

ESCOLA DE GUERRA NAVAL

CC (T) Gláucia Botelho de Figueiredo

GESTÃO DE DADOS COMO SUPORTE ÀS PERSPECTIVAS LÓGICA E COGNITIVA
DA DIMENSÃO INFORMACIONAL NO CONTEXTO DAS OPERAÇÕES DE
INFORMAÇÃO

Rio de Janeiro

2021

CC (T) Glaucia Botelho de Figueiredo

GESTÃO DE DADOS COMO SUPORTE ÀS PERSPECTIVAS LÓGICA E COGNITIVA
DA DIMENSÃO INFORMACIONAL NO CONTEXTO DAS OPERAÇÕES DE
INFORMAÇÃO

Monografia apresentada à Escola de Guerra
Naval, como requisito parcial para a conclusão
do Curso Superior.

Orientador: CMG (RM1-FN) Jorge Luís de
Araujo Mello

Rio de Janeiro
Escola de Guerra Naval
2021

AGRADECIMENTOS

Agradeço profundamente a Deus Pai que me sustentou e me fortificou para execução desse trabalho. Agradeço ao meu esposo Renato pela paciência e parceria. Aos meus filhos Nícollas e Bernardo agradeço pela motivação diária para que eu cumpra meu dever. Ao CMG (RM1-FN) Mello agradeço pela orientação, disponibilidade e celeridade no atendimento de minhas dúvidas. Ao SO (Ref) Rodrigues agradeço a pronta disponibilidade.

RESUMO

A informação é o conceito basilar das Operações de Informação que são executadas por ações realizadas na dimensão informacional. A dimensão informacional é composta por três perspectivas inter-relacionadas – física, lógica e cognitiva – englobando pessoas, organizações e sistemas que apoiam o processo de tomada de decisão. A tomada de decisões requer insumos confiáveis, representados por informações e conhecimentos. Na atual conjuntura mundial, as Tecnologias da Informação e Comunicações desempenham papel preponderante no volume e variedade de informações, bem como na velocidade com que elas são transmitidas, gerando problemas a serem contornados. Nesse contexto, a execução de atividades de gestão de dados nas perspectivas lógica e cognitiva da dimensão informacional pode orientar a organização e o cruzamento de informações para uso dos tomadores de decisão. Nesta monografia são analisadas as áreas de conhecimento da gestão de dados e suas tecnologias associadas que podem contribuir para as Operações de Informações. As análises identificaram diversas tecnologias, que podem ser utilizadas de forma combinada, capazes de gerar novos conhecimentos aos tomadores de decisão, facilitar a seleção e a difusão de informações, lidando com o dinamismo das fontes de informação que, sem o suporte tecnológico, são de difícil gestão. A aplicação de tecnologias associadas a gestão de dados às Operações de Informação proporciona melhorias na utilidade, usabilidade e qualidade das informações a serem utilizadas pelos tomadores de decisão.

Palavras-chave: Operações de Informação. Gestão de Dados. Dimensão Informacional.

LISTA DE TABELAS

TABELA 1 – Relações entre gestão de dados e as OpInfo	38
TABELA 2 – Tecnologias associadas às áreas de conhecimento da gestão de dados que apoiam as OpInfo	39

LISTA DE ABREVIATURAS E SIGLAS

EB	Exército Brasileiro
Cel	Coronel
DMBOK	<i>Data Management Body of Knowledge</i>
DW&BI	<i>Data Warehouse e Business Intelligence</i>
ICT	Instituições Científicas e Tecnológicas
MB	Marinha do Brasil
OpInfo	Operações de Informação
PLN	Processamento de Linguagem Natural
TIC	Tecnologias da Informação e Comunicações

SUMÁRIO

1	INTRODUÇÃO	7
2	A DIMENSÃO INFORMACIONAL E AS OPERAÇÕES DE INFORMAÇÃO	10
2.1	Dimensão Informacional	11
2.2	Operações de Informação	12
3	A GESTÃO DE DADOS E SUAS ÁREAS DE CONHECIMENTO	14
3.1	Gestão de Dados	14
3.2	Áreas de Conhecimento de Gestão de Dados	15
3.2.1	Fusão de Dados	17
3.2.2	Ciência de Dados	18
4	GESTÃO DE DADOS NA DIMENSÃO INFORMACIONAL	18
4.1	Gestão de Dados na Perspectiva Lógica da Dimensão Informacional	19
4.2	Gestão de Dados na Perspectiva Cognitiva da Dimensão Informacional	21
4.3	Gestão de Dados e as Operações de Informação	22
4.4	Tecnologias para Gestão de Dados como suporte à Dimensão Informacional	23
4.5	Exemplo de aplicação da Gestão de Dados nas Operações de Informação	29
4.6	Síntese da análise	31
5	CONCLUSÃO	32
	REFERÊNCIAS	34
	APÊNDICES	38
	ANEXOS	40

1 INTRODUÇÃO

Informação é a base para geração do conhecimento que será transmitido e utilizado para tomada de decisões (BRASIL, 2015). O 1º Volume da Doutrina de Operações Conjuntas do Ministério da Defesa (MD30-M-01) define que “informações são representações inteligíveis de objetos, estados ou acontecimentos nos domínios real, virtual ou intelectual humano” (BRASIL, 2020, p. 195), sem fazer distinção entre as informações processadas e não processadas por quaisquer tipos de processamento: manuais ou automatizados. Nesta monografia, será utilizado o termo “informação” para referenciar informações processados ou não processados, independente de sua fonte ou formato. No entanto, o Glossário das Forças Armadas (BRASIL, 2015) faz distinção entre informação, informação de combate, informação estratégica militar, informação geográfica e informação militar, criando, assim, categorias de informações utilizando como critérios suas fontes e seus usos.

A informação de combate, também chamada de informação operacional ou informação tática, refere-se aos dados do inimigo. A informação estratégica militar está relacionada à expressão militar do poder nacional para tomada de decisões dos altos escalões militares e civis. A informação geográfica está relacionada a dados que tem como referência a superfície terrestre. A informação militar refere-se ao planejamento (em tempo de paz) e a execução (em tempo de guerra) de operações militares (BRASIL, 2015).

Atualmente a geração, a coleta, o processamento, o armazenamento, a seleção e a disseminação de informações têm sido apoiados e, muitas vezes, viabilizados por meio de recursos, processos e técnicas de Tecnologias da Informação e Comunicações (TIC) (MINISTÉRIO DO PLANEJAMENTO, DESENVOLVIMENTO E GESTÃO, 2017). A evolução das TIC possibilitou o aumento do volume e da variedade de tipos, formatos e fontes de informações disponíveis e coletáveis para apoio aos tomadores de decisão. Sendo assim, desenvolver capacidades para lidar com a diversidade informacional, no ambiente operacional, torna-se um processo de suma importância.

O ambiente operacional engloba os aspectos condicionantes que circundam e influenciam a atuação das forças militares (BRASIL, 2015). Considerando que as informações são essenciais para determinar a forma como as decisões são tomadas, elas influenciam o ambiente operacional (BRASIL, 2018). Diante desse enfoque, pode-se dizer que as capacidades associadas a lidar com informações têm papel condicionante à atuação das forças

militares.

Segundo a Doutrina de Operações de Informações (EMA-335), o ambiente operacional contemporâneo pode ser analisado por meio de três dimensões: física, humana e informacional. A dimensão informacional é representada pelos indivíduos, organizações e sistemas utilizados para obter, produzir, difundir e atuar sobre informações (BRASIL, 2018). Destarte, a dimensão informacional propicia um ambiente operacional favorável à Força para tomada de decisão (BRASIL, 2015; BRASIL, 2018). Em função disso, a dimensão informacional deve ser considerada no planejamento militar. As ações realizadas na dimensão informacional constituem as Operações de Informação (OpInfo). A dimensão informacional é composta por três perspectivas inter-relacionadas – física, lógica e cognitiva – e engloba pessoas, organizações e sistemas que apoiam o processo de tomada de decisão (BRASIL, 2018).

Nas perspectivas lógica e cognitiva da dimensão informacional são incluídas as atividades para processamento de informações (BRASIL, 2018), com o propósito de gerar conhecimentos mais elaborados do que as informações em seu estado original, isto é, não processadas. Esses conhecimentos são subsídios para o processo de tomada de decisão. Nesse contexto, é possível pressupor algum nível de gerenciamento (HENDERSON; EARLEY, 2017). Gerenciamento de dados e informações inclui a habilidade de tomar iniciativas estratégicas (HENDERSON; EARLEY, 2017) e, nesse sentido, alinha-se com as OpInfo que almejam proteger os interesses estratégicos, que são de cunho político (BRASIL, 2018). Contudo, o volume e variedade de informações passíveis de uso pelas OpInfo podem gerar problemas a serem contornados. Informações não devidamente categorizadas, não definidas semanticamente, não contextualizadas e armazenadas sem políticas de segurança alertam para falhas de gestão de dados (HENDERSON; EARLEY, 2017). Nesta monografia, as expressões “gestão de dados”, “gerenciamento de dados”, “gestão de informações” e “gerenciamento de informações” serão utilizados de forma intercambiável.

Destarte, a presente monografia tem como objetivo geral identificar como a gestão de dados, como atividade-meio, pode fornecer suporte aos tomadores de decisões, considerando as perspectivas lógica e cognitiva da dimensão informacional, no contexto das OpInfo.

Para alcançar o objetivo geral deste estudo, os seguintes objetivos específicos devem ser atingidos:

I) analisar as áreas de conhecimento de gestão de dados que podem fornecer suporte à perspectiva lógica da dimensão informacional, contribuindo para que as

informações obtidas sejam processadas, categorizadas e armazenadas com segurança;

II) analisar as áreas de conhecimento de gestão de dados que podem fornecer suporte à perspectiva cognitiva da dimensão informacional, que podem ser utilizadas para apoiar a percepção, a avaliação, a análise e o cruzamento de informações a serem utilizadas pelos tomadores de decisão; e

III) sintetizar as tecnologias, associadas às áreas de conhecimento de gestão de dados, para suporte à dimensão informacional.

Embora a dimensão informacional possua também perspectiva física, que é a origem e o destino do ciclo das informações (BRASIL, 2018), ela não será incluída no escopo desta monografia, visto que este estudo se concentrará no apoio à gestão de dados para geração de conhecimento, o que não ocorre na perspectiva da dimensão informacional.

A relevância desse estudo está alinhada com a obra do Cel Alessandro Visacro¹ do Exército Brasileiro (EB) ao considerar a importância do ambiente informacional para conquista da vitória nas guerras pós-industriais (VISACRO, 2018). Portanto, face à necessidade de eficácia das OpInfo é relevante que não hajam falhas na gestão de dados. Falhas nesse tipo de gestão podem conduzir tomadores de decisões a análises incompletas.

Cabe enfatizar que este estudo têm características de uma assessoria de Estado-Maior na área de Tecnologia da Informação, que atua como atividade-meio, para corroborar com outras áreas de negócio, que têm como metas atingir objetivos que elevarão a vantagem estratégica de uma organização em relação a seus concorrentes. No contexto deste estudo, as organizações são representadas pelas OpInfo e seus propósitos de elevar sua capacidade informacional frente aos adversários.

Para este estudo, a metodologia adotada foi a pesquisa bibliográfica e documental. Para área de OpInfo, foram pesquisados manuais, doutrinas e outras publicações oficiais expedidas no âmbito do Ministério da Defesa (MD) e das Forças Armadas, primordialmente da MB. Ademais, foram pesquisados artigos e trabalhos científicos disponibilizados na Rede de Bibliotecas Integradas da Marinha (REDE BIM, 2021), no Repositório Institucional da Diretoria do Patrimônio Histórico e Documentação da Marinha (RI-DPHDM, 2021) e no Metabusador em Banco de Dados do Exército Brasileiro (EBConhecer, 2021) que incluam como palavras-chave “Operações de Informação”, “Dimensão Informacional” e “Capacidades

1 Cel Alessandro Visacro, graduado pela turma de 1991 da Academia Militar das Agulhas Negras, possui Curso de Infantaria da Escola de Aperfeiçoamento de Oficiais (1999); e Curso de Altos Estudos Militares da Escola de Comando e Estado-Maior do Exército (2008-2009). Foi Oficial de operações do 2º Batalhão de Força de Paz do 17º contingente brasileiro no Haiti (Porto Príncipe). Em outubro de 2020, o Cel Visacro assumiu o posto de Chefe da Divisão de Difusão do Centro de Doutrina do Exército (C Dou Ex), um dos setores do Comando de Operações Terrestres (COTER), em Brasília.

relacionadas à Informação” (em inglês: “*Information Operation*”, “*Informational Dimension*” e “*Information Related Capabilities*”).

Para a área de gestão de dados e tecnologias associadas foram pesquisados artigos científicos na base do IEEE *Xplorer*® (IEEEEXPLORE, 2021) que incluam como palavras-chave “Gestão de Dados”, “Gestão de Metadados”, “Interligação de Dados” e “Fusão de Dados” (em inglês: “*Data Management*”, “*Metadata Management*”, “*Data Interlinking*” e “*Data Fusion*”).

O restante desta monografia está organizada da seguinte forma. Na seção 2 é feita uma breve revisão da literatura sobre a dimensão informacional e OpInfo. Na seção 3 são apresentados os conceitos sobre a gestão de dados, suas áreas de conhecimento, além de conceituar fusão de dados e ciência de dados. Na seção 4 são analisadas as tecnologias, associadas às áreas de conhecimento de gestão de dados, que fornecem suporte às perspectivas lógica e cognitiva da dimensão informacional. E por fim, a seção 5 discute as conclusões do trabalho, suas contribuições, limitações e trabalhos futuros.

2 A DIMENSÃO INFORMACIONAL E AS OPERAÇÕES DE INFORMAÇÃO

Os decisores precisam estar cientes dos fatores e das condições que influenciam a execução de uma tarefa, mesmo que a influência se dê em períodos de tempo determinados. Em um ambiente operacional, a percepção dos decisores sobre os acontecimentos que se sucedem fornece as condições para focar no alcance do objetivo, criando a chamada consciência situacional (BRASIL, 2015).

Um baixo grau de consciência situacional, implica em falhas na tomada de decisões, refletindo em todo ambiente operacional, caracterizando um estado de desinformação (BRASIL, 2020), que pode levar a perda de foco no objetivo de uma tarefa ou até da missão.

Com o intuito de destacar a importância das informações no ambiente operacional, nesta seção são apresentadas conceituações sobre a dimensão informacional do ambiente operacional, no qual se executam as Operações de Informação. A seção 2.1 apresenta o conceito, as características e a importância da dimensão informacional. A seção 2.2 conceitua formalmente operações de informação, diferenciando-as de Guerra de Informações, explicitando as capacidades relacionadas à informação.

2.1 Dimensão Informacional

A publicação JP 3-13 – *Operation Information* apresenta a conceituação de ambiente de informação como “uma agregação de indivíduos, organizações e sistemas que coletam, processam, disseminam ou agem sobre informações” (UNITED STATES, 2014, p. ix). Com base na publicação JP 3-13, as publicações doutrinárias MD30-M-01 – Doutrina de Operações Conjuntas – 2ª edição (BRASIL, 2020), EB70-MC-10.213 – Manual de OpInfo do Exército Brasileiro (EB) – 2ª edição (BRASIL, 2019) e EMA-335 – Doutrina de Operações de Informações (BRASIL, 2018) acolheram a conceituação de *ambiente de informação*, porém renomearam o conceito para *dimensão informacional*. As publicações do Ministério da Defesa e das Forças Armadas Brasileiras consideram que a dimensão informacional é formada pelas perspectivas física, lógica e cognitiva. Essas perspectivas relacionam-se e interagem com os elementos da dimensão, isto é, com os indivíduos, as organizações e os sistemas (BRASIL, 2018; BRASIL, 2019; BRASIL, 2020).

A perspectiva física está relacionada com a infraestrutura física, *hardware*, objetos, centros de Comando e Controle, eventos e processos (civis e militares) formando uma grande rede. A perspectiva lógica refere-se à proveniência, armazenamento e difusão das informações, incluindo requisitos de proteção da informação. A perspectiva cognitiva abrange as análises realizadas por elementos humanos (os indivíduos) responsáveis pela tomada de decisão (BRASIL, 2018; BRASIL, 2019; BRASIL, 2020). Assim, a perspectiva física fornece a base para iniciação das ações relacionadas às informações. Na perspectiva lógica ocorre organização e tratamento de informações, viabilizando a geração de conhecimento na perspectiva cognitiva. A geração de conhecimento se dá por meio das análises dos indivíduos, apoiados ou não, por tecnologias. As decisões serão materializadas na perspectiva física, gerando um ciclo para as informações.

Naturalmente, as três perspectivas da dimensão informacional existem em função da informação, embora cada uma tenha uma maneira individual de percebê-la. No entanto, qualquer nível de comprometimento na lide com as informações, gerando um estado de desinformação, em quaisquer dos três pontos de vista distintos, de modo isolado ou combinado, gerará impactos negativos no processo decisório (BRASIL, 2020).

Os desdobramentos negativos frutos do estado de desinformação podem incluir a ineficiência dentro e fora do Teatro de Operações. Um estado de desinformação pode provocar: (i) o uso indevido de meios de combate; (ii) ineficiência de suporte logístico; (iii) redução da confiança entre as forças amigas; e (iv) pânico popular (BRASIL, 2020). O pânico

popular pode ser facilmente promovido nos dias atuais², por meio de mídias sociais que podem levar a êxodos de uma população para áreas de apoio ou outras localidades gerando caos nas fronteiras e na quantidade de itens de essenciais ao sustento humano e animal.

Face às consequências da desinformação, é fato de que ela deve ser eliminada. Nesse sentido, a dimensão informacional assume um nível estratégico, haja vista a necessidade de controlar a percepção dos diversos públicos-alvo, ou seja, compreender e influenciar a opinião dominante (BRASIL, 2019). A opinião e a percepção da sociedade em relação à realidade de um conflito, ou à iminência de conflito, pode influenciar as ações do decisor. Por isso, controlar a narrativa incorre na necessidade dos decisores, no âmbito da defesa nacional, em comunicar com eficiência, celeridade e qualidade (BRASIL, 2019).

Sendo a informação o cerne da dimensão informacional, é preponderante que ela seja preservada e que ela não seja destruída, dispensada, esquecida ou, até mesmo, substituída por uma outra informação sem critérios previamente estudados. Portanto, a dimensão informacional deve ser considerada no planejamento militar, incluindo o acompanhamento do conflito e sua manutenção após a cessação de ações no Teatro de Operações. Nesse cenário, surgem as Operações de Informação (BRASIL, 2018; BRASIL, 2019; BRASIL, 2020).

2.2 Operações de Informação

Para atingir os Objetivos Nacionais de Defesa, a MB emprega o Poder Naval. Para execução das tarefas do Poder Naval são efetuadas diferentes tipos de operações navais. Nesse contexto, situam-se as Operações de Informação (OpInfo) (BRASIL, 2017).

As OpInfo são um tipo de operação naval e, por isso, elas estão relacionadas a expressão militar do Poder Nacional (BRASIL, 2017). O Poder Nacional tem como propósito alcançar os Objetivos Nacionais, que são as metas de uma Nação ao longo de sua evolução histórico-cultural (BRASIL, 2015). Em função disso, o Ministério da Defesa em sua Doutrina de Operações Conjuntas (BRASIL, 2020) homogeneíza o conceito de OpInfo e o torna aplicável às três Forças Singulares. Além disso, a doutrina especifica o uso das OpInfo como suporte e/ou parte das Operações Conjuntas. Adicionalmente, na Doutrina de Logística Militar (BRASIL, 2016), o Ministério da Defesa reitera que os processos de logística conjunta, para uso nas Operações Conjuntas, requerem a utilização de informações

2 Na época da escrita desta monografia os dispositivos móveis pessoais, como celulares, eram corriqueiramente utilizados pelo público em geral para acessar mídias sociais, tais como redes de relacionamento social, que gerenciam ciclos de amizade entre quaisquer indivíduos e/ou organizações comerciais ou profissionais. Além disso, todo cidadão era capaz de criar conteúdos e disponibilizá-los ao público em geral, possibilitando a geração de verdadeiras lideranças a partir do mundo virtual, mas que se concretizavam em ações no mundo real.

continuamente para uma tomada de decisão eficaz.

A publicação JP 3-13 caracteriza OpInfo como o emprego integrado das capacidades relacionadas à informação, isto é, de competências para destabilizar os oponentes quando considerado o acesso e uso de informações (BRASIL, 2019) em apoio às outras operações militares para interromper, corromper ou usurpar a tomada de decisões de adversários e/ou adversários em potencial, enquanto protege a da própria força (UNITED STATES, 2014). No EMA-335 é ressaltado que as OpInfo podem compor o esforço principal de uma operação ou podem contribuir para o esforço de outros tipos de operações militares (BRASIL, 2018). Clarck (2010) afirma que as OpInfo viabilizam a aplicação de ações dissuasórias não letais e flexíveis, evitando um conflito com capacidades letais pelo uso de sistemas de armas. Nesse sentido, as OpInfo são conduzidas desde o tempo de paz, adequando a dimensão informacional. Caso as ações dissuasórias não gerem efeito, dando início a um conflito, as OpInfo são transformadas na Guerra de Informação (BRASIL, 2018), uma espécie de conflito não-armado focado em prejudicar a comunicação do adversário e as informações que eles utilizam como subsídios para tomada de decisões (BRASIL, 2019). Portanto, as OpInfo podem ser protagonistas ou coadjuvantes no ambiente operacional, intercalando-se nesses papéis.

Independente do papel desempenhado, as OpInfo são centradas nas capacidades relacionadas à informação. Operações Psicológicas, Comunicação Social, Guerra Eletrônica e Guerra Cibernética são capacidades relacionadas à informação (BRASIL, 2020). De fato, algumas dessas capacidades representam outros tipos de operações ou ações militares. Mais que isso, uma pode influenciar na outra. Como exemplo, uma Operação Psicológica, que busca influenciar a consciência situacional do inimigo e afetar suas escolhas, pode gerar outras ações de OpInfo. As atividades de Comunicação Social tem como propósito influenciar a opinião de um público-alvo: uma organização, a sociedade, a imprensa, as relações internacionais (BRASIL, 2018; BRASIL, 2020). Assim, embora a Comunicação Social não seja uma operação militar, ela representa um conjunto de ações que influenciam o rumo de um conflito. Logo, é essencial que haja uma sincronização de esforços para lidar com essas capacidades e com a execução simultânea dessas operações. Ademais, é importante destacar que embora as atividades de Inteligência³ não afetem diretamente a dimensão informacional, elas são uma componente importante para as OpInfo, pois as células de Inteligência obtêm dados de diferentes fontes (BRASIL, 2018).

3 Atividades de Inteligência englobam esforços para obter e analisar dados de modo a gerar conhecimentos para o país que, dada a importância, devem ser protegidos para sua uso eficiente (BRASIL, 2021a).

No tocante à informação em si, como motivação das OpInfo, o avanço da disponibilidade de acesso a TIC por diferentes públicos fomenta as comunicações e notícias quase que instantâneas por diferentes grupos e agências (estatais e não estatais). Como consequência da popularidade, o volume e variedade de informações está em plena expansão. No entanto, essa expansão, sozinha, não agrega valor às OpInfo. Compreender como as TIC afetam o ambiente operacional é o diferencial (RICHTER, 2009). Da mesma forma, é vital compreender como as informações afetam esse ambiente. Isso significa compreender o valor da informação em termos de sua utilidade e sua usabilidade, ou no sentido mais geral, em termos de seu uso (HILL *et al.*, 2000). Nesse contexto, ressalta-se a importância da gestão de dados para as operações militares (BRASIL, 2018).

Em face do exposto, otimizar o fluxo das informações a serem obtidas, tratadas, difundidas e utilizadas sob o enfoque das perspectivas da dimensão informacional e, por consequência, para efetiva execução das OpInfo, torna-se essencial a manutenção de um compromisso da dimensão informacional com a gestão de dados.

3 A GESTÃO DE DADOS E SUAS ÁREAS DE CONHECIMENTO

Nesta seção são apresentados os conceitos de gestão de dados e das áreas de conhecimento que organizam o processo de gestão, em consonância com o DAMA⁴ DMBOK2®. DMBOK® é sigla em inglês para *Data Management Body of Knowledge – second edition* (Corpo de Conhecimento de Gerenciamento de Dados – segunda versão). O corpo de conhecimento DMBOK® é o resultado do trabalho conjunto de profissionais técnicos e de negócios em todo o mundo, que iniciou-se em 1980 (DAMA INTERNATIONAL, 2021b). O DMBOK® estabelece um vocabulário comum para gerenciamento de dados. Além disso, ele é amplamente reconhecido hoje em dia como um recurso essencial para as organizações engajadas na promoção da gestão, administração e governança de dados. A seção 3.1 apresenta o conceito de gestão de dados. A seção 3.2 explica, brevemente, as onze áreas de conhecimento de gestão de dados, conceitua fusão de dados e ciência de dados e as situa na gestão de dados.

3.1 Gestão de Dados

A gestão é desenvolvida por meio de uma série de atividades. As atividades de

⁴ DAMA é a sigla para *The Data Management Association International*, organização fundada em 1980 para abordar os desafios do gerenciamento de dados. DAMA é uma associação sem fins lucrativos e totalmente voluntária. Mais detalhes em na referência Dama International (2021b).

gestão de dados são abrangentes, pois requerem habilidades técnicas, que envolvem o uso da tecnologia da informação, e não-técnicas, que estão relacionadas aos objetivos de negócio e sua estratégia (HENDERSON; EARLEY, 2017).

Segundo o DMBOK® a gestão de dados engloba planos, políticas e práticas para distribuir, controlar e proteger dados e informações. Além disso, a gestão de dados busca aumentar o valor dos dados e informações ao longo de seus ciclos de vida (HENDERSON; EARLEY, 2017).

Considerando os diferentes tipos de dados, há requisitos distintos de gestão, seja pelo formato (físico, digital), seja pelo conteúdo (texto, imagens, áudios, vídeos, etc.). Em função disso, a gestão de dados precisa lidar com diferentes tipos de dados. Nesse sentido, parte das áreas de conhecimento relacionadas à gestão de dados focam em aspectos de classificação e controle.

3.2 Áreas de Conhecimento de Gestão de Dados

A gestão de dados pode ser conduzida com base em diferentes níveis de abstração. O DMBOK® (HENDERSON; EARLEY, 2017) sugere que os níveis de abstração estejam segregados em onze áreas de conhecimento, que descrevem o escopo e contexto de um conjunto de atividades de gerenciamento de dados, conforme apresentado seguir:

- Governança de Dados → responsável pelo desenvolvimento de políticas e procedimentos para práticas de administração de dados alinhada com os objetivos estratégicos. Sendo conceitualmente abrangente, ela é necessária para consistência e balanceamento entre as demais áreas de conhecimento da gestão de dados.
- Arquitetura de Dados → engloba a representação dos dados em diferentes níveis de abstração, subsidiando controle dos ativos de dados e suas integrações.
- Modelagem e Projeto de Dados → projeta como os dados estão organizados e como eles se relacionam. Os modelos de dados podem possuir abstrações mais altas, como nos modelos conceituais, nos quais são reconhecidos os principais conceitos de um negócio/assunto, ou menos abstratos, como nos modelos que indicam detalhes de como os dados são armazenados.
- Armazenamento e Operações de Dados → inclui a implementação e o suporte a forma como será feito o armazenamento de dados.
- Segurança de Dados → define e implementa políticas e procedimentos para autenticação, autorização, acesso e auditoria de ativos de informação.
- Integração e Interoperabilidade de Dados → possibilitam o uso de dados distintos de

forma conjunta. Integrar e interoperar são atividades diferentes, mas com estreita relação e complementariedade. Integrar envolve a consolidação de dados, isto é, a preparação, transformação, adequação e interligação (inter-relação) de dados distintos, integrando-os. Interoperar engloba a habilidade de troca de dados entre diferentes soluções e/ou sistemas, sem transformar ou integrar dados, mas permitindo, também, o estabelecimento de interligações.

- Gerência de Documentos e Conteúdos → envolve o planejamento, implementação e controle de dados semi-estruturados (e-mails, páginas da Internet), não estruturados (documentos textos, PDFs) e dados binários (imagens, áudios, vídeos).
- Dados Mestre e de Referência → viabiliza o compartilhamento de ativos de dados em comum a diversas demandas e/ou sistemas, com foco na consistência.
- *Data Warehousing* (Armazém de Dados) e *Business Intelligence* (Inteligência do Negócio) → envolve o processo de suporte a tomada de decisões, com base na consulta e análise de dados históricos, ou seja, apoiada em fatos ocorridos no passado.
- Metadados → abrange a descrição dos dados, suas conceituações e seus relacionamentos, além de apontar o alinhamento ao negócio, de forma a viabilizar o entendimento do significado dos dados.
- Qualidade de Dados → emprega técnicas para assegurar que os dados atendem às demandas de seus consumidores.

Por meio da Governança de Dados as organizações implementarão as áreas de conhecimento adequadas à própria gestão de dados. A seleção das áreas de conhecimento deve estar alinhada aos objetivos estratégicos da organização. Assim, a implementação das áreas de conhecimento, de modo paulatino, alinha-se ao desenvolvimento crescente da maturidade em gestão de dados de uma determinada organização.

É importante ressaltar que a área de conhecimento de “*Data Warehousing* e *Business Intelligence*” fomenta o uso dos dados de forma conjunta, independente de suas proveniências. Todavia, como as fontes de dados podem não estar estruturadas no mesmo formato, processos de extração e de transformação são necessários. Da mesma forma, a área de conhecimento “Integração e Interoperabilidade de Dados” pode requerer processos de mapeamentos e ajustes para uso integrado dos dados, assim como fazer uso de fusão de dados. De acordo com Oxley e Thorsen (2004), integração é o processo de conectar sistemas, podendo necessitar a execução de tarefas de fusão de dados.

3.2.1 Fusão de Dados

A fusão de dados envolve complementar, correlacionar e estimar dados provenientes de diferentes fontes, tais como sensores, linhas de comunicação distintas, entre outras. Nesse sentido, dados de diferentes fontes de dados podem ser complementados para formar uma nova informação mais completa. Em outra situação, se mais de uma fonte de dados provê a mesma parte de uma informação, a fusão de dados pode aumentar a confiabilidade e a credibilidade da informação correlacionando-as e controlando as redundâncias. De forma mais ampla, dados de diferentes pontos de vista, da mesma fonte de dados ou de fontes distintas, podem ser fundidos gerando uma informação mais complexa, mais próxima da realidade (DE FARIAS *et al.*, 2019).

No que diz respeito à organização e armazenamento de dados, Mendes, Mühleisen e Bizer (2012) propõem que a fusão de dados pode resolver conflitos entre os valores diferentes atribuídos a um mesmo objeto, indivíduo ou situação provenientes de diferentes fontes de dados. Nesses casos, antes da fusão, um mapeamento precisa ser implementado (MENDES; MÜHLEISEN; BIZER, 2012). A Figura 1 apresenta um exemplo de três formas de representar a localização do município do Rio do Janeiro: (i) no formato Graus/Minutos/Segundos (Degrees/Minutes/Seconds (DMS)); (ii) em decimal, indicando latitude e longitude; e (iii) por meio de um identificador único de recurso para localização (Geo URI (*Uniform Resource Identifier*)).

DMS	22° 54' 10" S, 43° 12' 28" W
Decimal	-22.902778, -43.207778
Geo URI	geo:-22.902778,-43.207778

Figura 1 – Localização do município do Rio de Janeiro em três diferentes representações (GEOHACK, 2021c)

Em um processo de fusão de dados para controle de redundância, as três formas de representação da localização do município do Rio de Janeiro podem ser mapeadas em um conjunto de dados único. Isso possibilita que os interessados tenham mais confiabilidade nos dados.

Embora a fusão de dados não esteja descrita como uma das onze áreas de conhecimento do DMBOK®, ela refere-se ao tratamento de dados, de modo a transformá-los para ampliar, completar, contextualizar e definir suas semânticas. Isso significa que os dados passam por um processo de enriquecimento. Nesta monografia, a fusão de dados será considerada como mais um aspecto da gestão de dados, face ao seu compromisso no tratamento de dados.

O tratamento de dados como parte da gestão de dados fomenta possibilidades de análises avançadas, o que demanda processos avançados de gestão e governança de dados por meio de ciência de dados.

3.2.2 Ciência de Dados

O aumento do volume e da variedade de tipos, formatos e fontes de informações, como resposta à evolução das TIC, demandou o desenvolvimento de novas tecnologias, gerando um ciclo de evolução tecnológico com base em informações. Assim, para lidar com grandes volumes de dados, surgiram as tecnologias de *Big data*⁵. A exploração de *Big data* também está relacionado ao uso de técnicas de ciência de dados. A ciência de dados utiliza técnicas estatísticas, mineração de dados e aprendizado de máquinas para criação de modelos preditivos que possibilitam a exploração de padrões de conteúdos de dados, de forma a prever um comportamento futuro. Para analisar os modelos preditivos, ou seja, modelos de prognósticos, “o analista de dados usa o método científico (observação, hipótese, experimentação, análise e conclusão) para desenvolver e avaliar um modelo analítico ou preditivo” (ENAP, 2019, p.62). Sendo assim, enquanto a ciência de dados fornece ao tomador de decisões dados sobre prováveis eventos futuros, com base em predições, a dupla DW & BI fornece ao tomador de decisões subsídios para decidir com base em fatos históricos.

Assim, nesta monografia a ciência de dados será considerada como mais um aspecto da gestão de dados, face ao seu compromisso com a exploração e análise de dados.

Neste ponto, cabe enfatizar que a gestão de dados é aplicável a quaisquer tipos de organizações que almejem obter o melhor proveito das informações. Independente da forma como as organizações estejam estruturadas, o uso eficiente das informações envolverá controle, autonomia, autoridade e tomada de decisões tendo como base as informações.

4 GESTÃO DE DADOS NA DIMENSÃO INFORMACIONAL

A gestão de dados atua sobre as informações sob diferentes aspectos. A utilidade e a usabilidade das informações podem ser otimizadas por uma boa gestão. Por isso, faz-se necessária a análise de quais aspectos da gestão de dados podem apoiar a dimensão informacional.

Considerando os pressupostos teóricos da “Governança de Dados”, pode-se

5 “*Big data*” é uma expressão relativa à área da computação que trata da necessidade de lidar com grandes volumes de dados, obtidos de forma veloz e em variados formatos (ENAP, 2019).

considerar que, quando aplicada à dimensão informacional, essa governança deve ser guiada pelos objetivos estratégicos de uma OpInfo. Assim, significa dizer que governança de dados aplicada às OpInfo podem melhorar a gestão de informações, pois a governança permeia todas as áreas de conhecimento da gestão de dados.

A gestão de dados assume importância também quando considerado o fato de tratar da qualidade de dados. A disponibilização e uso de informações com qualidade nas OpInfo melhora a consciência situacional dos tomadores de decisão e daqueles envolvidos antes (no planejamento), durante o enfrentamento dos conflitos e após suas soluções, sejam estes armados ou não (BRASIL, 2020).

Nesse ponto, frisa-se que uma das características da dimensão informacional é o dinamismo, tendo a vista a necessidade de acompanhar a evolução tecnológica e os impactos dessa evolução em suas três perspectivas (BRASIL, 2020). A gestão de dados uma área em constante evolução, principalmente, com o advento do chamado *Big data*. Logo, sendo a dimensão informacional e a gestão de dados duas áreas dinâmicas em função dos avanços tecnológicos, uni-las pode ser, no mínimo, adequado.

Agregando a gestão de dados à dimensão informacional e, portanto, às OpInfo, na seção 4.1 são apresentadas as áreas de conhecimento de gestão de dados que podem apoiar as atribuições da perspectiva lógica da dimensão informacional, contribuindo para garantir que o processamento, a categorização, o armazenamento e o acesso às informações. Na seção 4.2 são apresentadas as áreas de conhecimento de gestão de dados que podem apoiar as atribuições da perspectiva cognitiva da dimensão informacional, contribuindo para apoiar a percepção, avaliação, análise e cruzamento de informações pelos tomadores de decisão. Na seção 4.3 apresenta as relações entre gestão de dados e as OpInfo. Na seção 4.4 são analisadas as tecnologias, associadas às áreas de conhecimento de gestão de dados, para suporte à dimensão informacional. Na seção 4.5 são apresentados exemplos de aplicação. Por fim, na seção 4.6 é apresentada uma síntese da análise.

4.1 Gestão de Dados na Perspectiva Lógica da Dimensão Informacional

Observando-se a abrangência da perspectiva lógica da dimensão informacional e as áreas de conhecimento da gestão de dados é possível se estabelecer um alinhamento de objetivos entre ambas.

A área de conhecimento “Armazenamento e Operações de Dados” fornece suporte aos aspectos ligados ao armazenamento das informações obtidas, tanto antes de qualquer tipo de processamento, quanto após serem processadas. O armazenamento adequado

para os diferentes tipos de estruturas de informações, isto é, para as distintas formas pelas quais as informações podem ser apresentadas, é a base para iniciação de qualquer outro processo ligado à informações. Como a perspectiva lógica da dimensão informacional também foca no armazenamento é possível estabelecer uma correspondência entre a área de conhecimento e a perspectiva. Adicionalmente, o armazenamento deve garantir a disponibilidade das informações aos interessados, assim como a preservação de seus conteúdos. Nesse sentido, a área de conhecimento “Segurança de Dados” corrobora com a preservação.

A “Segurança de Dados” abrange os aspectos lógicos para acesso às informações, como a implementação de mecanismos de autenticação, autorização de acesso e armazenamento de informações no modo criptografado. Como a perspectiva lógica da dimensão informacional tem atribuição de proteção das informações, alguma política (nesse caso, definida pelas OpInfo) podem determinar os requisitos de segurança apoiada pela área de conhecimento “Segurança de Dados”.

Mesmo após as definições de quem possui autorização de acesso às informações, é fato que as informações devem estar organizadas e descritas. Um diferencial para melhor eficácia no acesso às informações apoia-se na possibilidade de ter conhecimento sobre o assunto tratado pelas informações, sem precisar acessar todo o conteúdo da estrutura de armazenamento, ou seja, sem acessar as informações em si. Esse diferencial é viabilizado pela “Gestão de Metadados”, mais uma área de conhecimento da gestão de dados alinhada à perspectiva lógica da dimensão informacional.

Sendo os Metadados um conjunto de dados que descrevem outros dados, o grau de detalhamento das descrições dos dados varia de acordo com o gestor. Metadados podem incluir dados sobre a proveniência das informações, o assunto, a data da coleta, a fonte da coleta, a periodicidade de atualização, a forma como as informações estão organizadas (estruturadas), entre outras descrições (MILLER, 2011). Os metadados contribuem para contextualizar as informações e são aplicáveis para quaisquer tipos de organização de informações: (i) estruturadas e mantidas em sistemas informatizados de banco de dados; (ii) semi-estruturada, como conteúdos das páginas da Internet; (iii) não estruturadas, inseridas em documentos textos, confeccionados em editores de texto; e (iv) binárias, tais como uma imagem ou uma sequência de sinais obtidos de sensores organizados com alguma lógica, mas que não se caracterizam como um sistema informatizado de banco de dados. Significa dizer que Metadados aplicam-se também à área de conhecimento de “Gerência de Documentos e Conteúdos”, que lida com informações não estruturadas, semi-estruturadas e binárias. Assim,

“Gerência de Documentos e Conteúdos” aplica-se à perspectiva lógica da dimensão informacional.

A “Gestão de Metadados” é fator preponderante para a área de conhecimento “Integração e Interoperabilidade de Dados”. Nas OpInfo podem ser obtidas informações de múltiplas fontes tornando-se importante integrá-las e interoperá-las. Para a perspectiva lógica, a integração e interoperabilidade visa a execução de um conjunto de atividades para tornar possível que as informações sejam correlacionadas, descobrindo semelhanças, redundâncias ou formas distintas de representação de uma mesma informação. No entanto, como há situações nas quais técnicas de fusão de dados são necessárias, a perspectiva lógica pode, ainda, beneficiar-se dessa área de tratamento de dados para enriquecer as informações, tornando-as mais completas e confiáveis.

Confiabilidade é uma característica necessária às informações que serão utilizadas pelos tomadores de decisão. Embora as decisões façam parte da perspectiva cognitiva, é por meio da perspectiva lógica que as informações serão classificadas, tratadas e difundidas. Nesse sentido, a área de conhecimento “Qualidade de Dados” é aplicável na perspectiva lógica, pois auxilia na definição e seleção de quais requisitos as informações devem possuir para serem consideradas de qualidade. Os requisitos que determinarão a qualidade são definidos em literaturas especializadas como dimensões de qualidade e incluem, por exemplo, consistência, acurácia, oportunidade, entre outras⁶.

Em face do exposto, de um modo geral, o suporte da gestão de dados à perspectiva lógica engloba atividades para preparar o ambiente informacional para a perspectiva cognitiva.

4.2 Gestão de Dados na Perspectiva Cognitiva da Dimensão Informacional

Considerando que o ponto central da perspectiva cognitiva é o tomador de decisões humano, a gestão de dados deve apoiar a missão desses indivíduos. Assim, a gestão de dados deve alinhar-se às necessidades daqueles que têm a responsabilidade de difundir e atuar sobre as informações.

A área de conhecimento de “Gestão de Metadados” além de tornar possível aspectos de gestão atinentes à perspectiva lógica, também está relacionada com os objetivos da perspectiva cognitiva. Esse fato pode ser evidenciado pela necessidade do tomador de

⁶ Segundo o DMBOK®: (i) “Consistência” (*Consistency*) em qualidade de dados refere-se a uniformidade dos dados que tratam do mesmo elemento/assunto existentes no mesmo conjunto de dados ou em conjuntos de dados distintos; (ii) “Acurácia” (*Accuracy*) é definido como o grau de correteza do dado na representação de fatos/elementos do mundo real; e (iii) “Oportunidade” (*Timeliness*) refere-se ao grau de atualidade do dado.

decisões ter conhecimentos sobre as fontes de informações, para atribuir-lhes algum nível de credibilidade, por exemplo. Em sentido amplo, os Metadados são subsídios para que o tomador de decisões selecione quais conjuntos de informações utilizará em determinado contexto e objetivo.

Os Metadados também estão intrinsecamente ligados à área de conhecimento “*Data Warehousing* (Armazém de Dados) e *Business Intelligence* (Inteligência do Negócio)”, pois descrevem quais informações e como elas foram organizadas para operações de análise de informações. “*Data Warehousing* e *Business Intelligence*” (DW&BI) fornecem informações analíticas, no sentido de que, são extraídas de uma ou mais fontes de informações, transformadas e carregadas em uma nova estrutura de informações com múltiplas dimensões (múltiplos pontos de vista), que viabilizam consultas com variados níveis agregação de informações. Em função disso, fica evidente o alinhamento de DW&BI com a perspectiva cognitiva. Em suma, soluções de DW&BI disponibilizam ao tomador de decisões um ferramental para análise das informações por diferentes pontos de vistas, auxiliando na geração de novos conhecimentos com base em fatos históricos e, por consequência, ampliando suas percepções sobre o ambiente operacional.

De maneira semelhante, os Metadados também estão relacionados com a possibilidade de suporte da área de conhecimento de “Integração e Interoperabilidade de Dados”, das atividades de “Fusão de Dados” e “Ciência de Dados” na perspectiva cognitiva. A descrição de informações possibilita que o tomador de decisões compreenda as relações configuradas pelas tarefas de integração, fusão e ciência de dados para compreender cada informação ao seu dispor em sentido estrito.

Na perspectiva cognitiva, a “Ciência de Dados” atua como um conselheiro digital do tomador de decisões que, naturalmente, utilizará suas próprias experiências, cultura e conhecimentos para aceitar ou não as predições e “conselhos” gerados por meio de ciência de dados.

4.3 Gestão de Dados e as Operações de Informação

Com base na fundamentação teórica sobre as atribuições da gestão de dados e das atribuições das perspectivas lógica e cognitiva da dimensão informacional, é possível estabelecer algumas relações entre esses dois grupos de atribuições, conforme explicitado nas seções 4.1 e 4.2. Isto posto, em modo amplo, é possível identificar que a gestão de dados é aplicável às OpInfo.

Em resumo, as áreas de conhecimento “Armazenamento e Operações de Dados”,

“Segurança de Dados”, “Integração e Interoperabilidade de Dados”, “Gerência de Documentos e Conteúdos”, “Gestão de Metadados” e “Qualidade de Dados”, acrescida da “Fusão de Dados” atendem à perspectiva lógica da dimensão informacional. A perspectiva cognitiva da dimensão informacional é atendida pelas áreas de conhecimento “Integração e Interoperabilidade de Dados”, “*Data Warehousing* (Armazém de Dados) e *Business Intelligence* (Inteligência do Negócio)” e “Gestão de Metadados”, além só suporte da “Fusão de Dados” e da “Ciência de Dados”.

O APÊNDICE A apresenta as relações entre gestão de dados e as OpInfo. Embora a governança de dados seja um das áreas de conhecimento da gestão de dados, ela atua como uma guia para as outras áreas de conhecimento, controlando as decisões sobre as informações de uma organização. Nesta pesquisa, uma organização é representada pelas OpInfo que, assim como organizações, possuem objetivos estratégicos. Conforme pode ser observado no APÊNDICE A, as relações das áreas de conhecimento da gestão de dados e as perspectivas da dimensão informacional demonstram que métodos e técnicas aplicáveis aos diversos seguimentos que compõem a gestão de dados podem ser utilizados para apoiar os objetivos da dimensão informacional e, por consequência, as OpInfo. Esses métodos e técnicas são apoiados por tecnologias existentes mercado ou em desenvolvimento nos centros de pesquisa.

4.4 Tecnologias para Gestão de Dados como suporte à Dimensão Informacional

Cada área de conhecimento da gestão de dados fornece objetivos, diretrizes e resultados esperados de sua execução. Para efetivamente executar as diretrizes, há algumas tecnologias associadas. Uma tecnologia pode atender a mais de uma área de conhecimento ou tangenciá-las, ao fornecer subsídios de entrada para execução de outras áreas de conhecimento. O APÊNDICE B apresenta algumas tecnologias associadas às áreas de conhecimento da gestão de dados que apoiam as OpInfo.

De um modo geral, conforme pode ser observado no APÊNDICE B, o uso de tecnologias pode ser aplicado de forma conjunta com outras tecnologias que suportem a mesma área de conhecimento ou áreas afins. Significa dizer que as tecnologias podem ser utilizadas de forma isolada, combinada ou sequenciada.

Uma tecnologia associada às áreas de conhecimento de “Armazenamento e Operações de Dados” e “Segurança de Dados” e, de suma importância para OpInfo, inclui a implementação de Algoritmos Criptográficos de Estado. Dada a sua importância, os Algoritmos de Estado são formalmente definidos no inciso I do Art. 2º do Decreto 7.845, de 14 de novembro de 2012 como aqueles cuja lógica de implementação, assim como as chaves

utilizadas para cifrar e decifrar são estabelecidos pelo Estado. Sendo assim, esses algoritmos são de uso privativo de órgãos da administração pública no âmbito do Poder Executivo federal (BRASIL, 2012). Em face do exposto, o uso de algoritmos desenvolvidos por iniciativas privadas e que são disponibilizados ao público seja por meio de pagamento de uma licença de uso, seja obtido gratuitamente, não atendem demandas de segurança requerida para assuntos de Estado.

O uso da tecnologia de Algoritmos Criptográficos de Estado, apoia diretamente a área de conhecimento “Segurança de Dados”. Visto que as informações são essenciais às OpInfo, é preponderante que elas cumpram requisitos de segurança estabelecidos. A cifra de informações contribui com grande parte da segurança, pelo viés da perspectiva lógica, já que a segurança física, de acesso às informações e aos locais nos quais elas estão armazenadas é provida por outra perspectiva e por outras áreas. As tecnologias que implementam Algoritmos Criptográficos de Estado.

As tecnologias que implementam Algoritmos Criptográficos de Estado tangenciam a área de conhecimento “Armazenamento e Operações de Dados”. Como as informações a serem utilizadas pela OpInfo precisam ser armazenadas, como atribuição da perspectiva lógica, de modo a servirem de subsídios para uso na perspectiva cognitiva, deve-se considerar o armazenamento das informações de forma segura e acessíveis somente àqueles que tem necessidade de conhecer. Destarte, as informações cifradas são efetivamente armazenadas e decifradas para serem operacionalizadas. Em função disso, medidas de segurança devem ser estabelecidas. Nesse ponto, cabe destacar que o inciso XII do Art. 2º do Decreto 7.845 define medidas de segurança como aquelas que viabilizam a garantia do “sigilo, inviolabilidade, integridade, autenticidade e disponibilidade da informação classificada” (BRASIL, 2012, Art. 2, inciso XII).

Para identificar e definir a quais informações devem ser aplicadas medidas de segurança, faz-se necessário categorizá-las e contextualizá-las. As atividades e tecnologias associadas à área de “Gestão de Metadados” – aplicável tanto na perspectiva lógica, quanto na perspectiva cognitiva – podem fornecer suporte as OpInfo, pois viabilizam a adição de explicações às informações tais como, a data de coleta, se a informação é do tipo texto, imagem etc., a proveniência, incluindo se a informação já é de domínio público ou foi obtida por investigação, entre outras.

Dentre as tecnologias associadas à área de conhecimento “Gestão de Metadados” como suporte à perspectiva lógica é possível identificar o uso de uma das áreas de Inteligência

Artificial⁷ denominada Processamento de Linguagem Natural (PLN) para extração de informações de documentos textos e de páginas e conteúdos da Internet, por exemplo.

PLN pode ser aplicada a quaisquer tipos de dados na forma de texto para coleta de palavras-chaves que identifiquem seus conteúdos, de forma que o leitor poderá tomar conhecimento do que se trata um documento, sem ter que fazer a leitura completa dele ou até mesmo a leitura de um resumo. As palavras-chaves são descritores do documento e, por isso, são metadados, dados que descrevem outros dados e/ou informações. O uso de PLN para extração de metadados pode ser visto na pesquisa de Castro (2019).

O uso de PLN é bastante abrangente. PLN tem sido utilizado amplamente para buscas de informações em mídias digitais. Assim, as tecnologias existentes e implementadas podem servir à OpInfo para obtenção de informação de forma automatizada, com o intuito de facilitar e filtrar as informações que sejam úteis no âmbito de um determinado ambiente operacional. Logo, PLN é uma tecnologia alinhada à área de conhecimento “Qualidade de Dados”, pois torna-se útil para detecção de notícias falsas (*fake news*) contribuindo para a filtragem de informações que devem ser fornecidas aos tomadores de decisões. O uso de PLN para filtragem e seleção de informações pode ser consultado em Souza *et al.* (2020).

Ainda na perspectiva lógica, as tecnologias que implementam algum tipo de Vocabulário Controlado⁸ viabilizam a definição dos significados de termos, expressões, informações presentes tanto nos metadados quanto no próprio documento. Por conseguinte, essas tecnologias adequam-se às áreas de conhecimento de “Gerência de Documentos e Conteúdos” e “Gestão de Metadados”.

Com Vocabulários Controlados é possível sanar ambiguidades de significados que termos e informações possam possuir. Como exemplo, segundo o Dicionário Aurélio da língua portuguesa (FERREIRA, 2004), o termo “bateria” têm quatro definições: a primeira relacionada a um aspecto militar e outras três relativas, respectivamente, a itens de cozinha, à engenharia elétrica e a música. Todavia, a definição relacionada a assuntos militares difere completamente da definição dada pelo Glossário da Forças Armadas (BRASIL, 2015). Assim, no contexto da dimensão informacional, no qual atividades de difusão de informações são necessárias, será possível selecionar qual definição é a mais adequada para uma determinada

7 Inteligência Artificial (IA) é um conjunto de tecnologias para que computadores tenham autonomia e atuem como a mente humana. Tecnologias de IA atuam para obtenção e interpretação de informações, reconhecimento de padrões e regra e formulações de predições de eventos. IA possui vários seguimentos como o Processamento de Linguagem Natural (PLN), em inglês *Natural Processing Language* e Aprendizado de Máquinas, em inglês *Machine Learning* (GLIKSON; WOOLLEY, 2020)

8 Vocabulário Controlado é um conjunto organizado e finito de termos que representam um conhecimento e servem para organização de documentos de forma temática, ou seja, por assuntos. Exemplos de vocabulários controlados incluem Ontologias, Taxonomias e Tesouros (SMIT; KOBASHI, 2003; SOUZA, 2011).

informação, proporcionando clareza para o processo de seleção de informações a serem difundidas.

Em OpInfo, informações podem estar contidas em quaisquer tipos de documentos, incluindo documentos na forma de texto, imagens, áudios, vídeos etc. Para a área de conhecimento “Gerência de Documentos e Conteúdos” em suporte à perspectiva lógica da dimensão informacional, termos definidos em Vocabulários Controlados podem ser interligados aos documentos como parte da preparação, tratamento e implementação de controle desses documentos não-estruturados.

Para a área de conhecimento “Gestão de Metadados” em suporte à perspectiva lógica da dimensão informacional, vocabulários controlados possibilitam sanar ambiguidades e explicitar o significado dos metadados, conforme pode ser visto na pesquisa de TYGEL (2016). Na prática de mercado, há uma variedade de vocabulários controlados de livre acesso disponibilizados na Internet e que podem ser utilizado de modo online (por meio de acesso direto ao vocabulário) ou de modo offline (por meio da replicação do vocabulário para um local específico a definido pelo utilizador, podendo ser um local de uso exclusivo de uma organização). Além disso, vocabulários controlados podem ser criados por uma organização (SMIT; KOBASHI, 2003).

Para as OpInfo, embora o controle do uso de descritores seja atribuição da perspectiva lógica da dimensão informacional, ele é de pleno uso na perspectiva cognitiva, haja vista a necessidade do tomador de decisões selecionar os assuntos de interesse a uma situação, com eficiência e eficácia, e que podem corroborar para a tomada de decisões.

Para tomada de decisões, na perspectiva cognitiva da dimensão informacional, há, ainda, um outro conjunto de tecnologias: a Web Semântica⁹. Essa Web, provida e interligada considerando a semântica para definir as interligações, fomenta integração de dados, análise de dados e geração de conhecimentos.

Complementarmente, como suporte à perspectiva cognitiva nas áreas de “Gestão de Metadados” e “Integração e Interoperabilidade de Dados” o uso de Web Semântica e tecnologias associadas, como os Grafos de Conhecimento¹⁰ (*Knowledge Graphs*) possibilitam que o tomador de decisões navegue de uma informação a outra, identificando as inter-relações

9 Web Semântica é uma estrutura com base na Web (Internet) tradicional cujas ligações (*links*) vão além das ligações entre páginas da Internet. Na Web Semântica as ligações podem ser estabelecidas entre dados, que têm significados, e são estruturados de uma forma para serem automaticamente lidos (legíveis) por computadores (BERNERS-LEE; HENDLER; LASSILA, 2001).

10 Grafos de Conhecimento apresentam as relações entre um grupo de entidades (elementos do mundo real) por meio de grafos, cobrindo vários aspectos, de diferentes domínios de conhecimento, relacionados aos elementos presentes no grafo (PAULHEIM, 2016).

e, dessa forma, gere novos conhecimentos.

A Web Semântica e suas tecnologias associadas, quando aplicada no nível de metadados, tem como foco propiciar aos tomadores de decisões celeridade para selecionar as informações que serão analisadas e utilizadas para uma determinada situação no ambiente operacional. A celeridade é possível já que os metadados já apontam os assuntos e categorias das informações, facilitando a identificação do conteúdo por quem está buscando informações relacionadas a um assunto. O emprego conjunto de tecnologias associadas à Web Semântica e vocabulários controlados fornece um duplo benefício: a classificação dos termos dos metadados com base em Vocabulários Controlados que, após implementada, fornece subsídios para geração de interligações de dados por intermédio de metadados, ou seja, a interligação de metadados apontam as interligações de dados. O uso de tecnologias de Web Semântica e associado a vocabulários controlados para interligação de metadados pode ser consultado em Figueiredo (2020). Além disso, a aplicação de Web Semântica pode provisionar o emprego de tecnologias de Grafos de Conhecimento para a área de conhecimento “Integração e Interoperabilidade de Dados”.

De um modo geral, as ferramentas de suporte a Grafos de Conhecimento oferecem uma estrutura de grafo com suporte visual aos tomadores de decisões. Sobre essas estruturas podem ser implementadas ferramentas de análises de grafos que aplicam diversos tipos de métricas, como a influência de um elemento no comportamento de outros elementos apresentados no grafo, indicando, por exemplo, relações de dependência e de poder de um elemento sobre o outro. A aplicação de estruturas de grafo para geração de conhecimento pode ser consultada na pesquisa de Verona, Lopes e Campos (2018).

Adicionalmente, o emprego de tecnologias para a área de conhecimento “Integração e Interoperabilidade de Dados” é fundamental como suporte à perspectiva cognitiva da dimensão informacional, haja vista a necessidade de OpInfo em lidar com variedade e volume de informações provenientes de diferentes fontes de dados. Embora a mente humana possua amplas capacidades, as quais têm sido alvo de inúmeros estudos Inteligência Artificial, que tentam fazer com que máquinas imitem o funcionamento lógico da mente humana como já mencionado, é dificultosa a análise humana de grandes volumes de informações. Em função disso, tecnologias que empregam grafos para interligar informações, usualmente disponibilizam um rol de ferramentas para consultas personalizadas a esses grafos, em complemento a livre navegação por meio de cliques (seleções com o mouse) em elementos visuais. Exemplos de processos de integração e interligação de dados, bem como um ferramental de apoio aos usuários dessas ferramentas, podem ser observados na pesquisa

de Avelino (2020).

Cabe enfatizar que, além de integrações de dados via grafo, na prática de mercado há as tecnologias para integração e interoperabilidade de dados que incluem o desenvolvimento de serviços em alguma linguagem de programação, visões de dados disponibilizadas por Sistemas Gerenciadores de Banco de Dados, entre outros.

Ainda no tocante às questões de variedade de volume de informações, tecnologias empregadas no desenvolvimento de Ferramentas de Analíticas de Dados têm fundamental importância e estão alinhadas à área de conhecimento “*Data Warehousing* (Armazém de Dados) e *Business Intelligence* (Inteligência do Negócio)”. Nessas ferramentas é possível que o tomador de decisões perceba as informações por diferentes dimensões, podendo selecionar uma determinada informação e desdobrá-la, para compreender a informação com mais detalhes. Esse tipo de operação, no qual uma informação mais geral é detalhada para um nível mais específico, é chamado de análise *drill-down* (aprofundar). Da mesma forma, as ferramentas de análise permitem a operação inversa, quando uma informação mais detalhada é agregada, integrada, para gerar uma informação mais geral. Esse tipo de operação recebe o nome de *roll-up* (acumular, tornar-se mais geral) (KIMBALL; ROSS, 2013).

É importante compreender que as operações *drill-down* e *roll-up* são fundamentais para análises de volumes de informações pelos tomadores de decisão. Naturalmente, a geração dessas informações para análise, requerem algumas etapas de tratamento de dados, que podem incluir técnicas de fusão de dados, principalmente para informações provenientes de diferentes sensores e equipamentos distintos utilizados no ambiente operacional, de forma aumentar a confiabilidade e a credibilidade das informações coletadas. As tecnologias associadas à fusão de dados podem, ainda, fazer uso de vocabulários controlados, da área de “Gestão de Metadados” e “Qualidade de Dados” e alinham-se a tarefas atribuídas à perspectiva lógica da dimensão informacional.

Dentre as tecnologias associadas à exploração de dados por meio de “Ciência de Dados”, como suporte à perspectiva cognitiva, é possível identificar o uso de uma das áreas de Inteligência Artificial denominada Aprendizado de Máquinas. As tecnologias de Aprendizado de Máquinas viabilizam análises preditivas com base em avaliações estatísticas, como estimativas de probabilidade. A predição aumenta a percepção do tomador de decisões para responder a questões sobre o que é provável que aconteça em um determinado ambiente operacional.

4.5 Exemplo de aplicação da Gestão de Dados nas Operações de Informação

O ANEXO A é um extrato da publicação EMA-355 (BRASIL, 2018) contendo alguns questionamentos para orientar o Comandante, no papel de tomador de decisões, no planejamento do emprego das OpInfo no Ambiente Operacional. No item 6.1 da referência supracitada é apresentado um conjunto de questões relacionadas aos atores da dimensão informacional. Aplicando-se algumas das tecnologias de suporte à gestão de dados na dimensão informacional, é possível a obtenção de auxílio para algumas respostas.

Por exemplo, para o questionamento “Quais são as organizações que estão atuando na região?” (BRASIL, 2018, p. A-2), destacam-se as palavras “organizações” e “região” como objetivos a serem respondidos. Embasado nesses objetivos, poderiam ser aplicadas tecnologias de PLN como suporte à “Qualidade de Dados” e “Gestão de Metadados”. Com PLN é possível a varredura de um conjunto de documentos, páginas da Internet, bases de dados, entre outros, para obtenção automática de expressões e palavras-chaves que possam referenciar ou estar relacionadas às palavras “organizações” e “região”. Para fontes de dados de proveniência não-confiável, como alguns conteúdos da Internet, PLN pode auxiliar na detecção de notícias falsas e já sinalizar essas fontes de dados como não-confiáveis, contribuindo para a qualidade de dados. As palavras-chaves extraídas podem ser utilizadas como metadados para descrever os documentos dos quais foram extraídos e interligadas por meio de tecnologias da Web Semântica, como os Grafos de Conhecimento. Mais que isso, as palavras-chaves extraídas podem ser associadas a termos definidos em Vocabulários Controlados, de forma a garantir o significado correto para o contexto especificado. Por meio da interligação de metadados, o Comandante poderá identificar os documentos que tratam de assuntos correlatos. Mais que isso, o resultado da extração de palavras-chaves via PLN pode servir de insumo para interligação das próprias informações. Essas interligações podem ser disponibilizadas em Grafos de Conhecimento.

Ao submeter mais questionamentos a abordagem apresentada, de forma incremental, novas palavras-chaves serão extraídas. Como exemplo, na questão “Quais são as lideranças (regionais e internacionais) existentes que influenciam o conflito?” (BRASIL, 2018, p. A-2) destaca-se a palavra “lideranças”, que poderia servir de base para extração de palavras-chaves com o uso de PLN. Sabendo-se que essas “lideranças” tem seu escopo de atuação (regional e internacional), entende-se que a atuação dessas lideranças pode se dar na mesma região das organizações detectadas a partir do questionamento anterior, sobre “organizações” e “região”. Ao unir todas as informações extraídas em um mesmo Grafo de Conhecimento, poderiam ser utilizadas ferramentas de análises de grafos para se descobrir se

há influência das lideranças sob as organizações de uma mesma região. Logo, os grafos de conhecimento fornecerão suporte aos tomadores de decisão e podem ser armazenados com segurança lógica, por meio do emprego de Algoritmos Criptográficos de Estado.

As informações extraídas de documentos, páginas da Internet, bases de dados ou outros documentos podem também ser inseridas em bases de dados estruturadas para uso por meio de Ferramentas Analíticas de Dados. Nessas bases de dados as informações poderiam ser associadas a métricas, tais como a quantidade de vezes que as organizações tiveram influência em uma região nas quais ocorreram conflitos armados, por exemplo, permitindo que o Comandante pondere se trata-se de uma organização forte e já consolidada ou se está evoluindo ao longo do tempo. Essas bases de dados também podem ser protegidas por meio de controles de acesso que implementem Algoritmos Criptográficos de Estado.

Naturalmente, para cada questão, algum nível de análise por parte da equipe de OpInfo é requerido, seja para definição das palavras de base para uso de PLN, seja pela indicação de quais documentos devem ser utilizados para obtenção de respostas. Outrossim, Ferramentas Analíticas de Dados requerem análise crítica de decisores.

Um outro questionamento interessante a ser respondido pelo Comandante que planeja o emprego de OpInfo é “Qual é a opinião da população na região do conflito? Suas etnias? Suas religiões?” (BRASIL, 2018, p. A-2). A coleta de opinião pode ser realizada de diversas maneiras, utilizando diferentes ferramentas e, até mesmo, processos dependentes de ações manuais humanas. Todavia, dada a necessidade de celeridade para quaisquer assuntos no tocante a informações, o uso de tecnologias de coleta automática trará benefícios. Com foco nas ações de obtenção de informações no contexto da dimensão informacional, podem ser coletados comentários em mídias sociais de todos os tipos, de modo a formatar a opinião pública de uma sociedade por meio do emprego de PLN, associado a algoritmos de buscas em conteúdos Web e Vocabulários Controlados. A opinião pública será composta de opiniões de grupos organizados, lideranças políticas, imprensa, em uma mescla de opiniões nacionais e internacionais. A separação e organização desses conjuntos de informações obtidos na perspectiva lógica é apoiada pelo emprego de metadados categorizados com o uso de Vocabulários Controlados e tecnologias da Web Semântica para geração de interligações entre as informações coletadas. Além disso, o emprego de Aprendizado de Máquinas pode fornecer predições em tempo real, viabilizando ao Comandante uma visão da tendência das opiniões da população.

Um cenário completo de implementação da Gestão de Dados pode e deve ser criado em fases, garantindo o alcance de níveis de maturidade no que diz respeito à gestão da

informação. Dentre as tecnologias citadas há, ainda, um considerável número de possibilidades de emprego, dada as diversas abordagens que têm sido pesquisadas. Um conjunto de pesquisas nesta área, pode ser obtido por meio de buscas em bases de artigos científicos como o IEEE *Xplorer*® (IEEE XPLORE, 2021g) e o Portal de Periódicos da CAPES (PORTAL PERIÓDICOS CAPES, 2021d).

Portanto, a Gestão de Dados e tecnologias associadas não são uma panaceia para solucionar automaticamente e isoladamente todas as demandas de OpInfo. Todavia, de fato, ela atua como atividade-meio no suporte às OpInfo desde o planejamento a até a execução. As tecnologias auxiliam, mas ainda requerem intervenções por seres humanos, mesmo que em menor grau de esforço, já que até mesmo tecnologias de Inteligência Artificial requerem treinamento conduzidos por humanos.

4.6 Síntese da análise

As áreas de conhecimento de gestão de dados e as tecnologias que as implementam podem ser utilizadas de forma combinada para prover melhores resultados de preparação e uso das informações na dimensão informacional. Embora a dimensão informacional seja segmentada em três perspectivas, duas delas, especificamente, têm uma estreita relação: a lógica e a cognitiva. Isso posto, algumas áreas de conhecimento de gestão de dados são aplicáveis as duas perspectivas supramencionadas, assim como as tecnologias que as implementam. Sendo assim, em suma, frisa-se que: (i) uma mesma tecnologia pode ser aplicada a diferentes áreas de conhecimento que atendem a uma perspectiva da dimensão informacional; (ii) uma mesma tecnologia pode ser aplicada a uma mesma área de conhecimento que atende diferentes perspectivas da dimensão informacional; e (iii) uma mesma tecnologia pode ser aplicada a diferentes áreas de conhecimento de diferentes perspectivas da dimensão informacional.

Em síntese, Algoritmos Criptográficos de Estado aplicam-se às áreas de conhecimento “Armazenamento e Operações de Dados” e “Segurança de Dados” na perspectiva lógica da dimensão informacional. Processamento de Linguagem Natural como parte da Inteligência Artificial aplica-se às áreas “Gestão de Metadados” e “Qualidade de Dados” na perspectiva lógica da dimensão informacional. Vocabulários Controlados aplicam-se à área “Gerência de Documentos e Conteúdos” na perspectiva lógica da dimensão informacional e à área “Gestão de Metadados” na perspectiva lógica da dimensão informacional, embora com efetivo uso na perspectiva cognitiva. Web Semântica e Grafos de Conhecimento aplicam-se às áreas “Integração e Interoperabilidade de Dados” e “Gestão de

Metadados” nas perspectivas lógica e cognitiva da dimensão informacional. Ferramentas Analíticas de Dados aplicam-se à área de “*Data Warehousing* (Armazém de Dados) e *Business Intelligence* (Inteligência do Negócio)” na perspectiva cognitiva da dimensão informacional, embora sejam embasadas e dependam fundamentalmente de técnicas de preparação de dados que ocorrem na perspectiva lógica, como a fusão de dados. Aprendizado de Máquinas como parte da Inteligência Artificial aplica-se à “Ciência de Dados” na perspectiva cognitiva da dimensão informacional.

5 CONCLUSÃO

As OpInfo têm o desafio de lidar com informações de forma célere e com eficácia. Sendo a Tecnologia da Informação uma área que fornece suporte a outras áreas de atuação e que lida com informações para prover resultados úteis a seu usuário, a união de OpInfo com Tecnologia da Informação forma uma parceria harmônica.

Como a área de Tecnologia da Informação é ampla, nesta monografia foram apresentadas as tecnologias atinentes à gestão de dados (ou gestão de informações), haja vista a condição basilar da informação no contexto das OpInfo, que são permeadas pela dimensão informacional. As análises apresentadas consideraram que a informação é dinâmica, de tipos variados, de volume crescente e, por isso, de difícil gestão sem o suporte de tecnologias.

Na análise efetuada nesta monografia, foi evidenciada a melhoria da usabilidade, utilidade e qualidade das informações que servem às OpInfo proporcionada pelo emprego da gestão de dados, como atividade-meio, no suporte às OpInfo desde o planejamento a até a execução. A gestão de dados e sua implementação paulatina, em fases planejadas, por meio de tecnologias existentes possibilitam que as OpInfo ganhem destaque no suporte aos decisores no ambiente operacional, já que viabilizam a análise de grandes volumes informações, mesmo que apresentada de forma variada. As análises viabilizam, inclusive, a geração de novos conhecimentos.

Cabe ressaltar que a gestão de dados e suas tecnologias associadas apresentadas não englobam as variadas combinações e possibilidades que têm sido alvo de diversas pesquisas científicas. De fato, a monografia limitou-se a apresentar as possibilidades do que têm aplicação prática no mercado e de forma concisa. Além disso, este trabalho não se aprofundou no detalhamento da implementação das tecnologias por meio de diferentes abordagens, experimentos e geração de produtos, visto que essa tarefa abrangeria muitas conceituações teóricas de base da área de TIC, o que não cabe a uma assessoria no nível de

Estado-Maior.

Como sugestões de trabalhos futuros, sugere-se a convocação de pessoal especializado para desenvolvimento de projetos de pesquisas e inovação nas Instituições Científicas e Tecnológicas (ICT) da MB, capazes de aprofundar e ponderar as propostas apresentadas. Adicionalmente, sugere-se que as escolas e centros de aperfeiçoamento militares criem parcerias com as ICT da MB, das demais Forças Singulares ou extra-militar de forma a ampliar a troca de conhecimentos, fomentando a importância do aperfeiçoamento da gestão da informação. O desenvolvimento de aplicações para treinamento e simulações de OpInfo podem ser viabilizados por intermédio de pesquisadores no nível gerencial, e mão de obra no nível de graduação, sobretudo de mão de obra proveniente de alunos de institutos militares de nível superior com trabalhos práticos de fim de curso.

REFERÊNCIAS

AVELINO, Jones de Oliveira; CORDEIRO, Kelli de Faria; CAVALCANTI, Maria Cláudia Cavalcanti. **An RDF based approach for integrating data at different levels of abstraction.** *In*: Proceedings of the Brazilian Symposium on Multimedia and the Web (WebMedia '20). Association for Computing Machinery, New York, NY, USA, p. 81–88. 2020. DOI: <https://doi.org/10.1145/3428658.3430981>.

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. **A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities.** *Scientific American*, New York, v. 284, Issue 5, May 2001.

BRASIL. Agência Brasileira de Inteligência. 2021a. Disponível em: <<https://www.gov.br/abin/pt-br/assuntos/inteligencia-e-contrainteligencia>>. Acesso em: 18 jul. 2021.

BRASIL. Comando de Operações Terrestres. **EB70-MC-10.213**: manual de operações de informação. 2. ed. Brasília, DF, 2019.

BRASIL. **Decreto nº 7.845, de 14 de novembro de 2012.** Regulamenta procedimentos para credenciamento de segurança e tratamento de informação classificada em qualquer grau de sigilo, e dispõe sobre o Núcleo de Segurança e Credenciamento. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/decreto/d7845.htm>. Acesso em: 02 ago. 2021.

BRASIL. Estado-Maior da Armada. **EMA-305**: doutrina militar naval. 1. ed. Brasília, DF, 2017.

BRASIL. Estado-Maior da Armada. **EMA-335**: doutrina de operações de informação. 1. ed. Brasília, DF, 2018.

BRASIL. Ministério da Defesa. **MD35-G-01**: glossário das Forças Armadas. 5. ed. Brasília, DF, 2015.

BRASIL. Ministério da Defesa. **MD30-M-01**: doutrina de operações conjuntas – 1º Volume. 2. ed. Brasília, DF, 2020.

BRASIL. Ministério da Defesa. **MD42-M-02**: doutrina de logística militar. 3. ed. Brasília, DF, 2016.

CASTRO, Bruno Portes Costa de. **Semantic Enrichment and Exploration**: An approach developed for dataset tags of Open Data Portals. 2019. 85fl. Dissertação (Mestrado em Informática) – Instituto de Matemática, Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2019.

DAMA International. The Global Data Management Community. 2021b. Disponível em: <<https://www.dama.org/cpages/home>>. Acesso em: 18 jul. 2021.

DE FARIAS, Claudio *et al.* **Fusão de dados para Ambientes Inteligentes**. *In:* Minicursos da VI Escola Regional de Sistemas de Informação (ERSI 2019). UNIGRANRIO, Duque de Caxias, RJ, p.133-157. SBC: 2019. DOI: <https://doi.org/10.5753/sbc.488.5.5>

EBCconhecer. Metabusca em Banco de Dados do Exército Brasileiro. 2021f. Disponível em: <http://ebusca.eb.mil.br/vufind/>. Acesso em: 02 ago. 2021.

ENAP. Escola Nacional de Administração Pública. **Governança de Dados**. Módulo 3: Gestão Inteligente de Dados. Brasília, DF, 2019.

FERREIRA, Aurélio Buraque de Holanda. **Miniaurélio**: o minidicionário da língua portuguesa. 6.ed. Curitiba: Positivo, 2004. ISBN: 857472416-5

FIGUEIREDO, Glaucia Botelho de. **LigADOS**: Interligação de Conjuntos de Dados em Plataformas de Dados Abertos na Web Semântica. 2020. 138fl. Dissertação (Mestrado em Informática) – Instituto de Matemática, Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2020.

GEOHACK. Geohack – Rio de Janeiro. 2021c, Disponível em: https://geohack.toolforge.org/geohack.php?pagename=Rio_de_Janeiro¶ms=22_54_10_S_43_12_28_W_type:city_region:BR. Acesso em: 17 jul. 2021.

GLIKSON, Ella; WOOLLEY, Anita Williams: **Human Trust in Artificial Intelligence: Review of Empirical Research**. ANNALS, 14, 627–660, 2020. <https://doi.org/10.5465/annals.2018.0057>

HENDERSON, Deborah; EARLEY, Susan. **The DAMA guide to the data management body of knowledge (DAMA-DMBOK2)**. 2nd. ed. New Jersey: Technics Publications, 2017.

HILL, John M. D. *et al.* **Anticipatory planning in information operations**. *In:* Smc 2000 conference proceedings. 2000 IEEE International Conference on Systems, Man and Cybernetics. v. 4, vp. 2350-2355, 2000. Doi: 10.1109/ICSMC.2000.884342.

IEEE Xplore. **IEEE Xplore®**. 2021g. Disponível em: <https://ieeexplore.ieee.org/Xplore/home.jsp>. Acesso em: 03 ago. 2021.

KIMBALL, Ralph; ROSS, Margy. **The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling**. 3a.ed. Indianapolis: Wiley, 2013. ISBN: 978-1-118-53080-1

MENDES, Pablo N.; MÜHLEISEN, Hannes; BIZER, Christian. **Sieve: Linked Data Quality Assessment and Fusion**. *In:* Proceedings of the 2012 Joint EDBT/ICDT Workshops (EDBT-ICDT '12). Association for Computing Machinery, New York, NY, USA, p. 116–123. 2012. DOI: <https://doi.org/10.1145/2320765.2320803>

MILLER, S. J. **Metadata for digital collections: a how-to-do-it manual**. New York, NY: Neal-Schuman Publishers, 2011.

MINISTÉRIO DO PLANEJAMENTO, DESENVOLVIMENTO E GESTÃO. Secretaria de Tecnologia da Informação e Comunicação. **Guia de Governança de TIC do SISP v 2.0**. Brasília, 2017. Disponível em:

https://www.gov.br/governodigital/pt-br/sisp/documentos/arquivos/guia_de_governanca_de_tic_do_sisp_v_2-0.pdf. Acesso em: 7 maio 2021.

OXLEY, Mark E.; THORSEN, Steven N. **Fusion or integration: What's the difference?** Air Force Institute of Technology, Department of Mathematics and Statistics, Wright-Patterson AFB, OH, 2004. Disponível em: <<https://apps.dtic.mil/sti/citations/ADA522244>>. Acesso em: 02 ago. 2020.

PAULHEIM, Heiko. **Knowledge graph refinement: A survey of approaches and evaluation methods.** *Semantic Web*, 8(3), p. 489–508. 2016. doi:10.3233/sw-160218

Portal Periódicos CAPES. 2021d. Disponível em: <[https://www-periodicos-capes.gov-br.ez1.periodicos.capes.gov.br/index.php?>](https://www-periodicos-capes.gov.br.ez1.periodicos.capes.gov.br/index.php?>). Acesso em: 02 ago. 2021.

REDE BIM. Bibliotecas Integradas da Marinha (RedeBIM). Diretoria do Patrimônio Histórico e Documentação da Marinha. 2021e. Disponível em: <<http://www.redebim.dphdm.mar.mil.br/pergamum/biblioteca/index.php>>. Acesso em: 03 ago. 2021.

RICHTER, Walter E. **O Futuro das Operações de Informações.** *Military Review* (Maio-Junho 2009), p. 71-83. Disponível em: <<https://www.armyupress.army.mil/Journals/Edicao-Brasileira/Artigos-em-Destaque/2018/O-Futuro-das-Operacoes-de-Informacoes/>>. Acesso em: 06 fev. 2021

RI-DPHDM. Repositório Intitucional da Diretoria do Patrimônio Histórico e Documentação da Marinha. 2021f. Disponível em: <<https://www.repositorio.mar.mil.br/>>. Acesso em: 20 jul. 2021.

SMITH, Johanna Wilhelmina; KOBASHI, Nair Yumiko. **Como elaborar vocabulário controlado para aplicação em arquivos: como fazer VOL.10.** São Paulo: Arquivo do Estado/Imprensa Oficial do Estado, 2003. Disponível em: https://www.arqsp.org.br/arquivos/oficinas_colecao_como_fazer/cf10.pdf

SOUZA, Tatiana Almeida Rosali Fernandes. **O vocabulário controlado como instrumento e representação da informação na FINEP.** XII ENANCIB Encontro Nacional de Pesquisa em Ciência da Informação. Brasília, DF, 2011. Disponível em: <https://ridi.ibict.br/bitstream/123456789/88/1/RosaliTatianaEnancib2011b.pdf>

SOUZA, Marcelo Pereira de et al. **A Linguistic-Based Method that Combines Polarity, Emotion and Grammatical Characteristics to Detect Fake News in Portuguese.** *In: Proceedings of the Brazilian Symposium on Multimedia and the Web (WebMedia '20).* Association for Computing Machinery, New York, NY, USA, p. 217–224. 2020. DOI: <https://doi.org/10.1145/3428658.3430975>

TYGEL, A. **A Semantic Tags for Open Data Portals: Metadata Enhancements for Searchable Open Data.** Tese (Doutorado em Informática) – Instituto de Matemática, Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

UNITED STATES. Joint Chiefs of Staff. **JP 3-13: joint doctrine for information operations publication.** 2nd. ed. [Washington, D.C.]: Joint Electronic Library, 2014. Disponível em:

https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_13.pdf. Acesso em: 7 maio 2021.

VERONA, Leticia Dias; LOPES, Giseli Rabello; CAMPOS, Maria Luiza Machado. **Using Government Data to Uncover Political Power and Influence of Contemporary Slavery Agents in Brazil**. *In*: Workshop on Big Social Data and Urban Computing. Springer, Cham, 2018. DOI: https://doi.org/10.1007/978-3-030-11238-7_8

VISACRO, ALESSANDRO. **A Guerra na Era da Informação**. Rio de Janeiro: editora Contexto, 2018.

APÊNDICES

APÊNDICE A – Gestão de Dados e OpInfo

TABELA 1

Relações entre gestão de dados e as OpInfo

Gestão de Dados		OpInfo	
		Dimensão Informacional	
Áreas de Conhecimento		Perspectiva Lógica	Perspectiva Cognitiva
Governança de Dados	Arquitetura de Dados		
	Modelagem e Projeto de Dados		
	Armazenamento e Operações de Dados	X	
	Segurança de Dados	X	
	Integração e Interoperabilidade de Dados	X	X
	Gerência de Documentos e Conteúdos	X	
	Dados Mestre e de Referência		
	<i>Data Warehousing</i> (Armazém de Dados) e <i>Business Intelligence</i> (Inteligência do Negócio)		X
	Gestão de Metadados	X	X
	Qualidade de dados	X	
	Fusão de Dados	X	X
	Ciência de Dados		X

APÊNDICE B – Gestão de Dados e OpInfo

TABELA 2

Tecnologias associadas às áreas de conhecimento da gestão de dados que apoiam as OpInfo

Gestão de Dados		Tecnologias Associadas						
Áreas de Conhecimento		A	B	C	D	E	F	G
Governança de Dados	Armazenamento e Operações de Dados	X						
	Segurança de Dados	X						
	Integração e Interoperabilidade de Dados				X	X		
	Gerência de Documentos e Conteúdos			X				
	<i>Data Warehousing</i> (Armazém de Dados) e <i>Business Intelligence</i> (Inteligência do Negócio)						X	
	Metadados		X	X	X	X		
	Qualidade de Dados		X					
	Fusão de Dados			X				
	Ciência de Dados							X

Legenda:

	A – Algoritmos Criptográficos de Estado
	B – Inteligência Artificial – Processamento de Linguagem Natural
	C – Vocabulários Controlados
	D – Web Semântica
	E – Grafos de Conhecimento (<i>Knowledge Graphs</i>)
	F – Ferramentas Analíticas de Dados
	G – Inteligência Artificial – Aprendizado de Máquinas

ANEXOS

ANEXO A – Estudo das OpInfo no Ambiente Operacional (BRASIL, 2018)

1) Missão

- Como as Operações de Informação (OpInfo) podem contribuir para a missão da Força?
- No enunciado da missão, fica claro se as OpInfo exercerão o esforço principal?
- Se as OpInfo estiverem apoiando, como elas podem contribuir para o sucesso da Operação?

2) A Situação e sua Compreensão

- Quais são os Aspectos Gerais Políticos e Militares relevantes para as OpInfo?
- Há Forças Inimigas conduzindo OpInfo no Teatro/Área de Operações (TO/AOp)? Quem está coordenando as nossas ações com as demais Forças?
- Quais são as capacidades de OpInfo das próprias Forças e das Forças Amigas?
- Quais são as Hipóteses Básicas formuladas pelo Superior?
- A dimensão informacional é favorável às ações de nossas Forças?
- O adversário já está conduzindo suas OpInfo no TO/AOp? Tenho capacidade de realizar meu planejamento, ao mesmo tempo em que impeço as ações do adversário?

3) Capacidades de OpInfo que estão disponíveis para o Comandante

- Quais são os especialistas disponíveis no Estado-Maior?
- Há alguma capacidade para a qual teremos que buscar especialistas?
- Quais são as Capacidades Relacionadas à Informação (CRI) que estão subordinadas ao Comandante do Teatro de Operações (ComTO)?
- Dentre os especialistas, há alguém cursado em OpInfo?

4) Organização do Estado-Maior

- As OpInfo estão no esforço principal ou apoiando as demais ações?
- Há pessoal suficiente no Estado-Maior para dedicarem-se às OpInfo?
- Quem coordenará as OpInfo com as demais ações: o D-3 ou o Chefe do Estado-Maior?

Eles ficarão sobrecarregados?

5) Orientações do nível político e estratégico

- Quais são as orientações e efeitos desejados mencionados no Plano Estratégico de Emprego Conjunto das Forças Armadas (PEECFA)?
- Quais são as orientações das publicações de nível político e estratégicos (Política Nacional de Defesa, Livro Branco de Defesa Nacional etc.)?
- Qual é o Estado Final Desejado?
- Quais são os temas, mensagens e objetivos disponíveis?

6) Dimensão Informacional

6.1) Atores

- Quais são as lideranças (regionais e internacionais) existentes que influenciam o conflito?
- Quais são os objetivos e interesses dessas lideranças? Quais são suas vulnerabilidades?
- Pelo Diagrama de Relacionamento, quais são os relacionamentos que queremos neutralizar ou fortalecer?
- Quais são as organizações que estão atuando na região? Como são financiadas? Quais seus interesses e objetivos? Há nacionais que atuam nessas organizações?
- Qual é a opinião da população na região do conflito? Suas etnias? Suas religiões?

6.2) Sistemas de Informação

6.2.1) Adversários

- Quais são os equipamentos de Comando e Controle (C2) do adversário? Quais são as suas capacidades e vulnerabilidades?
- Qual é a localização das infraestruturas de C2 do adversário?
- Há empresas civis contratadas participando do tráfego de informações de C2?
- Quem são as pessoas que participam do processo de tomada de decisão do adversário? Essas pessoas podem ser cooptadas ou neutralizadas?
- Qual é a doutrina de C2 do adversário?

6.2.2) Nossas Forças

- Quais são as nossas capacidades e vulnerabilidades de nossos equipamentos de C2? Qual é a sua origem? O país fabricante é nosso aliado?

- Qual é a localização de nossas infraestruturas de C2? Qual é a proteção existente e está disponível?
- Há empresas civis contratadas participando do tráfego de informações de C2?
- Quem são as pessoas que participam do nosso processo de tomada de decisão? Essas pessoas podem ser cooptadas ou neutralizadas pelo adversário?
- O adversário tem conhecimento da nossa doutrina de C2?

6.3) Mídia

- Qual é a abrangência dos veículos de mídia (ou jornalistas) que têm influência na região? Quais são os seus interesses? Como são influenciados? Quem os financia?
- Qual é a credibilidade das mídias?
- Há mídias estatais?
- Quais são as mensagens e temas que estão sendo divulgados pelas mídias?
- Onde estão localizadas as infraestruturas de mídia na região?
- Pelo gráfico de influência de mídia, quais são os veículos de mídia que podemos utilizar?

7) Capacidades de OpInfo do adversário

- Quais são as CRI que estão disponíveis para o adversário?
- Qual é a experiência do adversário em OpInfo?
- Qual a filosofia e técnicas de influência do adversário?
- Quem são os aliados (atores, mídia, grupos, lideranças etc.) do adversário?
- Quais podem ser os objetivos e efeitos desejados do adversário?
- O adversário já está conduzindo suas OpInfo no TO/AOp? Ele está obtendo sucesso na dimensão informacional?

8) Possíveis efeitos a serem atingidos/contribuídos com as OpInfo

- Listar os efeitos que podemos obter com as OpInfo.
- Iniciar o preenchimento da tabela Efeito x Meio.