



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

**Uma Arquitetura da Informação para a Organização
de Informações de Eventos de Campo em Operações
Militares do Exército Brasileiro com o Emprego de
Ontologias**

Javan de Oliveira Cruz

Dissertação apresentada como requisito parcial para conclusão do
Mestrado Profissional em Computação Aplicada

Orientador

Prof. Dr. Marcio de Carvalho Victorino

Brasília
2022

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

dD278a de Oliveira Cruz, Javan
Uma Arquitetura da Informação para a Organização de
Informações de Eventos de Campo em Operações Militares do
Exército Brasileiro com o Emprego de Ontologias / Javan de
Oliveira Cruz; orientador Marcio de Carvalho Victorino. --
Brasília, 2022.
85 p.

Dissertação (Mestrado - Mestrado Profissional em
Computação Aplicada) -- Universidade de Brasília, 2022.

1. Arquitetura da Informação. 2. Ontologias. 3. Exército
Brasileiro. 4. Operações Militares. I. de Carvalho
Victorino, Marcio, orient. II. Título.



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

Uma Arquitetura da Informação para a Organização de Informações de Eventos de Campo em Operações Militares do Exército Brasileiro com o Emprego de Ontologias

Javan de Oliveira Cruz

Dissertação apresentada como requisito parcial para conclusão do
Mestrado Profissional em Computação Aplicada

Prof. Dr. Marcio de Carvalho Victorino (Orientador)
Faculdade de Ciência da Informação

Prof. Dr. Edison Ishikawa
Departamento de Ciência da Computação

Dr. Wallace Anacleto Pinheiro
Exército Brasileiro

Prof. Dr. Marcelo Ladeira
Coordenador do Programa de Pós-graduação em Computação Aplicada

Brasília, 03 de agosto de 2022

Dedicatória

Dedico este trabalho aos meus pais, Estanisia e Javan (*in memoriam*), exemplos de persistência e dedicação, que me passaram valores importantes que direcionam os meus passos.

A minha esposa Daniele, que ilumina a minha vida, compartilhando comigo momentos especiais. Aos meus filhos Milena e Gabriel, para os quais tenho a missão de conduzir e ensinar, porém são peças fundamentais para meu aprendizado diário.

Agradecimentos

Agradeço primordialmente à Deus, por me conceder muito mais do que mereço, me fortalecendo diariamente para chegar ao fim desse grande desafio.

À Daniele, minha amada esposa, companheira sempre presente durante toda o trajetória percorrida, por compreender as diversas horas de lazer despendidas para que esse trabalho pudesse ser concluído.

Ao meu amigo Luciano Sales, por toda amizade e pelo incentivo constante, sendo uma pessoa fundamental para minha participação nessa jornada.

Ao meu orientador, prof. Márcio Victorino, por todo o conhecimento compartilhado, pela solicitude e disponibilidade, bem como pelo direcionamento que me conduziu para conclusão desse curso.

Ao prof Ishikawa e Wallace, por terem aceitado participar da banca examinadora deste trabalho, bem como aos professores do Programa de Pós-Graduação em Computação Aplicada, pela transmissão de conhecimentos durante as disciplinas cursadas.

Ao Exército Brasileiro, por me conceder a oportunidade de realizar esse aperfeiçoamento profissional, disponibilizando horas de estudo em proveito dessa pesquisa.

Resumo

O Centro de Coordenação de Operações Móvel (CCOp Mv) é um projeto estratégico do Exército Brasileiro que está em desenvolvimento, tendo por objetivo apoiar um Grande Comando Operacional em situações de guerra e não guerra. O CCOp Mv será composto por tecnologias, sistemas e recursos diversos, de forma a possibilitar o alcance de seu objetivo. Dentre suas principais funcionalidades podem ser destacadas a obtenção e o fornecimento de informações de eventos de campo, que além de direcionarem a construção da consciência situacional, representam parâmetros essenciais para o processo de tomada de decisão das autoridades, em operações militares. Apesar dos avanços tecnológicos experimentados nos dias de hoje, a natureza dos conflitos atuais permite a percepção da relevância em obter dados por meio de observadores humanos. Porém, a simples obtenção de dados não é julgada suficiente para a formação de conhecimentos específicos sobre um determinado domínio, sendo necessária a construção de uma base de informações estruturada, que permita a recuperação de conhecimentos de forma otimizada. Nesse contexto, a proposta desta pesquisa é fornecer uma Arquitetura da Informação voltada à obtenção de dados por meio de observadores de campo, além de disponibilizar conhecimentos gerados, tendo como parâmetros a organização e a estruturação de relatórios informacionais, empregando ontologias para a sua representação semântica. Além disso, está sendo proposta a construção de um protótipo de ontologia, com o intuito de fornecer suporte à Arquitetura da Informação desenvolvida, uma vez que o domínio de eventos militares possui uma considerável especificidade. Como contribuições podem ser destacados o fornecimento de uma Arquitetura da Informação, que visa auxiliar na otimização da obtenção de informações em operações militares, bem como o protótipo de ontologia que foi desenvolvido, que poderá servir como base para o desenvolvimento outras aplicações dentro do Exército Brasileiro.

Palavras-chave: ontologia, Arquitetura da Informação, consciência situacional, tomada de decisão.

Abstract

The Centro de Coordenação de Operações Móvel (CCOp Mv) is a strategic project of the Brazilian Army that is under development, with the aim to support a Large Operational Command in war and non-war situations. CCOp Mv will be composed of different technologies, systems and resources, in order to achieve its objective. Among its main functionalities can be highlighted the acquisition and provision of information of field events, which in addition to directing the construction of situational awareness, represent the essential parameters for the decision-making process of authorities, in military operations. Despite the technological advances experienced today, the nature of current conflicts allows for the perception of the relevance of obtaining data through human observers. However, simply obtaining data is not considered sufficient for the formation of specific knowledge about a given domain, requiring the construction of a structured information base, which allows the recovery of knowledge in an optimized way. In this context, the proposal of this research is to provide an Information Architecture aimed at obtaining data through field observers, in addition to providing generated knowledge, having as parameters the organization and structuring of informational reports, employing ontologies for its semantic representation. Furthermore, the construction of an ontology prototype is being proposed, in order to support the developed information architecture, since the domain of military events has a considerable specificity. As contributions can be highlighted the provision of an Information Architecture, which aims to assist in the optimization of obtaining information in military operations, as well as the prototype of ontology that was developed, which could be used as a basis for the development of other applications within the Brazilian Army.

Keywords: ontology, Information Architecture, situational awareness, decision-making.

Sumário

1	Introdução	1
1.1	Definição do problema	2
1.2	Justificativa	3
1.3	Pergunta de pesquisa	4
1.4	Hipótese de pesquisa	4
1.5	Objetivo geral	4
1.6	Objetivos específicos	4
1.7	Estrutura do trabalho	5
2	Procedimento Metodológico	6
2.1	Teoria do Enfoque Meta-Analítico Consolidado (TEMAC)	7
2.1.1	TEMAC aplicado à temática Ontologias	8
2.1.2	TEMAC aplicado à temática Arquitetura da Informação	14
2.1.3	Considerações	19
3	Fundamentação Teórica	20
3.1	Representação do Conhecimento	20
3.1.1	Metadados	21
3.1.2	Taxonomia	22
3.1.3	Tesauros	23
3.1.4	Ontologias	24
3.1.5	<i>Web Ontology Language</i>	28
3.1.6	<i>Resource Description Framework</i>	30
3.2	Administração de Dados	33
3.3	Gestão da Informação	35
3.4	Usabilidade	37
3.5	Arquitetura da Informação	38
4	Trabalhos Relacionados	42
4.1	Ferramenta de relatório baseada em ontologia	43

4.2	Desenvolvimento de ontologia	45
4.3	Arquitetura para dados abertos conectados	47
5	Estudo de Caso	50
5.1	O Centro de Coordenação de Operações Móvel	50
5.2	Desenvolvimento de protótipo de ontologia	53
5.2.1	Determinação do domínio e do escopo da ontologia	54
5.2.2	Reutilização de ontologias existentes	54
5.2.3	Levantamento dos principais termos da ontologia	57
5.2.4	Definição de classes e a hierarquia de classes	59
5.2.5	Definição das propriedades das classes	59
5.2.6	Definição do detalhamento das propriedades das classes	60
5.2.7	Criação de instâncias	60
5.2.8	Validação do protótipo de ontologia	60
5.3	Proposta de Arquitetura da Informação	61
5.3.1	Recursos empregados	61
5.3.2	Concepção geral	62
5.3.3	Proposta de Arquitetura de <i>Software</i>	63
6	Resultados	70
6.1	Protótipo de Ontologia	70
6.2	Arquitetura da Informação	74
6.3	Análise da recuperação da informação	80
6.3.1	Recuperação da informação no sistema Pacificador	80
6.3.2	Recuperação da informação na Arquitetura da Informação proposta	81
6.3.3	Considerações	81
7	Conclusão	83
7.1	Contribuições	84
7.2	Trabalhos futuros	85
	Referências	86

Lista de Figuras

1.1	Portfolio Estratégico do Exército.	2
2.1	Etapas da TEMAC.	7
2.2	Número de citações ano a ano - Ontologias.	9
2.3	Nuvem de palavras-chave - Ontologias.	10
2.4	Co-citações - Ontologias.	12
2.5	<i>Coupling</i> - Ontologias.	13
2.6	Número de citações ano a ano - Arquitetura da Informação.	15
2.7	Nuvem de palavras-chave - Arquitetura da Informação.	16
2.8	Co-citações - Arquitetura da Informação.	17
2.9	<i>Coupling</i> - Arquitetura da Informação.	18
3.1	Representação Formal da OWL - Adaptado pelo autor.	30
3.2	Caracterização do RDF.	31
3.3	Representação de Tripla RDF.	32
3.4	Exemplo de Tripla RDF.	32
3.5	Elementos de dados - Modelo ER.	34
3.6	Administração de Dados - Adaptado pelo autor.	35
3.7	Processo de Gestão da Informação - Adaptado pelo autor.	36
3.8	Projeto de Construção VS Arquitetura da Informação - Adaptado pelo autor.	39
3.9	Modelo de Arquitetura da Informação - Adaptado pelo autor.	40
4.1	Descrição do processo de correspondência - Adaptado pelo autor.	44
4.2	Níveis de hierarquia de uma ontologia - Adaptado pelo autor.	46
4.3	Modelo da arquitetura de publicação de dados abertos conectados.	48
5.1	Descrição de incidente extraída do Sistema Pacificador.	52
5.2	Extrato de classes da ontologia PROTON - <i>Software Protégé</i>	55
5.3	Extrato de classes da ontologia COSMO - <i>Software Protégé</i>	55
5.4	Extrato de classes da ontologia JC3IEDM.	56
5.5	Grafo de Conhecimento simplificado.	58

5.6	Extrato de termos - MD33-M-02.	59
5.7	Mecanismo de Inferência - <i>Protégé</i>	61
5.8	Estrutura simplificada da Arquitetura da Informação.	62
5.9	Processo de Execução da Arquitetura de Software.	64
6.1	Grafo de Conhecimento - Registros do Sistema Pacificador.	71
6.2	Anotações - Protótipo OntoEB.	73
6.3	Extrato de Classes - Protótipo OntoEB.	73
6.4	Extrato de Classes - Protótipo OntoEB.	74
6.5	Eventos Registrados - Arquitetura de Software.	75
6.6	Consulta de classes da ontologia - Arquitetura de Software.	75
6.7	Refinamento da consulta de classes - Arquitetura de Software.	76
6.8	Propriedades da classe selecionada - Arquitetura de Software.	76
6.9	Consulta alternativa de classes da ontologia - Arquitetura de Software.	77
6.10	Tratamento para classe não encontrada - Arquitetura de Software.	77
6.11	Termo pendente utilizado pelo usuário - Arquitetura de Software.	78
6.12	Utilização de termo não encontrado - Arquitetura de Software.	78
6.13	Responsividade - Arquitetura de Software.	79
6.14	Relatório de Evento - Arquitetura de Software.	79

Lista de Tabelas

2.1	Frequência de palavras-chave - Ontologias.	10
2.2	Ranking de países que mais publicaram - Ontologias.	11
2.3	Autores que mais publicaram - Ontologias.	11
2.4	Análise de Co-citações - Núcleos de Pesquisa em Ontologias.	13
2.5	Análise de <i>Coupling</i> - Frentes de Pesquisa em Ontologias.	14
2.6	Frequência de palavras-chave - Arquitetura da Informação.	16
2.7	Ranking de países que mais publicaram - Arquitetura da Informação.	16
2.8	Autores que mais publicaram - Arquitetura da Informação.	17
2.9	Análise de Co-citações - Núcleos de Pesquisa em Arquitetura da Informação.	18
2.10	Análise de <i>Coupling</i> - Frentes de Pesquisa em Arquitetura da Informação.	19
4.1	Principais trabalhos relacionados com a pesquisa.	42

Lista de Abreviaturas e Siglas

AI Arquitetura da Informação.

AIfIA *Asilomar Institute for Information Architecture.*

C2 Comando e Controle.

CCOp Mv Centro de Coordenação de Operações Móvel.

CSV *Comma-separated values.*

DCMI *Dublin Core Metadata Initiative.*

DW *Data Warehouse.*

ER Entidades e Relacionamentos.

GLO Garantia da Lei e da Ordem.

IA Inteligência Artificial.

JC3IEDM *Joint Consultation, Command and Control Information Exchange Data Model.*

JSON *JavaScript Object Notation.*

MIP *Multilateral Interoperability Programme.*

NER *Named Entity Recognition.*

OEE Objetivos Estratégicos do Exército.

OTAN Organização do Tratado do Atlântico Norte.

OWL *Web Ontology Language.*

PEEx Plano Estratégico do Exército.

Ptf EE Portfólio Estratégico do Exército.

RDF *Resource Description Framework.*

RDF-S *Resource Description Framework Schema.*

SGBD Sistema de Gerenciamento de Banco de Dados.

SIEX Sistema de Extensão e Pós-graduação.

SIGRA Sistema de Graduação.

SIPES Sistema de Pessoal.

SIPPOS Sistema de Pós-Graduação.

SPARQL *SPARQL Protocol and RDF Query Language.*

SUS *System Usability Scale.*

TEMAC Teoria do Enfoque Meta-Analítico Consolidado.

TIC Tecnologia de Informação e Comunicações.

UnB Universidade de Brasília.

UnBGOLD *UnB Government Linked Open Data.*

URI *Uniform Resource Identifier.*

UX *User Experience.*

W3C *World Wide Web Consortium.*

XML *eXtensible Markup Language.*

Capítulo 1

Introdução

O Exército Brasileiro tem como missão contribuir para a garantia da soberania nacional, dos poderes constitucionais, da lei e da ordem, salvaguardando os interesses nacionais e cooperando com o desenvolvimento nacional e o bem-estar social [1]. Para tal, a cada quadriênio é disponibilizado o Plano Estratégico do Exército (PEEx), que tem por finalidade direcionar os investimentos da Força Terrestre para o alcance dos diversos objetivos estabelecidos.

Nesse contexto, foi consolidada a implantação do Portfólio Estratégico do Exército, que possui diversos Programas Estratégicos, destacados na Figura 1.1, que contribuem para o alcance de um ou mais Objetivos Estratégicos do Exército. Assim, é possibilitada a geração de capacidades necessárias para que o Exército Brasileiro cumpra sua missão institucional, de acordo com o previsto na Constituição Federal /88 e nas demais diretrizes constantes da normativa infraconstitucional, em particular na Estratégia Nacional de Defesa [2].

O Centro de Coordenação de Operações Móvel (CCOp Mv) é um Projeto em andamento, pertencente ao Programa Estratégico do Exército PROTEGER, que integra o Subportfólio Estratégico do Exército DEFESA DA SOCIEDADE.

Em sua aplicação prática, o CCOp Mv é concebido como um Centro de Comando e Controle configurado para apoiar, com recursos disponíveis, um Grande Comando Operacional, em situações de guerra e não guerra. Para o alcance de seus objetivos, será composto por diversas tecnologias, sistemas e recursos de Comando e Controle, possibilitando apoio em sistemas, equipamentos, softwares e acesso a sistemas táticos, estratégicos e críticos.

Um dos principais pilares do CCOp Mv é o fornecimento de informações relevantes que possibilitem a construção da Consciência Situacional, apoiando e direcionando a tomada de decisão.



Figura 1.1: Portfólio Estratégico do Exército (Fonte: [2]).

1.1 Definição do problema

Para um melhor entendimento do emprego do Centro de Coordenação de Operações Móvel e de seus sistemas integrados, é relevante citar o Pacificador. Trata-se de um sistema de apoio à decisão no âmbito da Força Terrestre Brasileira, concebido para situações de não guerra [3]. Possui emprego direto em apoio à operações de Garantia da Lei e da Ordem (GLO) e de defesa/segurança de Grandes Eventos, visando a formação da consciência situacional, o Comando e Controle (C2), a sincronização das ações entre os elementos envolvidos, bem como o tratamento de incidentes ocorridos.

O sistema Pacificador é utilizado em diversos eventos de grande relevância para o país, como ocorreu na Rio +20 (2012), Copa das Confederações (2013), Copa do Mundo (2014), Olimpíadas (2016), Copa América (2021), entre outros. O sistema dispõe de ambientes de Cenários e de Relatos, com o objetivo de permitir que o responsável por determinada operação tenha a percepção de todos os eventos que estão ocorrendo.

Mesmo com os avanços tecnológicos atuais, a natureza dos conflitos existentes nos dias de hoje aumenta a importância da obtenção de dados por observadores humanos. Assim, uma das ferramentas existentes no software é a de relato de incidentes por agentes de campo. Um Incidente é qualquer evento não planejado que possa comprometer a segurança do evento, requerendo tratamento adequado.

A simples obtenção de dados não é suficiente para a geração de conhecimento pleno sobre determinado domínio. A formação de uma base de informações semanticamente representada, com um formato adequado, torna-se de grande valor para a tomada de

decisão. Porém, dentro do Pacificador, a descrição de cada um dos incidentes relatados é entregue como texto livre, em linguagem natural, o que dificulta a extração dos conhecimentos necessários.

O conceito de ontologias vem sendo utilizado em diversos trabalhos relacionados à organização do conhecimento a partir de dados obtidos. Nos últimos anos, essa conceitualização tem recebido especial atenção, particularmente em abordagens de Arquitetura da Informação e para a representação semântica de dados.

A aplicabilidade da ontologia em uma Arquitetura da Informação proposta foi um dos principais focos das investigações deste trabalho, sendo um dos pilares para a solução do estudo de caso apresentado na pesquisa.

1.2 Justificativa

O gerenciamento de dados e informações é considerado um desafio, envolvendo o entendimento de conceitos pertencentes a diversos domínios de interesse. Disponibilizar uma estrutura que permita a organização e integração desses ativos é julgada uma condição necessária para o sucesso das ações de governança, nos mais diversos ambientes.

No que diz respeito às operações do Exército, há a necessidade de obtenção de informações confiáveis, que permitam a tomada de decisão mais ajustada por parte dos gestores. A partir da perspectiva científica, existe aplicabilidade da temática ontologia para a organização dos dados, aplicável a uma Arquitetura da Informação.

A proposta deste trabalho é fornecer uma estrutura voltada à obtenção de dados assistida e a disponibilização de informações, com fundamentos baseados na organização e estruturação de relatórios informacionais, empregando recursos de ontologias para a sua representação semântica. Nesse sentido, foi construída uma Arquitetura da Informação, que tem por objetivo auxiliar a obtenção de dados em operações militares, disponibilizando informações descritas semanticamente, permitindo melhor integração e análise. Adicionalmente, foi proposta a construção de um protótipo de ontologia, com o intuito de fornecer suporte à arquitetura da informação a ser construída, devido a necessidade de obtenção de termos e conceitos específicos para o fim a que se destina. Cabe destacar que a Arquitetura da Informação foi desenvolvida por meio do uso de tecnologias existentes para integração de dados e de ferramentas desenvolvidas especificamente para este fim.

Nesse contexto, esta pesquisa encontra justificativa na necessidade de obtenção de informações pertinentes e semanticamente representadas, permitindo a governança adequada de operações de não guerra no Exército Brasileiro. A ausência de entendimento dos dados coletados causa grande prejuízo para a tomada de decisão, tanto para aspectos operacionais quanto para administrativos. Para o desenvolvimento do trabalho

foram empregadas abordagens multidisciplinares capazes de proporcionar a organização, contextualização, armazenamento e recuperação efetiva da informação. Tais abordagens estão fundamentadas nos recursos convencionais utilizados para sistemas de informação somados aos recursos de organização da informação e do conhecimento.

1.3 Pergunta de pesquisa

Como organizar dados obtidos em campo para permitir a disponibilização de informações semanticamente descritas, a serem empregadas em apoio à governança e aos processos de tomada de decisão?

1.4 Hipótese de pesquisa

Acredita-se que com a Arquitetura da Informação proposta será facilitada a disponibilização de informações descritas semanticamente para a tomada de decisão, bem como será proporcionada uma melhor governança das atividades operacionais nas quais o CCOp Mv será empregado.

1.5 Objetivo geral

A partir do que foi exposto, este trabalho tem por objetivo disponibilizar uma Arquitetura da Informação para obtenção de dados de campo e fornecimento de informações, empregando recursos de ontologias para representação do conhecimento. Com isso, está sendo buscado o aprimoramento da governança e da tomada de decisão em situações de emprego do CCOp Mv, por meio de um ambiente propício para a análise das informações.

1.6 Objetivos específicos

- Realizar uma revisão sistemática da literatura para identificar publicações relacionadas às temáticas em estudo;
- Compreender o vocabulário utilizado no Exército Brasileiro, com destaque para operações de não guerra
- Identificar ontologias existentes, a serem empregadas em prol da Arquitetura proposta;
- Identificar e analisar diferentes arquiteturas da informação já utilizadas em outros trabalhos;

- Fornecer suporte ao relato de incidentes com eficiência e oportunidade;
- Desenvolver um protótipo de ontologia para emprego em prol da Arquitetura da Informação a ser disponibilizada;
- Desenvolver uma proposta de Arquitetura da Informação capaz de assistir a obtenção de dados e proporcionar a disponibilização de informações descritas semanticamente.

1.7 Estrutura do trabalho

O trabalho desenvolvido está estruturado em 7 capítulos, além deste, de acordo com o que segue:

- O **Capítulo 2** apresenta o procedimento metodológico empregado na pesquisa, que foi composto pela revisão sistemática da literatura baseado em uma metodologia específica. Foram contemplados os principais tópicos para formação da base teórica necessária à identificação de estudos que formaram a base do trabalho desenvolvido.
- O **Capítulo 3** aborda a fundamentação teórica do trabalho, abordando os principais conceitos e tecnologias empregados na pesquisa desenvolvida. Foram buscados conhecimentos relacionados à representação do conhecimento, abordando o entendimento sobre metadados, taxonomia, tesauros, ontologias, bem como sobre a estruturação das ontologias por meio do *Resource Description Framework* (RDF). Além disso, foi levantada uma visão geral sobre a administração de informações, sendo abordados conceitos relacionados à administração de dados, gestão da informação, usabilidade e sobre a Arquitetura da Informação, base fundamental deste trabalho.
- O **Capítulo 4** destaca trabalhos relacionados com esta pesquisa, quais os seus benefícios, suas limitações e seus objetivos.
- O **Capítulo 5** aborda um estudo de caso, que destaca as necessidades do Exército quanto às informações operacionais, apresenta a proposta de construção de um protótipo de ontologia e a Arquitetura da Informação proposta, além de abordar resultados obtidos com a pesquisa.
- O **Capítulo 6** apresenta os resultados que foram obtidos, por meio do desenvolvimento da pesquisa realizada.
- O **Capítulo 7**, por fim, apresenta a conclusão da pesquisa, bem como destaca contribuições a serem fornecidas e aponta possíveis propostas para trabalhos futuros.

Capítulo 2

Procedimento Metodológico

Uma pesquisa científica pode ser classificada de acordo com critérios estabelecidos, tomando por base a natureza, procedimentos, técnicas ou abordagens [4]. Este capítulo tem por objetivo apresentar a metodologia empregada no presente estudo, para o alcance do objetivo proposto.

Inicialmente, foi definido que a natureza é aplicada, caracterizada pela obtenção de conhecimentos a partir de aplicações práticas, relacionadas à solução de problemas que ocorrem no mundo real. A motivação para este tipo de pesquisa surge da necessidade de solução de problemas concretos [5]. A natureza aplicada é justificada pela proposição de uma Arquitetura da Informação que permita gerenciar informações obtidas em operações de campo, de forma a produzir e disponibilizar relatórios estruturados.

De acordo com Gil [6], pesquisas exploratórias são desenvolvidas com o objetivo de proporcionar uma visão mais aproximada, bem como o entendimento aprofundado sobre determinado objeto. Os objetivos desse tipo de pesquisa são desenvolver hipóteses, aumentar a familiaridade do pesquisador com determinado ambiente, além de modificar e clarificar conceitos levantados [7]. Sendo assim, quanto aos objetivos, a presente pesquisa é classificada como exploratória, a partir da necessidade de definir abordagens mais adequadas à Arquitetura da Informação proposta.

Com relação aos procedimentos adotados, este trabalho é classificado como um estudo de caso. Consiste em explorar situações da vida real as quais possuem limites não definidos claramente, bem como descrever situações relacionadas ao contexto no qual determinada investigação é desenvolvida [6]. Trata-se de um método de pesquisa sobre um assunto específico, que permite o aprofundamento de conhecimentos, de forma a fornecer subsídios que possibilitem a proposição de soluções para uma situação problema encontrada. Assim, será realizado um estudo de caso para dar suporte ao desenvolvimento do trabalho proposto.

Ainda em relação aos procedimentos, esta pesquisa é considerada bibliográfica. Esse

tipo de procedimento implica em estudar artigos, teses, livros, entre outras publicações, de forma a identificar contribuições para o estado da arte de temática a ser abordada. Para fundamentar o embasamento teórico, foi utilizada a Teoria do Enfoque Meta-Analítico Consolidado (TEMAC), proposta por Mariano e Santos [8], a ser descrita a seguir.

2.1 Teoria do Enfoque Meta-Analítico Consolidado (TEMAC)

A pesquisa científica busca produzir conhecimentos em diversas áreas, a partir de definições de determinado problema. Cabe ao pesquisador buscar o entendimento do que deve ser estudado, de forma a obter uma visão holística sobre o atual estado da arte das temáticas buscadas [9]. Devem ser verificados os trabalhos disponíveis na literatura relacionados ao objeto de estudo, de forma que seja possível produzir uma pesquisa que agregue conhecimentos adicionais à temática [10].

Com o objetivo de disponibilizar um formato robusto e eficiente, que permita realizar o levantamento sistemático da fundamentação teórica da pesquisa científica, foi desenvolvida por Mariano e Santos [8] uma metodologia intitulada Teoria do Enfoque Meta-Analítico Consolidado (TEMAC). De forma a destacar a importância do emprego de um método específico para o levantamento do que foi discutido anteriormente sobre o tema a ser estudado, é recomendável realizar a combinação de informações de fontes de dados adequadas, tornando possível a obtenção de materiais com qualidade e consistência para realização da pesquisa [9].

A TEMAC foi a metodologia utilizada para a obtenção das principais publicações que fundamentam o presente trabalho, fornecendo autenticidade para o arcabouço teórico construído. A metodologia consiste em três passos, estruturados de acordo com a Figura 2.1.

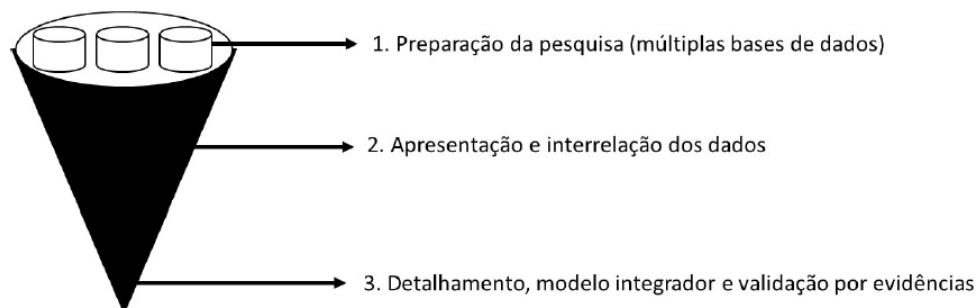


Figura 2.1: Etapas da TEMAC (Fonte: [8]).

Os passos da TEMAC são descritos conforme o que segue [8]:

1. Preparação da pesquisa - seleção de bases de dados, de palavras-chave, do espaço temporal e das áreas do conhecimento a serem pesquisados;
2. Apresentação e interrelação dos dados - análise de relevância de revistas, conferências e universidades, busca de maiores números de citações de trabalhos e de autores, áreas de conhecimento e países que mais publicam, frequência de palavras-chave, entre outros critérios de seleção do pesquisador. Cabe ressaltar que cada fator de análise apresentado tem por base princípios ou leis bibliométricas;
3. Detalhamento, modelo integrador e validação das evidências - nessa etapa, ocorre a consolidação integrada da pesquisa, validando o trabalho realizado amparado nas etapas realizadas anteriormente. Consiste da seleção dos principais autores, abordagens e linhas de pesquisa relacionados à temática escolhida.

Devido a alta complexidade e amplitude das principais áreas de pesquisa deste trabalho, foi realizada a aplicação da metodologia TEMAC sob duas perspectivas: Ontologias e Arquitetura da Informação. Foi verificado que o estudo unificado das duas temáticas, aplicado às pesquisas bibliográficas, gerou resultados pouco expressivos. A busca de trabalhos das duas áreas, de forma independente, proporcionou identificar publicações relevantes, o que serviu como alicerce para construção do arcabouço teórico que fomentou o desenvolvimento do trabalho. Sendo assim, a segunda etapa deste capítulo consiste em apresentar os resultados obtidos por meio da aplicação do TEMAC às temáticas propostas.

2.1.1 TEMAC aplicado à temática Ontologias

a. Preparação da pesquisa

Para o desenvolvimento da primeira etapa da metodologia, cumprindo o que nele é proposto, foram respondidas questões sobre termos de pesquisa, espaço temporal a ser empregado, seleção da base de dados a ser utilizada, bem como a delimitação das áreas de conhecimento relacionadas à pesquisa.

Foram utilizados, como *strings* de busca, os termos "*Ontology*" AND "*RDF*", uma vez que se buscou trabalhos que abordem Ontologias estruturadas por meio do *Resource Description Framework*. Foi tomada como base a busca por tópicos, que contemplam o título, resumo e palavras-chave dos trabalhos. De forma paralela, o termo "*Command and Control*" foi empregado em conjunto com outros termos em buscas realizadas, porém não foram encontrados resultados que agregassem ao que foi buscado inicialmente.

O espaço temporal foi de 10 anos completos, no período de 2011 e 2021. Foi empregada a *Web of Science* [11], considerada uma base de dados de reconhecida reputação, com

qualidade e disponibilidade de publicações. Foi também feita uma análise inicial dos resultados utilizando a base de dados *Scopus*, porém os resultados foram similares ao que foi obtido por meio da *Web of Science* [11], não sendo julgado oportuna a realização de uma nova aplicação da TEMAC com uma segunda base de dados. Além disso, foram delimitadas áreas de conhecimento conexas à pesquisa, de forma a otimizar a qualidade dos resultados obtidos.

Com a aplicação dos diversos parâmetros, como resultado da pesquisa realizada foram retornadas 448 publicações, empregadas para as análises da segunda e terceira etapa da TEMAC. Cabe salientar que a quantidade de artigos encontrados é suficiente para a aplicação da metodologia proposta, uma vez que na terceira etapa da metodologia, descrita nesta subseção, são realizadas análises que proporcionam a descoberta de artigos relacionados aos encontrados na busca inicial, enriquecendo em muito os resultados.

b. Apresentação e interrelação dos dados

Após o estabelecimento dos critérios de preparação da pesquisa, foi desenvolvida a apresentação e interrelação dos dados. O objetivo foi analisar a evolução no tema ano a ano, a frequência das principais palavras-chave, bem como identificar países e autores que mais publicaram na área.

- **Evolução do tema ano a ano:**

De forma a atestar a representatividade do tema, é possível perceber a evolução anual, dentro do período temporal estabelecido para a pesquisa (2011-2021), de citações dos artigos obtidos na busca realizada. A soma do número de citações até o ano de 2021 alcançou 2953, resultando em uma média de aproximadamente 295 citações anuais. A Figura 2.2 ilustra graficamente essa evolução.

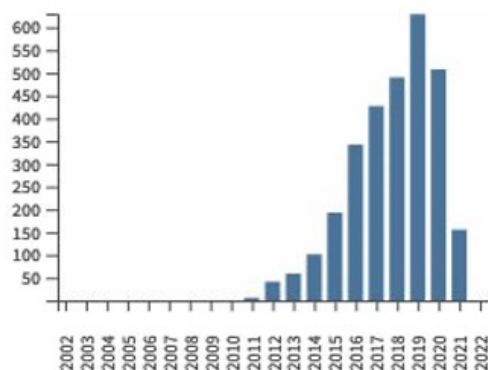


Figura 2.2: Número de citações ano a ano - Ontologias.

Pode ser percebido um aumento crescente de citações nos últimos anos, comprovando a expressividade da temática buscada. No ano de 2020, houve uma queda nesse número, pois possivelmente os dados ainda não haviam sido indexados à base de dados quando a pesquisa foi realizada.

- **Frequência de palavras-chave:**

Após verificar a evolução do tema ano a ano, foi realizada a análise das principais palavras-chave dos artigos resultantes da busca sobre Ontologias e RDF. Com isso, foi elaborada uma nuvem de palavras-chave, expressa na Figura 2.3, bem como foi levantada a frequência dessas palavras nos 448 artigos encontrados, conforme Tabela 2.1.



Figura 2.3: Nuvem de palavras-chave - Ontologias.

Tabela 2.1: Frequência de palavras-chave - Ontologias.

Palavra-chave	Frequência
Ontology	139
RDF	123
Semantic Web	84
Linked Data	64
Sparql	53
OWL	44

A análise das palavras-chave permite um primeiro entendimento sobre o tema pesquisado, antes da leitura dos artigos a serem selecionados. As palavras-chave mais frequentes estão relacionadas ao tema central da pesquisa, facilitando o direcionamento dos estudos, bem como constatando a eficácia da metodologia aplicada.

- **Países que mais publicaram:**

Foram verificados os países que mais pesquisaram a respeito do tema Ontologias e RDF. A Tabela 2.2 expressa o levantamento realizado.

Tabela 2.2: Ranking de países que mais publicaram - Ontologias.

Ranking	País	Publicações	% de 448
1	Estados Unidos	46	10.27
2	Alemanha	44	9.82
3	França	43	9.60
4	Itália	41	9.15
5	China	39	8.70
16	Brasil	11	2.46

Os Estados Unidos lideram as publicações dentro da temática em estudo, com mais de 10% do total de publicações no período. O Brasil, que ficou na 16^a posição, totaliza 11 publicações, equivalente a 2.46% do total de artigos encontrados.

- **Análise de autores:**

Finalizando a segunda etapa da metodologia, foi realizada a análise dos autores que mais publicaram sobre o objeto de estudo, representado na Tabela 2.3, verificando a relevância de suas publicações.

Tabela 2.3: Autores que mais publicaram - Ontologias.

Autores	Publicações	% de 448
Lehmann, Jens	8	1.79
Auer, Soren	8	1.79
Gangemi, Aldo	7	1.56
Nuzzolese, A. Giovanni	6	1.34
Hyvonen, Eero	5	1.12
Lohmann, Steffen	5	1.12
Peroni, Silvio	5	1.12

Cabe ressaltar que Lehmann e Aurer publicaram diversos trabalhos em conjunto, dentre os quais se destaca o trabalho intitulado *DBpedia – A large-scale, multilingual knowledge base extracted from Wikipedia* [12], que proporciona uma visão geral sobre a *DBpedia*, um projeto que tem por objetivo extrair conhecimento estruturado da *Wikipedia* e o torna disponível por meio de tecnologias de Web Semântica. O trabalho é julgado de grande relevância, totalizando 683 citações.

c. Detalhamento, modelo integrador e validação das evidências

Após a construção das primeiras impressões sobre a temática estudada, foram realizadas análises que permitem compreender as principais abordagens e linhas de pesquisa, bem como identificar autores relevantes que devem constar do trabalho.

Para o desenvolvimento da terceira etapa, foram desenvolvidas duas análises bibliométricas: Co-citação e *Coupling* (acoplamento bibliográfico). A análise de Co-citação

tem por objetivo verificar artigos que regularmente são citados em conjunto, criando um núcleo de abordagens que apoia o entendimento do que vem sendo estudado nos últimos anos. Nessa análise, os artigos mais citados pelas publicações encontradas na busca inicial são destacados, permitindo obter conteúdos de acentuada relevância, enriquecendo a pesquisa. No acoplamento bibliográfico, são determinados quais artigos encontrados na busca possuem literaturas base semelhante, tendo como premissa que artigos que citam trabalhos em comum, possuem similaridade. Assim, são identificadas frentes de pesquisa que estão sendo realizadas na atualidade [8]. Como base para a análise de co-citação, foi construída, com o apoio do *software VoS Viewer* (v1.6.15)¹, uma rede de co-citações, ilustrada na Figura 2.4.

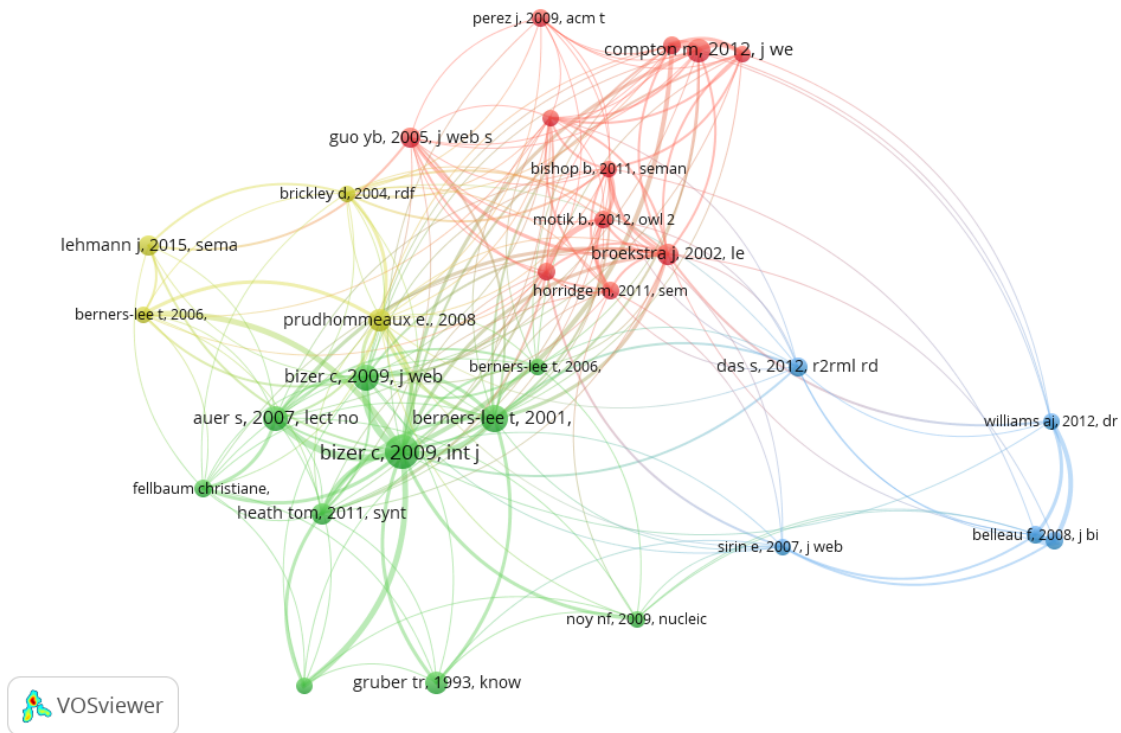


Figura 2.4: Co-citações - Ontologias.

Analisando a Figura 2.4, é possível verificar a existência de *clusters*, delimitados por cores específicas, que relaciona estudos realizados, transmitindo um bom grau de associação entre eles. A partir da verificação dos trabalhos de cada um deles, foi possível identificar núcleos de pesquisa, conforme a Tabela 2.4.

Dentre os trabalhos mais relevantes para a pesquisa realizada, além do realizado por Lehmann [12], explicitado na segunda etapa desse estudo, é possível indentificar as relaci-

¹<https://www.vosviewer.com>

Tabela 2.4: Análise de Co-citações - Núcleos de Pesquisa em Ontologias.

Cor	Núcleo de Pesquisa
Vermelho	Estruturação da informação com ontologias
Amarelo	Obtenção de conhecimento a partir de ontologias
Azul	Aplicação de ontologias na medicina
Verde	Conceitos de Ontologias e de Indexação

onadas aos conceitos de ontologias e de indexação. É possível destacar publicações como a de Bizer e Berners-Lee [13], de Gruber [14], bem como o trabalho de Berners-Lee [15]. Além dos trabalhos ressaltados, os autores identificados possuem diversas outras publicações, citadas por vários pesquisadores.

Para o desenvolvimento da análise de acoplamento bibliográfico, foi construído um mapa de calor, apresentado na Figura 2.5. Assim, foi possível relacionar os artigos mais recentes que retornaram na busca, dos últimos três anos, de acordo com suas citações em comum.

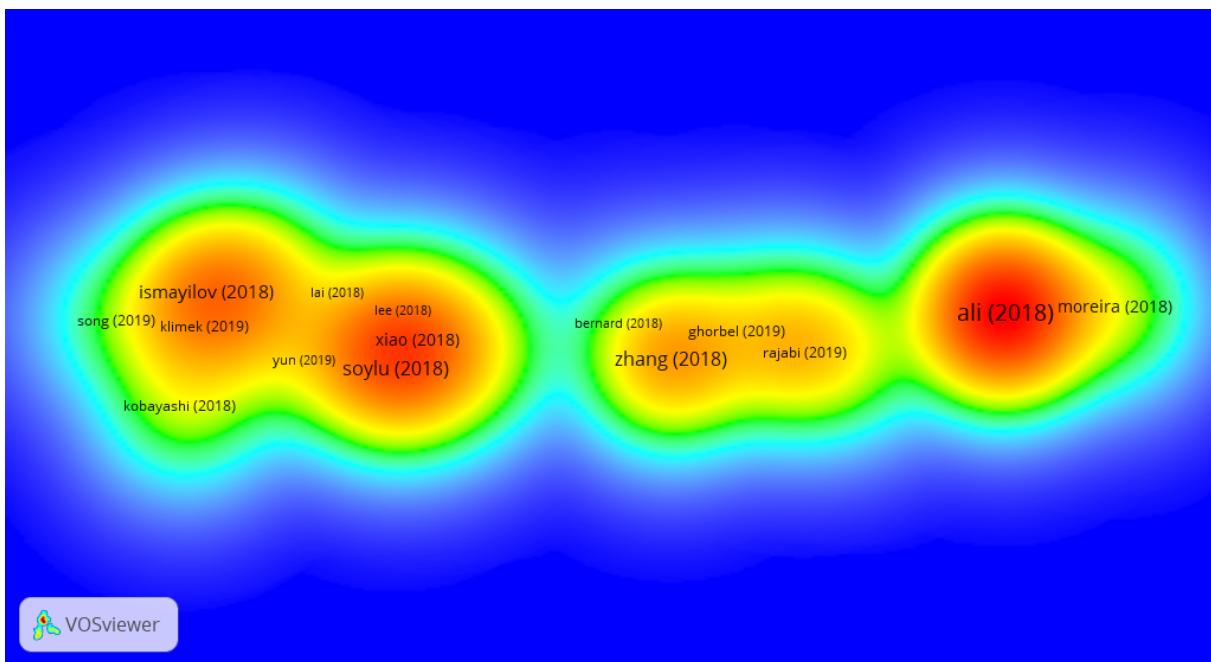


Figura 2.5: *Coupling* - Ontologias.

A partir da análise da Figura 2.5, foi possível verificar a existência de quatro *clusters*, de acordo com as zonas de calor. A partir deles, foram identificadas frentes de pesquisa, de acordo com a Tabela 2.5.

É possível destacar o trabalho realizado por Ismayilov *et al.* [16], realizado em conjunto com Lehmann, com o título *Wikidata through the eyes of DBpedia*, que apresenta uma continuidade ao trabalho realizado em 2015 [12]. O trabalho apresenta a *Wikidata*, uma

Tabela 2.5: Análise de *Coupling* - Frentes de Pesquisa em Ontologias.

Principais Autores	Núcleo de Pesquisa
Ismayilov[2018] e Klímek[2019]	Conhecimentos gerados a partir de ontologias
Soylu[2018] e Xiao[2018]	Consultas avançadas com <i>SPARQL</i>
Zhang[2018] e Ghorbel[2019]	Representação e recuperação de informações
Ali[2018] e Moreira[2018]	Aplicação de ontologias na medicina

base de conhecimento aberta e editável colaborativamente que fornece informações para a *Wikipedia/DBPedia*. Com isso, fica evidente a importância envidada pelos pesquisadores à representação do conhecimento por meio de ontologias, apresentando uma base de dados compreensiva e enriquecida pela própria comunidade.

2.1.2 TEMAC aplicado à temática Arquitetura da Informação

a. Preparação da Pesquisa

Para a temática voltada à Arquitetura da Informação, no desenvolvimento da primeira etapa da TEMAC, assim como na temática anterior, foram buscados os termos de pesquisa, espaço temporal utilizado, foi selecionada a base de dados considerada mais adequada, bem como foram delimitadas as áreas de conhecimento associadas ao trabalho.

Foi utilizado, como *string* de busca, o termo "*Information Architecture*", a partir de tópicos, que permeiam o título, resumo e palavras-chave dos trabalhos. O espaço temporal foi definido como 10 anos completos, sendo considerado o período de 2011 e 2021. Foi empregada a base de dados *Web of Science* [11]. Como na temática anterior, os resultados obtidos utilizando a base de dados Scopus foram similares aos obtidos por meio da *Web of Science* [11], não sendo adequada a realização de uma segunda aplicação da metodologia. Por fim, foram delimitadas áreas de conhecimento conexas à pesquisa, enriquecendo a qualidade dos resultados obtidos.

Como resultado da pesquisa realizada, foram retornados 244 trabalhos, utilizados para a realização das análises desenvolvidas nas etapas subsequentes da TEMAC. A quantidade de artigos encontrados, tomando por base a metodologia empregada, foi julgada suficiente, pois na análise de co-citações, realizada na terceira etapa da metodologia, foram contemplados artigos relacionados aos encontrados na busca inicial, ampliando a quantidade e qualidade dos resultados obtidos.

b. Apresentação e interrelação dos dados

Para o desenvolvimento da segunda etapa da TEMAC, seguindo o estabelecimento dos critérios de preparação da pesquisa, foi realizada a análise da evolução do tema ano a

ano, a frequência das principais palavras-chave, além de terem sido identificados países e autores que mais publicaram dentro da temática de pesquisa.

- **Evolução do tema ano a ano:**

A partir do período temporal delimitado (2011-2021), foi possível verificar a evolução do tema anualmente. A base para essa análise foi a quantidade de citações dos artigos obtidos por meio da busca realizada. O total de citações até o ano de 2021 foi de 901, sendo uma média de aproximadamente 90 citações anuais. A Figura 2.6 representa a evolução do tema ano a ano.

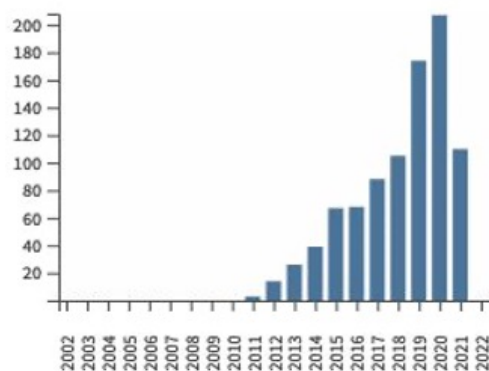


Figura 2.6: Número de citações ano a ano - Arquitetura da Informação.

A partir da análise da Figura 2.6, é possível verificar um aumento crescente da quantidade de citações nos últimos anos, atestando a representatividade da temática em estudo.

- **Frequência de palavras-chave:**

Prosseguindo na segunda etapa da TEMAC, foi desenvolvida a análise das principais palavras-chave dos artigos resultantes da busca relacionada à Arquitetura da Informação. Foi construída uma nuvem de palavras-chave, ilustrada na Figura 2.7, sendo a frequência dessas palavras, nos 244 trabalhos que retornaram a busca, expressada na Tabela 2.6.

A partir da verificação das palavras-chave mais empregadas pelos autores dos artigos e de suas respectivas frequências, é possível obter uma primeira aproximação sobre o temática em estudo. As palavras-chave mais frequentes representam o tema central do estudo, direcionando a pesquisa desenvolvida. Além disso, fica respaldado o emprego da metodologia TEMAC para o estudo em desenvolvimento.

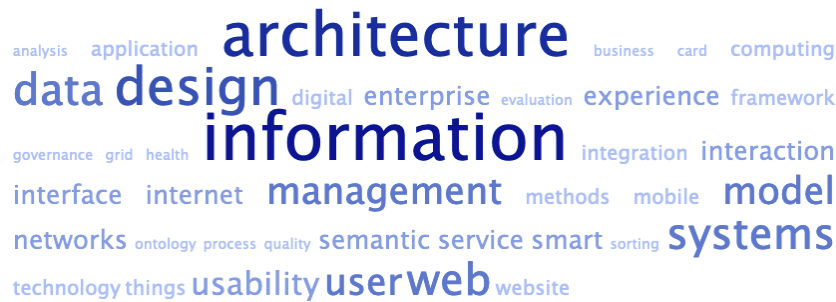


Figura 2.7: Nuvem de palavras-chave - Arquitetura da Informação.

Tabela 2.6: Frequência de palavras-chave - Arquitetura da Informação.

Palavra-chave	Frequência
Information Architecture	146
Design	89
Systems	55
Data Management	47
Web	44
User	42

- **Países que mais publicaram:**

Prosseguindo na segunda etapa da metodologia, foram identificados os países que mais pesquisaram a respeito do temática relacionada à Arquitetura da Informação. A Tabela 2.7 apresenta o ranking de países, por publicações.

Tabela 2.7: Ranking de países que mais publicaram - Arquitetura da Informação.

Ranking	País	Publicações	% de 448
1	Estados Unidos	54	22.13
2	China	33	13.53
3	Inglaterra	22	9.07
4	Espanha	15	6.15
5	Holanda	12	4.99
6	Itália	10	4.10
7	Portugal	10	4.10
8	Brasil	9	3.69

Como na temática anterior, os Estados Unidos possuem o maior número de publicações relacionadas à Arquitetura da Informação, com 22.13% do total de publicações no período de 2011 a 2021. O Brasil ficou com a 8ª colocação, com 11 publicações, representando 3.69% do total de artigos encontrados.

- **Análise de autores:**

Por fim, foram analisados os autores que mais publicaram sobre a temática em

estudo, além de ser verificada a relevância de suas publicações. Os dados estão representados na Tabela 2.8,

Tabela 2.8: Autores que mais publicaram - Arquitetura da Informação.

Autores	Publicações	% de 448
Dubey, Abhishek	8	3.28
Karsai, Gabor	8	3.28
Ahmad, Mahmood	5	2.05
Eisele, Scott	5	2.05
Odeh, Mohammed	5	2.05
Maria Brunetti, Josep	4	1.64

Dentre as publicações dos autores que mais publicaram, a com maior número de citações é a intitulada *RIAPS: Resilient Information Architecture Platform for Decentralized Smart Systems* [17], totalizando 27 referências.

c. Detalhamento, modelo integrador e validação das evidências

Dando prosseguimento à TEMAC aplicado à Arquitetura da Informação, após as duas etapas da metodologia que permite obter impressões iniciais sobre o tema da pesquisa, foram desenvolvidas as análises que compõe a terceira etapa, que permite a compreensão das principais abordagens e linhas de pesquisa relacionadas, além de identificar autores relevantes.

Para a elucidação do que é buscado na etapa em pauta, foram desenvolvidas as análises de Co-citação e de *Coupling*, assim como foi realizado na temática voltada às Ontologias. Destaca-se que na análise de co-citações são apresentados os artigos mais citados pelas publicações encontradas na busca inicial, permitindo a obtenção de conteúdos que enriquecem a pesquisa desenvolvida. Foi produzida uma rede de co-citações, expressa na Figura 2.8, com o apoio do *software VoS Viewer* (v1.6.15)².

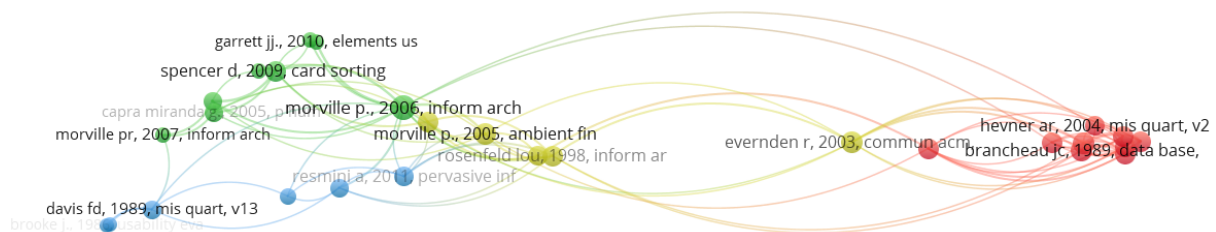


Figura 2.8: Co-citações - Arquitetura da Informação.

²<https://www.vosviewer.com>

Como na análise de co-citação, são verificados artigos que são citados em conjunto, a partir da análise da Figura 2.8, foi verificada a existência de quatro *clusters*, ilustrados por cores, que consolidam um bom grau de associação entre os trabalhos. A partir da verificação das publicações presentes em cada um deles, foi possível identificar áreas de pesquisa específicas, de acordo com a Tabela 2.9:

Tabela 2.9: Análise de Co-citações - Núcleos de Pesquisa em Arquitetura da Informação.

Cor	Núcleo de Pesquisa
Vermelho	Desenvolvimento e implementação de Arquitetura da Informação
Amarelo	Conceituação geral sobre Arquitetura da Informação
Azul	Usabilidade
Verde	Arquitetura para Desenvolvimento de Web Sites

Após verificar os trabalhos resultantes da análise de co-citação, foi possível identificar os que melhor se relacionam com a temática de estudo. Além das publicações que propiciaram a formação da rede de citações da Figura 2.8, foram encontradas outras publicações de Rosenfeld e Morville, relacionadas à Arquitetura da Informação. Os autores são julgados relevantes na área e, conseqüentemente, para a pesquisa. Se destacam nesta análise os trabalhos de Rosenfeld e Morville [18], Morville [19] e Morville e Rosenfeld [20], além do trabalho de Resmini e Rosati [21].

Para a análise de *coupling*, que permite determinar quais artigos possuem referências bibliográficas em comum, foi possível identificar frentes de pesquisa que estão sendo realizadas atualmente. Foi criado um mapa de calor, expresso na Figura 2.9, que torna possível relacionar artigos dos últimos três anos, resultantes da busca, que possuem citações semelhantes.

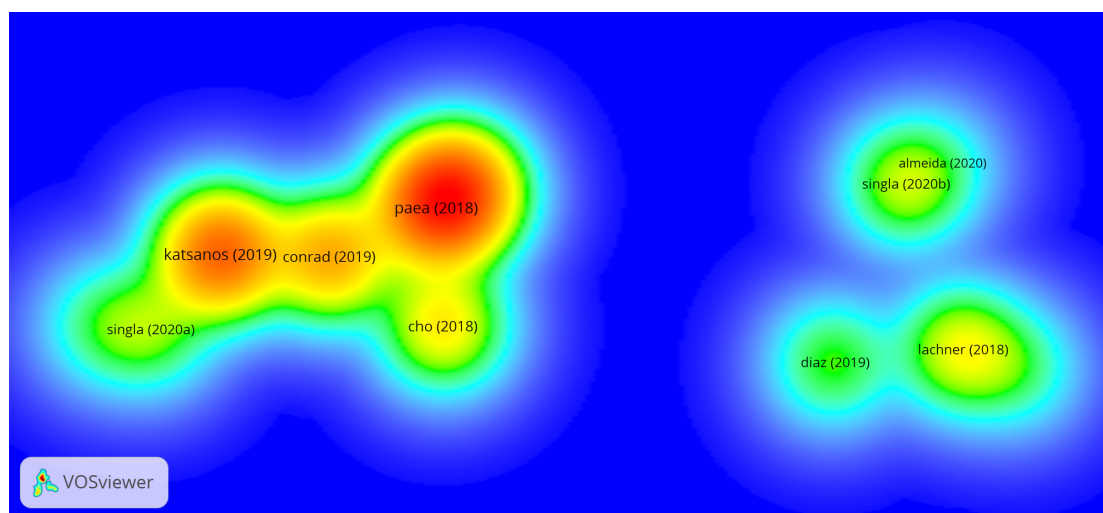


Figura 2.9: *Coupling* - Arquitetura da Informação.

Analisando a Figura 2.9, foi verificada a existência de três frentes de pesquisa recentes, tomando por base as zonas de calor, de acordo com a Tabela 2.10.

Tabela 2.10: Análise de *Coupling* - Frentes de Pesquisa em Arquitetura da Informação.

Principais Autores	Núcleo de Pesquisa
Almeida[2020] e Singla[2020]	Aplicabilidade de arquitetura da informação
Lachner[2018] e Yakunin[2018]	Usabilidade
Conrad[2019] e Katsanos[2019]	<i>Card Sorting</i>

A partir da verificação dos trabalhos, foi possível perceber o crescimento corrente da frente de pesquisa *Card Sorting* (Ordenação de cartões), uma das ferramentas utilizadas para estruturar a arquitetura da informação, a partir da experiência de usuários. Além disso, cabe destacar o trabalho de Almeida, *et al.*, intitulado *Toward a document-centered ontological theory for information architecture in corporations* [22], que versa sobre conhecimentos teóricos relacionados à Arquitetura da Informação, descrevendo recursos necessários para sua criação por meio de um modelo ontológico, apresentado conhecimentos úteis para a pesquisa realizada.

2.1.3 Considerações

Na pesquisa realizada, o emprego da Teoria do Enfoque Meta Analítico, uma metodologia de revisão bibliográfica sistemático, amparada em leis bibliométricas, contemplou a identificação de tendências dentro das temáticas de estudo, analisando a representatividade dos autores e das publicações encontradas. É importante ressaltar a terceira etapa da metodologia aplicada, que proporciona que os resultados encontrados não se restrinjam a busca inicial realizada na base de dados. Os artigos citados pelos trabalhos resultantes das buscas são agrupados, formando núcleos de abordagens. Com isso, são ampliadas a quantidade e, principalmente, a qualidade das publicações encontradas. A aplicação da metodologia proporcionou confiabilidade e credibilidade ao levantamento da literatura que foi empregada na fundamentação teórica descrita no trabalho, assegurando um estado da arte bem delimitado dos temas.

A TEMAC permitiu a obtenção das principais publicações utilizadas como arcabouço teórico deste trabalho, sendo utilizados artigos e autores identificados como os mais citados, durante a análise. Além disso, foram buscados outros trabalhos conexos considerados pertinentes, agregando positivamente aos resultados da pesquisa. A partir do estudo realizado, foi possível elencar os artigos e metodologias com considerável relevância nas áreas de Ontologias e de Arquitetura da Informação, construindo um embasamento teórico que respalda o desenvolvimento da pesquisa.

Capítulo 3

Fundamentação Teórica

Neste capítulo, é apresentada a Fundamentação Teórica dos principais conceitos e tecnologias consideradas relevantes para desenvolvimento desta pesquisa. A Seção 3.1 apresenta abordagens relacionadas à Representação do Conhecimento, sendo destacados conceitos sobre Metadados, Taxonomia, Tesouros, Ontologias, bem como sobre o sobre o *Resource Description Framework*. Na Seção 3.2 são apresentados conceitos relacionados à Administração de Dados, fazendo também uma necessária abordagem geral sobre a organização de dados, informações e conhecimentos. A Seção 3.3 são exploradas conceituações voltadas à Gestão da Informação, destacando sua importância em ambientes organizacionais. Na Seção 3.4 foram apresentadas definições relacionadas à Usabilidade e a sua importância para sistemas informacionais. Por fim, na Seção 3.5 foi realizada uma abordagem sobre Arquitetura da Informação, formando um alicerce de conceitos necessários ao entendimento de uma das temáticas principais do trabalho.

3.1 Representação do Conhecimento

A obtenção de informações confiáveis, oportunas e relevantes é um crescente desafio enfrentado por tomadores de decisão. Com o advento de tecnologias da informação, diversas possibilidades são disponibilizadas, de forma a facilitar a construção do conhecimento a partir das informações obtidas.

O conhecimento é uma fonte essencial do saber, que permite uma compreensão abrangente de um determinado domínio de interesse. Trata-se de um bem que deve ser armazenado de forma a possibilitar a sua recuperação e disponibilização organizada e objetiva. Para que o conhecimento seja armazenado, recuperado e entregue ao usuário, é necessário que esteja descrito segundo padrões que permitam a sua compreensão.

A representação do conhecimento pode ser definida como um conjunto de convenções sintáticas e semânticas que torna possível descrever coisas. Consiste na utilização

de linguagens específicas, frases ou números, que correspondem à descrição ou condição do mundo [23]. Em ambientes computacionais, é relevante que o conhecimento seja representado tomando por base as regras de uma linguagem específica, possibilitando o seu emprego para a tomada de decisão de forma adequada. Nesse contexto, a representação do conhecimento pode ser entendida como uma forma eficiente de se organizar informações, apoiando a solução de problemas com oportunidade.

Em um sistema da informação, a partir de um determinado domínio de interesse, vários são os instrumentos utilizados para representar o conhecimento ou apoiar a sua representação, sendo destacados e descritos nesta seção os que seguem: Metadados, Taxonomia, Tesouros e Ontologias.

3.1.1 Metadados

Metadados são geralmente definidos como ‘dados relativos aos dados’, o que é considerada uma definição muito ampla. Na Ciência da Informação, os metadados são geralmente considerados informações sobre um conjunto de dados em uma determinada representação disponibilizada. Na gestão do conhecimento e na Arquitetura da Informação, metadados significam informações sobre objetos, ou seja, informações sobre um documento, uma imagem, um módulo de conteúdo reutilizável, e assim por diante [24].

Os metadados são normalmente utilizados para armazenar informações úteis ao processo de recuperação da informação, sendo capazes de descrever o conteúdo de determinada informação [25]. O termo teve origem 1995, em um simpósio realizado em Dublin, Ohio, que deu origem à *Dublin Core Metadata Initiative* (DCMI) [26].

Em geral, metadados são melhor entendidos como qualquer declaração sobre um recurso de informação, independentemente de sua finalidade de uso, do vocabulário empregado ou de sua representação [24]. Eles descrevem a estrutura e a composição de um dado principal, apresentando suas especificações e a utilidade das informações disponíveis.

Um formato de categorização proposto por Taylor [27] subdivide os metadados administrativos, estruturais e descritivos. Os metadados administrativos tem por objetivo o gerenciamento, o suporte à tomada de decisão e a manutenção do registro das informações, disponibilizado, por exemplo, informações sobre requisitos de armazenamento e o processo de migração de informações digitais. Os estruturais estão relacionadas à estrutura do suporte físico da informação que está sendo descrita, como um arquivo digital, livro, fotografia, dentre outros. Já os metadados descritivos, são aqueles que discriminam características intelectuais do conteúdo de determinado documento [25].

É importante destacar que as organizações têm encontrado dificuldades para consolidar o gerenciamento de metadados em sistemas de informações, bem como em solucionar

problemas relativos à manutenção destes durante o tempo de vida de determinado objeto, uma vez que essas atividades são normalmente desenvolvidas de forma manual [28].

Algumas características de um documento, como título, autor e palavras-chave, podem ser extraídos automaticamente, enquanto outras, como as características relativas à semântica, normalmente necessitam de interação humana [29]. A tendência é realizar automaticamente esse processo de extração, com o uso de tecnologias como XML e RDF.

O emprego de metadados em um ambiente organizacional permite uma melhor compreensão de como os dados estão sendo gerados. Com uma boa gestão de metadados, torna-se possível relacionar contextos e realizar cruzamentos de dados, simplificando descrições e criando vocabulários específicos para determinado domínio. Além disso, garante maior eficiência para recuperar e analisar os dados disponibilizados.

3.1.2 Taxonomia

As taxonomias têm seu início com Carl Von Linné, que desenvolveu um sistema de classificação hierárquica para formas de vida no século XVIII, sendo a base para a moderna classificação zoológica e botânica e sistema de nomenclatura para espécies [24]. Ele é conhecido como ‘Pai da Taxonomia’, sendo seu sistema de nomeação, ordenação e classificação de organismos empregado até os dias atuais.

A Taxonomia pode ser definida como uma classificação sistemática, sendo conceituada no âmbito da Ciência da Informação como ferramenta voltada à organização intelectual [30]. Cabe salientar que, além de ser um componente de Ontologias, as taxonomias possuem amplo emprego em portais institucionais e em bibliotecas digitais, como um novo mecanismo de consulta, em conjunto com ferramentas de busca.

Com o emprego de taxonomia em um sistema da informação, é possível prover o agrupamento dos termos empregados, formando uma categorização que facilita encontrar o informações que melhor descrevem um objeto [24]. Em síntese, a taxonomia fornece mais informações sobre os conceitos, de forma a auxiliar o usuário a encontrar informações de interesse. A ferramenta permite o armazenamento e recuperação de informações, dentro de um sistema, de forma lógica, por meio de navegação estruturada.

As taxonomias vêm sendo usadas para a criação de metadados ou termos comuns para descrever um objeto, com foco na recuperação da informação e na categorização, como suporte de navegação e esquemas que organizam conteúdos das páginas na web [31].

É possível ainda definir a taxonomia como um sistema de classificação, que toma por base, normalmente, uma hierarquia de termos e conceitos, na qual os termos localizados nos níveis mais baixos representam os aspectos mais específicos do conteúdo [25]. Sendo assim, uma taxonomia apresenta seus termos organizados hierarquicamente, sendo

as classes nos níveis mais superiores e suas subclasses nos níveis mais inferiores, distribuídas em quantos níveis de especificidade sejam necessários.

As taxonomias desempenham um papel importante nas organizações, permitindo melhor compreensão e localização das informações disponíveis, além de perceber relacionamentos e correlações inerentes a essa informação [25]. As taxonomias geram redes semânticas comuns baseadas nas necessidades de negócios organizacionais, conectando os recursos humanos às informações.

Por fim, a partir do que foi apresentado, a taxonomia pode ser entendida como uma estrutura conceitual hierárquica que inclui todos os principais conceitos utilizados nos documentos na área de negócio considerada. O emprego isolado de taxonomias, em ambientes organizacionais, não é suficiente para a organização da informação, sendo adequado o emprego de tesouros para tal. Os tesouros proporcionam, além das hierarquias propostas pela taxonomia, relações não hierárquicas entre os termos e os conceitos existentes, sendo essa a temática a ser tratada na próxima subseção.

3.1.3 Tesouros

Como as taxonomias, os tesouros tem sido utilizado para descrever todos os tipos de estruturas de classificação de sujeitos. O tesouro estende a taxonomia, proporcionando maior capacidade de descrever o mundo, não só permitindo a organização em uma hierarquia específica, como também permitindo que outras declarações possam ser realizadas sobre as informações [24]. Em suma, o tesouro fornece um vocabulário rico para descrever os termos, sendo uma ferramenta importante para o processo de classificação de informações.

De forma a alcançar o entendimento do significado do termo tesouro, algumas definições podem ser destacadas:

- uma linguagem documentária dinâmica que contém termos relacionados semântica e logicamente, cobrindo de modo compreensivo um domínio do conhecimento [32];
- uma lista estruturada de termos associados empregada por analistas de informação e indexadores, para descrever um documento com a desejada especificidade, em nível de entrada, e para permitir aos pesquisadores a recuperação da informação que procuram [33];
- uma linguagem documentária, que apresenta algumas peculiaridades, pois sua hierarquia de assuntos possui uma relação associativa e sua estrutura não se baseia em conceito de palavras simplesmente, mas sim de termos conceituados e relacionados [34].

O tesouro possui cobertura abrangente para termos de um domínio específico do conhecimento, não havendo, portanto, um tesouro de domínio geral. Adicionalmente, deve ser dinâmico, permitindo alterações no significado e inserção de novos termos. Os tesouros são estruturas consolidadas e padronizadas por normas internacionais, como a ISO 2788, a ISO 5964 e a ANSI/NISO Z39.19-2003 que tratam da sua construção, formatação e manutenção [31].

Uma das características de maior importância do tesouro é sua capacidade de fornecer assistência para o processo de indexação e recuperação de documentos. A indexação pode ser definida como um processo intelectual que pressupõe que o acesso à informação documentária, por meio de termos ou códigos, deve ser o ponto de partida para selecionar os próprios documentos [35].

Em síntese, o tesouro é considerado um instrumento de controle terminológico eficaz para a organização do conhecimento e importante ferramenta no tratamento e recuperação da informação [34]. A identificação de conceitos relacionada à sistematização e representação de uma determinada área do conhecimento é de suma importância para os Sistemas de Informação. O Tesouro é apresentado como um instrumento dinâmico, possuidor de terminologia estruturada, capaz de fornecer assistência ao usuário para facilitar a busca por termos que representem determinado significado.

3.1.4 Ontologias

Uma das formas de estruturar informações, fazendo a sua representação de forma eficiente, é por meio do emprego de ontologias. Ontologia é um termo que tem origem na filosofia, que busca estudar a natureza do ‘ser’ e a ‘existência’. Sendo aplicada à ciência da computação, o significado desse termo pode ser definida como um conjunto de conceitos fundamentais e suas relações [36].

O termo ontologia é consolidado na comunidade de Engenharia do Conhecimento, sendo que muitas definições têm sido criadas, evoluindo e sendo modificadas ao longo do tempo [37]. As definições mais citadas, no contexto das áreas de Ciência da Informação e Ciência da Computação, tomam por base a proposta de Gruber [14], que afirma que ontologia é a ‘especificação de uma conceituação’. Essa definição é ampliada por Borst [38], que afirma que ontologia é ‘especificação formal e explícita de uma conceitualização compartilhada’.

Para Berners-Lee [15], considerado o criador da Web e o idealizador da Web Semântica, ontologia é um documento ou arquivo que formalmente define os relacionamentos entre termos.

Na área computacional, Guarino *et al.* [39] afirmam que ontologias são um meio de modelar formalmente a estrutura de sistemas, ou seja, as entidades e relacionamentos re-

levantes, oriundas de sua observação, sendo úteis para os seus propósitos. Elas oferecem a base necessária para solucionar problemas inerentes à construção de sistemas que utilizam bases de dados como forma de representar os dados formalmente [40] [41].

O uso de ontologias permite o estabelecimento de banco de dados semânticos, os quais possibilitam interfaces para usuários efetuarem pesquisas em bases de dados sem conhecimento da estrutura ou detalhe técnico a respeito das fontes de dados [42].

As ontologias ganharam grande popularidade na comunidade de Inteligência Artificial (IA) como uma forma de estabelecer um vocabulário formal para compartilhamento entre aplicações [43]. Nesse contexto da IA, ontologias são definidas como um artefato constituído de vocabulário específico, utilizado para descrever uma determinada realidade e um conjunto de suposições explícitas, relacionadas ao significado intencional das palavras [44]. Sua importância também tem sido reconhecida nas áreas de Representação do Conhecimento, Engenharia do Conhecimento, Modelagem da Informação, Recuperação da Informação, Tradução da Linguagem Natural, entre outras [45].

Uschold e Gruninger [46] destacam três categorias de uso para ontologias: comunicação entre pessoas e organizações, interoperabilidade entre sistemas e construção de sistemas. Também classificam ontologias de acordo com o grau de formalidade usado na definição dos termos, que podem ir de altamente informal (linguagem natural), passando pela semi-informal (linguagem natural restrita e estruturada de forma a reduzir ambiguidades), e formal (linguagem artificial definida formalmente), com termos definidos por meio de semântica formal, teoremas e provas de propriedades.

As principais propriedades de uma ontologia são consideradas o compartilhamento e a filtragem [47]. O compartilhamento toma por base o emprego de uma ontologia comum entre dois ou mais agentes diferentes. A filtragem é percebida sob o ponto de vista da abstração. Geralmente as pessoas consideram modelos da realidade, que, por definição, expressam somente uma parte da realidade, e a ontologia é quem define o que poderia ser extraído dessa realidade (as características mais relevantes para o domínio do problema), de modo a se construir um modelo para o sistema.

Diversas metodologias estão relacionadas ao desenvolvimento e manutenção de ontologias. A metodologia proposta por Noy e McGuinness [48] descreve os passos essenciais para o desenvolvimento de uma ontologia. Segundo as autoras, não há uma maneira ou metodologia exata para o desenvolvimento de ontologias. A abordagem deve ser iterativa, seguindo passos mais genéricos até alcançar o nível de refinamento almejado, destacando decisões de modelagem a serem tomadas, bem como os prós, contras e implicações de diferentes soluções. Esta metodologia será apresentada no próximo Capítulo.

Algumas razões para a criação de ontologias podem ser destacadas [48]:

- compartilhar o entendimento comum da estrutura da informação entre pessoas e agentes de software;
- possibilitar o reuso de um domínio do conhecimento;
- separar o domínio do conhecimento do conhecimento operacional;
- analisar um domínio do conhecimento; e
- fazer suposições explícitas de um domínio.

É impossível construir uma ontologia suficientemente rica para todos os fins e domínios, sendo algumas características desejáveis em uma ontologia [49]:

- aberta e dinâmica para se ajustar às mudanças e novos desenvolvimentos em um domínio. Uma ontologia deve ser aberta e dinâmica, tanto em termos de seus algoritmos quanto da sua estrutura;
- deve ser facilmente escalável, considerando um domínio amplo e adaptável a novos requisitos.
- deve ser interoperável, possibilitando a integração de várias ontologias;
- deve ter uma estrutura simples, limpa e modular para ser de fácil entendimento pelas pessoas, facilitando a sua manutenção;
- deve ser semanticamente consistente, guiando a escolha dos termos escolhidos; e
- não deve conter termos muito específicos para não tornar complexa a associação com as fontes de dados e as futuras integrações com outras ontologias.

Adicionalmente, Gruber [14], afirma que as mais importantes características que uma ontologia deve possuir são as seguintes:

Clareza: as definições devem ser objetivas. Sempre que for possível, uma definição deve ser declarada através de axiomas lógicos. Todas as definições devem ser documentadas com linguagem natural.

Coerência: caso uma sentença inferida a partir de axiomas contradiga uma definição ou exemplo dado, então a ontologia é incoerente.

Extensível: uma ontologia deve permitir que novos termos possam ser definidos para usos especiais baseados no vocabulário existente, de maneira que não seja requerida a revisão das definições previamente existentes.

Mínimo compromisso com implementação: a conceituação deve ser especificada no nível do conhecimento, isto é, sem depender de uma codificação particular no nível simbólico ou de codificação.

O emprego de ontologias propicia o entendimento compartilhado de um determinado domínio de interesse, que pode ser utilizado como uma estrutura unificada para solucionar vários tipos de problemas. Geralmente, é expressa por um conjunto de conceitos, suas definições e seus inter-relacionamentos [50]. Sendo assim, em termos gerais, uma ontologia é composta por classes que representam os conceitos dentro de determinado domínio/contexto, relações entre os conceitos do domínio especificado, instâncias que são derivadas das classes e axiomas pertencentes ao domínio, modelando as regras das instâncias [51].

Nesse contexto de composição de uma ontologia, Sales *et al.* [52] destacam a partes que compõe a sua estrutura:

Conceitos: ideias básicas sobre o que se pretende formalizar;

Classe e Subclasses: que podem estar organizadas em uma taxonomia;

Relações: que devem representar os tipos de interação entre as classes de um domínio;

Funções: são casos especiais de relações no qual os elementos dos relacionamentos são únicos para os elementos anteriores;

Axiomas: são teoremas que se declaram sobre as relações que devem cumprir todos elementos da ontologia; e

Instâncias: são utilizadas para representar objetos determinados de um conceito.

Cabe ressaltar que os conceitos e relações formam a base da ontologia, porém uma característica essencial das ontologias é a definição de axiomas. Propor uma taxonomia ou um conjunto de termos básicos, não constitui uma ontologia, devendo os axiomas proverem a definição a semântica dos termos. Esses axiomas especificam definições de termos na ontologia e restrições sobre sua interpretação [53].

As ontologias têm sido aplicadas em muitas maneiras diferentes, mas o núcleo de seu significado dentro ciência da computação é ser um modelo para descrever o mundo que consiste em um conjunto de objetos, propriedades e relacionamento. Além disso, há geralmente uma expectativa de que haja uma estreita semelhança entre o mundo real e as características do modelo em uma ontologia [24].

Ontologias propiciam um vocabulário controlado de uso geral, possibilitando o adequado entendimento de termos que descrevem os dados, além de explicitar tipos de entidades e a forma na qual estão relacionadas, atribuindo sentido e significado a elas. O uso

de ontologias possibilita a construção e uma relação organizada entre termos dentro de um domínio. Com isso, é favorecida a contextualização dos dados, tornando o processo de interpretação dos dados mais eficiente, além de facilitar a recuperação da informação por ferramentas computacionais [54].

Podem ser consideradas basicamente duas maneiras de representar ontologias: a representação formal e a segunda é a representação gráfica. A representação formal é utilizada para que as ontologias possam ser reconhecidas por computadores, enquanto a representação gráfica é empregada para entendimento humano [51]. Ambas são relevantes, uma vez que a falta de uma delas afeta a qualidade e o emprego uso da ontologia.

Quanto a representação formal, existem algumas linguagens para o trabalho com ontologias. Normalmente, usa-se lógica de predicados, lógica descritiva ou linguagens baseadas em *frames*. No entanto, linguagens mais expressivas (conhecidas como linguagens de descrição de ontologias) como a RDF, RDF-S e OWL, proporcionam a base para a criação de vocabulários utilizados para descrever entidades do mundo real e seus respectivos relacionamentos [13]. Tais linguagens serão descritas na próxima subseção.

A ontologia é usada para melhorar a comunicação humana ou entre computadores. De maneira concreta, a ontologia é usada para auxiliar a comunicação entre agentes humanos, para atingir a interoperabilidade entre sistemas computacionais, ou para aprimorar processos ou melhorar a qualidade de produtos de engenharia de software [55].

A partir do que foi apresentado, é possível inferir que ontologias são ferramentas úteis poderosas para a representação do conhecimento, proporcionando a organização das informações organizacionais. As ontologias viabilizam a comunicação e a disponibilização de informações de forma estruturada, dentro de um domínio de conhecimento. Com base na formalidade adotada para o emprego da ferramenta, a especificação do domínio apoia a eliminação de possíveis inconsistências e ambiguidades, tornando a informação representada uma fonte confiável de conhecimentos úteis, direcionando a tomada de decisão.

3.1.5 *Web Ontology Language*

Existem diversas linguagens para construção de ontologias, com suas respectivas funcionalidades. A *Web Ontology Language* (OWL) é o padrão mais recente de linguagens para ontologias, recomendada pelo *World Wide Web Consortium* (W3C). Sua primeira versão foi criada em 2004, sendo a sua versão corrente a OWL2, publicada em 2009 e atualizada em 2012 [56]. Trata-se de uma linguagem de ontologias para a sistemas Web, com significado definido formalmente. Ela fornece classes, subclasses, propriedades e relacionamentos, sendo estes dados armazenados como documentos. Atualmente, é a linguagem mais utilizada para representar ontologias formalmente, possuindo variantes da linguagem

que lidam com a escalabilidade e expressividade das ontologias e permite que aplicações com diferentes propósitos sejam construídas [36].

A linguagem foi projetada para possibilitar a utilização por aplicações que necessitem processar o conteúdo de informações, não se restringindo apenas à sua apresentação [57]. Para reduzir o problema de interpretação das representações dos dados, é adequada a utilização de mecanismos e linguagens de representação. Nesse contexto, a OWL explicita as relações entre conceitos, permitindo que tanto pessoas quanto máquinas possam compreender os conceitos que representam os dados disponibilizados, por meio do uso de descrições ontológicas [51]. Assim, no contexto das aplicações Web, uma das formas mais robustas para representação do conhecimento ocorre por meio do uso de ontologias e da OWL.

As classificações entre as sublinguagens da OWL podem ser descritas conforme o que segue [57]:

OWL-Lite: pode ser considerada a menos expressiva dentre as sublinguagens. É empregada onde apenas são necessárias restrições e uma hierarquia de classes simplificada;

OWL-DL: é direcionada aos usuários que necessitam do máximo de expressividade, sem perder o suporte computacional e o poder de raciocínio da linguagem. Permite computar, de forma automática, a hierarquia de classes, bem como verificar inconsistências na ontologia;

OWL-Full: fornece suporte para usuários que necessitam do máximo de expressividade, porém sem necessariamente permitir garantias computacionais. Cabe destacar que não é possível efetivar regras de inferências utilizando esta sublinguagem.

Ressalta-se que desenvolvedores de ontologias utilizando a *Web Ontology Language* devem considerar o emprego de uma sublinguagem que seja mais adequada as necessidades da aplicação a ser produzida.

Existem elementos básicos que compõe uma ontologia, conforme já explicitado na Subseção anterior. Quando da aplicação da modelagem de uma ontologia OWL, podem ser destacadas as suas classes (*owl:Class*), suas respectivas propriedades (*owl:ObjectProperty* e *owl:DataProperty*), seus indivíduos ou instâncias (*owl:Individual*), além de diversos outros elementos que permitem efetivar o relacionamento entre os componentes da ontologia.

Conforme o exposto na Figura 3.1, é possível verificar um exemplo de representação formal utilizando a OWL. No caso em questão, uma bicicleta (*bicycle*) é um tipo de veículo (*vehicle*) de transporte. Sendo assim, é definida a relação '*bicycle is-a vehicle*'. Adicionalmente, uma bicicleta pode possuir a especialização bicicleta esportiva (*sport bicycle*), bem como pode ser bicicleta para uso na cidade (*city bicycle*). Por fim, a linguagem permite destacar os atributos e propriedades da bicicleta, como a cor, entre outros.

```

<owl:Class rdf:ID="Vehicle">
  <rdfs:label>Vehicle</rdfs:label>
  <rdfs:SubClassOf rdf:resource="#Any" />
</owl:Class>
<owl:Class rdf:ID="sport_cycle">
  <rdfs:label>sport_cycle</rdfs:label>
  <rdfs:SubClassOf rdf:resource="#bicycle" />
</owl:Class>
<owl:Class rdf:ID="city_cycle">
  <rdfs:label>city_cycle</rdfs:label>
  <rdfs:SubClassOf rdf:resource="#bicycle" />
</owl:Class>
<owl:Class rdf:ID="bicycle">
  <rdfs:label>bicycle</rdfs:label>
  <rdfs:SubClassOf rdf:resource="#Vehicle" />
  <rdfs:SubClassOf>
    <owl:Restriction>
      <owl:cardinality rdf:datatype="http://www.w3.org/
        <owl:onProperty rdf:resource="#has_body_color" />
    </owl:Restriction>
  </rdfs:SubClassOf>
  <rdfs:SubClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#has_body_color" />
      <owl:allValuesFrom rdf:resource="#Color" />
    </owl:Restriction>
  </rdfs:SubClassOf>

```

Figura 3.1: Representação Formal da OWL - Adaptado pelo autor (Fonte: [51]).

É importante ainda destacar que a OWL não é uma linguagem de programação, sendo uma linguagem declarativa que descreve um determinado universo, permitindo inferir novas informações sobre o domínio em questão. Porém, o formato com o qual as ferramentas empregam a inferências sobre ontologias não configura parte do escopo da OWL. Além disso, a linguagem não deve ser confundida com um banco de dados. É necessário o entendimento de que documentos OWL registram informações (instâncias da ontologia) e, sendo assim, realizam o armazenamento de dados [58].

Por fim, a OWL é uma linguagem de marcação semântica para publicação e compartilhamento de ontologias. Ela toma por base especificações do *Resource Description Framework* (RDF)/*Resource Description Framework Schema* (RDF-S). Sendo assim, além de herdar as características do RDF, como a estrutura baseada em triplas, possui também uma semântica descrita no RDF-S [59]. Tais conceitos serão apresentados, de forma mais específica, na próxima Subseção.

3.1.6 *Resource Description Framework*

O RDF (*Resource Description Framework*) é reconhecido com um modelo que descreve recursos de informação na Web. Trata-se de uma linguagem de representação de informações, que proporciona a sua descrição formal, permitindo o seu acesso por máquinas [51].

Dentre os principais objetivos do RDF, destaca-se a construção de uma rede de informações formada por nós ligados semanticamente, dotados de informações oriundas das

mais diversas fontes [54]. De forma paralela, as ontologias são consideradas recursos empregados para efetivar uma relação organizada entre termos pertencentes a determinado domínio contextualizado, permitindo a interpretação e a recuperação de dados de forma mais eficiente [60]. Nesse contexto, o RDF apresenta um formato capaz de representar o conhecimento, empregando recursos de ontologias.

É importante salientar que um sistema da informação é considerado um espaço de informação no qual os itens de interesse precisam ser identificados. Sendo assim, é necessário que cada recurso possua um identificador único e global (URI - *Uniform Resource Identifier*), de forma a possibilitar a sua identificação [61]. Como exemplo, o URI `http://test.example.com/bsbcovid` poderia ser empregado como identificador único do recurso Relatório de Casos de Covid-19 em Brasília.

O URI proporciona uma maneira simplificada e única de identificar recursos em sistemas informacionais, sendo caracterizado por três aspectos [51]:

Uniformidade: utilização de recursos tanto no mesmo contexto quanto em contextos diferenciados;

Recurso: qualquer coisa que pode ser identificado por um URI, como um vídeo, imagem, serviço, documento, entre outros; e

Identificador: informação requerida para identificar e diferenciar um determinado recurso de qualquer outro.

É possível descrever esses recursos por diversas propriedades, denominadas descrição RDF. Essas propriedades possuem um tipo (*type*), bem como um valor (*value*), o que forma a referida descrição. A Figura 3.2 caracteriza o RDF, relacionando as propriedades RDF a um recurso, por meio de um URI. A relação RDF com URI é um conceito chave para referenciar e descrever um recurso de forma única e não ambígua [51].

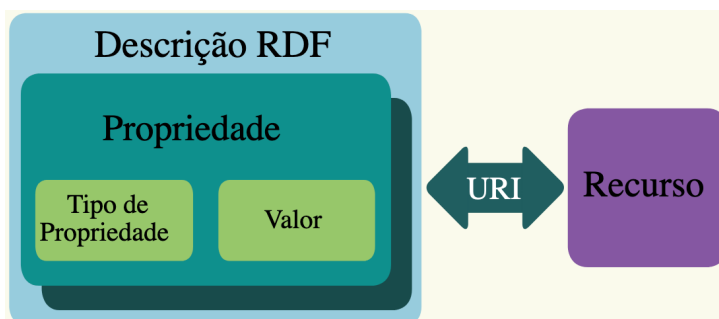


Figura 3.2: Caracterização do RDF (Fonte: [51]).

O descrição semântica fornecida pelo RDF é caracterizada por uma coleção de triplas que podem ser visualizadas como um multigrafo dirigido, sendo um modelo adequado para a representação de informações [41].

Uma tripla RDF é composta por três elementos: sujeito, predicado e objeto. Uma instância RDF representa a relação entre dois recursos, o sujeito e o objeto. O predicado representa a natureza da relação entre eles. A Figura 3.3 ilustra esse relacionamento.

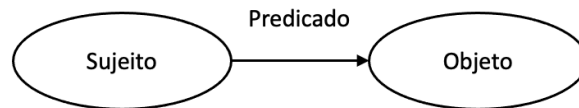


Figura 3.3: Representação de Tripla RDF.

Cabe ressaltar que esse grafo representativo normalmente é composto por diversas triplas que irão compor o documento RDF como um todo. Um exemplo simplificado de uma tripla RDF, indicando cada URI a ser empregado para identificar os recursos, pode ser visualizado na Figura 3.4.

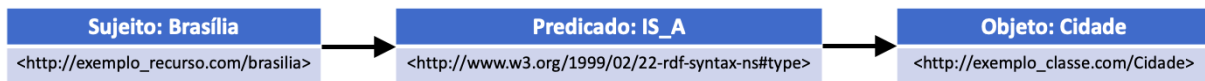


Figura 3.4: Exemplo de Tripla RDF.

Neste trabalho, a interação entre a Arquitetura da Informação implementada e a ontologia empregada seguiu os padrões do W3C (*World Wide Web Consortium*), empregando os seguintes recursos:

RDF (*Resource Description Framework*): modelo utilizado para representar as informações por meio de triplas;

RDF-S (*Resource Description Framework Schema*): extensão do RDF que permite a descrição semântica e a descrição de grupos de recursos e seus relacionamentos por meio de um vocabulário ampliado [51];

OWL (*Web Ontology Language*): linguagem utilizada para definir e instanciar ontologias, empregando suas classes, propriedades, bem como os seus relacionamentos;

SPARQL (*SPARQL Protocol and RDF Query Language*): linguagem padronizada para realização de consultas a grafos RDF.

3.2 Administração de Dados

Antes da abordagem do assunto desta seção, é relevante ter o entendimento sobre a conceituação de dado, informação e conhecimento, uma vez que esta pesquisa tem como ponto de partida a obtenção de dados, para estruturar informações e gerar conhecimentos necessários à tomada de decisão.

A representatividade dessa tríade possui conceituações bem definidas, conforme o que segue:

A cadeia conceitual que caracteriza a Ciência da Informação vai desde o dado à informação e conhecimento, de acordo com a ideia de muitos de seus autores, algumas vezes incluindo saber; num crescendo de complexidade, da forma bruta e primitiva do dado à sua elaboração como informação, e sua absorção, quando relevante, na estrutura cognitiva, transformando-se em conhecimento. Esta rede de conceitos poderá ter seu processo final na cultura, aqui considerando a incorporação dessas informações relevantes entre outras manifestações e produções e vivências do homem, individuais e coletivas [62].

Assim, algumas conceituações de interesse pode ser ressaltadas [63]:

- dado é um conjunto de registros qualitativos ou quantitativos, conhecido, que organizado, agrupado, categorizado e padronizado adequadamente transforma-se em informação;
- informação são dados organizados de modo significativo, sendo subsídio útil à tomada de decisão;
- o conhecimento são informações processadas por indivíduos, após um processo de compreensão análise e avaliação.

O ato de organizar a informação pode ser entendido como um tipo particular de uso da linguagem. Diversas vantagens e possibilidades podem ser obtidas com a organização da informação, destacando-se a utilização de constructos da linguística, como vocabulário, semântica e sintaxe, para a generalização de entendimentos e avaliação de conteúdos [64].

Organizar é considerada uma característica básica dos seres humanos, sendo o aprendizado humano baseado na capacidade de analisar e organizar os dados, as informações e os conhecimentos [27].

As informações necessitam ser organizadas para que seja possível realizar a sua recuperação [27]. É possível afirmar que o objetivo da organização da informação e do conhecimento é dar suporte ao fluxo do tratamento e da recuperação dos objetos informacionais [54].

A Administração de Dados pode ser definida como uma função da organização que tem por reponsabilidade desenvolver e administrar, de forma centralizada, estratégias, procedimentos, práticas e planos que possuem capacidade para disponibilizar dados corporativos, quando necessários, dotados de consistência, privacidade e integridade [65]. Trata-se do processo de gerenciamento de dados institucionais para proporcionar informações seguras e confiáveis, adequadas às necessidades organizacionais em todos os níveis.

Na administração de dados deve haver um foco na qualidade dos dados, bem como em seus usuários, sejam eles analistas de sistemas, projetistas, desenvolvedores, administradores de bancos de dados e gestores responsáveis pelos dados [25]. Isso possibilitará aos usuários a disponibilidade de dados com qualidade, prontos para empregos diversos.

O Elemento de Dados é a unidade fundamental dos dados a serem gerenciados por uma organização, sendo considerada indivisível contextos específicos [66].

As características do dado refletem o tipo de dado a ser armazenado por um elemento de dados, de acordo com o que é utilizado pela organização, seja caractere, numérico, imagem, entre outros. É possível subdividir um elemento de dados em classe de objeto, propriedade e representação [25].

As classes de objeto se referem aos dados que se deseja coletar e armazenar. São considerados conjuntos de ideias ou abstrações do mundo real que podem ser identificadas, possuindo propriedades e comportamentos seguem regras similares.

Já uma propriedade pode ser entendida como uma característica pertencente a todos os elementos de uma classe de objetos [25]. Como exemplo o termo *Idade* pode ser considerada uma propriedade da classe *Pessoa*.

A representação descreve como os dados são apresentados, ou seja, uma combinação de domínio de valores, tipos de dados e, se necessário, uma unidade de medida ou conjunto de caracteres. Em síntese, a representação especifica a forma como o dado é transcrito [25].

Os elementos de dados podem ser identificados por meio de um modelo de Entidades e Relacionamentos (ER), ou mesmo por modelos de classes usados no domínio de Orientação a Objetos [25]. Dessa forma, em um modelo ER, um atributo de uma entidade pode ser considerado equivalente a um elemento de dados. A Figura 3.5 mostra um exemplo de modelo de uma entidade com seus atributos, sendo os elementos de dados identificados.

Entidade	Atributos	Elementos de Dados
Local	Cidade Bairro Descrição	Local Cidade Local Bairro Local Descrição

Figura 3.5: Elementos de dados - Modelo ER.

A administração de dados participa diretamente da gestão dos sistemas de informação, apoiando a modelagem de dados, integrações, padrões e na documentação produzida. Além disso, pode especificar quesitos de segurança e privacidade [65]. A Figura 3.6 apresenta o ciclo que envolve os dados, a sua transformação em informação, a estruturação das informações em sistemas, bem como a interação da administração permeando todo o processo.

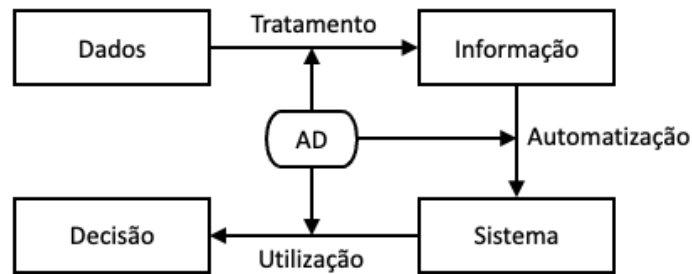


Figura 3.6: Administração de Dados - Adaptado pelo autor (Fonte: [65]).

3.3 Gestão da Informação

Após destacados aspectos julgados relevantes no que diz respeito à Administração dos Dados, serão apresentados conceitos relacionados à Gestão da Informação e sua importância em ambientes organizacionais.

A informação é considerada um componente presente nas mais diversas áreas de uma organização, sendo vital para tomadas de decisão. Sem a correta compreensão dos processos organizacionais, bem como sobre a transformação das informações em conhecimentos, as empresas enfrentam dificuldades para a percepção da importância de suas fontes de tecnologia da informação [67].

O trâmite de informações em ambientes organizacionais ocorre, frequentemente, sem o correto entendimento de seu impacto, valor, bem como de seu custo. Assim, é necessária uma gestão efetiva, proporcionando o controle adequado das informações de uma organização [68].

A gestão da informação pode ser identificada como uma rede de processos que tem por objetivo adquirir, criar, organizar, distribuir e utilizar as informações. Com isso, esse processo de gestão da informação é analisado como um ciclo contínuo de 6 (seis) processos que se relacionam, conforme destacado na Figura 3.7.

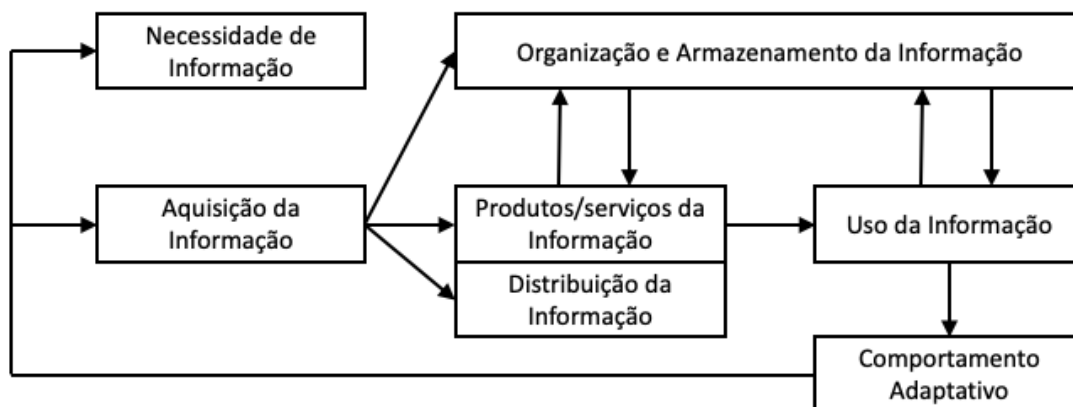


Figura 3.7: Processo de Gestão da Informação - Adaptado pelo autor (Fonte: [67]).

Nesse contexto, cada parte do processo de gestão da informação apresentado tem cada fase definida conforme o que segue [67]:

Necessidade de Informação: a partir de problemas, incertezas dentre questões que são resultantes de situações específicas, são identificadas as necessidades de informação. A origem dessa necessidade está relacionada não apenas às questões subjetivas, mas também à cultura organizacional, à execução de tarefas, aos objetivos traçados, entre outros;

Aquisição da Informação: dentro da gestão da informação, é considerada uma função crítica e complexa, buscando o equilíbrio entre duas demandas: as diversas necessidades de informação da organização e a seleção das informações que a organização dará atenção;

Organização e Armazenamento da Informação: parte da informação obtida ou criada é fisicamente organizada em arquivos, em um Sistema de Gerenciamento de Banco de Dados, dentre outras possibilidades, de forma a facilitar a distribuição. A maneira como a informação é armazenada está relacionada com a percepção e a representação do ambiente pela organização;

Produtos e Serviços de Informação: uma das principais funções da gestão da informação é garantir que as necessidades de informação sejam atendidas com um equilíbrio de produtos e serviços;

Distribuição da Informação: o objetivo da distribuir as informações é prover e facilitar a disseminação de informações, sendo essencial para proporcionar significado, construir conhecimento e a apoiar a tomada de decisão;

Uso da Informação: trata-se de um processo social dinâmico de pesquisa e construção, que tem como resultado a criação de significado, a construção do conhecimentos pertinentes, bem como a a seleção de padrões de ação. A informação é buscada e utilizada em todo processo de tomada de decisão, nos mais diversos níveis organizacionais.

Cabe destacar que os administradores de uma organização necessitam de uma visão abrangente para o tratamento das informações, de forma a possibilitar a percepção de alterações do ambiente interno e externo à organização, proporcionando a adaptação à realidade [69]. Esse tipo de abordagem da gestão informacional foi intitulada por *ecologia da informação*, possuindo quatro principais parâmetros: Integração dos diversos tipos de informação; Reconhecimento de mudanças evolutivas; Ênfase na observação e na descrição; e Ênfase no comportamento pessoal e informacional.

Dentro do que foi apresentado nessa seção do Capítulo, cabe ressaltar que a pesquisa desenvolvida não tem como objetivo um modelo de gestão informacional especificamente completo. Apesar disso, pretende-se apresentar uma Arquitetura da Informação que integre os principais conceitos de gestão da informação apresentados.

3.4 Usabilidade

A Experiência do Usuário (*User Experience - UX*), na abordagem de sistemas informacionais, pode ser entendida como um conjunto de atributos que determinam o nível de satisfação do usuário ao interagir com as ferramentas disponibilizadas, tanto no que diz respeito ao funcionamento estrutural, quanto em relação ao conteúdo obtido.

A definição formal de *User Experience* inclui três fatores que afetam diretamente a experiência: o sistema, o usuário e o contexto de uso [70]. A UX pode ser definida como a percepção e a resposta de uma pessoa, resultante do uso, ou antecipação do uso, de um produto, sistema ou serviço [71]. Nesse sentido, a experiência do usuário está relacionada a todos os fenômenos que ocorrem antes, durante e depois da interação com determinado produto ou serviço. De forma a serem alcançadas melhorias para otimizar a *User Experience*, é essencial obter um adequado entendimento sobre as necessidades dos usuários, seus objetivos, suas habilidades e limitações, proporcionando o desenvolvimento de aplicações que atendam as demandas a serem solucionadas.

Dentre os principais atributos da UX, tem papel destacado a Usabilidade. A primeira definição para o termo surgiu em 1991, descrito na norma ISO/IEC 9126 [72] sobre a qualidade de software, como sendo uma abordagem orientada ao produto e ao usuário. A partir dessa norma, a Usabilidade ultrapassou os limites do ambiente acadêmico da

Psicologia Aplicada e da Ergonomia, passando a fazer parte do vocabulário técnico de outras áreas do conhecimento [73].

A usabilidade surgiu por meio de raízes interconectadas com fatores e disciplinas como a computação gráfica, as interfaces humanas, os processos cognitivos, a engenharia industrial, entre outras diversas [74] .

Um dos fatores de grande relevância para a eficiência de sistemas de informação diz respeito a qualidade de sua interface gráfica. Os critérios de usabilidade fornecem parâmetros que permitem auferir o nível de eficiência da interface, bem como permite a percepção de como é realizada a interação entre os usuários e os respectivos sistemas informacionais [75].

O termo usabilidade pode ainda ser definido como a qualidade da interação dos usuários com uma determinada interface [76]. A qualidade está relacionada à facilidade de aprendizado, a facilidade de lembrar como realizar uma tarefa após algum tempo, a rapidez no desenvolvimento de tarefas, a baixa taxa de erros e a satisfação do usuário.

É importante destacar a importância da usabilidade para os usuários dos sistemas informacionais. A informação deve existir para servir aos usuários, sendo assim, é essencial melhorar cada vez mais a utilização das interfaces das ferramentas tecnológicas, de forma a permitir que os usuários recuperem as informações demandadas de maneira eficaz, eficiente e satisfatória [77]. Além disso, é importante conhecer as características dos usuários, de forma a mapear aspectos do sistema, definindo informações a serem disponibilizadas, bem como a forma como serão apresentadas.

No contexto da pesquisa realizada, o conceito de usabilidade é essencial para a Arquitetura da Informação construída, sendo uma estratégica relevante para a organização da informação. Assim, empregando princípios e boas práticas relacionados à usabilidade, acredita-se que será possível obter melhores resultados, com um maior nível de eficiência na execução das atividades propostas.

3.5 Arquitetura da Informação

O termo ‘Arquitetura da Informação’ (AI) remonta à década de 1990, época na qual Saul Wurman originalmente cunhou a expressão, mais precisamente em 1976. A ligação metafórica que relaciona a ‘arquitetura’ à ‘informação’ teve como origem a percepção de Wurman, em que as questões de coleta, organização e representação da informação possuem semelhanças em relação à concepção de edifícios por arquitetos [22]. Figura 3.8 apresenta esse relacionamento de um projeto de construção às visões de uma Arquitetura da Informação.

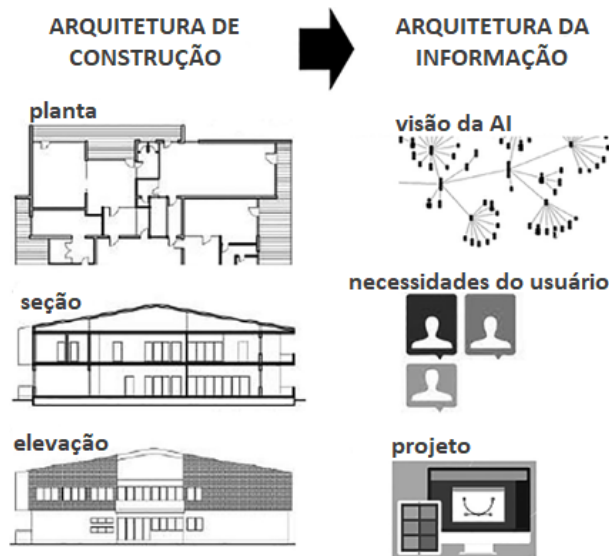


Figura 3.8: Projeto de Construção VS Arquitetura da Informação - Adaptado pelo autor (Fonte: [22]).

Não existe uma definição totalmente precisa sobre o significado ou constituição de uma Arquitetura da Informação. No entanto, vários pesquisadores que abordam o assunto descrevem uma grande quantidade e diversidade de definições aceitas na comunidade, conforme o que segue:

Wurman - define a Arquitetura da Informação como a construção de uma estrutura para auxiliar as pessoas na compreensão da informação [78];

Brancheau e Wetherbe - a AI consiste em um plano para modelagem dos requisitos informacionais de uma organização, provendo um modo de mapear as informações necessárias à própria organização [79];

Davenport - a AI é constituída por uma série de ferramentas que adaptam os recursos às necessidades da informação, estando relacionada aos comportamentos, processos e pessoal especializado, bem como outros aspectos da organização [69];

Morville - a Arquitetura da Informação consiste no projeto estrutural para o compartilhamento de ambientes informacionais [19];

Jonkers e Lankhorst - a Arquitetura da Informação é a organização fundamental de um sistema materializada em seus componentes, as relações entre seus componentes e entre componentes e o meio ambiente [80];

McGee e Prusak - uma AI deve ser capaz de representar o ambiente de informação de uma empresa, alcançando o equilíbrio entre as necessidades de informações e as limitações da tecnologia [81];

Bailey - AI é a arte e a ciência de organizar sistemas de informação para auxiliar o usuário a alcançar seus objetivos [82].

Ainda dentro desse contexto, Rosenfeld e Morville, autores também considerados pioneiros no assunto, realizaram um trabalho de desenvolvimento de uma Arquitetura da Informação para o desenho de ambientes digitais, principalmente para a Internet [18]. Essa arquitetura foi delineada no contexto das tecnologias emergentes para web, visando conceber ambientes online que permitem aos usuários encontrar informações de interesse.

Em uma Arquitetura da Informação, a tecnologia possui grande relevância. O objetivo da AI é a organização e armazenagem dos objetos informacionais em repositórios informacionais (bancos de dados, sistemas de arquivos, etc). A arquitetura deve ser provida de consistência, compartilhamento, documentação, privacidade e recuperação eficaz de seus conteúdos, sem se prender a técnicas específicas de modelagem de dados ou arquitetura de sistemas de informação [25].

No modelo apresentado na Figura 3.9, a Arquitetura da Informação é representada como uma interseção entre contexto, conteúdo e usuários [18]. Em uma organização, existe a necessidade de se conhecer os objetivos de negócio (contexto), estar consciente da natureza e do volume de informações existentes e de sua taxa de crescimento (conteúdo), bem como é necessário, também, entender as necessidades e os processos de busca do público-alvo (usuários) [25].

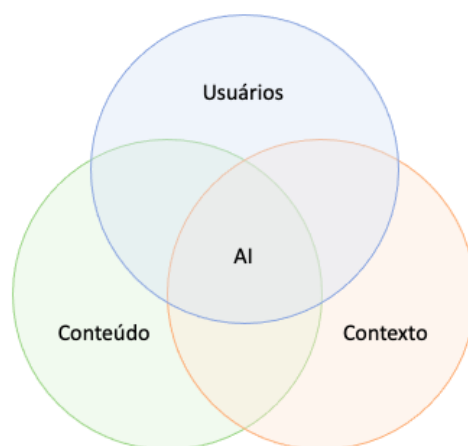


Figura 3.9: Modelo de Arquitetura da Informação - Adaptado pelo autor (Fonte: [18]).

Resmini e Rosati propuseram um vocabulário heurístico que deve ser empregado na construção de uma Arquitetura da Informação abrangente. Este vocabulário é composto por cinco princípios orientadores, conforme o que segue [21]:

Criação de lugares: capacidade de um modelo de Arquitetura da Informação ajudar os usuários a reduzir a desorientação, construir um senso de lugar e aumentar a legibilidade e a determinação de caminhos em ambientes digitais, físicos e entre canais;

Consistência: capacidade de um modelo abrangente de AI se adequar aos propósitos, contextos e pessoas para os quais foi projetado (consistência interna) e manter a mesma lógica ao longo de diferentes meios, ambientes e tempos em que atua (consistência externa);

Resiliência: capacidade do modelo de Arquitetura da Informação em moldar e adaptar-se aos utilizadores específicos, necessidades e busca de estratégias;

Redução: capacidade do modelo de AI em gerenciar grandes conjuntos de informações e minimizar o estresse e frustração associada à escolha um conjunto crescente de fontes de informação, serviços e bens;

Correlação: capacidade de um modelo de Arquitetura da Informação para sugerir conexões relevantes entre peças de informações, serviços e bens, de forma a auxiliar os usuários atingir objectivos explícitos ou estimular necessidades latentes.

O Instituto Asilomar para a Arquitetura de Informação (AIfIA) se dedica ao avanço do design de ambientes de Informações compartilhadas, por meio de pesquisa, educação, apoio e serviços. O instituto atesta três definições para Arquitetura da Informação [83]:

- o design estrutural de ambientes de informações compartilhadas;
- a ciência e arte de organizar e rotular Web sites, intranets, comunidades on-line e software para dar suporte à usabilidade e facilidade de encontrar informação; e
- uma comunidade de práticas emergente, focada em trazer princípios do design e arquitetura para o ambiente digital.

A partir do que foi exposto, é possível inferir que a Arquitetura da Informação está diretamente relacionada à representação semântica da informação, na organização de sua armazenagem, bem como na otimização de sua recuperação. Dentro de um ambiente que possui uma Arquitetura da Informação planejada, as informações são organizadas em uma estrutura compreensível, seguindo uma lógica adequada, levando em consideração a possibilidade de interação, garantindo maior qualidade em sua recuperação.

Capítulo 4

Trabalhos Relacionados

Este Capítulo tem por finalidade abordar os principais trabalhos que possuem relação com a pesquisa desenvolvida. Cabe ressaltar que as publicações que se relacionam com este trabalho não se restringem ao que será apresentado neste Capítulo. A partir da seleção realizada, foi possível obter subsídios que serviram como referenciais ao desenvolvimento da solução proposta desenvolvida.

Com o emprego da Teoria do Enfoque Meta-Analítico Consolidado, conforme foi apresentado no Capítulo 2, tornou-se possível realizar uma revisão bibliográfica adequada, de forma sistemática, resultando na base para o desenvolvimento da fundamentação teórica da pesquisa, apresentada no Capítulo 3. A aplicação da metodologia resultou na descoberta de diversas publicações de interesse, bem como autores relevantes, o que proporcionou o suporte necessário ao trabalho realizado. Além disso, outras publicações conexas foram selecionadas para a pesquisa, por complementarem os estudos relacionados aos diversos autores levantados.

Nesse contexto, selecionando as principais publicações relacionadas, foram consideradas aquelas que possuem temáticas próximas ao trabalho realizado, servindo como parâmetro inicial, que auxiliaram no direcionamento buscado para o desenvolvimento da pesquisa, conforme o apresentado na Tabela 4.1.

Tabela 4.1: Principais trabalhos relacionados com a pesquisa.

Autores	Títulos
Mårtenson, Horndahl e Kabir	<i>An Ontology-based Adaptive Reporting Tool</i> [40]
Noy e McGuinness	<i>Ontology Development 101: A Guide to Creating Your First Ontology</i> [48]
Victorino <i>et al.</i>	Arquitetura de Publicação de Dados Abertos Conectos Governamentais da Universidade de Brasília [54]

4.1 Ferramenta de relatório baseada em ontologia

A publicação de Mårtenson, Horndahl e Kabir [40] teve por objetivo realizar a descrição conceitual de uma ferramenta de criação de relatórios estruturados, baseados em uma interface de usuário adaptativa orientada por ontologia. O conceito estabelece a base para a implementação de um sistema de informação no terreno, que pode adaptar-se ao contexto da situação de informação, bem como às possíveis necessidades de informação de outros agentes do sistema de informação.

Os autores afirmam que informações referentes a um domínio cognitivo específico, entregues por humanos (normalmente chamadas de dados *soft*), têm como vantagem o seu alto valor informativo, mas possuem a desvantagem de serem entregues usualmente como texto livre, o que embora amigável para humanos, é menos adequado para exploração por meio de processamento automático. Assim, uma questão relevante no gerenciamento desses dados é a transformação de texto livre em um conteúdo estruturado. A estruturação da entrada de dados por humanos proporciona vantagens, como:

- linguagem é mais precisa, evitando que o usuário faça declarações inadequadas;
- formato é mais compacto, o que implica potencial para uma entrada mais rápida; e
- modelo de informação baseado em um entendimento compartilhado, evitando mal-entendidos e aumentando a interoperabilidade em nível semântico.

Nesse contexto, os autores propuseram um conceito geral de relatório humano estruturado, com base em uma interface de usuário adaptável orientada por ontologia. O conceito estabelece a base para a implementação de um sistema de relatório, que pode se adaptar ao contexto específico, bem como às possíveis necessidades de informação. Foram destacados alguns requisitos desejáveis ao sistema:

- deve ser intuitivo para um não especialista, que não é um engenheiro de ontologia nem um especialista de domínio;
- deve ser independente de domínio, ou seja, o sistema deve trabalhar com ontologias de diferentes domínios;
- a saída deve ser triplas RDF aderentes à ontologia;
- deve ser adaptável ao contexto da situação do relato; e
- deve ser adaptável às necessidades de informação de outros agentes de inteligência.

Um ponto de partida apresentado no trabalho é combinar atores, lugares e tipos de eventos observados. Se houver uma correspondência, o usuário poderá acessar informações

adicionais valiosas não relatadas ainda. O processo de correspondência também pode ser feito executando uma consulta SPARQL em uma ontologia específica. A descrição do processo de correspondência foi apresentado, conforme a Figura 4.1.

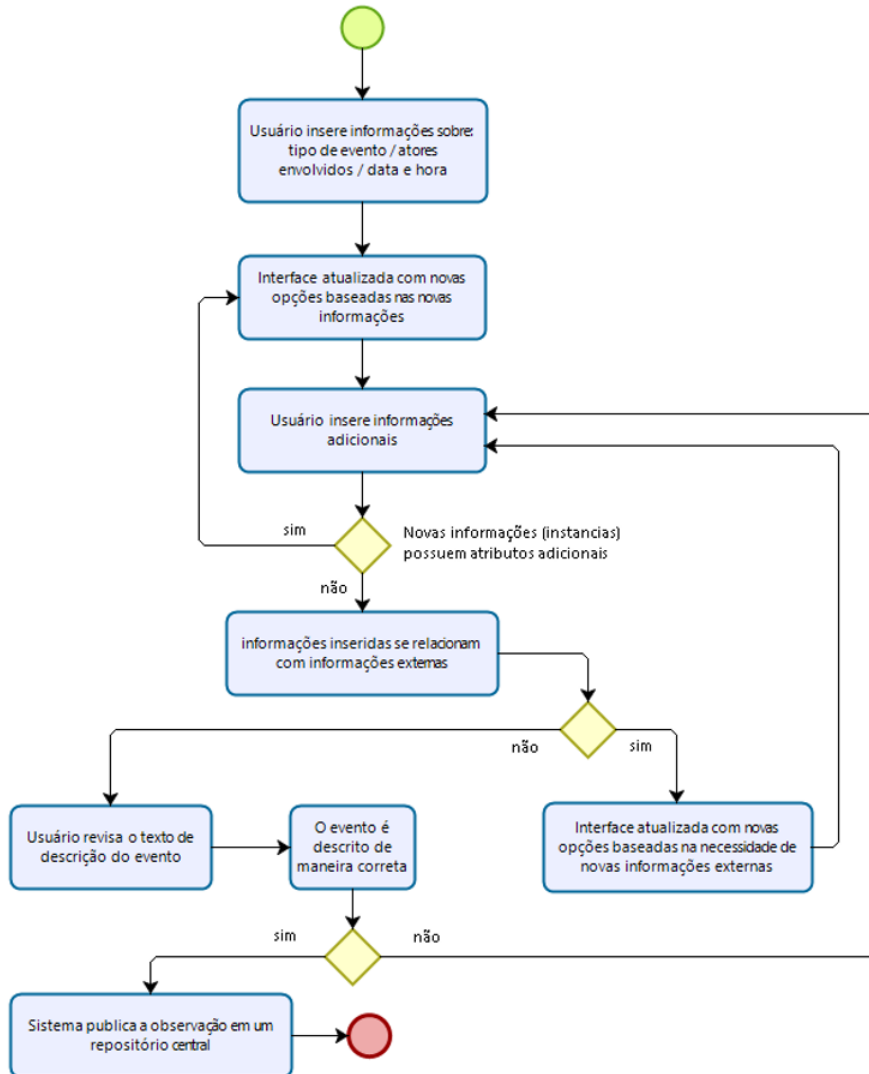


Figura 4.1: Descrição do processo de correspondência - Adaptado pelo autor (Fonte: [40]).

Além do design conceitual da ferramenta e da descrição de como deve ser realizada a combinação da descrição do evento com as necessidades de informações externas, os autores destacam que a interface deve ser adaptável aos diversos contextos.

No presente trabalho, foi desenvolvida uma arquitetura da informação, suportada por uma arquitetura de software, que possui em sua concepção referências da proposta apresentada neste trabalho relacionado. Porém, foi buscado um escopo mais abrangente, bem como uma aplicação prática, compreendendo conceitos adicionais que enriqueceram o que foi proposto.

4.2 Desenvolvimento de ontologia

Dentro do que está sendo realizado nesta pesquisa, não foi buscado o desenvolvimento de uma ontologia completa. A proposta foi construir um protótipo de uma ontologia, direcionada ao domínio de eventos observados em operações militares. Este protótipo foi julgado suficiente para possibilitar a utilização da arquitetura da informação que está sendo formulada.

Para o desenvolvimento do protótipo do modelo ontológico, foram buscadas direcionamentos adequados, que consolidassem boas práticas, de forma a proporcionar um resultado adequado ao que está sendo proposto. Nesse contexto, o trabalho publicado por Noy e McGuinness [48] foi tomado como base, pois fornece um guia abrangente para a criação de ontologias.

Segundo as autoras, é impossível cobrir todos os possíveis problemas que um desenvolvedor pode enfrentar ao construir uma ontologia. Além disso, afirmam que não há uma metodologia julgada correta para desenvolver ontologias. A proposta de seu trabalho é fornecer um ponto de partida que auxiliará o desenvolvedor na construção de ontologias. São apresentadas ideias a partir de experiências vividas em situações anteriores na construção de ontologias.

O processo proposto aponta para uma abordagem iterativa para o desenvolvimento de ontologias. Inicialmente, deve ser construído um esboço, seguido de uma revisão e refinamento. Ao longo do caminho, devem ser discutidas as decisões de modelagem a serem tomadas, bem como os prós, contras e implicações de diferentes possíveis soluções. Consolidando essa ideia, são enfatizadas regras gerais, julgadas assertivas para a construção de uma ontologia:

- não há uma maneira correta de modelar um domínio, pois sempre há alternativas viáveis. A melhor solução quase sempre depende do aplicativo e das extensões previstas;
- o desenvolvimento da ontologia é necessariamente um processo iterativo. A ontologia é criada, avaliada e depurada, sendo necessária uma revisão, sendo um processo que acompanha todo o ciclo de vida da ontologia; e
- conceitos na ontologia devem ser próximos aos objetos (físicos ou lógicos) e relacionamentos em seu domínio de interesse. Normalmente substantivos (objetos) ou verbos (relacionamentos), descrevendo o seu domínio.

O trabalho apresentado por Noy e McGuinness define passos considerados pertinentes para o desenvolvimento de ontologias, conforme o que segue:

Passo 1 - Determine o domínio e o escopo da ontologia: algumas questões devem ser respondidas, como: Que domínio será coberto pela ontologia? Para que a ontologia será utilizada? Que tipo de questões a ontologia deve responder? Quem irá utilizar e manter a ontologia?;

Passo 2 - Considere reutilizar ontologias existentes: quase sempre vale a pena considerar o que já foi desenvolvido, sendo verificado se é possível refinar e estender as fontes existentes para o domínio e tarefa específicos. Reutilizando o existente. Muitas ontologias já estão disponíveis em formato eletrônico e podem ser importados para um desenvolvimento de ontologia. Existem bibliotecas de ontologias reutilizáveis na Web e na literatura;

Passo 3 - Enumere termos importantes na ontologia: é útil escrever uma lista de todos os termos sobre os quais necessitam ser feitas declarações ou explicações para um usuário. É importante definir esses termos, as suas propriedades e o que é necessário explicitar sobre cada um;

Passo 4 - Defina as classes e a hierarquia de classes: Os autores Uschold e Gruninger [46] definem abordagens para o criação de classes e hierarquias: a *top-down*, que inicia com a definição de conceitos gerais e a sua posterior especialização; a *bottom-up* que inicia definindo conceitos mais específicos, agrupando-os em classes gerais; e uma combinação entre as duas, definindo primeiramente os conceitos principais e, em seguida, sua generalização e/ou especialização. A Figura 4.2 mostra um exemplo de diferentes níveis de generalidade de uma ontologia;

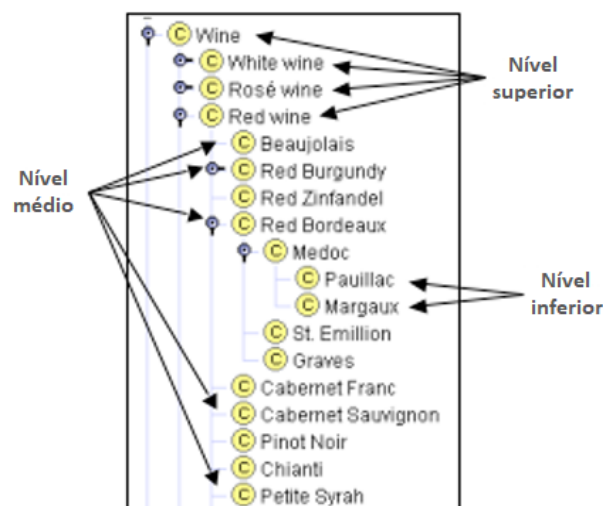


Figura 4.2: Níveis de hierarquia de uma ontologia - Adaptado pelo autor (Fonte: [48]).

Passo 5 - Defina as propriedades das classes (os *slots*): as classes, por si só, não fornecem informações suficientes para responder às questões de competência do domínio. Depois de definidas, deve ser descrita uma estrutura interna de conceitos das classes. Ou seja, uma classe Homem poderá possuir propriedades que descrevam suas características físicas como *Raça*, *Cor_Olhos*, *Cor_Cabelo*, entre outras;

Passo 6 - Defina as facetas dos *slots*: as propriedades das classes podem ter diferentes facetas que a descrevem, como o tipo de valor, valores permitidos, a cardinalidade, entre outros recursos dos valores que o *slot* pode assumir. Além disso, é necessário definir *domains* e *ranges* entre as classes e as propriedades;

Passo 7 - Crie instâncias: a última etapa é a criação de instâncias individuais de classes na hierarquia definida pela ontologia. Isso permitirá verificar o funcionamento e realizar a validação do que foi desenvolvido.

No guia apresentado, as autoras destacaram uma possível metodologia de desenvolvimento de ontologia, listando as etapas descritas no processo e abordando as complexas questões de definição de hierarquias de classe, propriedades de classes e instâncias. Sendo assim, na pesquisa que foi realizada, diversos conceitos apresentados no guia de Noy e McGuinness foram empregados, de forma que a construção do protótipo de ontologia proposto estivesse consonante com os objetivos traçados.

4.3 Arquitetura para dados abertos conectados

O trabalho publicado por Victorino *et al.* [54], em 2020, teve por objetivo apresentar uma Arquitetura da Informação, baseada em sistemas de organização da informação e do conhecimento, para dar suporte à publicação, desde a extração dos dados dos sistemas fontes até o consumo pelos usuários finais, de dados abertos conectados. A Arquitetura formulada foi verificada por meio de um estudo de caso utilizando dados abertos oriundos dos sistemas de informação da Universidade de Brasília.

Cabe evidenciar que para a conexão e integração semântica dos dados na arquitetura proposta, foram empregados metadados, ontologias e o *Resource Description Framework*, estabelecendo um elo entre os conceitos especificados. Assim, foi possível relacionar os vários conjuntos de dados disponibilizados, transformando-os em dados abertos conectados. Para dar suporte à Arquitetura da Informação formulada, foi proposta uma arquitetura de software, materializada por meio da implementação de uma ferramenta denominada *UnB Government Linked Open Data* (UnBGOLD). A Arquitetura construída pode ser visualizada por meio da Figura 4.3.

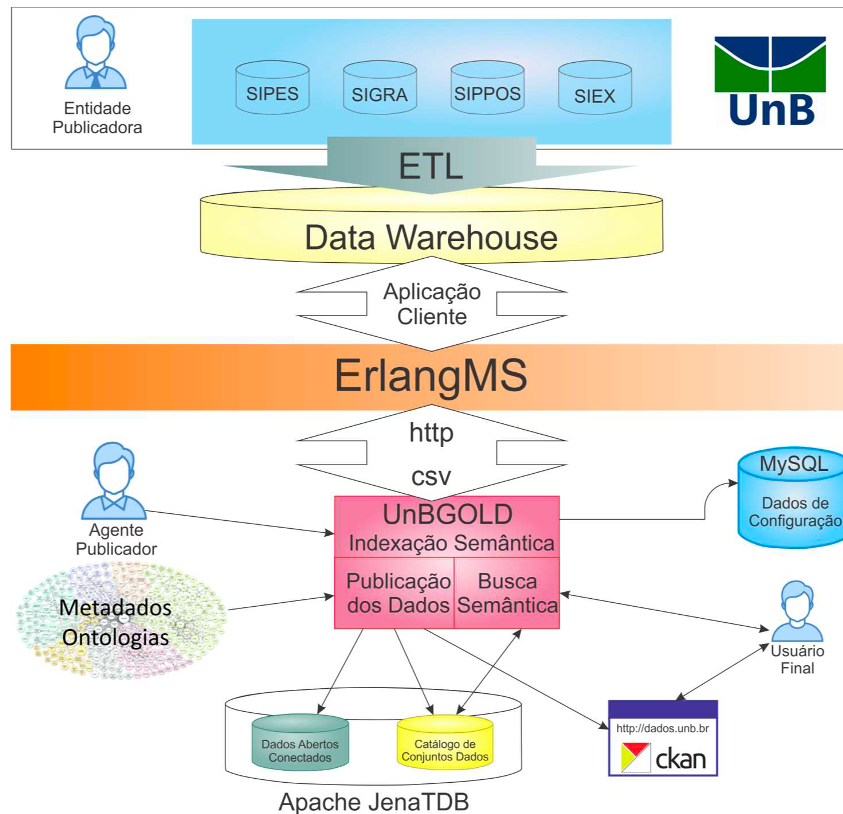


Figura 4.3: Modelo da arquitetura de publicação de dados abertos conectados (Fonte: [54]).

De forma sintática, a partir da parte superior da Figura 4.3, a arquitetura pode ser descrita como segue:

- as Entidades Publicadoras disponibilizam dados abertos para a comunidade. No caso da UnB, os dados são extraídos das bases de dados dos sistemas operativos que informatizam seus processos (SIPES, SIGRA, SIPPOS e SIEX);
- os dados são extraídos, transformados carregados em um *Data Warehouse* (DW) onde ficam disponíveis para o público interno da UnB realizar consultas integradas;
- os dados podem ser exportados em vários formatos por meio de requisição HTTP ao serviço ErlangMS. De forma geral, trata-se de software livre desenvolvido na UnB que extrai os dados do DW e retorna para o solicitante o conjunto de dados em formato já estruturado e aberto (no caso foi utilizado o CSV);
- após a geração dos CSV, os dados são preparados para consumo dos usuários finais por meio da UnBGOLD. A ferramenta é utilizada para configurar os parâmetros de publicação, gerar metadados, definir as ontologias utilizadas para conectar os

conjuntos de dados e gerar as triplas RDF, disponibilizando os dados abertos conectados;

- o Agente Publicador seleciona Metadados e Ontologias que possam descrever os dados. Neste processo, os dados são descritos em formato de triplas RDF e são armazenados em um banco de dados Apache Jena TDB;
- é gerado um banco de dados com os metadados selecionados para descrever as principais características dos conjuntos de dados, o Catálogo de Conjunto de Dados;
- configuram-se os parâmetros que proporcionarão que os dados sejam publicados automaticamente no portal de dados abertos da UnB que é uma instância da plataforma *CKAN*.

Em resumo, a ferramenta UnBGOLD, depois de configurada, proporciona a automação das republicações dos dados, gerencia o Catálogo de Conjunto de Dados e oferece aos usuários finais uma interface para a busca semântica aos dados. Esses usuários podem consumir os dados abertos conectados diretamente no portal dos dados abertos da UnB, onde poderão navegar ou baixar os dados para utilização. Também é disponibilizada uma interface de busca na ferramenta que recupera informações, por meio de consultas SPARQL ao banco de dados de triplas RDF.

Na pesquisa realizada, foram utilizadas abordagens apresentadas nesta Seção, uma vez que a publicação apresentada demonstra a aplicação prática de diversos conceitos relacionados à Arquitetura da Informação, suportada por uma arquitetura de software, onde foram utilizados metadados, ontologias e o *Resource Description Framework*. Tais conceitos são relacionados com o trabalho que foi sendo desenvolvido.

Capítulo 5

Estudo de Caso

Este capítulo tem por objetivo apresentar o estudo de caso da pesquisa realizada. A partir de um maior detalhamento do problema definido no Capítulo 1, será apresentada a proposta de trabalho a ser desenvolvida.

Para atingir o objetivo proposto, a Seção 5.1 tem em seu escopo a apresentação de uma ambientação relativa às operações e ao Comando e Controle no Exército, de forma a tornar compreensível o papel do CCOp Mv e o seu emprego. A Seção 5.2 contempla a apresentação das etapas de desenvolvimento do protótipo de ontologia, segundo uma metodologia e domínio específicos, para emprego na Arquitetura da Informação proposta no trabalho. A Seção 5.3 apresenta o desenvolvimento da Arquitetura da Informação, além da arquitetura de software que a ela fornecerá suporte, contando com um detalhamento do que está sendo proposto.

5.1 O Centro de Coordenação de Operações Móvel

O conceito operativo do Exército Brasileiro preconiza a máxima integração entre vetores militares e civis que buscam a unidade de esforços no ambiente interagências. Na concepção doutrinária para o emprego da Força Terrestre, são definidas as operações para casos de guerra, sejam ofensivas e defensivas, bem como operações para casos de não guerra, destacando-se operações de cooperação e coordenação com as diversas agências civis e militares, governamentais e não governamentais.

Conforme apresentado no Capítulo 1, este trabalho tem foco no suporte a operações de não guerra. Em operações de coordenação e cooperação com agências, há uma maior dificuldade para o provimento da interoperabilidade entre os diversos órgãos envolvidos. Isso se deve, principalmente, à diversidade de informações e de meios tecnológicos, podendo tornar custoso o processo decisório. Com isso, o princípio da interoperabilidade

deve ser priorizado no planejamento, uma vez que é latente, no combate moderno, esforços conjuntos demandados.

Salienta-se que em operações militares, a capacidade de gerenciamento dos meios de comando e controle é considerada um dos pilares para a obtenção do sucesso. Essa capacidade é decorrente de eficiência e integração horizontal e vertical dos procedimentos, meios tecnológicos e decisões, garantindo a consciência situacional constantemente atualizada para todos os comandantes envolvidos em determinada operação [3].

Nesse contexto, buscando obter uma consciência situacional integrada, o Exército implementou os Centros de Coordenação de Operações fixos, de forma a proporcionar a tomada de decisão dos escalões superiores por meio de informações oportunas. Complementando as funcionalidades desses Centros, de forma mais flexível, foi concebido o Projeto Centro de Coordenação de Operações Móvel.

Conforme abordado anteriormente, o CCOp Mv foi concebido com a finalidade de ser um Centro de Comando e Controle configurado para apoiar, com recursos disponíveis, um Grande Comando Operacional em situações de guerra e não guerra. As principais missões do CCOp Mv são [84]:

- proporcionar as funcionalidades necessárias de Tecnologia de Informação e Comunicações (TIC), essenciais ao desenvolvimento do Comando e Controle e do apoio à decisão do escalão considerado, por intermédio de um conjunto de meios, processos e serviços;
- contribuir para a obtenção da Superioridade de Informação e para a formação da consciência situacional pretendida em operações, preconizadas na Doutrina Militar de C2;
- disponibilizar informações oportunas, precisas e relevantes aos tomadores de decisão, civis ou militares, governamentais ou não governamentais, nacionais ou multinacionais, no nível Grande Comando; e
- acessar, processar e disseminar um ininterrupto fluxo de informação baseado no conceito de flexibilidade.

Com o avanço tecnológico, uma tendência natural foi a digitalização do campo de batalha, aprimorando o planejamento e o acompanhamento das operações militares. Assim, foram realizados investimentos para o desenvolvimento de *softwares* de apoio à decisão, visando a sincronização de ações e o tratamento de incidentes, provendo uma melhor consciência situacional das operações.

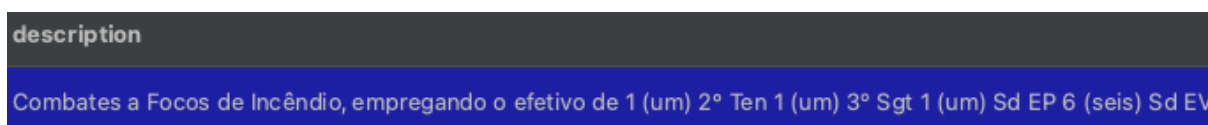
Atualmente, um dos principais sistemas de apoio à decisão no Exército é o Pacificador, concebido para operações de não guerra. Seu emprego foi implementado inicialmente

em 2011, durante os 5º Jogos Mundiais Militares. Desde então, o *software* vem sendo empregado em grandes eventos e operações de GLO, tendo viés relacionado às operações de não guerra ou de coordenação com agências.

Uma das ferramentas existentes no Pacificador é a de relato de incidentes. Esses incidentes são identificados como eventos não planejados com potencial para comprometer a segurança de eventos, devendo receber um tratamento adequado. Eles estão presentes no *software* como elementos de controle, surgindo à medida que os problemas são identificados, sendo necessário o tratamento adequado pelos Centros de Coordenação de Operações, sejam fixos ou, futuramente, móveis.

As informações registradas pela ferramenta de relato de incidentes, do sistema Pacificador, são armazenadas em um documento JSON, com uma estrutura que contempla diversos campos de dados. Após a análise desses dados, foi percebido que não é contemplada uma categorização para os incidentes, bem como o campo ‘descrição’, que tem por objetivo detalhar o incidente relatado, é inserido pelo usuário como texto livre, em linguagem natural, sem uma estruturação padronizada. Cabe destacar que em versões anteriores do *software*, havia uma categorização dos incidentes, a ser preenchida no momento do relato observado. Porém, a funcionalidade deixou de ser utilizada, por não atender de forma plena as necessidades dos usuários.

De forma a exemplificar uma informação não estruturada, extraída de um registro de incidente do *software* relatado por usuário, um dado do campo ‘descrição’ pode ser visualizado na Figura 5.1. A informação não é explicitamente precisa, principalmente para uma recuperação efetiva, onde seria dificultada a aplicação de um filtro para a obtenção de informações específicas.



The image shows a screenshot of a JSON object. The key 'description' is highlighted in a dark grey box. The value is a string: 'Combates a Focos de Incêndio, empregando o efetivo de 1 (um) 2º Ten 1 (um) 3º Sgt 1 (um) Sd EP 6 (seis) Sd EV'. The text is displayed on a blue background.

Figura 5.1: Descrição de incidente extraída do Sistema Pacificador.

Para o desenvolvimento do estudo de caso, esta pesquisa visa propor uma possível solução a um problema real que foi apresentado, uma vez que informações estruturadas proporcionarão um processo decisório mais adequado por parte do Comando de uma operação. O que se busca é fornecer uma Arquitetura voltada à obtenção assistida e para a disponibilização de dados, com fundamentos baseados na organização e estruturação da informação, empregando técnicas que serão apresentadas no decorrer deste Capítulo. Por fim, é importante evidenciar que a proposta apresentada contempla naturalmente a aplicabilidade junto aos CCOp fixos ou móveis.

5.2 Desenvolvimento de protótipo de ontologia

Com a proposta de construir uma Arquitetura da Informação suportada por um modelo ontológico, foram buscadas ontologias disponíveis que pudessem atender ao domínio específico relacionado. Não foram encontradas ontologias com uma cobertura ideal para os principais termos buscados, o que tornou plausível a construção de um protótipo de ontologia, de forma a atender ao que se pretendia alcançar neste trabalho.

Como mencionado anteriormente, não foi buscada a construção de uma ontologia completa e complexa, por ser julgado inviável no escopo desta pesquisa. A proposta é voltada ao desenvolvimento de um protótipo de ontologia adequado para utilização em proveito da Arquitetura da Informação construída. Este protótipo foi intitulado ‘OntoEB’.

Apesar de não fazer parte do escopo deste trabalho, sendo uma indicação para oportunidades futuras, um dos benefícios do emprego de ontologias em uma Arquitetura da Informação, que representa vantagens em relação ao emprego de outras tecnologias, é a possibilidade de trabalhar com regras de inferência para melhoria dos resultados e das análises a serem realizadas. No exemplo apresentado a seguir é demonstrada uma regra de inferência, aplicável em um possível caso de uso. Um dos resultados poderia ser a formação de uma rede de relacionamentos entre organizações militares, a partir de inferências:

SE		
Região Militar	apoia	Brigada
Companhia	pertence_a	Brigada
ENTÃO		
Companhia	eh_apoiada_por	Região Militar
SABENDO QUE		
11ª RM	eh	Região Militar
3ª Bda Inf Mtz	eh	Brigada
6ª Cia Com	eh	Companhia
11ª RM	apoia	3ª Bda Inf Mtz
6ª Cia Com	pertence_a	3ª Bda Inf Mtz
ENTÃO		
6ª Cia Com	eh_apoiada_por	11ª RM

Por fim, destaca-se que o protótipo de ontologia foi construído em língua portuguesa, utilizando o *software Protégé*, um editor de ontologias gratuito e de código aberto. Além disso, seguiu padrões pautados em boas práticas, em busca de um resultado adequado à proposta. O guia para desenvolvimento de ontologias de Noy e McGuinness [48], descrito no Capítulo 4, foi o parâmetro utilizado para tal. Nesse contexto, a proposta de cons-

trução do protótipo de ontologia desta pesquisa, seguiu as etapas descritas nas próximas subseções.

5.2.1 Determinação do domínio e do escopo da ontologia

O foco desta pesquisa está voltado à eventos militares observados em campo. Foi verificada a necessidade de entendimento de um domínio específico, voltado à operações militares, que configuram a área de maior atuação dos Centros de Coordenação de Operações.

Sendo assim, o protótipo de ontologia tem como domínio as operações militares de não guerra das quais o Exército Brasileiro participa. O foco está na observação de eventos de campo, sendo necessária a descrição de um incidente específico, o seu local de ocorrência e os seus participantes. O protótipo foi utilizado para dar suporte à Arquitetura a Informação desenvolvido, tendo como premissa cobrir os principais termos a serem utilizados na busca pelos usuários, durante a entrada de dados. Com isso, foi contemplada a obtenção de dados assistida, visando proporcionar a estruturação de relatórios com informações dos eventos observados.

5.2.2 Reutilização de ontologias existentes

Existem milhares de ontologias empregadas na *World Wide Web*, compostas por diversos conceitos e relacionamentos entre eles. Algumas dessas ontologias possuem um domínio específico de aplicação, enquanto outras são mais primitivas, de domínio geral, conhecidas como ontologias superiores (*upper ontologies*), possuindo terminologias relevantes para diversas áreas do conhecimento. As ontologias superiores estão se tornando uma tecnologia-chave voltadas para a integração de conhecimentos heterogêneos oriundos de fontes diversas [85]. Cabe ressaltar que esse tipo de ontologia é um importante ponto de partida para criação de ontologias específicas.

Após o entendimento do domínio e do escopo de aplicação do modelo ontológico, foram buscadas ontologias superiores com potencial para reaproveitamento, com classes, subclasses e propriedades úteis para o desenvolvimento do protótipo de ontologia. As ontologias que mais se destacaram na busca realizada foram a PROTON e a COSMO. Ambas ontologias possuem versões disponíveis para download por meio da base de dados aberta da DBpedia [86].

A PROTON fornece considerável cobertura geral de conceitos necessários para uma ampla gama de tarefas, incluindo a anotação semântica, indexação e recuperação de documentos. Possui como principais características a independência de domínio, definições lógicas leves, alinhamento com padrões populares, bem como boa cobertura de entidades nomeadas e de domínios concretos (ou seja, pessoas, organizações, locais, números, datas,

endereços). Algumas classes e subclasses da ontologia PROTON, relacionadas com locais de interesse, podem ser visualizadas na Figura 5.2.



Figura 5.2: Extrato de classes da ontologia PROTON - *Software Protégé*.

A ontologia COSMO é formada pela combinação de uma rede de ontologias, resultado de um trabalho colaborativo que tem por objetivo representar elementos básicos. Com a aplicação da ontologia é facilitada a criação de elementos de uma ontologia de domínio específico, com a combinação dos elementos primitivos. Ela é composta por um conjunto abrangente de conceitos (classes, relações, funções, instâncias). Foram destacadas na Figura 5.3 algumas classes, que discriminam unidades militares, da ontologia COSMO.



Figura 5.3: Extrato de classes da ontologia COSMO - *Software Protégé*.

Após o levantamento de ontologias superiores com potencial para reaproveitamento, foi realizada a busca por ontologias de domínio específico, que fossem compostas por um vocabulário militar. Foram então encontradas informações sobre o *Multilateral Interoperability Programme* (MIP) [87], um consórcio composto principalmente por países da OTAN, que desenvolve especificações técnicas para Sistemas de Informação de Comando e Controle, em apoio a operações militares. O órgão possui uma abordagem para estruturação informações complexas, com linhas de base desenvolvidas ao longo dos anos.

Em uma das linhas de base do MIP, iniciada em 2006, foi construído o *Joint Consultation, Command and Control Information Exchange Data Model* (JC3IEDM) [87], um modelo de dados destinado à troca de informações sobre comando, controle e comunicações. O modelo contempla a descrição de conceitos, baseada em XML, de todas as suas entidades, atributos e relacionamentos. Esse modelo vem sendo aperfeiçoado e até os dias atuais.

Matheus e Ulicny [88] desenvolveram um trabalho que, a partir do vocabulário do modelo JC3IEDM, foi realizada a criação de uma ontologia, com um considerável quantitativo de classes e propriedades. A Figura 5.4 apresenta algumas classes propostas pela ontologia, que discriminam ‘locais’.



Figura 5.4: Extrato de classes da ontologia JC3IEDM.

Em síntese, nesta subseção foram destacadas ontologias superiores, bem como uma ontologia de domínio específico, as quais foram oportunas para o reaproveitamento neste trabalho.

5.2.3 Levantamento dos principais termos da ontologia

A Extração de Informações é uma técnica empregada para obter informações estruturadas de textos não estruturados. Com isso, torna-se possível, a partir de um texto bruto, extrair informações relevantes.

O *Named Entity Recognition* (NER) é uma subtarefa da Extração de Informações, que tem por objetivo identificar entidades em um texto não estruturado e atribuí-las a uma lista de entidades predefinidas. A lista de entidades pode ser padrão ou particular, se treinarmos nosso próprio modelo linguístico para um conjunto de dados específico. As entidades identificadas podem ser, como exemplo: Pessoas, Eventos, Organizações, Locais, etc.

Um grafo de conhecimento é uma forma de apresentar os dados que resultam da Extração de Informações. Uma das formas de representar esses grafos é por meio de triplas, compostas por um sujeito, um predicado e um por objeto, seguindo o conceito semelhante ao RDF.

Os principais termos empregados no desenvolvimento do protótipo de ontologia deste trabalho estão relacionados às operações militares de não guerra. Para um melhor levantamento destes termos, foi realizada a análise da base de dados do sistema Pacificador, mais especificamente no que diz respeito aos registros das descrições dos eventos, cadastradas pelos usuários.

Para realizar a análise da base de dados do Pacificador, foi empregada a a Extração de Informações, empregando o *Named Entity Recognition* (NER), técnica considerada adequada para o levantamento das expressões relevantes a serem cobertas pelo protótipo. A partir dos dados extraídos, foi possível construir grafos de conhecimento simplificados, com os principais termos que descrevem as informações obtidas por meio da base de dados analisada.

Cabe destacar que neste trabalho foi efetivada uma adaptação da técnica descrita, em busca de resultados adequados. O objetivo não foi diretamente extrair conhecimentos da base de dados, mas buscar os termos mais adequados para construção de um protótipo de ontologia, de forma a atender ao domínio específico do estudo que foi realizado.

Para a materialização da técnica descrita, foi desenvolvida uma aplicação que possibilita a obtenção de termos e conceitos de textos não estruturados. O modelo foi construído utilizando a biblioteca *spaCy* (v3.0.3)¹, escrita em *Python* (v3.8)², utilizada para processamento avançado de linguagem natural. É amplamente utilizada ao fim que se destina, por possuir boa velocidade de processamento, adequada usabilidade, documentação completa disponível, bem como suporte à linguagem PT-BR. Para a construção do grafo, em

¹<https://spacy.io>

²<https://www.python.org>

formato de triplas, foi utilizada a biblioteca *NetworkX* (v2.5)³, bem como para a plotagem dos grafos de conhecimento foi utilizada a biblioteca *Matplotlib* (v3.3.4)⁴.

Como exemplo de um teste inicial, foi utilizando como entrada da aplicação o texto:

```
text = "Três bandidos assaltaram uma loja no bairro do Méier. " \
      "A Patrulha prendeu dois assaltantes em flagrante. " \
      "A Patrulha apreendeu ainda duas pistolas e drogas. "
```

Como resultado, um grafo de conhecimento simplificado foi construído, apresentado na Figura 5.5.

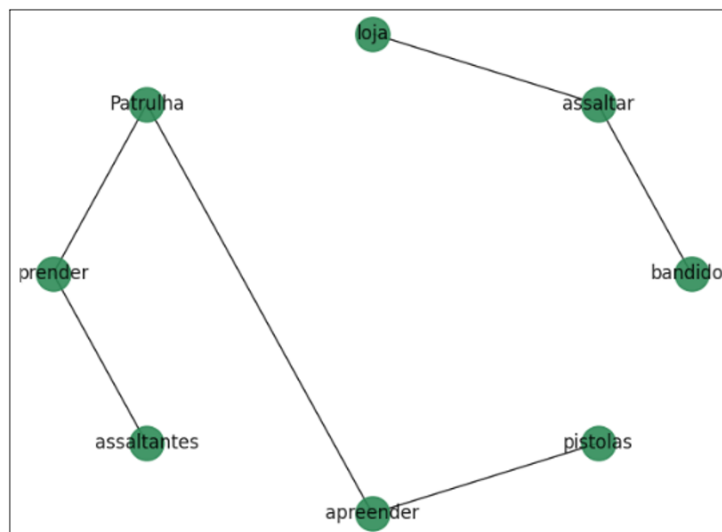


Figura 5.5: Grafo de Conhecimento simplificado.

A partir do grafo exposto, poderão ser extraídas triplas, conforme o que segue:

```
Patrulha prender assaltantes
Patrulha apreender pistolas
bandido assaltar loja
```

Os diversos termos que compõe as triplas obtidas podem direcionar a composição dos elementos que envolverão um evento observado:

- Natureza do Evento - Assalto, Apreensão, Prisão
- Local do Evento - Loja
- Atores do Evento - Patrulha, Bandido, Assaltante

³<https://networkx.org>

⁴<https://matplotlib.org>

É importante ressaltar que o resultado obtido por meio da aplicação desenvolvida, utilizando a base de dados do sistema Pacificador, está sendo apresentado no Capítulo 6.

Ainda nesta etapa de levantamento dos principais termos da ontologia, foi realizada a análise do Manual de Abreviaturas, Siglas, Símbolos e Convenções Cartográficas das Forças Armadas (MD33-M-02) [89], publicado pelo Ministério da Defesa. O Manual possui um vasto vocabulário que cobre centenas de siglas e seus respectivos significados, que são utilizados no âmbito do Exército Brasileiro. A partir da obtenção deste vocabulário, foi possível enriquecer o protótipo de ontologia com termos adequados, alinhados ao domínio estudado. Um pequeno extrato de termos descritos no Manual foram exemplificados por meio da Figura 5.6.

Operação de Paz	Op Paz
Operação Defensiva	Op Def
Operação Ofensiva	Op Ofs
Operação sob Condições Especiais	Op Cnd Esp
Operação Terrestre	Op Ter
Operações Aéreas Militares	OAM
Operações Cibernéticas	Op Ciber
Operações Complementares	Op Cmpl
Operações Conjuntas	Op Cj
Operações de Apoio	Op Ap

Figura 5.6: Extrato de termos - MD33-M-02.

Com isso, complementando o reaproveitamento de ontologias existentes, realizado na etapa anterior, nessa fase foi possível realizar o levantamento de termos considerados pertinentes para o protótipo de ontologia.

5.2.4 Definição de classes e a hierarquia de classes

A partir da definição de termos e conceitos consonantes, que contemplam o vocabulário geral utilizado na construção do protótipo de ontologia, foram criadas as classes, as subclasses e as suas respectivas hierarquias. Para essa definição, foi utilizada uma abordagem *top-down*, que tem como passo inicial a definição de conceitos mais gerais a serem descritos, contemplando as classes da ontologia, seguida de uma posterior especialização desses conceitos, que formatam as diversas subclasses criadas.

5.2.5 Definição das propriedades das classes

Após a definição das classes e subclasses e do estabelecimento de suas hierarquias, foram descritas propriedades que fornecem o detalhamento adequado de cada uma dessas classes. A partir dos resultados obtidos da análise do domínio específico, da reutilização

de ontologias existentes, bem como do levantamento dos principais termos da ontologia, com a utilização da técnica de Extração de Informações, não só as classes foram explicitadas, como também foi facilitado o entendimento das propriedades dos termos levantados, proporcionando robustez ao protótipo de ontologia.

5.2.6 Definição do detalhamento das propriedades das classes

Com a definição de quais seriam as propriedades das classes da ontologia, foram definidos detalhamentos essenciais dessas propriedades, buscando identificar os relacionamentos adequados entre as propriedades e as classes e subclasses do protótipo. Essas propriedades podem assumir valores literais, valores fixos, bem como apontar para o relacionamento entre diferentes classes e subclasses. A partir dos conhecimentos obtidos, foram definidos os domínios (*domains*) e os alcances (*ranges*), que proporcionam o relacionamento adequado entre as classes, subclasses e propriedades do protótipo de ontologia.

5.2.7 Criação de instâncias

Por fim, esta etapa visa a criação de instâncias individuais com base nas classes/subclasses definidas no protótipo de ontologia. Esse passo foi realizado com a aplicação direta do protótipo de ontologia em suporte à Arquitetura da Informação. Por meio da arquitetura de software concebida, foi possível criar instâncias para a ontologia, de forma que um determinado conjunto de instâncias são consolidadas como as partes componentes do relatório um evento observado. Ou seja, como resultado, essas instâncias criadas após a observação de um evento são exibidas em conjunto, em um relatório estruturado.

5.2.8 Validação do protótipo de ontologia

Complementando as etapas estabelecidas por Noy e McGuinness [48] para criação de ontologias, foi realizada a validação do que foi desenvolvido neste trabalho, mesmo com a proposta aqui apresentada ser a de construir um protótipo de ontologia.

O *software Protégé*, utilizado para construção do protótipo, possui um mecanismo de inferência (*reasoner*) conhecido por HermiT, que permite a realização de validações na ontologia, conforme disposto na Figura 5.7.

Podem ser destacados relevantes serviços providos pelo citado mecanismo de inferência [90]:

- verificação de relacionamentos entre classes e subclasses - é possível realizar testes para identificar se uma classe seria ou não uma subclasse de outra classe, ou seja, as

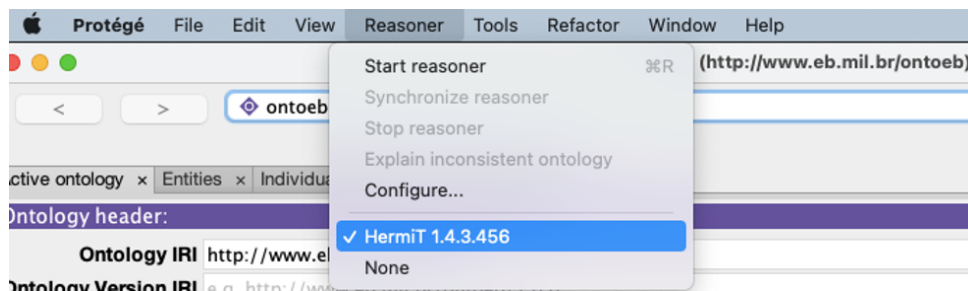


Figura 5.7: Mecanismo de Inferência - *Protégé*.

suas descrições das classes são utilizadas para verifica se há uma relação superclasse/subclasse entre elas. Com isso, é possível inferir automaticamente a hierarquia de classes da ontologia;

- verificação de inconsistências - a partir das propriedades descritivas de uma classe, o mecanismo de inferência tem a capacidade de verificar a possibilidade de uma classe possuir instâncias. Se não for possível que uma classe possua instâncias, ela é considerada inconsistente.

Nesse contexto, é pertinente destacar que foram realizadas as validações fornecidas pelo *software Protégé*, de forma a apoiar a construção do protótipo de ontologia de forma consistente e adequada ao que está sendo proposto.

5.3 Proposta de Arquitetura da Informação

O escopo desta pesquisa está direcionado à administração dos dados obtidos por agentes observadores de campo, visando proporcionar informações seguras e confiáveis, adequadas às necessidades demandadas. A proposta foi a de organizar as informações obtidas, empregando recursos diversos, buscando uma gestão efetiva dessas informações. O gerenciamento dessas informações tem por objetivo produzir conhecimentos essenciais para a efetiva tomada de decisão.

5.3.1 Recursos empregados

De forma a materializar o que foi apresentado, permeando a administração dos dados obtidos, bem como a gestão da informação, foi proposta uma Arquitetura da Informação, composta por recursos diversos como metadados, taxonomia, tesouros, ontologia e usabilidade. Esses recursos, presentes na arquitetura desenvolvida, podem ser descritos como segue:

- os metadados serão empregados para descrever o conteúdo, servindo diretamente para a recuperação da informação;
- as taxonomias foram utilizadas para permitir a navegação pelo conteúdo disponibilizado, facilitando o entendimento das informações disponibilizadas, bem como proporcionando suporte para buscas por informações;
- os tesouros permitiram localizar termos que representem significados específicos buscados pelo usuário;
- as ontologias permitem a indexação dos dados coletados e o aprimoramento das buscas realizadas pelos usuários; e
- o usuário irá interagir com interfaces implementadas com um direcionamento voltado à usabilidade.

5.3.2 Concepção geral

A Arquitetura da Informação proposta busca a representação semântica da informação, em uma estrutura compreensível, bem como a organização de seu armazenamento, de forma a otimizar a recuperação dessas informações. A estrutura simplificada do que foi proposto pode ser visualizada na Figura 5.8.

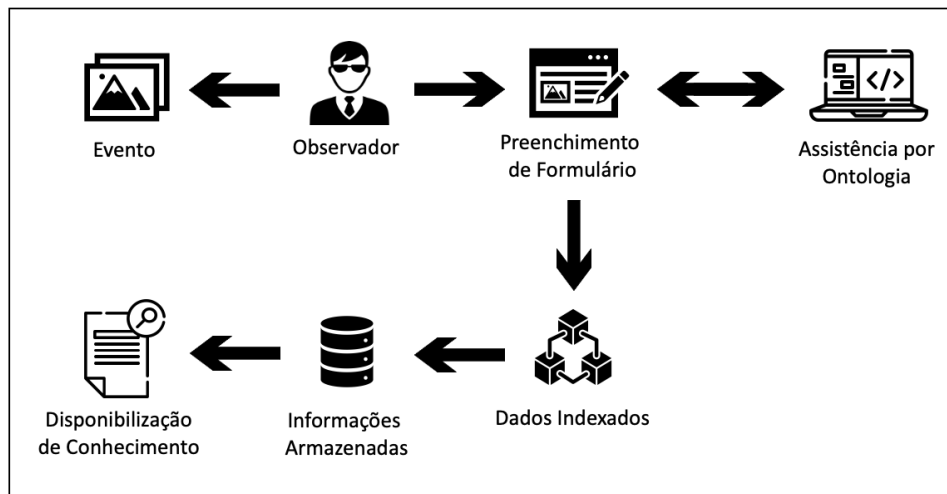


Figura 5.8: Estrutura simplificada da Arquitetura da Informação.

Dentro do proposto como estrutura da Arquitetura da Informação, o exposto na Figura 5.8 pode ser descrito de acordo com o que segue:

- um observador de campo realiza a identificação de determinado evento que necessita ser registrado;

- o observador de campo utilizará o formulário disponibilizado, concebido em um formato assistido por ontologia;
- os dados inseridos por meio do formulário são indexados, utilizando a ontologia;
- os dados indexados passarão por um tratamento, tendo como resultado informações a serem armazenadas em uma estrutura de relatório;
- será disponibilizada a recuperação de relatórios que consolidam as informações geradas, para fornecer melhores condições para a tomada de decisão.

5.3.3 Proposta de Arquitetura de *Software*

Para dar o suporte necessário à Arquitetura da Informação, foi proposta uma arquitetura de software. Para tal, foi desenvolvido um sistema de informação, que tem como objetivo assistir a coleta dos dados, a organização das informações obtidas, a sua representação semântica, bem como proporcionar o armazenamento e a recuperação dos conhecimentos a serem gerados.

A proposta de arquitetura de software foi construída utilizando a linguagem de programação *Python* (v3.8)⁵, empregando o *framework Flask* (v1.1.4)⁶, que provê um modelo de desenvolvimento para web utilizando a linguagem *Python*. O *framework* possui um núcleo base, que permite que o projeto a ser desenvolvido funcione apenas com os recursos necessários à sua execução, sendo porém extensível, o que possibilita a adição de novas bibliotecas e funcionalidades, conforme a necessidade. Isso torna o desenvolvimento simplificado e direcionado às funcionalidades buscadas. Sendo assim, como a interface do usuário é baseada no preenchimento e no envio de formulários, na lógica da exibição desses formulários e dos registros a serem realizados, foi utilizada a biblioteca *RDFLib* (v6.0.0)⁷, de forma a permitir o processamento e a indexação dos dados. A biblioteca possibilita a execução das seguintes funcionalidades principais:

- o carregamento de ontologias através de arquivos ou da web;
- a criação de novas triplas semânticas;
- a persistência da ontologia através de arquivo ou banco de dados;
- a execução de consultas SPARQL.

Dentro dessa arquitetura de *software*, para a estruturação do relatório do evento observado, foi concebida a construção de uma estrutura inicial fixa, que padroniza a entrada

⁵<https://www.python.org>

⁶<https://flask.palletsprojects.com>

⁷<https://rdflib.dev>

de dados. Assim, o formulário de entrada de dados, a ser preenchido pelo observador do evento, é composto por três subdivisões principais, baseadas nos seguintes elementos: a natureza do evento, os atores envolvidos com o evento e o local de ocorrência do evento. Essa subdivisão padronizada tem por finalidade facilitar a indexação dos dados, permitindo a formação de três triplas semânticas principais para determinado evento observado, sendo o sujeito o evento a ser instanciado e o objeto cada um dos elementos descritos:

- Evento possui Natureza
- Evento possui Ator
- Evento possui Local

Na concepção da arquitetura de software proposta, para cada um dos elementos descritos (natureza, atores e local), será seguido um processo de execução de atividades, materializado por meio da Figura 5.9.

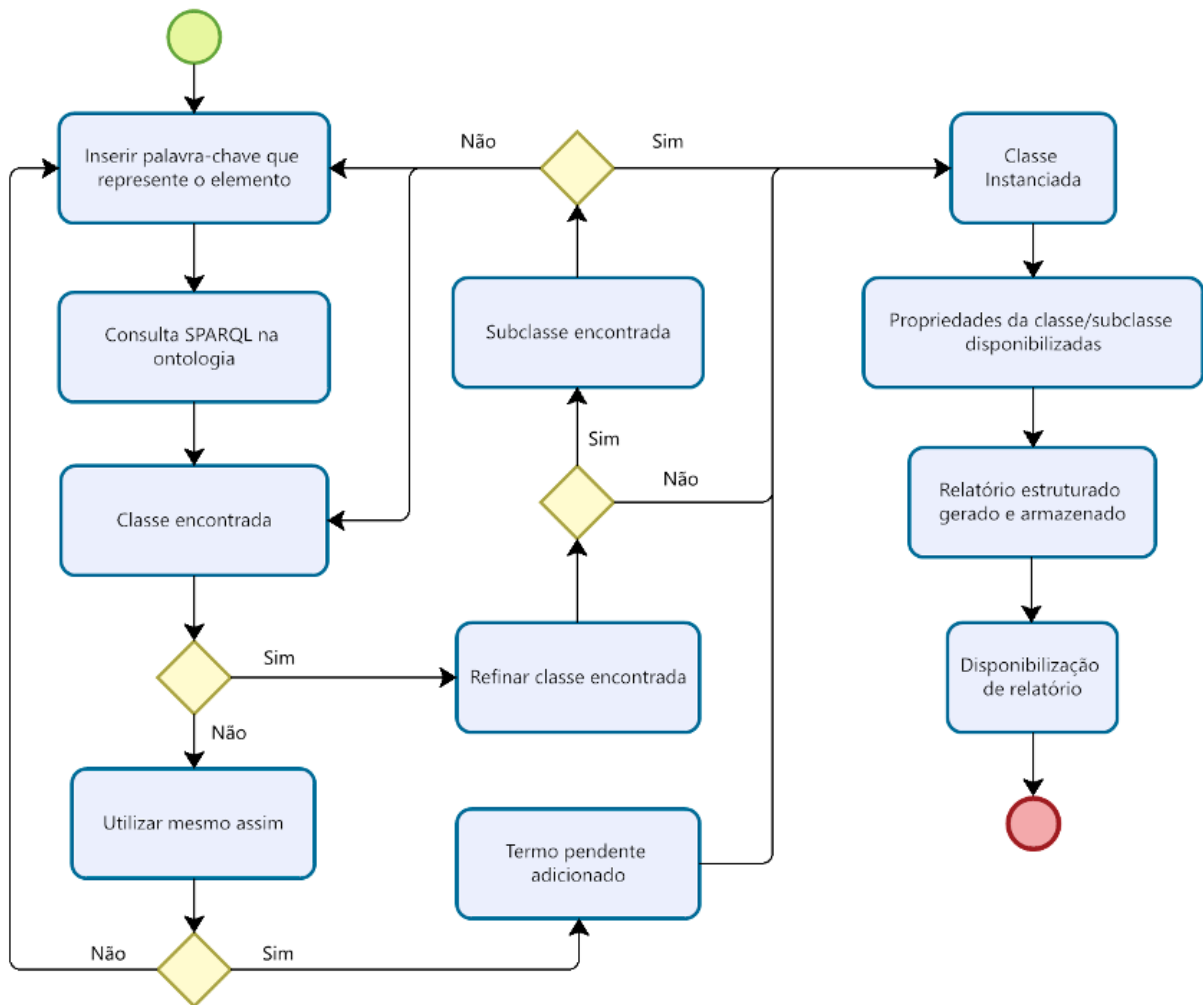


Figura 5.9: Processo de Execução da Arquitetura de Software.

Sendo assim, para cada elemento será buscada uma classe (recurso *owl:Class*) da ontologia empregada que melhor o represente. A busca por essa classe é iniciada por uma palavra-chave inserida pelo observador de campo, que resultará em uma *query* SPARQL na ontologia. Essa palavra-chave deverá sintetizar, em um único termo, a definição do elemento do evento. De forma a tornar a consulta mais amigável ao usuário, as *query* são realizadas sobre os *labels* e os *comments* das classes, recursos que rotulam e descrevem os termos da ontologia em linguagem natural, utilizando-se para isso métodos de busca com expressões regulares (*regular expressions - regex*).

Após o resultado da busca, caso o sistema não encontre a classe da ontologia relacionada, será fornecida ao usuário a opção de utilizar a própria palavra-chave utilizada como entrada. Se o usuário não optar pelo uso do termo, será direcionado para inserir uma nova palavra-chave, porém se aceitar a utilização, poderá fazê-lo naturalmente, sendo prosseguido o fluxo do sistema, considerando o termo como uma classe instanciada. De forma complementar, o termo utilizado será adicionado a um repositório de termos pendentes. Estes termos poderão ser analisados por um especialista em ontologias, capaz de verificar a pertinência de inserção do novo termo na estrutura da ontologia.

Retornando à busca inicial, caso o usuário identifique a classe de interesse, realizará a sua seleção. Após a seleção da classe inicial, o usuário poderá realizar o seu refinamento, caso julgue necessário. Para isso, o sistema buscará as subclasses da classe selecionada (recurso *rdfs:subClassOf*), de forma que sejam disponibilizadas ao usuário, para seleção. Se a subclasse não for julgada adequada, o usuário terá a opção de continuar utilizando a classe inicial ou poderá iniciar uma nova busca, inserindo outra palavra-chave que descreva o elemento.

Com a classe/subclasse encontrada, esta será instanciada, efetivando o relacionamento com o elemento que compõe o evento (natureza, ator ou local). Com a classe da ontologia instanciada, o formulário exibirá então as suas propriedades (recurso *owl:DatatypeProperty*), permitindo ao usuário o preenchimento dos dados necessários. Essas propriedades dos elementos do evento são as informações que descrevem em detalhes a classe selecionada.

Ao final do processo, o sistema registra o novo evento. Os três elementos integrantes da ocorrência (natureza, atores e local) serão detalhados por meio das classes/subclasses instanciadas, bem como a partir das descrições das respectivas propriedades. As informações geradas são armazenadas, estruturadas em relatórios, sendo disponibilizadas para recuperação, o que viabiliza consultas por usuários e autoridades.

De forma geral, a partir de uma palavra-chave de entrada é realizada uma consulta SPARQL a partir dos *labels* e *comments* das classes da ontologia. Com isso, serão retornadas as principais classes que refletem a busca. O Código 5.1 apresenta a consulta construída com a finalidade de buscar classes a partir de termos especificados pelo usuário.

```

1 def busca_classes(palavra_chave):
2     query_str = (
3         """
4         SELECT DISTINCT ?class ?label ?comment ?result
5         WHERE {
6         ?class a owl:Class.
7         {?class rdfs:label ?label.
8         FILTER(
9             regex(?label,
10                """
11                + "\"" + palavra_chave + "\""
12                +
13                """
14                , "i")
15            )}
16         UNION
17         {?class rdfs:comment ?comment.
18         FILTER(
19             regex(?comment,
20                """
21                + "\"" + palavra_chave + "\""
22                +
23                """
24                , "i")
25            )}
26         ?class rdfs:label ?result
27         }
28         """
29     )
30     qres = OntologiaSup.onto.query(query_str)
31     classes = []
32     for row in qres:
33         classes.append(str(row.asdict()['result']))
34     classes = list(dict.fromkeys(classes))
35     classes = OntologiaSup.ordena(classes, palavra_chave)
36
37     return classes

```

Código 5.1: Consulta SPARQL para classes da ontologia

Como será possível realizar o refinamento dessa busca por meio da obtenção também de subclasses das classes da ontologia, a partir de uma classe inicial anteriormente encontrada, foi proposta uma consulta SPARQL para dar suporte a esta finalidade, conforme disposto no Código 5.2.

```
1 def busca_subclasses(classe):
2     query_str = (
3         """
4         SELECT DISTINCT ?class ?label ?subclass ?sublabel
5         WHERE {
6         ?class a owl:Class.
7         ?class rdfs:label
8         """
9         + "\"" + classe + "\"" + OntologiaSup.onto_lang + "."
10        +
11        """
12        ?subclass rdfs:subClassOf ?class.
13        ?subclass rdfs:label ?sublabel.
14        }
15        """
16    )
17    qres = OntologiaSup.onto.query(query_str)
18    subclasses = []
19    for row in qres:
20        subclasses.append(str(row.asdict()['sublabel']))
21    return subclasses
```

Código 5.2: Consulta SPARQL para subclasses da ontologia

Além do desenvolvimento das consultas SPARQL para busca de classes e subclasses, foi construída uma *query* específica para buscar as propriedades da classe/subclasse instanciada, conforme destacado no Código 5.3

```
1 def busca_propriedades(classe):
2     query_str = (
3         """
4         SELECT DISTINCT ?proplabel
```

```

5         WHERE {
6             ?class a owl:Class.
7             ?class rdfs:label
8             """
9             + "\"" + classe + "\"" + OntologiaSup.onto_lang + "."
10            +
11            """
12            ?properties rdf:type owl:DatatypeProperty .
13            ?properties rdfs:domain ?class .
14            ?properties rdfs:label ?proplabel .
15        }
16        """
17    )
18    qres = OntologiaSup.onto.query(query_str)
19    labels = []
20    for row in qres:
21        labels.append(str(row.asdict()['proplabel']))
22    return labels

```

Código 5.3: Consulta SPARQL para propriedades das classes/subclasses da ontologia

Dando continuidade ao desenvolvimento, foram envidados esforços para a autenticação de usuários ao *software*, sendo considerada um parâmetro essencial para a segurança de um sistema de informação. Para a abordagem, foi utilizada a biblioteca *flask-login*, presente no próprio *framework flask*.

Para tornar o sistema de autenticação mais específico e adequado às necessidades, foi implementado, de forma complementar, o controle de acesso dos usuários por meio de perfis. Com isso foi possível restringir as funcionalidades do sistema aos tipos de usuários específicos, conforme o que segue:

- Administrador - possui acesso garantido à todas as funcionalidades do sistema, sem restrições;
- Usuário - possui apenas permissão para registrar novos eventos observados, bem como editar ou excluir os seus próprios eventos, além de visualizar eventos de outros usuários. Não possui permissão para editar ou excluir eventos de outros usuários;
- Autoridade - possui apenas acesso para visualização de todos os eventos registrados, possibilitando a análise e apoiando o processo de tomada de decisão;

- Especialista em Ontologia - possui apenas a permissão para visualizar todos os eventos registrados, bem como acesso ao repositório de termos pendentes, de forma a possibilitar a análise e a decisão de inserção do novo termo na ontologia.

Após a definição das necessidades do *software* desenvolvido e a realização das atividades descritas anteriormente, foi realizada a construção da interface do usuário, apresentada no Capítulo seguinte, onde são destacados os resultados deste trabalho. Com o objetivo de alcançar uma melhor usabilidade, foi buscada a construção de uma interface simples e objetiva, de forma a proporcionar agilidade, facilidade e oportunidade para o registro de uma nova ocorrência observada pelo usuário. Adicionalmente, um dos parâmetros base foi atender aos requisitos de responsividade, proporcionando a adequação da interface aos variados tamanhos de tela dos dispositivos existentes. O objetivo é garantir mobilidade ao sistema, uma vez que a coleta dos dados em campo poderá ser realizada a partir de equipamentos diversos.

Dentre as funcionalidades buscadas para o sistema, de forma a permitir uma arquitetura flexível e configurável, não há uma restrição para a utilização apenas do protótipo de ontologia que foi desenvolvido com o propósito específico para este trabalho. É possível que sejam empregadas outras ontologias no *software* desenvolvido, caso seja uma necessidade. Podem ser buscadas ou desenvolvidas outras ontologias com termos e conceitos considerados adequados para utilização.

Em síntese, a Arquitetura da Informação proposta tem como produto final viabilizar a construção de relatórios estruturados, a partir de dados coletados por agentes de campo, por meio de um sistema de informação suportado por ontologia. Sendo assim, os conhecimentos a serem armazenados poderão ser recuperados pelas autoridades, visando proporcionar uma maior consciência situacional e otimizar, com mais assertividade, o processo de tomada de decisão.

Capítulo 6

Resultados

Este Capítulo tem por finalidade apresentar os resultados obtidos no decorrer deste trabalho, a partir das diversas etapas apresentadas anteriormente. Sendo assim, seguem duas subseções que explicitam os resultados alcançados por meio da construção do protótipo da Ontologia, bem como a partir do desenvolvimento da Arquitetura da Informação concebida.

6.1 Protótipo de Ontologia

No Capítulo anterior, foi destacado o processo de desenvolvimento do protótipo da ontologia. Após a definição do escopo, específico para operações militares, foi realizado o reaproveitamento de ontologias, superiores e de domínio específico, bem como foi realizada a obtenção de um vocabulário de termos militares, por meio de um manual do Ministério da Defesa.

Após a realização dos passos destacados, foi desenvolvida uma aplicação que tem como objetivo extrair termos empregados em operações militares já realizadas, a partir dos registros existentes no *software* Pacificador. Com isso, foi possível refinar os termos componentes do protótipo de ontologia, estando de acordo com um vocabulário já utilizado em outras situações relacionadas com a temática deste trabalho.

Com a realização dos testes iniciais, a aplicação foi empregada com os registros do Pacificador, sendo os dados de entrada a descrição dos eventos cadastrados na base de dados do sistema. É importante salientar que a base de dados foi considerada ‘suja’, com muitos dados irrelevantes e registros incompletos, o que dificultou parcialmente o processo realizado. Como resultado, foi possibilitada a geração de grafos, conforme o exemplo disposto na Figura 6.1:

Da análise do grafo, diversas triplas puderam ser extraídas, formando conjuntos de termos úteis para complementar a formação do vocabulário empregado no protótipo de ontologia, conforme o que segue:

indivíduo	disparar	pistola
tropa	deslocar	meios
pelotão	realizar	higienização
militares	apoiar	eleições
militar	apoiar	itinerário
exército	prestar	apoio
pelotão	identificar	incêndio
tropa	combateu	incêndio
militar	informar viatura	chegada
Tropa	fiscalizar	local
PBCE	deter	suspeitos
pelotão	retornar	missão
militares	apoiaram	transporte
esquadrão	participar	atividade
GC	realizar fluvial	patrulhamento
batalhão	distribuir	alimentos
empresa	transportar	madeira
pelotão	reconhecer	área
batalhão	identificar	desmatamento
exército	participar covid-19	operação
GC	participar	operação
batalhão	apoiar	eleições
patrulha	apreender ilegal	armamento
Pelotão	reconhecer	incêndio
grupo	realizar	patrulhamento
militares	descarregar básico	cestas
empresa	cometer ambiental	crime
PBCE	apreender	drogas

Após a formação do vocabulário a ser utilizado, foram definidas as classes, subclasses e as respectivas hierarquias. Além disso, foram definidas as propriedades das classes e os seus relacionamentos.

Como resultado, foi possível construir o protótipo de ontologia, atingindo um dos objetivos do trabalho. A Figura 6.2, extraída do *software Protégé*, destaca as anotações que identificam detalhes da ontologia.

Annotations +

rdfs:label
Ontologia do Exército Brasileiro

dc:title
OntoEB

dc:rights
A ontologia OntoEB está em desenvolvimento. Este protótipo inicial é para uso exclusivo do Exército Brasileiro.

dc:creator
Javan Cruz

rdfs:comment
ONTOEB (Ontologia do Exército Brasileiro)

dc:subject
Ontologia de domínio específico na área militar, composta por um vocabulário direcionado para utilização em operações militares.

owl:versionInfo
"1.0"

Figura 6.2: Anotações - Protótipo OntoEB.

A partir do vocabulário obtido durante as etapas do desenvolvimento do protótipo de ontologia, foram construídas um total de 383 classes. A Figura 6.3 e a Figura 6.4 apresentam extratos dessas classes.

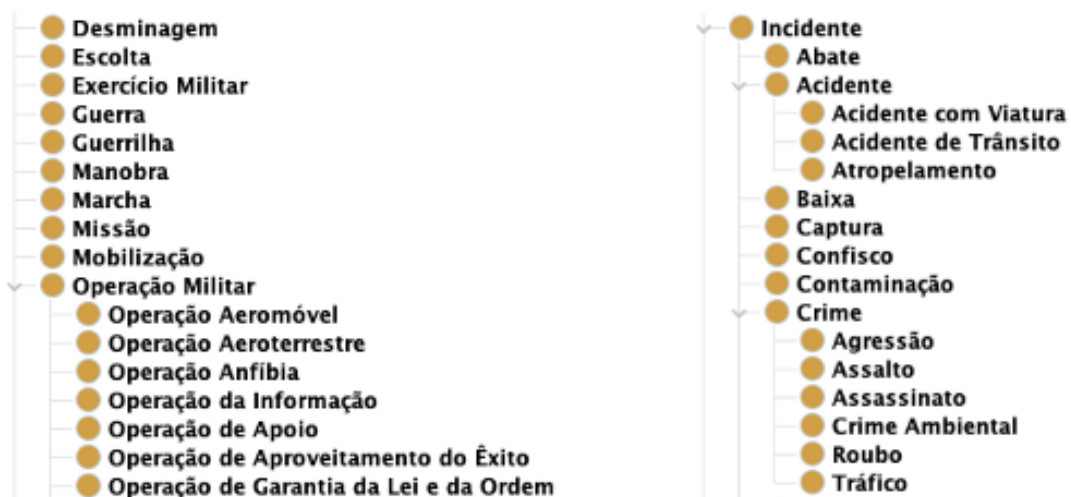


Figura 6.3: Extrato de Classes - Protótipo OntoEB.

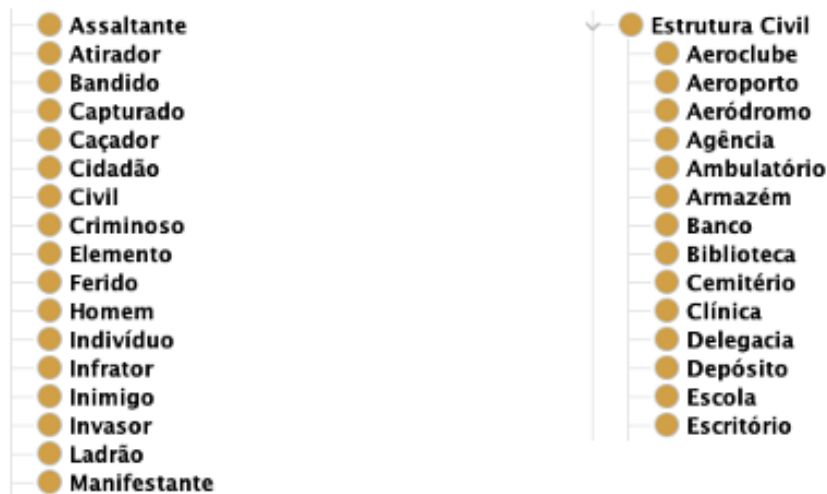


Figura 6.4: Extrato de Classes - Protótipo OntoEB.

Por fim, o protótipo de ontologia foi considerado suficiente para a aplicação em proveito da Arquitetura da Informação construída, estando de acordo com o que foi proposto como objetivo deste trabalho.

6.2 Arquitetura da Informação

No Capítulo anterior, foram apresentados os parâmetros e as tecnologias utilizados para o desenvolvimento da arquitetura de software, que fornece suporte à Arquitetura da Informação proposta. O objetivo foi construir um sistema com uma interface de fácil acesso e intuitiva, permitindo a flexibilidade adequada, além de proporcionar maior oportunidade para o registro por parte do observador, já que os eventos necessitam ser registrados com agilidade e devem ser computados com riqueza de detalhes.

Os conceitos de usabilidade foram buscados ao longo do desenvolvimento do *software*. Dentro desse contexto, princípios e boas práticas foram estratégias relevantes para a interface construída, com o objetivo de propiciar melhores resultados na execução da proposta apresentada.

A Figura 6.5 representa a tela inicial do sistema, que destaca os eventos já registrados, compostos por dados diversos sobre as ocorrências, bem como apresenta ações de visualização, edição e deleção de cada ocorrência. É válido destacar que essas ações são garantidas a partir dos perfis de acesso, discriminados no Capítulo anterior.

Um novo registro no sistema é realizado a partir da entrada de dados que especifiquem os elementos da ocorrência. Pode ser exemplificado um caso no qual o usuário pretende cadastrar um evento que tem como natureza ser uma ‘Operação de Garantia da Lei e

Eventos Registrados						
Consultar						
Mostrar <input type="text" value="10"/> registros						Buscar: <input type="text"/>
ID	Descrição	Data	Horário	Usuário	Ações	
1	Bloqueio	04/02/2022	16:30	1° Sgt Alfa		
2	Incêndio	03/02/2022	14:50	3° Sgt Bravo		
3	Acidente com Viatura	11/02/2022	13:00	1° Ten Charlie		
4	Invasão	07/03/2022	23:30	1° Sgt Alfa		
5	Manifestação	16/03/2022	14:00	1° Ten Charlie		

Figura 6.5: Eventos Registrados - Arquitetura de Software.

da Ordem’. A Figura 6.6 apresenta a busca utilizando a palavra-chave ‘operação’, sendo ‘Operação Militar’ uma das classes resultantes da busca.

NATUREZA DO EVENTO

Insira uma palavra-chave

Buscar Classes

Resultado

Operação Militar
▼

Refinar

Selecionar

Limpar

Figura 6.6: Consulta de classes da ontologia - Arquitetura de Software.

O sistema permite o refinamento da classe, obtendo as suas subclasses, conforme apresentado na Figura 6.7. Foram listadas as diversas subclasses relacionadas com a classe ‘Operação Militar’, conforme especificado no escopo da ontologia. Dentre os resultados torna-se possível selecionar a classe ‘Operação de Garantia da Lei e da Ordem’, estando de acordo com a necessidade do usuário.

The image shows a web interface for refining search results. At the top, there is a dropdown menu labeled 'Resultado' with the selected item 'Operação Militar'. Below this are two buttons: 'Refinar' (highlighted with a red box) and 'Selecionar'. Underneath, a section titled 'Novo Resultado' displays a list of search results. The first item, 'Operação Aeromóvel', is checked with a green checkmark. The item 'Operação de Garantia da Lei e da Ordem' is highlighted with a blue background, indicating it is the current selection. Other items in the list include 'Operação Aeroterrestre', 'Operação Anfíbia', 'Operação de Apoio', 'Operação de Aproveitamento do Êxito', 'Operação Defensiva', 'Operação Especial', 'Operação de Incursão', 'Operação de Infiltração', 'Operação da Informação', 'Operação de Inteligência', 'Operação Ofensiva', 'Operação Psicológica', 'Operação de Reconhecimento', 'Operação de Retardamento', and 'Posto de Bloqueio e Controle de Estradas'.

Figura 6.7: Refinamento da consulta de classes - Arquitetura de Software.

Prosseguindo, com a seleção da classe de interesse, o sistema realizará a busca por propriedades dessa classe, que serão fornecidas ao usuário, para utilização. A figura Figura 6.8 especifica a propriedade ‘Descrição do Evento Militar’, relacionada a classe selecionada.

The image shows a web interface for viewing the properties of a selected class. At the top, there is a dropdown menu labeled 'Novo Resultado' with the selected item 'Operação de Garantia da Lei e da Ordem'. Below this is a button labeled 'Selecionar'. Underneath, a section titled 'PREENCHER PROPRIEDADES' contains a label 'Descrição do Evento Militar' and an empty text input field. Below the input field is a button labeled 'Salvar'.

Figura 6.8: Propriedades da classe selecionada - Arquitetura de Software.

Uma outra possibilidade de busca da classes seria a utilização da palavra-chave ‘GLO’, terminologia militar que representa uma Operação de Garantia da Lei e da Ordem. Como

foi descrito anteriormente, o sistema realiza a busca por *labels* e *comments* das classes. Neste caso, a partir de um dos *comments* da classe buscada ('glo') foi possível encontrar a classe de interesse, conforme apresentado na Figura 6.9.



NATUREZA DO EVENTO

Insira uma palavra-chave

Buscar Classes

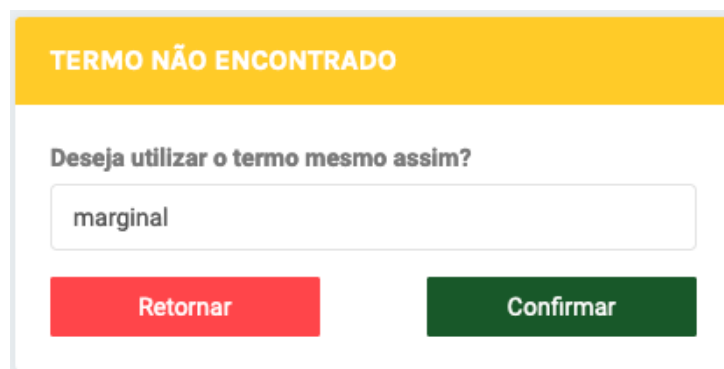
Resultado

Refinar **Selecionar**

Limpar

Figura 6.9: Consulta alternativa de classes da ontologia - Arquitetura de Software.

Uma funcionalidade do sistema, destacada na descrição do fluxograma de arquitetura de software (Figura 5.9), é o tratamento de termos não encontrados durante a busca por classes. Como exemplo, foi buscada uma classe por meio da palavra-chave 'marginal'. Como a classe não foi encontrada, o sistema permite ao usuário decidir ou não por sua utilização, conforme a Figura 6.10



TERMO NÃO ENCONTRADO

Deseja utilizar o termo mesmo assim?

Retornar **Confirmar**

Figura 6.10: Tratamento para classe não encontrada - Arquitetura de Software.

Caso o usuário confirme a utilização do termo, poderá continuar o registro do evento normalmente, sendo a a palavra-chave inserida considerada como uma classe instanciada pelo sistema, de acordo com a Figura 6.11

Insira uma palavra-chave

Buscar Classes

Resultado

Figura 6.11: Termo pendente utilizado pelo usuário - Arquitetura de Software.

Adicionalmente, o termo utilizado pelo usuário será adicionado a um repositório composto por termos pendentes, conforme disposto na Figura 6.12. Desta forma, será possibilitado a um especialista em ontologias realizar a análise do conteúdo do repositório, tomando a decisão pela aceitação ou rejeição dos referidos termos como novas classes na ontologia.

Lista de Termos Pendentes		
Aguardando Análise		
#	Termo Pendente	Ação
1	marginal	

Figura 6.12: Utilização de termo não encontrado - Arquitetura de Software.

Conforme destacado anteriormente, a interface do sistema é responsiva, ou seja, a visualização do conteúdo disponibilizado é adaptável aos diversos dispositivos, independentemente do tamanho de tela. A Figura 6.13 representa a visualização da busca por classes a partir do sistema acessado por meio de um dispositivo móvel.

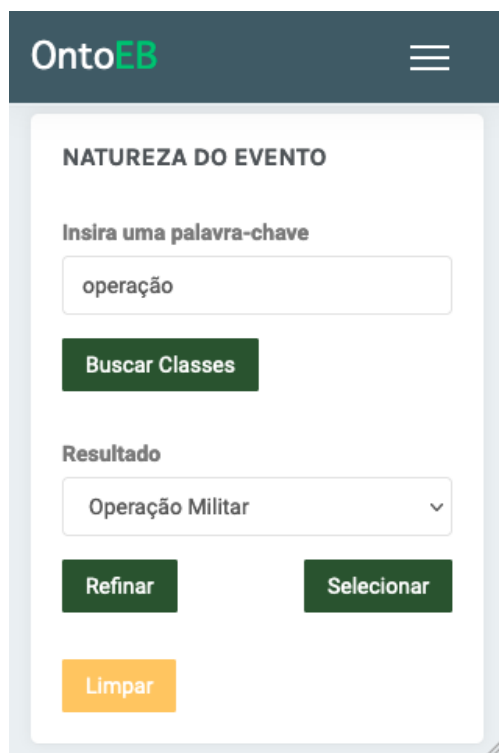


Figura 6.13: Responsividade - Arquitetura de Software.

Após o registro de todos os elementos que contemplam uma ocorrência observada por um usuário, o sistema possibilita visualizar detalhes do relatório consolidado. Na Figura 6.14 é exemplificado um evento de ‘bloqueio de via por caminhoneiros’.

Dados Gerais do Evento	
Data	04/02/2022
Hora	16:30
Usuário	1º Sgt Alfa

Natureza do Evento		Atores do Evento		Local do Evento	
Natureza	Bloqueio	Ator	Grupo de Caminhoneiros	Local	Rodovia
Descrição do Evento	Bloqueio de via por caminhoneiros	Identificação da Entidade	Diversos caminhoneiros do Distrito Federal e de Goiás	Bairro	Não há
				Cidade	Santa Maria/DF - Valparaíso de Goiás/GO
				Estado	DF - GO
				Identificação da Via	BR-040
				Ponto de Referência	KM 0 - Divisa entre os estados

Figura 6.14: Relatório de Evento - Arquitetura de Software.

Sintetizando o que foi apresentado como resultado obtido por meio da construção da Arquitetura da Informação, o que se buscou foi o desenvolvimento de um *software* direcionado à facilitar o registro de observações sobre os eventos, com o suporte fornecido pela representação dos dados por meio de ontologias. Os relatórios produzidos pelo sistema tem como objetivo fornecer informações com oportunidade, de forma a atender as necessidades buscadas.

6.3 Análise da recuperação da informação

O objetivo desta Seção é propor uma análise comparativa entre o modelo de recuperação da informação do Pacificador, sistema atualmente empregado pelo Exército Brasileiro, em apoio às operações militares, e a proposta de Arquitetura da Informação, descrita no escopo desta pesquisa realizada.

6.3.1 Recuperação da informação no sistema Pacificador

O sistema Pacificador é utilizado atualmente em apoio aos Centros de Coordenação de Operações, para coleta e disponibilização de dados sobre incidentes observados, em operações militares do Exército Brasileiro. Após um observador de campo analisar um determinado evento, empregará a ferramenta de relato de incidentes do sistema, para registrar o conteúdo obtido.

Os registros inseridos no sistema são disponibilizados para recuperação, a partir de uma tabela única, sendo cada linha referente a um incidente, com colunas que representam todos os dados coletados (data, local, descrição, entre outros). Em síntese, os milhares de registros, bem como seus respectivos dados coletados, são apresentados em uma interface única, que consolida todas as informações.

Para realizar a recuperação de informações específicas, de forma direcionada à necessidade do usuário, na mesma interface que disponibiliza todos os registros é disponibilizado um sistema de buscas, concebido em um formato simplificado. A partir dos dados registrados, já expostos na tabela, é possibilitada a realização de consultas, nas quais as palavras-chave utilizadas para as buscas devem corresponder exatamente aos dados dispostos na interface. Sendo assim, se por exemplo for realizada uma consulta utilizando as palavras-chave “acidente viatura”, havendo algum registro que tenha em sua descrição “acidente de viatura”, este não será exibido, por não ser efetivada a correspondência exata entre os temas de busca e o que consta entre os dados registrados.

6.3.2 Recuperação da informação na Arquitetura da Informação proposta

Conforme o que foi apresentado no decorrer deste trabalho, a Arquitetura da Informação proposta está direcionada à administração dos dados obtidos, bem como à organização de informações. Após a observação de determinado evento, o usuário utilizará o sistema desenvolvido para o registro dos dados, em um formato assistido por ontologia. O objetivo é buscar a gestão efetiva dos registros realizados, que proporcionará o suporte para a estruturação de relatórios, resultando em maior eficiência para recuperação e análise dos dados disponibilizados.

As informações sobre os eventos, coletadas por meio do sistema, são organizadas por meio de relatórios, disponíveis para recuperação pelos usuários, a partir da ferramenta de consulta, fornecida pelo *software*. As informações essenciais dos incidentes são dispostas em uma interface, que disponibiliza ações a serem realizadas sobre cada dos registros. Dentre as ações, existe a opção de verificar detalhes, que apresenta o relatório completo de informações sobre cada um dos eventos.

Para recuperar informações registradas, o software possui um sistema de buscas, que permite listar incidentes específicos, de acordo com as necessidades de informação. O usuário tem a opção de realizar buscas, utilizando palavras-chave, o que ocasiona uma consulta a partir de todos os dados dos eventos registrados, não apenas sobre os dados constantes da interface inicial. No caso, as consultas também consideram as classes/sub-classes instanciadas da ontologia, bem como as suas respectivas descrições de propriedades de dados, referentes a cada incidente registrado. Como resultado, são retornados os eventos específicos, de acordo com os termos de entrada. Assim, sendo realizada uma consulta empregando as palavras-chave “bloqueio via”, os registros que possuírem, por exemplo, “bloqueio de via” em seu escopo, serão exibidos, permitindo a recuperação dos relatórios correspondentes.

6.3.3 Considerações

De acordo com o que foi apresentado nesta seção, em uma análise puramente sintática, é possível chegar à conclusão de que há vantagens na realização da recuperação da informação empregando a Arquitetura da Informação proposta neste trabalho, quando comparada ao que ocorre no sistema Pacificador, já que as consultas são mais abrangentes, não sendo relacionadas apenas à buscas por expressões exatas, permitindo que sejam contemplados resultados mais compreensivos.

Para o estabelecimento de medidas comparativas para recuperação da informação entre os dois sistemas apresentados, uma possibilidade seria a aplicação de métricas quan-

titativas aos resultados de buscas por incidentes, produzidos a partir de consultas dos usuários. Dentre as diversas métricas de avaliação de recuperação da informação, podem ser destacadas a precisão e a cobertura (revocação), descritas conforme o que segue:

Precisão: capacidade que determinado sistema possui de recuperar apenas itens relevantes, a partir de consulta realizada. De forma geral, seria o quociente, em percentual, entre registros relevantes encontrados por meio da busca e o total de registros recuperados pela busca;

Cobertura: capacidade que determinado sistema possui de recuperar todos os itens relevantes, resultante de consulta realizada. Em resumo, seria o quociente, representado de forma percentual, entre os registros relevantes encontrados pela busca e o total de registros relevantes armazenados no sistema.

Para tais avaliações comparativas baseadas em métricas, em igualdade de condições entre os sistemas, seria pertinente a preparação de bases de dados conhecidas, comuns para ambas as aplicações, compostas pelos mesmos registros de eventos. Além disso, seria adequada a criação de um ambiente de desenvolvimento do Pacificador, pois a realização de testes em ambiente de produção não é oportuna. Com isso, acredita-se que poderiam ser obtidos resultados isentos de possíveis vieses, com métricas que refletiriam a comparação real entre os sistemas.

O Centro de Desenvolvimento de Sistemas do Exército, órgão que colaborou diretamente com a pesquisa realizada, foi consultado sobre a viabilidade de realização dessa análise comparativa entre os sistemas, a partir dos parâmetros estabelecidos. Porém, dentro do tempo disponível, não seria possível realizar essa avaliação, uma vez que o Centro possui diversas atividades correntes, e não teria disponibilidade em alocar pessoal para apoio ao solicitado, uma vez que o responsável por esta pesquisa não possui autonomia para trabalhar diretamente com o Pacificador. Sendo assim, os testes de avaliação, baseados em métricas, foram entendidos como oportunidade para possíveis trabalhos futuros.

Capítulo 7

Conclusão

O projeto Centro de Coordenação de Operações Móvel, pertencente ao Programa Estratégico do Exército PROTEGER é uma realidade para a Força Terrestre. Dentre seus objetivos precípuos podem ser evidenciados o de contribuir para a obtenção de informações visando a formação da consciência situacional de uma operação militar, como também a disponibilização dessas informações aos tomadores de decisão, com oportunidade e precisão.

A situação problema levantada nesta pesquisa destaca a dificuldade em produzir conhecimentos a partir de textos não estruturados, coletados a partir dos relatórios do sistema Pacificador, *software* empregado atualmente em proveito dos Centros de Coordenação de Operações. Nesse contexto, foi desenvolvida a proposta desse trabalho, que tem por objetivo disponibilizar uma Arquitetura da Informação voltada à aquisição de dados em operações, bem como ao fornecimento das informações obtidas, utilizando recursos de ontologias para representação do conhecimento.

Para proporcionar o alcance deste objetivo, o trabalho foi permeado por diversas etapas, dentre as quais se destacam o desenvolvimento de um protótipo de ontologia, voltado para o domínio de operações militares, bem como a construção de uma proposta de Arquitetura da Informação, para assistir a obtenção de dados e prover informações descritas semanticamente, proporcionando o fornecimento de relatórios estruturados.

O protótipo de ontologia foi construído por meio de um faseamento, a partir das boas práticas dispostas no Capítulo 5. Após a determinação do domínio e do escopo da ontologia, foi efetivado o aproveitamento de ontologias existentes, bem como a obtenção dos principais termos a serem utilizados. Com isso, foi proporcionada a construção de classes e de sua hierarquização, além de definição e do detalhamento de suas respectivas propriedades. Para tornar o processo de construção mais assertivo, foi realizada a validação da ontologia por meio de um mecanismo de inferência, de forma a identificar possíveis inconsistências. O resultado foi a disponibilização de um protótipo de ontologia (intitulado

OntoEB), considerado adequado para utilização em prol da Arquitetura da Informação fornecida.

A Arquitetura da Informação proposta foi construída a partir de diversos recursos e tecnologias disponíveis. Para sua materialização, foi desenvolvida uma arquitetura de *software*, escrita em linguagem de programação *Python* (v3.8)¹, empregando o *framework Flask* (v1.1.4)², que possibilitou a utilização do *Python* para o desenvolvimento web. Para permitir a utilização de ontologias em proveito do sistema desenvolvido, foi utilizada a biblioteca *RDFLib* (v6.0.0)³. Além disso, buscando atender aos requisitos que favorecem a Experiência do Usuário, foi tomada por base a construção de uma interface direcionada à oportunidade e à facilidade para o registro de eventos pelos usuários. Por fim, um dos requisitos buscados foi a responsividade, de forma que a interface do sistema pudesse ser adequada aos diferentes tamanhos de tela dos dispositivos, garantindo flexibilidade para o emprego efetivo do sistema em diversos equipamentos.

O resultado obtido foi o fornecimento de uma Arquitetura da Informação, suportada por um protótipo de ontologia, com a capacidade de registrar as ocorrências observadas em operações de forma oportuna. Com isso torna-se possível resolver o problema levantado em sua origem, ou seja, no momento da obtenção dos dados, oferecendo o suporte necessário para a construção dos relatórios a serem disponibilizados para as autoridades. Assim, torna-se favorecida a construção da consciência situacional necessária, bem como o provimento de informações organizadas para apoiar com maior assertividade o processo de tomada de decisão.

7.1 Contribuições

As contribuições principais, contempladas por este trabalho, podem ser salientadas conforme o que segue:

- o fornecimento de um protótipo de Ontologia para utilização pelo Exército Brasileiro, uma vez que não foram identificados estudos relacionados a modelos ontológicos na Instituição;
- a disponibilização de uma proposta de Arquitetura da Informação para viabilizar o registro de dados de eventos de campo, de forma assistida, bem como para estruturação de relatórios informacionais, a partir dos dados obtidos.

¹<https://www.python.org>

²<https://flask.palletsprojects.com>

³<https://rdflib.dev>

7.2 Trabalhos futuros

A partir da abordagem apresentada neste trabalho, acredita-se que alguns assuntos, além do que foi tratado na Seção 6.3.3, tem potencial para direcionar novas atividades de pesquisa e aplicações práticas.

Uma possibilidade futura, seria a realização de melhorias relacionadas com a utilização do *software* desenvolvido, para execução de tarefas pelos usuários. Poderia ser efetivada a aplicação de uma escala numérica de usabilidade, método utilizado para verificar o nível de usabilidade de um sistema de forma mais concreta. O *System Usability Scale* (SUS) representa um desses métodos, focado em efetividade, eficiência e satisfação. O SUS é composto basicamente por um questionário 10 perguntas, para as quais o usuário responderá em uma escala de 1 a 5 pontos. Os resultados obtidos são analisados e permitem obter uma métrica que avalia o nível de usabilidade do sistema desenvolvido. Assim, seria possível verificar se o *software* está de acordo com o estabelecido ou se há a necessidade de serem buscadas melhorias para utilização pelos usuários.

Conforme já apresentado, neste trabalho foi proposta a construção de um protótipo de ontologia. De forma a possibilitar a sua evolução ao longo do tempo, o protótipo poderá ser aprimorado posteriormente, a partir do compartilhamento com especialistas em ontologias. Poderão ser descritas novas classes, propriedades e relacionamentos, enriquecendo o que foi produzido, proporcionando o emprego do modelo ontológico em diversas outras aplicações.

Além disso, a partir do que foi pretendido nesta pesquisa, conforme salientado no Capítulo 5, não foi contemplado o emprego de regras de inferência utilizando o protótipo de ontologia desenvolvido. Acredita-se que, como trabalho futuro, tal conteúdo poderia ser explorado, complementando o que foi construído, além de enriquecer o conteúdo disponibilizado pela arquitetura da informação proposta.

Referências

- [1] Brasil: *Exército brasileiro: Missão e visão de futuro*. Disponível em <http://www.eb.mil.br/missao-e-visao-de-futuro> Acessado em: Julho. 12, 2021., Julho 2021. 1
- [2] Exército, Escritório de Projetos do: *Portfólio estratégico do exército*. Disponível em <http://www.epex.eb.mil.br/index.php/texto-explicativo> Acessado em: Julho. 13, 2021., Julho 2021. 1, 2
- [3] Nóbrega, Gildenildo Paulino da: *Sistemas Militares de Comando e Controle do Exército Brasileiro nas Operações*. Escola de Comando e Estado-Maior do Exército, 2019. 2, 51
- [4] Wazlawick, Raul: *Metodologia de pesquisa para ciência da computação*. Elsevier Brasil, 2017. 6
- [5] Vergara, Sylvia C.: *Projetos e relatórios de pesquisa*. Atlas, 2006. 6
- [6] Gil, Antônio C.: *Métodos e técnicas de pesquisa social*. Atlas, 2008. 6
- [7] Marconi, Marina A. e Lakatos, Eva M.: *Fundamentos da metodologia científica*. Atlas, 2003. 6
- [8] Mariano, Ari Melo e Maíra Rocha Santos: *Revisão da literatura: Apresentação de uma abordagem integradora*. Em *XXVI Congreso Internacional AEDEM / 2017 AEDEM International Conference - Economy, Business and Uncertainty: ideas for a European and Mediterranean industrial policy*, páginas 427–443. Academia Europea de Dirección y Economía de la Empresa (AEDEM), 2017. 7, 8, 12
- [9] Mariano, Ari Melo, Rosario García Cruz e Jorge Arenas Gaitán: *Meta análises como instrumento de pesquisa: Uma revisão sistemática da bibliografia aplicada ao estudo das alianças estratégicas internacionais*. Em *Congresso internacional de Administração: Gestão Estratégica: inovação colaborativa e competitividade*, página 12. Universidade Estadual de Ponta Grossa (UEPG), 2011. 7
- [10] Correa, Patricio Ramírez e Rosario García Cruz: *Meta-análisis sobre la implantación de sistemas de planificación de recursos empresariales (erp)*. Em *Journal of Information Systems and Technology Management*, páginas 225–273, 2005. 7
- [11] Analytics, Clarivate: *Web of science*. Disponível em <http://www.webofscience.com> Acessado em: Junho. 7, 2021. 8, 9, 14

- [12] Lehmann, Jens, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N. Mendes, Sebastian Hellmann; Mohamed Morsey, Patrick Van Kleef e Soren Auer: *Dbpedia – a large-scale, multilingual knowledge base extracted from wikipedia*. *Semantic Web*, 6(2):167–195, 2015. 11, 12, 13
- [13] Bizer, Christian, Tom Heath e Tim Berners-Lee: *Linked data - the story so far*. *International Journal On Semantic Web And Information Systems*, 5(3):1–22, 2009. 13, 28
- [14] Gruber, Thomas R.: *A translation approach to portable ontology specifications*. *Knowledge acquisition*, 5:199–220, 1993. 13, 24, 26
- [15] Berners-Lee, Tim, James Hendler e Ora Lassila: *The semantic web*. *Scientific American*, 284(5):34–43, 2001. 13, 24
- [16] Ismayilov, Ali, Dimitris Kontokostas, Sören Auer, Jens Lehmann e Sebastian Hellmann: *Wikidata through the eyes of dbpedia*. *Semantic Web*, 9(4):493–503, 2018. 13
- [17] Eisele, Scott, Istvan Mardari, Abhishek Dubey e Gabor Karsai: *Riaps: Resilient information architecture platform for decentralized smart systems*. Em *2017 IEEE 20th International Symposium on Real-Time Distributed Computing (ISORC)*, páginas 125–132. Springer International Publishing, 2017. 17
- [18] Rosenfeld, Louis e Peter Morville: *Information Architecture for the World Wide Web*. O’Reilly and Associates, 1998. 18, 40
- [19] Morville, Peter: *Ambient findability*. O’Reilly Media, 2005. 18, 39
- [20] Morville, Peter e Louis Rosenfeld: *Information Architecture for the World Wide Web: Designing Large-Scale Web Sites*. O’Reilly Media, 2006. 18
- [21] Resmini, Andrea e Luca Rosati: *Pervasive Information Architecture: Designing Cross-Channel User Experiences*. Morgan Kaufmann, 2011. 18, 41
- [22] Almeida, Maurício B., Eduardo R. Felipe e Renata Barcelos: *Toward a document-centered ontological theory for information architecture in corporations*. *Journal Of The Association For Information Science And Technology*, 71(11):1308–1326, 2020. 19, 38, 39
- [23] Sowa, John F.: *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks Cole Publishing Co., Pacific Grove, CA, 2ª edição, 2000. 21
- [24] Garshol, Lars Marius: *Metadata? thesauri? taxonomies? topic maps! making sense of it all*. *Journal Of Information Science*, 30(4):378–391, 2004. 21, 22, 23, 27
- [25] Victorino, Marcio Carvalho: *Organização da Informação para dar Suporte à Arquitetura Orientada a Serviços: Reuso da Informação nas Organizações*. Doutorado em ciência da informação, Universidade de Brasília, Brasília, 2011. 21, 22, 23, 34, 40

- [26] (DCMI), Dublin Core Metadata Initiative: *Dublin core metadata element set, version 1.1: Reference description*. Disponível em <https://www.dublincore.org/specifications/dublin-core/dces/> Acessado em: Agosto. 05, 2021., 2008. 21
- [27] Taylor, Arlene G.: *The Organization of Information*. Libraries Unlimited, 2004. 21, 33
- [28] Ferreira, Eveline Cruz Hora Gomes: *Geração Automática de Metadados: uma Contribuição para a Web Semântica*. Doutorado em sistemas eletrônicos, Escola Politécnica da Universidade de São Paulo, São Paulo, 2006. 22
- [29] Greenberg, Jane, Kristina Spurgin e Abe Crystal: *Final report for the omega (automatic metadata generation applications) project*. Disponível em <https://repository.arizona.edu/handle/10150/106026> Acessado em: Agosto. 06, 2021., 2005. 22
- [30] Campos, Maria Luiza Almeida e Hagar Espanha Gomes: *Taxonomia e classificação: o princípio de categorização*. DataGramZero, 9(4), 2008. 22
- [31] Carlan, Eliana e Marisa Brascher Basílio Medeiros: *Sistemas de organização do conhecimento na visão da ciência da informação*. Revista Ibero-Americana de Ciência da Informação, 4(2):53–73, 2011. 22, 24
- [32] Cavalcanti, Cordelia R.: *Indexação e tesouro: metodologia e técnicas*. Associação de Bibliotecários do Distrito Federal (ABDF), Brasília, 1978. 23
- [33] Gomes, Hagar Espanha: *Manual de Elaboração de Tesouros Monolíngue*. Programa Nacional de Bibliotecas de Ensino Superior, Brasília, 1990. 23
- [34] Jesus, Jerocir Botelho Marques de: *Tesouro: um instrumento de representação do conhecimento em sistemas de recuperação da informação*. Disponível em <http://repositorio.febab.org.br/items/show/4075> Acessado em: Agosto. 09, 2021., 2002. 23, 24
- [35] Robredo, Jaime: *Documentação de hoje e de amanhã: uma abordagem revisitada e contemporânea da Ciência da Informação e de suas aplicações biblioteconômicas, documentárias, arquivísticas e museológicas*. Report, Brasília, 2005. 24
- [36] Mizoguchi, Riichiro: *Tutorial on ontological engineering - part 3: Advanced course of ontological engineering*. New Generation Computing, 22(2):198–220, 2004. 24, 29
- [37] Corcho, Oscar, Mariano Fernández-López e Asunción Gómez-Pérez: *Methodologies, tools and languages for building ontologies. where is their meeting point?* Data & Knowledge Engineering, 46(1):41–64, 2003. 24
- [38] Borst, Willem Nico: *Construction of engineering ontologies for knowledge sharing and reuse*. Tese de Doutorado, Institute for Telematica and Information Technology, University of Twente, Enschede - The Netherlands, 1997. 24

- [39] Guarino, Nicola, Daniel Oberle e Steffen Staab: *What is an ontology?* Handbook on Ontologies - International Handbooks on Information Systems, páginas 1–17, 2009. 24
- [40] Christian Mårtenson, Andreas Horndahl e Ziaul Kabir: *An ontology-based adaptive reporting tool*. Em *6th International Conference on Semantic Technologies for Intelligence, Defense, and Security (STIDS) 2011*, páginas 85–88. CEUR Workshop Proceedings, 2011. 25, 42, 43, 44
- [41] Leal, José Paulo: *Using proximity to compute semantic relatedness in rdf graphs*. Computer Science And Information Systems, 10(4):1727–1746, 2013. 25, 32
- [42] Köhler, Jacob, Stephan Philippi e Matthias Lange: *Semeda: ontology based semantic integration of biological databases*. Bioinformatics, 19(18):2420–2427, 2003. 25
- [43] Noy, Natalya F.: *Semantic integration: A survey of ontology-based approaches*. ACM SIGMOD Record, 33:65–70, 2004. 25
- [44] Guarino, Nicola: *Formal ontology and information systems*. Em *In: PROCEEDINGS OF FOIS'98. Amsterdam, Netherlands: IOS Press*, páginas 3–15, 1998. 25
- [45] Guarino, Nicola: *Formal ontology, conceptual analysis and knowledge representation*. International Journal Of Human-Computer Studies, 43(5-6):625–640, 1995. 25
- [46] Uschold, Mike e Michael Gruninger: *Ontologies: principles, methods and application*. The Knowledge Engineering Review, 11(2):93–136, 1996. 25, 46
- [47] BÉZIVIN, Jean: *Who's afraid of ontologies?* Disponível em https://flaterco.com/pur1s/oopsla98_bezivin/ Acessado em: Agosto. 16, 2021., 1998. 25
- [48] Noy, Natalya F. e Deborah L. McGuinness: *Ontology Development 101: A Guide to Creating Your First Ontology*. Stanford University, 2001. 25, 42, 45, 46, 53, 60
- [49] Hwang, Chung Hee: *Incompletely and imprecisely speaking :using dynamic ontologies for representing and retrieving information*. Em *Proceedings of the 6th International Workshop on Knowledge Representation meets Databases (KRDB'99)*, página 13, 1999. 26
- [50] Uschold, Mike: *Building ontologies: Towards a unified methodology*. Em *16th Annual Conference of the British Computer Society Specialist Group on Expert Systems*, página 18, 1996. 27
- [51] Isotani, Seiji e Ig Ibert Bittencourt: *Dados Abertos Conectados*. Nnovatec Editora, São Paulo, 2015. 27, 28, 29, 30, 31, 32
- [52] Sales, Luana Farias, Maria Luiza Almeida Campos e Hagar Espanha Gomes: *Ontologias de domínio: Um estudo das relações conceituais e sua aplicação*. Disponível em <http://repositorios.questoesemrede.uff.br/repositorios/bitstream/handle/123456789/490/Luana%20Farias-Sales.pdf?sequence=1> Acessado em: Agosto. 19, 2021., 2006. 27

- [53] Almeida Falbo, Ricardo de: *Integração de conhecimento em um ambiente de desenvolvimento de software*. Doutorado em engenharia de sistemas e computação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1998. 27
- [54] Carvalho Victorino, Marcio de, Luiz Martins, Maristela Holanda e Rodrigo Fonseca: *Arquitetura de publicação de dados abertos conectados governamentais da universidade de Brasília*. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, 25:1–25, 2020. 28, 31, 33, 42, 47, 48
- [55] Uschold, Mike e Robert Jasper: *A framework for understanding and classifying ontology applications*. Em *Proceedings of the IJCAI-99 workshop on Ontologies and Problem-Solving Methods (KRR5)*, página 12, 1999. 28
- [56] Standards, W3C Semantic Web: *Web ontology language (owl)*. Disponível em <https://www.w3.org/OWL/>. 28
- [57] Mcguinness, Deborah L. e Frank Van Harmelen: *Owl web ontology language overview*. Disponível em <http://www.w3.org/TR/2004/REC-owl-features-20040210/>. 29
- [58] Hitzler, Pascal, Markus Krötzsch, Bijan Parsia, Peter F. Patel-Schneider e Sebastian Rudolph: *Owl 2 web ontology language primer (second edition)*. Disponível em <https://www.w3.org/TR/owl2-primer/>. 30
- [59] Bechhofer, Sean, Simon Harper e Darren Lunn: *Sadle: Semantic annotation for accessibility*. Em *5th International Semantic Web Conference (ISWC 2006)*, páginas 101–115, 2006. 30
- [60] Victorino, Marcio, Maristela Terto de Holanda, Edison Ishikawa, Edgard Costa Oliveira e Sammohan Chhetri: *Transforming open data to linked open data using ontologies for information organization in big data environments of the brazilian government: the brazilian database government open linked data dbgoldbr*. *Knowledge Organization*, 45(6):443–466, 2018. 31
- [61] Bizer, Christian e Richard Cyganiak: *Resource description framework (rdf)*. Disponível em <https://www.w3.org/RDF/> Acessado em: Agosto. 11, 2021., 2014. 31
- [62] Pinheiro, Lênia Vânia Ribeiro: *Processo evolutivo e tendências contemporâneas da ciência da informação*. *Informação e Sociedade: Estudos*, 15(1):13–48, 2005. 33
- [63] Russo, Mariza: *Fundamentos de Biblioteconomia e Ciência da Informação*. Editora E-papers, 2010. 33
- [64] Svenonius, Elaine: *The Intellectual Foundation of Information Organization*. MIT Press, 2009. 33
- [65] Barbieri, Carlos: *Modelagem de Dados*. IBPI Press, 1994. 34, 35
- [66] International Organization for Standardization (ISO): *ISO/IEC 11179-5: Information technology — Specification and standardization of data elements*, 1995. 34

- [67] Choo, Chun Wei: *Organização do Conhecimento*. Senac - São Paulo, 2003. 35, 36
- [68] Davenport, Thomas H.: *Reengenharia de processos: como inovar na empresa através da tecnologia da informação*. Campus, 1994. 35
- [69] Davenport, Thomas H.: *Ecologia da informação: por que só a tecnologia não basta para o sucesso na era da informação*. Futura, 2000. 37, 39
- [70] Bank, Chris e Jerry Cao: *The Guide to UX Design Process & Documentation*. UXPin, 2015. 37
- [71] International Organization for Standardization (ISO): *ISO 9241-210:2010 - Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems*, 2010. 37
- [72] Organização Internacional de Normalização (ISO): *ISO/IEC 9126-1 Engenharia de software – Qualidade de produto – Parte 1: Modelo de qualidade*, 2003. 37
- [73] Dias, Cláudia: *Métodos de avaliação de usabilidade no contexto de portais corporativos: um estudo de caso no senado federal*. Mestrado em ciência da informação, Universidade de Brasília, Brasília, 2001. 38
- [74] Bustamante, Antonio Montes de Oca Sánchez de: <http://scielo.sld.cu/pdf/aci/v12n6/aci04604.pdf>. Disponível em <http://scielo.sld.cu/pdf/aci/v12n6/aci04604.pdf> Acessado em: Julho. 21, 2021., 2004. 38
- [75] Bohmerwald, Paula: *Uma proposta metodológica para avaliação de bibliotecas digitais: usabilidade e comportamento de busca por informação na biblioteca digital da puc-minas*. *Ciência da Informação*, 34(1):95–103, 2005. 38
- [76] Bevan, Nigel: *Usability is quality of use*. Em *Proceedings of the 6th International Conference on Human Computer Interaction*, páginas 349–354. Anzai and Ogawa (Elsevier), 1995. 38
- [77] Kafure, Ivette e Murilo Bastos Cunha: *Usabilidade de ferramentas tecnológicas para o acesso à informação*. *Revista ACB: Biblioteconomia em Santa Catarina, Florianópolis*, 11(2):273–282, 2006. 38
- [78] Wurman, Richard Sasul: *Information Architects*. Graphic Press Inc., 1996. 39
- [79] Brancheau, James C. e James C. Wetherbe: *Information architectures: methods and practice*. *Information Processing and Management*, 22(6):453–463, 1986. 39
- [80] Jonkers, Henk, Marc M.Lankhorst, Hugo W. L. Ter Doest, Farhad Arbab, Hans Bosma e Roel J. Wieringa: *Enterprise architecture: Management tool and blueprint for the organisation*. *Information Systems Frontiers (Springer Science and Business Media LLC)*, 8(2):63–66, 2006. 39
- [81] Mcgee, James e Laurence Prusak: *Gerenciamento Estratégico da Informação*. Campus, Rio de Janeiro, 11ª edição, 1994. 40

- [82] Bailey, Samantha: *Information architecture: a brief introduction*. Disponível em <https://www.iainstitute.org/sites/default/files/bailey-iaintro.pdf> Acessado em: julho. 25, 2021., Março 2003. 40
- [83] Informação (AIFIA), Instituto Asilomar para a Arquitetura de: *25 teses para arquitetura da informação*. Disponível em http://archive.iainstitute.org/pt/25_teses.html Acessado em: Agosto. 02, 2021., 2008. 41
- [84] Exército Brasileiro: *Portaria n.º 258-EME, de 31 de outubro de 2018: Aprova a Compreensão das Operações (COMOP) n.º 02/2018 - O Centro de Coordenação de Operações Móvel dos Grandes Comandos em Apoio às Operações Básicas da Força Terrestre*, 2018. 51
- [85] Mascardi, Viviana, Valentina Cordi e Paolo Rosso: *A comparison of upper ontologies*. Em *WOA 2007 - 8th AI*IA/TABOO Joint Workshop "from Objects to Agents": Agents and Industry: Technological Applications of Software Agents*, páginas 55–64. CEUR Workshop Proceedings, 2007. 54
- [86] DBpedia: *Dbpedia archive*. Disponível em <https://databus.dbpedia.org/ontologies/> Acessado em: Agosto. 30, 2021., Agosto 2021. 54
- [87] Systematic: *Multilateral interoperability programme*. Disponível em <https://systematic.com/en-gb/industries/defence/domains/land/interoperability/mip/> Acessado em: Janeiro. 04, 2022. 56
- [88] Matheus, Christopher J. e Brian Ulicny: *On the automatic generation of an owl ontology based on the joint c3 information exchange data model*. Em *12th International Command and Control Research and Technology Symposium*, 2007. 56
- [89] Ministério da Defesa: *Manual de Abreviaturas, Siglas, Símbolos e Convergências Cartográficas das Forças Armadas*, 2021. 59
- [90] Miroir, Jean Claude: *Guia Prático de Construção de Ontologias OWL*. Universidade de Brasília, 2018. 60