

CENTRO UNIVERSITÁRIO EURÍPIDES DE MARÍLIA
FUNDAÇÃO DE ENSINO “EURÍPIDES SOARES DA ROCHA”
BACHARELADO EM SISTEMAS DE INFORMAÇÃO

**PROCESSO DE ANÁLISE QUANTITATIVA DE
EVENTOS CRIMINAIS UTILIZANDO
ABORDAGEM SEMÂNTICA**

GUSTAVO MARTTOS CÁCERES PEREIRA

ORIENTADOR: PROF. DR. LEONARDO CASTRO BOTEGA

Marília - SP
Novembro/2017

CENTRO UNIVERSITÁRIO EURÍPIDES DE MARÍLIA
FUNDAÇÃO DE ENSINO “EURÍPIDES SOARES DA ROCHA”
BACHARELADO EM SISTEMAS DE INFORMAÇÃO

**PROCESSO DE ANÁLISE QUANTITATIVA DE
EVENTOS CRIMINAIS UTILIZANDO
ABORDAGEM SEMÂNTICA**

GUSTAVO MARTTOS CÁCERES PEREIRA

Monografia apresentada ao Centro Universitário Eurípides de Marília como parte dos requisitos necessários para a obtenção do grau de Bacharel em Sistemas de Informação.
Orientador: Prof. Dr. Leonardo Castro Botega

Marília - SP
Novembro/2017



CENTRO UNIVERSITÁRIO EURÍPIDES DE MARÍLIA - UNIVEM
MANTIDO PELA FUNDAÇÃO DE ENSINO "EURÍPIDES SOARES DA ROCHA"

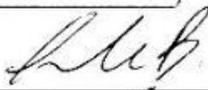
BACHARELADO EM SISTEMAS DE INFORMAÇÃO

Gustavo Marttos Caceres Pereira

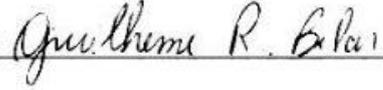
Análise Quantitativa de Eventos Criminais Utilizando Abordagem Semântica

Banca examinadora da monografia apresentada ao Curso de Bacharelado em Sistemas de Informação do UNIVEM/F.E.E.S.R., para obtenção do Título de Bacharel em Sistemas de Informação.

Nota: 10.0 (dez)

Orientador: Leonardo Castro Botega 

1º. Examinador: Fabio Piola Navarro 

2º. Examinador: Guilherme Rodrigues Bilal 

Marília, 27 de novembro de 2017.

Dedico à minha família pelo apoio incondicional.

AGRADECIMENTOS

A Deus, aos meus pais Dorival A. Pereira e Renata M. Cáceres, à minha avó Terezinha das Graças Marttos, ao meu tio Fábio M. Cáceres e à minha bisavó Aparecida R. Marttos (*in memoriam*) por permitirem que eu esteja aonde estou hoje. Vocês são a minha base.

Ao meu Orientador Prof. Dr. Leonardo Castro Botega pela paciência, dedicação e orientação dada.

Aos Professores Jorge Luiz B. Maciel Jr., Elvis Fusco, Fabio P. Navarro e Geraldo Pereira Jr. pelos valiosos ensinamentos.

À minha querida e amada amiga Alice Moraes pelo incentivo e apoio dados ao longo de vários anos.

À minha querida e amada amiga Jordana Nogueira pela companhia nesse trajeto com inúmeros aprendizados e ensinamentos.

Aos meus amigos Leandro Cazarini, Cinthia Silman e Taynara Sene.

Aos amigos e colegas do grupo de pesquisa GIHC.

Ao UNIVEM por proporcionar um ambiente excelente e incrível, o qual me conquistou desde a primeira vez que entrei.

A todos que influenciaram de alguma forma para que eu tenha conseguido esta conquista.

Um homem não pode fazer o certo numa área da vida, enquanto está ocupado em fazer o errado em outra. A vida é um todo indivisível.

Mahatma Gandhi

RESUMO

O gerenciamento de informações de riscos utilizando dados criminais apresenta desafios associados à aquisição de Consciência Situacional, tais como a dinamicidade, heterogeneidade, variedade e o grande volume de dados. Além disso, representar adequadamente as informações no domínio criminal, tais como dados sobre vítimas, criminosos, locais e a própria situação de crime, pode contribuir para processos de análise quantitativa de dados, tornando assim a tomada de decisão ser mais assertiva. Ao mesmo tempo, há ainda a necessidade do desenvolvimento de modelos semânticos que representem a realidade das situações de ambientes de riscos, o que produz ainda mais desafios para a quantificação de entidades e características relevantes. Portanto, este trabalho tem por objetivo apresentar o processo de desenvolvimento de quantificação de dados em ontologias para o domínio de gerenciamento de riscos, visando suportar a extração de dados específicos para ampliar a capacidade de obtenção de Consciência Situacional e permitir que as decisões tomadas, baseadas nestas informações, sejam mais assertivas. Para tal, será empregada a análise de tarefas dirigidas por objetivos, a análise de vocabulários e propriedades no contexto criminal a fim de restringir sua interpretação e o desenvolvimento do modelo semântico. Resultados indicam que o processo desenvolvido é capaz de mensurar informações relevantes dentro de uma ontologia dedicada ao gerenciamento de riscos, permitindo que o operador utilize a informação para adquirir e manter seu nível de Consciência Situacional e melhorar a ação de tomada de decisão.

Palavras-chave: Consciência Situacional; Ontologia; Análise Quantitativa; Gerenciamento de Riscos.

ABSTRACT

The management of risk information using criminal data presents challenges associated with the acquisition of Situational Awareness, such as dynamicity, heterogeneity, variety and large data volume. In addition, adequately representing information in the criminal domain, such as data on victims, criminals, places and the actual crime situation, can contribute to quantitative data analysis processes, thus making decision making more assertive. At the same time, there is still a need for the development of semantic models that represent the reality of situations in risk environments, which produces still more challenges for the quantification of entities and relevant characteristics. Therefore, this paper aims to present the process of a development of data quantification in ontologies for the domain of risk management, aiming at support the extraction of specific data to increase the capacity to obtain Situational Awareness and allow the decisions taken to be more assertive. For this, the analysis of tasks directed by objectives, the analysis of vocabularies and properties in the criminal context will be used in order to restrict their interpretation and the evaluation of semantic weights and values of these vocabularies. Results indicate that the developed process is capable of measuring relevant information within an ontology dedicated to risk management, allowing the operator to use the information to acquire and maintain their level of Situational Awareness and to improve decision making.

Keywords: Situational Awareness, Ontology, Quantitative Analysis, Risks Management.

LISTA DE FIGURAS

Figura 2.1 – Modelo de Consciência Situacional de Endsley	18
Figura 3.1 – Exemplo de Ontologia	25
Figura 3.2 – Proposta para Captura de Intervalo de Tempo	32
Figura 3.3 – Valores e Relações baseados em Intervalo de Tempo	32
Figura 3.4 – Primeiro momento no campo de batalha	33
Figura 3.5 – Segundo momento no campo de batalha	33
Figura 4.1 – Quantificação Universal	36
Figura 4.2 – Quantificação Existencial	37
Figura 4.3 – Representação de Ontologia de Filme	41
Figura 5.1 – Fluxograma da metodologia utilizada	47
Figura 5.2 – Parte do GDTA gerado a partir dos resultados da PMESP	48
Figura 5.3 – Vocabulários utilizados de acordo com a relevância no domínio	49
Figura 5.4 – Metodologia 101 de Noy e McGuinness	50
Figura 5.5 – Ontologia de domínio criminal desenvolvida	52
Figura 5.6 – Exemplo de Instância da Ontologia	54

LISTA DE TABELAS

Tabela 4.1 – Diretores de Cinema com Identificação	39
Tabela 4.2 – Filmes de Cinema com Identificação	39
Tabela 4.3 – Resultado da Quantificação Sintática - Exemplo 1	40
Tabela 4.4 – Resultado da Quantificação Sintática - Exemplo 2	40
Tabela 5.1 – Resultado Quantificado da Proposição 1	56
Tabela 5.2 – Resultado Quantificado da Proposição 2	57
Tabela 5.3 – Resultado Quantificado da Proposição 3	58
Tabela 5.4 – Resultado Quantificado da Proposição 4	60

LISTA DE CÓDIGOS-FONTE

Código-Fonte 4.1 – Exemplo de RDF	38
Código-Fonte 4.2 – Consulta SQL do Exemplo 1	39
Código-Fonte 4.3 – Consulta SQL do Exemplo 2	40
Código-Fonte 4.4 – Consulta SPARQL do Exemplo 1	42
Código-Fonte 4.5 – Consulta SPARQL do Exemplo 2	42
Código-Fonte 4.6 – Consulta SPARQL do Exemplo 3	43
Código-Fonte 4.7 – Consulta SPARQL do Exemplo 4	43
Código-Fonte 5.1 – Prefixos utilizados para as consultas SPARQL no estudo de caso	55
Código-Fonte 5.2 – Consulta SPARQL da Proposição 1 para retornar a relação de situações onde a vítima possui algum objeto	56
Código-Fonte 5.3 – Consulta SPARQL da Proposição 1 para retornar a quantidade por situações onde a vítima possui algum objeto	56
Código-Fonte 5.4 – Consulta SPARQL da Proposição 2 para retornar a relação de locais públicos ou privados das situações de furtos e roubos	57
Código-Fonte 5.5 – Consulta SPARQL da Proposição 2 para retornar a quantidade por locais públicos ou privados das situações de furtos e roubos	57
Código-Fonte 5.6 – Consulta SPARQL da Proposição 3 para retornar a relação de vítimas com veículos furtados ou roubados em locais públicos	58
Código-Fonte 5.7 – Consulta SPARQL da Proposição 3 para retornar a quantidade por locais públicos onde vítimas tiveram seus carros furtados ou roubados	58
Código-Fonte 5.8 – Consulta SPARQL da Proposição 4 para retornar a relação de situações onde há objetos roubados ou furtados	59
Código-Fonte 5.9 – Consulta SPARQL da Proposição 4 para retornar a quantidade por situações onde há objetos roubados ou furtados	60

LISTA DE ABREVIATURAS E SIGLAS

GDTA - *Goal-Directed Task Analysis*

OWL - *Web Ontology Language*

PMESP - *Polícia Militar do Estado de São Paulo*

RDF - *Resource Description Framework*

RDF-S - *RDF Schema*

SAW - *Situational Awareness*

SPARQL - *Simple Protocol and RDF Query Language*

SQL - *Structured Query Language*

W3C - *World Wide Web Consortium*

XML - *eXtensible Markup Language*

SUMÁRIO

	LISTA DE CÓDIGOS-FONTE	9
	1 INTRODUÇÃO	13
1.1	Contexto	13
1.2	Motivação e Objetivos	14
1.3	Metodologia de Desenvolvimento do Trabalho	14
1.4	Organização do Trabalho	15
	2 CONSCIÊNCIA SITUACIONAL	16
2.1	Consciência Situacional	16
2.1.1	Modelo de Consciência Situacional de Endsley	17
2.1.2	Problemas para a obtenção e manutenção de Consciência Situacional	18
2.1.3	Uso de Consciência Situacional em Ambientes Críticos	20
2.2	Considerações Finais	21
	3 ONTOLOGIAS EM SUPORTE A SAW	22
3.1	Ontologias	22
3.1.1	Estrutura de uma Ontologia	23
3.1.2	Classificação das Ontologias	25
3.1.3	Vocabulários	27
3.1.4	Metodologias de Desenvolvimento	27
3.1.5	Representação de Ontologias	30
3.2	Ontologia como apoio à SAW	30
3.2.1	Ontologias de Domínio	33
3.3	Considerações Finais	34
	4 QUANTIFICAÇÃO DE DADOS SEMÂNTICOS	35
4.1	Introdução à Quantificação de Dados	35
4.1.1	Quantificadores Universais	36
4.1.2	Quantificadores Existenciais	37
4.2	Quantificação em Modelos Estruturados e Sintáticos	38
4.3	Quantificação em Modelos Semânticos	40
4.4	Considerações Finais	43
	5 PROCESSO DE ANÁLISE QUANTITATIVA DE EVENTOS CRI- MINAIS UTILIZANDO ABORDAGEM SEMÂNTICA	45

5.1	Estado da Arte em Abordagens de Análise Quantitativa em Modelos Semânticos	45
5.2	Metodologia	47
5.2.1	Levantamento de Requisitos	47
5.2.2	Análise de Vocabulários	47
5.2.3	Desenvolvimento de Ontologia de Domínio	49
5.3	Estudo de Caso	51
5.3.1	Criação das Instâncias	53
5.3.2	Definição de Proposições	53
5.3.3	Construção de Consultas SPARQL	55
5.3.3.1	Proposição 1 - Relação das situações onde a vítima possui algum objeto .	55
5.3.3.2	Proposição 2 - Quantidade de locais públicos ou privados das situações de furtos e roubos	57
5.3.3.3	Proposição 3 - Relação de vítimas com veículos furtados ou roubados em locais públicos	58
5.3.3.4	Proposição 4 - Relação de situações onde há objetos roubados ou furtados	59
	6 CONCLUSÕES	61
6.1	Publicações	62
	REFERÊNCIAS	63

Capítulo 1

INTRODUÇÃO

Neste Capítulo é apresentada uma introdução referente aos temas que compõem a estrutura deste trabalho, bem como a motivação e os objetivos, além da metodologia adotada.

1.1 Contexto

Consciência Situacional (do inglês *Situational Awareness* - SAW) é um conceito fundamental para auxiliar a tomada de decisão em ambientes complexos e dinâmicos em uma variedade de domínios, entre eles o de gerenciamento de riscos (BOSSÉ; ROY; WARK, 2007).

Operadores de sistemas de gerenciamento de riscos, como os que lidam com dados criminais, e que estão constantemente sob alta pressão e expostos a uma gama de informações sensíveis, precisam manter seus níveis de SAW elevados para assim sustentar o melhor retrato de uma situação crítica e tomar a decisão mais assertiva, evitando prejuízos à vida, ao patrimônio e também ao meio ambiente. Por ser um estado cognitivo do operador, a SAW não garante que o operador tomará a melhor decisão, entretanto garante melhores subsídios para que possa melhorá-la.

Para suportar a obtenção e manutenção de SAW, foi constatado que o uso de modelos semânticos, mais especificamente ontologias, quando aplicados para suportar sistemas de gerenciamento de riscos, podem contribuir para uma melhor assertividade nas inferências úteis à tomada de decisão (MATHEUS; KOKAR; BACLAWSKI, 2003).

A tendência dos sistemas dinâmicos é utilizar ontologias ou outros modelos semânticos para que os dados sejam representados. Entretanto, esta ação gera novos desafios para quantificar os dados necessários para que assim sejam transformados em informações que apoiem à tomada de decisão dos operadores de sistemas de

gerenciamento de riscos.

1.2 Motivação e Objetivos

A complexidade e a dinâmica de um ambiente de avaliação de riscos aumentam proporcionalmente conforme a quantidade de variáveis existentes nesse ambiente, tornando a aquisição e manutenção de SAW processos mais difíceis de serem atingidos.

Partindo dessa situação, foi identificado que a quantificação é um excelente subsídio para evoluir a SAW de operadores e conseqüentemente melhorar a decisão por ele tomada. Fluit, Sabou e Harmelen (2006) apontaram a dificuldade dos operadores em fornecer termos que melhor descrevam as suas necessidades de informação. Esta dificuldade aumenta quando expressões lógicas simples são utilizadas, tornando o processo de quantificação mais complicado. A quantificação é fundamental para a estruturação de um processo de análise crítica sobre os dados.

O objetivo geral do presente trabalho é contribuir com a SAW de operadores de sistemas de gerenciamento de riscos, utilizando a análise quantitativa como processo para gerar subsídios necessários a fim de tomar a decisão mais assertiva.

Por sua vez, o objetivo específico é desenvolver um processo de quantificação de informações de riscos, organizadas e representadas semanticamente em uma ontologia de domínio, sendo exclusivamente voltada para o contexto criminal.

1.3 Metodologia de Desenvolvimento do Trabalho

A pesquisa do presente trabalho tem característica experimental, exploratória, quantitativa e dirigida por um estudo de caso e é fundamentada nas bases teóricas da matemática sobre quantificadores e seus derivados, bem como em uma ontologia desenvolvida a partir da metodologia proposta por Noy e McGuinness (2001).

Com o propósito de atingir os objetivos, o trabalho foi dividido nas etapas a seguir:

- Pesquisa e levantamento de metodologias e processos de quantificação semântica;
- Compreensão do domínio proposto para desenvolvimento da ontologia;
- Levantamento de requisitos necessários do domínio aplicando a Análise de Tarefas Dirigida a Objetivos (GDTA) nos operadores de sistemas de gerenciamento de riscos;

- Análise de vocabulários relevantes com o auxílio de um especialista no domínio criminal;
- Desenvolvimento de ontologia de domínio, baseada na Metodologia 101 de Noy e McGuinness (2001);
- Estudo de caso para validar a ontologia e executar consultas SPARQL a fim de obter os resultados que validem a presente monografia.

Os dados utilizados para a conclusão de testes e estudo de caso foram Boletins de Ocorrências, fornecidos pela Secretaria de Segurança Pública do Estado de São Paulo e pela Polícia Militar do Estado de São Paulo, em consonância com a Lei de Acesso à Informação (Lei 12.527/2011 e Decreto Estadual nº 58.052/2012).

1.4 Organização do Trabalho

No Capítulo 2 é apresentado os conceitos de SAW e sua relação com os ambientes de riscos. O Capítulo 3 aborda os conceitos do tema, as principais metodologias utilizadas para desenvolvimento de uma ontologia, as classificações e seis respectivos tipos de representação, bem como sua estrutura apoiada a SAW e em um domínio específico. Já o Capítulo 4 apresenta os conceitos teóricos e metodológicos da quantificação de dados, base deste trabalho, assim como a quantificação em modelos estruturados, sintáticos e semânticos. O Capítulo 5 apresenta o Estado da Arte, a metodologia de trabalho em detalhes, seguida de um Estudo de Caso como demonstração e validação com o propósito de comprovar a necessidade da abordagem do tema deste trabalho. Por fim, o Capítulo 6 remete às conclusões obtidas com este trabalho.

Capítulo 2

CONSCIÊNCIA SITUACIONAL

Neste Capítulo serão apresentados os conceitos de Consciência Situacional (*Situational Awareness – SAW*), referenciando o modelo desenvolvido por Endsley (ENDSLEY, 1988).

Serão abordados os principais problemas que podem ocorrer com um modelo mal elaborado fazendo referência ao modelo citado acima, além do seu uso em sistemas dinâmicos e complexos, como os de domínios militares.

2.1 Consciência Situacional

O termo consciência situacional (SAW) mostrou-se como um importante conceito que auxilia na dinâmica da tomada de decisão humana (BOSSÉ; ROY; WARK, 2007). A SAW permite que as pessoas estejam cientes do que está acontecendo ao seu redor em determinado momento e que entendam como essa informação pode ser utilizada no presente e no futuro, tornando a tomada de decisão e a ação eficazes. Geralmente, ela pode ser relacionada com os objetivos existentes em uma função, sendo aplicada a ambientes com situações operacionais.

Além de relacionada, ela é estudada em diversas áreas, como educação e operações de centrais elétricas, além daquelas que utilizam sistemas críticos¹, como operações militares e aviação.

Um exemplo rotineiro de SAW é um motorista, o qual precisa estar ciente de todas as ações tomadas, como quando acelerar e frear, quando deve parar no semáforo e principalmente estar atento aos pedestres e carros que estejam ao seu redor (ambiente dinâmico) e observar as condições do tráfego para que assim ele possa escolher o melhor trajeto, se possível.

¹ Sistemas que manipulam informações que visam a segurança de pessoas e patrimônios (BOTEGA, 2016).

Sua definição formal foi contemplada como "a percepção dos elementos no ambiente dentro de um volume de tempo e espaço, a compreensão de seu significado e a projeção de seu status no futuro próximo"(ENDSLEY, 1988).

A complexidade e a dinâmica do ambiente aumentam proporcionalmente, conforme a quantidade de variáveis existentes no ambiente, tornando a aquisição e manutenção de SAW processos mais difíceis.

Para Endsley (1995b), dentro de ambientes dinâmicos, a necessidade de tomadas de decisões rápidas é demasiadamente grande, principalmente dentro de um espaço de tempo bastante estreito, cujas tarefas dependem de análises contínuas e que estejam atualizadas.

É possível qualificar a SAW de um indivíduo por meio de suas capacidades, experiências, preconceitos, metas e objetivos e carga de trabalho, entretanto até mesmo os mais qualificados podem tomar decisões errôneas caso sua SAW esteja imprecisa.

2.1.1 Modelo de Consciência Situacional de Endsley

Dentro de diversas definições, o modelo proposto mais aceito pela comunidade de fusão de dados foi o de Endsley (1988). Esse modelo tem sido utilizado em estudos de diversas áreas como pretexto para desenvolver o processo de SAW apoiado por sistemas, principalmente aqueles considerados críticos.

Existem três níveis, que são influenciados direta e indiretamente por fatores internos e externos, os quais os indivíduos podem atingi-los conforme o seu desenvolvimento dentro do ambiente, como mostra a Figura 2.1.

- Nível um: Percepção dos elementos no ambiente. Este é o primeiro passo para se adquirir a SAW, o qual é necessário, dentro do ambiente, perceber a dinâmica, o estado e os atributos de elementos pertinentes.
- Nível dois: Compreensão da situação atual. Baseado na síntese dos elementos encontrados no ambiente, é necessário criar uma relação entre eles, desenvolvendo um contexto de acordo com os objetivos e metas, e priorizando as importâncias e o que essa informação combinada representa.
- Nível três: Projeção de estados futuros. Ter ciência do que os elementos são e o que eles significam perante a situação atual, é o último passo para o desenvolvimento de SAW. Nesta etapa é necessário ter a capacidade de prever quais serão as ações futuras desses elementos perante o objetivo principal.

Não é possível, entretanto, alcançar o Nível 3 sem ter total compreensão dos níveis anteriores. Predições requerem experiências do domínio em que se situa, portanto um indivíduo necessitaria de um sistema que o apoiasse na tomada de decisões.

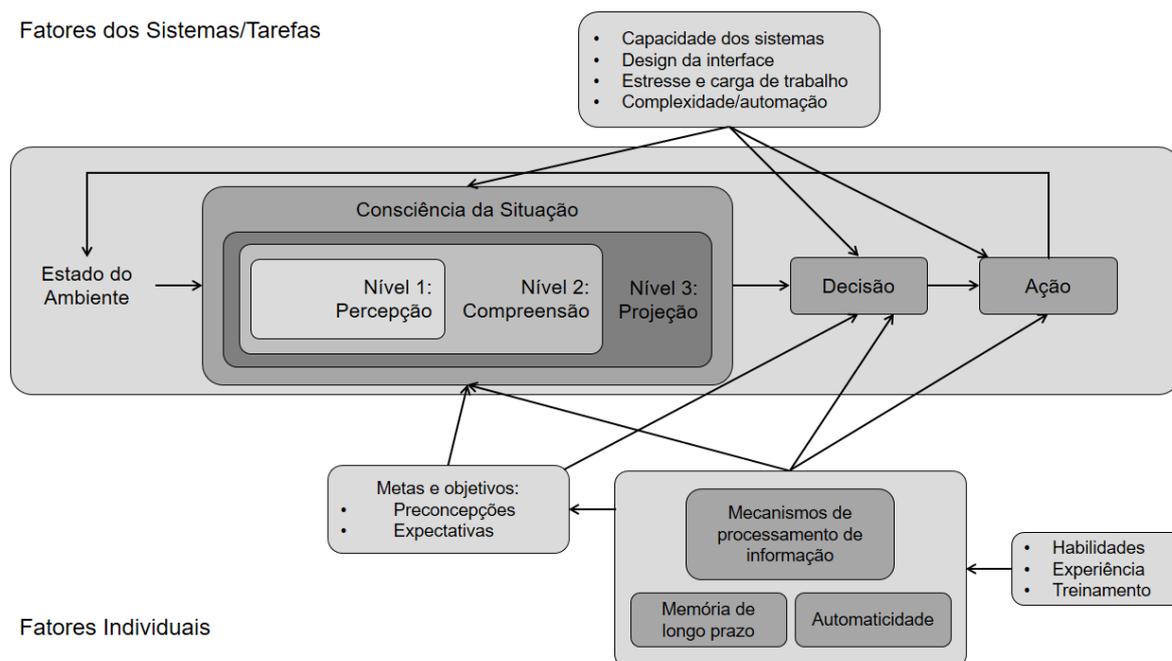


Figura 2.1 – Modelo de Consciência Situacional de Endsley, traduzido de (ENDSLEY, 1988)

Inúmeras variáveis devem ser consideradas para se obter SAW, entre elas os elementos, os quais devem ser claros, ou seja, identificar as reais necessidades de que o indivíduo precisa perceber e entender. A segunda variável é o tempo, pois SAW não é construída imediatamente, ela é dinâmica e de natureza temporal, totalmente dependente dos conhecimentos e experiências adquiridos a curto e longo prazos. Por fim, o espaço também é considerado, uma vez que os dados podem ter sua relevância variada conforme o tempo.

SAW é um importante insumo tanto para formar uma base sólida para a tomada de decisão, quanto para impactar no processo em si.

Endsley (1995b) considera que na ausência de um modelo adequado, os indivíduos muitas vezes não conseguem resolver um novo problema, mesmo que eles precisassem aplicar a mesma lógica que a usada para um problema familiar.

2.1.2 Problemas para a obtenção e manutenção de Consciência Situacional

Adquirir e manter a SAW de operadores de sistemas críticos, como uma cabine de um avião, é uma tarefa complexa. Pelo fato de o ambiente ser dinâmico e a quantidade de variáveis a ser processada pelos operadores ser grande, Endsley (1988) apresentou oito problemas que podem atrapalhar a manutenção da SAW:

1. Encapsulamento de atenção: o grande volume de informações provenientes de um sistema complexo junto a uma situação crítica faz com que o operador tenha sua atenção dispersa.
2. Limitações de memória: para se operar sistemas complexos, é necessário que o operador tenha domínio do ambiente e que possa contar com suas memórias de curto e longo prazo, pois é com base nelas que ele poderá interagir com outros ambientes.
3. Fatores estressantes: determinados ambientes podem fazer com que o operador aumente o seu nível de estresse e ansiedade, podendo causar fadiga. Logo, é comum que seja apresentado uma baixa disponibilidade de recursos cognitivos.
4. Sobrecarga de dados: operadores de sistemas críticos possuem uma limitação cognitiva do quanto de dados eles podem absorver, porém a dinamicidade de dados cria uma necessidade de consumir informações rapidamente, interferindo no fluxo de dados que o operador está consumindo.
5. Informações desnecessárias: sistemas e ambientes complexos e dinâmicos que não têm um bom design orientado a SAW tiram o foco do operador de seu objetivo.
6. Complexidade: quanto mais recursos e regras forem aplicadas, mais complexo se torna o conhecimento do sistema, o que pode resultar em riscos à vida, ao patrimônio ou ao meio ambiente.
7. Modelos mentais vagos: a criação de modelos mentais por parte do operador é importante, porém dentro de novas situações, os modelos adquiridos em situações passadas podem interferir nas interpretações das novas informações, o que consequentemente afeta os Níveis 2 e 3 do Modelo de Consciência Situacional.
8. Automações de processos: apesar da automação ser uma das etapas importantes para a elaboração de uma boa interface orientada a SAW para sistemas críticos, às vezes, ela pode fazer com que o operador não tenha total ciência do que está acontecendo, fazendo com que ele não tenha mais controle, prejudicando a SAW.

Falhas humanas decorrente às decisões tomadas são constantemente referenciadas em investigações que só aconteceram devido aos erros que os sistemas apresentaram. Segundo Endsley (1995a), dentre os acidentes que ocorreram no campo da aviação, 51,6% ocorreram devido falhas na aquisição de SAW ocasionando perda de vidas e dentre 35,1% dos não fatais sabe-se que 80-85% são atribuídos ao erro humano.

2.1.3 Uso de Consciência Situacional em Ambientes Críticos

As primeiras discussões a respeito de SAW ocorreram dentro da comunidade de aviação na Primeira Guerra Mundial (ENDSLEY; GARLAND, 2000). Inicialmente, SAW foi estudada no campo da aviação (pilotos e controladores de tráfego aéreo), depois no campo militar (comando e controle), medicina (anestesiologia), entre outros.

Para Botega (2016) é irrefutável que a dependência de metas e objetivos existentes em funções e trabalhos específicos e da necessidade e importância da SAW como apoio às tomadas de decisões se aplica a quase todos os campos de atuação.

No domínio da aviação, os controladores de tráfego aéreo costumam se referir à SAW como "a imagem da situação", que é o modelo mental da situação no qual se baseiam suas decisões (BOTEGA, 2016 apud ENDSLEY, 1988). "Ter a imagem da situação" para os controladores é o mesmo que atingir os Níveis 1 e 2 conforme subseção 2.1.1. "Perder a imagem da situação" é o mesmo que não conseguir atingir o Nível 3, portanto é dito como um dos maiores perigos para os controladores de tráfego aéreo, uma vez que estes se tornam inaptos de prever a evolução da situação.

É importante ressaltar que mesmo operadores mais experientes e qualificados podem não atingir o Nível 3 em determinadas situações por fatores diversos, como informação incompleta ou imprecisa e esgotamento mental (fadiga, estresse).

Em ambientes emergenciais - cujas falhas podem provocar danos ao meio ambiente, perda de vidas e danos ao patrimônio -, como controle de tráfego aéreo (detecção e prevenção de problemas, planejar e indicar melhores rotas), salas de emergências hospitalares (status de paciente, com base nas leituras de sensores e exames e seguir o melhor procedimento conforme o estado do paciente) e comando e controle policial/militar (posicionamento de tropas, disponibilidade de armamentos), é necessário adquirir, manter e melhorar a SAW constantemente e para isso é preciso que existam sistemas que apoiem sua obtenção.

O objetivo desses sistemas é fornecer informações precisas, de forma rápida e em um formato correto aos operadores. Eles devem ser elaborados de modo que auxiliem uma tomada de decisão mais assertiva ao operador, considerando o tempo como fator limitador. No domínio militar, pode-se constatar o uso de SAW e a aplicação de seus níveis da seguinte forma:

- Nível 1 de SAW: Percepção do operador sobre o posicionamento e condicionamento físico de sua equipe, armamento disponível e a quantidade de tropas inimigas.
- Nível 2 de SAW: Compreensão sobre o avanço de tropas inimigas, possibilidade de deslocamento ou combate.

- Nível 3 de SAW: Visão de um futuro próximo em relação à tropa do inimigo, se sua atual posição representa uma ameaça ou risco iminente à equipe.

Botega (2016) analisou o domínio da PMESP (Polícia Militar do Estado de São Paulo) em relação ao atendimento a chamadas realizadas ao serviço de emergências 190:

"Uma má SAW afeta o modelo mental do humano operador ao atender uma chamada de emergência, o que leva a uma má compreensão da situação e a uma tomada de decisão comprometida."

Operadores dedicados ao atendimento de emergências devem ser capazes de identificar vítimas e criminosos, principalmente suas características, aspectos físicos e comportamentos por meio das informações dadas pelo denunciante. Isto permite que os operadores consigam adquirir e manter uma SAW mais confiável e consequentemente tomar decisões mais assertivas, como quantas viaturas deslocar e quais outros operadores dentro da cadeia hierárquica escalar.

2.2 Considerações Finais

Neste Capítulo foram apresentados os conceitos de consciência situacional, como adquirir e manter - juntamente com seus respectivos problemas - e como funcionam os Níveis de SAW baseado no Modelo de Endsley (1988).

Vale enfatizar a importância de SAW na tomada de decisão em ambientes críticos, pois ela representa o conhecimento que um operador possui sobre este ambiente, incorporando a percepção de elementos primordiais, a compreensão desses em relação aos objetivos do operador e a projeção de futuros estados do ambiente com base em seu entendimento.

No Capítulo seguinte serão apresentados os conceitos de Ontologia no cenário de Inteligência Artificial, bem como suas características e como ela pode apoiar a SAW.

Capítulo 3

ONTOLOGIAS EM SUPORTE A SAW

Neste Capítulo serão abordados os conceitos de Ontologia, sua estrutura e classificação, seu relacionamento com a Web Semântica, os vocabulários utilizados para construí-la, as metodologias de desenvolvimento e as representações de ontologias.

Também será apresentada sua relação com Consciência Situacional e como que a Ontologia pode auxiliar neste processo.

3.1 Ontologias

A palavra ontologia é derivada da Metafísica, sendo o estudo mais geral do Ser, Existência ou Realidade. A palavra é originária de duas outras palavras: *onto*, que significa "o Ser" e *logia*, "estudo ou conhecimento". Sendo assim, conforme Schiessl (2007), ontologia significa o "estudo ou conhecimento do Ser, dos entes ou das coisas tais como são em si mesmas, real e verdadeiramente". Entretanto, na área da Inteligência Artificial, ontologia é um meio utilizado para que se possa organizar informações a fim de criar representações formais de conhecimento.

Uma definição computacional mais completa pode ser vista em Gómez-Pérez e Benjamins (1999), a qual os autores conceituam ontologia como um conjunto de termos hierarquicamente ordenados para descrever um domínio que possa ser utilizado como princípio para uma base de conhecimento. Guarino (1998) permite entender o conceito por meio de outra perspectiva, a qual ele define ontologia como uma lógica que explica o significado pretendido de um vocabulário formal, ou seja, seu compromisso ontológico com uma conceitualização particular do mundo.

Uma ontologia possibilita a criação de um nível genérico definido por meio do conhecimento do domínio, a qual visa fornecer uma compreensão comumente acordada, que pode ser reaproveitado e compartilhado entre outras aplicações e grupos.

Santarém Segundo (2015) afirma que:

Utilizar ontologias é uma das maneiras de se construir uma relação organizada entre termos dentro de um domínio, favorecendo a possibilidade de contextualizar os dados, tornando mais eficiente e facilitando o processo de interpretação dos dados pelas ferramentas de recuperação da informação.

As ontologias tornam-se de fato computacionais a partir do momento que elas são implementadas, algo além da conceitualização, viabilizando o desenvolvimento de aplicações mais robustas e inteligentes.

3.1.1 Estrutura de uma Ontologia

Durante o desenvolvimento de uma ontologia, deve-se discutir e entender os conceitos essenciais do domínio com o qual está trabalhando, pois assim a definição de terminologias, ou seja, o vocabulário, torna-se um processo mais fácil de ser concluído.

Isotani e Bittencourt (2015 apud MIZOGUCHI, 2004) dizem que uma ontologia é composta pelo seguinte:

- Um conjunto de conceitos fundamentais resultantes da inter-relação do conhecimento básico presente em um determinado domínio. Pode ser utilizado um vocabulário especializado para representar tais conceitos.
- O corpo de conhecimento, que descreve o domínio utilizando os conceitos fundamentais. Ele é constituído por:
 - Uma hierarquia (classe/subclasse) resultante das relações *is-a* (é um(a)) entre conceitos, ou seja, uma taxionomia.
Os conceitos podem ser abstratos (*e.g.*, força), concretos (*e.g.*, ser humano), elementares (*e.g.*, elétron) ou compostos (*e.g.*, átomo), reais ou fictícios.
Um exemplo mais completo de taxionomia é o conceito *homem* estar hierarquicamente abaixo do conceito *pessoa*.
 - Um conjunto de relações necessárias entre conceitos além das relações *is-a* (como, *part-of* (parte de)).
O relacionamento entre os conceitos *pessoa* e *carro* resultam no relacionamento de *ser-dono*, por exemplo.
 - Uma axiomatização de restrições semânticas entre esses conceitos e relações.

Os axiomas são as regras pertinentes ao domínio em questão. Um exemplo de axioma é afirmar que toda *pessoa* tem uma mãe, considerando *pessoa* como um conceito.

Gómez-Pérez e Benjamins (1999) consideram que funções e instâncias também devem compor o corpo de conhecimento, portanto os demais componentes são os seguintes:

- Um conjunto de funções que estabelecem uma relação única entre dois elementos.

Pode-se dar como exemplo a função *ser_pais_biológicos*, onde os conceitos *homem* e *mulher* estão relacionados ao conceito *pessoa*.

- Um conjunto de instâncias que fazem parte do conhecimento prévio existente na ontologia.

Uma instância nada mais é do que a concretização de uma classe, ou seja, ela representa um indivíduo único em uma hierarquia.

Considerando que a classe *SerHumano*¹ possui como subclasses as classes *Homem* e *Mulher*. Uma instância válida de *Homem* seria *João_da_Silva*.

Para torná-lo único, só pode haver um João da Silva na ontologia.

Uma ontologia é considerada como bem desenvolvida e de maior qualidade a partir do momento que o domínio com o qual se esteja trabalhando possa ser explicado e suas propriedades fundamentais dos conceitos sejam representadas explicitamente. "Quanto mais ontológica for a ontologia, melhor". (MIZOGUCHI, 2004).

Isotani e Bittencourt (2015) afirmam que ao criar uma ontologia, além da definição de termos e conceitos, deve-se:

- fazer clara diferenciação entre as funções e conceitos fundamentais;
- identificar o uso adequado das relações, principalmente as relações *part-of* e *is-a*;
- evitar a herança múltipla;
- discernir o que é atributo e o que é propriedade; e
- analisar cada decisão importante tomada, com o propósito de produzir uma ontologia de qualidade.

É possível verificar um exemplo na Figura 3.1, a qual apresenta a hierarquia, as propriedades e as relações de uma ontologia.

¹ Por convenção, os nomes das classes devem ser substantivos, considerando a primeira letra da palavra interna em maiúscula. (NOY; MCGUINNESS, 2001).

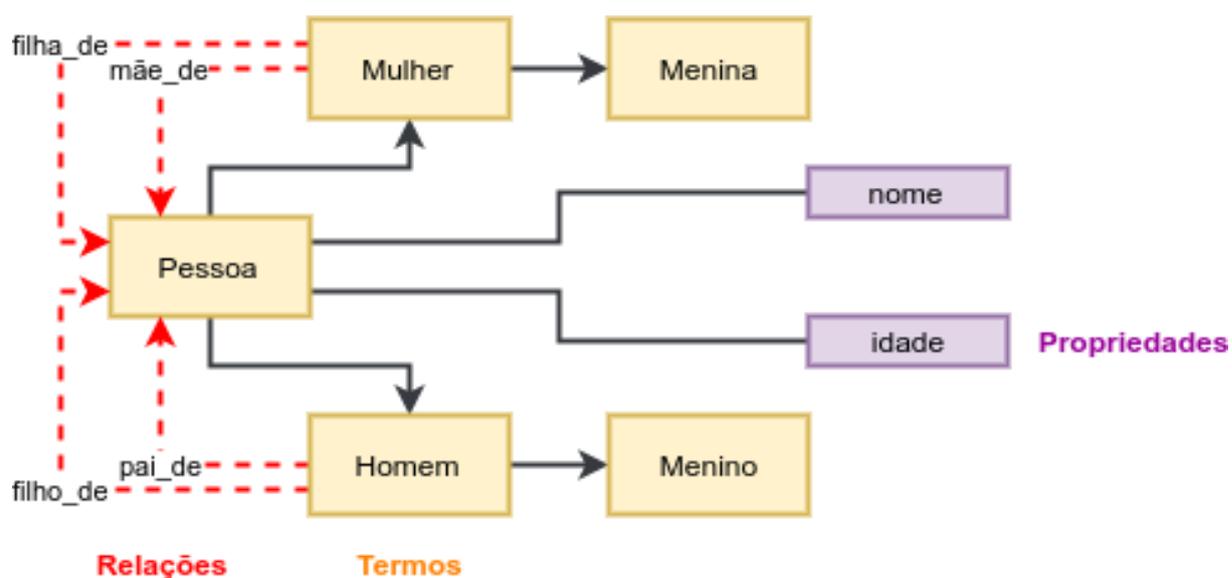


Figura 3.1 – Exemplo de Ontologia, adaptado de (VIEIRA; LOPES, 2010)

3.1.2 Classificação das Ontologias

Ontologias não apresentam sempre a mesma estrutura, entretanto é possível identificar tipos comuns e bem definidos em grande parte delas.

Almeida e Bax (2003) afirmam que mesmo não existindo um consenso entre as estruturas das ontologias, pode-se observar que os tipos apresentados possuem semelhanças entre suas funções.

Na literatura, foram propostos alguns tipos de ontologias, como aqueles que se relacionam: ao grau de formalidade do vocabulário (USCHOLD; GRUNINGER, 1996), à função (MIZOGUCHI; VANWELKENHUYSEN; IKEDA, 1995), à estrutura e conteúdo da conceitualização (HEIJST; SCHREIBER; WIELINGA, 1997) e à aplicação (JASPER; USCHOLD et al., 1999). A relação abaixo descreve cada tipo:

- **Quanto ao grau de formalidade**

- Ontologias altamente informais: remetem livremente em linguagem natural.
- Ontologias semi informais: remetem em linguagem natural de forma restrita e estruturada.
- Ontologias semi formais: remetem em linguagem artificial definida formalmente.
- Ontologias altamente formais: remetem em termos definidos com semântica formal, teoremas e provas.

- **Quanto à função**

- Ontologias de domínio: reutilizadas no domínio, proporcionam vocabulário sobre conceitos, seus relacionamentos, sobre regras e atividades que os regem.
- Ontologias de tarefa: proporcionam um vocabulário sistematizado, particularizando tarefas que podem ou não estar no mesmo domínio.
- Ontologias gerais: contém um vocabulário relacionado a eventos, tempo, espaço, comportamento etc.

- **Quanto à estrutura e conteúdo da conceitualização**

- Ontologias de alto nível: descrevem os conceitos gerais que se relacionam a todos os elementos da ontologia os quais são independentes do domínio.
- Ontologias de domínio: descrevem o vocabulário relacionado ao domínio (por exemplo: medicina, aviação).
- Ontologias de tarefa: descrevem uma atividade ou tarefa, como por exemplo diagnósticos ou compras, mediante termos especializados previamente inseridos na ontologia.
- Ontologias terminológicas: definem termos que serão utilizados a fim de representar o conhecimento em um domínio.
- Ontologias de informação: definem a estrutura das tuplas no banco de dados, como os esquemas relacionais.
- Ontologias de modelagem no conhecimento: definem conceitualizações de conhecimento, são aprimoradas para utilização no domínio do conhecimento que descrevem e possuem uma estrutura interna semanticamente rica.
- Ontologias de aplicação: contêm as definições utilizadas para modelar o conhecimento em uma aplicação.
- Ontologias genéricas: os conceitos que as definem são considerados genéricos e comuns.
- Ontologias de representação: especificam as conceitualizações que estão por trás da formalidade de representação do conhecimento.

- **Quanto à aplicação**

- Ontologias de autoria neutra: uma aplicação é desenvolvida em uma única língua e em seguida convertida para utilização em diversos sistemas, reutilizando as informações.

- Ontologias como especificação: uma ontologia é desenvolvida para um domínio específico, a qual é utilizada para documentar e dar manutenção no desenvolvimento de softwares.
- Ontologias de acesso comum à informação: esta ontologia torna a informação inteligível caso o vocabulário seja inacessível, o que proporciona o compartilhamento dos termos.

3.1.3 Vocabulários

Uma ontologia possui um vocabulário para que se possa descrever uma realidade específica. Além disso, este deve ser capaz de capturar os conceitos e relações em um determinado domínio e um conjunto de axiomas, restringindo sua interpretação.

Noy e McGuinness (2001) defendem que uma ontologia define um vocabulário comum para aqueles que necessitam compartilhar informações em um domínio, incluindo definições que possam ser interpretadas por agentes computacionais desde os conceitos básicos do domínio até a relações entre eles.

Os vocabulários devem ser criados após a entendimento do domínio, constituindo a semântica e conceitos. Compreendê-lo como um todo proporciona um conhecimento amplo das características e propriedades de suas classes. Quanto mais extenso for o vocabulário, melhor será a descrição dos relacionamentos que acontecem dentro do domínio.

3.1.4 Metodologias de Desenvolvimento

A construção de ontologias carece de um processo iterativo, com revisões constantes, a fim de sempre mantê-la como referência de qualidade. Entretanto, na literatura, há a ausência de uma metodologia unificada e respaldada por bases teóricas, sendo considerada um dos problemas em relação ao desenvolvimento de ontologias, pois cada equipe normalmente segue seus próprios princípios, critérios de design e fases ao longo do processo.

Gómez-Pérez e Benjamins (1999) apontam que caso a ontologia a ser desenvolvida seja para pequena escala, algumas atividades podem ser desconsideradas, caso contrário deve-se garantir que ela esteja completa, revisada e correta perante ao domínio. Neste último caso, uma abordagem metodológica pode auxiliar sua elaboração.

A metodologia proposta por Grüninger e Fox (1995) é baseada na experiência adquirida na construção de uma ontologia cuja modelagem é empresarial por meio da estrutura do projeto TOVE (TOronto Virtual Enterprise). Possui seis etapas:

- Captar cenários de motivação.
Com base no histórico de problemas ou meios que não foram desenvolvidos em ontologias existentes, deve-se prover uma série de possíveis cenários a fim de encontrar uma solução.
- Formular questões de competência informal.
Dado cenário motivacional, um conjunto de questões será formulado com base naquilo que a ontologia deverá responder.
- Definir a terminologia.
Uma vez formuladas as questões citadas acima, é preciso especificar a terminologia da ontologia para identificar os objetos no domínio em questão.
- Formular questões de competência formal.
Nesta etapa é distinto o que a ontologia será capaz de resolver, isto é, ela deve estar preparada para representar e prover soluções perante uma série de competências que outra ontologia não tenha conseguido resolver.
- Especificar axiomas.
É importante compreender o significado do uso dos axiomas para ter ciência das limitações da ontologia. Esta etapa seria a mais difícil, caso as questões de competência formal não tivessem sido formuladas.
- Avaliar a ontologia.
Deve-se verificar se a ontologia provê uma solução para todas as questões formuladas.

Uschold e King (1995) elaboraram uma metodologia, a qual é a mais utilizada no âmbito empresarial - baseando-se em Grüninger e Fox (1995) -, e consiste em quatro etapas:

- Identificar a proposta e o escopo da ontologia.
É preciso identificar a real necessidade do desenvolvimento da ontologia, partindo do princípio de qual é a intenção daqueles que a utilizarão.
- Construir a ontologia.
Por meio da captura e codificação do conhecimento, deve-se integrar com outras ontologias já existentes.
- Avaliar a ontologia.
- Documentar ações tomadas.
Isto é, incluir todas as técnicas, métodos, princípios e diretrizes, separando-as por estágio de desenvolvimento.

Fernández-López, Gómez-Pérez e Juristo (1997) desenvolveram uma metodologia chamada *Methontology*, a qual consiste nas seguintes etapas:

- Especificar o requisito.
Antes de construir a ontologia é necessário planejar quais serão as principais tarefas a serem realizadas, como serão organizadas, qual o tempo necessário para conclusão e quais recursos serão alocados (humanos e tecnológicos).
- Conceitualizar o domínio do conhecimento
A fim de garantir o conhecimento absoluto sobre a proposta e o escopo, as seguintes perguntas devem ser respondidas: "por que esta ontologia está sendo feita?" e "quais são os usos previstos e usuários finais?".
- Formalizar o modelo conceitual em uma linguagem formal.
Ao adquirir conhecimento suficiente, deve-se conceituar o problema e a solução.
- Implementar um modelo formal.
- Fazer a manutenção das ontologias já implementadas.
É uma etapa a qual deve ser seguida com muito cuidado, sempre se orientando por meio de diretrizes.

Por fim, Noy e McGuinness (2001) desenvolveram a Metodologia 101, a mais utilizada atualmente. Ela é composta por sete etapas:

- Determinar o domínio e o escopo da ontologia.
Algumas questões devem ser levantadas, como "qual o domínio que a ontologia cobrirá?", "quem usará e manterá a ontologia?", "quais serão os usos da ontologia?" e "quais tipos de questões a ontologia deveria responder?". É importante ressaltar que as respostas podem mudar ao longo do processo de design.
- Reutilizar ontologias existentes.
Isotani e Bittencourt (2015) defendem que esta etapa é essencial e jamais deverá ser negligenciada. O reuso também pode ocorrer por meio de conceitos específicos.
- Enumerar termos importantes na ontologia.
É importante ressaltar todos os termos que serão apresentados àqueles que utilizarão a ontologia.
- Definir as classes e sua hierarquia.
A abordagem a ser considerada nesta etapa depende da equipe de desenvolvimento e principalmente da visão que se tem do domínio.
- Definir as propriedades das classes.
As questões da primeira etapa não poderão ser respondidas sem que haja informação suficiente, considerando apenas o conjunto de classes, logo é necessário definir seus conceitos.
- Definir as restrições de cada propriedade e classe.
Nesta etapa pode-se definir a cardinalidade, os tipos de valores e os domínios abrangentes.

- Criar instâncias.

Podem aparecer instâncias obrigatórias (farão parte da própria ontologia), instâncias de testes (criadas com objetivo de testar a ontologia) e instâncias mapeadas (fonte externa de dados mapeada de acordo com a ontologia).

3.1.5 Representação de Ontologias

Há duas maneiras de representar as ontologias, sendo elas formal e gráfica. A primeira é utilizada para que as ontologias possam ser consumidas por agentes computacionais, enquanto a segunda é utilizada para compreensão humana. Isotani e Bittencourt (2015) ratificam que ambas são importantes, uma vez que a falta de uma dessas representações afeta a qualidade/uso da ontologia.

As linguagens mais populares para descrever ontologias são OWL (*Web Ontology Language*), RDF (*Resource Description Framework*) e RDF-S (*RDF Schema*), sendo estes dois últimos desenvolvimentos pelo W3C² (*World Wide Web Consortium*).

OWL é a linguagem mais utilizada para formalizar ontologias para que os agentes computacionais possam compreender e responder a consultas. (MIZOGUCHI, 2004).

O RDF permite criar triplas que contém um nó sujeito, uma relação chamada de predicado e o nó objeto. Essa tripla permite que tanto pessoas quanto agentes computacionais possam compreender esta informação. O RDF-S possui a mesma estrutura que o RDF, acrescentando a definição de *tags* e a taxionomia. Um fator limitante do RDF-S é o apoio falho ao raciocínio computacional dos dados disponíveis na Internet. (ISOTANI; BITTENCOURT, 2015).

3.2 Ontologia como apoio à SAW

A fim de satisfazer os severos requisitos de uma ontologia que apoia a SAW, ela precisa estar de acordo com três aspectos:

- Deve ser capaz de representar objetos e relacionamentos, bem como suas evoluções ao longo do tempo.
- Expressar qualquer evolução de objetos e relacionamentos.
- Possibilitar a implementação em um domínio existente.

Uma ontologia é valiosa somente se a maioria da comunidade aceitar seus principais conceitos e estrutura. (KOKAR; MATHEUS; BACLAWSKI, 2009).

² Veja <https://www.w3.org/>

Para utilizar SAW, especialmente para tomada de decisões, é preciso reconhecer diversas situações, avaliar o impacto sobre os objetivos, relacionar propriedades a situações particulares e comunicar as descrições destas às demais pessoas. Com isto, existem dois requisitos adicionais em relação às representações situacionais: (1) elas podem ser classificadas por tipos de situação, e (2) elas podem ser tratadas como objetos, físicos ou conceituais.

Um aspecto importante que não pode deixar de ser considerado é o de atributos e relacionamentos, os quais devem ser associados com valores que possam mudar ao longo do tempo. Devido as várias formas que as situações podem ser abordadas, principalmente aquelas caracterizadas como emergenciais, estas devem exigir robustez e solidez da solução de SAW.

Seria vantajoso a ontologia responder conforme o intervalo de tempo é processado. Informações recebidas sobre a posição de um jato a partir de um sistema de radar eletrônico devem ser registradas em intervalos de tempo muito curtos, assim como o estado de vida de um paciente em um hospital, enquanto que a localização de um campo minado informado por observadores humanos exige uma atualização menos frequente.

A ontologia precisa representar as relações e os valores dos atributos conforme possam ser atualizados com frequência - alta ou baixa -, quando necessário. Matheus, Kokar e Baclawski (2003) propuseram definir uma classe que deve capturar os tempos iniciais e finais, os quais um valor ou relação de atributo é válido, conforme demonstra a Figura 3.2.

Já a Figura 3.3 demonstra como os valores e relações de atributos são associados ao agrupamento de intervalos de tempo, proposta por Matheus, Kokar e Baclawski (2003). Esta abordagem garante que a captura de tempos iniciais e finais não tenham redundância.

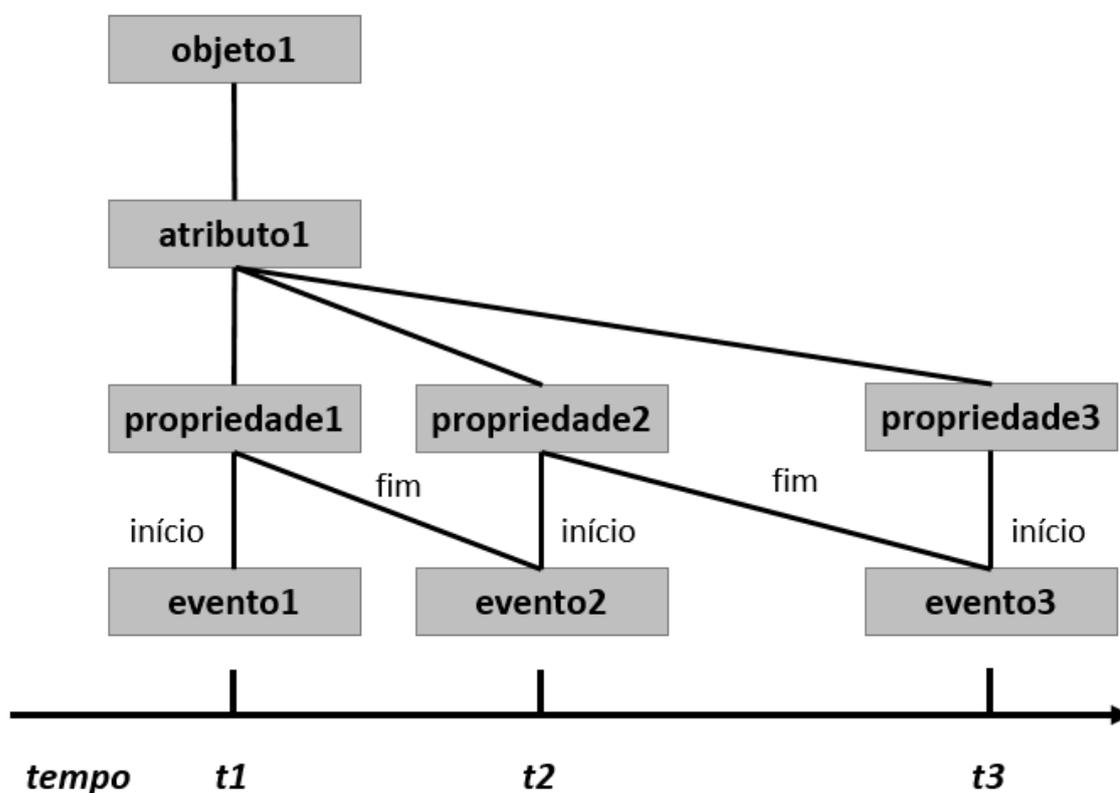


Figura 3.2 – Proposta para Captura de Intervalo de Tempo, traduzido de (MATHEUS; KOKAR; BACLAWSKI, 2003)

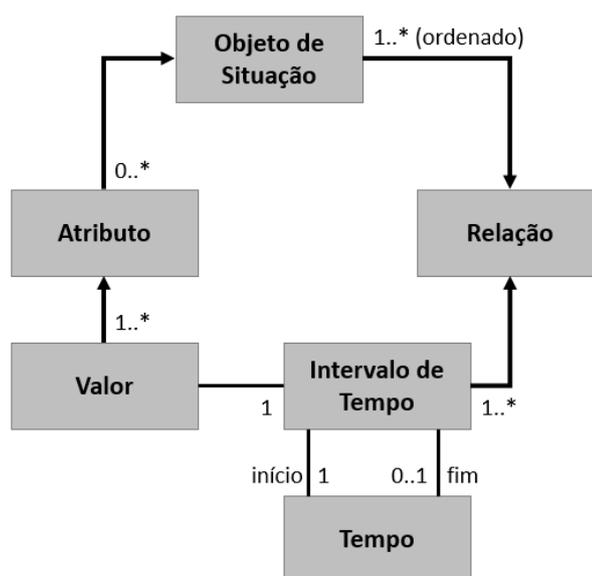


Figura 3.3 – Valores e Relações baseados em Intervalo de Tempo, traduzido de (MATHEUS; KOKAR; BACLAWSKI, 2003)

3.2.1 Ontologias de Domínio

As ontologias que se apoiam em SAW devem ser projetadas para que possam ser expandidas para atender necessidades específicas de determinado domínio.

Considere um cenário de campo de batalha que tem dois momentos os quais descrevem o avanço de dois pelotões de tanques de equipes opostas (azul e vermelha): a Figura 3.4 demonstra o primeiro, enquanto a Figura 3.5 demonstra o segundo.

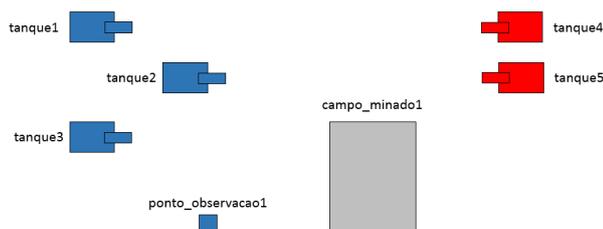


Figura 3.4 – Primeiro momento no campo de batalha, traduzido e adaptado de (MATHEUS; KOKAR; BACLAWSKI, 2003)

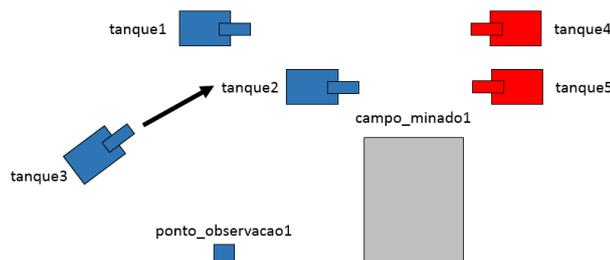


Figura 3.5 – Segundo momento no campo de batalha, traduzido e adaptado de (MATHEUS; KOKAR; BACLAWSKI, 2003)

No primeiro momento, há três tanques azuis (*tanque1*, *tanque2* e *tanque3*), um posto de observação da tropa azul (*ponto_observacao1*), um campo minado próximo às duas tropas (*campo_minado1*) e dois tanques vermelhos (*tanque4* e *tanque5*). Ambos os pelotões de tanques estão indo de encontro ao outro, entretanto caso o *tanque3* siga na direção horizontal continuamente, passará pelo *campo_minado1*.

No segundo momento, o *ponto_observacao1* deve alertar em tempo hábil o *tanque3* sobre o *campo_minado1*, o qual deverá mudar seu trajeto para que não passe pelo campo.

Desta forma, é possível notar o emprego de SAW entre aqueles que operam o ponto de observação e o terceiro tanque. A ontologia deve ser responsável por criar os objetos de situação para que defina as noções abstratas das doutrinas utilizadas no domínio em questão.

3.3 Considerações Finais

Neste Capítulo foram apresentados os conceitos de ontologia, bem como sua estrutura, sua classificação, o estudo de vocabulários, formas de representação e metodologias de desenvolvimento de ontologias.

A ontologia é uma forma de representação de dados e, portanto, ela pode ser utilizada como uma base para o desenvolvimento de outras ontologias (ontologia de topo) ou ter um domínio especificado. Esta última, quando apoiada à SAW, deve ter sua contextualização bem consolidada a fim de ter um vocabulário fiel ao domínio.

No Capítulo seguinte será apresentado os conceitos de quantificação de dados e quantificadores, bem como os processos de quantificação em modelos estruturados, sintáticos e semânticos.

Capítulo 4

QUANTIFICAÇÃO DE DADOS SEMÂNTICOS

Neste Capítulo serão abordadas as diferenças entre a quantificação em modelos semânticos e relacionais, assim como suas vantagens e desvantagens, além das possibilidades que a quantificação semântica oferece para se obter melhores subsídios para a tomada de decisão mais assertiva.

4.1 Introdução à Quantificação de Dados

A quantificação tornou-se um importante meio de análise em situações diversas que possuam uma alta demanda de dados. Ela é necessária para que haja a compreensão do todo, possibilitando comparações, avaliações, previsões e motivações para possíveis melhorias dentro do domínio de trabalho. Para tal, Basili (1985) diz que é necessário validar métricas e modelos específicos em cada ambiente com o objetivo de garantir que eles atendam a todas as características apropriadamente.

As métricas podem ser objetivas ou subjetivas. A primeira corresponde aos resultados absolutos, como por exemplo - dentro do domínio de gerenciamento de riscos - saber quantos crimes de roubo ou furto ocorreram em determinado período. Já a segunda trata-se de resultados obtidos por meio de estimativas, onde pode-se utilizar como exemplo - ainda dentro do mesmo domínio - a estimativa de crimes de roubo ou furto em locais privados dentro de uma amostragem de dados. Basili (1985) descreve as métricas subjetivas como aquelas que são utilizadas quando não se sabe como quantificar uma métrica objetiva.

A fim de auxiliar na quantificação são utilizados quantificadores lógicos que são aplicados a variáveis e a expressões. Estes são definidos por proposições, as quais podem ser simples ou compostas. Considera-se que a primeira não contenha nenhuma outra proposição em sua formação, enquanto a segunda pode possuir duas

ou mais (FILHO, 2002). São exemplos de expressões consideradas como quantificadores:

- Pelo menos um professor da universidade tem doutorado;
- Nenhum crime é perfeito;
- Algum animal é selvagem;
- Existe uma pessoa que é careca;
- Toda árvore possui folhas.

4.1.1 Quantificadores Universais

O quantificador universal, representado pelo símbolo \forall , é utilizado para representar afirmações universais, ou seja, que possam ser lidas com "para todo", "para qualquer" ou "qualquer que seja". Segundo Gluz (2002), este quantificador deve ser aplicado em uma sentença aberta $P(x)$ definida para um conjunto A . Considerando que todos os elementos de A satisfazem $P(x)$, a sentença torna-se verdadeira.

Para todo $x \in A, P(x)$ é verdadeira, ou

qualquer que seja o $x \in A$, tem-se que $P(x)$ é verdadeira.

A expressão $\forall x \in \mathbb{N}, x + 2 > x$ pode servir como exemplo de proposição, onde \forall é o quantificador, x é a variável, \mathbb{N} é o conjunto de números naturais e $x + 2 > x$ é a sentença da expressão. Ela é verdadeira, pois para qualquer que seja $x \in \mathbb{N}$, é verdade que $x + 2 > x$.

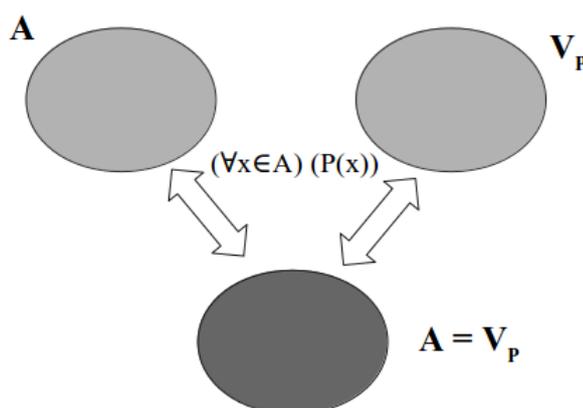


Figura 4.1 – Quantificação universal, domínio e conjunto-verdade (V_P), adaptado de (GLUZ, 2002)

A Figura 4.1 é a representação gráfica da relação entre o conjunto-verdade V_P de $P(x)$ quando V_P for igual a A , afirmando que $(\forall x \in A)(P(x)) \Leftrightarrow V_P = A$.

Enquanto $P(x)$ é uma sentença aberta, a sentença quantificada $(\forall x \in A)(P(x))$ não é mais aberta, pois se torna uma proposição simples que deve retornar um valor lógico, ou seja, verdadeiro ou falso.

4.1.2 Quantificadores Existenciais

Por sua vez, o quantificador existencial, representado pelo símbolo \exists , é utilizado para representar afirmações existenciais, ou seja, que possam ser lidas com "existe", "existe um", "existe pelo menos um" ou "para algum". Assim como o quantificador universal, o existencial deve ser aplicado em uma sentença aberta $P(x)$ definida para um conjunto A , porém, caso o conjunto não seja vazio, existe algum elemento de A que satisfará $P(x)$, tornando-se verdadeira (GLUZ, 2002).

Para algum $x \in A, P(x)$ é verdadeira, ou

existe pelo menos um $x \in A$, no qual $P(x)$ é verdadeira.

A expressão $\exists x \in \mathbb{N}, x^2 = 4$ é um exemplo de proposição que resultará em verdadeiro, onde \exists é o quantificador, x é a variável, \mathbb{N} é o conjunto de números naturais e $x^2 = 4$ é a sentença aberta da expressão.

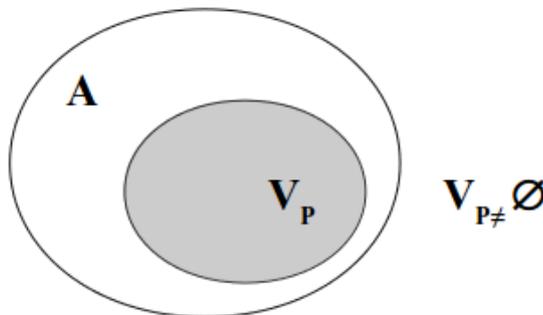


Figura 4.2 – Quantificação existencial, domínio e conjunto-verdade vazio (V_P), adaptado de (GLUZ, 2002)

A Figura 4.2 é a representação gráfica de que o conjunto-verdade V_P não pode ser vazio, afirmando que a expressão $(\exists x \in A)(P(x))$ é equivalente a $V_P \neq \emptyset$, portanto $(\exists x \in A)(P(x)) \Leftrightarrow V_P \neq \emptyset$.

Assim como quantificador universal, embora $P(x)$ seja uma sentença aberta, ao utilizar a sentença quantificada $(\exists x \in A)(P(x))$ ela não fechada, criando uma proposição simples que pode ser verdadeira ou falsa no domínio de A , dependendo de V_P ser ou não vazio.

4.2 Quantificação em Modelos Estruturados e Sintáticos

O processo de quantificação de dados é o mesmo utilizado em modelos sintáticos, estruturados e semânticos, uma vez que todos implementam a base matemática dos quantificadores lógicos apresentados conforme subseção 4.1.1 e subseção 4.1.2.

Segundo Broekstra, Kampman e Harmelen (2002), modelos sintáticos e estruturados são regradados por expressões que percorrem toda sua estrutura de dados, as quais são definidas em formato de árvore e os retornos de suas consultas são chamados de nós. Entretanto, os modelos semânticos são definidos por grafos, como por exemplo o RDF, e ambos os seus nós (objetos) e suas arestas (propriedades) possuem rótulos, os quais permitem que haja a criação de um contexto formado por metadados. O RDF abaixo servirá de exemplo para os casos de consulta em um modelo sintático ou estruturado, porém aproveitando apenas a sua estrutura XML (*eXtensible Markup Language*).

```
<rdf:Description rdf:about="http://dbpedia.org/resource/Quentin_Tarantino">
  <s:name>Quentin Tarantino</s:name>
  <s:directorOf rdf:resource="http://dbpedia.org/page/Pulp_Fiction" />
</rdf:Description>

<rdf:Description rdf:about="http://dbpedia.org/page/Pulp_Fiction">
  <s:label>Pulp Fiction</s:label>
</rdf:Description>
```

Código-Fonte 4.1 – Exemplo de RDF

Uma consulta em um modelo sintático ou estruturado não seria capaz de criar e/ou retornar a relação entre um recurso chamado "*Quentin Tarantino*" e outro chamado "*Pulp Fiction*" sem o conhecimento prévio necessário da localização dos nós desses respectivos recursos, pois a interpretação da consulta sintática seria algo como "retorne todos os elementos onde a descrição seja igual a Quentin Tarantino, onde este contenha um atributo cujo valor esteja presente em outro recurso o qual a descrição seja Pulp Fiction".

Considerando a Tabela 4.1 como o conjunto de dados sobre diretores de cinema (conjunto *D*) e a Tabela 4.2 como o conjunto sobre filmes (conjunto *F*), torna-se possível criar relações graças às colunas identificadoras (colunas "ID" de ambas as tabelas), permitindo que quantificações básicas sejam realizadas.

Caso esses dados estivessem armazenados em algum banco de dados relacional, seria possível criar proposições a fim de quantificá-los, como por exemplo:

- Total de filmes dirigidos por Quentin Tarantino;

- Total de filmes que possam ter sido dirigidos por Robert Rodriguez ou Guy Ritchie.

ID	Nome
1	Quentin Tarantino
2	Robert Rodriguez
3	Guy Ritchie
4	Joseph Lawrence

Tabela 4.1 – Diretores de Cinema com Identificação - Conjunto D / Tabela "diretores"

ID	Diretor	Filme
1	1	Pulp Fiction
2	1	The Hateful Eight
3	1	Kill Bill: Volume 1
4	1	Kill Bill: Volume 2
5	1	Inglourious Basterds
6	2	El Mariachi
7	2	Sin City
8	2	From Dusk Till Dawn
9	3	The Man from U.N.C.L.E.

Tabela 4.2 – Filmes de Cinema com Identificação - Conjunto F / Tabela "filmes"

O primeiro exemplo pode ser representado pela expressão $(\forall x \in F)(P(x))$, onde x é o diretor, a qual deve ser igual a "Quentin Tarantino", o que satisfaz o conjunto-verdade $V_P = \{ "PulpFiction", "TheHatefulEight", \dots \}$. Nesse caso, a expressão utiliza o quantificador universal, uma vez que a proposição deixa claro que deseja retornar *todos* os filmes dirigidos por Quentin Tarantino. Já o segundo exemplo, agora utilizando o quantificador existencial, é representado pela expressão $(\exists y \in F)(Q(y))$, onde y , assim como x na expressão anterior, é o diretor a ser retornado, que deve ser "Robert Rodriguez" ou "Guy Ritchie".

As consultas SQL (*Structured Query Language*) para essas duas situações seriam:

```
# Exemplo 1
SELECT D.Nome AS Diretor, COUNT(F.Id) AS "Total Filmes"
FROM filmes AS F
INNER JOIN diretores AS D ON (F.Diretor = D.Id)
WHERE D.Id = 1
GROUP BY D.Id;
```

Código-Fonte 4.2 – Consulta SQL do Exemplo 1

```
SELECT D.Nome AS Diretor, COUNT(F.Id) AS "Total Filmes"  
FROM filmes AS F  
INNER JOIN diretores AS D ON (F.Diretor = D.Id)  
WHERE D.Id IN (2, 3)  
GROUP BY D.Id;
```

Código-Fonte 4.3 – Consulta SQL do Exemplo 2

Os resultados retornados são apresentados na Tabela 4.3 e Tabela 4.4, respectivamente.

Diretor	Total Filmes
Quentin Tarantino	5

Tabela 4.3 – Resultado da Quantificação Sintática - Exemplo 1

Diretor	Total Filmes
Robert Rodriguez	3
Guy Ritchie	1

Tabela 4.4 – Resultado da Quantificação Sintática - Exemplo 2

Apesar dos resultados quantificados estarem corretos, foi necessário ter o conhecimento prévio daquilo que se desejava saber (a quantidade de filmes dirigidos por Quentin Tarantino e que foram dirigidos por Robert Rodriguez ou Guy Ritchie), portanto todas as métricas quantitativas previamente definidas foram objetivas.

Broekstra, Kampman e Harmelen (2002) afirmam que modelos sintáticos e estruturados não aproveitam totalmente os metadados para facilitar consultas, o que possibilitaria a ampliação da gama de quantificações a serem realizadas.

4.3 Quantificação em Modelos Semânticos

Diferente dos modelos citados acima, a quantificação em modelos semânticos permite que além de utilizar os dados instanciados em uma ontologia, por exemplo, também é possível aproveitar os metadados de classes, propriedades e atributos a fim de criar métricas objetivas e subjetivas.

A partir do momento que é viável utilizar os metadados de recursos para se criar consultas e extrair valores quantificáveis, os resultados são mais flexíveis pelo fato de serem semânticos, portanto estão dentro de um contexto. Este contexto faz referência a um conjunto de instâncias com metadados comuns (MACGREGOR; KO, 2003).

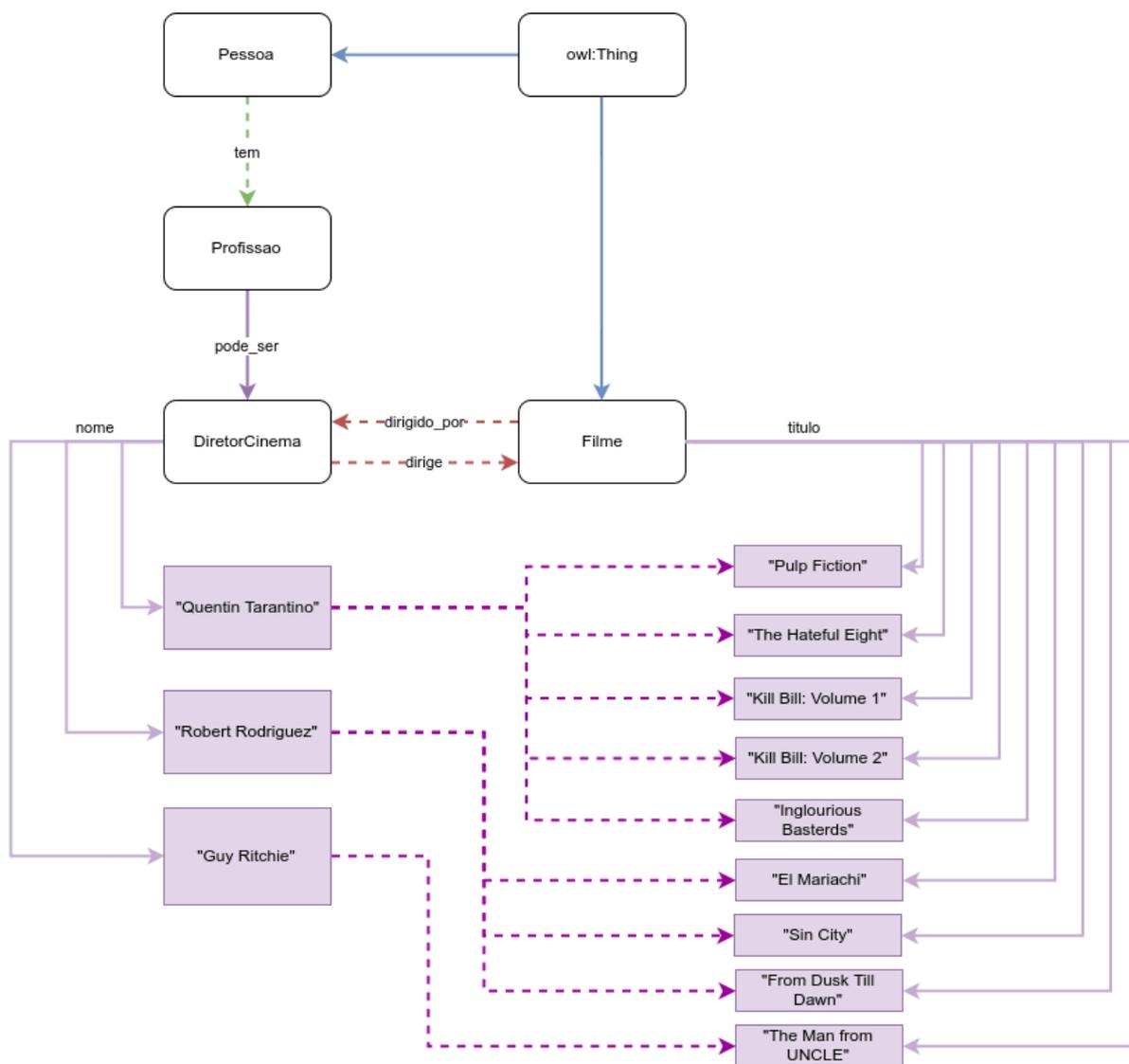


Figura 4.3 – Representação de Ontologia de Filme para demonstração de exemplos, elaborado pelo próprio autor

A Figura 4.3 representa uma ontologia a qual servirá para demonstração dos exemplos de quantificação semântica. As classes definidas são Pessoa, Profissão ("Profissao"), Diretor de Cinema ("DiretorCinema") e Filme. A relação entre Pessoa e Profissão se dá pela propriedade "tem", onde esta última classe também permite que exista a relação de que "pode ser"um Diretor de Cinema. Além disso, há a relação de um Diretor de Cinema dirigir um Filme, assim como a relação oposta também ser válida. Os retângulos roxos representam as instâncias de Diretor de Cinema com o atributo "nome"e Filme com o atributo "titulo", respectivamente.

Utilizar o contexto durante a quantificação em modelos semânticos é vantajoso pelo fato de se aproveitar os metadados dos recursos disponíveis. MacGregor e Ko (2003) afirmam que é aconselhável identificar um conjunto de predicados que forneçam um padrão para representar os demais metadados mais comuns. No caso, os

predicados presentes na ontologia são os que formam as relações entre as classes.

São passíveis de quantificação as proposições abaixo:

- Total de pessoas que dirigiram filmes;
- Total de filmes dirigidos por pessoas cuja profissão seja diretor de cinema;
- Total de filmes dirigidos por Quentin Tarantino;
- Total de filmes que possam ter sido dirigidos por Robert Rodriguez ou Guy Ritchie.

Eis abaixo as consultas SPARQL (*Simple Protocol and RDF Query Language*) que permitirão que as proposições acima tornem-se verdadeiras, aproveitando inclusive os predicados. Os resultados são demonstrados abaixo de cada consulta em formato de comentário, linha a qual é precedida por uma cerquilha.

```
1 SELECT (COUNT(DISTINCT ?diretorCinema) AS ?total)
2 WHERE {
3     ?pessoa _:tem ?profissao .
4     ?profissao _:pode_ser ?diretorCinema .
5 }
6 ORDER BY DESC(?total)
7 # Resultado: 3
```

Código-Fonte 4.4 – Consulta SPARQL do Exemplo 1

```
1 SELECT (COUNT(DISTINCT ?filme) AS ?total)
2 WHERE {
3     _:Pessoa _:tem ?profissao .
4     ?profissao _:pode_ser ?diretorCinema .
5     ?filme _:dirigido_por ?diretorCinema .
6 }
7 GROUP BY ?filme
8 ORDER BY DESC(?total)
9 # Resultado: 9
```

Código-Fonte 4.5 – Consulta SPARQL do Exemplo 2

```
1 SELECT (COUNT(DISTINCT ?diretor) AS ?total)
2 WHERE {
3     ?filme _:dirigido_por ?diretor .
4     ?diretor _:nome 'Quentin Tarantino' .
5 }
6 GROUP BY ?diretor
7 ORDER BY DESC(?total)
8 # Resultado: 5
```

Código-Fonte 4.6 – Consulta SPARQL do Exemplo 3

```
1 # Exemplo 4
2 SELECT (COUNT(DISTINCT ?diretor) AS ?total)
3 WHERE {
4     ?filme _:dirigido_por ?diretor .
5     OPTIONAL {
6         ?diretor _:nome 'Robert Rodriguez' .
7     } .
8     OPTIONAL {
9         ?diretor _:nome 'Guy Ritchie' .
10    }
11 }
12 GROUP BY ?diretor
13 ORDER BY DESC(?total)
14 # Resultado de Robert Rodriguez: 3
15 # Resultado de Guy Ritchie: 1
```

Código-Fonte 4.7 – Consulta SPARQL do Exemplo 4

Os resultados gerados pelo terceiro e quarto exemplos devem ser iguais em relação aos resultados dos exemplos sintáticos, não somente por se tratarem dos mesmos dados, mas para assegurar que haja consistência nas informações geradas. Nessa situação, é notável o quão importante é a contextualização daquilo que se deseja quantificar.

As consultas semânticas são capazes de organizar dentro de uma hierarquia os contextos dos recursos os quais estão sendo analisados, facilitando a compreensão daquele que necessita de dados quantificados com o propósito de obter a melhor ação para a tomada de decisão.

4.4 Considerações Finais

Neste Capítulo foram apresentados e demonstrados os conceitos de quantificação de dados utilizando os quantificadores universal e existencial, assim como a forma

que eles interagem com as proposições propostas a fim de obter resultados quantificados. Dados quantificados geram informações que viabilizam a aplicação das devidas ações nas tomadas de decisões.

No Capítulo seguinte será apresentado o processo de análise quantitativa de eventos criminais utilizando modelos semânticos com o objetivo de apoiar a SAW de operadores de sistemas de gerenciamento de riscos com base em um estudo de caso visando a validação do trabalho.

Capítulo 5

PROCESSO DE ANÁLISE QUANTITATIVA DE EVENTOS CRIMINAIS UTILIZANDO ABORDAGEM SEMÂNTICA

Neste Capítulo serão abordados a entrevista com a PMESP para levantamento e identificação de requisitos utilizando a metodologia GDTA (*Goal-Directed Task Analysis*), o desenvolvimento da ontologia de domínio para o gerenciamento de riscos, declaração de instâncias para formar o estudo de caso e o processo de aplicação da análise quantitativa nesse estudo.

5.1 Estado da Arte em Abordagens de Análise Quantitativa em Modelos Semânticos

Ao decorrer da revisão bibliográfica foi constatado que recentemente o interesse por modelos semânticos cresceu, principalmente pelo uso e implementação de ontologias, as quais fornecem uma estrutura com que se pode trabalhar com o contexto do domínio. A aplicação da quantificação nesses modelos é um meio de se obter resultados mais precisos, porém é um assunto pouco retratado por ser complexo e/ou variar de domínio para domínio.

Noy e McGuinness (2001) propõem em seu trabalho uma metodologia para o desenvolvimento de ontologias de topo, que contém informações relevantes e controladas sobre classes e vocabulários, e de domínio, as quais são capazes de serem aplicadas e utilizadas em sistemas. Foi discutido a respeito do uso de ontologias como base de conhecimento a partir de instâncias individuais, sendo definidas a partir de classes, organizadas em taxonomias, as quais devem ter em sua hierarquia propriedades significativas. Este trabalho é considerado importante por nortear a elaboração

deste trabalho, uma vez que esta foi a metodologia adotada para o desenvolvimento da ontologia aqui utilizada.

Asaro et al. (2003) modelaram uma ontologia de domínio para atender aos critérios das leis criminais italianas, sendo um dos objetivos a identificação e sugestão de hipóteses dirigidas ao juiz, além da tentativa de acoplar os comportamentos dos envolvidos durante o julgamento. Entretanto, esta ontologia não chegou a ser de fato implementada, considerando que até o presente momento não houve nenhuma publicação que diz respeito ao seu desenvolvimento.

Baumgartner et al. (2010) desenvolveram uma ontologia de domínio apoiada em SAW a fim de prover dados relevantes para sistemas de gerenciamento de trânsito. Seu diferencial é a introdução do conceito de espaço-tempo entre os relacionamentos das propriedades, o que provê a reusabilidade das mesmas, uma vez que estas podem fornecer informações relevantes para os operadores, evoluindo seu nível de SAW. Os autores contaram com a colaboração de um especialista no domínio para definir as situações de interesse, além de auxiliar na elaboração de vocabulários. A ontologia proposta é capaz de lidar com situações de engarrafamentos e incidentes, como acidentes e motoristas que dirigem na contramão. Testes de performance foram feitos após diversos estudos de casos terem sido realizados comprovando a eficiência do modelo.

Hepp, Leukel e Schmitz (2007) apresentaram uma metodologia de análise quantitativa em modelos semânticos utilizando pesos e valores, levando em consideração o grau de especificidade dos vocabulários. Os atributos e propriedades recebem um peso semântico com base em sua relevância. A análise é fundamentada em que uma propriedade, quando muito utilizada, é geralmente menos específica do que uma propriedade pouco utilizada. Os valores são calculados somando todos os pesos e atribuídos às classes, que são iguais aos valores recíprocos de suas frequências de uso. Esta metodologia, porém, é útil quando o domínio não possui situações que possam interferir em outras.

Com base nesses e outros trabalhos que serviram como base bibliográfica para o desenvolvimento deste, é possível identificar que certas metodologias não atendem a todos os domínios devido o grau de granularidade, como é o caso da metodologia proposta por Hepp, Leukel e Schmitz (2007). Fazendo uma analogia ao tema abordado no presente trabalho, uma situação de roubo é diferente de uma situação de furto e, portanto, sua relevância é variável para cada caso, não podendo ter um peso semântico fixado em suas propriedades.

5.2 Metodologia

Para cumprir o objetivo proposto pelo trabalho, a metodologia aplicada se apoia em: (1) empregar a metodologia GDTA (*Goal-Directed Task Analysis*) para revelar informações úteis a respeito do domínio; (2) avaliação de vocabulários e associação de relacionamentos semânticos; (3) desenvolvimento da ontologia de domínio; (4) declaração de instâncias para validação; e (5) elaboração do estudo de caso aplicando a quantificação ao realizar consultas SPARQL. O fluxograma apresentado na Figura 5.1 representa as etapas do processo.

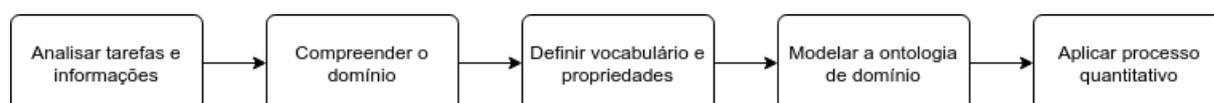


Figura 5.1 – Fluxograma da metodologia utilizada, elaborado pelo próprio autor

5.2.1 Levantamento de Requisitos

A primeira etapa a ser realizada é empregar a metodologia de Análise de Tarefas Dirigidas por Objetivos (*Goal-Directed Task Analysis - GDTA*), útil para revelar as informações necessárias para que operadores humanos tomem decisões ao longo do uso do sistema de avaliação de situações de risco, bem como as tarefas que os mesmos devem realizar.

Esta metodologia se concentra na obtenção de uma descrição precisa dos requisitos de SAW, não apenas focada em quais dados os operadores precisam, mas sobre como os dados devem ser combinados e integrados entre si para apoiar a tomada de decisão (BOLSTAD et al., 2002).

Para entender e conhecer o domínio de gerenciamento de riscos, um questionário foi elaborado e encaminhado a PMESP para que fosse respondido. A partir das respostas obtidas foi possível identificar quais informações eram relevantes para o operador, gerando assim o GDTA. Parte da análise é apresentada na Figura 5.2, a qual demonstra um de seus objetivos, que é a identificação de problemas criminais.

5.2.2 Análise de Vocabulários

Conforme descrito na subseção 3.1.3, após compreender o domínio os vocabulários devem ser criados a fim de descrever os relacionamentos que podem ocorrer dentro do domínio. Os termos foram selecionados de acordo com a sua relevância perante o operador do sistema de gerenciamento de riscos. O vocabulário foi construído baseado em registros oficiais das bases de dados da PMESP, sendo validado junto

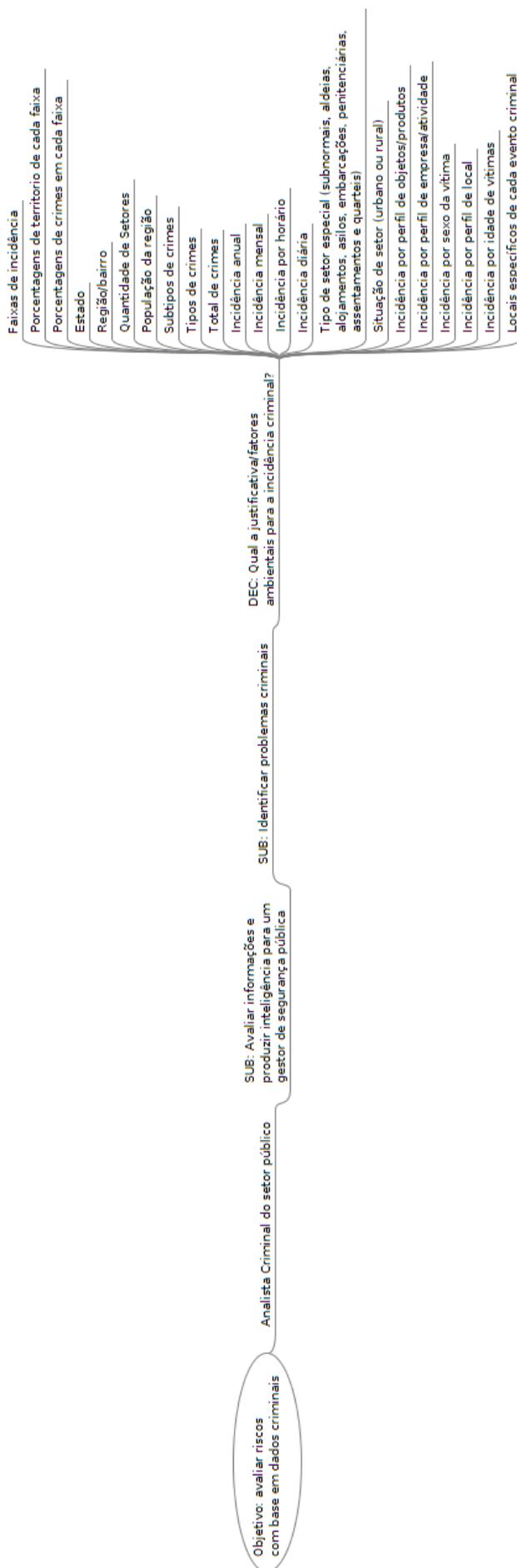


Figura 5.2 – Parte da análise de tarefas dirigida a objetivos gerado a partir dos resultados obtidos no questionário encaminhado à PMESP

a um especialista criminal, estabelecendo um nível de granularidade relevante para o domínio.

A Figura 5.3a e Figura 5.3b apresentam os vocabulários utilizados para os relacionamentos entre objetos (classes) e as propriedades de dados (instâncias), respectivamente.

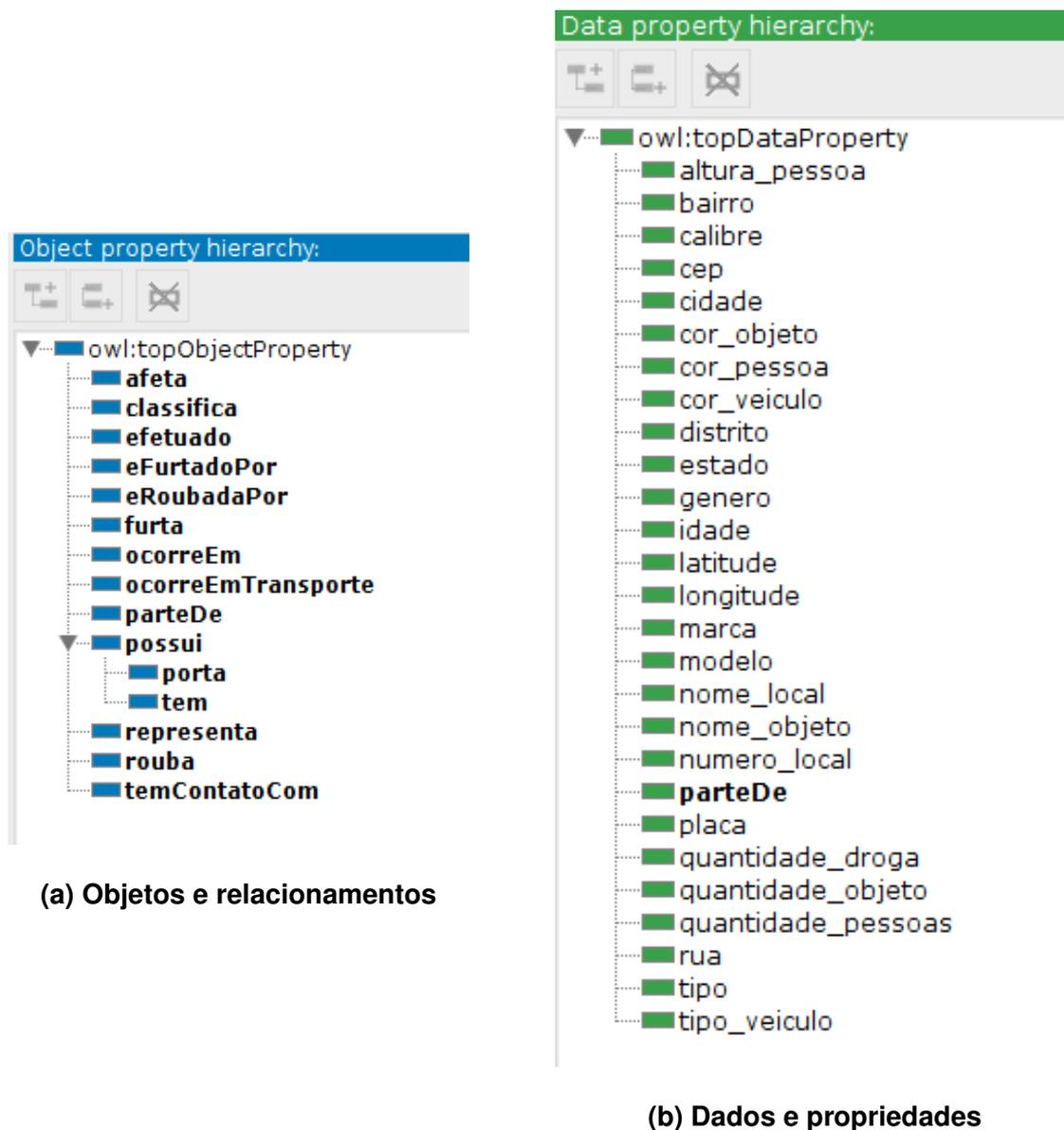


Figura 5.3 – Vocabulários utilizados de acordo com a relevância no domínio

5.2.3 Desenvolvimento de Ontologia de Domínio

A metodologia empregada para o desenvolvimento da ontologia de domínio aplicada neste trabalho foi a Metodologia 101 de Noy e McGuinness (2001), a qual foi abordada e conceituada no Capítulo 3 e é representada pela Figura 5.4.

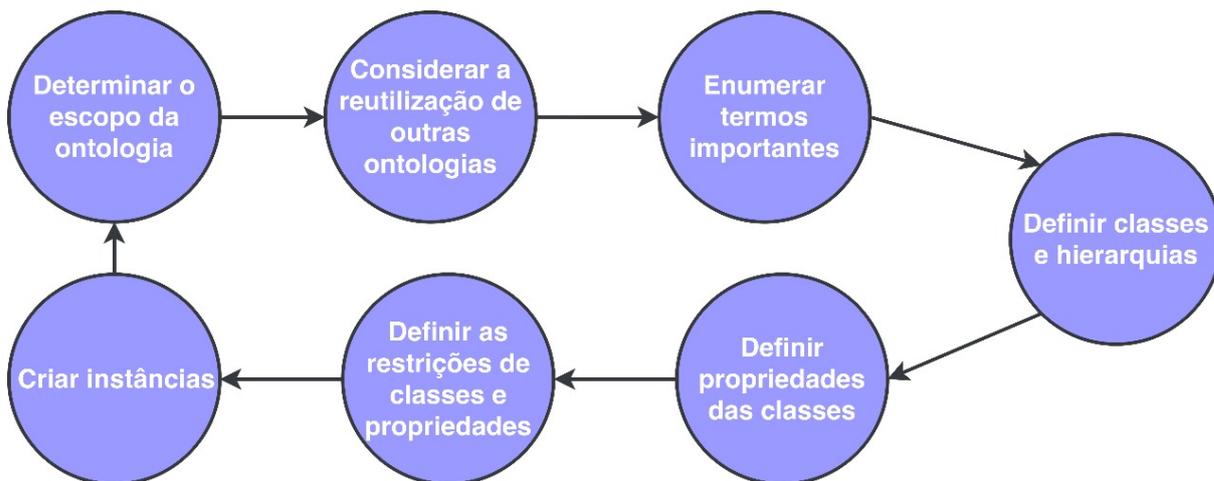


Figura 5.4 – Ciclo da Metodologia 101 de Noy e McGuinness, traduzido e adaptado de (NOY; MCGUINNESS, 2001)

Baseando-se no ciclo de sete etapas dessa metodologia, eis a descrição de cada etapa:

1. Determinar o escopo da ontologia.
O domínio de aplicação e estudo dessa ontologia são situações de risco, para este trabalho situações de roubo e furto especificamente.
2. Considerar o reuso de outras ontologias.
Nas pesquisas realizadas não foi encontrada nenhuma ontologia sobre o domínio que auxiliasse no desenvolvimento desta, principalmente considerando a utilização da Língua Portuguesa e a organização de Problemas Criminais adotada.
3. Enumerar termos relevantes.
A realização desta etapa foi através da análise das respostas obtidas do questionário aplicado à operadores de sistemas críticos, além de especialistas de segurança pública e privada, junto a uma análise de denúncias presentes em bases oficiais, assim foi feito um levantamento dos principais termos encontrados e estes passaram pela validação de um especialista.
4. Definir classes, hierarquias e propriedades das classes.
Baseado no questionário aplicado, foi possível identificar as informações essenciais para caracterizar uma situação de risco no domínio criminal. Com o resultado desse questionamento, foram definidas as classes principais como: Criminoso, Vítima, Local e Objeto. Dentro de cada uma dessas foram criadas classes filhas mais específicas assim como as propriedades que as caracterizam.
5. Definir restrições de classes e propriedades.
Com a definição das classes e propriedades foi necessário definir as restrições do domínio, como por exemplo: "vítima pode ter um objeto", a qual recebe a

cardinalidade 1 – N dentro da ontologia;

6. Criar instâncias.

A ontologia foi alimentada com diversas instâncias que caracterizam as informações do domínio, estas serão apresentadas na próxima seção.

A ontologia desenvolvida é apresentada na Figura 5.5.

Sabe-se que esta ontologia de domínio passará por diversas modificações e melhorias com o propósito de inferir melhores resultados e, portanto, o ciclo citado acima se repetirá a partir da segunda etapa. A ontologia deve ser atualizada de acordo com a necessidade do grau de informação a ser recuperado.

5.3 Estudo de Caso

Ao decorrer deste trabalho, mais especificamente após o desenvolvimento da ontologia, um ambiente simulado e controlado foi utilizado para que as instâncias fossem criadas a partir de registros oficiais fornecidos pela PMESP. Tais registros permitiram que este trabalho fosse validado, possibilitando que operadores de sistemas críticos utilizassem os resultados quantificados, evoluindo sua SAW e aprimorando a tomada de decisão.

Algumas ocorrências são demonstradas abaixo, as quais fizeram parte do conjunto para a análise quantitativa.

"Comparece neste plantão policial a vítima informando que deixou seu veículo estacionado e ao retornar constatou que haviam furtado do interior do veículo uma pochete contendo os seguintes documentos: RG, CIC, título de eleitor, cartão de crédito e conta corrente do Banco Itaú, cartão de crédito do Banco Real, cartão de crédito do Banco Unibanco, cartão do Carrefour, Rener, Riachuelo, C&A, Sorocred, um talão de cheque do Banco Real e Unibanco, um aparelho de CD e 48 CDs e documento do veículo. Nada mais."

"Presente os representantes das empresas vítimas informando que na data dos fatos, oito indivíduos encapuzados e portando armas de fogo adentraram o local, que é uma área da Prefeitura onde estão sendo construídas casas e, mediante grave ameaça, subtraíram 52 rolos de cabo flexível preto 1.5mm, 60 rolos de cabo flexível azul 1.5mm, 30 rolos de cabo flexível verde 4.0mm, 72 interruptores 2p+tomada, 86 interruptores simples, 72 tomadas para telefone, 72 interruptores duplos, tudo isso avaliado em R\$ 14.576,00, além de quatro aparelhos

Nextel. Nada mais."

"Comparece neste plantão policial, a vítima informando que estacionou seu veículo na garagem de sua residência na noite de ontem e na data de hoje, ao sair, notou que o mesmo havia sido violado e de seu interior fora subtraído sua carteira de couro, cor preta, contendo os seguintes documentos: RG, CPF, CNH, documentos do veículo Ford Fiesta, título de eleitor, cartão de crédito Visa Banespa, cartão de crédito American Express, cartão cidadão da CEF, cartão Carrefour, talão de cheques do Banespa, carteira de registro no MEC. Nada mais."

"Comparece a vítima alegando que meliantes mediante arrombamento da porta do local dos fatos adentraram o mesmo e subtraíram um vídeo cassete da marca Panasonic, um videogame Tok Toy, um micro system da marca CCE, roupas diversas e alguns pares de tênis. Nada mais."

5.3.1 Criação das Instâncias

Além das ocorrências apresentadas acima, outras também tiveram suas instâncias criadas com os seus respectivos relacionamentos entre os objetos, o essencial para que consultas básicas sejam criadas. A Figura 5.6 representa a instância de uma das ocorrências, mais especificamente a primeira apresentada nos exemplos acima.

A ocorrência instanciada acima foi classificada como Furto¹ e, portanto, é possível observar que a Vítima "tem"² um Veículo e vários Documentos e foi "furtada" por um Criminoso. Esta situação permitiu que fossem criadas as relações de que um Furto "afeta" uma Vítima e "é efetuado por" um Criminoso. A maioria das ocorrências disponíveis envolveram roubo ou furto de documentos e, portanto, o estudo de caso deste trabalho será apoiado nessas instâncias.

5.3.2 Definição de Proposições

Conforme apresentado no Capítulo 4, o primeiro passo para se obter dados quantificados é a definição de proposições, as quais tornam-se o objetivo de consulta com o propósito de resgatar as informações necessárias.

Neste estudo de caso, as proposições abaixo serão demonstradas baseadas no perfil de ocorrências, ou seja, características relevantes para um operador:

¹ Os sujeitos estão declarados como maiúsculo com o propósito de representar as classes as quais foram instanciadas.

² Está entre aspas para representar os predicados da relação entre os objetos, como por exemplo "tem", "furtado por", "possui" etc.

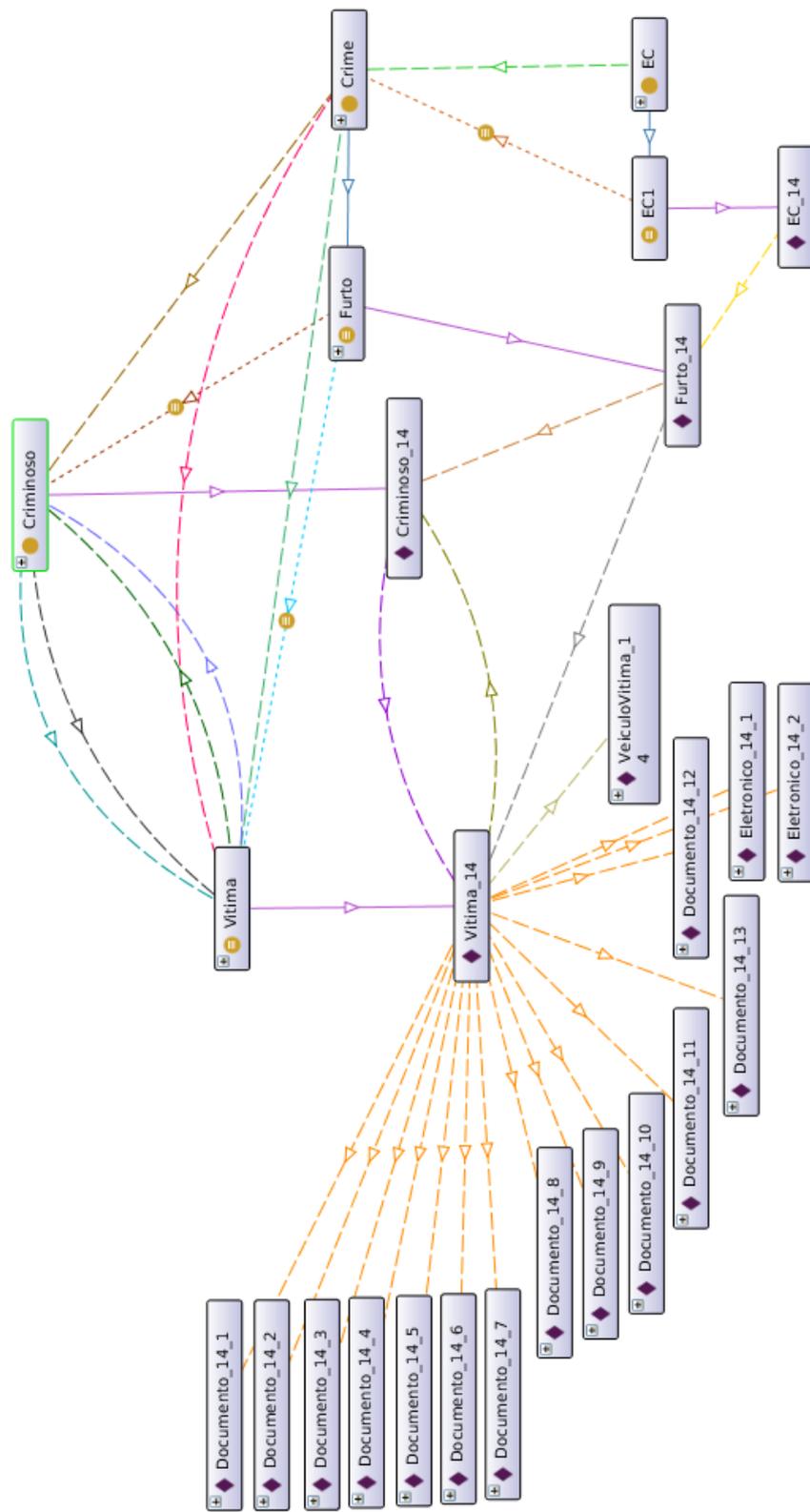


Figura 5.6 – Exemplo de Instância de Ontologia

1. Qual a relação das situações (ou seja, furtos e roubos) onde a vítima possui algum objeto?
2. Qual a quantidade de locais públicos ou privados das situações de furtos e roubos?
3. Qual a relação das situações onde uma vítima teve seu veículo roubado ou furtado em um local público?
4. Qual a relação de situações onde há objetos roubados ou furtados?

As proposições serão utilizadas para construir as suas respectivas consultas SPARQL, a fim de obter os resultados esperados.

5.3.3 Construção de Consultas SPARQL

As consultas SPARQL possibilitam a obtenção de resultados os quais podem ser quantificados e analisados, permitindo que os operadores de sistemas de gerenciamento de riscos tomem a decisão mais assertiva. Baseando-se nas proposições da subseção acima, serão apresentados duas consultas para cada proposição, sendo a primeira o retorno dos dados e a segunda a quantificação dos mesmos a fim de comprovar sua origem, além do resultado da análise quantitativa.

Considere os prefixos abaixo para todas as consultas apresentadas ao longo deste estudo de caso.

```
1 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
2 prefix dados: <http://localhost/crimevoc#>
3 PREFIX furto: <http://localhost/furto#>
4 PREFIX roubo: <http://localhost/roubo#>
5 prefix local: <http://localhost/local#>
```

Código-Fonte 5.1 – Prefixos utilizados para as consultas SPARQL no estudo de caso

5.3.3.1 Proposição 1 - Relação das situações onde a vítima possui algum objeto

As consultas de retorno geral e quantificado utilizadas para atender a proposição são apresentadas a seguir:

```

1 # Resultado geral.
2 SELECT ?tipo_situacao ?tipo_objeto
3 WHERE {
4   ?situacao dados:afeta ?vitima .
5   ?situacao rdf:type ?tipo_situacao .
6   FILTER regex(str(?tipo_situacao), '(furto|roubo)+#', 'i') .
7   ?vitima dados:possui|dados:tem ?objeto .
8   ?objeto rdf:type ?tipo_objeto .
9   FILTER regex(str(?tipo_objeto), 'objeto#', 'i') .
10 }
11 ORDER BY ASC(?tipo_situacao)

```

Código-Fonte 5.2 – Consulta SPARQL da Proposição 1 para retornar a relação de situações onde a vítima possui algum objeto

```

1 # Resultado quantificado.
2 SELECT ?tipo_situacao (count(distinct ?objeto) as ?total)
3 WHERE {
4   ?situacao dados:afeta ?vitima .
5   ?situacao rdf:type ?tipo_situacao .
6   FILTER regex(str(?tipo_situacao), '(furto|roubo)+#', 'i') .
7   ?vitima dados:possui|dados:tem ?objeto .
8   ?objeto rdf:type ?tipo_objeto .
9   FILTER regex(str(?tipo_objeto), 'objeto#', 'i') .
10 }
11 GROUP BY ?tipo_situacao
12 ORDER BY ASC(?tipo_situacao)

```

Código-Fonte 5.3 – Consulta SPARQL da Proposição 1 para retornar a quantidade por situações onde a vítima possui algum objeto

tipo_situacao	total
furto:Furto	"36"^^ xsd:integer
roubo:Roubo	"19"^^ xsd:integer

Tabela 5.1 – Resultado Quantificado da Proposição 1

Na Tabela 5.1 nota-se o resultado quantificado por cada situação, ou seja, trinta e seis documentos furtados e dezenove roubados, totalizando cinquenta e cinco documentos pertencentes às vítimas, independentemente de quem seja. Não foram retornadas as situações de crime em si, mas sim o tipo de situação, ou seja, a classe do objeto instanciado.

5.3.3.2 Proposição 2 - Quantidade de locais públicos ou privados das situações de furtos e roubos

As consultas de retorno geral e quantificado utilizadas para atender a proposição são apresentadas a seguir:

```

1 # Resultado geral.
2 SELECT ?tipo_local ?local
3 WHERE {
4   ?situacao dados:ocorreEm ?local .
5   ?local rdf:type ?tipo_local .
6   FILTER regex(str(?tipo_local), 'local#', 'i') .
7 }
8 ORDER BY ASC(?tipo_local)

```

Código-Fonte 5.4 – Consulta SPARQL da Proposição 2 para retornar a relação de locais públicos ou privados das situações de furtos e roubos

```

1 # Resultado quantificado.
2 SELECT ?tipo_local (count(distinct ?local) as ?total)
3 WHERE {
4   ?situacao dados:ocorreEm ?local .
5   ?local rdf:type ?tipo_local .
6   FILTER regex(str(?tipo_local), 'local#', 'i') .
7 }
8 GROUP BY ?tipo_local
9 ORDER BY ASC(?tipo_local)

```

Código-Fonte 5.5 – Consulta SPARQL da Proposição 2 para retornar a quantidade por locais públicos ou privados das situações de furtos e roubos

tipo_local	total
local:Comercio	"3"^^ xsd:integer
local:Privado	"2"^^ xsd:integer
local:Publico	"4"^^ xsd:integer
local:Residencial	"3"^^ xsd:integer

Tabela 5.2 – Resultado Quantificado da Proposição 2

É notável que o resultado quantificado não trouxe somente a quantidade de situações que ocorreram em locais públicos ou privados cujas classes de instância são Público e Privado, respectivamente. Os resultados de dados retornados na Tabela 5.2 demonstram que as classes de instâncias são Comércio e Residencial, pois estas são subclasses de Local Privado. A semântica identifica o relacionamento hierárquico e o retorna igualmente, a fim de prover dados mais significativos.

5.3.3.3 Proposição 3 - Relação de vítimas com veículos furtados ou roubados em locais públicos

As consultas de retorno geral e quantificado utilizadas para atender a proposição são apresentadas a seguir:

```

1 # Resultado geral.
2 SELECT ?vitima ?veiculo
3 WHERE
4 {
5   ?situacao dados:afeta ?vitima .
6   ?situacao rdf:type ?tipo_situacao .
7   FILTER regex(str(?tipo_situacao), '(furto|roubo)+#', 'i') .
8   ?situacao dados:ocorreEm ?local .
9   ?local rdf:type local:Publico .
10  ?vitima dados:possui ?veiculo .
11  ?veiculo rdf:type objeto:Veiculo .
12 }

```

Código-Fonte 5.6 – Consulta SPARQL da Proposição 3 para retornar a relação de vítimas com veículos furtados ou roubados em locais públicos

```

1 # Resultado quantificado.
2 SELECT ?tipo_situacao (count(?situacao) as ?total)
3 WHERE
4 {
5   ?situacao dados:afeta ?vitima .
6   ?situacao rdf:type ?tipo_situacao .
7   FILTER regex(str(?tipo_situacao), '(furto|roubo)+#', 'i') .
8   ?situacao dados:ocorreEm ?local .
9   ?local rdf:type local:Publico .
10  ?vitima dados:possui ?veiculo .
11  ?veiculo rdf:type objeto:Veiculo .
12 }
13 GROUP BY ?tipo_situacao
14 ORDER BY ASC(?tipo_situacao)

```

Código-Fonte 5.7 – Consulta SPARQL da Proposição 3 para retornar a quantidade por locais públicos onde vítimas tiveram seus carros furtados ou roubados

tipo_situacao	total
furto:Furto	"1"^^ xsd:integer
roubo:Roubo	"1"^^ xsd:integer

Tabela 5.3 – Resultado Quantificado da Proposição 3

Dentro da amostragem analisada no ambiente controlado, apenas duas vítimas tiveram seus veículos furtados ou roubados em um local público, conforme a Tabela 5.3. Nesse contexto pode-se concluir que nem todas as ocorrências de furto ou roubo de veículo possuem um local instanciado na ontologia.

Caso as linhas 8 e 9 do resultado quantificado fossem comentadas, novos dados seriam retornados. Nesse segundo exemplo, o dado a ser retornado para a situação de furto seria oito, não alterando a quantidade de situações de roubo. Remover essas linhas da consulta é o mesmo que retirar a condição de que obrigatoriamente uma situação de furto ou roubo deve ocorrer em um local público.

5.3.3.4 Proposição 4 - Relação de situações onde há objetos roubados ou furtados

As consultas de retorno geral e quantificado utilizadas para atender a proposição são apresentadas a seguir:

```

1 # Resultado geral.
2 SELECT ?tipo_situacao ?situacao ?tipo_objeto ?objeto
3 WHERE {
4   ?situacao dados:afeta ?vitima .
5   ?situacao rdf:type ?tipo_situacao .
6   FILTER regex(str(?tipo_situacao), '(furto|roubo)+#', 'i') .
7   OPTIONAL {
8     ?vitima dados:possui|dados:tem ?objeto .
9     ?objeto rdf:type ?tipo_objeto .
10    FILTER regex(str(?tipo_objeto), 'objeto#', 'i') .
11  } .
12 }
13 ORDER BY ?tipo_situacao

```

Código-Fonte 5.8 – Consulta SPARQL da Proposição 4 para retornar a relação de situações onde há objetos roubados ou furtados

```

1 # Resultado quantificado.
2 SELECT ?tipo_situacao ?tipo_objeto (count(distinct ?objeto) as ?total)
3 WHERE {
4   ?situacao dados:afeta ?vitima .
5   ?situacao rdf:type ?tipo_situacao .
6   FILTER regex(str(?tipo_situacao), '(furto|roubo)+#', 'i') .
7   ?vitima dados:possui|dados:tem ?objeto .
8   ?objeto rdf:type ?tipo_objeto .
9   FILTER regex(str(?tipo_objeto), 'objeto#', 'i') .
10 }
11 GROUP BY ?tipo_situacao ?tipo_objeto
12 ORDER BY ?tipo_situacao DESC(?total)

```

Código-Fonte 5.9 – Consulta SPARQL da Proposição 4 para retornar a quantidade por situações onde há objetos roubados ou furtados

É observável na Tabela 5.4 a existência de dados relevantes para uma análise, como o tipo de objeto mais furtado ser da classe Documento, enquanto o mais roubado pertence a classe Objeto. Os dados classificados como Objeto na coluna *tipo_objeto* são aqueles os quais não puderam ter o seu tipo real identificado no momento da criação das instâncias.

tipo_situacao	tipo_objeto	total
furto:Furto	objeto:Documento	"24"^^ xsd:integer
furto:Furto	objeto:Veiculo	"8"^^ xsd:integer
furto:Furto	objeto:Objeto	"3"^^ xsd:integer
furto:Furto	objeto:Acessorio	"2"^^ xsd:integer
furto:Furto	objeto:Eletroeletronico	"2"^^ xsd:integer
roubo:Roubo	objeto:Objeto	"13"^^ xsd:integer
roubo:Roubo	objeto:Acessorio	"1"^^ xsd:integer
roubo:Roubo	objeto:Documento	"1"^^ xsd:integer
roubo:Roubo	objeto:Veiculo	"1"^^ xsd:integer

Tabela 5.4 – Resultado Quantificado da Proposição 4

Capítulo 6

CONCLUSÕES

O presente trabalho teve como objetivo o desenvolvimento de um processo de quantificação de informações em situação de risco, por meio de abordagem semântica. Este, por sua vez, estava englobado em um objetivo mais generalizado, que foi caracterizado pela obtenção e manutenção de SAW de operadores de sistemas de gerenciamento de riscos.

Ao decorrer deste trabalho e do estudo de caso apresentado, foi constatado que por meio de consultas específicas, de acordo com a necessidade do operador, é possível utilizar as informações geradas para mensurar aquilo que se deseja saber baseando-se nos dados que estão contidos nas ocorrências mediante as instâncias da ontologia.

Os resultados quantificados dentro de um modelo semântico permitiram demonstrar que a contextualização é fundamental para que se tenha uma noção do todo, como foi o caso do exemplo dado no estudo de caso referente à relação de locais públicos ou privados onde ocorreram furtos e roubos. Ficou claro que a semântica permitiu que as subclasses fossem demonstradas no resultado, sem ter a necessidade de alterar as consultas SPARQL com comandos adicionais para incluir tais subclasses. Pelo fato de ser um domínio complexo e de requerer constante estudo, os resultados apresentados são comprobatórios de que o processo de quantificação se faz necessário para a tomada de decisão mais assertiva por operadores.

Devido a aleatoriedade de eventos que podem existir dentro do domínio, falhas podem ocorrer perante a quantificação, o que pode levar o operador de sistema de gerenciamento de riscos tomar a decisão errada, podendo causar prejuízos à vida e ao patrimônio. Portanto, como parte dos trabalhos futuros o processo proposto deverá ser aperfeiçoado unindo os resultados quantificados à qualidade de informação. Quanto mais confiável e completa for a informação quantificada gerada, melhor será o subsídio para a evolução dos níveis de SAW do operador e este poderá tomar uma decisão

mais assertiva.

Deste modo se conclui que gerar insumos para sistemas que demandam de análises quantitativas torna-se um instrumento indispensável para cumprir com a aquisição e manutenção do nível 2 de SAW, ou seja, a compreensão da situação atual. Isto permite que o operador utilize tais informações para atingir o nível 3 de SAW, a projeção de estados futuros em uma situação de risco.

6.1 Publicações

Nesta seção são apresentadas as publicações científicas relacionadas ao tema da monografia:

- *Análise Quantitativa de Eventos Criminais Utilizando Abordagem Semântica*. Gustavo Marttos Cáceres Pereira, João Henrique Martins e Leonardo Castro Botega. I Workshop de Informação, Dados e Tecnologia, UFSC, Santa Catarina, Brasil, 2017.
- *Análise Quantitativa de Eventos Criminais Utilizando Abordagem Semântica*. Gustavo Marttos Cáceres Pereira e Leonardo Castro Botega. VII Congresso de Pesquisa Científica: Inovação, Sustentabilidade, Ética e Cidadania, Marília, Brasil, 2017.

REFERÊNCIAS

ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. *Ciência da Informação, Brasília*, SciELO Brasil, v. 32, n. 3, p. 7–20, 2003.

ASARO, C. et al. A domain ontology: Italian crime ontology. In: *Proceedings of the ICAIL 2003 Workshop on Legal Ontologies & Web based legal information management*. [S.l.: s.n.], 2003.

BASILI, V. R. *Quantitative Evaluation of Software Methodology*. [S.l.], 1985.

BAUMGARTNER, N. et al. Beaware!—situation awareness, the ontology-driven way. *Data & Knowledge Engineering*, Elsevier, v. 69, n. 11, p. 1181–1193, 2010.

BOLSTAD, C. A. et al. Using goal directed task analysis with army brigade officer teams. In: SAGE PUBLICATIONS SAGE CA: LOS ANGELES, CA. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. [S.l.], 2002. v. 46, n. 3, p. 472–476.

BOSSÉ, É.; ROY, J.; WARK, S. *Concepts, Models, and Tools for Information Fusion*. [S.l.]: Artech House, Incorporated, 2007. (Artech House intelligence and information operations library). ISBN 9781596930810.

BOTEGA, L. C. *Modelo de Fusão Dirigido por Humanos e Ciente de Qualidade de Informação*. 247 p. Tese (Doutorado) — UFSCar - Universidade Federal de São Carlos, 2016.

BROEKSTRA, J.; KAMPMAN, A.; HARMELEN, F. V. Sesame: A generic architecture for storing and querying rdf and rdf schema. In: SPRINGER. *International semantic web conference*. [S.l.], 2002. p. 54–68.

ENDSLEY, M. R. Design and evaluation for situation awareness enhancement. In: SAGE PUBLICATIONS. *Proceedings of the human factors and ergonomics society annual meeting*. [S.l.], 1988. v. 32, n. 2, p. 97–101.

ENDSLEY, M. R. A taxonomy of situation awareness errors. *Human factors in aviation operations*, v. 3, n. 2, p. 287–292, 1995.

ENDSLEY, M. R. Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, SAGE Publications, v. 37, n. 1, p. 32–64, 1995.

- ENDSLEY, M. R.; GARLAND, D. Theoretical underpinnings of situation awareness: A critical review. *Situation awareness analysis and measurement*, Lawrence Erlbaum Associates, Mahwah, NJ, p. 3–32, 2000.
- FERNÁNDEZ-LÓPEZ, M.; GÓMEZ-PÉREZ, A.; JURISTO, N. Methontology: from ontological art towards ontological engineering. American Association for Artificial Intelligence, 1997.
- FILHO, E. de A. *Iniciação à lógica matemática*. [S.l.]: NBL Editora, 2002.
- FLUIT, C.; SABOU, M.; HARMELEN, F. V. Ontology-based information visualization: toward semantic web applications. In: *Visualizing the semantic web*. [S.l.]: Springer, 2006. p. 45–58.
- GLUZ, J. C. Introdução às lógicas probabilísticas. p. 40–44, 2002.
- GÓMEZ-PÉREZ, A.; BENJAMINS, R. Overview of knowledge sharing and reuse components: Ontologies and problem-solving methods. In: *IJCAI AND THE SCANDINAVIAN AI SOCIETIES. CEUR WORKSHOP PROCEEDINGS*. [S.l.], 1999.
- GRÜNINGER, M.; FOX, M. S. Methodology for the design and evaluation of ontologies. Citeseer, 1995.
- GUARINO, N. Formal ontology and information systems. In: *Proceedings of FOIS*. [S.l.: s.n.], 1998. v. 98, n. 1998, p. 81–97.
- HEIJST, G. V.; SCHREIBER, A. T.; WIELINGA, B. J. Using explicit ontologies in kbs development. *International journal of human-computer studies*, Elsevier, v. 46, n. 2-3, p. 183–292, 1997.
- HEPP, M.; LEUKEL, J.; SCHMITZ, V. A quantitative analysis of product categorization standards: content, coverage, and maintenance of ecl@ ss, unspsc, eotd, and the rosettanet technical dictionary. *Knowledge and Information Systems*, Springer, v. 13, n. 1, p. 77–114, 2007.
- ISOTANI, S.; BITTENCOURT, I. I. *Dados Abertos Conectados: Em busca da Web do Conhecimento*. [S.l.]: Novatec Editora, 2015.
- JASPER, R.; USCHOLD, M. et al. A framework for understanding and classifying ontology applications. In: *Proceedings 12th Int. Workshop on Knowledge Acquisition, Modelling, and Management KAW*. [S.l.: s.n.], 1999. v. 99, p. 16–21.
- KOKAR, M. M.; MATHEUS, C. J.; BACLAWSKI, K. Ontology-based situation awareness. *Information fusion*, Elsevier, v. 10, n. 1, p. 83–98, 2009.
- MACGREGOR, R. M.; KO, I.-Y. Representing contextualized data using semantic web tools. In: *PSSS*. [S.l.: s.n.], 2003.
- MATHEUS, C. J.; KOKAR, M. M.; BACLAWSKI, K. A core ontology for situation awareness. In: *Proceedings of the Sixth International Conference on Information Fusion*. [S.l.: s.n.], 2003. v. 1, p. 545–552.
- MIZOGUCHI, R. Part 3: Advanced course of ontological engineering. *New Generation Computing*, Springer, v. 22, n. 2, p. 193–220, 2004.

MIZOGUCHI, R.; VANWELKENHUYSEN, J.; IKEDA, M. Task ontology for reuse of problem solving knowledge. *Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing*, IOS press Amsterdam, v. 46, p. 59, 1995.

NOY, N. F.; MCGUINNESS, D. L. *Ontology development 101: A guide to creating your first ontology*. [S.l.]: Stanford knowledge systems laboratory technical report KSL-01-05 and Stanford medical informatics technical report SMI-2001-0880, Stanford, CA, 2001.

SANTARÉM SEGUNDO, J. E. Web semântica, dados ligados e dados abertos: uma visão dos desafios do Brasil frente às iniciativas internacionais. *Tendências da Pesquisa Brasileira em Ciência da Informação*, v. 8, n. 2, 2015. Acesso em: 07 junho 2017. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/view/207/272>>.

SCHIESSL, J. M. Ontologia: o termo e a idéia 10.5007/1518-2924.2007 v12n24p172. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, v. 12, n. 24, p. 172–181, 2007.

USCHOLD, M.; GRUNINGER, M. Ontologies: Principles, methods and applications. *The knowledge engineering review*, Cambridge Univ Press, v. 11, n. 02, p. 93–136, 1996.

USCHOLD, M.; KING, M. *Towards a methodology for building ontologies*. [S.l.]: Artificial Intelligence Applications Institute, University of Edinburgh Edinburgh, 1995.

VIEIRA, R.; LOPES, L. Processamento de linguagem natural e o tratamento computacional de linguagens científicas. *EM CORPORA*, p. 183, 2010.