UNIVERSIDADE ESTADUAL DO MARANHÃO
GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO

MURILO CASTELO BRANCO ARAÚJO

A IMPORTÂNCIA DA APLICAÇÃO DE MÉTODOS MULTICRITÉRIOS EM SISTEMAS PARA MELHORIA DA SEGURANÇA DA INFORMAÇÃO E DA TOMADA DE DECISÃO NO AMBIENTE ORGANIZACIONAL DO BIG DATA

SÃO LUÍS - MARANHÃO
2023

MURILO CASTELO BRANCO ARAÚJO

A IMPORTÂNCIA DA APLICAÇÃO DE MÉTODOS MULTICRITÉRIOS EM SISTEMAS PARA MELHORIA DA SEGURANÇA DA INFORMAÇÃO E DA TOMADA DE DECISÃO NO AMBIENTE ORGANIZACIONAL DO BIG DATA

Monografia apresentada à Universidade Estadual do Maranhão como requisito parcial à obtenção do título de Bacharelado em Engenharia de Computação.

Orientador: Prof. Me. Raimundo de Carvalho Silva Neto.

SÃO LUÍS - MARANHÃO

2023

Araújo, Murilo Castelo Branco.
A importância da aplicação de métodos multicritérios em sistemas para melhoria da segurança da informação e da tomada de decisão no ambiente organizacional do big data / Murilo Castelo Branco Araújo. – São Luís, 2023.
69 f

Monografia (Graduação em Engenharia da Computação) - Universidade Estadual do Maranhão, 2023.

Orientador: Prof. Me. Raimundo de Carvalho Silva Neto.

1.Análises multicritério. 2.Big data. 3.Segurança da informação. 4.Tomada de decisão. I.Título.

CDU: 004.62:004.056

Elaborado por Giselle Frazão Tavares - CRB 13/665

MURILO CASTELO BRANCO ARAÚJO

A IMPORTÂNCIA DA APLICAÇÃO DE MÉTODOS MULTICRITÉRIOS EM SISTEMAS PARA MELHORIA DA SEGURANÇA DA INFORMAÇÃO E DA TOMADA DE DECISÃO NO AMBIENTE ORGANIZACIONAL DO BIG DATA

Monografia apresentada à Universidade Estadual do Maranhão como requisito parcial à obtenção do título de Bacharelado em Engenharia de Computação.

Aprovado em 30 de Janeiro de 2023.

BANCA EXAMINADORA

Prof. Me. Raimundo de Carvalho Silva Neto
Orientador

Documento assinado digitalmente
gov.br MARCOS JOSE DOS PASSOS SA
Data: 02/03/2023 10:16:36 -0300
Verifique em https://verificador.iti.br

Prof. Me. Marcos José dos Passos Sá

Prof. Me. Pedro Brandão Neto

AGRADECIMENTOS

Agradeço primeiramente à Deus que me deu sabedoria e força para chegar até aqui, também agradeço aos meus pais pelo apoio que sempre me deram durante a minha carreira como estudante.

Agradecimento especial a minha namorada, Thaynara, pelo amor e apoio incondicional durante todo o meu curso. Obrigado por acreditar em mim e me incentivar a continuar. Sem você, esse trabalho não teria sido possível.

Aos meus amigos Daniel, Ítalo, João Pedro, João Vitor, Leonardo, Luis Felipe, Luiz Gabriel, Marco Antonio, Matheus, Thalyson e Welington pelos bons momentos e companheirismo durante todo o curso.

Agradeço também aos meus amigos Adler, Átila, José Pedro, Marcos Guilhon e Thales que me ajudaram bastante no decorrer do curso. Em especial gostaria de agradecer à Nizar e Lucas que além da amizade, trilharam praticamente todo o curso junto comigo, sempre mantendo o bom humor mesmo com todas as dificuldades que tivemos, guardarei esses momentos com carinho.

Agradeço também a todos os professores que tive ao longo da graduação, aos que não lecionam mais na UEMA e também aos que já se foram. Todos tiveram um papel importantíssimo para a minha formação.

Por fim agradeço ao Prof. Me. Raimundo de Carvalho Silva Neto, um excelente educador, pela orientação para a conclusão deste trabalho.

LISTA ABREVIATURAS E SIGLAS

ACM - Associação de Máquinas de Computação

CODASYL - Conferência sobre as Linguagens de Sistemas de Dados

IBM - Máquinas de Negócios Internacionais

AMD - Análise Multicritério de Tomada de Decisão

AHP (*Analytical Hierarchy Process*) - Processo de hierarquia analítica

MCDM - Métodos de decisão multicritério

PCP - Planejamento e Controle da Produção

LISTA DE ILUSTRAÇÕES

LISTA DE FLUXOGRAMA

LISTA DE TABELAS

RESUMO

Nas últimas décadas, ferramentas de software estão em desenvolvimento a fim de facilitar o complexo processo de análise de dados, propondo ambientes integrados ou pacotes especializados para atender necessidades particulares. Como boas decisões dependem de informações relevantes e precisas, a seleção inadequada de um pacote de software pode resultar em decisões estratégicas erradas com subsequente perda econômica para o negócio. As mudanças e inovações encontra-se a Segurança da Informação, que passou a se preocupar ainda mais com a proteção dos dados e não só com o armazenamento, entende-se por segurança da informação todo o conteúdo ou dado valioso para um indivíduo/organização, que consiste na capacidade de armazenamento ou transferência de dados, de um determinado propósito. A metodologia utilizada para o desenvolvimento deste estudo foi uma pesquisa qualitativa, através de uma revisão bibliográfica sistemática da literatura. O objetivo geral do presente trabalho é analisar a importância da implantação do *Big Data* para melhorar a aplicação de métodos multicritérios em sistemas para melhoria da segurança da informação na prevenção de ameaças no ambiente organizacional. Conclui-se que o objetivo geral da pesquisa que é propor um modelo de análise multicritério como alternativa para suporte na tomada de decisões gerenciais no contexto do *Big Data* foi alcançado, esclarecendo até que ponto a aplicação de análise multicritério contribui de forma efetiva para o suporte na tomada de decisões gerenciais em uma organização.

Palavras-chave: Análises Multicritério, Big Data, Segurança da Informação, Tomada de decisão.

ABSTRACT

In the last decades, software tools have been developed to facilitate the complex process of data analysis, proposing integrated environments or specialized packages to meet particular needs. As good decisions depend on relevant and accurate information, improper selection of a software package can result in wrong strategic decisions with subsequent economic loss to the business. The changes and innovations are Information Security, which has become even more concerned with data protection and not only with storage, information security means all content or valuable data for an individual/organization, which consists of the ability to store or transfer data, for a particular purpose. The methodology used for the development of this study was a qualitative research, through a systematic literature review. The general objective of the present work is to analyze the importance of implementing Big Data to improve the application of multicriteria methods in systems to improve information security in the prevention of threats in the organizational environment. It is concluded that the general objective of the research, which is to propose a model of multi-criteria analysis as an alternative to support managerial decision-making in the context of Big Data, was achieved, clarifying to what extent the application of multi-criteria analysis effectively contributes to the support in making managerial decisions in an organization.

Keywords: Multicriteria Analysis, Big Data, Information Security, Decision Making.

SUMÁRIO

1 INTRODUÇÃO

O crescimento da quantidade e diversidade de informações nos últimos anos levou a um conjunto de dados maior do que os gerenciáveis pelas ferramentas convencionais de gerenciamento manual. O alto volume de dados é consequência direta dos avanços da tecnologia e da informatização da vida moderna, e tem promovido interesse de muitas organizações quanto a extração de percepções concisas e proveitosas para objetivos específicos (AGGARWAL, 2015).

A tecnologia nunca esteve tão presente na sociedade como nos dias atuais. Não apenas pessoas são responsáveis por produzir informações, equipamentos eletrônicos também se tornaram grandes criadores de dados, como por exemplo, registros de logs de servidores, sensores que são instalados nos mais variados contextos, entre uma infinidade de aplicações. Mensurar o volume de todos estes registros eletrônicos não é uma tarefa fácil.

O rápido aumento na quantidade de softwares para mineração de dados disponíveis no mercado tornou a tecnologia mais acessível para usuários finais não especialistas, no entanto criou um ponto de decisão crítico no processo de tomada de decisão nos negócios. Ao considerar a aplicação da mineração de dados, os usuários empresariais enfrentam o desafio de selecionar uma ferramenta adequada às necessidades e expectativas organizacionais.

Nesse contexto de mudanças e inovações encontra-se a Segurança da Informação, que passou a se preocupar ainda mais com a proteção dos dados e não só com o armazenamento. Em contrapartida os cibercriminosos também mudaram seu modo de atuar, empregaram novas técnicas de invasão enquanto que as técnicas de proteção ainda são reativas, mas vem melhorando significativamente, pois a adoção de novas tecnologias como o *Big Data* impulsionaram o investimento no setor de segurança da informação em sistemas de controle de estoque, que passaram a utilizar novas maneiras de se proteger.

Entende-se por segurança da informação todo o conteúdo ou dado valioso para um indivíduo/organização, que consiste na capacidade de armazenamento ou transferência de dados, de um determinado propósito. Nos últimos anos as tecnologias de informação e comunicação têm evoluído de forma rápida, fazendo com que as organizações tenham maior eficiência nas tomadas de decisão. Devido

a este fato as chances de uma empresa não usar segurança de informação tornou-se praticamente nula (CAMILO, 2018).

Em 2013 foi o inicial da comercialização de tecnologias de *Big Data* em segurança, uma tendência que vai remodelar as abordagens de segurança, as soluções e os gastos em relação aos próximos anos, segundo a RSA, subsidiária da EMC, a *big data* será um condutor para uma grande mudança em toda a indústria da segurança da informação e exigirá o uso de soluções orientadas à inteligência (BARTON, 2015).

De acordo com Zikopoulos e Eaton (2017), o termo *Big Data* se aplica a todo este potencial de dados que não são passíveis de análise ou processamento através dos métodos e ferramentas tradicionais. Por muito tempo várias empresas tinham a liberdade de ignorar o uso de grande parte deste volume de informações, pois não havia como armazenar estes dados a um custo benefício aceitável.

Big Data é um conjunto de dados extremamente amplos que necessitam de ferramentas especialmente preparadas para lidar com grandes volumes, informações nos meios que possa ser encontrada, analisada e aproveitada em tempo ágil. Projeta-se que, dentro de três a cinco anos, as ferramentas de análise de dados irão evoluir a ponto de permitir uma variedade de capacidades de previsão avançadas e controles em tempo real automatizado (ELMASRI, 2015).

Do lado da Segurança da Informação, a *big data* pode transformar os objetivos de proteção empresarial, ajudando a suportar as demandas de segurança e elevando controle e indicadores para níveis superiores, visto que as organizações experimentam nos dias atuais (ELMASRI, 2018).

O *Big data* no mercado pode oferecer milhares de oportunidades no cruzamento de dados, por gerar informações relevantes para diferentes públicos, porém, exige um sistema personalizado e com alta capacidade de segurança para que os dados sejam capturados e analisados sem riscos para as organizações. Além de tecnologias focadas, o sucesso da análise das informações pela *Big Data* ocorre também pela contratação de pessoas que saibam manipular tanto as máquinas quanto transformar os dados em informações de valor (GEREMIA, 2018).

A tomada de decisão está vinculada às rotinas dos seres humanos de escolher dentre as alternativas disponíveis, a que mais se adere aos requisitos definidos pelo contexto da decisão, a complexidade de uma decisão está diretamente relacionada à quantidade de critérios a serem ponderados. Para trazer

mais clareza, transparência, e apoio ao processo decisório, métodos matemáticos podem ser aplicados em diversas áreas do conhecimento.

O processo decisório não estruturado configura-se como uma situação que requer análise sucessiva e tem sido conceituado como um processo que é revestido de ambiguidades, estando direcionado à resolução de problemas que, em essência, são envoltos em incertezas e são corriqueiros nos níveis tático e estratégico.

A visão de empresas de diferentes portes e diferentes propostas tecnológicas sobre a aplicabilidade da tecnologia *big data* nas tarefas de tomada de decisão estruturada e não estruturada, tentando compilar os procedimentos balizadores desses tipos de processo decisório, que estivessem sendo seguidos nessas organizações, no intuito de observar mudanças proporcionadas pelo uso desse monumental acervo de dados disponibilizado

A necessidade de melhoria das técnicas de gestão tecnológica dentro das empresas, evidenciado por *Big Data* são indicadores de mudanças na estratégia organizacional das empresas, visando a adequação dos métodos e processos, o tema contribuirá para fornecer possibilidades de melhoria na administração das tecnologias e inovações voltadas para a gestão estratégica e tecnológica das organizações.

O presente estudo tem como objetivo geral analisar a importância da implantação do *Big Data* para melhorar a aplicação de métodos multicritérios em sistemas para melhoria da segurança da informação na prevenção de ameaças no ambiente organizacional e como objetivos específicos: Demonstrar os três conceitos do *Big data* no gerenciamento segurança; abranger as seis orientações do *Big Data* na solução adequada para os problemas de segurança da informação no contexto da tomada de decisão; descrever a visão geral dos conceitos referências aos dados e ao método AHP; verificar aplicação da análise multicritério de tomada de decisão no problema da segurança da informação em sistemas.

A motivação para realização deste trabalho parte das organizações que têm buscado respostas para fazer frente às necessidades decorrentes dos contínuos usos de novas fontes de dados digitais e apontam para uma evolução denominada tecnologia *big data* que quando associada às tecnologias de análise multicritério de tomada de decisão na segurança da informação, pode ser uma grande aliada para as organizações ampliarem sua posição no mercado em que atuam.

A fim de contribuir para o aumento da confiabilidade dos dados e informações geradas nos ambientes *Big Data*, este estudo se propõe a realizar uma análise com foco na característica veracidade em *Big Data***,** apresentando as relações deste aspecto com os mecanismos tecnológicos de segurança da informação.

Este trabalho é relevante para uma organização no ponto de visto prático no uso de métodos multicritérios nas decisões gerenciais diante do *big data* permite melhor adaptação às mudanças, evitando as falhas de processo que culminam em paradas de linhas e podem significar, em atrasos de produção, retrabalho, ineficiência, desperdícios de insumos, indisponibilidade de equipamentos, horas extras e estoques altos, incorrendo em prejuízos que atingem a empresa financeiramente e podem até determinar sua falência.

Frente aos argumentos exposto, entende-se que a presente pesquisa propõe desenvolver um modelo para aplicação da análise multicritério, oferecendo suporte nas futuras tomadas de decisões juntamente com o *big data*, para garantir um controle mais preciso e rápido dos defeitos e minimizar os trabalhos operacionais, bem como as margens para erros de tratamento de dados, há a necessidade de implementar uma ferramenta de *Big data* que seja capaz de cumprir com as demandas acima mencionadas, garantindo às partes interessadas, atualização automática da análise de multicritério em tomada de decisão.

No ponto de vista acadêmico, o trabalho envolve importantes fundamentos apresentados ao autor durante o curso que reforçam o envolvimento entre a teoria e a prática no mercado de trabalho. Dessa forma demonstrando a importância da aplicação de métodos multicritérios em sistemas para melhoria da segurança da informação e da tomada de decisão no ambiente organizacional do *big data*, não foram encontrados durante a pesquisa bibliográfica.

Com isso, pretende-se, com essa pesquisa, fazer uma análise para compreender as possibilidades de se aplicar uma arquitetura de Auxílio de Multicritérios no contexto *big data* em empresas. Esta arquitetura deve ter ênfase em uma estrutura projetada para atender de forma correta o processo de extração, tratamento, e análise de dados, para, por fim, auxiliar de forma efetiva o processo de tomada de decisão.

Este documento consiste em cinco capítulos. O primeiro contempla a introdução do trabalho, onde se disserta tanto sobre a contextualização quanto sobre o problema de pesquisa e os objetivos almejados (geral e específicos). Logo

em seguida, são apresentadas as justificativas para a realização deste trabalho e a estrutura da dissertação.

O segundo capítulo aborda a seleção e justificativa da Metodologia de Pesquisa utilizada para construir o processo de investigação deste trabalho. Também foram considerados neste capítulo as classificações da pesquisa no que diz respeito a natureza da pesquisa, a abordagem adotada, o tipo de objetivo almejado e o método científico utilizado.

No terceiro capítulo é também realizada a fundamentação teórica dos assuntos apontados, a fim de buscar conhecimento na literatura desenvolvida até o momento sobre o processo de abordagem e mineração de dados no contexto da segurança da informação.

No capítulo quatro é realizado a discussão sobre a aplicação da análise de multicritério de tomada de decisão no contexto *big data* relata sobre o *Big Data* em controles de dados em análise multicritério de tomada de decisão assim analisando a estrutura real e ágil na segurança da informação em gerenciamento de dados no *Big Data*.

Por último, no capítulo cinco, são explicitadas as considerações finais do trabalho, englobando: conclusões, limitações, principais dificuldades e proposições de estudos futuros.

2 REFERENCIAL TEÓRICO

2.1 A importância da aplicação de métodos multicritérios em sistemas para melhoria da segurança da informação

2.1.1 Processo de abordagem de mineração de dados

O termo “mineração de dados” é relativamente recente, sendo referido pela primeira vez na década de 1980. Como um campo de estudo, a mineração de dados surgiu há poucas décadas com um marco notório: a primeira Conferência ACM (*Association for Computing Machinery*) sobre Descoberta de Conhecimento em Bancos de Dados, realizada em 1995 no Canadá. Os avanços neste campo foram acompanhados pelo desenvolvimento de ferramentas de software relacionadas, fortemente baseadas em áreas associadas como estatística, inteligência artificial, aprendizado de máquina e pesquisa de banco de dados (CALDERS e CUSTERS, 2016; MIKUT e REISCHL, 2011).

A mineração de dados é uma etapa do processo de descoberta de conhecimento em bancos de dados que consiste na aplicação de algoritmos de descoberta e análise para produzir uma enumeração particular de padrões ou modelos nos dados (MIKUT e REISCHL, 2011). Aggarwal (2015, p. 1) a define como “o estudo de coleta, limpeza, processamento, análise e obtenção de percepções úteis de dados”.

De acordo com Mor *et al*., (2020), o processo de mineração de dados é dividido em duas etapas: preparação ou pré-processamento de dados e mineração de dados. No pré-processamento ocorre a limpeza, integração, seleção e transformação dos dados, enquanto na mineração ocorre a avaliação de padrões e a representação do conhecimento. De forma geral, após uma adequada preparação dos dados, diferentes modelos podem ser construídos de acordo com o objetivo da pesquisa e interpretados por meio de softwares específicos.

O banco de dados costuma ser definido como a extração automatizada de padrões que representam o conhecimento implicitamente armazenado em grandes bancos de dados como *data warehouses*, *web* e outros repositórios de informações massivas ou fluxos de dados. No entanto, vale ressaltar que, ao contrário da estatística, na qual as informações são coletadas especialmente a fim de testar uma

determinada hipótese ou estimar parâmetros de um modelo, a mineração de dados geralmente é originada a partir de dados que não foram coletados com o propósito de análise, mas como um subproduto de um sistema (CALDERS e CUSTERS, 2016).

A maioria das ferramentas modernas de mineração de dados são arquiteturas de fluxo de dados baseadas em software, algumas dessas ferramentas apresentam ambientes gráficos integrados que permitem a inserção, a conexão e a movimentação de componentes visuais, o modelo mais comum desses dispositivos inclui a extração, transformação e apresentação dos dados.

A implementação de procedimentos adicionais, como pacotes especializados e *add-ons*, normalmente tem pouca importância para os usuários particulares e iniciantes, entretanto, usuários empresariais e mais avançados geralmente dispõem do código do modelo e linguagens de programação para melhorá-lo (JOVIĆ *et al*., 2018).

Vários modelos de processos de mineração de dados foram desenvolvidos nas últimas décadas. Embora cada um opere com suas particularidades, seus princípios básicos e estrutura geral são essencialmente os mesmos. Além disso, os pacotes de software fornecem uma diversidade de recursos personalizáveis que podem ser ajustados para atender às necessidades específicas das organizações (JADHAV e SONAR, 2019).

Por essa razão, conforme o número de ferramentas disponíveis continua a crescer, a escolha da mais adequada torna-se cada vez mais difícil. Nesse sentido, métodos de apoio multicritério à decisão são utilizados para selecionar o software mais adequado aos objetivos de uma organização.

2.1.2 História do Banco de dados

Com a correria do dia a dia é difícil parar para pensar em como as coisas eram feitas em outrora e no porquê de terem se transformado tanto, principalmente quando trata-se de questões relacionadas ao armazenamento e o acesso de dados um dos quesitos mais importantes dentro da tecnologia.

O armazenamento e acesso a dados era bastante “grosseiro”, isso no final da década de 50 e no início da década de 60. Os bancos de dados que conhecemos hoje começaram a ser projetados pelo Departamento de Defesa dos Estados Unidos

da América. Em 1957, essa instituição inaugurou a *Conference on Data Systems Languages* (Conferência sobre as Linguagens de Sistemas de Dados ou CODASYL), com o objetivo de desenvolver linguagens de programação de computador. A CODASYL se destacou pela criação da linguagem de programação COBOL, embora poucos saibam que o CODASYL também foi responsável pela criação do primeiro banco de dados moderno (CAMILO, 2018).

No final da década de 60, as grandes empresas viram uma enorme evolução da informática, onde já não era mais possível controlar a quantidade de processos que vinham sendo utilizados pelos usuários. Foi então que surgiu o banco de dados, criado pela a empresa IBM (International Business Machines) que é uma empresa dos Estados Unidos voltada para a área de informática.

O objetivo de criar os bancos de dados seria fazer com que estes fossem responsáveis pelo controle, organização e armazenamento de dados, fazendo com que houvesse uma diminuição no grande número de pessoas dentro das empresas, pois já não haveria a necessidade de várias pessoas para armazenar e organizar os arquivos.

Em 1963, duas divisões do Departamento de Defesa dos Estados Unidos da América formaram uma conferência intitulada *"Development and Management of a Computer-Centered Data Base"* (Desenvolvimento e Gerenciamento de um Banco de Dados para Computadores), onde o termo *database* (banco de dados ou base de dados) foi concebido e definido como se segue: um conjunto de arquivos (tabelas), onde um arquivo é uma coleção ordenada de registros (linhas), e um registro consiste em uma ou mais chaves e dados (CAMILO, 2018).

Desde a década de 60 a criação de banco de dados vem sendo utilizada até hoje e durante essa grande jornada foram feitas várias pesquisas e muitas transformações aconteceram e outros tipos de banco de dados foram surgindo (Peter) Chen (CAMILO, 2018).

Segundo Elmasri e Navathe (2018) um Banco de Dados representa as características do mundo real, sendo chamado de mini-mundo, possui os dados armazenados de uma maneira lógica e coerente. Para Geremia (2018) um banco de dados é uma coleção de dados ou registros relacionados.

Um banco de dados (BD) pode ser definido como uma coleção de dados inter-relacionados, armazenados de forma centralizada ou distribuída, com redundância controlada, para servir a uma ou mais aplicações. Para Korth (2017),

que vai além na sua definição diz que um banco de dados é uma coleção de dados inter-relacionados, representando informações sobre um domínio específico, ou seja, sempre que for possível agrupar informações que se relacionam e tratam de um mesmo assunto, posso dizer que tenho um banco de dados.

Sousa e Moraes (2018) afirmam que no que se refere aos dados e informações:

> Os termos dados e informação estão intrinsecamente relacionados, sendo usados da mesma forma em muitas ocasiões, porém há uma diferença sutil que os faz diferentes: o dado é uma referência não elaborada, algo não interpretado, não classificado, não estruturado, não ajustado a um contexto.

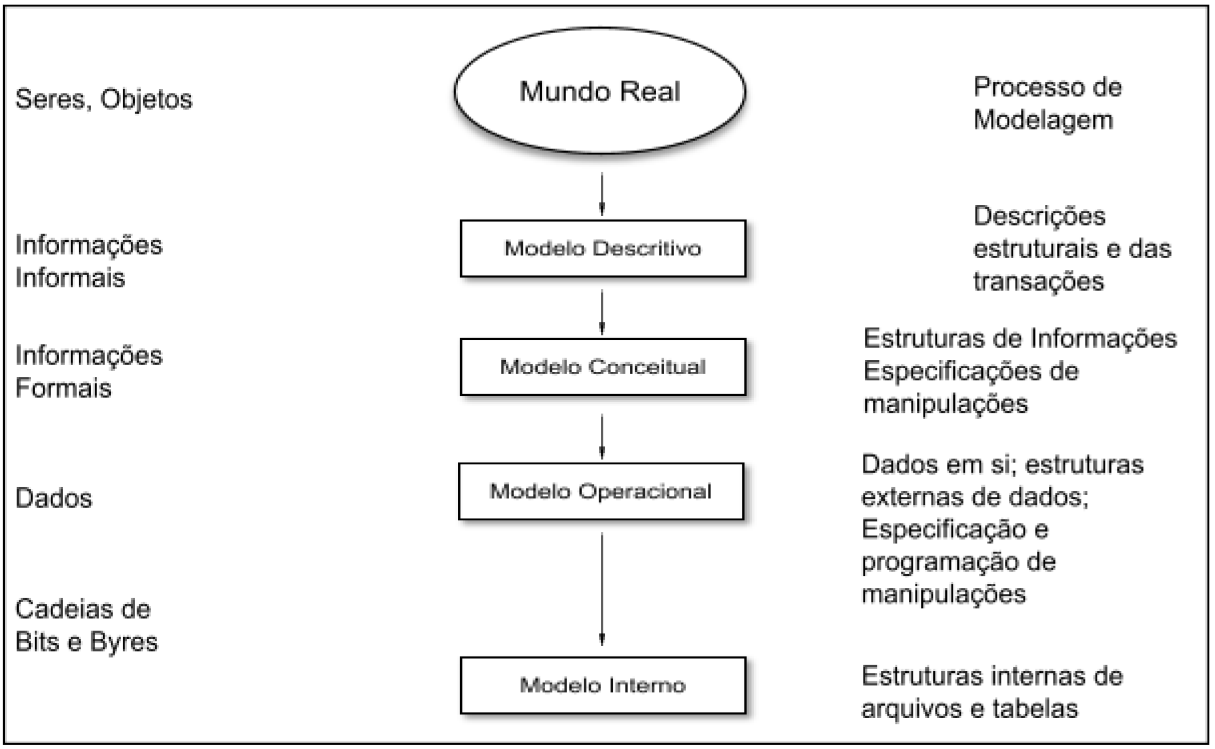
Os dados reunidos segundo uma determinada estrutura lógica nos trazem a informação, são tratados como sinônimos e alguns autores distinguem dados como os valores fisicamente armazenados em um banco de dados e informação como o significado que esses valores têm para o usuário. De uma maneira conceitual pode-se dizer que o banco de dados é o coração de muitos sistemas de informação, pois são formados por modelos de informações originadas do mundo real.

Os sistemas de gerenciamento de banco de dados surgiram para solucionar principalmente problemas de incoerência de informações, inconsistência de dados, compartilhamento de informações e dependências de dados de aplicações. Segundo Elmasri (2018), um sistema de gerenciamento de banco de dados (SGBD) é um sistema ou conjunto de sistemas que tem como principal objetivo o controle de acesso, armazenamento, restauração e organização de informações.

O armazenamento de dados é um fator estratégico para uma empresa moderna. Sua importância se revela quando se avalia o intenso uso de tecnologia para comunicação e realização de atividades corporativas que, em conjunto, levam a um crescimento do volume de dados a ser gerido. A quantidade de dados gerados vem tomando proporções gigantescas e este crescimento não vai diminuir, muito pelo contrário, a tendência é de aumentar gradativamente.

Pode-se observar na figura 1 os modelos que mostram a estrutura da informação desde o seu processo de modelagem, onde são captadas do mundo real o que se quer do banco de dados, até a descrição de sua estrutura física de armazenamento, onde e como estarão localizados no banco de dados.

Figura 1 - Modelos de Concepção de um Banco de Dados



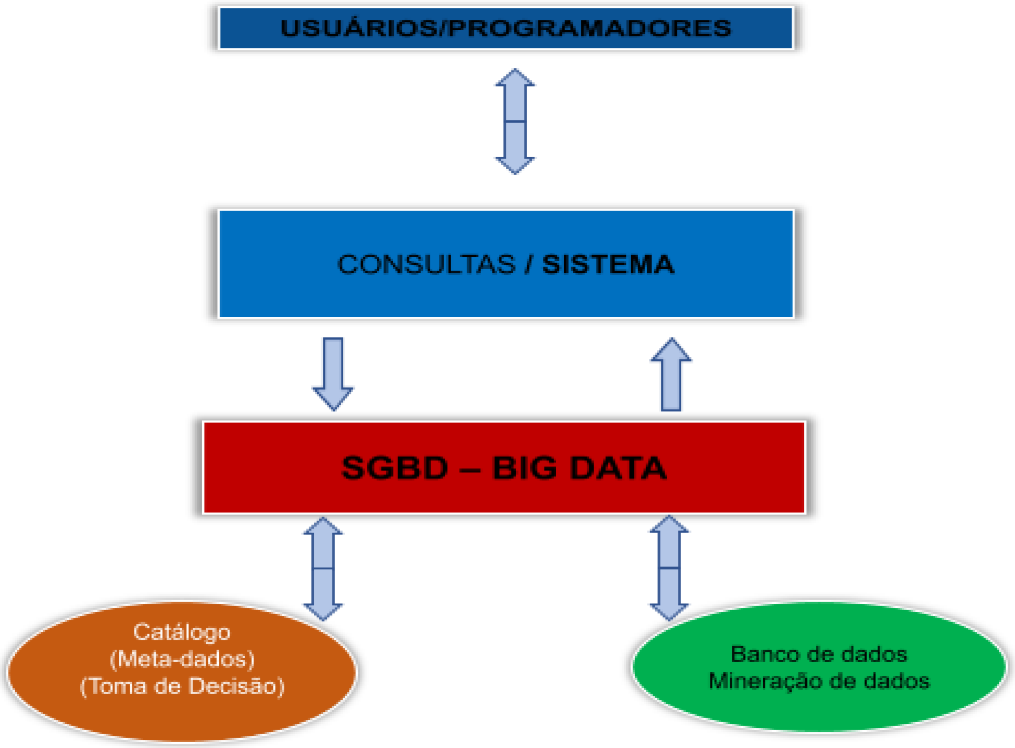
Fonte: Sousa (2018)

Os Banco de dados são coleções de dados interligados entre si e organizados para fornecer informações. Para entender o que é um banco de dados é muito importante saber a diferença entre dados e informações. Segundo Date (2019), dados são fatos brutos, em forma primária, e muitas das vezes podem não fazer sentido sozinhos. As informações consistem no agrupamento de dados de forma organizada para fazer sentido, gerar conhecimento.

Segundo Kenneth e Jane (2018), disponibilizar uma interface para programas e usuários externos acessem o banco de dados, controlar o acesso a informações, manter cópias dos dados para recuperação de uma possível falha, garantir transações no banco de dados, enfim, sem o SGBD o banco de dados não funciona

Resumidamente o SGBD é um conjunto de software para gerenciar (definir, criar, modificar, usar) um BD e garantir a integridade e segurança dos dados. O SGBD é a interface entre os programas de aplicação e o BD, isso pode ser observado na figura 2.

Figura 2 - Funcionamento do SGBD em Sistemas de tomada de decisão.



Fonte: Sousa (2018)

Como pôde ser observado na figura 2 tudo que se faz no banco de dados passa pelo SGBD, pois ele é o responsável por salvar e encriptar dados, monitorar o acesso à informação, reter cópias dos dados, assegurar transações no banco de dados, ou seja, o banco de dados necessita do SGBD para que possa funcionar.

Sousa (2018) destaca as principais vantagens para a utilização de um SGBD: controle de redundância de dados; controle de acesso; armazenamento de dados; existência de múltiplas interfaces para os usuários; representação de relacionamentos complexos entre dados; manutenção de restrições de integridade; e recuperação de falhas.

Apesar das vantagens no uso do SGBD, Elmasri e Navathe (2018) citam algumas situações em que esse sistema pode envolver custos altos e desnecessários, que normalmente não ocorreriam no processamento tradicional de arquivos. Investimentos iniciais altos em *hardware*, *software* e treinamento; generalidade que o SGBD fornece para a definição e processamento de dados;

custos elevados para oferecer segurança, controle de concorrência, recuperação e funções de integridade.

Elmasri e Navathe (2015) afirmam ainda que problemas adicionais podem surgir se os projetistas do banco de dados e o administrador do banco de dados não projetarem o banco de dados de maneira adequada ou se a aplicação não for implementada apropriadamente. Sendo assim, indicam o uso de arquivos convencionais nas seguintes situações: o banco de dados e suas aplicações são simples, bem definidas e sem previsão de mudanças; há requisitos de tempo real para alguns programas difíceis de serem atendidos por causa da sobrecarga do SGBD.

2.2 Definição de Segurança Da Informação

Segurança da informação é um tema atual em constante discussão nas mais diversas organizações, seja governo, educação, indústria, comércio ou serviços; visto que as organizações se utilizam da Tecnologia da Informação para apoiar e gerar negócios, aliados aos benefícios da internet. Desse modo, independentemente do segmento de mercado, *core business* ou porte, todas as organizações sempre usufruirão da informação, objetivando melhor produtividade, redução de custos, ganho na participação do mercado, aumento de agilidade, competitividade e apoio à tomada de decisão (SÊMOLA, 2016).

Segurança da informação, conforme Beal (2018), é o processo de proteção da informação das ameaças a sua integridade, disponibilidade e confidencialidade. Soares (2017) define segurança da informação como “uma área do conhecimento dedicada à proteção de ativos da informação contra acessos não autorizados, alterações indevidas ou sua indisponibilidade”

A ISO/IEC 17799:2005, em sua seção introdutória define segurança da informação como a proteção da informação de vários tipos de ameaças para garantir a continuidade do negócio, minimizar o risco ao negócio, maximizar o retorno sobre os investimentos e a oportunidade de negócio (CAMILO, 2018). Assim, podemos definir segurança da informação como a área do conhecimento que visa à proteção da informação das ameaças a sua integridade, disponibilidade e confidencialidade a fim de garantir a continuidade do negócio e minimizar os riscos, maximizar o retorno sobre os investimentos e as oportunidades de negócio.

Beal (2018) e Sêmola (2016) asseguraram que o objetivo da informação é preservar os ativos de informação quanto a sua confidencialidade, integridade e disponibilidade. A Confidencialidade da informação é a garantia de que somente pessoas autorizadas terão acesso a ela, protegendo-a de acordo com o grau de sigilo do seu conteúdo; Integridade da informação que tem como objetivo garantir a exatidão da informação, assegurando que pessoas não autorizadas possam modificá-la, adicioná-la ou removê-la, seja de forma intencional ou acidental.

Disponibilidade significa garantir que a informação possa ser obtida sempre que for necessário, isto é, que esteja sempre disponível para quem precisar dela no exercício de suas funções; a legalidade garante que a informação foi produzida em conformidade com a lei; a autenticidade garante que num processo de comunicação que os remetentes sejam o que dizem ser e que a mensagem ou informação não foi alterada após seu envio ou validação (CARUSO, 2019).

Ainda dentro dos objetivos Sêmola (2016) acrescenta outros dois a legalidade e a autenticidade. O primeiro garante a informação produzida em conformidade com a lei e o segundo garante que nenhum processo de comunicação do remetente seja exatamente o que é colocado na mensagem ou informação sendo alterada após o seu envio ou avaliação.

Para Beal (2018), o objetivo da legalidade é melhor classificado como organizacional, assim como o objetivo de uso legítimo da informação citados por outros autores, pois deles derivam os requisitos de segurança da informação. Quanto ao objetivo: *autenticidade,* a autora entende como necessário somente quando é usado num processo de transmissão de informações, e estabelece alguns objetivos adicionais relativos à segurança da comunicação:

- Integridade do conteúdo;
- Irretratabilidade da comunicação:
- Autenticidade do emissor e do receptor;
- Confidencialidade do conteúdo;
- Capacidade de recuperação do conteúdo pelo receptor.

Sêmola (2016) adverte sobre a expressão “Segurança da informação”, dizendo que por si só, é um termo ambíguo, podendo assumir dupla interpretação:

segurança como uma prática adotada, objetivo a ser alcançado. Sêmola (2016, p. 85) cita exemplos:

> Segurança como “meio” a segurança da informação visa garantir a confidencialidade, integridade e disponibilidade da informação, a impossibilidade de agentes participantes em transações ou na comunicação repudiem a autoria de suas mensagens, a conformidade com a legislação vigente e a continuidade dos negócios;
> Segurança como “fim” - a segurança da informação é alcançada por meio de práticas e políticas voltadas a uma adequada padronização operacional e gerencial dos ativos, e processos que manipulem e executem a informação.

Deve-se tomar cuidado com a definição de segurança pela confusão corrente do termo com risco, privacidade e confiança, no caso da confiança explicam: “confiança engloba e significa muito mais do que segurança, confiança é o pilar de sustentação de qualquer negócio ou empreendimento, tradicional ou eletrônico, dentro ou fora da internet, sendo a segurança um dos seus principais construtos (ELMASRI, 2018).

Existem alguns termos relacionados à gestão da segurança da informação que merecem atenção, de acordo com Caruso (2019, p. 45), são eles:

> Ativo: tudo aquilo que tem valor para a organização; ameaça: expectativa de acontecimentos acidental ou proposital, causado por um agente, que pode afetar um ambiente, sistema ou ativo de informação. É algo normalmente externo ao ativo que se quer proteger (falha de energia, fogo, vírus); vulnerabilidade: fragilidade que poderia ser explorada por uma ameaça para concretizar um ataque. Está associada ao próprio ativo, podendo ser decorrente de uma série de fatores, como falta de treinamento, falta de manutenção, falha nos controles de acesso etc.; impacto: efeito ou consequência de um ataque ou incidente para a organização; ataque: evento decorrente a exploração de uma vulnerabilidade por ameaça. Exemplos de ataque: digitação incorreta de dados pelo usuário vazamento de informações, inclusão indevida no sistema de pagamento de compra fictícia; incidente: fato (ou evento) decorrente da ação de uma ameaça, que explora uma ou mais vulnerabilidades, levando à perda de princípios da segurança da informação: confidencialidade, integridade e disponibilidade.

O ataque deriva de uma ameaça inteligente e é uma ação deliberada contra as políticas de segurança do sistema, aproveitando-se de uma vulnerabilidade. Os ataques contra a integridade e disponibilidade irão contra a confiabilidade da informação, e contra o suporte ao serviço ou a destruição da informação. Seja como for, com certeza as maiores ameaças estão dentro da própria organização, de forma que um *firewall* pode ser insuficiente (GEREMIA, 2018).

O preparo da defesa contra ataques ao controle de acesso pode ser mais simples, porém em outros casos, o caminho está primeiramente na detecção do ataque e até mesmo no alerta sobre a situação. A defesa por níveis de acesso pode incluir a necessidade de mais uma autorização ou tipo de autorização. Em alguns casos a defesa contra os ataques pode significar perda da certa privacidade do usuário, uma vez que seu perfil e ações podem estar sendo armazenadas pelo controle (KORTH, 2017).

Ataques que se aproveitam de vulnerabilidades técnicas de sistemas de aplicações normalmente precisam de algum tipo de ajuda de dentro, esta ajuda pode não ser intencional por parte do usuário interno, seja por conta de comportamento impróprio, seja por conta de curiosidade quanto aos limites de seus direitos no sistema. Quando é possível para o usuário baixar programas da internet, ele deve estar consciente de que pode estar sendo um portal para ataques. Sites na Internet que não são confiáveis devem fazer parte da lista de bloqueio (GEREMIA, 2018).

2.2.1 O preço das informações

Em uma empresa o seu bem mais valioso são as informações relacionadas com os bens de consumo ou serviços prestados por ela (CARUSO e STEFFEN, 2019), e pela alta capacidade que dados, informação e conhecimento têm de adicionar valor a processos, produtos e serviços, esses constituem recursos cada vez mais críticos para o alcance da missão e dos objetivos organizacionais. Consequentemente, as informações críticas para o negócio precisam ser protegidas contra as ameaças que podem levar à sua destruição, indisponibilidade temporária, adulteração ou divulgação não autorizada (BEAL, 2018).

A informação é um recurso que tem valor para a organização e deve ser bem gerenciada e utilizada, é necessário garantir que ela esteja sendo disponibilizada apenas para pessoas que precisam dela para o desempenho de suas atividades profissionais. Com base nas informações o autor alerta que o maior risco para as informações é a organização não se preocupar com a segurança e como os possíveis ataques a ela, achando que isso acontece apenas com os outros (CARUSO E STEFFEN, 2019).

O ser humano sempre buscou o controle sobre as informações que lhe eram importantes, até mesmo na mais remota antiguidade, o que mudou foram as formas de registro e armazenamento dessas informações, que nos dois últimos séculos passaram a ter importância crucial para as organizações humanas. Devido ao modo arcaico de registro de informações na antiguidade, era natural que o controle de disseminação das informações tornassem o acesso às mesmas restritas e a uma minoria sempre ligada ao grupo que dominava o poder econômico e político da sociedade (SOUSA, 2018).

Os primeiros suportes para registro de informações foram as paredes das habitações humanas; por si só implica um conjunto de consequências, restrições de acesso físico, de transferências para terceiros ou para outro local e de pessoal capacitado. Além da “imobilidade” das informações que ninguém na época detinha o conhecimento necessário para reconhecê-las. Em meados do século XX a alfabetização se universalizou, o que possibilitou a várias pessoas o acesso à informação (CARUSO e STEFFEN, 2019).

Para os autores “não há organização humana que não seja dependente da tecnologia de informações”, fato que se potencializou em função de evolução da informática com o acúmulo de grande quantidade de informações em pequenos espaços, essa característica acarreta consequências graves para essas mesmas organizações, por facilitar os ataques de pessoas não autorizadas (TABUENA, 2017).

Independente do setor da economia em que a empresa atue, as informações estão relacionadas com seus processos de produção e de negócios, políticas estratégicas, de marketing, cadastros de clientes, dados contábeis e financeiros etc. E não importa o meio de armazenamento, elas são de valor inestimável não só para a empresa que as gerou como também para seus concorrentes (VACCA, 2017).

2.3 O *Big Data* em controles de dados na Análise Multicritério de Tomada de Decisão

Nas últimas décadas, ferramentas de software estão em desenvolvimento a fim de facilitar o complexo processo de análise de dados, propondo ambientes integrados ou pacotes especializados para atender necessidades particulares. Como boas decisões dependem de informações relevantes e precisas, a seleção

inadequada de um pacote de software pode resultar em decisões estratégicas erradas com subsequente perda econômica para o negócio (JADHAV e SONAR, 2019).

Devido aos diversos fatores de impacto na seleção de inúmeros dados por segundo, o processo segurança é considerado um problema complexo de tomada de decisão multicritério, neste contexto, os critérios conflitantes precisam de uma avaliação cuidadosa para garantir uma decisão correta, confiável e segura de acordo com os objetivos da organização para a segurança dos seus dados.

Uma organização ao saber como utilizar seus dados poderá melhorar seu produto, criar estratégia de marketing mais eficiente, cortar gastos, produzir mais, evitar desperdício de recursos, superar um concorrente, disponibilizar um serviço a um cliente de maneira satisfatória.

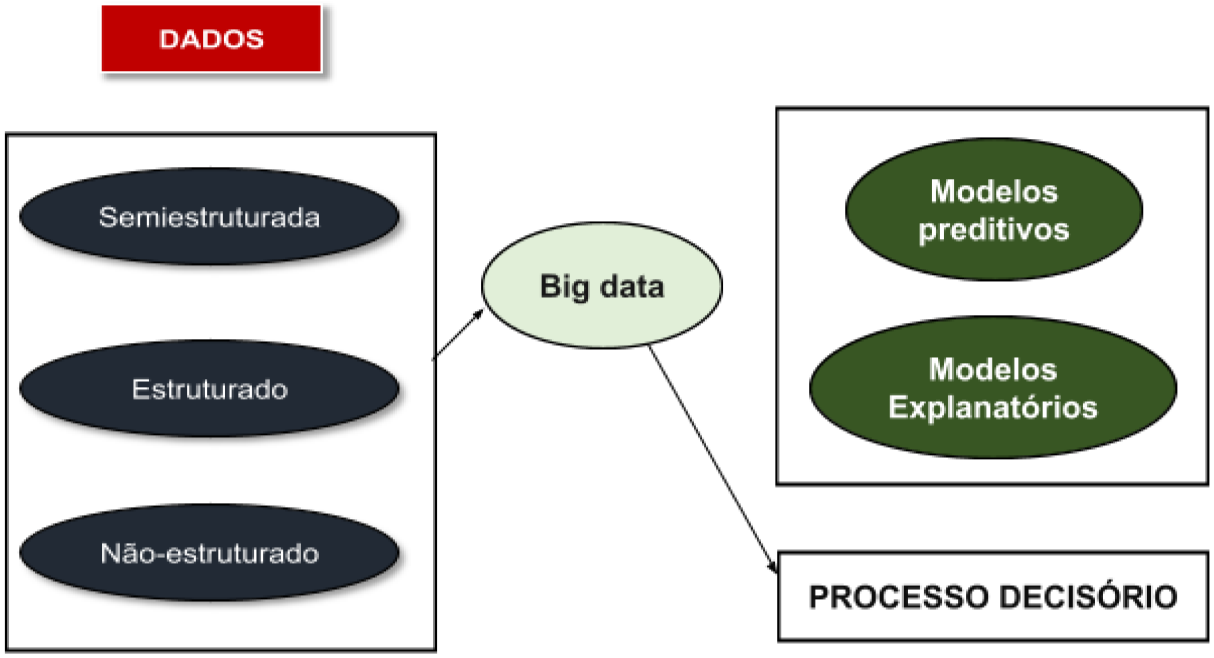
Atualmente os avanços computacionais permitem guardar, organizar e analisar dados mais facilmente e com maior frequência, a cada dia que se passa, a sociedade progride tecnologicamente, beneficiando logicamente, várias áreas da nossa sociedade. A quantidade de dados gerados diariamente nas redes sociais e com um enorme número de sites na Web voltado a oferecer compras on-line, pois quanto maior for o número de informatização pelas lojas, será necessário de gerenciar seus estabelecimentos físicos, e neste sentido uma segurança cada vez mais eficiente.

A longo prazo, a *big data*, aponta ainda as previsões, mudará a natureza dos controles convencionais de segurança, como anti-malware, prevenção de perda de dados e *firewalls*. O *Big Data* tem como foco a segurança e a privacidade dos dados e informações disponíveis, pois ao mesmo tempo em que aumentamos o acesso a novos dados, acabamos induzindo a um maior controle da vida dos usuários. Assim, ao passo que o Big Data disponibiliza dados sobre localizações, preferências e opiniões dos indivíduos para permitir novos conhecimentos e possibilidades, há efeitos sobre a liberdade, segurança ou privacidade.

Além disso, os softwares modernos de mineração de dados precisam oferecer atributos que os tornem capazes de lidar com dados heterogêneos, adaptáveis às diferentes demandas e acessíveis no mercado. Para tanto, neste trabalho, o método AMD (Análises Multicritério de tomada de decisão) ajudará a determinar o julgamento correto na seleção do software na segurança mais adequada com base em requisitos específicos selecionados pelos tomadores de decisão.

A tecnologia *big data* vem se estabelecendo como um assunto de grande interesse devido ao seu potencial analítico, mas, em contraponto, impõe novos desafios às organizações, em especial no que tange ao volume de dados gerados por meio da rede e à utilização desses dados em benefício próprio, dentre esses desafios, a tecnologia *big data* requer das organizações a criação de estratégias mais consistentes para traçar o efetivo uso dos dados.

Figura 3 - Representatividade de *big data* para o processo decisório



Fonte: Baseado em Davenport (2018).

Dentre esses desafios, a tecnologia *big data* requer das organizações a criação de estratégias mais consistentes para traçar o efetivo uso dos dados. Ora, supõe-se que no processo decisório, *big data* tem grande representatividade, vez que adiciona novas fontes de dados a modelos preditivos e explanatórios necessários ao raciocínio dos decisores, considerando suas necessidades, de acordo indica-se na figura 3 acima. Dessa forma, a fim de facilitar os processos de segurança da informação e decisórios no setor organizacional, o presente estudo visa contribuir com uma técnica estruturada de solução de problemas de segurança da informação.

Não há consenso do conceito de *Big Data* utilizado no mercado, pois cada organização considera esse fenômeno por alguma perspectiva, no entanto, a essência do termo é contemplada por muitas organizações referências (KORTH,

2017). Para o desenvolvimento deste trabalho, é importante que este conceito esteja claro, portanto, é necessária a formulação de um conceito único. As organizações *Decision Support Systems e McKinsey Global Institute* utilizam a dimensão de volume de dados em seus conceitos, ressaltando que há um desafio no que se refere à gestão e análise desses dados.

No entanto, como considera o *Gartner Group* e consta no estudo elaborado pelo IDC, há outras duas dimensões muito importantes ao considerar *Big Data*: velocidade e variedade, o fenômeno *Big Data*, com o conceito mais amplo, o qual contempla as três dimensões, é fundamental para incrementar os resultados das organizações, principalmente, os resultados de vendas e segurança da informação.

As novas tecnologias aplicadas da IBM na tomada de decisão, em opinião da IBM (2014), agregaram outras duas dimensões que complementam a relevância do fenômeno e no seu efeito no processo de tomada de decisão, são elas: valor e veracidade. Segundo ele, veracidade é importante “porque precisamos ter certeza que os dados fazem sentido e são autênticos”, uma segurança da informação não pode ser embasada em dados duvidosos

O volume é a dimensão mais comum nos conceitos de *Big Data*, visto que o fenômeno vem chamando atenção pela acelerada escala em que os volumes cada vez maiores de dados são criados pela sociedade” (TANKARD, 2017). Preimesberger (2021) simplifica esta dimensão contabilizando-a: “terabytes, petabytes e eventualmente exabytes” de dados criados por seres humanos e por máquinas. Além disso, o mesmo autor ressalta a dificuldade de “armazenar, proteger e tornar acessível”.

Segundo Tankard (2017), é um desafio gerar e armazenar esse grande volume de dados com as ferramentas tradicionais. Os desafios técnicos também são levantados por Geremia (2018), assim como os sociais, visto que deve haver uma mudança para um mundo em que grandes volumes de conjuntos de dados são rotineiramente publicados, variáveis envolvidas: volume de informações; acessibilidade de informações; mudança do comportamento em função do volume

A variedade é a existência de dados estruturados e não estruturados é abordada por Korth (2017) para exemplificar diferentes tipos de dados, os dados estruturados são os utilizados frequentemente nos sistemas das organizações, como bancos de dados, arquivos, arquivos sequenciais e com relação de importância; os dados semi estruturados são heterogêneos, ou seja, ora estão dispostos em padrão,

ora estão em outro, dificultando a sua manipulação; há também os dados não estruturados que são os conteúdos digitais de diversas mídias como vídeos, imagens, áudios, conteúdo de e-mails, entre outros (CIO, 2012).

Uma pesquisa realizada pelo IDC afirma que quase 90% dos dados do universo digital são considerados dados não estruturados. Dessa forma, é extremamente relevante que as organizações atentem para esse tipo de dado, visto que tem uma representatividade considerável. A mesma pesquisa projeta um crescimento de 48% da quantidade de informações no ano de 2012, com relação ao ano anterior Segundo Taurion (2012) os dados estruturados são a imensa maioria.

McAfee e Brynjolfsson (2019) e Preimesberger (2021) apresentam uma variedade de dados através de e-mails, registro de cartões de créditos e de sensores, informações de redes sociais, bem como fotos, arquivos audiovisuais e gráficos. Variáveis envolvidas: variedade de fontes de informação; variedade de tipos de dados.

A velocidade é a exigência de que a tomada de decisão seja próxima do mercado, que é mais ágil e dinâmico, são fatores que ressaltam a importância da dimensão velocidade. Mídias sociais e microblogs auxiliam na propagação mais rápida de informações, sejam elas organização ou dos próprios consumidores, positivas ou negativas. McAfee e Brynjolfsson (2019) e Preimesberger (2021) também trabalham como o argumento de que é importante saber trabalhar com a velocidade, pois pode ser um limitador da análise, podendo inclusive inviabilizar uma operação, caso um software não receba cargas em tempo real, quando a velocidade dos dados é esta. Segundo Taurion (2012) muitas vezes precisamos agir em tempo real exigindo um processamento que acompanhe esta velocidade.

O valor, segundo Webber (2019), é a qualidade de dados que exige a exatidão, integridade, consistência e relevância", e ainda ressalta a dependência da qualidade de informações percebida pelo usuário nas suas necessidades. A confiança e a experiência para extração de valor na análise rigorosa de dados do *Big Data*. Tankard (2017) destaca a clareza como uma característica importante para atribuir valor às informações.

Variáveis envolvidas: resultado agregado a partir das análises das informações; qualidade das informações; valor financeiro para ter dados de qualidade. É importante ressaltar a importância da qualidade dos dados para análise

adequada do *Big Data*, visto que dados de alta qualidade são pré-requisito para auxiliar as organizações a adequarem-se às mudanças necessárias.

É necessário ter certeza que os dados fazem sentido e são autênticos. Veracidade conforme o dicionário Priberam, é a qualidade daquilo que é verdadeiro e exato. Seguindo o mesmo significado considera “credibilidade” como outra dimensão da qualidade da informação. Informações verdadeiras têm relação com informações exatas, íntegras, consistentes e relevantes – características da qualidade da informação – podendo então ser utilizadas pelos gestores ou responsáveis para responder aos desafios estratégicos e operacionais (WEBER *et. al.,* 2019).

2.4 *Big Data* e a Segurança da Informação

O *Big Data* é um ponto positivo e atrativo para todos os setores econômicos e sociais. Entretanto, deve-se explorar também os pontos negativos da utilização e implementação do *Big Data* em organizações, tendo como foco a segurança e a privacidade dos dados e informações disponíveis, pois ao mesmo tempo em que aumentamos o acesso a novos dados, acabamos induzindo a um maior controle da vida dos usuários (TABUENA, 2017).

Assim, ao passo que o *Big Data* disponibiliza dados sobre localizações, preferências e opiniões dos indivíduos para permitir novos conhecimentos e possibilidades, há efeitos sobre a liberdade, segurança ou privacidade. Quando pensamos de maneira macro, percebemos que informações sobre nossos hábitos estão disponíveis de muitas formas, seja com eletrodomésticos que possuem medidores sobre o nosso consumo, revelando nossas práticas diárias, seja pelas declarações nas redes sociais.

De fato, com o avanço do *Big Data* fica mais suscetíveis, e mesmo sem perceber a restrição da nossa liberdade e ao aumento dos riscos, o mesmo trabalha para disponibilizar às empresas o maior nível de segurança e privacidade dos seus dados, e muitas vezes para que isso ocorra outros níveis da cadeia são expostos, sejam clientes ou profissionais.

O *Big Data* ainda possui grandes barreiras a ultrapassar na questão da privacidade dos dados, principalmente quando se tratam de dados pessoais ou de empresas que necessitam de sigilo. Por diversas vezes, organizações se utilizam

de dados pessoais como e-mails, fotos publicadas e pesquisas online para promover novos produtos e serviços que possam atingir o interesse do usuário. Sabe-se que as informações disponíveis na internet nada mais são que os próprios dados disponibilizados pelos usuários, e são por meio desses dados que as empresas utilizam novas fontes de análise e promoção dos seus produtos e serviços (PREIMESBERGER, 2021).

Apesar de ser por meio desses dados disponíveis que muitas vezes conquistamos aquilo que desejamos, sejam produtos ou serviços, nossa segurança fica vulnerável às práticas que utilizam o *Big Data* como estratégia para alavancar o negócio, e muitas vezes dados pessoais podem ficar expostos para qualquer usuário, colocando em risco informações importantes, dados bancários, dos membros familiares, localização e entre outros.

Quanto mais as análises do *Big Data* nas tomada de decisão perante a segurança da informação se mostraram promissoras as organizações e ao público, mais confiança os usuários terão em disponibilizar informações e mais visíveis estarão perante aos dados analisados, ou seja, por ser um fenômeno conhecido e confiável por muitos, os usuários tendem a achar que estão controlando os dados, quando na verdade por muitas vezes o *Big Data* programa os dados que mais temos acesso, diminuindo a privacidade e segurança do usuário.

A segurança de informações e de dados é outro fator de extrema importância para o funcionamento do *Big Data* em empresas, já que sem esses requisitos as organizações não conseguem ter a confirmação que o trabalho que será realizado e as ideias que irão surgir ficarão armazenadas sem o risco de empresas concorrentes terem acesso (SCHMARZO, 2016).

A segurança dos dados pode ser analisada de duas maneiras, se os dados estão seguros quando são armazenados em certo local, sendo o grande receio o caso de alguém fora da organização consiga acesso sem grandes dificuldades. Outro ponto é se os dados extraídos são confiáveis e seguros, muitos usuários e funcionários de organizações não acreditam totalmente que o *Big Data* forneça informações tão precisas e atualizadas, e optam por realizar a extração de dados e realizar análises de maneira primitiva, ou como estão habituados desde o início (CRUZ, 2018).

Os dados e informações estão disponíveis em qualquer lugar e a qualquer momento, cabe a sociedade e as empresas restringi-los e utilizá-los da melhor

maneira, beneficiando ambos e gerando crescimento econômico. O *Big Data* ainda está em fase de amadurecimento e neste sentido é necessário controlá-lo para que não sejamos controlados por ele; por isso se faz necessário o uso das melhores técnicas e dos melhores profissionais, além da consciência da proteção da confidencialidade das informações utilizadas para não gerar riscos aos usuários.

2.4.1 *A Estrutura Real e ágil do Big Data na Tomada de Decisão no contexto Segurança da Informação*

O efeito combinado desses fatores nos ambientes organizacionais torna o gerenciamento de segurança e tomada de decisão muito mais complexo, com muito mais interdependência e um escopo mais amplo de responsabilidade. À medida que mais processos de negócios se tornam digitalizados, as equipes de segurança têm a oportunidade e o desafio de coletar e gerenciar mais dados. Os investimentos estão cada vez maiores em ferramentas de gerenciamento de registros, gerenciamento de vulnerabilidade, gerenciamento de identidade e gerenciamento de configuração (TABUENA, 2017).

Uma estratégia real de *Big Data* para o gerenciamento de segurança deve abranger todos estes três aspectos para solucionar adequadamente os problemas à mão: infraestrutura, ferramentas analíticas e inteligência. Abaixo na figura 4 veremos os três pilares de *big data* em gerenciamento de segurança:

Figura 4. Os pilares de *big data* em gerenciamento de segurança

Fonte: Sousa (2018).

Para extrair valor dos dados coletados, obter eficiência da atividade de gerenciamento de ameaças e usar a atividade conformidade para orientar a tomada de decisões, as equipes de segurança precisam adotar uma abordagem de *big data* para o gerenciamento de segurança, segundo Weber (2019, p. 148), isso significa ter:

Uma infraestrutura ágil de scale out para responder às mudanças no ambiente de TI e às ameaças em evolução, o gerenciamento de segurança precisa dar suporte a novas iniciativas de negócios que afetam as organizações, de novos aplicativos a novos modelos de entrega como mobilidade, virtualização, computação em nuvem e terceirização.

A infraestrutura de gerenciamento de segurança/tomada de decisão deve ser capaz de coletar e gerenciar dados de segurança em escala corporativa e deve ser dimensionada de acordo com as exigências atuais das empresas, física e economicamente, isso significar fazer “scale out” em vez de “scale up”, pois a centralização de todos esses dados será praticamente impossível, além disso, a infraestrutura precisa se estender com mais facilidade para adaptar-se a novos

ambientes e desenvolver-se rapidamente para dá suporte à análise das ameaças em evolução.

Ferramentas de lógica e visualização que dão suporte a especialidades de analistas de segurança, os profissionais de segurança exigem ferramentas analíticas especializadas para dar suporte a seu trabalho, alguns analistas exigem ferramentas para facilitar a identificação de eventos com alguns detalhes do suporte. Os gerentes podem exigir visualizações de alto nível e análise de tendências de medidas-chaves, os analistas de malware precisam de arquivos suspeitos reconstruídos e ferramentas para automatizar os testes desses arquivos.

Inteligência contra ameaças para aplicar técnicas analíticas de dados às informações coletadas, as organizações exigem uma exibição do ambiente atual de ameaças externas para correlação com as informações reunidas da própria organização, essa correlação é importante para os analistas obterem uma compreensão clara dos indicadores de ameaças atuais e sobre o que procurar. Embora as técnicas avançadas, como lógica preditiva e inferência estatística, sejam provavelmente importantes no futuro, é importante que as equipes de segurança comecem a se concentrar em abordagens básicas e em fases.

Comece pela implementação de uma infraestrutura de dados de segurança que possa crescer com você. Isso envolve a implementação de uma arquitetura que seja capaz de coletar informações detalhadas sobre registros, sessões de rede, vulnerabilidades, configurações e identidades, e também inteligência humana sobre o que os sistemas fazem e como eles funcionam (TABUENA, 2017).

Embora você possa começar pequeno, o sistema precisa basear-se em uma arquitetura sólida e distribuída para garantir o dimensionamento à medida que seus requisitos evoluem, o sistema deve dar suporte aos domínios lógicos de confiança, inclusive jurisdições legais, bem como dados para unidades de negócios ou diferentes rapidez e facilidade (por exemplo, mostrar todos os registros, sessões de rede e resultados de verificação de determinado endereço IP e sua comunicação com um sistema financeiro de produção).

Implemente ferramentas analíticas básicas para automatizar as interações humanas repetitivas, normalmente, um objetivo de curto prazo é criar um modelo que irá correlacionar informações visualmente a fim de reduzir o número de etapas necessárias para reunir todas essas informações em uma exibição (por exemplo, mostrar todos os registros e sessões de rede que envolvem sistemas que dão

suporte ao processamento de transações de cartão de crédito e que sejam vulneráveis a um ataque já verificado em outras partes da empresa).

Crie visualizações e resultado que deem suporte às principais funções de segurança, alguns analistas precisarão ver apenas os eventos mais suspeitos com algum detalhe de suporte, os analistas de malware precisarão de uma lista priorizada de arquivos suspeitos e dos motivos pelos quais eles são suspeitos, os analistas de perícia forense de rede precisarão de resultados detalhados de consultas complexas.

Outros precisarão revisar relatórios de conformidade agendados ou relatórios gerais usados para identificar tendências ou áreas para aprimoramento no sistema, o sistema também precisa ser aberto para permitir que outro sistema acesse os dados e use-os para agir contra um invasor, como colocar em quarentena ou intensificar o monitoramento sobre o eles estão fazendo (TABUENA, 2017).

Adicione mais métodos analíticos inteligentes, só nesse ponto é que a lógica mais complexa pode ser aplicada aos dados para dar suporte a essas funções, essa lógica pode incluir uma combinação de técnicas analíticas, como regras definidas para identificar um comportamento provavelmente ruim ou conhecido como bom.

Também pode incorporar técnicas avançadas de linha de base e criação de perfil comportamental que implementam técnicas estatísticas mais avançadas, como a inferência bayesiana ou a modelagem preditiva. Essas técnicas analíticas podem ser usadas juntas para criar um “modelo de influência” (um modelo que combina indicadores diferentes para “classificar” os problemas que o sistema identificou, a fim de levar o analista às áreas que exigem atenção mais urgente).

Aprimore o modelo continuamente, depois que o sistema estiver em funcionamento, ele precisará ser ajustado continuamente para responder aos vetores de ameaças em desenvolvimento e às alterações para organização, o sistema precisará ter habilidade para ajustar regras e modelo para eliminar o ruído, consumir dados adicionais dentro e fora da organização e incorporar funções de autoaprendizagem para aumenta o sucesso geral do sistema.

Em cada ponto o sistema precisará aproveitar a inteligência externa como informações para o modelo, o sistema também deve ser capaz de facilitar a colaboração acerca do conhecimento compartilhado, o sistema deve compartilhar resultados de consulta ou inteligência não estruturado publicamente, ou em um

modelo controlado com comunidades de interesse confiáveis ou com base em quem precisa saber, de acordo figura 5:

Figura 5. Etapas para a implementação de *big data* no gerenciamento de segurança e análise de multicritério de tomada de decisão (AMD)

Fonte: Sousa (2018).

O Security Brief apresenta seis orientações que auxiliam as organizações a começarem o planejamento para a transformação de dados orientados a *Big Data*, de seus conjuntos de ferramentas e operações de segurança e análise de multicritério de tomada de decisão (AMD), como parte de um programa de segurança orientado à inteligência. São elas:

Desenvolver uma estratégia de segurança cibernética integral: as organizações devem alinhar as suas capacidades de segurança por trás de um programa e uma estratégia de segurança cibernética holística que é personalizada para os riscos, ameaças e exigências específicas da organização.

Estabelecer uma arquitetura de dados compartilhados para a Segurança da Informação: a análise de *Big Data* requer informações coletadas de diversas fontes em diferentes formatos, uma arquitetura única permitindo que todas as informações sejam capturadas, indexadas, normalizadas, analisadas e compartilhadas, é um objetivo lógico (PREIMESBERGER, 2021).

Migrar dos produtos de ponta para uma arquitetura de segurança unificada: as organizações precisam pensar estrategicamente sobre quais produtos de segurança

elas continuarão a utilizar por vários anos, pois cada produto apresentará a sua própria estrutura de dados que deve ser integrada num quadro de análise unificado para a segurança (ALMEIDA, 2017).

Procurar por ferramentas de segurança de *Big Data* abertas e escalonáveis: as organizações devem garantir que os investimentos em curso sobre os produtos de segurança favoreçam as tecnologias que utilizam abordagens baseadas em análises ágeis e não ferramentas estáticas baseadas em assinaturas de ameaças ou limites de rede. As novas ferramentas da *Big Data* devem oferecer a flexibilidade da arquitetura para mudar conforme o negócio, tanto TI como as ameaças evoluem (MCAFEE, 2019).

Reforçar as competências da ciência da computação, enquanto as soluções de segurança prontas serão emergentes: as equipes de segurança podem não estar prontas para operá-las, a análise de dados é uma área em que uma equipe talentosa está ausente do mercado de trabalho, engenheiros de Tecnologia da Informação com conhecimentos especializados em segurança de Banco de Dados são escassos e sua contratação permanecerá em alta demanda por mais algum tempo (MANIYKA, 2020).

Como resultado, muitas organizações provavelmente se transformarão em parceiras ou terceirizadas para complementar as capacidades de análise de segurança interna. Potencializar a inteligência de ameaça externa, aumentar os programas de análise de segurança interna com os serviços de inteligência de ameaças externas e avaliar os dados sobre ameaças de fontes confiáveis e relevantes, após elaborado, a *Big Data* será a referência por trás de um modelo de melhor eficiência e rentabilidade nos negócios. A segurança desta aplicação não deve ser subestimada (KORTH, 2017).

2.5 A Aplicação da Análise Multicritério de Tomada de Decisão (AMD)

Na década de 70, diante da ineficiência de alguns modelos convencionais de pesquisa operacional para realizar análises de problemas gerenciais, surgiram novos estudos centrados em soluções construtivas. O surgimento da análise multicritério de apoio à decisão nesse período buscou atender essa necessidade.

A metodologia de Apoio Multicritério à Decisão (AMD) consiste em um conjunto de métodos e técnicas para auxiliar pessoas e organizações a tomarem

decisões considerando a multiplicidade de critérios envolvidos. Nesse contexto, as análises multicritério abordam técnicas qualitativa-quantitativas situadas entre as abordagens puramente exploratórias e pouco estruturadas de tomada de decisão, como o *brainstorm* e grupos de discussão, e os modelos quantitativos rigidamente estruturados de pesquisa operacional, voltados à otimização de funções-objetivo sujeitas a um conjunto de restrições (JANNUZZI, 2019).

Além disso, uma diferença expressiva entre a metodologia multicritério e as metodologias tradicionais de avaliação se dá pelo grau de incorporação dos valores subjetivos dos decisores nos modelos de avaliação, permitindo que uma mesma alternativa seja analisada de diferentes formas ao considerar os critérios de valor individuais de cada especialista. Por essa razão, na perspectiva de apoio à decisão, a consideração da subjetividade na construção de modelos de avaliação, além da objetividade inerente ao processo, constitui uma das principais vantagens dos atuais modelos multicritérios sobre os modelos clássicos de pesquisa operacional (VILLELA, 2019).

O AHP (*Analytical Hierarchy Process*) é um método de tomada de decisão multicritério proposto na década de 1970, pelo pesquisador americano Thomas L. Saaty, utilizado extensivamente para analisar e estruturar problemas complexos de decisão (HANINE *et al*., 2016).

Segundo Leite e Freitas (2016), o método AHP possui como vantagem ser mais conhecido academicamente e no meio empresarial. Além disso, este método consegue representar claramente as preferências dos decisores principalmente em situações em que predominam restrições qualitativas e o grupo de decisão é composto por pessoas com interesses e visões divergentes.

Nesse contexto, o AHP reduz o estudo dos sistemas complexos a uma sequência de comparações, aos pares, de fatores identificados criteriosamente. O método se caracteriza pela capacidade de analisar um problema e propor uma tomada de decisão por meio da construção de níveis hierárquicos, decompondo-o em níveis de atributos. Como os valores dos julgamentos das comparações paritárias são baseados em experiência, intuição e dados físicos, o AHP pode lidar com aspectos qualitativos e quantitativos de um problema de decisão, uma das maiores vantagens do método (BLOCK, 2017).

O método AMD é amplamente utilizado em uma variada gama de situações de decisão e é aplicado em disciplinas como administração, indústria, manufatura,

saúde e educação. Além disso, o método é um dos mais utilizados e difundidos no mercado mundial, uma vez que apresenta aplicação simples e intuitiva (MARAM *et al*., 2019).

Devido à presença de diversos fatores que influenciam a escolha de um software de mineração de dados, a tomada de decisão multicritério tem se mostrado uma técnica poderosa e adequada para resolver este tipo de problema de seleção (HANINE *et al*., 2016). Além disso, Jadhav e Sonar (2019), em seu artigo sobre a avaliação e a seleção de pacotes de software, concluem que o AHP é o método mais utilizado como técnica de avaliação para este problema.

Devido à popularidade do método e à validação de que este é amplamente utilizado para seleção de softwares, o AHP foi escolhido como técnica AMD para resolver o problema deste estudo. Isto posto, o método AHP é composto por seis passos (BLOCK, 2017):

a) Definir o objetivo;
b) Definir as alternativas;
c) Definir os critérios/subcritérios relevantes para o problema de decisão;
d) Avaliar a importância relativa de cada critério/subcritério;
e) Avaliar as alternativas em relação aos critérios/subcritérios; e
f) Determinar a avaliação global de cada alternativa.

De modo geral, o AHP apresenta um problema decisório que pode ser estruturado de maneira hierárquica (Figura 1). Nesse sentido, o topo da hierarquia aponta para o objetivo geral da análise e o nível seguinte indica os critérios considerados. Na próxima camada da estrutura hierárquica, estes critérios podem ser decompostos em sub critérios.

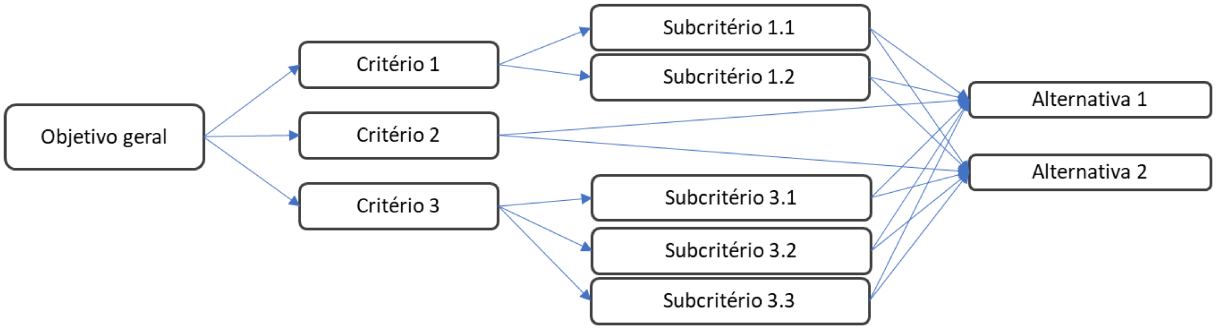
Por fim, no último nível, são encontradas as alternativas consideradas na análise, posicionadas na base para que cada uma seja analisada individualmente, somente sob a óptica das últimas ramificações diretamente associadas da estrutura. Dessa forma, um problema decisório complexo é subdividido em problemas menores abordados separadamente que, em seguida, são agregados para chegar na solução final do problema complexo.

Portanto, de acordo com Kriksciuniene *et al.* (2019), a aplicação do AHP significa que são exercidas quatro etapas principais: modelagem de problemas na

forma hierárquica; avaliação dos pesos dos fatores por comparações entre pares; agregação de peso no vetor de prioridade; e classificação das alternativas de decisão.

Vale ressaltar que, para definição dos critérios e subcritérios utilizados no problema de seleção do software de mineração de dados, foi utilizada como base a lista de critérios genéricos recomendados para uso em avaliações de software, criada por Jadhav e Sonar (2019). A partir dessa listagem, os decisores elegeram os critérios e subcritérios mais relevantes para a análise de acordo figura 6.

Figura 6 - Estrutura hierárquica do AHP.



Fonte: Autor da Pesquisa (2022).

Por fim, as matrizes de comparação paritária, construídas com base na escala fundamental de Saaty (Tabela 1), as razões de consistência e os pesos de prioridade são calculados com o auxílio da *AHP Priority Calculator*, uma ferramenta virtual que torna o processo das comparações aos pares mais automatizado e ágil, disponibilizada pela *Business Performance Management Singapore* (BPMSG).

Tabela 1 - Escala fundamental de Saaty.

ESCALA	AVALIAÇÃO	DESCRIÇÃO
Igual importância	1	Os dois critérios contribuem igualmente para os objetivos
Moderada importância	3	A experiência e o julgamento favorecem um critério levemente em relação ao outro
Mais importante	5	A experiência e o julgamento favorecem um critério fortemente em relação ao outro
Muito importante	7	Um critério é fortemente favorecido em relação a outro e pode ser demonstrado na prática
Extrema importância	9	Um critério é favorecido em relação ao outro com o mais alto grau de certeza
Valores intermediários	2, 4, 6 e 8	Quando se procura condições de compromisso entre duas definições, é necessário acordo

Fonte: Adaptado de Gomede e Barros (2015).

Para dar objetividade à tomada de decisão, após o fim da segunda guerra mundial, ocorreu um grande esforço para desenvolver métodos estritamente matemáticos para encontrar a melhor solução para um problema. Esses métodos fazem parte da otimização clássica, e estão atrelados à procura do valor máximo ou mínimo de uma função, sendo excessivamente rígidos nas decisões.

Os métodos de apoio à tomada de decisão envolvem quatro problemáticas: a) encontrar a melhor alternativa; b) agrupar as alternativas dentro de classes bem definidas, c) ranquear as alternativas em ordem de preferência e d) descrever como cada alternativa atende a todos os critérios simultaneamente, conforme proposto.

Quanto às técnicas para se tomar decisões a literatura aponta para uma gama acerca dessa temática. Neste trabalho foram apontadas as técnicas de tomada de decisão em três categorias, conforme citado por Yan *et al.,* (2016) nomeadamente: Técnicas multicritério de tomada de decisão (MCDM), técnicas de programação matemática (MP) e técnicas de inteligência artificial (AI).

A utilização de métodos multicritérios no apoio à tomada de decisão vem ganhando espaço nas organizações. Segundo Paganotti (2018), isso deve-se ao fato de que o planejamento embasado em análise multicritérios, tem como fundamento a definição de metas, variáveis e expectativas bem definidas, permitindo uma decisão mensurada, que por si só auxilia na construção de uma plataforma de gestão e controle com mais qualidade.

Ao se aplicar o método multicritérios no contexto de segurança da informação para avaliar serviços, o desafio é identificar e estabelecer o indicador de eficiência de cada critério, que pode ser extraído com base nos sistemas de gerenciamento da empresa, com vistas a alinhar as diretrizes do serviço às metas da organização. Deve-se ainda, estabelecer uma escala de importância para cada critério, de forma a representar dentro do sistema a flexibilidade que a variável pode suportar, pois, dos métodos de coleta das variáveis e dos pesos, dependerá a credibilidade em expressar para o sistema, seu limite de tolerância com a flexibilidade de cada critério.

Técnicas quantitativas que permitem realizar a avaliação de diversas alternativas considerando múltiplos critérios simultaneamente são conhecidas como métodos de decisão multicritério (ou métodos MCDM – *Multicriteria Decision*

Making), são vistos como ferramentas matemáticas, eficazes para resolução de problemas em que existem critérios conflitantes (RODRIGUEZ, 2020).

A vantagem proveniente da utilização de métodos multicritérios na segurança da informação, ocorre pelo fato de que não há, em geral, decisões que sejam simultaneamente ótimas sob todos os pontos de análise ressalta que os métodos MCDM são um conjunto de ferramentas para abordar difíceis decisões auxiliando gestores em situações de incerteza, complexidade e objetivos conflitantes. Como consequência da utilização desse método obtém-se a melhor seleção possível (SKINNER, 2019).

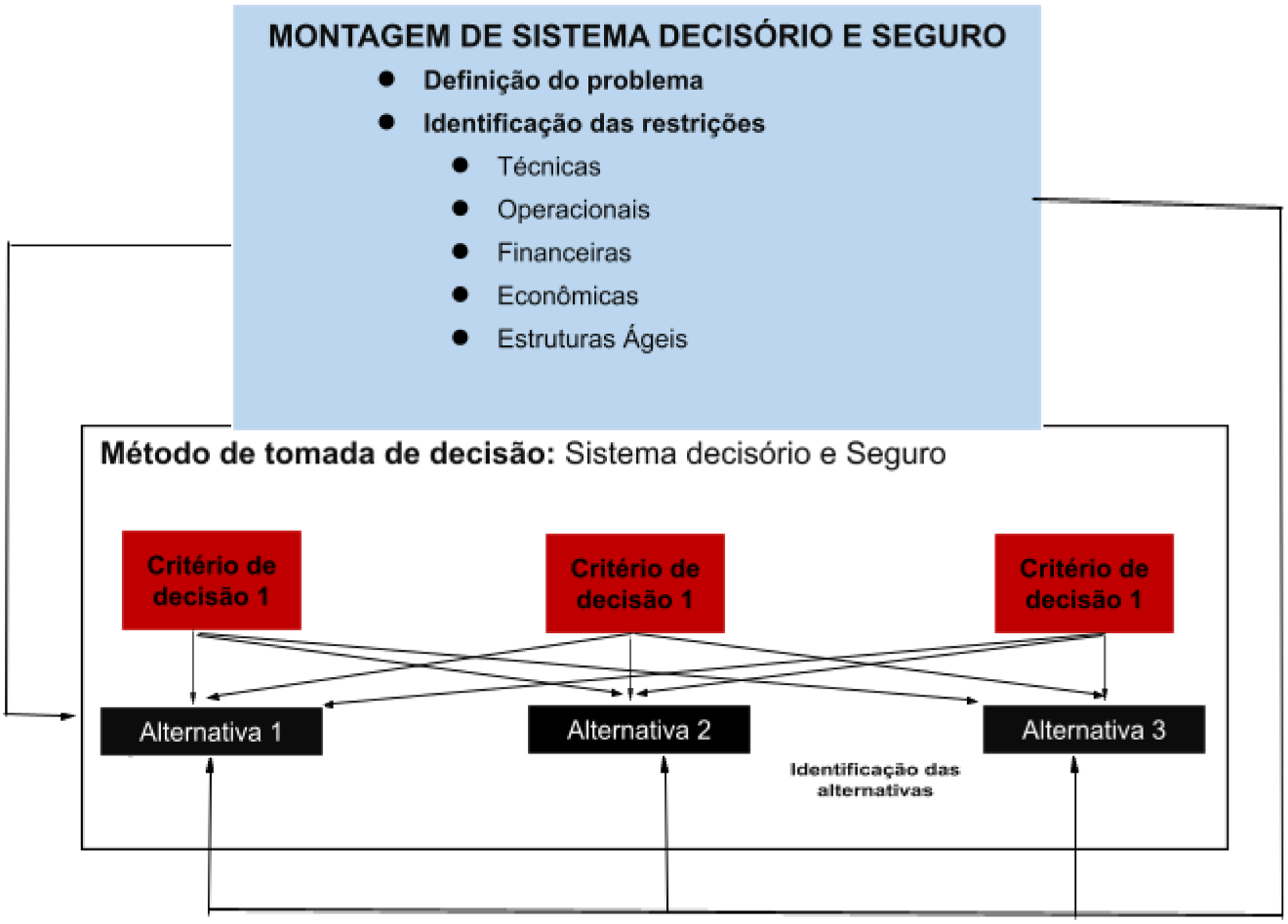
Os métodos multicritérios de tomada de decisão consideram diversos aspectos e avaliam as ações por meio de um conjunto de critérios, derivando de cada conjunto uma função matemática, que possuem como funcionalidade medir o desempenho de cada ação. Dessa forma, grandes esforços e avanços significativos foram feitos para o desenvolvimento de várias metodologias MCDM para solucionar diferentes tipos de problemas do mundo real

Como parâmetro o índice de segurança, que é definido como a relação entre a capacidade de carga disponível e a necessária, o peso de cada parâmetro é calculado a partir do vetor de prioridade, obtido através da média aritmética dos valores de cada linha da matriz normalizada, a aplicação do AMD, permite identificar a partir da perspectiva dos especialistas consultados, os pesos para cada critério no ranqueamento do *Big Data*, segundo a importância de cada um deles. A distribuição de pesos que foi adotada para este ranqueamento poderia ser totalmente modificada, de acordo com as preferências de análise desejadas pelo decisor (GOMES, 2018).

O grande problema é a definição das informações que devem ser geradas e, principalmente, a forma de integrá-las e de utilizá-las, assim, cabe às organizações o trabalho de selecionar, processar, analisar e escolher os dados que irão gerar as informações com veracidade e oportunidade que possam contribuir para com o processo decisório

Na figura 7 encontra-se um sistema decisório para o método multicritério de tomada de decisão na segurança da informação.

Figura 7 – Montagem do sistema decisório nos métodos de tomada de decisão na segurança.



Fonte: Adaptado de Andreotti (2019)

Podemos perceber de acordo a figura 7, a resolução de um problema multicritério de tomada de decisão na segurança da informação inicia-se com a definição do problema, a identificação das restrições, os critérios e as alternativas a serem avaliadas e selecionadas pelo tomador de decisão. Isto ocorre por meio do cruzamento dos critérios com as alternativas e dos critérios com o objetivo final.

2.6 Modelagem e Simulação do processo decisório no Big Data

A seleção de uma ferramenta de UTM é o seu desempenho, que está relacionado à velocidade de tratamento dos pacotes que trafegam através das suas interfaces de rede. Geralmente, o UTM está posicionado na borda da rede, filtrando todo tráfego entre a rede interna da empresa e a Internet e, caso não estiver

corretamente dimensionado pode tornar-se um gargalo, prejudicando a velocidade da conexão dos usuários ou deixando ameaças passarem (MESQUITA, 2018).

A modelagem de processos decisórios no Big Data, pressupõe que exista uma abstração em relação à realidade, acarretando uma simplificação sobre o funcionamento do sistema real. Os processos de negócio podem ser modelados a partir de um conjunto de atividades que têm por objetivo representar um processo de negócio existente (GOMES, 2015).

Os modelos podem ser classificados como estáticos ou dinâmicos, determinísticos ou estocásticos e ainda, contínuos ou discretos. Modelos estáticos visam representar o estado de um sistema em um instante, ou que em suas formulações não considera a variável tempo. Os modelos dinâmicos são formulados para representar as alterações de estado do sistema ao longo de um tempo definido na simulação. Os modelos determinísticos são os que em suas formulações não fazem uso de variáveis aleatórias, já os modelos estocásticos podem empregar uma ou mais variáveis com essas características (LIMA, 2017).

De uma forma diferente, mas com impactos igualmente significativos, as decisões nas organizações também são importantes para direcionar, obter vantagem competitiva e permitir a sua sobrevivência diante de cenários favoráveis ou adversos. Embora com o decorrer do tempo a experiência leve a tomar decisões mais acertadas e com mais rapidez, os problemas que exigem a decisão por um determinado caminho, também tendem a se modificar (GOMES, 2019).

As decisões programadas no Big Data, têm a característica de ser previamente determinadas, ou seja, são tomadas com consciência anterior para a sua execução, onde foram pesados os benefícios que a sua adoção trará. São geralmente aplicadas, a problemas repetitivos; as decisões não programadas têm a característica de ocorrer, na maioria das vezes, em decorrência de um acontecimento não previsto, a um problema que não é familiar, que pode ser benéfico ou maléfico para a organização (GOMES, 2018).

3 METODOLOGIA

3.1 Classificação da Pesquisa

A metodologia utilizada para o desenvolvimento deste estudo foi uma pesquisa qualitativa, através de uma revisão bibliográfica sistemática da literatura, por meio de teses, artigos, monografias e livros que relatam sobre o tema abordado e estão disponíveis gratuitamente em sites na internet.

Rousseau, Manning e Denyer (2018) suporta a abordagem estruturada da pesquisa bibliográfica, afirmando que ela deve ser um processo de acumulação abrangente, de análise transparente e de interpretação reflexiva de todos os estudos empíricos pertinentes a uma questão específica. Dessa forma, é destacado a importância em definir os princípios objetivos e garantir o rigor na elaboração de uma pesquisa bibliográfica.

Segundo Gunther (2016), uma vantagem da pesquisa qualitativa é utilizar “dados que ocorrem naturalmente para encontrar sequências em que os significados dos participantes são exibidos e, assim, estabelecer o caráter de algum fenômeno.

Gil (2018) afirmou que a pesquisa qualitativa deve ser utilizada para estudar um “fenômeno no seu contexto natural”, sem que o pesquisador tenha controle das variáveis presentes no caso a ser estudado. Esta pesquisa se refere à uma abordagem qualitativa, onde deseja pegar as essências do problema e acredita-se que será capaz de verificar a qualidade da temática em questão.

A revisão bibliográfica sistemática da literatura é habitualmente considerada como evidência de alta qualidade. É uma espécie de pesquisa, que segue protocolos específicos, e que busca entender e dar coerência a um grande *corpus* documental, especialmente, verificando o que se aplica e o que não se aplica num dado contexto (GALVÃO; RICARTE, 2019).

Deste modo, a revisão bibliográfica de literatura sistemática possui alto nível de evidência e se constitui em um importante documento para tomada de decisão nos contextos públicos e privados além de ser uma pesquisa científica composta por seus próprios objetivos, problemas de pesquisa, metodologia, resultados e conclusão, não se constituindo apenas como mera introdução de uma pesquisa maior, como pode ser o caso de uma revisão de literatura de conveniência.

Neste trabalho a revisão bibliográfica foi realizada de forma prévia ao projeto. Esta etapa almeja estudar o que foi descoberto e/ou proposto até o momento dentro da interseção entre seleção de projetos na indústria e tomada de decisão, para ser utilizado como base metodológica e procedimental a fim de propor um roteiro para seleção.

Dessa forma, esta revisão bibliográfica tem como objetivo responder às seguintes perguntas:

1) Quais são os principais autores no tema e em quais áreas se concentram os estudos?
2) Quais os principais métodos de análise multicritério à decisão no contexto *big data* que estão sendo utilizados para selecionar ou priorizar projetos?
3) Como estão sendo tratadas as avaliações dos especialistas no processo de tomada de decisão no contexto *big data* nas organizações?

Para desenvolver a pesquisa bibliográfica de forma estruturada, replicável e eficiente, será utilizada a abordagem proposta por Ertz e Leblanc-Proulx (2019), que se baseia na metodologia apresentada por Zhao e Strotmann (2015), com a diferença de que o primeiro é melhor adaptado para o contexto organizacional e implementa técnicas de visualização do mapeamento da bibliometria.

3.2 Procedimento de pesquisa

O roteiro de seleção de artigos permeia por quatro etapas: (1) delimitação do campo de estudo; (2) seleção das bases de dados, palavras-chave e critérios de pesquisa; (3) extração, limpeza e formatação e, por fim, a (4) análise de citação, cocitação, coocorrência de palavras-chave e visualização (Figura 8). Os artigos mais relevantes, perante a análise final da bibliometria, serão explicitados através da breve apresentação dos problemas estudados e métodos utilizados, que irão servir como guia para a definição dos métodos a serem utilizados no presente trabalho.

Figura 8 – Seleção de artigos.

Fonte: Autor da pesquisa (2022).

Os pontos focais do estudo podem ser avaliados pela centralidade dos nodos em uma rede de concorrência de palavras-chave, onde os nodos são as palavras-chaves e as arestas representam a interligação entre os nodos. Já a análise de coautoria, também utiliza uma representação em rede, sendo os autores representados pelos nodos e as arestas a relação existente entre os coautores de um trabalho.

Esta análise, por sua vez, permite identificar diversos padrões de interesse, como o número de artigos que um autor escreve, com quantas pessoas este autor escreveu um artigo, a distância entre autores na rede e como o padrão de colaboração varia entre temas e com relação ao tempo.

No próximo capítulo, encontra-se detalhado o procedimento de aplicação da pesquisa, desde a obtenção dos dados até o processo de validação do roteiro construído.

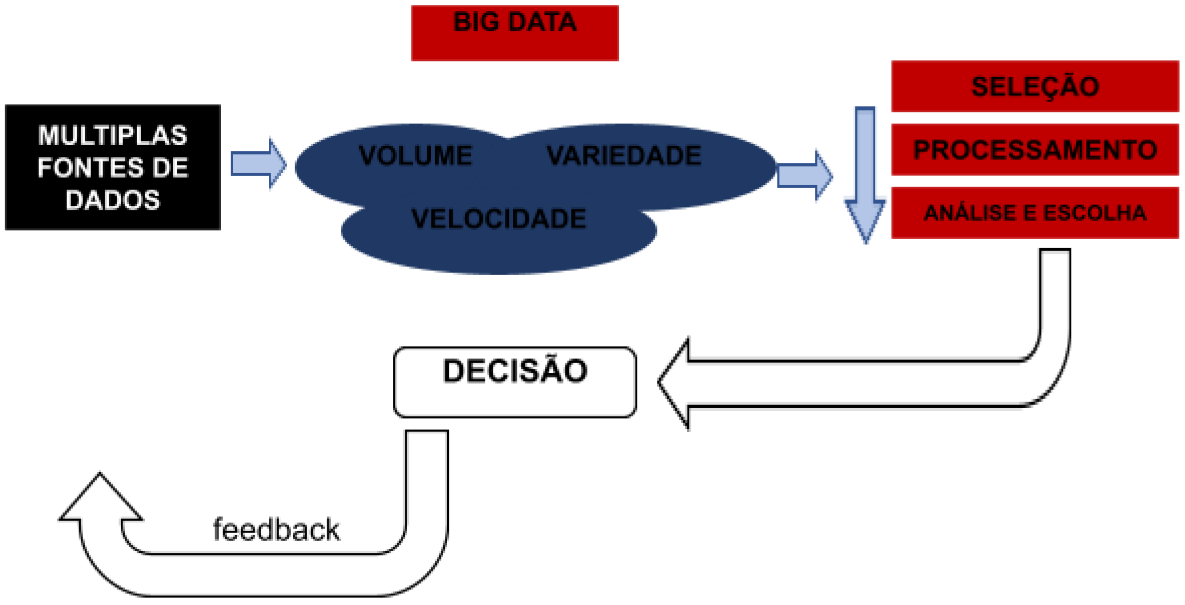
4 RESULTADOS E DISCUSSÃO

Nesta etapa, foi desenvolvido a questão do desempenho das organizações no mercado para definir de que forma com que as mesmas irão trabalhar com esse volume e essa variedade de informações, mediante a aplicação de métodos multicritérios aplicados pelo Big Data na tomada de decisão no ambiente organizacional.

Como a diversidade e quantidade dos dados é muito grande, naturalmente, a análise destes, traz uma gama enorme de informações para as empresas. Com o aumento da quantidade e acuracidade de informações disponíveis, gerando insights novos a partir da descoberta de dados, temos uma tendência de melhoria de decisões gerenciais, podendo gerar um diferencial competitivo.

A tomada de decisão é processo necessário para dar resposta a um problema em que alternativas de escolha são propostas para possíveis soluções que venham a gerar os melhores resultados para as organizações". Nesse mesmo sentido organizacional, Raskin (2020) analisa a tomada de decisão como o processo de responder a um problema, procurando e selecionando uma solução ou ação que irá criar valor, de acordo figura 9 abaixo:

Figura 9 – Ciclo de Decisão na Segurança da Informação do Big Data



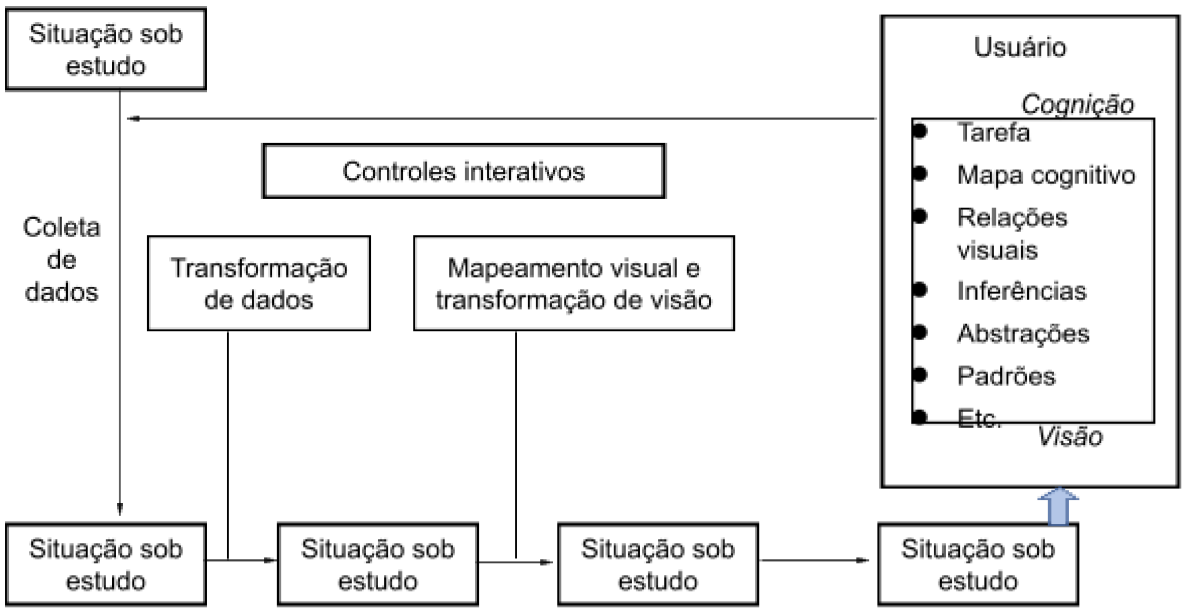
Fonte: Autor da pesquisa (2022)

Os benefícios da tomada de decisão orientada por dados têm sido demonstrados conclusivamente. A formulação estratégica de qualquer negócio sempre é feita a partir das informações disponíveis e, portanto, nenhuma estratégia consegue ser melhor que a informação da qual é derivada. Um dos pontos mais importantes no processo de gerar valor para as companhias na área de ciência de dados é o processo de modelagem. Nesta etapa diversos modelos (centenas ou milhares) são treinados com o uso de diversas técnicas de inteligência artificial.

O propósito de todas essas modelagens é encontrar as inter-relações entre as variáveis (colunas) com o uso de inteligência artificial, e gerar outputs como: predições: probabilidade de ocorrência de um evento dado um conjunto de características (leia mais sobre análise preditiva); previsões: projeções de como será o futuro de séries temporais; análise de perfis: identificação de personas, categorização entre indivíduos similares dentro do conjunto de dados, o que permite encontrar também grupos de outliers, os quais possuem características muito diferentes dos demais; criação de cenários: identificar grupos que impactam positiva ou negativamente em um target, uma variável a ser explicada.

A formulação da visualização de dados através do Big Data, pode ser elaborada a partir de alguns modelos de referência. No processo apresentado na figura 10, por exemplo, propõe-se a coleta dos dados brutos, seguido de sua organização, depois a escolha de uma estrutura visual para, então, ser efetuada a montagem da visualização, possivelmente no modelo de painéis interativos.

Figura 10 – Processo de elaboração de visualização de informações na Segurança da informação



Fonte: Adaptado de Valiati (2018).

Nesse modelo, é simples traçar um paralelo entre o processo de obtenção do conhecimento oferecido pelo Big Data, apresentado na sessão anterior, com o processo de elaboração da visualização de informações. Os dados brutos, exemplificando por esse contexto, são os dados extraídos de dispositivos com sensores que geram dados, de redes sociais, de dados de navegação na web, pesquisas em ferramentas de busca, cookies gerados nos sites, conversas por dispositivos de mensagem, vídeos postados em diversos servidores, hábitos de e-commerce e por aí vai.

A formulação estratégica de qualquer negócio sempre é feita a partir das informações disponíveis e, portanto, nenhuma estratégia consegue ser melhor que a informação da qual é derivada dois cenários macro de aplicações de tomadas de decisão orientada por dados, usando os princípios do Big Data: (1) as descobertas realizadas a partir de dados e (2) decisões repetitivas e em grande escala. O caso (1) está mais próximo do que hoje é chamado de *Advanced Analytics,* onde a empresa adquire novas informações apenas "olhando" para os dados.

A aplicação de Big Data como ferramenta de apoio à tomada de decisão dentro de uma organização, não depende somente do fato de se possuir dados para analisar, mas também se deve considerar como uma das principais “partes” recursos humanos com talentos destinados a exploração dos dados, ou seja, pessoas capazes de explorar os diversos volumes e fontes de dados com o objetivo de responder perguntas até então sem respostas, pessoas às quais dominem sim as tecnologias necessárias, mas que também entendam do negócio e saibam apresentar os resulta obtidos por meio do processo de análise dos dados.

O Big Data, no ato de métodos multicritérios, informa que é uma medida base para a efetividade da avaliação, ou seja, permite estabelecer um julgamento de preferência entre as ações. Os critérios podem ser metas, alvos ou objetivos almejados. Desta forma, torna-se necessária a análise de decisão que tem como objetivo ajudar o gestor a escolher as melhores alternativas.

Os critérios que são usados na análise de um conjunto de alternativas, na maioria das vezes são conflitantes, dificultando a resolução do problema. Portanto, a existência de uma metodologia de apoio à tomada de decisão torna-se um fator essencial. O interesse das empresas em adotar métodos de apoio à decisão tem

crescido rapidamente nos últimos anos devido ao surgimento de diversas técnicas com apoio computacional que tornaram os processos de tomada de decisão mais simplificados.

O processo decisório por meio da sugestão de alternativas ao decisor ao invés de apresentar uma solução para seu problema elegendo uma única alternativa verdadeira. Portanto, diante do agente de decisão haverá um conjunto de alternativas de soluções possíveis, no qual, ele irá escolher a melhor alternativa para solucionar o seu problema.

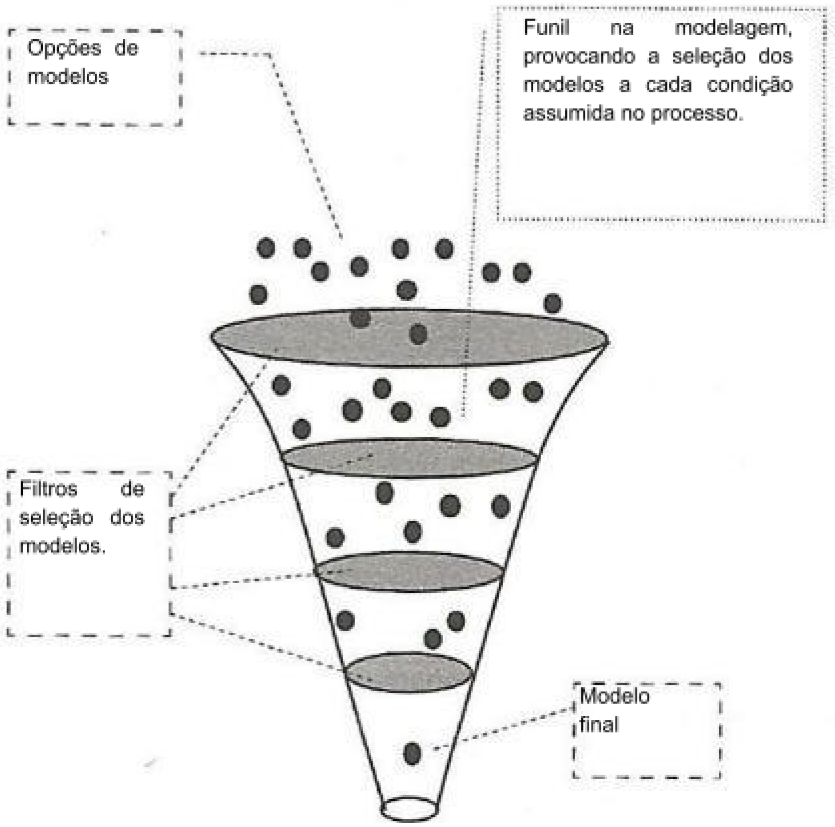
Em razão da existência de diversos métodos multicritérios é imprescindível a escolha de um método para resolução de problemas de decisão. Essa escolha deve ser adequada às características do problema em questão. Neste âmbito, devem-se considerar também as problemáticas de referência, para que seja possível alcançar o resultado esperado em determinado problema.

Com o objetivo de dar suporte a uma situação de tomada de decisão na prática, a seguir é apresentada uma abordagem para o processo de construção de um modelo e resolução de um problema de decisão. Vistos os principais elementos envolvidos no processo de tomada de decisão, este tópico faz uma integração dos demais assuntos apresentados, possibilitando ao decisor o desfecho para o processo decisório

O procedimento para resolução de problemas de decisão inclui a escolha do método mais apropriado para abordar o problema de decisão no processo de multicritérios do Big Data, em função de suas características fundamentais. Para o desenvolvimento da modelagem de um problema deve-se observar que há muitas possibilidades que levam a diversos modelos aplicáveis. Essas possibilidades estão associadas a diferentes hipóteses ou formas distintas de se estabelecerem os diversos ingredientes do problema, tais como: a forma do espaço de ações, a escolha dos atributos, etc.

A Figura 11 apresenta o processo de modelagem, nesta ilustração considera-se um funil por onde passam todos os possíveis modelos para o problema de decisão.

Figura 11 – Funil da obtenção do modelo de multicritérios do Big Data



Fonte: Adaptado de Almeida (2019)

Como podemos observar na figura 11, o filtro elimina, a cada decisão tomada no processo de multicritérios no sistema pelo analista, algumas possibilidades de modelos. Essas decisões consistem numa abordagem escolhida numa etapa qualquer do processo, ou em hipóteses assumidas em relação ao problema em estudo, ou ainda, outros fatores em relação à decisão analisada. Na passagem por cada filtro, há um número menor de possíveis formas de representar o problema, isto é, os modelos, até chegar a um único modelo (o modelo final) que será utilizado para análise do problema de decisão.

Dessa forma, os sistemas de Planejamento e Controle da Produção (PCP) desempenham um importante papel na busca contínua da melhoria no uso dos recursos de produção. Paralelamente, a modelagem de problemas de decisão, sob a ótica de múltiplos critérios, tem tido uma profusão no âmbito da engenharia de produção.

Como identificado na figura 11 acima, a consolidação de duas vertentes do conhecimento: a tomada de decisão multicritério, em inglês *Multicriteria Decision Making* (MCDM); e o auxílio multicritério à decisão (AMD), em inglês *Multicriteria*

Decision Aid (MCDA). A primeira vertente (MCDM) dedica-se à modelagem de problemas de decisão e segurança da informação baseada no emprego de modelos matemáticos à otimização, os quais requerem que o modelo seja robusto o suficiente para suportar.

Como ferramenta de auxílio a decisões complexas, os métodos MCDA buscam explicitar e mensurar a subjetividade do problema e agregar os objetivos com julgamentos de valor, auxiliando o decisor a entender o problema e as suas soluções, podem tratar dessa subjetividade, devido permitirem que os tomadores de decisão apresentem suas preferências e valores de forma explícita.

Um critério ou subcritério, além de ser visto como uma representação de um objetivo, também pode ser definido como uma função que mensura o desempenho adquirido no objetivo representado. Desta forma, pode-se definir o critério como uma função g (ou v) sobre o conjunto A, essa função representa as preferências do decisor conforme um objetivo (ou ponto de vista).

Para que uma família de critérios possa desempenhar de forma adequada sua função de apoiar um processo decisório, estabelecendo preferências sobre um conjunto de alternativas, algumas propriedades básicas de coerência precisam ser respeitadas. No critério verdadeiro a estrutura de preferência associada é uma pré-ordem completa. A estrutura de pré-ordem completa ocorre quando um par de relações binárias (a,b), em um conjunto de elementos A, corresponde à noção intuitiva de classificação em que existe a possibilidade de empate por similaridade. Essa estrutura consiste nas seguintes propriedades: a e b são exaustivas e mutuamente excludentes; b é assimétrica e transitiva; e a é simétrica e transitiva

O semicritério é assim denominado quando a estrutura de preferência associada é uma semi-ordem, que corresponde ao modelo limiar. Nestes modelos, existe uma faixa (limiar) de indefinição nos valores para aceitação de uma relação de preferência; por exemplo, o limiar de indiferença estabelece uma faixa de indefinição para que se concretize a relação de indiferença".

No procedimento de análise de multicritérios na segurança da informação utilizado pelo Big Data, consistem na transformação de valores das consequências, em cada critério, em valores em uma escala, geralmente de 0 a 1, assumindo uma função linear. Ou seja, esses valores entre 0 e 1, representam a menor ou maior satisfação do decisor em relação ao critério.

De acordo com a tabela 2, podemos avalizar a função de 3 procedimentos de normalização de um decisão em segurança da informação.

Tabela 2 – Procedimento de normalização de uma decisão em segurança da informação

PROCEDIMENTO 1	vj(ai) = [vj(ai) – Min vj(ai)] / [Max vj(ai) – Min vj(ai)], em que vj = critério, ai = alternativa, Min = valor mínimo, Max = valor máximo.
PROCEDIMENTO 2	vj(ai) = vj(ai) / [Max vj(ai)], isto é, divisão pelo valor máximo.
PROCEDIMENTO 3	vj(ai) = vj(ai) / ∑i vj(ai)], divisão pela soma.

Fonte: Autor da pesquisa (2022)

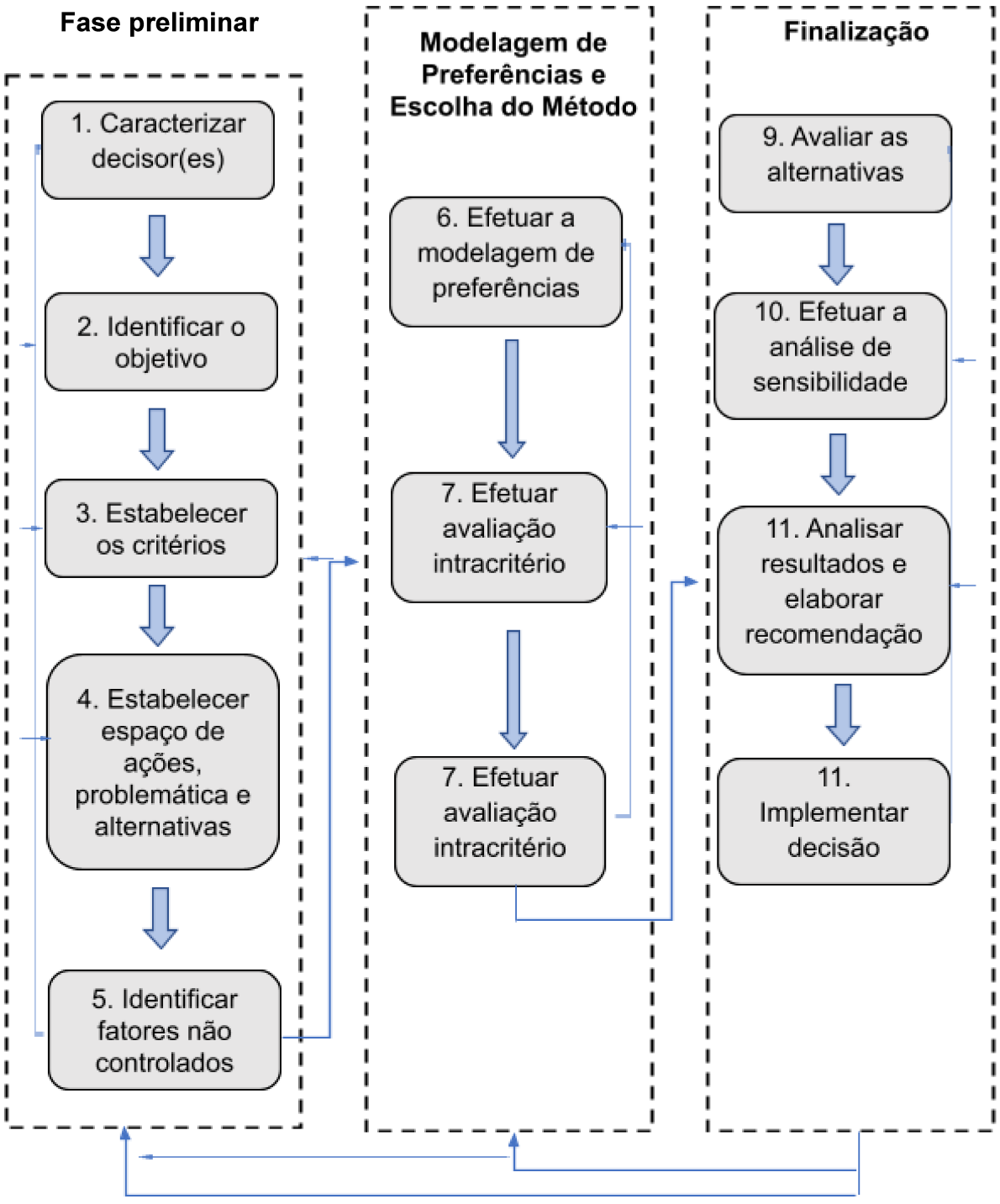
O cuidado ao utilizar um procedimento de normalização é importante porque em algumas situações tal procedimento pode influenciar os resultados obtidos, os procedimentos de normalização podem modificar a origem ou a unidade das escalas originais, então, é necessário utilizar um procedimento que seja compatível com o método MCDA a ser usado.

O estudo de multicritérios na segurança da informação mediante o Big Data, utiliza o método AHP, que é um dos primeiros métodos voltado ao ambiente de decisão multicritério e possivelmente o mais utilizado em todo o mundo. O método AHP propõe a estruturação do problema de decisão em níveis hierárquicos e através de julgamentos do decisor sobre critérios e alternativas, obtém-se um conjunto de pontuações ou pesos gerais.

O AHP utiliza uma abordagem hierárquica para estabelecer os critérios e identificar as alternativas, ou seja, esta forma de estruturação hierárquica dos critérios possibilita ao decisor, de maneira simplificada, uma maior compreensão e avaliação do seu processo e participação na estruturação do problema de decisão.

Para essa condução, o processo de tomada de decisão em um sistema de segurança, passa por diversas fases, abaixo no fluxograma 1, apresenta a fase preliminar, modelagem de preferências e escolha do método, finalização, estes procedimentos é representado abaixo:

Fluxograma 1 – Procedimento para resolução de um problema de decisão da segurança da informação do Big Data.



Fonte: Autor da pesquisa (2022)

A primeira fase, denominada fase preliminar, possui cinco etapas onde são estruturados os elementos básicos para a formulação do problema de decisão, estes elementos podem influenciar de forma decisiva o modelo final, que vai ser construído para análise do problema, conforme visualizado no fluxograma 1.

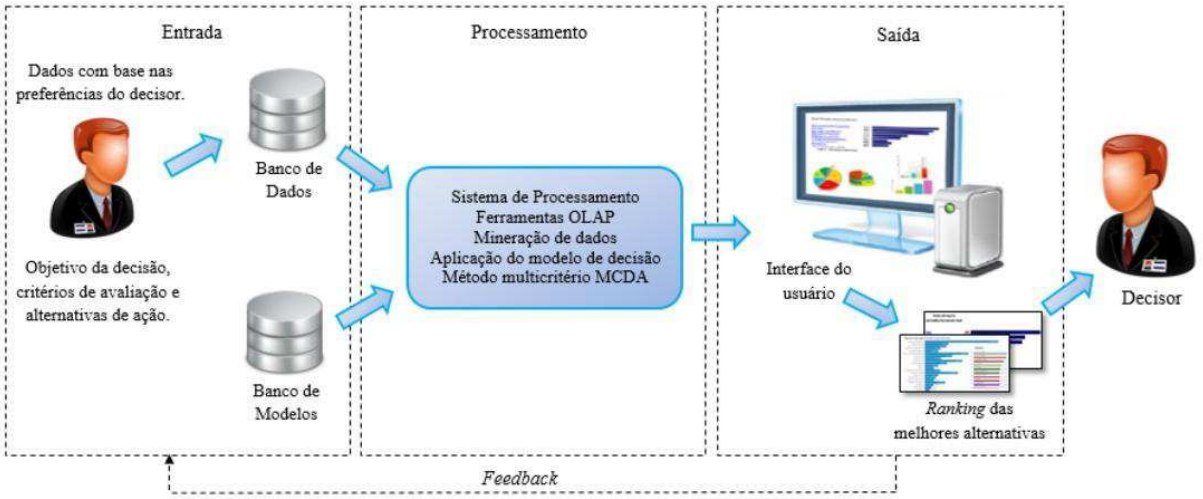
Na segunda fase, são desenvolvidas as etapas relativas à modelagem de preferências. Ao final dessa fase o método de decisão é escolhido e o modelo de decisão é construído, estando pronto para ser utilizado, ressaltando que revisões podem ser feitas, pois, como já foi visto, a abordagem de refinamentos sucessivos possibilita retornar da fase três para revisar algumas etapas nesta segunda fase. E por fim, na terceira fase têm-se o modelo consolidado, o desenvolvimento das etapas finais para a resolução do problema e a implementação da ação recomendada.

Com base em métodos multicritérios de apoio à decisão é possível desenvolver uma segurança de informações que irão apoiar as diversas etapas do processo decisório, desde a obtenção de informações e criação de modelos de tomada de decisão até a escolha das alternativas de solução para um determinado problema. Estes métodos associados à segurança da informação, auxiliam o decisor na busca da melhor solução com uma margem mínima de erros.

É importante ressaltar que dentre os sistemas de informações no Big Data existentes destacam-se os sistemas de apoio à decisão, devido às suas características, já apresentadas neste trabalho. Desta forma, métodos MCDA têm sido amplamente implantados em Sistemas de Apoio à Decisão. A metodologia MCDA tem como principal objetivo o auxílio à tomada de decisões complexas que envolvem múltiplos critérios e que, de certa forma, são critérios conflitantes. Essa metodologia permite que problemas complexos sejam abordados de maneira mais abrangente e realista, logo torna-se possível a modelagem de diversos fatores envolvidos no processo decisório.

A arquitetura do sistema de apoio à decisão associado a métodos multicritérios de apoio à decisão (MCDA) é composta também por banco de dados; banco de modelos; sistema de processamento e interface do usuário, na figura 12, a distinção está nos objetivos inseridos em cada componente da arquitetura e nos resultados apresentados.

Figura 12 – Arquitetura do Sistema de apoio à decisão associado a métodos multicritérios de apoio à decisão (MCDA)

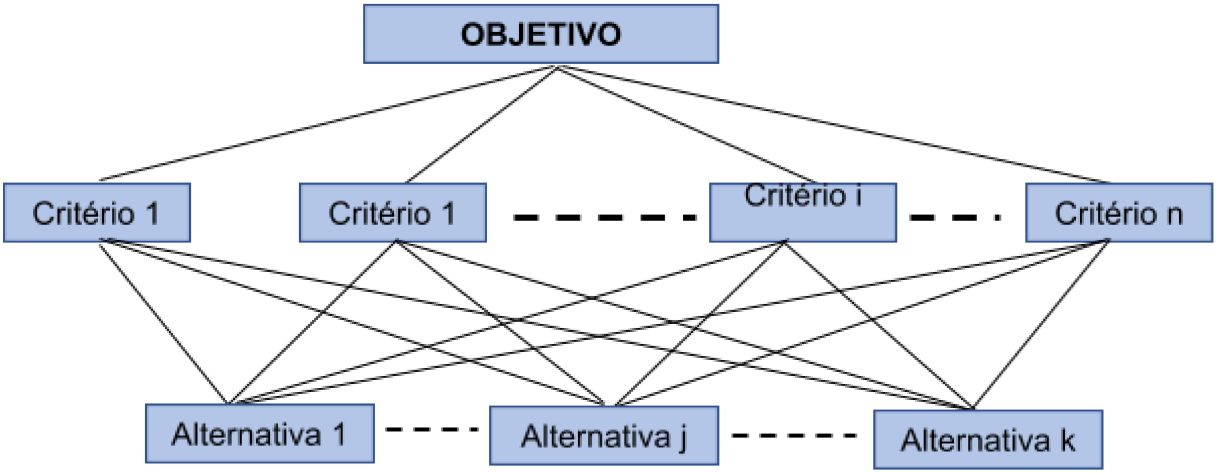


Fonte: Autor da pesquisa (2022)

Esta é a principal característica que distingue os sistemas de apoio à decisão associados a métodos MCDA dos sistemas de apoio à decisão tradicional: a possibilidade de agregar subjetividade ao processo decisório, isto é, todo o processo de tomada de decisão é construído embasado nas preferências do decisor. Desta forma, o decisor desempenha um papel fundamental na entrada destes dados. Com o objetivo definido e os critérios já estabelecidos, o decisor irá atribuir pesos (grau de importância) para cada critério, assim, as alternativas serão avaliadas de acordo com o grau de importância estabelecido pelo tomador de decisão.

Dessa forma, utilizando o método Big Data, pode lidar com aspectos qualitativos e quantitativos na segurança da informação, ou seja, estruturando a parte de multicritérios, ou seja, podem realizar um problema de decisão de acordo o fluxograma 2:

Fluxograma 2 – Hierarquia geral do Big data no processo de multicritérios



Fonte: Adaptado de Saaty e Vargas (2018)

Na representação do fluxograma 2, acima a importância relativa de um elemento em relação a outro, uma escala de avaliação adequada é introduzida, também chamada "Escala Fundamental de Saaty". Como podemos observar ela define os valores de 1 a 9 atribuições a decisões para a segurança da informação em comparação por pares de elementos em cada nível em relação a um critério no nível superior seguinte, como podemos perceber na tabela 3:

Tabela 3 – Escala fundamental de Saaty no modelo de Multicritérios

VALOR	DEFINIÇÃO	EXPLICAÇÃO
1	Importância igual	Duas atividades contribuem igualmente para o objetivo.
3	Importância fraca	Experiência e julgamento levemente a favor de uma atividade sobre a outra
5	Importância essencial ou forte	Experiência e julgamento fortemente a favor de uma atividade sobre a outra.
7	Importância demonstrada	Uma atividade é fortemente favorecida e sua dominância é demonstrada na prática.
9	Importância absoluta	A evidência favorece uma atividade em relação à outra com ordem de afirmação a mais alta possível.
2, 4, 6, 8	Valores intermediários	Quando se procura uma melhor relação de compromisso.
Recíprocos dos números não-zero acima	Se uma atividade / tem um dos números não-zero acima associados a ela quando comparada a uma atividade j, então j tem o valor recíproco quando comparado com i	----------
Número racionais	Razões fora da escala	Se a coerência precisa ser forçada obtendo-se n valores numéricos para toda a matriz

Fonte: Adaptado de Saaty e Vargas (2018)

Por meio de vários experimentos e pela utilização do Big Data na prática, a habilidade da Escala Fundamental para capturar informação e retratar, precisamente, a intensidade de preferência de um indivíduo está demonstrada. Num modelo em relação a prioridade de uma alternativa depende das outras alternativas que está a ser comparada. Se o resto é fraco, ele terá uma alta prioridade. Se o resto são fortes, sua prioridade será baixa. Isso pode levar a algo chamado "inversão do ranqueamento". Esse é um comportamento perfeitamente aceitável em um mundo relativo onde as coisas são interdependentes.

Analisando a crítica do alto número de comparações, o método Big Data diz respeito ao esforço para a tomada de decisão, que pode ser medido pelo número de comparações necessárias. Por exemplo, para uma decisão com 9 alternativas e 5 critérios, a aplicação do método Big Data necessitará de 190 comparações. Com o propósito de reduzir o número de comparações necessárias, "o que permitirá ao

grupo focar-se no debate e não na trabalhosa tarefa de preencher, por completo, cada matriz de comparações".

De acordo com os problemas de decisão na segurança da informação, surgem no momento em que existem pelo menos duas alternativas a serem escolhidas. Ou seja, essa escolha é pautada pelo propósito de atender múltiplos objetivos, muitas vezes conflitantes entre si. A fim de alcançar a resolução de um problema multicritério, apresenta-se a necessidade da construção de um modelo de decisão multicritério. Um modelo de decisão multicritério equivale a uma representação formal e com simplificação do problema de decisão com múltiplos objetivos enfrentados pelo tomador de decisão.

O estabelecimento de critérios válidos e abrangentes é uma etapa fundamental do processo de construção do modelo de decisão, visto que os mesmos servirão de base para a avaliação das alternativas previamente definidas, influenciando diretamente na qualidade da decisão a ser tomada. De acordo o fluxograma 2, relata sobre os critérios fundamentais para seleção de uma ferramenta de UTM é o seu desempenho, que está relacionado à velocidade de tratamento dos pacotes que trafegam através das suas interfaces de rede.

Geralmente, o UTM está posicionado na borda da rede, filtrando todo tráfego entre a rede interna da empresa e a Internet e, caso não esteja corretamente dimensionado pode tornar-se um gargalo, prejudicando a velocidade da conexão dos usuários ou deixando ameaças passarem.

Uma resposta do mercado da segurança da informação no Big data ao problema da defesa em camadas foi a criação de ferramentas, que agreguem todos estes pontos díspares de soluções de segurança em um único produto, conhecidas como sistema de gerenciamento unificado de ameaças (Unified Threat Management - UTM).

Ressaltando o gerenciamento unificado de ameaças (UTM), ou seja, um ferramenta que unifica diversos recursos de segurança em um único dispositivo na rede. O grande desafio da proteção de dados numa infraestrutura que lida com Big Data é que os controles de acesso à informação devem ser aplicados de acordo com o dado em si, e não somente aos sistemas e aplicações que o armazenam. Uma das formas de se atingir este controle é protegendo os dados mais críticos e sensíveis, tornando-os inelegíveis, por meio de técnicas já bastante utilizadas comumente, como por exemplo, a criptografia.

Dessa forma a análise de da aplicação de métodos multicritérios em sistemas para melhoria da segurança da informação e da tomada de decisão mediante ao big data, foram criadas com o objetivo de proporcionar uma forma mais conveniente de implementar o conceito da defesa em camadas, uma vez que há um único produto para configurar, gerenciar e monitorar. Toda a inspeção e análise dos pacotes da rede são feitas apenas uma vez, e as informações são compartilhadas entre os múltiplos recursos de segurança para aumentar a precisão das detecções.

CONCLUSÃO

Inicialmente, foi realizada uma busca por informações esclarecedoras referente ao conceito *big data* no contexto de aplicação de análise multicritério de tomada de decisão, para comprovar com fatos, a real necessidade de uma implementação. A partir das pesquisas, foram apresentadas tecnologias e ferramentas as quais poderiam vir a contribuir com o desenvolvimento da solução, de encontro à parte, foram apresentados estudos e delimitações para centrar o real foco do trabalho

Ao mesmo tempo que vivemos em um contexto de alta competitividade entre as diversas empresas, vivemos também um tempo onde o crescimento do volume de dados tem ocorrido de forma exponencial, tornando este cenário uma grande oportunidade de extrair conhecimento, informações e insights por meio da análise destes dados utilizando o Big Data.

Sua utilização poderá possibilitar a realização de análises descritivas as quais buscam compreender situações e eventos em tempo real, análises preditivas que buscam prever cenários futuros por meio da identificação e análise de padrões de comportamento, análises prescritivas que podem indicar ações a serem tomada mediante determinados cenários e situações e pôr fim a análise diagnóstica que busca a compreender as causas de um comportamento ou evento e os detalhes real situação de determinados eventos.

A presença de diversos fatores influencia a escolha do *Big Data* como recurso de segurança da informação e tomada de decisão, a aplicação da análise de multicritério se mostra expressiva para resolver este tipo de problema. O método AMD, amplamente utilizado em uma variada gama de disciplinas, é um dos métodos de apoio multicritério a decisão mais aplicados devido à sua execução simples e intuitiva. Ademais, a existência de diversas referências acerca do uso do AMD para a avaliação e a seleção de softwares sustentou a aplicação do método como o mais adequado para solucionar o problema deste estudo.

Os outros dois métodos foram escolhidos para efetivamente modelar o problema (AHP e PROMETHEE), pois possuem estruturas lógicas e de tratamento de dados bem definidas, são de fácil aplicação, podem ser aplicados tanto em grupo como individualmente e possuem transparência nos resultados. Assim, compreende-se que as respostas obtidas por meio destes métodos são de grande

importância para auxiliar o decisor no processo de tomada de decisão do Big *Data* no contexto de segurança da informação.

Com isso se pode concluir-se que o objetivo geral da pesquisa que é propor um modelo de análise multicritério como alternativa para suporte na tomada de decisões gerenciais no contexto do *Big Data* foi alcançado, esclarecendo até que ponto a aplicação de análise multicritério contribui de forma efetiva para o suporte na tomada de decisões gerenciais em uma organização.

REFERÊNCIAS

ALMEIDA, Adiel Teixeira de. **O conhecimento e o uso de métodos multicritério de apoio à decisão.** Pernambuco: Ed universitária da UFPE, 2017.

ANDREOTTI, M. M. **Método multicritério de tomada de decisão**: aplicação na Segurança da informação. Atlas, Rio de Janeiro. 2015.

AGGARWAL, Charu C. **Data Mining: The Textbook.** Nova Iorque: Springer, 2015.

BEAL, Adriana. **Segurança da informação:** princípios e melhores práticas para a proteção dos ativos de informação nas organizações. São Paulo: Altas, 2018.

BLOCK, Luiz Felipe. **Modelo de Decisão Aplicando um Método Multicritério de Apoio à Decisão para Apurar os Votos do Desfile das Escolas de Samba do Rio de Janeiro**. Trabalho de Conclusão de Curso – Ponta Grossa: Universidade Tecnológica Federal do Paraná, 2017.

CALDERS, Toon; CUSTERS, Bart. **What is data mining and how does it work?.** Studies in Applied Philosophy, Epistemology and Rational Ethics, v. 3, p. 27–42, 2013.

CAMILO, Márcio da Silva. **Uma breve história dos bancos de dados.** Rio de Janeiro: Altas, 2018. Disponível em: <http://www.sirmacstronger.eti.br/bd/introdbd.php>. Acesso em: 5 de janeiro 2022.

CARUSO, Carlos A. A.; STEFFEN, Flávio Deny. **Segurança em Informática e de informações** – São Paulo: Editora SENAC São Paulo, 2019.

CRUZ, Eduardo Picanço.; BARRETO, Cesar Ramos.; FONTANILHAS, Carlos Navarro. **O processo decisório nas organizações no contexto Big Data.** 1. ed. Curitiba: Intersaberes, 2018.

DATE, C.J Date. **Introdução à sistema de banco de dados**/ C.J Date: Tradução de Daniel Vieira – Rio de janeiro: Elsevier, 2019.

ELMASRI, Ramez; NAVATHE, Shamkant B. **Sistema de Banco de dados**. 4ª ed. São Paulo. Pearson Addison Wesley, 2015.

ELMASRI, Ramez; NAVATHE, Shamkant B. **Sistema de Banco de dados**. 6ª ed. São Paulo. Pearson Addison Wesley, 2018.

GEREMIA, Juliana. **Tutorial de Introdução a Banco de Dados**. Niterói, Rio de Janeiro. Dezembro, 2018.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa.** 4. ed. [S.l.]: Atlas, 2017.

GOMEDE, Everton; BARROS, Rodolfo Miranda de. **Utilizando o método Analytic Hierarchy Process (AHP) para priorização de serviços de TI:** um estudo de caso. VIII Simpósio Brasileiro de Sistemas de Informação, p. 408–419, 2015.

GOMES, Luiz F.A. **Da tomada de decisão à decisão:** agregando valor através dos métodos multicritérios em sistema da informação no contexto do Big Data. Revista de ciência e Tecnologia Política e Gestão para a Periferia, Recife, v. 2, n. 2, p. 117 – 139, 2018.

GOMES, C.F.S.; COSTA, H.G. **Aplicação de métodos multicritério ao problema de escolha de modelos de pagamento eletrônico por cartão de crédito.** Ubatuba-SP: XLIII-SBPO, Ago/2019.

GOMES, C. F. S.; SOARES, V. M. S. **Metodologias de análise e estruturação de problemas para auxílio à tomada de decisão:** Comparação e uma nova proposta. Revista Pesquisa naval, n. 14, p. 109-119. 2018.

GUNTHER, L. **Metodologia Científica**. 7. ed. São Paulo: Loyola, 2016.

LIMA Jr, F.R.; OSIRO, L.; CARPINETTI, L.C.R. **Métodos de decisão multicritério para seleção de fornecedores: um panorama do estado da arte.** Gestão e Produção, São Carlos, v. 20, n. 4, p. 781-801. 2017.

JADHAV, Anil S.; SONAR, Rajendra M. **Evaluating and selecting software packages:** A review. Information and Software Technology, v. 51, n. 3, p. 555–563, 2019.

JANNUZZI, Paulo de Martino; MIRANDA, Wilmer Lázaro de; SILVA, Daniela Santos Gomes da. **Análise Multicritério e Tomada de Decisão em Políticas Públicas:** Aspectos Metodológicos, Aplicativo Operacional e Aplicações. Belo Horizonte: INFORMÁTICA PÚBLICA, p. 69-87, 2019.

JOVIĆ, A.; BRKIĆ, K.; BOGUNOVIĆ, N. **An overview of free software tools for general data mining.** Opatija, Croácia: 37th International Convention on Information and Communication Technology, Electronics and Microelectronics, p. 1112–1117, 2018.

KENNETH Laudon e Jane Laudon, Kenneth e Jane. **Sistemas de Informações Gerenciais.** 9ª ed: Pearson. p. 163,164. 2018.

KORTH, Henry F., SILBERSCHATZ, Abraham. **Sistema de Banco de Dados**. 2ª ed. rev. – São Paulo: Makron Books, 2017.

KRIKSCIUNIENE, Dalia; SAKALAUSKAS, Virgilijus; LEWANDOWSKI, Roman. **Process mining of periodic rating scale survey data using analytic hierarchy process.** Springer International Publishing, v. 339, 2019.

LEITE, I. M. S.; FREITAS, F. F. T. **Análise comparativa dos métodos de Apoio multicritério a decisão:** AHP, ELECTRE e PROMETHEE. XXXI Encontro Nacional de Engenharia De Produção: Rio Grande do Sul, 2016.

MARAM, Venkataramana; SULTAN, Sultan Juma; OMAR, Mohd Faizal Bin; BOMMISETTY, Venkata Naga Ramakumar. **Selection of software in**

manufacturing operations using analytic hierarchy process. AIP Conference Proceedings, v. 2138, 2019.

MESQUITA, M.; COSTA, H. G.; LUSTOSA, L.; SILVA, A. S. **Planejamento e Controle da Produção.** 2 ed. Rio de Janeiro: Elsevier. 2018. 376 p.

MOR, Navdeep; SOOD, Hemant; GOYAL, Tripta. A Critical Review on Use of Data Mining Technique for Prediction of Road Accidents. In: SINGH, Vijendra; ASARI, Vijayan K.; KUMAR, Sanjay; PATEL, R.B**. Computational Methods and Data Engineering: Proceedings of ICMDE 2020**. v. 2. p. 141-149. Singapura: Springer, 2020.

MCAFEE, Andrew, BRYNJOLFSSON, Erik. **Big Data:** The Management Revolution. Harvard Business Review, outubro, 2019.

SÊMOLA, Marcos. **Gestão da Segurança da Informação:** uma visão executiva. Rio de Janeiro: Campus, 2016.

SOUSA, Paulo de Tarso Costa; MORAES, Sérgio A. S. **Sistemas de Bancos de Dados.** São Paulo: Atlas, 2018.

SCHMARZO, Bl. **Big Data:** Understanding how Data Powers Big Business. John Wiley & Sons, Inc., Hoboken, NJ, USA, 2016.

SKINNER, D. C. **Introduction to decision analysis:** A Practitioner's Guide to Improving Decision Quality. 3 ed. Florida: Published, 2019.

PAGANOTTI, José Antonio. **Processo decisório.** São Paulo: Pearson Education do Brasil - Coleção Bibliográfica Universitária Pearson, 2018.

PREIMESBERGER, Chris. **Big ideas about big data**. EWeek, 15 aug. 2021.

TABUENA, José. **What internal auditors should know about big data**. Compliance week, 2017.

TANKARD, Cezar. **Você realmente sabe o que é big data?** Blog da IBM, Rio de Jneiro. 2017. Disponível em: https://www.ibm.com/developerworks/mydeveloperwoerks/blogs/ctaurion/entry/voce_realmente_sabe_o_que_e_big_data?lang_em. Acesso em: 15 de janeiro de 2023.

RODRIGUEZ, M. V.; FERRANTE, A. J. **A Tecnologia de Informação e Mudança Organizacional.** Rio de Janeiro. Infobook, 2020.

VACCA, JOHN R. **Computer Forensics:** Computer Crime Scene Investigation. Massachusetts, USA: Charles River Media, 2017

VILLELA, Flávia Ribeiro. **Análise multicritério para a definição do índice de qualidade de fornecimento de energia elétrica por uma distribuidora.** Dissertação de Mestrado - Rio de Janeiro: PUC, 2019.

YAN, Xuefeng. et al. **Qualitative and Quantitative Integrated Modeling for Stochastic Simulation and Optimization.** J. Applied Mathematics, v. 2013, p. 831273:1 – 831273:12, 2016.

WEBER, K.*et al.,* **One size does not fit ali:** a contingency approach to data governance. ACM J. Data Informa. Quality 1, 1, Article 4, junho de 2019.