awareness and skills [R] Talent management [R] Management of technology [R] Lack of top management support [R] Shortage of consultancy services</p>
9	<p>[R] Lack of understanding [R] Concerns over data security and privacy [I]Information quality [I]System quality</p>

10	<p>[R] Supply (lack of software editors specialized in SME, shortage in skilled employees on the job market and high cost of salaries) [R] Demand (increased competition and price due to high demand in service providers) [R] Technical (unsuitable IT infrastructure for Data Analytics adoption, lack of formal data management processes) [R] Managerial (lack of managerial support in Data Analytics adoption process, tensions between top management and employees) [R] Strategic (inability to identify strategic use cases for Data Analytics, misalignment between business strategy and IS) [R] Cultural (lack of organizational culture or leadership supporting Data Analytics use, lack of trust conservatism or skepticism against digital technologies) [R] Skills (lack of knowledge about big data and Data Analytics, lack of Data Analytics-skilled employees and professional training) [R] Financial (high cost of adopting data analytics and maintaining IS infrastructure) [R] Legal (lack of legal knowledge to comply with GDPR legislation)</p>
----	---

APÊNDICE B - FATORES MAPEADOS POR ESTUDO E AGRUPAMENTO

	Estudos									
Fatores	1	2	3	4	5	6	7	8	9	10
Software quality	Software errors	Lack of BI system flexibility				Trialability			System quality	Lack of software editors specialized in SME,
	Low speed of the product	"User friendly" BI system								
	Limited data export functionalities	The kind of BI technology and tools								
	Problematic conflation of data									
	Structural fixing									
Software security	Inadequate security function							Inability to assess and address digital risks	Concerns over data security and privacy	Lack of legal knowledge to comply with GDPR legislation

Estudos

Fatores	1	2	3	4	5	6	7	8	9	10
Software suitability		Suitability BI for users' business needs Integration BI system with operational systems	BIS is a part of ERP					Lack of suitable software		
System implementation price		The price of BI system and its implementation	Cost							High cost of salaries [for skilled employees] High cost of adopting data analytics and maintaining IT infrastructure
Management of technology								Management of technology		
Organizational culture			Organizational culture							

	Estudos									
Fatores	1	2	3	4	5	6	7	8	9	10
Observance of available budget/resources	Level of observance of budgets	Exceeded BI implementation budget			Limited financial resources Poor knowledge of available funds and data tools			Financial barriers		Increased competition and price due to high demand in service providers
Awareness of benefits		Lack of knowledge about BI system and its capabilities	Expected benefits Perception of strategic value	Limited knowledge on how BI analytics can support production goals Lack of skills to see how BI can help increase revenue	Low awareness of open data benefits Inadequate knowledge of data value and its potential	Observability	Relative advantage of BD			

Estudos

Fatores	1	2	3	4	5	6	7	8	9	10
Support for development	Insufficient support Employee training	Lack of manager's supporting relations between business vision and BI system	Management support	Current IT support did not fully support the processes of the area		Top management support	Top management support	Lack of top management support		Unsuitable infrastructure for Data Analytics adoption, Lack of formal data management processes) Lack of managerial support in Data Analytics adoption process, Tensions between top management and employees

	Estudos									
Fatores	1	2	3	4	5	6	7	8	9	10
Management interest				Limited interest from executives and owners on how to use BI as a decision support				Lack of managerial awareness and skills		Lack of organizational culture or leadership supporting Data Analytics use,
Successful project management		<p>Ineffective BI project management</p> <p>Insufficient responsiveness on changing users' requirements</p> <p>Uncoordinated BI project activities</p> <p>Different viewpoints on BI system implementation</p> <p>Changeability and developing BI system</p>		<p>Lack of KPIs, with focus on more company specific KPIs</p> <p>Risk of data overflow</p>				<p>Lack of common standards</p> <p>Change management</p>		Misalignment between business strategy and IS

	Estudos									
Fatores	1	2	3	4	5	6	7	8	9	10
Existence of references		Past cooperation with a BI supplier	Project champion							
Existence of external support			External support			External support				
Existence of external pressure							Government regulation	Regulatory issues		
Organizational readiness			Organizational readiness			Organizational readiness	Organizational readiness			
Existence of an area				When IT is outsourced, the companies have difficulties getting the support needed						

Estudos

Fatores	1	2	3	4	5	6	7	8	9	10
Employee qualification	Personnel using the BI solution are not qualified enough	Lack of well defined business problem No determining the expectation of BI user Misunderstanding of knowledge provided by BI Well defined business problem and processes Inadequate up-to-date technical knowledge		Lack of skills to use BI analytics in general				Ethical concerns in data use Talent management	Lack of understanding	Lack of knowledge about big data and Data Analytics, Lack of professional training

APÊNDICE C - FORMULÁRIO ESTADO DA PRÁTICA

Survey - Data Analytics em micro e pequenas empresas de software brasileiras

Esta pesquisa objetiva coletar dados de forma anônima acerca do uso de análise de dados em micro e pequenas organizações de software brasileiras, como parte de pesquisa na área de engenharia de software.

 O tempo de preenchimento deste questionário é de aproximadamente 5 minutos.

A sua participação nesta pesquisa consistirá em responder o questionário a seguir com base na sua experiência dentro da organização atual.

Esta pesquisa está sendo realizada pela graduanda Livia Silva Marques sob orientação do Prof. Dr. Jean Hauck do Curso de Graduação em Sistemas de Informação da Universidade Federal de Santa Catarina.

A participação é gratuita e voluntária.

Esta é uma pesquisa de opinião pública com participantes não identificados (conforme artigo 1º da Resolução CNS n.º 510 de 2016). Nesta pesquisa não são coletados quaisquer dados que permitam a identificação dos participantes.

 Qualquer dúvida, favor entre em contato via e-mail liviasmqs@gmail.com.

liviasmqs@gmail.com [Mudar de conta](#)



 Não compartilhado

*** Indica uma pergunta obrigatória**

Confirmo meu interesse em participar desta pesquisa tendo lido e aprovado o ***** Termo de Consentimento disponível em <https://bit.ly/3UbV7WS>

Sim

Não

[Próxima](#)

 Página 1 de 3

[Limpar formulário](#)

Survey - Data Analytics em micro e pequenas empresas de software brasileiras

liviasmqs@gmail.com [Mudar de conta](#)



Não compartilhado

* Indica uma pergunta obrigatória

Sobre a micro/pequena organização em que o processo ocorreu

5 perguntas de múltipla escolha e 1 aberta sobre as características da empresa.

Qual é a sua experiência com Data Analytics? *

- Menos de 1 ano
- Entre 1 e 5 anos
- Entre 5 e 10 anos
- Mais de 10 anos

Qual a quantidade de funcionários da organização? *

Responda considerando a quantidade de funcionários que a organização tinha no momento em que implementavam o processo de Data Analytics em questão - não necessariamente a quantidade atual.

- Menos de 10
- Entre 10 e 99
- Entre 100 e 500
- Mais de 500

Em que região do Brasil está situada a organização? *

Responda considerando a organização em que o processo de Data Analytics em questão foi implementado - não necessariamente a sua empresa atual.

Escolher

Quantos anos de atividade a organização tem/tinha? *

Responda considerando os anos de atividade que a organização tinha no momento em que implementavam o processo de Data Analytics em questão - não necessariamente a idade atual.

- Entre 0 e 2
- Entre 2 e 5
- Entre 5 e 10
- Mais de 10

Qual é/era seu cargo na organização? *

Responda considerando o seu cargo na organização tinha no momento em que implementavam o processo de Data Analytics em questão - não necessariamente o seu cargo atual atual.

- CEO
- CTO
- CDO
- CIO
- Engenheiro/Analista/Cientista/Arquiteto de dados
- Gerente de projetos de dados
- Desenvolvedor de Software
- Outro: _____

- CTO
- CDO
- CIO
- Engenheiro/Analista/Cientista/Arquiteto de dados
- Gerente de projetos de dados
- Desenvolvedor de Software
- Outro: _____

Qual o segmento de atuação da organização? *

Em caso de desenvolvimento de software, para qual área se desenvolve?

- Consultoria (inclui serviços de consultoria e aconselhamento relativos à Tecnologia da Informação)
- Serviços (empresas da área de saúde, transportes, educação, turismo, entretenimento e demais serviços)
- Indústria (inclui todas as empresas de manufatura, seja de manufatura discreta ou de transformação)
- Comércio (Inclui todas as empresas que atuem no comércio, varejo ou distribuição)
- Agroindústria (inclui empresas fabricantes de maquinário agrícola, adubos e fertilizantes, usinas e cooperativas, empresas de trading, exportação de carnes e processamento de alimentos)
- Governo (instituição voltada à administração pública, seja em nível municipal, estadual ou federal)
- Finanças (inclui empresas públicas e privadas, bancos, empresas de seguros, cartões de crédito, corretora de valores, etc.)
- Óleo e Gás (inclui empresas públicas e privadas relacionadas ao setor de óleo, gás e mineração)
- Outro: _____

[Voltar](#)

[Próxima](#)

 Página 2 de 3 [Limpar formulário](#)

Sobre o processo de Data Analytics na organização

4 perguntas de múltipla escolha e 1 pergunta aberta sobre como se dá o processo de Data Analytics e seus impactos.

A organização possui/possuía pelo menos um responsável por análise de dados? *

Responda considerando o cenário da organização no momento em que implementavam o processo de Data Analytics em questão.

- Sim, temos mais de 1 pessoa com foco em dados
- Sim, temos 1 pessoa com foco em dados
- Sim, informalmente - não é parte de suas funções formais, mas assume esse papel
- Não, não há nenhuma referência em dados

Quais fases do processo de Data Analytics foram efetivamente executadas pela organização? *

Considere que o processo de Data Analytics compreende todas as fases descritas, mas nem sempre todas elas são executadas, e eventualmente podem até acontecer em uma ordem diferente ou com outros nomes.

	Não executa	Executa parcialmente	Executa completamente
Entendimento de negócio	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Entendimento dos dados disponíveis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Preparação dos dados (tratamento)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Modelagem de dados (aplicação de técnicas e/ou algoritmos de modelagem de dados)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

técnicas e/ou algoritmos de modelagem de dados)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Avaliação dos dados modelados	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Lançamento (entrega de dados gerando valor)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Marque todos os fatores que, de acordo com a sua experiência na organização, *
você considera críticos para o sucesso do processo de Data Analytics

Mesmo que todos os itens pareçam relevantes, considere a sua própria experiência para selecionar aqueles essenciais para que o processo de Data Analytics possa existir eficientemente. Marque tanto os fatores que fizeram o seu processo ter algum nível de sucesso quanto os que, devido a sua falta, impossibilitaram que isso acontecesse.

- Acesso a software de qualidade
- Suporte técnico para implementação
- Suporte organizacional (da gerência) para implementação
- Um bom gerenciamento do processo de dados
- Orçamento adequado disponível
- Todos os impactados terem conhecimento sobre os benefícios
- Uma boa etapa de preparação de dados
- Clareza sobre necessidade do processo
- Estrutura para qualificação dos funcionários
- Equipe com confiança em resultados apresentados por meio de dados
- Existência de profissionais qualificados (com experiência) em dados
- Inexistência de uma resistência psicológica quanto ao uso de dados
- Inexistência da crença de que o tópico é complexo demais
- Outro: _____

Qual metodologia foi/é utilizada pela organização para implementar o processo de Data Analytics? *

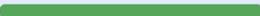
- Nenhuma metodologia específica foi utilizada como base
- CRISP-DM
- SEMMA
- KDD
- Outro: _____

Quais impactos (positivos ou negativos) foram observados na organização com a implementação do processo de Data Analytics? *

Sua resposta _____

[Voltar](#)

[Enviar](#)

 Página 3 de 3

[Limpar formulário](#)

APÊNDICE D - SPREADSHEET DO AGRUPAMENTO DE IMPACTOS

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Quais impactos (positivos ou negativos) foram observados na organização com a implementação do processo de Data Analytics?												
		Outros	Tomada de decisão	Identificar pontos de atenção	Entendimento de negócio	Identificar oportunidades	Eficiência operacional	Negativos	Redução de custos	Confiança	Geração de valor	Alinhamento com objetivos estratégicos	
2	Fundamentação das pesquisas e melhoria nos processos.	Fundamentação das pesquisas e melhoria nos processos.	Impactos relevantes no dia-a-dia que dão suporte a tomadas de decisão com maior embasamento.	Identificação dos pontos de atenção em relação ao negócio. Áreas que precisam ser revistas para garantir o crescimento sustentável da empresa.	Maior autonomia para o time utilizar dados, gerando maior entendimento do negócio.	- Identificação de novas oportunidades de mercado, resultando no reajuste de preços de alguns de nossos produtos para abranger um público maior e assim captar mais recursos	Esse esforço tem gerado um grande ganho em eficiência operacional	Como ponto negativo, a falta de processos rotineiros e de entendimento da importância de se seguir este trabalho fez com que a análise de dados não fizesse parte do dia a dia das áreas.	Em muitos cenários, as empresas conseguem reduzir custos ao ter uma visão completa dos indicadores.	Entrega de confiança nas projeções e processos.	Geração de produtos de dados de valor.	Clareza na interpretação dos dados mantém equipe alinhada aos objetivos corporativos	
3	Impactos relevantes no dia-a-dia que dão suporte a tomadas de decisão com maior embasamento. Direcionamento do desenvolvimento do produto com base em dados coletados. Maior autonomia para o time utilizar dados, gerando maior entendimento do negócio.	Ampliação de dados de empresas e clientes através de modelos de predição	Direcionamento do desenvolvimento do produto com base em dados coletados.	O compartilhamento dos dados fez com que toda empresa adquirisse conhecimento em relação a qual momento estamos e o que precisa ser feito a partir dali, o que aumentou o engajamento e o trabalho em equipe.	Maior conscientização da equipe, que passou a ter uma visão mais clara quanto aos impactos gerados a cada lançamento de novas versões do produto.	Fornecer sugestões de melhorias no processo.	maior eficiência operacional	o custo de implementação em casos nos quais a empresa não está preparada para usufruir da forma correta dos dados.	Limitação dos investimentos.	Confiabilidade dos indicadores	Criação de valor com dados disponíveis e subutilizados.	Alinhamento dos objetivos estratégicos da empresa (OKRs vs KPIs)	
4	De forma geral foram observados impactos primariamente positivos, incluindo: - Maior conscientização da equipe, que passou a ter uma visão mais clara quanto aos impactos gerados a cada lançamento de novas versões do produto. - Tomadas decisões estratégicas com base nos resultados históricos e projeções realizadas - Identificação de novas oportunidades de mercado, resultando no reajuste de preços de alguns de nossos produtos para abranger um público maior e assim captar mais recursos	Melhorias no processo de desenvolvimento de software	Tomadas decisões estratégicas com base nos resultados históricos e projeções realizadas	O processo ajudou / ajuda ao facilitar uma visão mais ampla dos problemas / questões	Melhor clareza da saúde financeira da empresa	enxergarem indicadores potenciais para aumento de receita e engajamento,	otimiza processos operacionais	As reuniões sem o DA parece não ter eficácia.	redução de custos	Informações oficiais validadas prontas para o consumo.	Agregação de valor com base nos dados		
5	Identificação dos pontos de atenção em relação ao negócio. Áreas que precisam ser revistas para garantir o crescimento sustentável da empresa. O compartilhamento dos dados fez com que toda empresa adquirisse conhecimento em relação a qual momento estamos e o que precisa ser feito a partir dali, o que aumentou o engajamento e o trabalho em equipe. Como ponto negativo, a falta de processos rotineiros e de entendimento da importância de se seguir este trabalho fez com que a análise de dados não fizesse parte do dia a dia das áreas.	Reestruturação de processos internos	embasamento comprobatório a respeito das ações que o time precisa tomar para se tornar melhor (mais ágil)	e embasamento para pontos de atenção	[embasamento para] pontos prósperos de negócios	identificação de novas oportunidades de negócios.	e a empresa ganhou com um aumento de produtividade geral e alinhado com a estratégia.	necessidade de investir em infraestrutura e treinamento, além de questões relacionadas à privacidade de dados.	Custo, tempo de construção da solução e dificuldade na precificação para os clientes	Agora é certeza que o número é correto e único	Geração de valor inovando processos		

Acesso em:

<https://docs.google.com/spreadsheets/d/1gHylgcz1I5v9x7ZDaZomztcNARM9ZfXOcByEh5vIkJg/edit?usp=sharing>

APENDICE E - GOOGLE FORMS COM AVALIAÇÃO DO GUIA

Avaliação do Guia para implementação de Data Analytics em micro e pequenas empresas brasileiras de software

Após analisar o guia, responda as questões a seguir em relação a sua abrangência, aplicabilidade e cobertura. Depois, fique à vontade para deixar suas percepções sobre erros/falhas no guia e comentários gerais nas questões abertas.

Abrangência: utilidade do conteúdo proposto em múltiplos contextos

Aplicabilidade: qualidade e utilidade prática do conteúdo do guia

Cobertura: extensão do conteúdo em relação ao tema

liviasmqs@gmail.com [Mudar de conta](#)



Não compartilhado

* Indica uma pergunta obrigatória

Qual a sua formação? *

Sua resposta

Qual o seu cargo atual e como se relaciona com o contexto de dados? *

Sua resposta

Quanto tempo de experiência você possui na área de dados? *

Sua resposta

Sobre a **abrangência** do guia: *

Discordo
totalmente

Discordo

Concordo

Concordo
totalmente

Compartilhar

Sobre a **abrangência** do guia: *

	Discordo totalmente	Discordo	Concordo	Concordo totalmente
O guia pode ser aplicado em organizações de diferentes setores (desde que tenha uma área de software)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O Guia pode ser adaptável a diferentes métodos ágeis escolhidos para o processo de data analytics	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O Guia pode ser adaptado a organizações de diferentes tamanhos que trabalham com desenvolvimento de software	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Sobre a **aplicabilidade** do guia: *

	Discordo totalmente	Discordo	Concordo	Concordo totalmente
O Guia aborda práticas que podem realmente ser observadas na prática	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O Guia aborda orientações e conteúdos que podem efetivamente ser	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

O Guia aborda práticas que podem realmente ser observadas na prática

O Guia aborda orientações e conteúdos que podem efetivamente ser aplicados

Sobre a **cobertura** do guia: *

Discordo totalmente

Discordo

Concordo

Concordo totalmente

A seção de descrição fornece embasamento necessário para entendimento do conteúdo do Guia

O conteúdo do Guia proposto para gestão de riscos está alinhado aos princípios e valores ágeis

O Guia apresenta atividades, técnicas, ferramentas e demais componentes suficientes para sua aplicação prática

Caso tenha encontrado, liste elementos faltantes ou incorretos encontrados no Guia.

Sua resposta

Caso deseje, utilize esse espaço para deixar considerações/contribuições gerais relacionadas ao Guia.

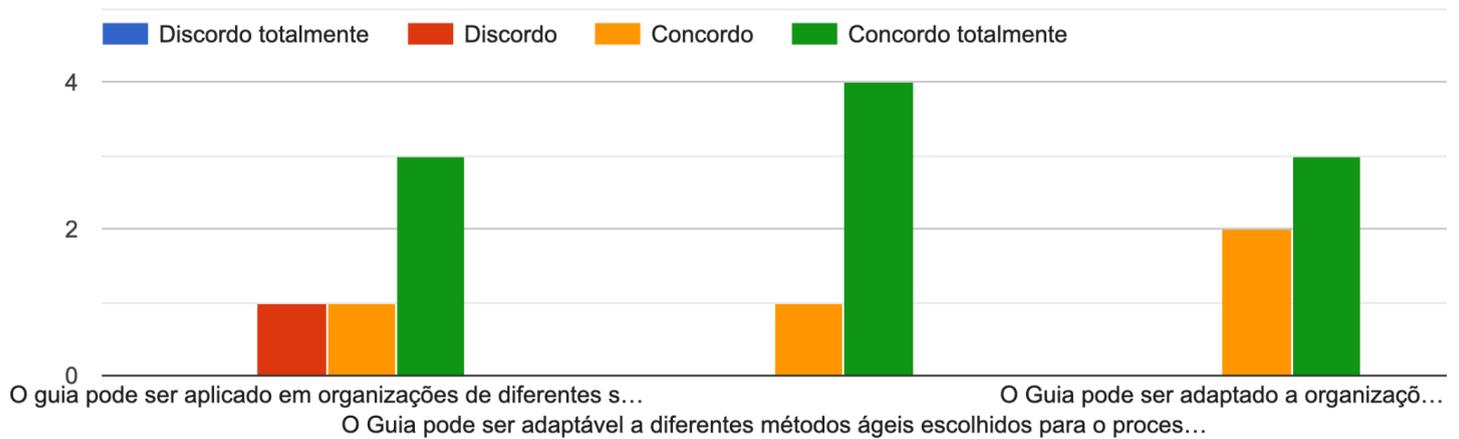
Sua resposta

Enviar

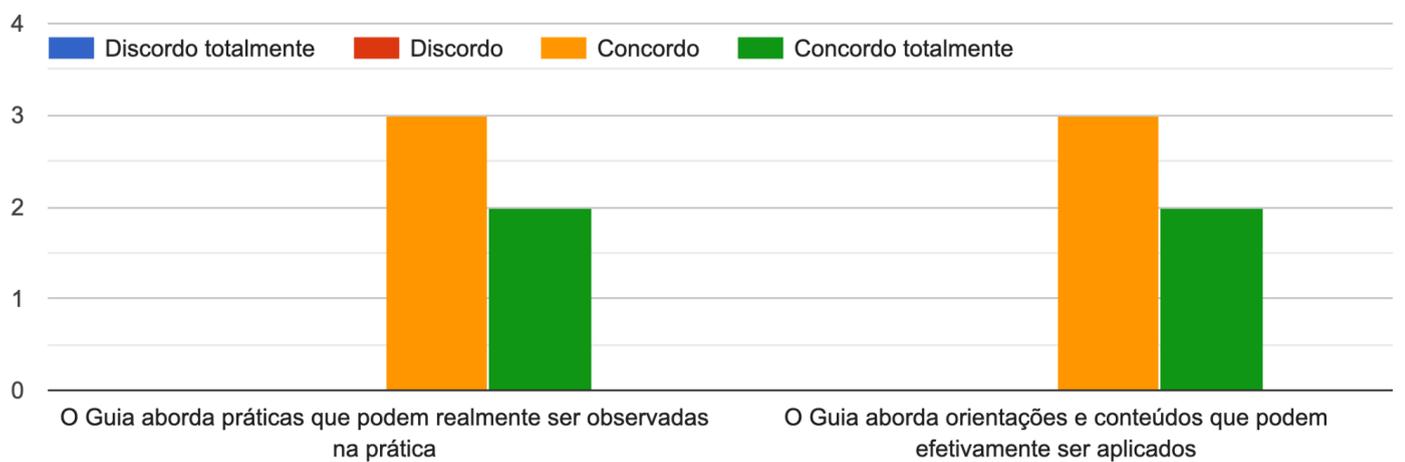
[Limpar formulário](#)

APÊNDICE F - GRÁFICOS DE RESPOSTAS DO PAINEL DE ESPECIALISTAS

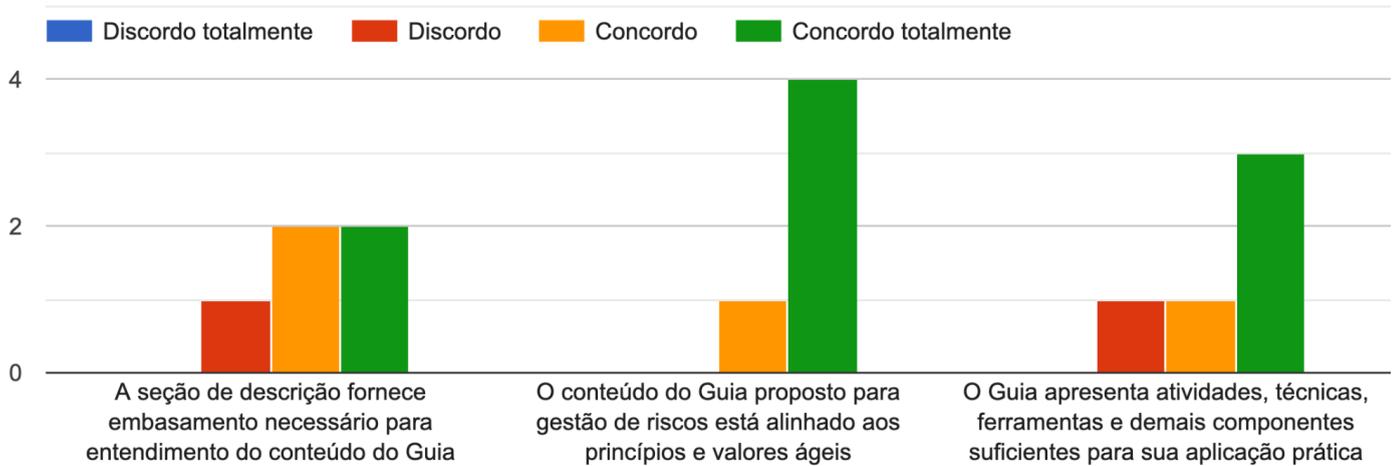
Sobre a abrangência do guia:



Sobre a aplicabilidade do guia:



Sobre a cobertura do guia:



Caso tenha encontrado, liste elementos faltantes ou incorretos encontrados no Guia.

3 respostas

Eu fiquei um pouco curioso para entender melhor quais são esses frameworks de colaboração ágil.

Melhor descrição sobre os processos necessários para aplicação do método

Considerando que a informática enquanto ciência está alicerçada em 3 pilares básicos, software, hardware e recursos humanos, seria interessante observar a possibilidade de incrementar a seção "Qualificação dos profissionais" - "Percebe-se habilidade do time para compartilhar informação e coordenar atividades", o fato de inserir esta parte: "Oportunidades para desenvolvimento de competências são definidas e oferecidas" neste ponto da seção, pode minimizar a importância do processo contínuo e necessário do desenvolvimento de competências. Sugiro pensar na possibilidade de abrir nesta seção um item que aborde o processo de capacitação profissional e talvez identificar entre outras coisas se isto é realizado periodicamente.

Na seção "Clareza sobre objetivos - Métricas definidas", sugiro rever o texto e incluir o termo "suficientes", portanto, ficaria assim "Dados objetivos e/ou subjetivos são identificados e suficientes para associar às métricas e respondê-las em uma forma quantitativa";

Na seção "Preparação de dados" seria interessante rever o nome da seção para "Tratamento dos dados", isto porque a mesma não aborda apenas as etapas preliminares dos dados mas também a formatação e integração.

Caso deseje, utilize esse espaço para deixar considerações/contribuições gerais relacionadas ao Guia.

2 respostas

No geral, ficou muito bom. Parabéns!

O guia tem uma linguagem acessível e uma metodologia de fácil compreensão. O fato de permitir um resultado com visualização gráfica do levantamento feito, facilita muito a "leitura" do cenário e o "start" de providências e encaminhamentos necessários.

APÊNDICE G - ARTIGO ACADÊMICO

Proposta de Guia para Iniciativas de Data Analytics em Micro e Pequenas Empresas Brasileiras

Lívia Silva Marques¹

¹Departamento de Informática e Estatística – Universidade Federal de Santa Catarina (UFSC) Florianópolis – SC – Brazil

livia.marques@grad.ufsc.br

Abstract. *This study identifies critical success factors for implementing Data Analytics in Brazilian micro and small software companies. Through a systematic literature review and a survey, five key factors were prioritized: data process management, data preparation, process clarity, professional qualification, and organizational support. A guide based on the Essence approach was developed to provide best practices, self-assessment, and progress monitoring. The findings emphasize the importance of Data Analytics for organizational success and improved process execution.*

Resumo. *Este estudo investiga fatores críticos para implementar Data Analytics em micro e pequenas empresas brasileiras de software, destacando cinco principais: gerenciamento de dados, preparação dos dados, clareza do processo, qualificação profissional e suporte organizacional. Com base nisso, foi desenvolvido um guia online utilizando a abordagem Essence, oferecendo boas práticas, autoavaliação e monitoramento. Os resultados reforçam a importância do Data Analytics para o sucesso organizacional e a execução eficiente dos processos.*

1. Introdução

A transformação digital tem marcado os últimos anos, impulsionada pelo crescente impacto da ciência e análise de dados. O avanço tecnológico e o uso de dispositivos conectados à internet têm possibilitado a geração de grandes volumes de dados, permitindo às organizações aplicar análises e obter informações estratégicas para tomada de decisão e melhora no desempenho (Baijens et al., 2020).

O Business Intelligence (BI) destaca-se como a principal ferramenta para aplicações analíticas, especialmente nos últimos anos (Bhatiasevi e Naglis, 2018). Estudos mostram que a adoção de BI traz benefícios significativos, como vantagem competitiva e maior performance organizacional (Owusu, 2017; Anwar et al., 2018).

Pequenas e médias empresas (SMEs) têm crescido expressivamente, representando uma parte importante da economia global, tanto em geração de empregos quanto em receita (Bhardwaj, 2022). Apesar de sua relevância, muitas SMEs enfrentam dificuldades e falham anualmente devido à falta de uso estratégico de recursos, como dados (Tarek & Adel, 2016; Wang & Wang, 2020).

A competitividade crescente pressiona SMEs a adotar BI e Data Analytics para acompanhar seus concorrentes (Bhatiasevi e Naglis, 2018). Embora exista uma relação comprovada entre o uso de análise de dados e a melhora no desempenho empresarial

(O'Connor & Kelly, 2017), muitos negócios ainda resistem a investir nessa prática, seja por desconhecimento ou outras barreiras (Bhardwaj, 2022).

Este estudo propõe um mapeamento sistemático da literatura para identificar os principais fatores críticos na implementação de Data Analytics em SMEs. Além disso, busca compreender os desafios enfrentados por pequenas empresas brasileiras de software por meio de um questionário estruturado. O objetivo final é desenvolver um guia que auxilie essas empresas na adoção bem-sucedida de iniciativas de Data Analytics.

1.1. Método de pesquisa

Este trabalho foi dividido em duas fases: a primeira, inicial e teórica, traz a conceituação do problema de pesquisa e uma coleta de dados referente ao estudo do Estado da Arte. A fase seguinte, de caráter mais prático, traz um *Survey* (Pinsonneault e Kraemer, 1993), que identifica os principais fatores influentes no processo de Data Analytics em micro e pequenas empresas brasileiras de software, além da elaboração de um Guia, que pretende auxiliar as organizações a evoluírem nesses fatores com base em práticas reconhecidas pela literatura.

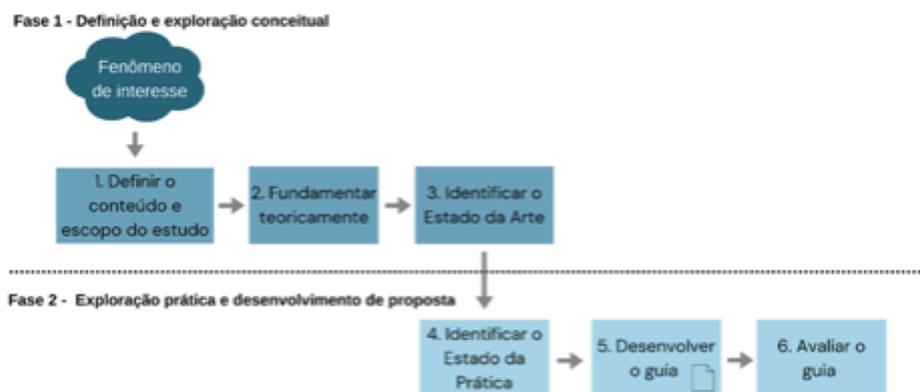


Figura 1. Metodologia de pesquisa

A Figura 1, acima, apresenta o método de pesquisa detalhado nas atividades que o compõem.

1.1. O processo de data analytics

Gökalp (2021) identifica "Data Analytics" como sendo uma área que compreende toda a pipeline do processo de dados, necessária para gerar ideias de ações a partir de dados brutos. Segundo o autor, existem frameworks conhecidos em Data Analytics que englobam, de forma geral, seis processos: entendimento de negócio, entendimento de dados, preparação de dados, construção de modelo, avaliação e lançamento.

Holtz (2024, p.02) diz que "equipes de ciência de dados que combinam uma implementação flexível do CRISP-DM com abordagens abrangentes de gerenciamento de projetos ágeis baseadas em equipe provavelmente obterão os melhores resultados",

sugerindo que uma combinação de modelos com metodologias de gerenciamento de projetos ágeis pode ser ideal, a fim de juntar aspectos próprios do processo de Data Analytics com práticas que levem a uma execução eficiente.

1.2. Forças impulsionadoras e restritivas

Rockart (1979) apresenta fatores críticos de sucesso como o conjunto de questões importantes a que se deve estar atento para obtenção de êxito. Nas organizações, a análise de campo de forças de Lewin é usada para identificar as forças impulsionadoras e as forças contrárias a determinada mudança proposta. Com isso, é possível criar um plano de ação para restringir os fatores contrários e investir nos impulsionadores.

Trazendo os conceitos de forças impulsionadoras e forças contrárias para o contexto da implementação de processos de análise de dados, alguns autores, como Bhardwaj (2022), apresentam pesquisas sobre o tema conforme revisão sistemática da literatura. Os fatores encontrados são categorizados em grandes grupos que podem ter certa sobreposição, mas ainda assim facilitam a visualização.

1.3. Essence

O Essence é um padrão para criação, uso e melhorias de práticas e métodos de engenharia de software que se propõe a focar apenas nas informações essenciais para otimizar os processos de engenharia, fornecendo práticas leves e adaptáveis para cada contexto (OMG, 2018).

A Linguagem apresentada pelo Essence descreve uma forma para escrita de práticas e elementos do processo de software. Por ser apresentada de maneira simples, padronizada e com elementos independentes, ela pode ser facilmente entendida e compartilhada, além de permitir diferentes combinações de seus elementos para caracterizar os processos de uma organização de forma customizada, e não como um framework fechado e inflexível.

Os elementos chave do Essence considerados relevantes para esse trabalho são apresentados no Quadro 1: alphas, estados do alpha, *patterns* e recursos.

Símbolo identificador	Elemento	Descrição
	Alpha	Aspecto chave ou elemento do qual se quer monitorar o progresso; o resultado em que se quer manter o foco.
	Estados do Alpha	Especificação do estado de progresso de um Alpha. Podem ser quebrados em uma lista de itens a serem alcançados para passar ao próximo estado (ou verificar o estado atual).
	<i>Pattern</i>	Estrutura genérica existente para permitir descrever algum elemento adicional que se relaciona com outro elemento Essence.
	Recursos	Fonte de informação ou conteúdo (como site ou livro) que serve de apoio para aprofundamento e implementação.

Quadro 1. Alguns elementos chave do Essence

2. Estado da Arte

Foi analisado o estado da arte quanto aos fatores que influenciam no sucesso de iniciativas de Data Analytics. Assim, para reunir esses estudos e apresentar um panorama geral do que já foi abordado pela literatura, realizou-se um Mapeamento Sistemático da Literatura (MSL), com base em Petersen et al., 2008.

A partir do MSL, foram identificados 10 artigos relevantes: Scholz et al. (2010), Olszak e Ziemba (2012), Puklavec, Oliveira e Popovič (2014), Gudfinnsson e Strand (2017), Mohamed e Weber (2020), Maroufkhani et al. (2020), Alsyof et al. (2022), Willetts, Atkins e Stanier (2022), Perdana et al. (2022) e Justy et al. (2023).

De modo geral, os fatores relacionados à qualidade de software, à existência de suporte para implantação do sistema de analytics, à qualificação de funcionários e à qualidade do gerenciamento do projeto de dados demonstraram ser os mais relevantes. Além disso, considerou-se que há pouco contexto nos estudos em relação à metodologias utilizadas ou nos critérios para considerar que a implementação dos processos teve sucesso ou não.

3. Estado da Prática

Para entender como as empresas de desenvolvimento de software, no Brasil, experienciam o processo de Data Analytics e a fim de corroborar com as descobertas do estado da arte, foi realizado um *survey* com profissionais de pequenas e médias empresas de desenvolvimento de software, seguindo atividades definidas por Pinsonneault e Kraemer (1993).

Ao total, o estudo teve 133 respondentes e o resultado evidenciou os fatores mais populares entre os respondentes. Os fatores que mais de 50% da amostra considerou como mais importante foram cinco: “Um bom gerenciamento do processo de dados”, “Uma boa etapa de preparação de dados”, “Clareza sobre a necessidade do processo”, “Existência de profissionais qualificados (com experiência em dados)”, “Suporte organizacional (da gerência) para implementação”.

De modo geral, os resultados do mapeamento dos fatores relevantes no Estado da Arte e no Estado da Prática convergiram. Especialmente se considerarmos que os fatores são provenientes de agrupamentos - que podem carregar aspectos relacionados paralelos - pode-se dizer que a maioria dos 5 fatores identificados pelo estado da prática estão entre os 5 principais fatores mapeados pelo estado da arte.

Outro aspecto indicado pelos resultados do *survey* é a predominância do uso de uma abordagem *ad-hoc* quanto às metodologias/abordagens utilizadas para implementação do processo da Data Analytics, seguindo a tendência identificada nos estudos do Estado da Arte e indicando que o processo de Data Analytics ainda não foi contemplado com uma abordagem validada que o atenda completamente.

4. Guia para implementação do processo de data analytics

A construção do Guia é baseada na abordagem Essence e é cada um dos cinco fatores mais relevantes identificados no levantamento do estado da prática (seção 4.4.2) é definido como um *Alpha* do Essence, ou seja, um fator essencial cuja evolução deve ser monitorada pela organização.

Cada Alpha possui estados, compostos por um checklist que marca o progresso no Alpha. Além disso, recursos externos auxiliam na evolução do Alpha. Esses conteúdos foram buscados em práticas já consolidadas na literatura, como a norma ISO 15504 (2012) e o People Capability Maturity Model (2001).

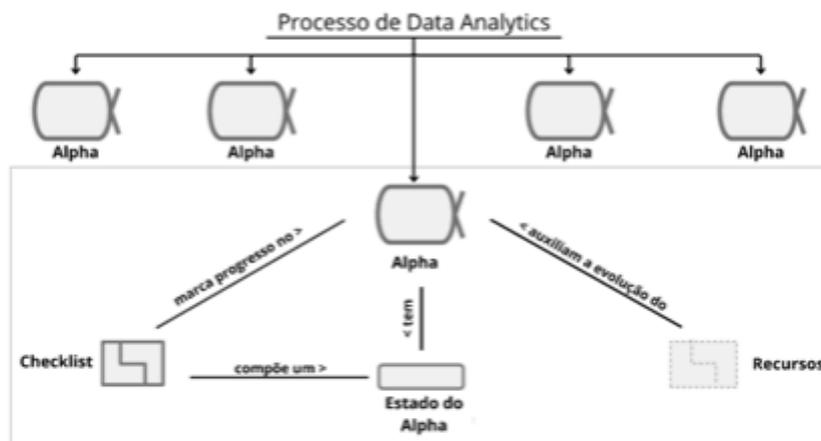


Figura 2. Estrutura do Guia

A partir do guia, foi desenvolvida uma ferramenta de autoavaliação para que as organizações possam, dinamicamente, identificar seu status atual em relação aos 5 fatores críticos e monitorar seu progresso. A ferramenta pode ser acessada pelo link de compartilhamento do template, que já leva o usuário à criação da sua cópia pessoal: [Guia para implementação de Data Analytics em micro e pequenas empresas brasileiras de software](#).

A partir do gráfico gerado, a organização tem uma representação visual do estágio em que se encontra em cada área. Também foram adicionados links para os recursos externos de apoio à evolução de cada Alpha (todos usados como referência para a construção do Guia), para que as organizações possam conhecer a teoria, visando a evolução de suas práticas.

5. Conclusão

Considera-se que o objetivo geral deste trabalho foi alcançado, gerando o guia que serve de contribuição tanto para o público alvo do estudo - micro e pequenas organizações brasileiras de software - quanto para a área de pesquisa em que se encontra.

As análises realizadas neste trabalho indicaram algumas ideias sobre o processo de Data Analytics, como a necessidade de investigar o uso de frameworks para

gerenciamento do processo de Data Analytics, e a dificuldade em mensurar resultados do processo em questão.

5.1. Trabalhos futuros

Vários espaços de pesquisa se mostram possíveis a partir desse estudo:

- a avaliação prática do Guia gerado
- a evolução do Guia com mais ferramentas fornecidas pela abordagem Essence
- um recorte do estudo para considerar aspectos geográficos
- uma análise temporal que identifique a mudança dos fatores ao longo do tempo
- a análise desses fatores em organizações que não são de software, considerando o impacto das micro e pequenas empresas brasileiras de diversos setores

Referências

- BAIJENS, Jeroen; HELMS, Remko; KUSTERS, Rob. Data Analytics Project Methodologies: Which One to Choose?. In: Proceedings of the 2020 International Conference on Big Data in Management. 2020. p. 41-47. Disponível em <https://dl.acm.org/doi/10.1145/3437075.3437087>. Acesso em 5 de novembro de 2023.
- BHATIASEVI, Veera; NAGLIS, Michael. Elucidating the determinants of business intelligence adoption and organizational performance. *Information development*, v. 36, n. 1, p. 78-96, 2020. Disponível em <https://doi.org/10.1177/0266666918811394>. Acesso em 9 de novembro de 2023.
- OWUSU, Acheampong. Business intelligence systems and bank performance in Ghana: The balanced scorecard approach. *Cogent Business & Management*, v. 4, n. 1, p. 1364056, 2017. <https://doi.org/10.1080/23311975.2017.1364056>.
- BHARDWAJ, Shikha. Data analytics in small and medium enterprises (SME): a systematic review and future research directions. *Information Resources Management Journal (IRMJ)*, v. 35, n. 2, p. 1-18, 2022. 10.4018/IRMJ.291691.
- TAREK, Bel Hadj; ADEL, Ghodbane. Business intelligence versus entrepreneurial competitive intelligence and international competitiveness of North African SMEs. *Journal of International Entrepreneurship*, v. 14, p. 539-561, 2016. doi:10.1007/s10843-016-0194-8.
- PINSONNEAULT, Alain and KRAEMER, Kenneth. 1993. Survey Research Methodology in Management Information Systems: An Assessment. *Journal of Management Information Systems* 10, 2 (1993), 75–105.
- WANG, Shouhong; WANG, Hai. Big data for small and medium-sized enterprises (SME): a knowledge management model. *Journal of Knowledge Management*, v. 24, n. 4, p. 881-897, 2020. doi:10.1108/JKM-02-2020-0081.
- O'CONNOR, Christina; KELLY, Stephen. Facilitating knowledge management through filtered big data: SME competitiveness in an agri-food sector. *Journal of Knowledge Management*, v. 21, n. 1, p. 156-179, 2017. doi:10.1108/JKM-08-2016-0357.

- GÖKALP, Mert Onuralp. Data Science Capability Maturity Model. 2021.
- HOLTZ, Nick. What is CRISP DM? 28 abr. 2024. Disponível em: <https://www.datascience-pm.com/crisp-dm-2/>. Acesso em: 2 jul. 2024.
- ROCKART, J.F., 1979, "Chief executives define their own data needs", Harvard Business Review, v. 2, pp. 81-93.
- OMG - OBJECT MANAGEMENT GROUP. Essence- Kernel and Language for Software Engineering Methods. Version 1.2, outubro 2018. Disponível em <https://www.omg.org/spec/Essence/1.2/About-Essence>.
- Scholz, Patrick & Schieder, Christian & Kurze, Christian & Gluchowski, Peter & Böhringer, Martin. (2010). Benefits and Challenges of Business Intelligence Adoption in Small and Medium-Sized Enterprises. 18th European Conference on Information Systems, ECIS 2010.
- OLSZAK, Celina M.; ZIEMBA, Ewa. Critical success factors for implementing business intelligence systems in small and medium enterprises on the example of upper Silesia, Poland. *Interdisciplinary Journal of Information, Knowledge, and Management*, v. 7, p. 129-150, 2012. DOI: 10.28945/1584. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84862019547&doi=10.28945%2f1584&partnerID=40&md5=766ffe1471dfd5f9b6d74eb86b0ac340>.
- Puklavec, B., Oliveira, T., & Popovič, A. (2014). Unpacking Business Intelligence Systems Adoption Determinants: An Exploratory Study of Small and Medium Enterprises. *Economic and Business Review*, 16(2). <https://doi.org/10.15458/2335-4216.1278>
- K. Gudfinnsson and M. Strand, "Challenges with BI adoption in SMEs," 2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA), Larnaca, Cyprus, 2017, pp. 1-6, doi: 10.1109/IISA.2017.8316407.
- M. Mohamed and P. Weber, "Trends of digitalization and adoption of big data & analytics among UK SMEs: Analysis and lessons drawn from a case study of 53 SMEs," 2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Cardiff, UK, 2020, pp. 1-6, doi: 10.1109/ICE/ITMC49519.2020.9198545.
- Parisa Maroufkhani, Ming-Lang Tseng, Mohammad Iranmanesh, Wan Khairuzzaman Wan Ismail, Haliyana Khalid, Big Data Analytics adoption: Determinants and performances among small to medium-sized enterprises, *International Journal of Information Management*, Volume 54, 2020, 102190, ISSN 0268-4012, <https://doi.org/10.1016/j.ijinfomgt.2020.102190>. (<https://www.sciencedirect.com/science/article/pii/S026840122030178X>).
- Alsyouf, Adi & Lutfi, Abdalwali & Almaiah, Drmohammed & Alrawad, Mahmaod. (2022). Factors Influencing the Adoption of Big Data Analytics in the Digital Transformation Era: Case Study of Jordanian SMEs. *Sustainability*. 14. 10.3390/su14031802.

Willets, Matthew & Atkins, Anthony & Stanier, C.. (2022). Quantitative Study on Barriers of Adopting Big Data Analytics for UK and Eire SMEs. 10.1007/978-981-16-2937-2_23.

Arif Perdana, Hwee Hoon Lee, SzeKee Koh, Desi Arisandi, Data analytics in small and mid-size enterprises: Enablers and inhibitors for business value and firm performance, *International Journal of Accounting Information Systems*, Volume 44, 2022, 100547, ISSN 1467-0895, <https://doi.org/10.1016/j.accinf.2021.100547>. (<https://www.sciencedirect.com/science/article/pii/S146708952100049X>)

Théo Justy, Estelle Pellegrin-Boucher, Denis Lescop, Julien Granata, Shivam Gupta, On the edge of Big Data: Drivers and barriers to Data Analytics adoption in SMEs, *Technovation*, Volume 127, 2023, 102850, ISSN 0166-4972, <https://doi.org/10.1016/j.technovation.2023.102850>. (<https://www.sciencedirect.com/science/article/pii/S016649722300161X>)