

UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
CAMPUS DE MARÍLIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

**Modelo de Transformação e Representação de
Informações para Suportar Projeções Situacionais**

Valdir Amancio Pereira Junior

Marília, 2020

UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
CAMPUS DE MARÍLIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

Valdir Amancio Pereira Junior

Modelo de Transformação e Representação de Informações para Suportar Projeções Situacionais

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da informação da Universidade Estadual Paulista “Júlio de Mesquita Filho” - Campus Marília como requisito para a obtenção do título de mestre em Ciência da Informação.

Orientador: Prof. Dr. Leonardo Castro Botega

Financiadora: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior e Conselho Nacional de Desenvolvimento Científico e Tecnológico (processo nº 132000/2019-9)

Marília, 2020

P436m Pereira Junior, Valdir Amancio
Modelo de Transformação e Representação de
Informações para Suportar Projeções Situacionais / Valdir
Amancio Pereira Junior. -- Marília, 2020
108 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista
(Unesp), Faculdade de Filosofia e Ciências, Marília
Orientador: Leonardo Castro Botega

1. Transformação da Informação. 2. Representação da
Informação. 3. Qualidade da Informação. 4. Projeção
Situacional. 5. Consciência Situacional. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da
Faculdade de Filosofia e Ciências, Marília. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

VALDIR AMANCIO PEREIRA JUNIOR

MODELO DE TRANSFORMAÇÃO E REPRESENTAÇÃO DE INFORMAÇÕES
PARA SUPORTAR PROJEÇÕES SITUACIONAIS

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação, da Universidade Estadual Paulista – Campus de Marília, como requisito para a obtenção de título de mestre em Ciência da Informação.

Área de Concentração: Informação, Tecnologia e Conhecimento.

Linha de Pesquisa: Informação e Tecnologia.

Data da defesa: 12 de maio de 2020.

BANCA EXAMINADORA

Leonardo Castro Botega (Orientador)
Docente do Programa de Pós-Graduação em Ciência da Informação da UNESP/FFC

José Eduardo Santarém Segundo
Docente do Programa de Pós-Graduação em Ciência da Informação da UNESP/FFC

Paulo César Guerreiro da Costa
Docente do Departamento de Systems Engineering and Operations Research da George Mason University/EUA

Suplentes

Ricardo César Gonçalves Sant'Ana
Docente do Programa de Pós-Graduação em Ciência da Informação da UNESP/FFC

Seiji Isotani
Docente do Departamento de Sistemas de Computação da USP/ICMC.

Agradecimentos

Agradeço primeiramente ao todo, a cada instante vivido que me proporcionou alegrias, dores e aprendizados, moldando minha mente, corpo e espírito até o presente momento.

Aos meus pais, Elisângela e Valdir. Vocês são a minha base, meus guias e meus faróis em meio às tempestades.

Aos meus irmãos, André e Pedro. Formamos um trio único e espetacular, salvo as peculiaridades de cada existência nesse plano, sentimentos e visões, por mais que haja desentendimentos, juntos sempre seremos mais fortes. Ao Chester e a Kyojin. Demonstrando amor de forma inigualável a nós, mesmo quando não merecemos.

Ao meu orientador e amigo Prof. Dr. Leonardo Castro Botega. Por todos os ensinamentos e orientações, acadêmicas e do mundo, pelo tempo que nos conhecemos, reuniões, congressos, viagens, happy hours e churrascos.

Ao meu amigo Marttos. Por toda a parceria de longa data, pelas conversas filosóficas, risadas e apoio em todos os momentos, histórias não faltam.

Ao meu amigo Cássio. Que mostrou como é importante rir de tudo, passar pelas dificuldades e que os amigos surgem em diversas situações. Tudo passa.

Ao meu amigo Mapelli. Pelas conversas, risadas e bons momentos.

Ao meu amigo Benini. Mostrou como as amizades têm nuances e que a vida pode ter prioridades diferentes.

Ao meu amigo Gian. Pelos conselhos e risadas sempre garantidas.

À banca composta por Prof. Dr. José Eduardo Santarém Segundo e Prof. Dr. Paulo Costa. Por toda receptividade no mundo acadêmico e gerando oportunidades únicas de aprendizado e discussões.

Aos docentes e discentes do PPGCI da Unesp e dos grupos de pesquisa. Pelo acolhimento e possibilidade de discussões e novos conhecimentos.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Pelo apoio financeiro.

De forma geral, agradeço a todos amigos de longa data, passageiros ou de momento que de alguma forma contribuíram durante este tempo.

Resumo

Consciência Situacional (SAW) refere-se ao nível de consciência que um indivíduo ou equipe detém sobre uma situação. No domínio criminal, falhas da SAW podem induzir operadores humanos a cometer erros na tomada de decisão e gerar riscos à vida e ao patrimônio. Neste contexto, domínios críticos são afetados por problemas de qualidade da informação que podem emergir e propagar em diversos momentos do ciclo informacional, especialmente quando a fonte dos dados são seres humanos. Além disso, a literatura apresenta soluções limitadas e incompletas para a concepção e representação de projeções situacionais. Desta maneira, o objetivo deste trabalho é apresentar a construção do modelo Tree inspirado pela complexidade do domínio criminal, baseado em ferramentas computacionais de fusão semântica e avaliação de qualidade de informações, vislumbrando gerar subsídios para melhor representar e recuperar informações de projeções situacionais e conseqüentemente entregar aos analistas criminais melhores insumos para SAW. De forma a validar o modelo foi desenvolvida uma prova de conceito envolvendo situações criminais críticas, cujos resultados apontam a capacidade de criar associações semânticas entre informações e representar a projeção situacional em um modelo ontológico, gerando insumo informacionais mais completos.

Palavras-chave: Consciência Situacional, Projeção Situacional, Transformação de Informações, Representação da Informação, Ontologia, Qualidade da Informação, Fusão de Informações

Abstract

Situational Awareness (SAW) refers to the level of consciousness that an individual or team holds about a situation. In the criminal field, SAW failures can induce human operators to make mistakes in decision-making and pose risks to life and property. In this context, critical domains are affected by information quality problems that can emerge and propagate at various times in the information cycle, especially when the data source is human beings. In addition, the literature presents limited and incomplete solutions for the conception and representation of situational projections. Thus, the objective of this work is to build a model of transformation and representation of information inspired by the complexity of the criminal domain, based on computational tools of semantic fusion and evaluation of information quality, envisioning generate subsidies to better represent and recover information from situational projections and consequently deliver to criminal analysts better inputs for SAW. In order to validate the model, a proof of concept was developed involving critical criminal situations, whose results indicate the ability to create semantic associations between information and represent situational projection in an ontological model, generating more complete information inputs.

Keywords: Situational Awareness, Situational Projection, Information Transformation, Information Representation, Ontology, Information Quality, Information Fusion.

Lista de Quadros

Quadro 1: Classes da URBANITY, descrevendo o que cada classe presente representa para a situação.....	51
Quadro 2: Resultados obtidas nas buscas por vocabulários relacionados a projeção na plataforma <i>Linked Open Vocabularies</i>	56
Quadro 3: Propriedades que representam a projeção de classes e instâncias.....	57

Lista de Figuras

Figura 1: Modelo de Consciência Situacional de Endsley.....	24
Figura 2: Exemplo de ontologia aplicada ao domínio criminal.....	33
Figura 3: Modelo Tree.....	48
Figura 4: Representação gráfica da Ontologia URBANITY.....	53
Figura 5: Passos da Metodologia 101.....	54
Figura 6: Adesão da classe <i>Projection</i> no Modelo de Representação.....	57
Figura 7: Modelo Quantify.....	59
Figura 8: Processo de fusão semântica de informações baseado no Modelo Quantify.....	62
Figura 9: Situação instanciada com as informações do relato criminal 1.....	65
Figura 10: Situação instanciada com as informações do relato criminal 2.....	66
Figura 11: Situação instanciada com as informações do relato criminal 3.....	66
Figura 12: Ontologia da situação instantânea após a transformação e representação, com as informações de todos os relatos associados a um índice de qualidade.....	69
Figura 13: Representação de um resultado de fusão com projeção informacional.....	70
Figura 14: Etapas da prova de conceito para validar o modelo.	73
Figura 15: Destaque para a camada Coleta de Dados Provindos de Inteligência Humana.....	76
Figura 16: Amostra de dados disponíveis no Portal de Transparência da SSP-SP.	77
Figura 17: Saída da camada de coleta com dados.....	77
Figura 18: Destaque para a camada Transformação dos Dados em Instâncias.....	78
Figura 19: Processos internos à camada de Transformação de Dados.....	78
Figura 20: Processos de limpeza e padronização das informações coletadas.....	79
Figura 21: Saída obtida dos procedimentos de limpeza e padronização na camada de Transformação.....	80
Figura 22: Trecho do primeiro relato transformado em OWL para ser consumido pela URBANIT.....	81
Figura 23: Trecho do segundo relato transformado em OWL para ser consumido pela URBANITY.....	82
Figura 24: Trecho do terceiro relato transformado em OWL para ser consumido pela URBANITY.....	82
Figura 25: Instância do primeiro relato e seus elementos já importados para a URBANITY.....	83
Figura 26: Instância do segundo relato e seus elementos já importados para a URBANITY.....	83
Figura 27: Instância do terceiro relato e seus elementos já importados para a URBANITY.....	84
Figura 28: Função SPARQL para listagem dos elementos informacionais por instância.....	84
Figura 29: Resultado da Função SPARQL para listar elementos informacionais.....	85

Figura 30: Representação da instância da avaliação de qualidade com DQV e o relacionamento a situação.....	85
Figura 31: Destaque para a camada Fusão Semântica de Informações.....	86
Figura 32: Processos internos à camada de Fusão de Informações.....	86
Figura 33: Função de busca semântica de informações para fusão.....	87
Figura 34: Resultado busca semântica de informações para fusão.....	88
Figura 35: Função de associação semântica de informações correlatas.....	89
Figura 36: Resultado da função SPARQL de associação semântica de informações.....	89
Figura 37: Resultado da fusão de informação em OWL.....	90
Figura 38: Representação das instâncias.....	90
Figura 39: Representação do resultado da fusão com avaliação de qualidade utilizando o DQV.....	91
Figura 40: Destaque para a camada Representação Semântica das Situações.....	92
Figura 41: Função SPARQL para listagem de instâncias e propriedades como insumos para a camada de Projeção.....	92
Figura 42: Recorte do resultado obtido ao aplicada a função SPARQL para representar a situação.....	93
Figura 43: Destaque para a camada Fusão Semântica de Informações.....	94
Figura 44: Processos internos à camada de Projeções Situacionais.....	95
Figura 45: Função SPARQL de listagem de propriedades.....	95
Figura 46: Processos internos à camada de Projeções Situacionais.....	96
Figura 47: Resultado do procedimento de criação de instâncias de projeção para a situação.....	96
Figura 48: Relação entre o nível informacional da projeção e temporalidade implícita.....	97
Figura 49: Representação das instâncias de projeção informacional em relação a fusão de informações.....	98

Sumário

Resumo	6
Abstract.....	7
Lista de Quadros	8
Lista de Figuras	9
1. Introdução.....	13
1.1. Contextualização e Justificativa	13
1.2. Motivação	14
1.3. Problemática	15
1.4. Objetivos.....	16
1.5. Metodologia.....	16
1.6. Trabalhos Relacionados.....	18
1.7. Organização da Proposta	20
2. Consciência Situacional e Projeções Situacionais	22
2.1. Consciência Situacional.....	22
2.2. Modelo de Consciência Situacional de Endsley	23
2.3. Fatores Influentes no Processo	25
2.4. Domínios de Risco e Consciência Situacional	26
2.5. Suporte à Consciência Situacional e Projeções Situacionais.....	28
2.6. Considerações Finais	29
3. Ontologias como Suporte à Consciência Situacional.....	30
3.1. A Web Semântica e Ontologias	30
3.2. O Papel da Ontologia para Representar Informações Contextuais.....	32
3.3. Ontologias como Suporte a Consciência Situacional	33
3.4. Considerações Finais	34
4. Sistemas de Fusão de Dados e Informações.....	36
4.1. Fusão de Dados e Informações	36
4.2. Fusão de Informações e a Inteligência Humana	37
4.3. Adequação dos Métodos de Fusão para Suportar Informações em Alto-Nível ...	39
4.4. Qualidade de Dados e Informações na Fusão de Informações	41

4.4.1. Incorporando a Qualidade de Dados e Informações nos Processos da Fusão de Informações	42
4.5. Considerações Finais	43
5. Modelo de Transformação e Representação de Informações para Suportar Projeções Situacionais.....	45
5.1. O Modelo Tree.....	45
5.2. Ontologia URBANITY: A Ontologia Criminal	50
5.2.1. Avanços na URBANITY para suporte às Projeções Situacionais.....	54
5.3. Modelo Quantify.....	58
5.4. Avaliação de Qualidade utilizando Metodologia IQESA.....	59
5.5. Avanços para um processo de Fusão Semântica de Informações	61
5.6. Demonstração do Modelo Tree	63
6. Prova de Conceito do Modelo Tree.....	73
6.1. Coleta de Dados.....	75
6.2. Transformação	77
6.3. Fusão.....	86
6.4. Representação	91
6.5. Projeção	93
6.6. Considerações sobre Modelo Tree e Prova de Conceito	98
7. Conclusões	100
7.1. Trabalhos Futuros	103
Referências	105

1. Introdução

O presente Capítulo se dedica a apresentar uma introdução sobre os temas que embasam o arcabouço científico e tecnológico deste trabalho, junto ao domínio estudado, bem como apresentar a motivação, a problemática, os objetivos, a metodologia utilizada para o desenvolvimento do trabalho, os resultados esperados e a organização dos capítulos.

1.1. Contextualização e Justificativa

Consciência Situacional (*Situational Awareness – SAW*) é um processo cognitivo importante para tomadores de decisão em diversas áreas e se refere ao nível de consciência que um indivíduo ou equipe detém sobre uma situação. Especificamente, SAW diz respeito à percepção da presença e disposição de entidades de interesse em um ambiente, digital ou analógico, à compreensão do significado e importância de suas ações individuais e coletivas no espaço-tempo e à projeção de seus status em um futuro próximo (KOKAR; ENDSLEY, 2012, p. 91).

No domínio criminal, SAW é um fator crucial para, com base em dados coletados de múltiplas fontes, processados e devidamente representados, revelar tendências, ocorrências e a possibilidade de riscos iminentes. Uma SAW limitada pode comprometer a compreensão sobre o estado real de um ambiente de interesse, levando a uma má tomada de decisão, podendo resultar em consequências para vidas, patrimônios e até o meio ambiente. Embora o acesso a informações sobre as situações não possa garantir SAW e a melhor tomada de decisão de um humano, uma informação de qualidade pode colaborar com tomadores de decisão a fim de eliminar a incerteza, manter um conhecimento superior e atingir a projeção situacional (ROGOVA; BOSSE, 2010, p. 2).

Conseguir estimar a evolução de uma situação e suas informações, atingindo a projeção situacional, é ainda mais desafiador ao considerar o contexto de dados criminais no Brasil, levando em conta o grande volume e heterogeneidade dos mesmos. Outro fator crítico é a qualidade dos dados e suas respectivas fontes, em sua maioria de inteligência humana, neste caso, relatos de crimes às centrais de atendimento das forças de segurança. Tipicamente tais dados são incompletos, desatualizados, inconsistentes e influenciados por fatores culturais, afetando os processos informacionais e computacionais que coletam, organizam,

recuperam, processam e representam os dados e estimulam a SAW do analista (BOTEGA *et al.*, 2019; BOTEGA *et al.*, 2017).

Considerando o contexto e desafio apresentados, a Ciência da Informação (CI) se posiciona como Ciência essencial para viabilizar e organizar a comunicação entre diversas ciências e disciplinas, de acordo com seu caráter interdisciplinar, Saracevic (1995) já aponta que CI pode melhorar as formas como as informações são utilizadas. Suplementado por Borko (1968), que traz a CI como forma de investigar as formas, comportamentos e fluxos informacionais, além de considerar o processamento da informação, visando otimizar a acessibilidade e usabilidade. Questões essenciais ao considerar a qualidade da informação, uma vez que este conceito é subjetivo e diretamente ligado ao domínio estudo, sendo que a avaliação dos índices de qualidade da informação irá depender de parâmetros e padrões das ciências associadas ao domínio, não apenas a informação em si.

Sobre essa visão interdisciplinar e embasados na CI, diversos trabalhos vêm buscando formas de superar as problemáticas de coletar, armazenar, transformar, processar, inferir, representar, recuperar e apresentar dados relacionados à domínios de emergência (BOTEGA *et al.*, 2019; PEREIRA JUNIOR; PEREIRA; BOTEGA, 2019; SILVA *et al.*, 2018; PEREIRA *et al.*, 2017; BOTEGA; *et al.*, 2017).

Adicionalmente, tais processos afetam diretamente os dados que são consumidos por um humano que irá tomar decisões, agregando diversas ciências junto a CI, como a Ciência da Computação, Psicologia, Ciência Cognitiva e disciplinas como Fusão de Dados e Informações, Ontologias, Qualidade da Informação, Recuperação de Informações entre outras.

Desta maneira, este trabalho propõe um modelo de transformação e representação de informações em suporte à obtenção de projeções situacionais. Para tal, é discutida um arranjo tecnológico, envolvendo também SAW, ontologias, fusão de informações, avaliação da qualidade da informação.

1.2. Motivação

Adquirir e manter SAW, ao ponto de atingir a projeção situacional, visando proporcionar a melhor informação para a tomada de decisão é desafiador, principalmente ao considerar um domínio crítico como o contexto brasileiro de dados criminais, levando em conta o grande volume, heterogeneidade e qualidade dos mesmos.

Fatores humanos associados ao operador, junto a fatores informacionais e relacionados à sistemas de informação podem influenciar negativamente a análise e compreensão das informações de um ambiente, considerando que apenas o acesso do humano as informações não garantem sua SAW, projeção situacional e a futura tomada de decisão. Estes fatores discriminam a necessidade do desenvolvimento e adesão de processos informacionais e computacionais com a capacidade de manipular, qualificar e apresentar informações que dão suporte à SAW.

1.3. Problemática

Diversos processos, técnicas e ferramentas computacionais podem contribuir para a melhoria da informação utilizada em sistemas de avaliação de situações e apoio à tomada de decisão. Todavia, é necessária uma atenção maior aos processos informacionais e ao valor atribuído às informações, uma vez que sistemas complexos como os citados acima, que influenciam diretamente na SAW e ações do humano colocando em risco vidas, patrimônio e ambiente, demandam de informações de confiança e de qualidade, logo refletem na forma como a informação é coletada, organizada, recuperada, processada, representada e apresentada.

Junto à necessidade de atentar-se aos processos informacionais, está presente a dificuldade em lidar com as informações provenientes de inteligência humana, ainda mais dentro de um domínio crítico, como o criminal, agregando e melhorando valor e qualidade a tais informações de maneira a conseguir construir e manter SAW no humano, levando este a elaborar uma projeção situacional.

Diante de tais problemas pode-se estabelecer as seguintes questões de pesquisa: Como são os dados sobre crimes no Brasil? Como esses dados são manipulados hoje? Como são persistidos e propagados? Qual a necessidade uma integração e processamento destes? Por que os dados são falhos e sensíveis? Qual a confiança para tomada de decisão sobre os dados como estão persistidos? É possível estimular a SAW e projeção situacional através dos dados em seu formato original.

Tais questões indicam a necessidade de desenvolver ou incrementar processos informacionais e computacionais que tenham a capacidade de lidar com as dificuldades do domínio e as informações que este gera.

1.4. Objetivos

O presente trabalho tem como objetivo geral a criação de um modelo capaz de transformar e representar informações com suporte semântico e de qualidade da informação o qual permita a obtenção e melhoria das informações a partir de processos e inferências semânticas, que contribuam para processos informacionais de recuperação, transformação, representação e inferência.

Especificamente, busca-se:

- Estruturar um processo informacional que empregue Fusão de Informações Semânticas, Avaliação da Qualidade de Informações e a Representação de Informações Semânticas utilizando técnicas já consolidadas na literatura;
- Aplicar o Modelo Quantify (*Quality-aware Human-driven Information Fusion Model*) para dar suporte ao processo de transformação e fusão de informações e orientar a forma como serão inferidas e avaliadas as informações provenientes de humanos;
- Aplicar a metodologia IQESA (*Information Quality Assessment Methodology in the Context of Emergency Situational Awareness*) como o processo de avaliação de qualidade das informações, sendo complementar ao Modelo Quantify que prevê a avaliação e uso da qualidade;
- Customizar a ontologia URBANITY (*Human-driven Brazilian Crime Ontology*) para que seja possível, além de representar informações de domínios críticos junto aos seus índices de qualidade, também representar projeções informacionais.

1.5. Metodologia

Este trabalho se coloca como uma pesquisa exploratória e experimental, devido à natureza e condição do mesmo, que se encaixa entre áreas muito consolidadas, a Ciência da Informação e a Ciência da Computação. Essa associação entre as duas áreas cria um cenário específico que demanda uma revisão detalhada da literatura para a estruturação dos conceitos que envolvem o tema abordado.

O trabalho também apresenta uma natureza qualitativa, isto devido ao domínio tratado e questionamentos do trabalho, onde apenas uma abordagem qualitativa, visando apenas números e estatísticas, não seria suficiente para atingir os objetivos propostos e sanar questões.

Ou seja, em um domínio crítico, como o de crimes no território brasileiro, abordado nesta pesquisa, a interpretação e associação deste, deve estar a mais próxima possível do ambiente de cada sujeito ou agente imerso, assim, exigindo uma abordagem qualitativa, tanto nos métodos e técnicas, para que seja possível executar o trabalho e atingir os objetivos levantados.

A abordagem qualitativa, não restringe ou impede a aplicação futura de método e técnicas quantitativas, isto visando a geração de novas informações, como indicadores, índices e estatísticas acerca do domínio estudado, sendo possível alinhar os resultados obtidos com a aplicação das diferentes naturezas de pesquisa. Todavia, este trabalho se restringe apenas à abordagem qualitativa.

A execução deste se divide nas seguintes etapas:

1. Levantamento e revisão bibliográfica em busca de teorias, conceitos e definições para sustentar as temáticas de SAW, fusão de informações, avaliação de qualidade, web semântica e ontologias. Junto foi realizada o levantamento de trabalhos relacionamentos, buscando trabalhos que se aproximam de cada uma das temáticas descritas acima;
2. Coleta de dados criminais através da Lei de Acesso à Informação, junto a análise desses dados com o intuito de identificar os principais elementos informacionais e problemas de qualidade mais comuns;
3. Análise do modelo de fusão de informação Quantify buscando identificar a capacidade do mesmo em ser aplicado como motor de transformação de informações e sua aderência com SAW, dados provindos de inteligência humana, processos de avaliação de qualidade, semântica e ontologias;
4. Análise da metodologia de avaliação de qualidade IQESA com o intuito de validar a competência em promover a análise contínua de qualidade dentro de um modelo de transformação, uma vez que a mesma já está relacionada ao Modelo Quantify em alguns trabalhos;
5. Ampliação da ontologia URBANITY, trazendo novas classes, propriedades, instâncias e axiomas, para sanar as necessidades informacionais de representação dos dados analisados neste trabalho e também a suprir a capacidade de representar projeções situacionais;
6. Especificação do modelo de transformação e representação de informações provenientes de inteligência humana, capaz de produzir e que seja ciente de índices de qualidade, junto a capacidade de criar representações de projeções

situacionais. Estando intrínseco ao modelo os processos de coleta, fusão semântica de informações e ontologias;

7. Instanciação das informações e situações na ontologia, possibilitando os procedimentos de análise, busca, transformação e representação das informações quanto a fusão e projeção.
8. Realização de prova de conceito, buscando validar as análises e ampliações realizadas para criar o modelo, assim como validar o próprio modelo junto à suas camadas e procedimentos internos, onde a URBANITY já ampliada seja aplicada.

1.6. Trabalhos Relacionados

Os trabalhos listados como relacionados nesta seção fazem parte do levantamento e revisão bibliográfica feitos no período de junho de 2018 a abril de 2020. Para este aporte foi utilizada principalmente a base do Google Scholar, que além de conter o próprio repositório, faz a indexação de outras bases como Scielo, Elsevier, Springer e IEEE (*Institute of Electrical and Electronics Engineers*).

Como parâmetros para as buscas foi definido o intervalo de 2005 e 2017 para buscas dos trabalhos, utilizando principalmente dos termos: “fusão de dados e informações semânticas”, “ontologias de domínio crítico e emergência”, e “consciência situacional preditiva”, junto aos termos em inglês: “*semantic data and information fusion*”, “*critical and emergency domain ontologies*” e “*predictive situational awareness*”.

Dentre os resultados obtidos foram selecionados os trabalhos a seguir, sendo distribuídos da seguinte forma entre as bases: 5 (cinco) foram encontrados na base do Google Scholar, 3 (três) na base do IEEE e 2 (dois) na base da Springer.

Atualmente, há um grande interesse no uso de sensores não-físicos para promover uma visão holística do mundo. Entre esses sensores está o uso da inteligência humana para observar um cenário, que inclui o conhecimento como fontes qualitativas de dados e informações, direta ou indiretamente. As informações da inteligência humana estão sujeitas a tendências, imprecisões e omissões. Interpretar e utilizar essas informações requer a dedução do significado semântico, que muitas vezes não pode ser inferido por funções automatizadas. Exige a participação humana na interpretação e no uso da informação. Exemplos destes desafios incluem a interpretação de relatos feitos por testemunhas de sistemas de chamada de emergência (VINCEN *et al.* 2009).

Recentemente, surgiram abordagens semânticas para a fusão de dados, com o uso de ontologias como parte do processo de representação e inferência. No trabalho de McGuinness (2003), as primeiras classificações e organizações são apresentadas utilizando ontologias como parte das aplicações (ontologia de domínio), bem como o uso de ontologias externas (ontologia superior), que prestam serviços de vocabulários e glossários controlados, contendo informações de diferentes classes, com restrições sobre tipos e valores.

Além da evolução do papel humano durante os processos do modelo de fusão de dados, conforme relatado por Blasch *et al.* (2013), também foram consideradas questões referentes aos dados e à qualidade da informação. O objetivo é integrar soluções para a avaliação de emergências em conjunto com abordagens de fusão de dados, para melhores visualizações, interfaces de usuário e integração de dados (SOUZA *et al.* 2015) (BOTEGA *et al.* 2017).

Ao mesmo tempo, houve uma evolução dos modelos e algoritmos de fusão, além de novas áreas de aplicação que passaram a fazer uso de tais abordagens, especialmente no contexto de cenários críticos. Eles necessitam de informações extremamente confiáveis, como aplicações relacionadas ao domínio da gestão de emergências (VINCEN *et al.* 2009) (BLASCH *et al.* 2013).

Carvalho *et al.* (2010), fazem uso de ontologias para representar áreas críticas caracterizadas pela presença de incerteza. Para tanto, realizou-se um estudo de técnicas capazes de representar a semântica complexa presente nessas áreas, juntamente com uma análise da incerteza que eles contêm. Empregando também Redes Bayesianas com o intuito de fortalecer as relações.

Costa *et al.* (2005) traz a possibilidade do uso de ontologias probabilísticas (PROWL) e uma rede bayesiana. A rede Bayesiana foi capaz de representar relações entre múltiplas entidades presentes em uma ontologia e a probabilidade e incerteza da informação. O trabalho apresenta detalhes relevantes sobre os elementos de uma ontologia e como a utilização de uma ontologia probabilística influencia a construções de ontologias OWL.

Costa *et al.* (2010) apresentam alguns dos principais requisitos de Sistemas de Consciência Situacional Preditiva (*Predictive Situation Awareness Systems - PSAW*), para isso os autores trazem uma abordagem baseada no levantamento e análise de trabalho relacionados que sustentam a proposta, junto a apresentação de uma arquitetura computacional que contempla o processamento distribuído, principalmente de redes Bayesianas. O trabalho também traz a preocupação sobre as diversas incertezas e problemas que envolvem os processos de inferências, porém o mesmo não traz o foco sobre a

necessidade da avaliação de qualidade dos dados e informações, impossibilitando o uso desses índices como parâmetros nos processos computacionais e informacionais.

Costa *et al.* (2012) utilizaram a fusão de dados em modelos Bayesianos altamente expressivos, de modo que é possível processar informações de alto nível em um ambiente de fusão de dados. Além disso, os modelos de captura e classificação de dados são comparados, relacionando os pontos positivos e negativos entre eles e como fatores de incerteza podem surgir e afetar os dados, considerando o papel do ser humano e como ele irá interagir e reagir aos dados. Técnicas e métodos que têm a capacidade de interpretar e representar essa incerteza dos dados organizados em uma ontologia, foram analisados quanto às relações probabilísticas existentes entre as classes. Finalmente, o trabalho também emprega técnicas de redes de múltiplas entidades Bayesianas, uma técnica altamente expressiva e útil para uma grande quantidade e diversidade de dados para realizar o processo de fusão.

Kim *et al.* (2013), realizaram uma análise de métodos de captura de dados e informações Web heterogêneas, utilizando o conceito de dados vinculados, o que permitiu a conexão entre dados através de algum contexto comum, garantindo a máxima captação de extensão e recuperação de dados em na Web. Neste trabalho também é citada a tecnologia RDF que permite a classificação e estruturação de um documento capturado. Eles podem ser reutilizados por outras tecnologias, como SPARQL, que permitiu a execução da consulta no RDF. Isso permite maior processamento e manipulação de dados capturados da Web. Essas técnicas permitem que o operador humano realize questões altamente complexas e receba a melhor resposta possível no contexto da fusão de dados e informações com maior integração de dados vinculados.

1.7. Organização da Proposta

A presente proposta está organizada em capítulos e seções, buscando a melhor organização das ciências, modelos, processos e abordagens apresentadas e utilizadas no decorrer do trabalho.

O Capítulo 2 se dedica a apresentação do conceito de SAW, bem com o modelo proposto por Endsley, o qual prevê a ação de fatores externos, são também colocadas questões acerca das projeções situacionais, além de tratar da ação da SAW em domínios de risco e quais mecanismos dão suporte a este processo.

No Capítulo 3 são apresentar os conceitos e definições a cerca de Ontologias, trazendo-as como uma forma de representação da informação, além de tratar como as

Ontologias e a Web Semântica podem se relacionar, trazendo todos os benefícios e meios relacionados à aplicação de ontologias no suporte à tomada de decisão.

O Capítulo 4 traz a Fusão de Dados e Informações, demonstrando suas capacidades e contribuições para os processos informacionais, refletindo diretamente na qualidade das informações, uma vez que apresenta a capacidade de buscar, analisar e integrar informações de fontes heterogêneas. Modelos de fusão também preveem a participação humana como fonte e parte do ciclo informacional, trazendo variações para possibilitar o ciclo de vida das informações, como o uso de ontologias em seus processos internos. O Capítulo traz também o uso direto da qualidade da informação pela fusão.

O Capítulo 5 apresenta a proposta, trabalhos relacionados, junto ao conjunto de modelos e processos que inspiram e dão suporte direto ao modelo proposto, como a ontologia URBANITY, o Modelo Quantify, a metodologia de avaliação de qualidade IQESA, apresentando a ideia do modelo de transformação e representação de informações para auxiliar nas projeções situacionais, ao final do Capítulo são apresentados os resultados preliminares.

O Capítulo 6 traz a Prova de Conceito, a qual busca validar a aplicação do modelo proposto e seus respectivos métodos dentro do contexto de situações críticas. Apresentando as fases do processamento de informações, assim tais informações evoluem e se comportam entre cada camada do modelo.

Os Capítulos 7 e 8, se dedicam respectivamente aos Trabalhos Futuros e Conclusões do trabalho.

2. Consciência Situacional e Projeções Situacionais

Esse capítulo traz dois conceitos deste trabalho, que estão diretamente relacionados, o processo cognitivo de Consciência Situacional (*Situational Awareness* - SAW) e a Projeção Situacional, estado cognitivo derivado da consciência situacional que representa a capacidade humana de projetar como situações e informações estarão em um futuro baseado no seu estado mental e informações dispostas.

2.1. Consciência Situacional

Cunhado por Endsley (1988), o termo consciência situacional (SAW) refere-se a um conceito fundamental que diz respeito ao processo decisório de cada indivíduo, colocando-a como a percepção das entidades que compõem um ambiente, considerando o tempo e espaço, buscando a compreensão do significado e contexto em que se encontram, assim como realizar a projeção. A autora também posiciona SAW como uma rotina, um processamento de informações dirigida pela aquisição e uso destas, estando assim, presente desde atividade operacionais e técnicas, até atividades comuns ao dia a dia.

Porém o humano, pode estar imerso em um ambiente que a condição de estar alta pressão e exposto a uma gama relevante de informações sensíveis, que precisa ser analisado com a devida atenção do indivíduo para que não haja má interpretação e conseqüentemente, pode manter a assertividade da tomada de decisões. A SAW corresponde a um processo cognitivo do indivíduo. Não garante que isso produzirá a melhor imagem da situação e, conseqüentemente, tomar a decisão mais adequada. No entanto, isso poderia melhorar a qualidade da tomada de decisão.

A SAW de um indivíduo pode ser qualificada através de suas experiências, capacidades, carga de trabalho, metas e objetivos, entre outros fatores. No entanto, mesmo os indivíduos mais qualificados podem tomar decisões equivocadas se o seu SAW é impreciso. Devido a isso, Endsley (2016) argumenta que, a fim de ser mais propensos a fazer uma tomada de decisão mais definitiva e mais adequada, é necessário melhorar continuamente a SAW, melhorando assim o processo de tomada de decisão.

Devido aos cenários críticos e dinâmicos, o domínio militar foi um dos primeiros domínios a aprofundar as pesquisas sobre SAW. Todavia esse conceito não fica restrito, sendo aplicado em outros como, controle de tráfego aéreo, diagnóstico médico, análise de

mercado financeiro, monitoramento de desastres e outros, principalmente aqueles críticos ao tempo e que envolvem muitos fatores, como vida e patrimônio (ENDSLEY, 2001; KOKAR; ENDSLEY, 2012, p. 93).

Apesar das diferenças entre tais domínios, é possível identificar um fator comum, são ambientes dinâmicos que têm seus estados e informações alteradas independente da ação de um agente, seja humano ou máquina. Isto faz com que SAW seja um processo contínuo e incremental, exigindo atenção total do humano para que possa estar ciente de todo ambiente e possa realizar ações baseadas no conhecimento que construiu.

Partindo dessas considerações, pode-se posicionar SAW como um dos pilares da tomada de decisão. Entretanto, ter SAW não é o suficiente para construir e suportar uma boa decisão, considerando ambiente dinâmicos e as respectivas variáveis envolvidas, além de SAW é necessário auxiliar o humano a atingir a Projeção Situacional. Assim, para que haja maior probabilidade de uma tomada de decisão eficaz e concisa com uma situação, é necessário melhorar de forma contínua a SAW (ENDSLEY; JONES, 2012; BOTEGA *et al.*, 2019).

2.2. Modelo de Consciência Situacional de Endsley

Neste trabalho será adotado o modelo de Endsley (1988), pois este modelo é dirigido por dados e informações. Este modelo é composto por três níveis, com elementos que se relacionam, considerando, ações externas de outras tarefas e sistemas, como demonstrado na Figura 1. Os níveis podem ser alcançados gradualmente de acordo com a evolução da situação e melhor entendimento pelo indivíduo em relação a ele. Os níveis de SAW são:

- Nível um: Percepção dos elementos no ambiente. É coerente com a percepção dos elementos relevantes existentes, bem como suas características.
- Nível dois: Entendendo a situação atual. Com base na composição dos elementos encontrados, é necessário criar uma relação, contextualizando-as de acordo com os objetivos e objetivos respectivos.
- Nível três: Projeção de futuros Estados. Este nível refere-se a estar ciente do que os elementos são e o que eles significam para a situação real. Para atingir esse nível, é necessário ter a capacidade de prever quais serão os futuros Estados dos elementos. O nível três é conseguido somente com sucesso quando há uma boa compreensão da situação, que é incumbido no nível precedente. Fazer projeções requer experiência no domínio no qual o indivíduo está definido.

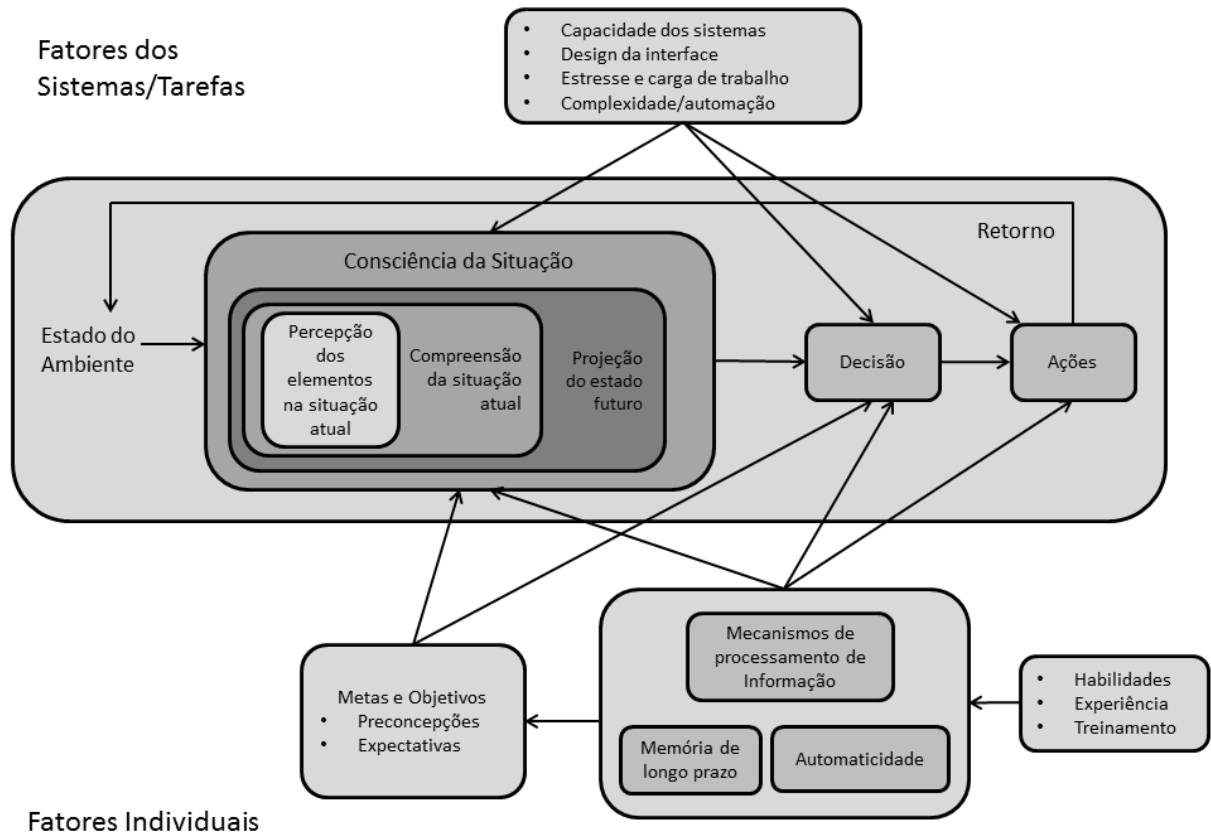


Figura 1 - Modelo de Consciência Situacional de Endsley. Fonte: Adaptado de Endsley, 1998.

Só é possível alcançar o Nível três, listado acima e que se refere a Projeção Situacional, através de uma boa compreensão das informações que embasam a situação, coletadas e estruturadas no Níveis um e dois. Outros fatores podem afetar atingir a Projeção Situacional, como o viés do humano que consome as informações, o sistema de informação que está sendo utilizado, como as informações são apresentadas via uma interface orientada a SAW, assim como fatores intrínsecos aos dados e informações, como a sua qualidade, forma como estão representados e até mesmo sua fonte. Estes fatores podem levar a falhas em todos os Níveis de SAW, distanciando o humano ainda mais da Projeção Situacional e tomada de decisão.

Ao analisar a Figura 1, pode-se notar que o modelo proposto por Endsley (1988) tem um núcleo de processamento das informações, sendo os Níveis 1, 2 e 3, sendo que estes sofrem ações de sistemas externos, estando disposto a fatores e informações do ambiente, humanos e sistemas informacionais/computacionais. Dentre estes fatores externos, é possível identificar Fatores Individuais, que estão diretamente relacionados ao humano que desenvolve o processo de SAW e também Fatores dos Sistemas e Tarefas, que dizem

respeitos a capacidade dos sistemas informacionais conseguirem lidar com o volume e complexidade das informações necessárias para o processo cognitivo humano.

2.3. Fatores Influentes no Processo

Tendo que SAW é um modelo mental do operador sobre uma determinada situação em um ambiente, e baseado nas informações que ele constrói dentro desse modelo, decide o que irá fazer naquele instante. Esse modelo mental ou núcleo de SAW é diretamente alimentado por fatores externos ao indivíduo, tarefas, sistemas e ambiente físico (KOKAR; ENDSLEY, 2012).

Sendo assim, temos que tais fatores podem afetar o processo de construção e alcance de SAW do humano, facilitando ou dificultando o mesmo. Sistemas de apoio a este processo influenciam por meio de sua interface e funcionalidades, isso conforme cada tarefa executada, em torno de sua complexidade, nível de atenção, carga de trabalho e estresse provocados no indivíduo. Ainda existem outros fatores que afetam de maneira significativa este processo, como ambiente físico, com ruídos, luminosidade, temperatura, entre outros (ENDSLEY, 2001; STANTON *et al.*, 2001).

Assim como o próprio indivíduo pode facilitar ou dificultar o seu processo de construção de SAW, pois este processo está diretamente relacionado a sua capacidade e habilidade em executar tarefas relacionadas ao domínio ou operação dos sistemas, nível de experiência, objetivos e metas. Que acabam por guiar a atenção do mesmo para pontos específicos no ambiente, sendo que cada indivíduo realiza esse processo de uma forma distinta (BOTEGA *et al.*, 2019; KOKAR, ENDSLEY, 2012).

Para isso, sistemas e mecanismos de processamento de informação tentam se enquadrar e suprir a forma como humanos, buscam e processam informações para SAW. Através da construção de sistemas que estimulam o indivíduo a manter a atenção distribuída diante de múltiplas fontes de informação, priorizando algumas informações ou detalhes mais relevantes, que apresentam maior impacto no núcleo de processamento de SAW.

Mas ainda existem alguns fatores limitantes intrínsecos à SAW que podem interferir negativamente o processo de SAW, desde a aquisição até a evolução. Pode-se encontrar oito fatores relacionados a essas limitações, sendo eles (ENDSLEY, 2012):

- Gargalo de atenção: Representa a incapacidade do humano em prestar atenção em mais uma informação em um dado instante.

- Limitação de memória: Humano com memória de curto prazo limitada, empregada no processo que forma a figura mental da situação atual com base na experiência e conhecimento do usuário.
- Fatores de estresse: Carga de trabalho, ansiedade e exaustão sobrecarregam e prejudicam a cognição.
- Excesso de dados: Mudança de dados rápida de dados dinâmicos dentro dos sistemas, dificultando o processo de interpretação dos humanos para manterem-se atualizados.
- Sugestões mal inseridas: Má disposição ou destaque das informações, guiando o operador para uma SAW imprecisa e fora do contexto real.
- Complexidade do sistema: Funcionalidades em excesso e maldispostas dentro dos sistemas, criam um obstáculo para a evolução do modelo mental sobre a utilização deste.
- Modelos mentais incorretos: Modelos mentais criados pelo operador que são fortemente embasados em situações familiares passados do mesmo, porém não estão condizentes com o cenário atual.
- Automação: Dependência extrema que humano pode desenvolver sobre a utilização do sistema.

2.4. Domínios de Risco e Consciência Situacional

SAW é um processo cognitivo realizado por todos humanos quando necessitamos compreender um ambiente ou situação para realizar uma tomada de decisão. Onde entende-se por situação o conjunto de humanos, elementos e fatores, variando conforme o domínio aplicado, mas compartilhando a característica de relacionamento entre tais componentes. Este fato amplia o domínio de pesquisa e estudos de SAW para quase todos os domínios, qualquer que envolva humanos, informações e a necessidade de tomada de decisão baseada nas informações.

Por isso, pesquisas que abordam essa temática são facilmente encontradas em grandes áreas críticas como a medicina, militar ou comando e controle, navegação aérea, marítima e terrestre, manutenção de equipamentos, entre outras (ENDSLEY; JONES, 2012). Em tais domínios, a falta ou erro de SAW pode provocar a perda de vidas humanas e danos irreversíveis a patrimônios.

Trabalhos como de Gaba *et al.* (1995), que tratam SAW no domínio da aviação e anestesiologia, confirmam todo o potencial de abstração e aplicação de SAW em diversos

domínios. Sendo que ambos, aviação e anestesiologia, apresentam dinamicidade, complexidade, um fluxo contínuo e intenso de informações e uma taxa de risco. E também relacionam quais aspectos de uma situação o humano deve manter seu foco e atenção, sendo três aspectos: sinais sutis, situações em evolução e elementos de conhecimento especial.

Outra área que demonstra o quão importante e crítica é a SAW é a medicina, Shelton *et al.* (2012) relaciona essa diretamente a segurança e vida do paciente, onde médicos e enfermeiros devem ter uma visão holística e atualizada do status do paciente a todo instante, através de acompanhamento dos sensores e exames. Compreender essas informações e associar o seu estado atual com uma evolução é essencial para executar o procedimento mais recomendado.

Em um estudo realizado por Botega *et al.* (2019), demonstra que SAW também é largamente aplicado ao domínio militar, devido à necessidade de tomada de decisões precisas em ambientes altamente dinâmicos e críticos, principalmente em relação ao tempo. De maneira que a qualidade da ação final é afetada pelo nível de consciência de um único indivíduo ou equipe sobre uma situação.

Operadores de sistemas voltados para o atendimento de emergências, como relatos de crimes às centrais de atendimento das forças de segurança, devem ter a capacidade de perceber algumas entidades básicas para o cenário e seus estados. No domínio criminal, por exemplo, pode-se ter vítimas e criminosos, suas condições físicas, descrições, atitudes e comportamentos. Considerando também as relações entre as entidades e o ambiente inserido, ainda sobre o serviço citado, pode-se caracterizar um crime, sabendo de uma pessoa está ameaçando ou ferindo outra de alguma forma, e por último, como a situação pode decorrer, novos ataques ou fuga (BOTEGA *et al.*, 2019).

Outros sistemas críticos que fazem proveito de SAW, porém mudando o modo de operação, são Sistemas de Suporte à Tomada de Decisão, que buscam auxiliar especialistas, consultores e analistas a compreender situações e realizar projeções situacionais em torno de um determinado ambiente, normalmente urbano, como indústrias, empresas privadas, espaços públicos e residências, possibilitando uma análise do nível de perigo e risco nestes ambientes.

Por exemplo, para realizar uma análise de risco em torno de um tipo de ambiente onde se deseja instalar uma empresa de um determinado ramo, para um especialista realizar esta análise sobre este ambiente, é necessário que ele tenha conhecimento de todo histórico de ações criminosas ou que ofereceram algum risco naquele local, com o mesmo, ou parecido, tipo de estabelecimento. Sendo que esse histórico é composto de diversos tipos de

dados, como históricos policiais, imagens de câmeras, registros de redes de comunicação, entre outros.

Este processo de análise do histórico de eventos de um local para determinar o nível de risco oferecido, é também uma ação de extrema dificuldade e que exige um nível de SAW muito alto do analista, a projeção situacional, onde precisa compreender diversos tipos de informação, como estas se relacionam e neste caso, ainda considerar um fator temporal, que é o quando ação aconteceu, como esse fator reflete no momento atual e o que pode acontecer em um futuro próximo. Levando em consideração que são informações de múltiplas fontes heterogêneas.

A projeção situacional é a capacidade de gerar uma estimativa sobre um conjunto de possibilidades futuras de uma situação, baseado nos elementos e padrões atuais de uma situação no presente. A capacidade de projetar situações é essencial pois SAW não é apenas sobre o que está acontecendo no momento presente, mas também sobre eventos e ações que vão acontecer a seguir. Considerando que o humano não consegue influenciar no presente de uma situação, apenas no seu futuro, tornando assim os elementos de uma situação atual insumos para a projeção situacional e compreensão de estados futuros (ROY, 2001).

Considerando que a falta de informações pode acarretar falhas graves durante o processo de SAW, que por sua vez afeta a ação de um humano, que compromete vidas e bens ao tomar uma decisão. Com foco em tais erros e na complexidade do processo de SAW, surgem mecanismos de apoio ao processo de SAW e projeção situacional, tratados na próxima Seção.

2.5. Suporte à Consciência Situacional e Projeções Situacionais

Partindo de erros, falhas e alta complexidade do processo e modelo de construção de SAW e tomada de decisão, surgiram diversas tecnologias, algoritmo e soluções diversas que auxiliaram de alguma forma todo esse processo. As principais soluções foram Sistemas de Apoio a SAW e tomada de decisão, todavia, dependendo do domínio e dados dos quais tratados, o sistema se tornava quase ineficiente, apenas trabalhando uma interface computacional, sem considerar a qualidade da informação que estava presente, afetando assim todo o processo de SAW (GABA *et al.*, 1995; ENDSLEY, 2001; HALL *et al.*, 2006; ENDSLEY; JONES, 2012).

Com os avanços na área aplicados ao domínio e trabalhos científicos relacionados, soluções novas e mais eficiente surgiram, agora tratando sistemas informacionais que englobam todo o processo perante as informações que serão apresentadas, desde a coleta a

partir de sensores físicos e não físicos, passando por processos de tratativa, transformação, representação, inferências e apresentação, os quais podem agir diretamente sobre o processo, alterando parâmetros e até mesmo complementando a informação, fazendo com que o resultado final seja mais condizente com a realidade (VINCEN *et al.*, 2009).

Alguns desses sistemas adotam modelos de transformação, como a fusão de informações, e modelo de representação, especialmente ontologias, que se consideram tanto o papel do humano durante o ciclo de vida dos dados e informações, quanto a melhor forma de lidar com informações provenientes de inteligência humana, garantindo uma sinergia entre os processos e fases, facilitando o desenvolvimento destes. E afeta também o resultado que é demonstrado ao humano tomador de decisão e sua decisão, trazendo informações mais coerentes sobre da situação, atualizada e precisa (BOTEGA *et al.*, 2019; PEREIRA JUNIOR; PEREIRA; BOTEGA, 2019; BOTEGA *et al.*, 2017; PEREIRA *et al.*, 2016).

2.6. Considerações Finais

Neste capítulo foi apresentado as definições e conceitos sobre SAW e projeção situacional, junto ao modelo de Endsley, o qual dispõe SAW como um processo incremental e segmentado em níveis, os quais podem sofrer ações de agente externos, como sistemas informacionais ou até mesmo fatores internos, como um viés ou experiência do humano. Este modelo embasa os demais assuntos do capítulo que é a contribuição da SAW para domínios de risco e como algumas pesquisas estão se posicionando, por fim, traz a capacidade de sistemas informacionais darem suporte ao processo de SAW e tomada de decisão, relacionando o modelo de Endsley com processos e artefatos informacionais, como ontologias e processos de fusão de informações, temas os quais serão abordados e descritos nos próximos capítulos.

3. Ontologias como Suporte à Consciência Situacional

Neste Capítulo será dedicado às ontologias, sua estrutura, classificação, relacionamento com a Web Semântica, os vocabulários utilizados para construí-la e metodologias de desenvolvimento e as representações de ontologias. Também serão apresentados sua relação com Consciência Situacional e como que a Ontologia pode auxiliar neste processo.

3.1. A Web Semântica e Ontologias

A Web Semântica, Web 3.0, ou mesmo Web de Dados é uma evolução da Web atual, uma Web de documentos. As informações na Web atual não representam seu valor real, sendo não relacionadas ao contexto e, portanto, irrelevante. Isso dificulta sua distribuição, acesso e interoperabilidade. Para Berners-Lee (2001), a web semântica não é uma Web separada, mas uma extensão do atual. Nele a informação é dada com um significado bem definido, permitindo uma melhor interação entre computadores e pessoas.

Junto com a evolução da teia semântica, surgiu outro conceito para atender as demandas da aplicação semântica à Web: o conceito de ontologias. A Web de dados precisa de formas e modelos mais ricos e expressivos de lidar com informações, e é aí que entram as ontologias.

Juntamente com a representação do significado e do contexto, as ontologias na Web atuam também na reutilização e recuperação de informações através da formalização, padronização e contextualização da informação. De acordo com Isotani e Bittencourt (2015), ontologias "[...] fornecem um quadro conceitual comum no qual podemos desenvolver bases de conhecimento compartilháveis e reutilizáveis. E em segundo lugar, facilitam a interoperabilidade e a fusão de informações, o que possibilita a criação de aplicações computacionais poderosas e inteligentes".

Visando à construção de uma ontologia, alguns elementos essenciais podem ser observados e, quando interligados, representam todo o conhecimento e poder contidos em uma ontologia. Para Kiryakov *et al.* (2005), uma ontologia pode ser representada pela notação $O = \{C, R, I, A\}$, onde cada letra do conjunto representa um elemento essencial para a formalização das ontologias. Cada letra é definida abaixo:

- C - um conjunto de classes que representam os conceitos do domínio que a ontologia representa, trazendo a essência de cada indivíduo nesse contexto.
- R - representa o conjunto de relacionamentos (ou Propriedades) estabelecidos entre os conceitos listados, essas associações criam as regras essenciais e que estruturam a base do domínio representado.
- I - conjunto de instâncias, são as materializações das classes, trazendo um conceito ou abstração (classe) para uma de suas formas no ambiente real, ou conceitos e definições da própria ontologia, por exemplo, um fator de qualidade da ontologia propriamente dita.
- A - é o conjunto de axiomas relacionados ao domínio essenciais para representar o domínio em um ambiente virtual, restrições de modelagem e regras contidas em classes, relacionamentos e instâncias.

No entanto, para ser viável, a web semântica e as ontologias devem estar associadas a alguns meios tecnológicos e computacionais, como linguagens de programação e de marcação. Três das principais tecnologias são o OWL (*Web Ontology Language*), o RDF (*Resource Description Framework*), o RDF-S (*Resource Description Framework Schema*) e o SPARQL. Todas essas linguagens são para representação e descrição formal de ontologias (ISOTANI; BITTENCOURT, 2015).

Uma das principais linguagens adotadas para representar e expressar os conhecimentos presentes em uma ontologia é a *Web Ontology Language* (OWL), que já faz parte de outra grande área de pesquisa, a Web Semântica, que junto as ontologias tornam possível representar incertezas de uma informação. Outras duas tecnologias que em conjunto OWL, fazem possível a implementação de soluções baseadas na Web Semântica, são o RDF e SPARQL. Onde RDF é uma forma de descrição de entidades e vocabulários, instanciando classes e propriedades de uma determinada ontologia, possibilitando utilizar propriedades e conceitos de uma ontologia dentro de uma aplicação em tempo de execução, tendo o SPARQL como uma linguagem de consulta e manipulação de RDFs instanciados na aplicação e presentes em outras fontes de dados externas (BAADER; HORROCKS; SATTLER, 200; KIM *et al.*, 2013).

Destes OWL é o mais aplicado e tem o maior poder de representação das formalidades de uma ontologia, especialmente quando se trabalha com a web semântica. Um dos fatores que leva a isso é que o OWL é baseado em duas outras especificações: RDF e RDF-S. Ele herda a estrutura do RDF, trabalhando com triplos e identificadores de recurso uniforme (URI), juntamente com a capacidade de descrição semântica contida no RDF-S.

Ele tem SPARQL como uma linguagem RDF-S de consulta e manipulação instanciada no aplicativo e presente em outras fontes de dados externas. É importante ressaltar que o OWL não é uma linguagem de programação, mas sim uma linguagem destinada a declarar recursos, seguindo uma certa representação (HITZLER *et al.*, 2012).

Quando consideramos o domínio ou cenário em que uma ontologia é inserida, outra forma de representar os conceitos e relacionamentos, tão importante quanto a representação formal, é a representação gráfica. Essa representação permite a interpretação visual das classes, relacionamentos (ou Propriedades) e instâncias da ontologia em questão. As ontologias representam ambientes reais e complexos. Possibilitar essa avaliação por indivíduos é essencial, pois traz robustez e maior confiança à ontologia.

3.2. O Papel da Ontologia para Representar Informações Contextuais

Conforme Schiessl e Brascher (2017), ontologia significa o "estudo ou conhecimento do ser, dos entes ou das coisas tais como são em si mesmas, real e verdadeiramente". Todavia, ontologias começaram a ser empregadas em pesquisas na Ciência da Informação e Ciência da Computação, sendo utilizadas como meio para organizar e representar informações e contexto de um ambiente, a fim tornar possível representar computacionalmente situações do mundo real (RAMALHO, 2006).

Uma das definições mais utilizadas para definir uma ontologia é a de Gruber (1993), no qual o autor traz a ontologia como sendo uma especificação explícita de uma conceitualização, onde se interpreta que a especificação explícita é concebida através da representação de conceitos e relacionamentos utilizando símbolos e conceitualizações, isto considerando que uma ontologia é tida como um modelo abstrato sendo construída e modificada conforme o domínio.

Guarino (1998) traz outra definição para ontologia, a qual permite entender o conceito como uma lógica que explica o significado pretendido de um vocabulário formal, ou seja, seu compromisso ontológico com uma conceitualização particular do mundo. Para Mizoguchi (2004), uma ontologia é um conjunto de conceitos e suas relações que capturam e formalizam a compreensão dos seres humanos e sua interpretação de um ambiente real. Sendo esta representação compreensível para outros agentes informacionais, humanos ou máquinas.

Para Santarém Segundo (2015) a utilização de ontologias é uma das formas que possibilita a construção de relações organizadas, isto considerando os termos que estão

presentes em um dado domínio, o que torna propício contextualizar dados, além de tornar um fator facilitador para as ferramentas e processos de interpretação e recuperação de informações.

Pode-se então compreender uma ontologia como uma formalização de conceitos e suas relações com o mundo real, para que ele possa ser representado e compreendido por computadores e sem perder a expressividade essencial sobre o domínio que eles representam. Ou seja, mantém o verdadeiro significado da informação em qualquer meio que se baseia no modelo formal construído e usado. (ISOTANI; BITTENCOURT, 2015).

Na Figura 2, pode-se notar dois dos principais recursos presentes em uma ontologia, classes (círculos), e propriedades (setas com termos), onde as classes representam entidades e ativos presentes em situações e as propriedades representam como estas classes podem estar relacionadas, como uma age ou é afetada por outra. Onde a representação dessas classes e propriedades dentro de uma situação específica, representam o real contexto da ação e todos os fatores envolvidos.

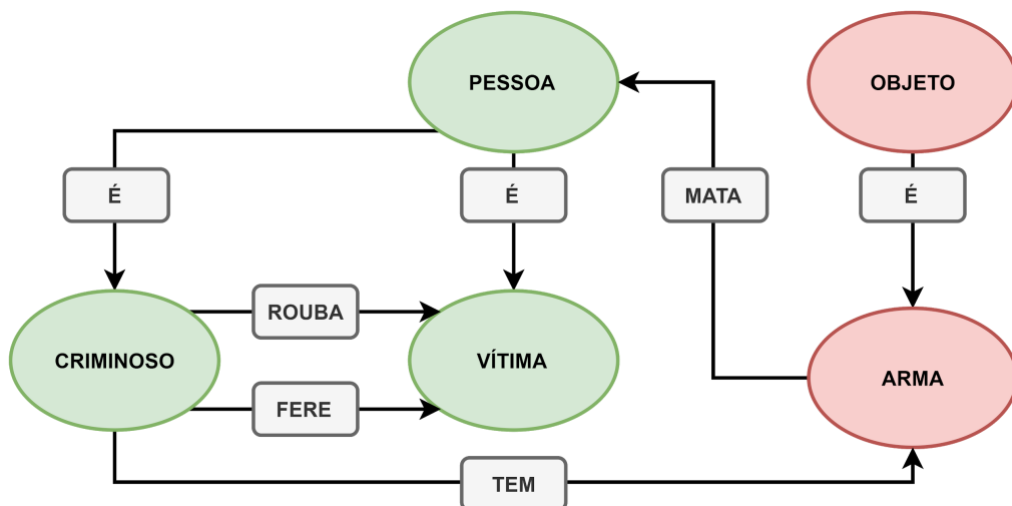


Figura 2 – Exemplo de ontologia aplicada ao domínio criminal. Fonte: o próprio autor.

3.3. Ontologias como Suporte a Consciência Situacional

A qualidade e a forma como as informações estão representadas podem influenciar diretamente no processo de aquisição e manutenção de SAW de um indivíduo e, conseqüentemente, na projeção situacional e futura tomada de decisão, a qual demanda de insumos informacionais precisos, atuais e completos para que seja assertiva.

Entretanto, sabe-se que a origem de tais insumos informacionais não possuem a devida qualidade, estrutura e representação aceitáveis perante às recomendações de um

especialista do domínio de gerenciamento de riscos e atendimento de emergências, pois as fontes informacionais são heterogêneas, complexas e com grande volume, tornando a informação no momento de origem precária e, somada aos fatores de estresse e modelos mentais dos indivíduos em uma situação de emergência, inviável de ser levada em consideração, o que pode gerar riscos à vida, ao patrimônio e ao meio ambiente.

Aplicar uma ontologia como forma de suporte e melhoria às informações de um domínio crítico pode influenciar diretamente a SAW, pois a essência da ontologia é possibilitar a descoberta de novas informações por meio de inferências e regras definidas em axiomas, fornecendo critérios e parâmetros baseados em índices qualitativos, uma vez que a qualidade reflete o valor e o quão confiável, atual, precisa e compreensível é uma informação.

Todas as classes, propriedades e outros elementos de uma ontologia remetem à percepção humana e modelos mentais sobre um domínio e suas informações, logo representar informações com uma ontologia ciente de qualidade é interessante, pois a contextualização informacional oriunda da ontologia somada aos fatores qualitativos pode resultar em melhores insumos para que o indivíduo melhore continuamente sua SAW.

Silva *et al.* (2018) desenvolveram uma ontologia, a qual possui meios para avaliar a qualidade da informação, objetivando produzir melhores insumos informacionais para o desenvolvimento de SAW. Os autores elaboraram um estudo de caso envolvendo um atendimento a uma situação de incêndio florestal, o qual foi utilizado para demonstrar a aplicabilidade da ontologia junto à gestão da qualidade, contribuindo positivamente para uma SAW dos operadores humanos envolvidos.

3.4. Considerações Finais

Neste Capítulo foram apresentados o que são ontologias, como estas se estruturam, um pouco sobre a sua classificação. Também foi tratada a relação entre ontologias e a web semântica, trazendo questões de como tornar possível a construção de um ambiente semântico dentro da web e quais tecnologias podem colaborar com a implementação e uso.

Ontologias são uma forma de representação de dados, portanto, estas podem ser utilizadas como uma base para o desenvolvimento de outras ontologias (ontologias 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.

O Capítulo seguinte trará questões sobre a Fusão de Dados e Informações e como este processo pode colaborar tanto para a melhoria da qualidade da informação e influenciar as

informações que irão estimular a SAW de humanos, tendo modelos e processos adaptáveis e que permitem a associação de outras formas de inferência.

4. Sistemas de Fusão de Dados e Informações

Este capítulo tratará sobre Fusão de Dados e Informações, demonstrando conceitos, técnicas, modelos e aplicações deste processo e como pode ser aplicado de maneira a colaborar com o domínio de situações de riscos, assim trazer contribuições aos processos informações que manipulam, inferem, produzem e consomem informações.

4.1. Fusão de Dados e Informações

Caracterizada como um processo de transformação e inferência de dados e informações, a Fusão tem o intuito de estimar e prever estados de entidades dentro de um sistema, tendo como objetivo principal maximizar o valor da informação. Isso com redução da dimensão dos dados, incorporação de valor à informação, evolução da representatividade e produção de insumo para o processo de construção do conhecimento sobre situações. Esse processo também afeta o processo de consciência situacional, estimulando de maneira mais eficaz um humano sobre uma dada situação em um ambiente (BLASCH *et al.*, 2013; STEINBERG, 1999).

Segundo LLINAS *et al.* (2004), a Fusão de Dados e Informações é um processo multinível e multifacetado de associação de dados de múltiplas fontes, com o intuito de captar e processar informações visando um resultado mais significativo, ao invés de consumi-las de fontes individuais, sem fusão.

Considerando o cenário tecnológico atual, com o número crescente de aplicações para os mais diversos fins. As quais produzem um número gigantesco de informações, impactando consideravelmente todo o fluxo de dados e informações. Se torna necessário algumas soluções para lidar com essa quantidade e formas distintas de dados, sendo nesse ponto que a fusão encaixa como campo de estudo multidisciplinar, aplicando conceitos e técnicas de diversas disciplinas: processamento digital de sinais, estimativa estatística, teoria de controle, inteligência artificial, e métodos numéricos clássicos.

Em um primeiro instante a definição de Fusão de Dados era dada como um conjunto de métodos para lidar com dados e informações de uma única ou diversas fontes, realizando associações, correlações e combinações de múltiplos entres estes. Com intuito de obter estimativas mais precisas sobre um determinado objeto e possíveis relações entre estes, sendo

o processo um ciclo contínuo de estimativas e validações, assim como a aplicabilidade de fontes externas adicionais ou alterações no processo (HALL; JORDAN, 2010).

Sistemas de Fusão de Dados e Informações compartilham de objetivos em comum, independente do domínio de aplicação, sendo estes a compreensão e predição do estado físico de uma determinada entidade. Essa análise é feita a partir das propriedades dessas, características, atividade, posição espacial e movimentação, todos considerando um determinado tempo, passado, presente ou futuro (LLINAS *et al.*, 2004). Por exemplo, se o objetivo é analisar e deduzir o estado de um objeto, é importante não considerar apenas o objeto. Pode ser de extrema relevância determinar os atributos relacionados ao ambiente que o objeto está inserido e seu contexto, assim como as formas de interação com demais fatores presentes.

Como grande parte das linhas de pesquisa aplicadas, um dos domínios que mais incentivou e subsidiou os avanços e larga implantação de novas linhas de pesquisa, fornecendo desafios reais e complexos, foi o domínio militar. Com a Fusão de Dados e Informações o principal apoio e incentivo partiu do Departamento de Defesa Norte Americano (DoD), variando as aplicações entre sistemas de rastreamento e detecção automática de alvos (ATR), identificação de tropas amigas, inimigas e neutras (IFFN) e variantes de Comando, Controle, Comunicação e Inteligência (C3i), visando auxiliar no processo de construção de SAW de operadores diante situações de ameaça contra vida e patrimônios (HALL; JORDAN, 2010).

Assim como existem diversos modelos propostos para representar SAW, existem alguns para Fusão de Dados e Informações. Entretanto, nesse trabalho será abordado o modelo JDL de Fusão de Dados, apresentado na próxima Seção, esse modelo é uma abordagem clássica, amplamente aplicado como referência para entendimento da área e conceitos correlatos. Serão apresentados seus níveis de atuação, relação com humanos durante o processo e com SAW.

4.2. Fusão de Informações e a Inteligência Humana

O início dos estudos sobre processos de Fusão de Dados e Informações foram impulsionados pelo domínio militar, com o intuito de analisar e monitorar fatores militares, veículos, tropas, sistemas de armamento, entre outras. Mas não apenas coletar e apresentar esses dados, mas adicionalmente tratar esses dados de maneira a reduzir grandes volumes

desses dados, deixando apenas os mais relevantes e realizando medições melhores, tendo como resultado informações mais significativas.

As primeiras soluções propostas, baseavam-se em processos e modelos voltados para o emprego exclusivo de sensores físicos e inteligência proveniente de sinais (SIGINT), algumas destas fontes são: sensores físicos dos mais variados tipos, radares, entre outros. Restringindo assim a atuação dos modelos, processos e Sistemas de Fusão de Dados à mineração e associação de dados resultantes da coleta de diversos dados destes sensores.

Esse modo de fusão é tido como Fusão de Informação de Baixo-Nível (*Low-Lever Information Fusion* - LLIF), que como dito se limita a estimar e associar dados de baixo nível provindos de sensores. Um fator de extrema importância, mas que é desconsiderado ou tratado com menor relevância nos Sistemas de Fusão de Dados que adotam essa classe de processo, é a participação de um humano durante o processo, que vai contra um dos principais objetivos que é ser insumo para o processo de construção de SAW e futura tomada de decisão (SALERNO, 2002).

Esse é uma característica que está mudando com os avanços dos estudos de aplicação de Fusão nos mais diversos domínios, com demandas diferentes e também estudos para melhorar a interação humano-computador. Parte desses avanços se deve pela crescente demanda de utilização de sensores não físicos, como: banco de dados diversos que detêm conhecimento de um domínio específico, informações provindas de redes de comunicação mundial (*World Wide Web*) e principal o uso da inteligência humana (HUMINT) como input do processo e método de análise de um cenário. A adoção dessas novas fontes de informações dentro dos processos do Modelo JDL e SAW, garantem um resultado melhor para o humano, promovendo uma visão holística de um cenário (BLASCH; VALIN, 2012; ENDSLEY, 2001).

A adoção de tais sensores não físicos, apresenta grandes avanços para Sistemas de Fusão de Dados, mas também grandes desafios, pois a informação HUMINT está propícia a erros, tendências, imprecisão e omissões. Logo para ser possível adotar técnicas de fusão baseadas nessas informações, se torna necessário um nível mais alto de interpretação da informação, requer uma dedução da semântica em torno daquela informação. Onde, na maioria dos casos, não poder ser inferido e manipulado por funções automatizadas, necessitando do humano de forma ativa para interpretação e uso dessas informações. Um exemplo claro e comum desses desafios, são denúncias feitas por testemunhas ou vítimas à serviços de risco, como o 190 (VINCEN *et al.*, 2009; SALERNO, 2002).

A partir deste momento, é possível observar uma alteração no papel do humano perante Sistemas de Fusão de Dados quanto a passividade no Modelo de SAW. Onde esta passa a exercer ações importantes durante o processo. Chegando em um ponto relevante, discutido por Nilsson (2012), que seria a necessidade e oportunidade do trabalho colaborativo entre humano e processos de Fusão de Informações. Junto à essas mudanças de perspectivas, pesquisadores e intuições imersas no universo da fusão, começaram a levantar questionamentos sobre Fusão de Informações em Alto-Nível (*High-Level Information Fusion* - HLIF), a real participação do humano durante os processos, tratativas e processos as quais as entidades, provindas de sensores físicos e não físicos, são submetidas, com o intuito de adaptar e suportar as novas formas de inferência, classificação e mineração mais alto nível (COSTA *et al.*, 2005; BLASCH; VALIN, 2012; SALERNO, 2002).

Botega *et al.* (2016), demonstra e propõe um novo modelo de fusão de informações, onde já considera o papel ativo de humanos durante o processo de fusão e tomada de decisão, com foco em sistemas críticos para análise de situações de risco, sendo o Modelo Quantify (*Quality-aware Human-driven Information Fusion Model*). Sendo todo o modelo pensado para auxiliar o humano na construção de SAW durante a operação de sistemas críticos analisando cenários complexos, como sistemas de análise e gerenciamento de risco. Para isso, existe o emprego de diversas tecnologias, modelos e métodos para satisfazer as necessidades do humano especialista como consumidor do produto final, SAW, e também autor, fazendo um trabalho colaborativo síncrono entre homem e computador.

4.3. Adequação dos Métodos de Fusão para Suportar Informações em Alto-Nível

Costa et al (2012), faz uma distinção clara e direta entre processos LLIF e HLIF de fusão, onde LLIF infere e combina informações provindas de sensores para identificar, classificar e localizar objetos isoladamente. Enquanto HLIF lida com informações relacionadas a diversos objetos, considerando também o contexto que estão inseridos, com o objetivo de definir situações complexas, realizar medições sobre possíveis intenções de atores.

Um fato que se tem claro é que seres humanos têm uma capacidade muito mais avançada do que os computadores, em analisar situações ou cenários e relacionar estas, ou partes destas, em uma imagem da situação. Mas ainda apresenta um grande problema para todos os domínios, altos índices de erros, falhas e imprecisões (COSTA *et al.*, 2012).

Pode-se exemplificar esse fato através de um cenário de uma situação de risco, onde precisa-se levantar suposições e culpados de um sequestro que aconteceu a três dias. Um investigador consegue observar arquivos de vídeo e imagem, e achar de maneira rápida e ágil algum relacionamento entre esses artefatos e traçar alguma relação prévia entre entidades ativas no vídeo e que podem estar nas imagens também, mas em estados diferentes. Como um homem de boné vermelho passando em frente a um banco e no momento seguinte saindo com o carro em alta velocidade. Baseado em seus conhecimentos pessoais e profissionais acumulados, ele chega a essa conclusão de maneira discreta e implícita, que pode ser o início da solução de um crime.

Para um computador processar as mesmas informações e inferir as mesmas condições, este necessita possuir o mesmo tipo de conhecimento. Mas para tornar esse todo esse conhecimento passível de entendimento de um computador, é necessário um conjunto de regras e formalismos, para que assim posso fazer o uso da melhor forma possível sempre que necessário (KIM *et al.*, 2013; COSTA *et al.*, 2012).

A principal e mais atual tecnologia para formalizar e explicitar conhecimentos sobre um determinado domínio, são as Ontologias. Auxiliando na organização, propagação, acesso, entre outros. Mas a principal contribuição das ontologias para o uso computacional é poder fazer com que máquinas tenham acesso e consigam organizar conhecimento e informações de uma maneira mais formalizada, mais próxima de como o conhecimento humano é estruturado, e até mesmo gerar um conhecimento melhor, em parceria com outras técnicas, como inteligência artificial e *machine learning*. McGuinness (2003), Blasch *et al.* (2012), são alguns dos grandes autores que adotam ontologias em alguns dos seus trabalhos para tornar possível a realização de processos para Fusão de Informações de Alto Nível.

Algoritmos de Fusão de Informações de Alto-Nível que adotam ontologias para inferir e representar conhecimentos, realizam estas ações em cima da presença e organização destes recursos, conforme disponibilidade. Considerando que as ontologias representam uma visão holística da situação, onde tentam prever todos os cenários possíveis dentro de um ambiente (NOY; McGUINNESS, 2001).

O emprego dessas técnicas em conjunto, são capazes de dar o poder necessários para Sistemas de Fusão de Dados de elaborar ambientes e situações mais complexos e realizar as devidas inferências antes de expor o resultado ao humano, sendo também mais flexível as possíveis interações humanas dentro do processo de fusão.

4.4. Qualidade de Dados e Informações na Fusão de Informações

O conceito de qualidade é variável e subjetivo, alinhando as necessidades, ações e objetivos de cada domínio, caracterizando-se como parâmetros para que o produto ou resultado esteja de acordo com as necessidades. Como o produto analisado está livre de problemas e/ou falhas, ele permite que os fatores dependentes do produto analisado sejam executados com êxito. Como a qualidade é inerente a vários domínios e pode ser aplicada a dados, informações e outros produtos, os problemas e as dimensões da qualidade também são mutáveis (OLETO, 2006).

Abordar diretamente a qualidade dos dados visa alinhá-lo com as necessidades do domínio de que eles vêm ou são aplicados. Os dados que não descrevem fielmente os elementos do ambiente real interferem na efetividade dos Sistemas de Informação (SI) e podem gerar falhas no próprio sistema. Em atividades do mundo real, a qualidade dos dados é um elemento que possibilita a SAW e a tomada de decisões em um domínio (PEREIRA JUNIOR; PEREIRA; BOTEGA, 2019).

Para Todoran *et al.* (2015), novas questões e problemas relacionados à quantidade e qualidade dos dados estão surgindo em um ambiente de Tecnologias de Informação e Comunicação (TIC) em rápida mutação. Os autores também fazem uma comparação entre as tarefas que o sistema de informação realizou e seus requisitos, destacando o uso de tarefas comuns no passado, como problemas de classificação e processamento de dados homogêneos, e como os requisitos e as tarefas evoluíram, se tornando cada vez mais complexas, mesmo lidando com a SAW.

Considerando essas mudanças nos sistemas de informação, juntamente com as mudanças no meio ambiente, o espectro informativo se expande consideravelmente. Existem diferentes fontes de dados de diferentes tipos de sensores, tais como sensores físicos, robôs e algoritmos de mineração, sensores de imagem e sensores humanos (ou seja, de inteligência humana). Com a crescente quantidade de dados heterogêneos sendo capturados e armazenados nas mais diversas coleções de dados, lidar com todos esses dados em toda a sua amplitude não é mais o único problema a ser resolvido. Considerar, avaliar e melhorar a qualidade também se tornaram uma necessidade e preocupação.

De acordo com o English (2009), nas últimas décadas foi destacado um novo conceito, a interferência da qualidade dos dados na execução e na capacidade dos SI. A má qualidade impacta negativamente várias facetas do ambiente de informação, a partir do custo

do SI, através da representação através da interface gráfica e afetando níveis mais críticos, como a construção da SAW e tomada de decisão.

4.4.1. Incorporando a Qualidade de Dados e Informações nos Processos da Fusão de Informações

Mapear entidades complexas, como seres humanos e suas interações com o mundo real, é um processo desafiador. A natureza dinâmica e complexa das interações entre pessoas, objetos e lugares exige o uso de técnicas computacionais compreensivas para revelar seus Estados ao longo do tempo.

O processo que os sistemas de fusão usam para entender as interações humanas começa com a busca e determinação de quais entidades estão presentes em um cenário real. Em seguida, os Estados das entidades são determinados, formados por suas características físicas, posição, orientação e outros dados relevantes para o domínio. Finalmente, os sistemas de fusão estabelecem as possíveis relações entre entidades, relacionando cada contexto de entidade e estado entre si. Essas relações podem ajudar os humanos e os sistemas a entender as situações.

Alcançar o SAW é um processo desafiador que os sistemas de avaliação da situação procuram apoiar. Com o advento dos modelos SAW, especialmente o modelo Endsley (1988; 2001), novos modelos de IF, arquiteturas e processos emergiram visando apoiar o desenvolvimento de sistemas de avaliação de situação.

Nesse contexto, o conhecimento da qualidade da informação pelos sistemas de fusão de informações surge como um recurso complementar ao processo de inferência de situações. Além disso, os sistemas de fusão de informações atuam como tecnologia para apoiar os modelos e processos existentes para a aquisição e manutenção da SAW.

Os benefícios de conhecer a qualidade da informação em um processo de fusão de informações em cada um dos possíveis níveis de inferência do modelo JDL são a confiabilidade das fontes de dados e a efetividade dos algoritmos de preparação de dados (nível 0), a completude e exatidão da identificação de objetos (nível 1), a integridade de uma relação entre objetos (nível 2), a assertividade de uma projeção (nível 3), o refinamento do processo (nível 4), que pode incluir desde a preparação dos dados e abranger questões de interação humano-computador com a interface e representações gráficas. O conhecimento da qualidade da informação não só pode apoiar as decisões em cada nível, mas também influenciar e parametrizar as rotinas de inferência interna.

Além de contribuir para a operacionalização dos mecanismos internos dos níveis de fusão de dados, o conhecimento sobre a qualidade dos dados e informações também pode contribuir para as relações entre os níveis de fusão, ou seja, para ajudar a determinar e direcionar informações em virtude das saídas e entradas desejadas para cada nível. Índices de qualidade operam como alavanca para determinar a utilidade da informação ao longo do processo de um nível para outro. Esta rotina contribui ainda mais para preencher a lacuna entre as inferências de fusão de informações de baixo nível e de alto nível (BLASCH, 2013).

Entre os desafios de incorporar dados e qualidade da informação em um processo de fusão, podemos destacar o papel da qualidade da informação e a gestão da dinâmica da informação para a definição da qualidade da informação. Em relação ao primeiro desafio, sabe-se que os processos informatizados de um sistema de fusão de informações (por exemplo, mineração, integração e correlação) inferem novas informações de forma distribuída, assíncrona e dinâmica. Para colaborar uns com os outros, esses processos devem ter um mecanismo que qualifica cada novo dado ou informação produzida com um indicador de qualidade (metadados de qualidade). Assim, a parametrização do processo ganha uma nova variável (diferente de atributos ou objetos) que deve ser considerada toda vez que um novo processo de fusão é executado. Esta rotina contribui para a informação de qualidade atingindo os níveis superiores do processo.

Além disso, ao ajudar a parametrizar a automação, o segundo desafio é representar adequadamente a qualidade da informação e estimular as interações dos operadores em interfaces de usuário especializadas e visualizações dedicadas ao SAW. Para estimular visualmente a percepção dos operadores na busca por padrões e relacionamentos, é necessário usar pistas ou sugestões que qualifiquem as informações. Esses sinais ajudam a justificar o comportamento humano e explicam por que as informações são aceitas ou não e, conseqüentemente, podem ajudar a orientar os operadores a melhorarem a qualidade da informação por meio de um processo de refinamento contínuo.

4.5. Considerações Finais

Neste Capítulo foram apresentados o que são ontologias, como estas se estruturam, um pouco sobre a sua classificação. Também foi tratada a relação entre ontologias e a web semântica, trazendo questões de como tornar possível a construção de um ambiente semântico dentro da web e quais tecnologias podem colaborar com a implementação e uso.

Ontologias são uma forma de representação de dados, portanto, estas podem ser utilizadas como uma base para o desenvolvimento de outras ontologias (ontologias 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.

O Capítulo seguinte trará questões sobre a Fusão de Dados e Informações e como este processo pode colaborar tanto para a melhoria da qualidade da informação e influenciar as informações que irão estimular a SAW de humanos, tendo modelos e processos adaptáveis e que permitem a associação de outras formas de inferência.

5. Modelo de Transformação e Representação de Informações para Suportar Projeções Situacionais

O presente trabalho tem como principal proposta um modelo de transformação e representação de informações capaz de contribuir e dar suporte à humanos no processo de projeção situacional. Para isto foi feito um levantamento de trabalhos relacionados que pudessem contribuir para a consolidação da proposta, buscando por trabalhos que abordam algumas das temáticas deste trabalho, como SAW, representação de informações através de ontologia, modelo e processos de fusão que são aplicados como processo de transformação neste trabalho e por fim, trabalhos que fazem o uso de índices de qualidade de dados e informações.

5.1. O Modelo Tree

Esta seção se dedica a apresentar o Modelo de Transformação e Representação de Informações para Suporte à Projeções Situacionais (*information Transformation and REpresentation modEl for situational projections support - Tree*), que possam auxiliar o humano no processo de construção de SAW e tomada de decisão. Para isso, são necessários alguns módulos capazes de lidar com as informações da devida maneira, neste caso, capazes de lidar com dados provenientes de humanos, problemas de qualidade intrínsecos aos dados, informações semânticas e SAW.

Este modelo cria um arranjo técnico informacional que pode contribuir tanto para o avanço das ciências, quanto para a melhoria da qualidade da informação e principalmente para o processo de SAW, projeção situacional e tomada de decisão. Onde traz todos os avanços dos processos encontrados nos trabalhos relacionados, corrigindo suas falhas e absorvendo os pontos de sucesso.

Considerando que um modelo representa um arranjo organizado de técnicas, tecnologias e até mesmo outros modelos, com intuito de atender as necessidades informacionais de um domínio, aplicação ou problemática. Ao considerar a estruturação de um modelo que prevê a transformação e representação de informações que irão auxiliar o humano na construção de projeções situacionais, este vai de encontro com alguns pontos da CI apontados por Borko (1968), como a coleta, organização, armazenamento, recuperação, interpretação e transformação da informação.

A literatura já apresenta trabalhos que discutem e trazem modelos aplicando a fusão semântica de informações, índices de qualidade e ontologias. Alguns destes também começam a considerar processos que possibilitem a projeção situacional ou a organização de demais processos para que possam auxiliar o humano a atingir a projeção situacional, como apresentado por Pereira Junior, Pereira e Botega (2019). Porém tais trabalhos consideram a avaliação dos índices de qualidade como um processo interno, que cria insumos para os processos de inferência, representação e recuperação, não dando a devida base à qualidade dos dados, sendo que essa deve estar presente em todos os processos que manipulam a informação.

Botega *et al.* (2019) também apresenta um modelo de fusão de informações (Quantify), capaz de fazer análises sintática e semântica de informações, junto a avaliação de qualidade visando melhorar SAW. Sendo que a qualidade aparece como uma das bases do modelo, porém não é discutido qual a sua relevância para o modelo e como esta influenciou no arranjo tecnológico que foi pensado, além de não considerar a capacidade de produzir, representar e avaliar informações sobre um futuro, as projeções informacionais, dando indícios de que é possível sanar problemas informacionais ao considerar a qualidade de dados dentro dos processos que manipulam informações.

Com isso o presente trabalho adota técnicas, processos e modelos discutidos nas próximas seções, como a aplicação da *Human-driven Brazilian Crime Ontology* (URBANITY), ontologia capaz de representar situações criminais que tem como principal fonte relatos humanos sobre crimes, junto a índices de qualidade e a metodologia *Information Quality Assessment Methodology in the Context of Emergency Situational Awareness* (IQESA), desenvolvida com o foco em avaliar e representar os índices de qualidade para que sejam aplicados junto ao processo de avaliação de informações. Agregando também técnicas de fusão semântica de informações que consideram critérios semânticos e de qualidade, aplicando este durante os processos de busca e associação de informações. (PEREIRA JUNIOR; PEREIRA; BOTECA, 2019; BOTECA *et al.*, 2019; MELO; BOTECA; SANTARÉM SEGUNDO, 2017).

Considerando os apontamentos feitos por Melo, Botega e Santarém Segundo (2017), que traz os problemas de qualidade não relacionando apenas aos dados diretamente, mas considera também a estrutura provida para sua publicação, sendo que este pode gerar falhas e dificuldades no acesso, assim como impossibilitar o devido consumo, resultados que embasam e direcionam as discussões deste trabalho para trazer a qualidade como processo e ativo e participativo dentro de todo o modelo. Junto com Pereira Junior, Pereira e Botega

(2019), onde afirmam que a melhoria na qualidade dos dados pode influenciar positivamente a SAW, permitindo assim que humanos atinjam a projeção situacional através do consumo dos dados de melhor qualidade.

No âmbito da transformação e inferência de informações utilizando de ontologias e aplicando a avaliação contínua de qualidade, tem sido um tema amplamente abordado em alguns trabalhos, Botega *et al.* (2019) demonstram um processo consolidado de fusão de dados e informações semânticas, Pereira Junior, Pereira e Botega (2019) apresentam avanços sobre os processos de busca e associação de informações dentro de um processo de fusão de informações, demonstrando quais tecnologias e metodologias podem colaborar para os processos, buscando as projeções situacionais.

Entretanto, os trabalhos citados acima não abordam um ponto relevante que relaciona todos os processos apresentados até agora: SAW, fusão de informações, qualidade e ontologias, sendo este ponto a projeção situacional dentro dos processos informacionais e computacionais descritos no trabalho. Em outras palavras, explorar a capacidade de buscar, associar, inferir e representar informações que representem tendências ou um futuro, que possam colaborar diretamente com o nível de SAW de um humano viabilizam uma tomada de decisão mais assertiva.

Ao analisar esses pontos, é possível perceber que a projeção situacional é induzida através das informações que são apresentadas ao humano, juntamente com os conhecimentos que o mesmo já detém (KOKAR; ENDSLEY, 2012). Com isso, pode-se notar que ambos, humanos e processos informacionais, devem ser capazes de lidar com a projeção, a qual para o humano é um processo comum e intrínseco. Mas para meios informacionais, lidar com informações complexas e de alto nível semântico não é trivial, uma vez que existem diversos processos inerentes.

Este trabalho propõe uma forma de lidar com a projeção de informações no meio digital partindo de como as informações são representadas, ou seja, expandir a ontologia URBANITY para possibilitar a representação de informações sobre o futuro, ou que indiquem possíveis evoluções da informação e aplicar a mesmo dentro de um modelo que suporte a transformação e representação de informação. Através de estudos e ensaios, foi identificado que é possível por meio da criação de novas classes, propriedades e axiomas, a criação de projeções situacionais na ontologia. Assim, uma vez que o modelo ontológico é capaz de representar tal estado, demais processos informacionais podem se aproveitar destas expansões.

A partir destas considerações foi construído o Modelo Tree. A Figura 3, apresenta o modelo, o qual é composto por sete camadas de procedimento, sendo: “Coleta de Dados”, “Transformação dos Dados”, “Fusão Semântica”, “Representação das Situações”, “Inferências de Projeções Situacionais”, “Avaliação de Qualidade” e “Modelo de Representação”.

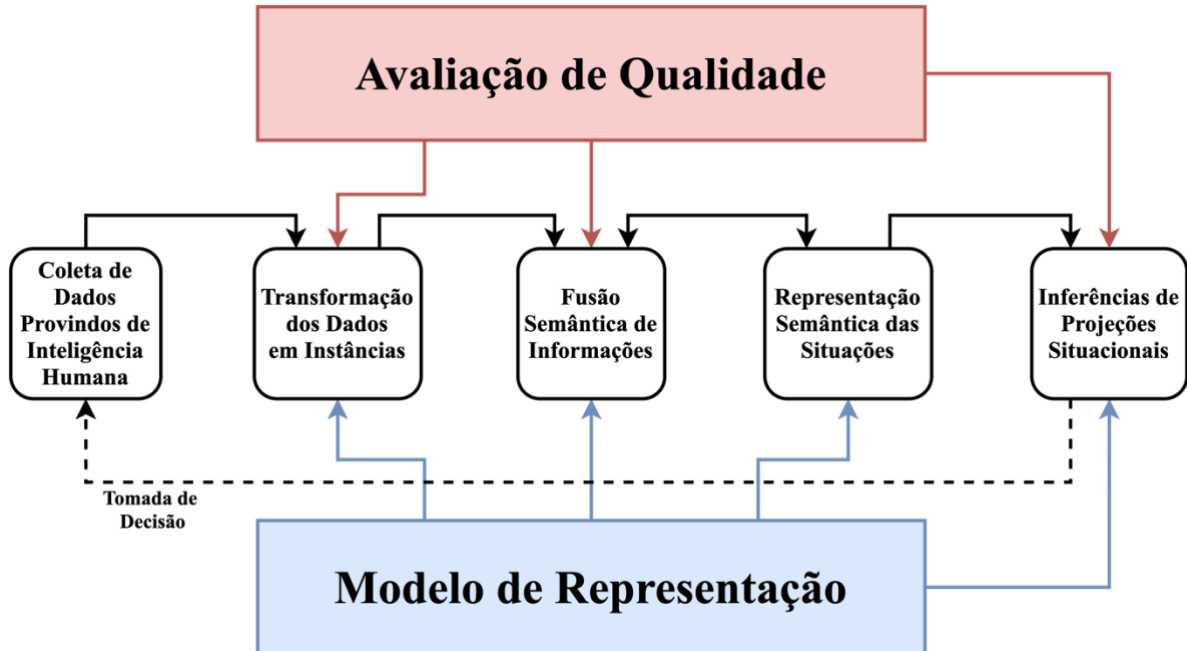


Figura 3 - Modelo Tree. Fonte: o próprio autor.

O modelo Tree tem como objetivo gerar melhores insumos informacionais, agregando semântica e índices de qualidade, que deem suporte a SAW e projeções situacionais em domínios críticos, especialmente domínios que lidam com dados providos de inteligência humana e que sofram com problemas de qualidade, gerando incerteza e afetando a capacidade de interpretação e tomada de decisão dos humanos. Sendo que o processo é suportado pela avaliação de qualidade e representação em ontologia.

Pode-se destacar entre as principais capacidades do modelo Tree os mecanismos para:

- Coletar dados providos de humanos, principalmente da web e realizar o seu tratamento;
- Avaliar e representar a qualidade de dados e informações durante os processos de transformação, fusão e projeção;
- Transformar dados estruturados e não estruturados em instâncias ontológica, seguindo um modelo de representação;

- Suportar a representação contínua das informações e suas transformações pela fusão e projeção de informações.

A operação do modelo em questão se dá de forma cíclica e iterativa, possibilitando a dinamicidade entre as camadas e processos internos, destacando a interação e iteração direta entre as camadas “Fusão Semântica de Informações” e “Representação Semântica das Situações”.

O modelo Tree se apropria de recomendações de modelos anteriores e supera limitações dos mesmos, onde é possível notar que a qualidade não está apenas relacionada à avaliação e representação de índices de qualidade, ela permeia todo o ciclo de vida das informações dentro de um modelo, servindo como parâmetros de busca, capacidade de inferência dentro de uma ontologia e até mesmo estimulam o usuário final em alguns domínios. Por isso a necessidade de considerar a mesma a partir do momento de estruturação de um modelo que siga uma abordagem informação perante as bases da CI.

Com a intuito de validar a discussão e alcançar os primeiros passos na direção de concretizar um modelo de transformação e representação de informações totalmente calçado por índices de qualidade, foi elaborado o modelo demonstrado na Figura 3, associando todos os possíveis módulos do modelo e destacando como os índices de qualidade e o módulo de qualidade, estarão diretamente associados a todos. Neste caso considerando desde a coleta dos dados, o processamento e transformação realizados por um processo de fusão de informações semânticas, a avaliação de qualidade e um modelo de representação de dados, que no caso será aplicado a URBANITY, devido ao domínio considerado.

Ao analisar o modelo proposto, é possível notar a presença contínua, desde a transformação dos dados coletados em instância seguindo o Modelo de Representação, aqui representado por Transformação de Dados em Instâncias, os quais os dados já são avaliados, assim, o processo de fusão já se torna capaz de inferir e produzir novas informações baseado no conteúdo, semântica e índices de qualidade das informações.

Os resultados produzidos por esse módulo são novamente avaliados, podendo trazer novas dimensões e métricas ainda não consideradas. As novas informações e seus respectivos índices de qualidade, são representados aqui na “URBANITY” que novamente tem a capacidade de interagir com a avaliação de qualidade para além de poder representar semanticamente os índices de qualidade, através de vocabulários específicos, pode reavaliar alguns índices de qualidade, como a consistência e completude, agora sobre uma visão

semântica, possibilitando inferir propriedades e dimensões ainda não consideradas, enriquecendo mais o resultado da fusão.

5.2. Ontologia URBANITY: A Ontologia Criminal

Assim como o Modelo Quantify prevê a avaliação contínua dos índices de qualidade, desde sua aquisição até a representação por meio da interface do usuário, permite também a possibilidade de fazer a representação da informação, através da sintaxe usando o relacionamento estruturas, tais como bases de dados, ou através de análise semântica, através de ontologias. Neste caso, tratando-se de um processo semântico, onde a representação das informações deve ser capaz de representar o maior e melhor contexto possível, as ontologias serão adotadas.

A ontologia aqui trabalhada, assume o nome de URBANITY (*Human-driven Brazilian Crime Ontology*), já em desenvolvimento, pelo GIHC (Grupo de Interação Humano-Computador). A URBANITY está inserida dentro do contexto de informações criminais no Brasil. O foco da mesma é representar informações oriundas de inteligência humana, neste caso, Boletins de Ocorrências (BOs). Junto a capacidade de representar e inferir índices de qualidade nos registros criminais que estarão presentes, trazendo mais dinamismo e possibilidade de inferências, juntamente com a metodologia de avaliação da qualidade IQESA (PEREIRA JUNIOR; PEREIRA; BOTEGA, 2019).

Segundo Pereira Junior, Pereira e Botega (2019) a ontologia em questão foi construída através da análise de conteúdo de dados criminais, mais especificamente, Boletins de Ocorrência (BOs) sobre roubo e furto acessado através da lei n° 12,527, o acesso à Lei de Acesso à Informação (LAI), sendo regulado pelo Decreto n° 7,724, pois apresentam informações em maior número e relevância, como dados históricos e descritivos sobre os acontecimentos. Os quais trazem detalhes importantes sobre o ambiente real e possibilita identificar e caracterizar os principais elementos que compõem e estão presentes em situações de crime.

Com o devido acesso à ontologia em questão e usando o software Protégé, a ontologia foi importada para o devido estudo e realizar as alterações necessárias para satisfazer os requisitos de projeção do presente trabalho. Uma representação gráfica também foi gerada, visando possibilitar uma fácil interpretação, tornando mais viável encontrar falhas, realizar correções e evoluir a ontologia.

Contendo diversas classes que representam os principais elementos de uma situação criminal, a URBANITY traz classes como Criminoso, Vítima, Objeto Roubado, Arma e Local, o Quadro 1 traz detalhes das principais classes da ontologia e que compõem uma situação. Além disso, há propriedades que relacionam as classes e criam uma dependência semântica, fator que justifica a não aplicação de um banco de dados, que por definição criaria uma relação de dependência entre classes, entretanto sem expressar contextos e relacionamentos semânticos.

Dentre as propriedades existentes na ontologia adotada, alguns exemplos são: foge, fere, rouba, tem. Por exemplo, a propriedade “tem” relaciona Vítima com Objeto Roubado, logo a vítima pode ter um objeto. Outro exemplo é dado pela propriedade “foge”, a qual relaciona Criminoso com Local. A Figura 4 traz a representação gráfica da URBANITY, demonstrando classes (retângulos com círculo amarelo), propriedades (setas) e algumas instâncias das classes (retângulos com losango roxo). É importante ressaltar que a estrutura da ontologia está em inglês devido ao perfil do grupo que a desenvolveu.

Quadro 1: Classes da URBANITY, descrevendo o que cada classe presente representa para a situação.

Descrição das Classes da Ontologia	
Classe	Descrição
<i>Situation</i>	Classe de nível mais alto que representa toda a situação possível de ser representada utilizando esta ontologia, sendo específica pela classe Crime
<i>Crime</i>	Esta classe é responsável por definir quais são as possíveis situações a serem representadas, no caso, os crimes de Roubo e Furto
<i>Robbery</i>	Classe que representa a especificação de situações de Roubo, representado quais as propriedades com as demais classes que caracterizam um roubo para o domínio
<i>Theft</i>	Classe que representa a especificação de situações de Furto, representado quais as propriedades com as demais classes que caracterizam um roubo para o domínio

<i>Person</i>	Responsável por agrupar os possíveis tipos de pessoa que estão presentes nas situações de Roubo e Furto, no caso, Criminoso, Vítima e Denunciante
<i>Criminal</i>	A classe Criminoso representa o elemento a Pessoa que atua nos crimes, realizando determinadas ações sobre a Vítima, na maior parte das vezes realiza a ação com algum elemento de Arma, subclasse da classe Objeto
<i>Victim</i>	Representa a Pessoa que sofre com a ação representada, Roubo ou Furto, sendo influenciada pelas propriedades com criminoso, na maior parte das vezes está associada a algum objeto que sofre a ação em conjunto
<i>Denouncer</i>	Classe que representa o papel de alguém que apenas relatou uma situação, em alguns dos casos a classe Vítima assume esse papel em conjunto
<i>Local</i>	Classe responsável por representar os locais da situação, estando diretamente relacionada às classes de Roubo e Furto
<i>Object</i>	Demonstra qualquer objeto que esteja relacionado às ações de Roubo e Furto, tanto os do Criminoso, principalmente arma, quanto da Vítima
<i>QualityGraph, Quality Measurement Data, DataSet, Category, Concept, Observation, Metric, MeasureProperty e Dimension</i>	Classes utilizadas para representar informações dos índices de qualidade avaliados, pertencentes ao vocabulário <i>Data Quality Vocabulary</i> (DQV)
<i>Education, Entertainment e Transport</i>	Classes externas utilizadas para enquadrar os tipos de locais
<i>Drug, Appliance, Document, Vehicle, Munition, Weapon, Clothing, Accessory e Electronic</i>	Subclasses de Objeto utilizadas para especificar qual tipo está presente na Situação

Fonte: o próprio autor.

Todavia, a URBANITY em sua estrutura original não prevê ou suporta classes, propriedades e axiomas que dão a projeção situacional, assim como a representação de tais projeções. Tendo isso, a próxima Seção descreve os avanços feitos na URBANITY que foram necessário para alcançar uma forma de conseguir representação a projeção de informações e situações dentro da ontologia.

5.2.1. Avanços na URBANITY para suporte às Projeções Situacionais

A expansão da ontologia URBANITY para lidar com projeções situacionais segue a metodologia de Noy e McGuinness (2001), a metodologia 101 (*Ontology Development 101*). De acordo com Isotani e Bittencourt (2015), esta é a metodologia mais adotada para projetos de construção de uma ontologia, devido à sua simplicidade de elaboração, que pode ser aplicada a vários casos ou também aceitar mudanças e ajustes de acordo com a necessidade de cada domínio.

A metodologia é composta por sete passos muito claros e diretos que orientam o processo de construção, e esse processo de sete etapas é contínuo e não é realizado apenas uma vez, fazendo um processo iterativo e robusto. A Figura 5 demonstra claramente as sete etapas, sua interação e a possibilidade de iteração.

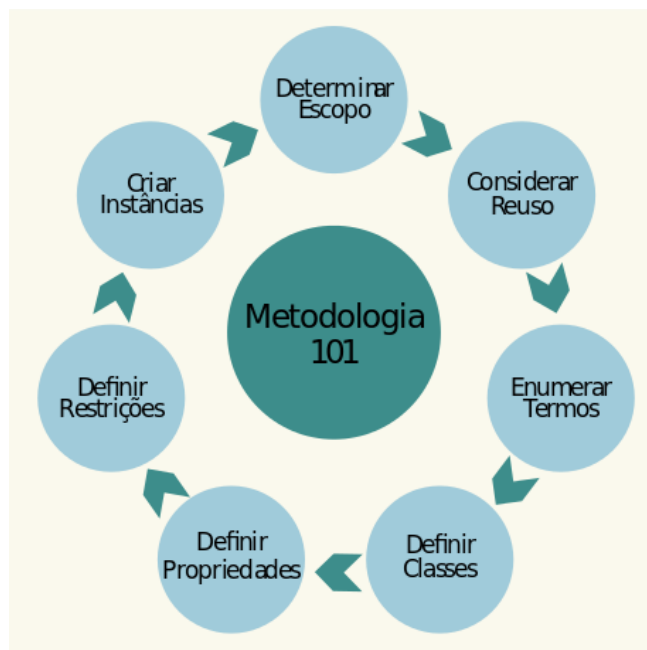


Figura 5 - Passos da Metodologia 101. Fonte: Isotani e Bittencourt, 2015.

Partindo do domínio muito bem definido, focado em aderir a capacidade de projeção situacional, o próximo passo foi a expansão da ontologia através do software Protégé,

permitindo o desenvolvimento através de uma interface computacional e no final exportar o modelo alterado, com suas classes, relacionamento, instâncias e axiomas, em formato OWL.

Estabelecidos o domínio e a forma como a expansão devem acontecer, seguindo a Metodologia 101, o próximo passo é considerar o reuso de ontologias e vocabulários já existentes que contemplem as necessidades do trabalho. Para isso foram feitas buscas na plataforma *Linked Open Vocabularies* – LOV (<https://lov.linkeddata.es/dataset/lov/>), a qual categoriza diversas ontologias, vocabulários e instâncias presentes na web permitindo buscar, baixar e utilizar estes elementos.

Para as buscas foram focadas em encontrar vocabulários sobre projeção, predição e tomada de decisão, termos relacionados ao desenvolver do projeto e objetivo de expansão da ontologia, para isso foi aplicado o filtro de busca “*vocabulary*”. Os termos buscados foram: “*decision*”, “*decision-making*”, “*projection*” e “*prediction*”. Porém ao analisar a descrição, classes e propriedades de cada um dos resultados obtidos, nota-se que não viável a adesão de nenhum dos vocabulários na URBANITY, uma vez que não agregam contribuições relevantes. Sendo que a busca pelo termo “*projection*” não apresentou resultado e não aparece listado no quadro. Conforme detalhado no Quadro 2.

Com a possibilidade de reuso de ontologias e vocabulários já existentes descartadas, tem-se como próxima passo levantar e enumerar termos relevantes, neste caso o foco principal da expansão do modelo ontológico é dar suporte a projeção de informações. Para isso foi considerado o domínio definido e junto a um especialista foram listados termos que podem representar o avanço das informações para um futuro próximo (projeção), assim como a própria situação.

Notou-se que há termos comuns independente do domínio, principalmente aqueles relacionados a situações genéricas com tempo verbal no futuro, como poderia, poderá, vai, acontecerá, porém quando se analisa termos para indicar ideia de tempo futuro para o domínio específica, percebe-se que os termos são os mesmo já levantamos para a construção da ontologia inicial (sem projeção), porém conjugados em seu tempo futuro.

Porém a criação desses termos como propriedades promovendo os relacionamentos das classes e instâncias, pode gerar dualidade, duplicidade e ambiguidade em diversos casos, afetando a capacidade e confiabilidade do modelo de representação. Com isso percebeu-se que junto aos termos genéricos que dão noção de futuro, a criação de uma classe no modelo que represente a presença da projeção é mais válida, possibilitando o uso em situações genéricas dentro do domínio.

Quadro 2: Resultados obtidas nas buscas por vocabulários relacionados a projeção na plataforma *Linked Open Vocabularies*.

Resultados <i>Linked Open Vocabularies</i>		
Termo de Busca	Vocabulário	Descrição e Motivo
<i>Decision</i>	<i>Decision ontology</i>	O foco da ontologia é a descrição e representação de decisões e tomada de decisões, não focando na representação do cenário, seus elementos e como a situação era ou possa evoluir.
	<i>Decision Provenance ontology</i>	O foco da ontologia é a descrição e representação de decisões e tomada de decisões, não focando na representação do cenário, seus elementos e como a situação era ou possa evoluir.
<i>Decision-making</i>	<i>Decision ontology</i>	O foco da ontologia é a descrição e representação de decisões e tomada de decisões, não focando na representação do cenário, seus elementos e como a situação era ou possa evoluir.
	<i>Decision Provenance ontology</i>	O foco da ontologia é a descrição e representação de decisões e tomada de decisões, não focando na representação do cenário, seus elementos e como a situação era ou possa evoluir.
	<i>EduProgression Ontology</i>	Esta ontologia é focada no domínio educacional apenas, apresentando a capacidade de representar progressões educacionais no sistema educacional francês.
	<i>Schema.org vocabular</i>	A Schema.org é uma ontologia focada em promover a representação e formalização de dados estruturados na internet. Não agregando elementos informacionais sobre tomada de decisão e/ou projeções.
	<i>Agreements ontology</i>	A ontologia é construída para modelar e representar acordos judiciais ou contratos sociais, como, leis, contratos e licenças, buscando criar e explicitar relações entre acordos.
<i>Prediction</i>	<i>Energy Efficiency Prediction Semantic Assistant</i>	Ontologia para o domínio energético, desenvolvida para a predição de eficiência energética, porém mesmo lidando com a predição está fica implícita e não traz elementos ontológicos que possam contribuir com o trabalho.

Fonte: o próprio autor.

Dessa forma, o principal avanço dentro da ontologia URBANITY para gerar suporte a representação de projeções situacionais é a classe “*Projection*” e as propriedades “*has projection*” e “*is projection*”, o Quadro 3 a seguir descreve ambas as propriedades e a Figura 6 traz como ocorre o relacionamento entre as classes já existentes na ontologia com a *Projection* e também já demonstra um exemplo de duas instâncias da classe Vítima, onde uma representa a projeção.

Quadro 3: Propriedades que representam a projeção de classes e instâncias.

Propriedades de projeção	
Propriedade	Descrição
<i>has projection</i>	Propriedade que relaciona todas as classes de elementos informacionais que caracterizam uma situação de crime com a classe “ <i>Projection</i> ”, assim criando uma relação direta da classe com a possibilidade de projeção
<i>is projection</i>	Propriedade que relaciona diretamente uma classe com sua projeção, ou seja, vai relacionar a instância de uma classe no momento atual com a sua possível projeção.

Fonte: o próprio autor.

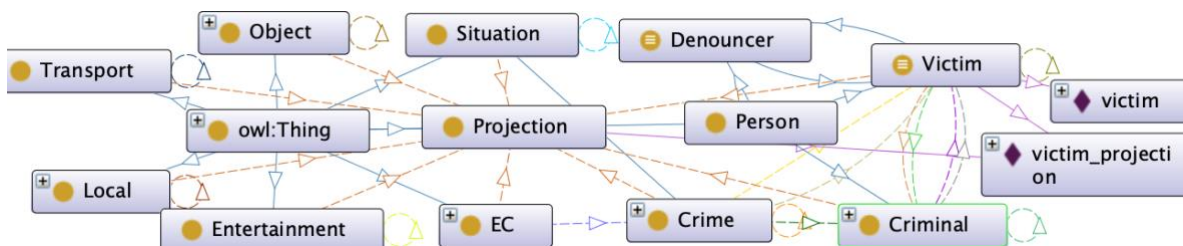


Figura 6 – Adesão da classe *Projection* no Modelo de Representação. Fonte: o próprio autor.

Seguindo com os componentes do modelo Tree, a próxima Seção dedica-se a apresentar o modelo Quantify de fusão de informações, o qual serviu de base para criação da camada de Fusão Semântica de Informações do modelo proposto neste trabalho. Sendo que o modelo Quantify, vai de encontro aos objetivos e capacidades do modelo Tree ao considerar a avaliação e representação de índices de qualidade, representação do conhecimento e dados provindos de inteligência humano.

5.3. Modelo Quantify

Os principais diferenciais deste modelo são a combinação de abordagens sintáticas e semânticas para avaliar os dados e o uso da gestão da qualidade da informação ao longo do processo de fusão. O Quantify consiste em seis processos internos: aquisição de dados, avaliação da qualidade de dados e informações, avaliação de objetos, avaliação de situação, representação de informações e interfaces de usuário, conforme apresentado na Figura 7 (BOTEGA *et al.*, 2019; BOTEGA *et al.*, 2017).

Este modelo tem como objetivo orientar o desenvolvimento de sistemas de fusão de informações, dedicados a apoiar a avaliação de situações que ocorrem em cenários complexos em tempo real, especialmente quando é difícil adquirir informações confiáveis. Esses cenários complexos compreendem entidades altamente complexas que interagem e se relacionam entre si para formar situações, que evoluem no tempo e no espaço (STEINBERG; CHRISTOPHER; FRANKLIN, 2008; BLASCH; PLANO, 2005).

Segundo Botega *et al.* (2019) as principais características do modelo estão os mecanismos projetados para:

- Gerenciar dados e qualidade da informação (inferir, representar e mitigar) em contextos locais e globais do processo fusão de informações em níveis baixos e altos de abstração;
- Apoiar os operadores na melhoria da sua percepção e compreensão da situação e na orientação e refinação das informações;
- Parametrizar processos automatizados da rotina fusão de informações usando informações situacionais qualificadas em processos de fusão sintática e semântica.

Este modelo compreende também:

- Um processo completo de avaliação da situação em cenários complexos em tempo real, com processos internos de aquisição, processamento, representação e refinamento de situações que têm dados fornecidos por humanos produzidos por fontes heterogêneas;
- Uma operação cíclica, iterativa e interativa, que permite aos operadores acompanhar as mudanças na situação;
- Um conjunto de mecanismos para gerenciar dados e qualidade da informação para avaliar cada nova informação inferida, para enriquecer a representação situacional, parametrizar processos e orientar o operador na tarefa de refinamento.

- Uma abordagem de fusão semântica que utiliza modelos semânticos (ontologias) com qualidade de dados associada para melhorar a constatação de informações sinérgicas em dados fornecidos por humanos.

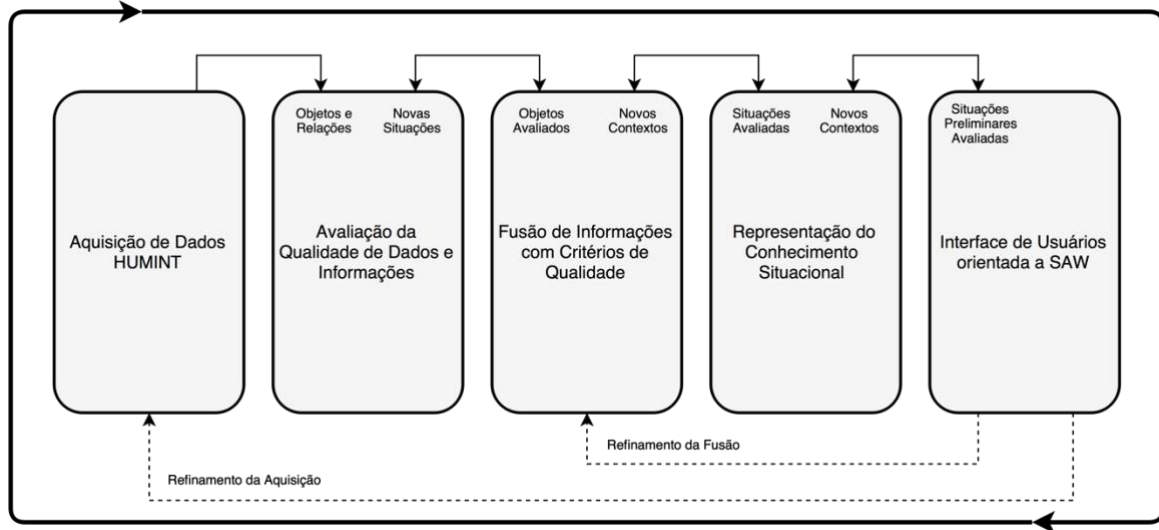


Figura 7 - Modelo Quantify. Fonte: BOTEGA *et al.*, 2017.

Indo de encontro ao modelo Quantify de fusão, que prevê uma camada de avaliação de qualidade e o modelo proposto neste trabalho que também aplica a avaliação contínua de qualidade. A próxima Seção do trabalho se dedica a explicar a metodologia IQESA de avaliação de qualidade e como está irá atuar dentro do modelo em questão e a forma como tal metodologia lida com as dimensões e métricas de qualidade.

5.4. Avaliação de Qualidade utilizando Metodologia IQESA

O objetivo da metodologia IQESA é buscar a melhor maneira de avaliar e representar os índices de dados e qualidade da informação, a fim de possibilitar o uso da qualidade em um processo de avaliação da informação. Isso sempre visa a construção e manutenção da SAW em situações de contexto de risco e emergência, como o domínio criminoso, abordado neste trabalho. A metodologia IQESA, que especifica rotinas de avaliação e representação da qualidade de dados e informações, no contexto de consciência situacional de emergências, apresenta as seguintes capacidades:

- Rastrear as mudanças de qualidade através do sistema de fusão de informações.
- Acompanhar a evolução dos modelos de fusão de informações.

- Apresentar aos usuários uma estimativa atualizada e confiável da qualidade da informação em processos de inferência.
- Especificar rotinas de avaliação em sinergia com os processos de fusão e transformação de informações, os quais afetam a qualidade.

Parte-se do pressuposto que mudanças na informação têm um impacto direto na avaliação da qualidade. A medida que a informação é processada por um sistema de fusão, esta pode evoluir e se torna mais rica, demandando novas avaliações. O processo IQESA consiste em três etapas: (1) o levantamento dos requisitos de qualidade dos dados e informações, (2) a definição das funções e métricas para quantificar as dimensões de qualidade e (3) a representação da informação situacional (BOTEGA *et al.* 2019).

No entanto, a metodologia centra-se no aumento da percepção e no apoio à compreensão situacional, ações que remetem, respectivamente, aos níveis 1 e 2 da SAW, não considerando durante o processo o nível 3, a projeção situacional. Com isso e considerando o modelo de Endsley (1988), onde para o ser humano chegar à projeção, é necessário que ele tenha obtido o nível 2 de SAW. A abordagem do presente trabalho em relação à adoção do IQESA é integrar o mesmo dentro do Quantify (Botega *et al.* 2019), propondo uma mudança e adequação do modelo para que seja possível realizar a fusão semântica com os índices de qualidade visando atingir a projeção situacional (Botega *et al.* 2016).

De acordo com Botega *et al.* (2019), a IQESA já está preparada para lidar com um fluxo contínuo de informações a serem avaliadas e enviadas a outros processos informativos ou humanos. Ao considerar a avaliação da qualidade como parte do ciclo de vida de um processo de avaliação da situação, as informações precisam passar por outros processos, como a aquisição, ou coleta de informações, antes de serem avaliados. Pode então passar através dos dados e da fusão da informação. Após a avaliação, os dados podem ser enviados para outros processos como a representação de informações, neste caso, a ontologia URBANITY ou a interface do usuário.

5.5. Avanços para um processo de Fusão Semântica de Informações

Esta seção elucidada e consolida os avanços para um processo de fusão semântica de informação. A Figura 8 representa os avanços da representação original do Quantify na Figura 7 e fornece detalhes sobre o processo interno.

A representação do processo é mostrada como o ciclo de vida das informações a partir da aquisição de dados, através dos processos de avaliação de objetos, avaliação da situação, avaliação da qualidade, representação de informações e, finalmente, a interface do usuário. Lembrando que a interface do usuário não é o foco deste trabalho, só está presente para demarcar o fluxo de informações.

Uma das contribuições deste trabalho é o alinhamento da recuperação da informação e da avaliação semântica das informações com a metodologia IQESA e a ontologia URBANITY, utilizando OWL, RDF, RDF-S e SPARQL. Habilitando o algoritmo de fusão de dados e informações, já mostrado pelo Modelo Quantify para suportar o processamento semântico junto com índices de qualidade. Nesta abordagem totalmente semântica, os processos internos de fusão, “AVALIAÇÃO SEMÂNTICA DE OBJETOS” e “AVALIAÇÃO SEMÂNTICA DA SITUAÇÃO” sofrer alterações na forma como eles lidam com os dados. A fim de obter, processar, inferir e representar informações semânticas, isso se torna ainda mais dependente de uma ontologia, neste caso a ontologia URBANITY. Várias consultas em ontologia via SPARQL foram necessárias para lidar com toda a troca de informações com o uso de RDF e RDF-S.

Ao analisar o processo, é importante notar que o ciclo de vida das informações, neste caso, começa com "AQUISIÇÃO DE DADOS", o quadrado mais à esquerda, este processo coleta e faz o pré-processamento dos dados. Isso representa a entrada de todos os dados necessários, denominados "dados e informações fornecidos pela inteligência humana". Neste ponto, os dados recolhidos ainda são heterogêneos, não seguindo qualquer modelo, padrão ou formatação. Mesmo que os dados sejam sobre relatos de um crime ou de um arquivo de áudio transcrito, não é possível analisá-los apropriadamente. Após a aquisição de dados, "limpeza de dados" é realizada. Isso removerá todos os acentos e pontuação desnecessária para que uma análise semântica do texto possa ocorrer. Todo o texto é convertido em minúsculas.

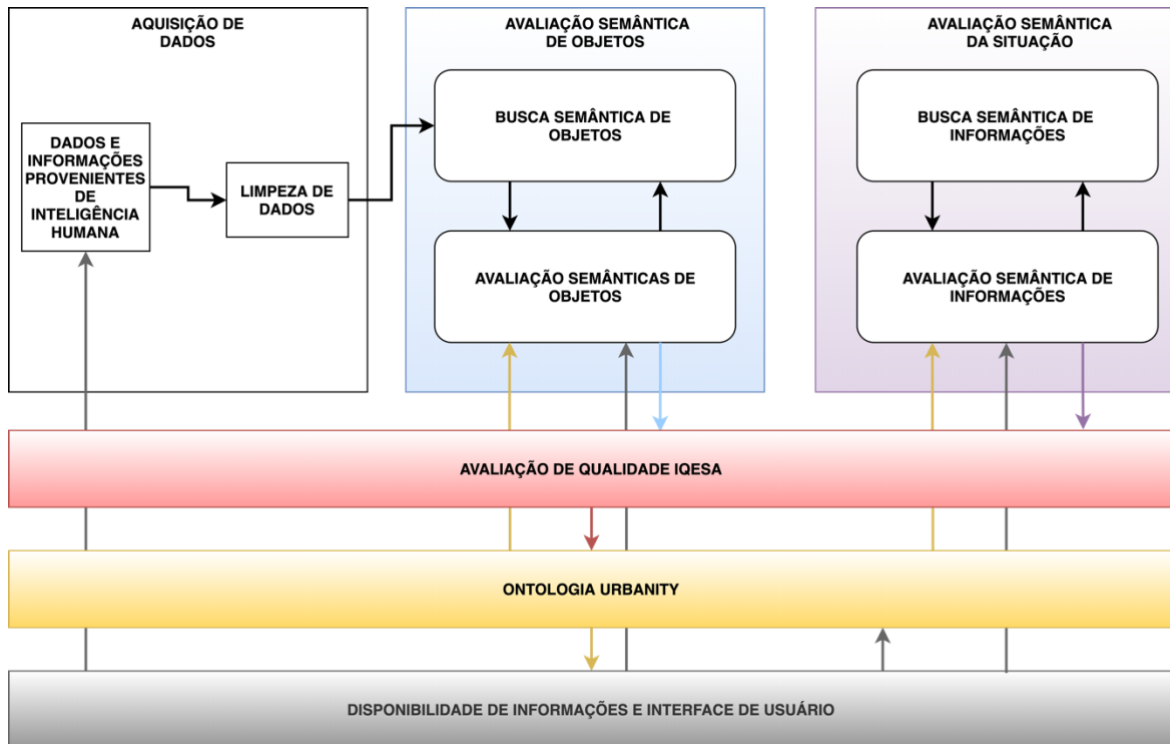


Figura 8 – Processo de fusão semântica de informações baseado no Modelo Quantify.

Fonte: o próprio autor.

Após a limpeza e padronização, o texto passa por "AVALIAÇÃO SEMÂNTICA DE OBJETOS", que identificará e classificará os objetos e os elementos encontrados, nos relatos de crimes. O primeiro processo interno desta etapa, "busca semântica de objetos" segmenta os textos em blocos, cada peça se refere a um termo, momento em que o relatório é transformado um objeto de informação, mais especificamente um objeto RDF. O segundo processo interno, "avaliação semântica de objetos" analisa cada um dos objetos criados na etapa anterior e os classifica, dando-lhes uma possível classe e contexto. Neste ponto, o processo consulta a ontologia URBANITY, através de uma pesquisa SPARQL. Alinha as informações com o modelo ontológico, agregando classes, propriedades, instâncias e axiomas da ontologia, parâmetros que serão utilizados para a fusão e inferência.

O próximo passo é a avaliação da qualidade, submetendo o resultado da avaliação à "AVALIAÇÃO DA QUALIDADE IQESA", que avaliará as informações de acordo com as dimensões e métricas definidas para cada objeto. A representação dos índices de qualidade já deve seguir as normas esperadas pela ontologia, possibilitando que a qualidade seja representada, recuperada e avaliada apropriadamente em processos futuros. Com os índices de qualidade gerados, o objeto informativo é encaminhado para a ontologia URBANITY,

responsável por representar e armazenar as informações já analisadas com os dados de qualidade.

Quando a informação é adequadamente representada na ontologia, o processo de informação pode ser seguido por três procedimentos distintos: (1) apresentar as informações ao usuário, quadro de inferência cinzenta, (2) retornando à "AVALIAÇÃO SEMÂNTICA DE OBJETOS", onde a entrada seria revista, se, por exemplo, não for de qualidade satisfatória, conforme definido pelo usuário ou (3) indo para o processo de "AVALIAÇÃO SEMÂNTICA DA SITUAÇÃO", onde o objeto RDF-S passa a ser o processo de "avaliação semântica de informações", mas agora buscando criar uma imagem mais completa tendo em conta outras situações e objetos de informação previamente analisados. O processo de "busca semântica de informações", que é responsável pela busca semântica dentro da ontologia, utiliza como parâmetro o RDF-S já criado. Isso torna possível fazer pesquisas mais precisas, que levam em consideração o contexto e as propriedades das informações atuais.

Se houver informações de situações semelhantes, uma nova situação será criada. Um novo objeto de informação é criado, contendo informações referentes a situações anteriores e com o objetivo de ser mais completo e de maior qualidade. Este ciclo de busca e comparação é a fusão de informações. Novas informações estão associadas a objetos dentro da ontologia, gerando novos objetos informativos com mais informações, o que pode melhorar os índices de qualidade e alimentar de volta para a ontologia. A nova situação ou objeto de informação, criado, é submetido novamente à "AVALIAÇÃO DA QUALIDADE IQESA" e "ONTOLOGIA URBANITY" completando o ciclo de informação.

Ao analisar o ciclo de informação, a avaliação da qualidade e a ontologia foram utilizadas em vários momentos, caracterizando um processo de avaliação dinâmica. Isso possibilita o uso de elementos semânticos e índices de qualidade, também representados na ontologia, para fazer pesquisa, avaliação, comparação e Associação de informações. Colocando a qualidade e a ontologia no cerne do processo de busca e avaliação de informações é possível criar e usar algoritmos de tamanho semântico para lidar com informações.

5.6. Demonstração do Modelo Tree

A prova de conceito é baseada na avaliação que suporta a consciência situacional de um crime real, mais especificamente um crime da extorsão, relatado ao serviço da resposta da emergência das polícias. As informações relatadas foram submetidas a cada etapa do

modelo e seus algoritmos. Os resultados são analisados com base na aquisição, mineração, fusão e representação de informações relevantes úteis para a tomada de decisões. Complementando-os, existem outros processos cruciais para a implementação do IF, como identificação de objetos, representação de informações e avaliação de qualidade.

A Avaliação da Situação começa com a identificação de entidades e objetos presentes nos relatos, que se baseia em técnicas de PNL, utilizando um vocabulário rico da linguagem, focado no domínio de gerenciamento de emergências, mais especificamente a análise do tempo real dados criminais. Após a identificação de entidades e objetos, as informações, comunicadas através de objetos JSON, começam a ser instantâneas em pequenas ontologias (Figuras. 10, 11 e 12) que mais tarde irão compor uma situação. O módulo de representação de informações é usado aqui para apoiar os processos de fusão de informações durante a execução, mantendo a versão mais atual das informações situacionais.

Cada resultado também é submetido ao módulo de avaliação de qualidade, que neste momento só avalia a completude das entidades e objetos identificados. Os dados retornarão a este módulo cada vez que novas inferências ou qualquer tipo de transformação de informação é feita.

Assim, o fluxo de execução segue o ciclo: identificação de objetos, representação, avaliação de qualidade e de volta ao módulo de representação. Os seguintes relatos foram submetidos ao Modelo, e os resultados são apresentados e discutidos abaixo, destacando objetos de interesse.

Relato Criminal 1: A vítima parou o veículo em um semáforo. Então, ela foi surpreendida por dois indivíduos em uma motocicleta preta. O passageiro da moto bateu no vidro do carro com uma arma. Ameaçou a vítima e subtraiu seu veículo. Os criminosos escaparam. A vítima não tem condições para descrever os autores em detalhes.

Relato Criminal 2: Dois caras em uma moto escura roubaram um carro no semáforo no cruzamento entre as ruas Mooca e Taquari, apontou um revólver, e pegou o carro da mulher. Um deles estava de casaco vermelho.

Relato Criminal 3: Dois homens roubaram um carro cinza na rua da Mooca para o lado do Banco Santander. Os homens estavam em uma motocicleta preta, um deles estava em uma jaqueta azul e jeans. Eles partiram na direção do Hospital Villa-Lobos.

No início do processo, as informações não são analisadas diretamente, ou seja, pela forma como os termos são escritos, cada relato é estruturado em instâncias da ontologia URBANITY. Cada um dos relatos descritos acima será uma instância diferente, dependendo da situação que representa.

Esse processo de instanciação semântica começa a partir da identificação de elementos armazenados na camada de representação de informações. Em seguida, para cada relatório, as consultas SPARQL são feitas na ontologia de domínio para cada elemento previamente identificado. Essas consultas buscam identificar qual classe ou conjunto de classes da ontologia melhor representa os elementos e as prováveis propriedades que os caracterizam, mesmo que não sejam explícitas no relatório. O resultado deste processo é um exemplo de uma situação, que não representa necessariamente um crime.

Neste ponto, temos vários casos pertencentes aos mesmos eventos sem relações entre eles. O próximo passo é realizar consultas SPARQL, agora considerando um contexto local, ou seja, dentro de cada classe, para identificar propriedades comuns de instâncias de elementos, para tornar possível inferir novos objetos, propriedades e atributos, a fim de definir uma situação para cada um Relatório. Estes exemplos são mostrados em Figuras 9, 10, e 11.

Nessas figuras, os retângulos representam as instâncias, e os círculos representam os atributos. As setas e links representam as conexões e suas propriedades entre classes e atributos, definidas de acordo com um vocabulário que também foi desenvolvido para o domínio e mostrado nas seguintes figuras.

As instâncias são atribuídas de acordo com as classes que representam para facilitar a compreensão da situação. Rosa representa objetos relacionados à vítima; vermelho representa objetos relacionados ao criminoso; verde estão relacionados com o site; roxo é a vítima; azuis são criminosos; e branco representa o exemplo da situação.

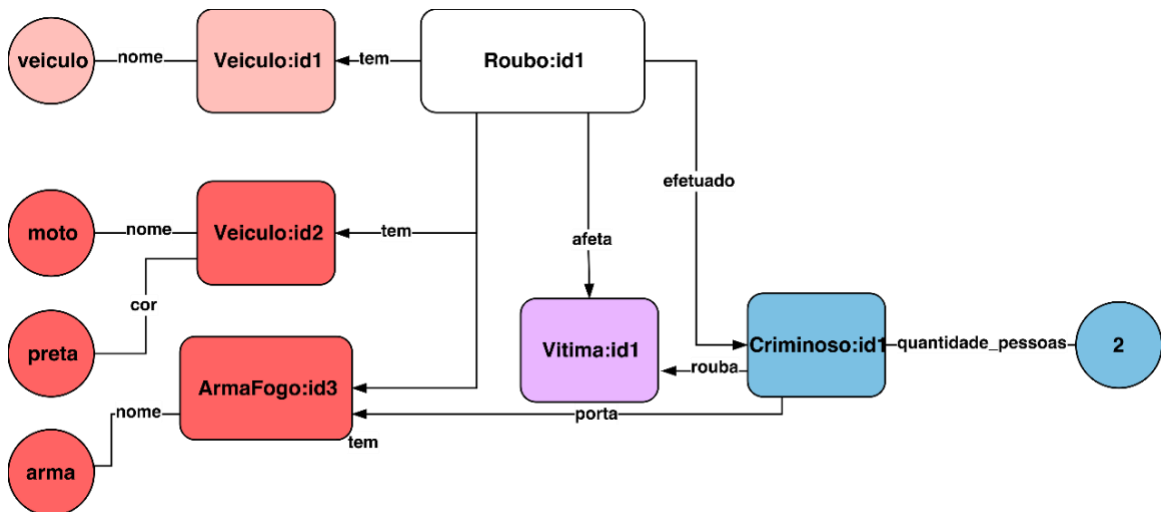


Figura 9 - Situação instanciada com as informações do relato criminal 1. Fonte: o próprio autor.

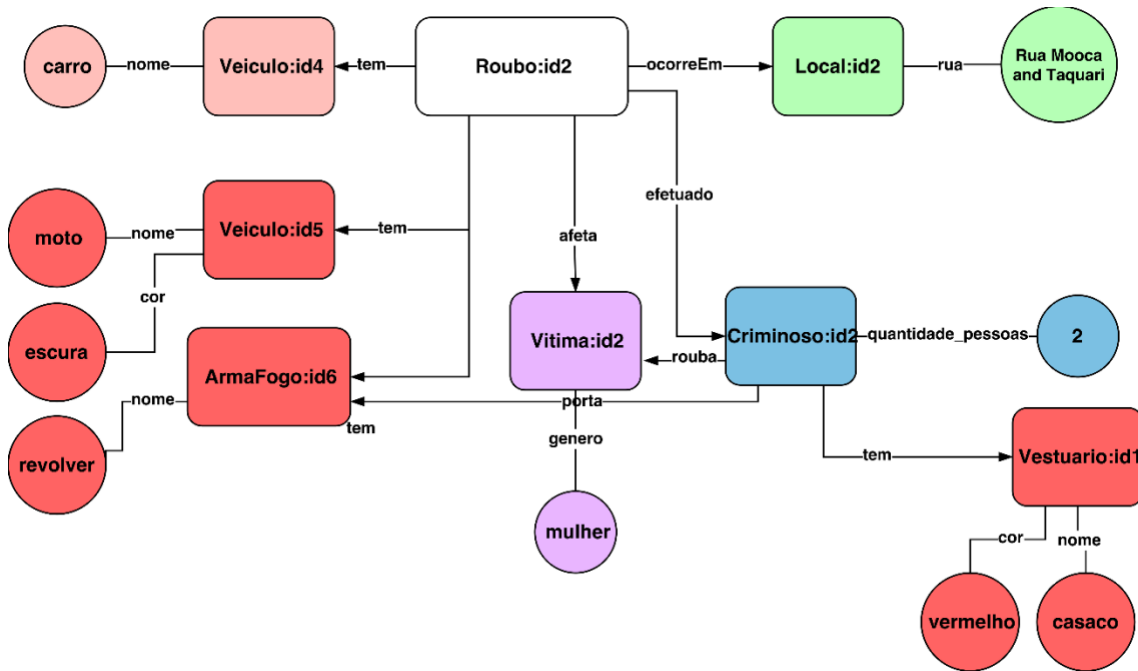


Figura 10 – Situação instanciada com as informações do relato criminal 2. Fonte: o próprio autor.

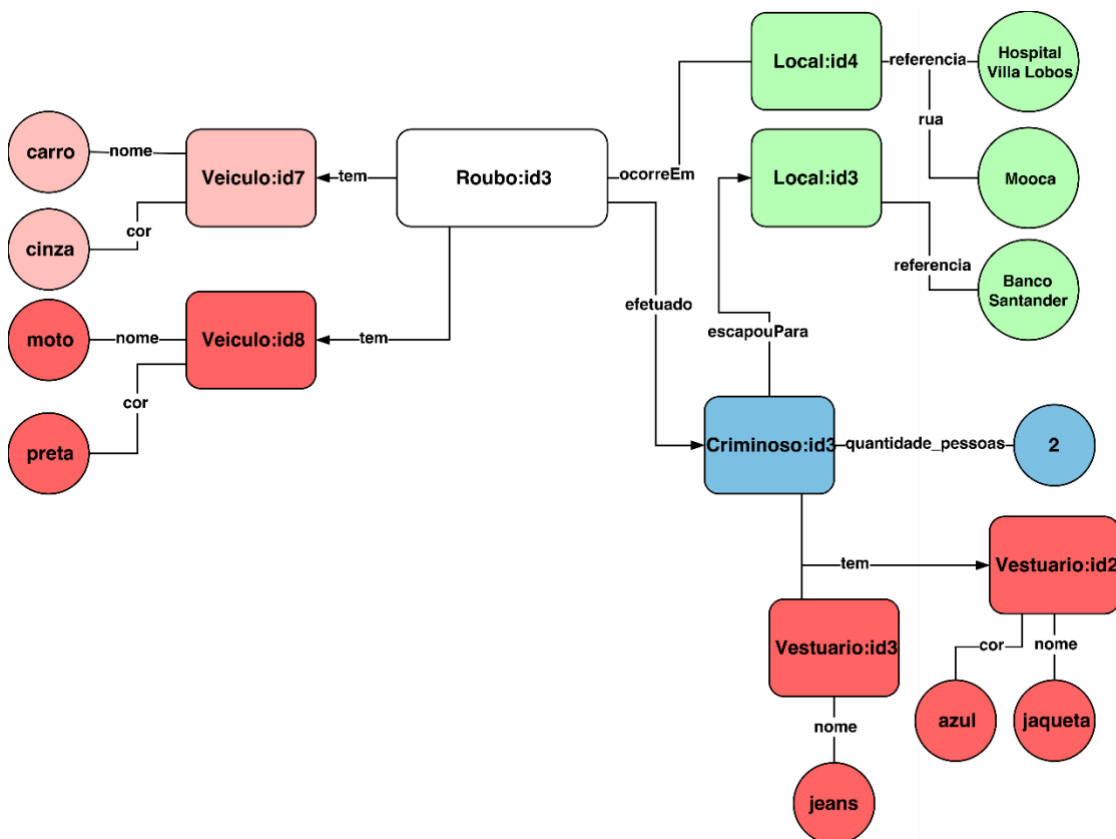


Figura 11 – Situação instanciada com as informações do relato criminal 3. Fonte: o próprio autor.

O processamento começa com a análise dos elementos encontrados no Relato 1, sendo vítima, veículo, dois indivíduos, motocicleta preta, passageiro de motocicleta, com uma arma, ameaçada, subtraída e seu veículo.

Com base nesses termos, as consultas são feitas para a ontologia, que retornará em qual classe eles se encaixam, ou se são apenas propriedades. Neste caso, um conjunto de classes é devolvido com vítima, criminoso e objeto e algumas propriedades, como ameaçadas e subtraídas, o que pode representar um roubo.

Com consultas SPARQL para uma rica ontologia, também é possível fazer associações entre termos distintos. Por exemplo, o Relatório 1 não tem o termo “criminoso”, e ainda uma classe criminosa foi identificada, porque na ontologia, um dos termos instantâneos que caracteriza um criminoso e está presente no Relato 1 é “indivíduos”.

Em outro exemplo, percebemos uma propriedade “porta” de um criminoso por uma arma de fogo, indicando que os criminosos têm uma arma, mesmo que não haja termo “empunhava” no relatório. Isso ocorre porque o termo “com” está no relatório, e através da ontologia, é possível estabelecer uma covariância entre eles. Ao final da análise semântica, as entidades estão correlacionadas de acordo com propriedades, dependências ou relações encontradas nos relatos e previstas no vocabulário. O resultado (Figura 9) para o Relato 1 também persiste na camada de representação de informações.

Essa rotina de instanciação da ontologia com objetos inferidos, atributos e propriedades referentes a uma única situação ocorre para todos os relatos submetidos ao processo semântico. Como é possível observar em Figuras 9, 10 e 11, os três relatos parecem se referir à mesma situação, e fragmentos dele foram descritos por pessoas diferente presentes na situação.

Essa relação entre as situações encontradas nos relatos é facilmente compreendida por um ser humano, que deduz as situações e as interpreta em um conhecimento situacional (o que se sabe sobre uma situação). No entanto, em um cenário mais complexo, os seres humanos tornam-se propensos a erros e não podem absorver todas as características de uma situação.

Neste ponto, a fusão semântica combina novos relatos com instâncias de RDFs (informações previamente processadas armazenadas na camada de representação de informações). O objetivo é construir um modelo computacional muito próximo de um modelo mental humano, através de ontologias e vocabulários. Esse processo utiliza as

informações na camada de representação de informações, após o processo de identificação semântica.

A fusão semântica é muito semelhante à identificação semântica, mas em um nível mais alto de importância, uma vez que suas inferências são feitas usando instâncias de todos os relatos. Além disso, a fusão semântica considera todas as propriedades apresentadas em cada instância interna dos elementos. Ao final da fusão semântica, teremos um novo conjunto de situações possíveis que variam na organização e presença dos elementos e propriedades. Cada uma dessas situações possíveis é salva na camada de representação e tem sua qualidade avaliada.

Em seguida, a inferência das informações utilizando de critérios qualitativos é realizada, principalmente avaliando as melhorias nos índices de qualidade. A situação que apresenta os índices de melhor qualidade e que satisfaça os critérios eliciados nos requisitos, como a presença de algum elemento específico, será eleita como a situação final resultante da fusão. Mais uma vez, esta situação é persistiu pela ontologia na camada de representação.

A fusão semântica dos relatos analisados permite obter uma informação com maior valor agregado pela junção dos termos encontrados. Neste caso com os três relatos, a informação fundida permite a identificação das roupas de ambos os criminosos, as características do objeto roubado, a arma utilizada pelos criminosos e veículo de fuga, bem como o local onde ocorreu o crime. Os exemplos mostram que os relatos de criminalidade por si só não apresentam essas situações explicitamente.

O resultado da fusão semântica é mostrado na Figura 13, que usa as mesmas cores de antes, com a diferença de que a cor amarela representa as informações que foram fundidas a partir dos três relatos. Além da presença de mais instâncias e atributos na situação fundida, também é possível notar novas propriedades, que não foram explícitas ao considerar cada relatório separadamente.

Outro ponto a ser observado é que a qualidade da informação pode ser avaliada na ontologia para permitir a verificação de quais informações têm melhor qualidade e, assim, decidir o que deve ser usado na fusão ou não. Para esta avaliação, usamos a ontologia de Vocabulário de Qualidade de Dados (DQV), este permite a criação de instâncias de categorias, dimensões e métricas para mapear medições de qualidade.

Para cada instância da ontologia, é criada uma instância de gráfico de qualidade, fazendo a conexão entre todas as medidas de qualidade aplicadas a essas informações instantâneas. A Figura 12 mostra o resultado da fusão semântica com avaliação de qualidade, mostrando as dimensões de consistência e corrente para a situação de roubo avaliadas.

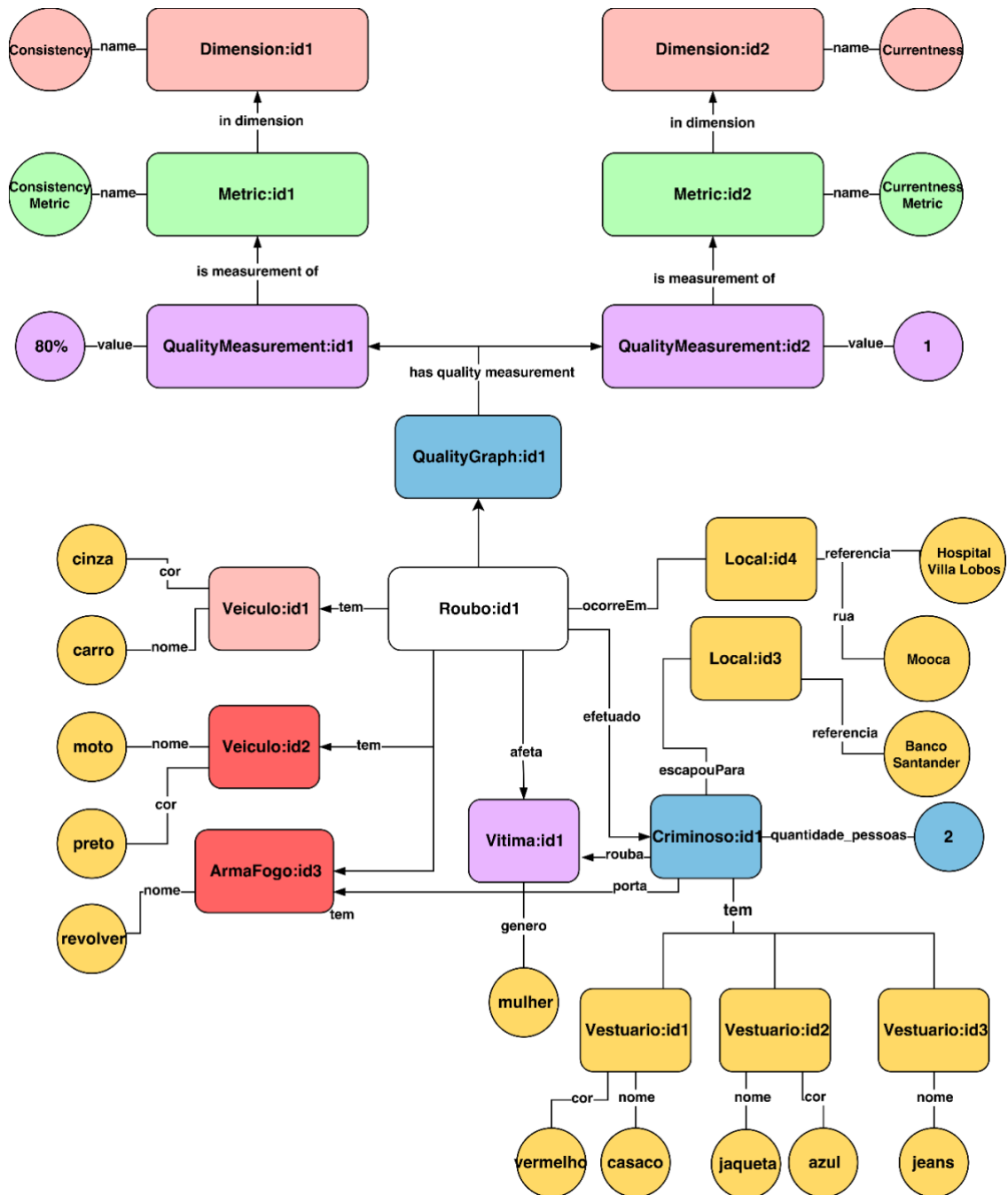


Figura 12 – Ontologia da situação instantânea após a transformação e representação, com as informações de todos os relatos associados a um índice de qualidade. Fonte: o próprio autor.

Com o processo de projeção de informações sendo antecipado pela ontologia, torna-se viável a execução de diversos outros processos informacionais que poderão colaborar para a projeção situacional do humano, tais como inferência de novas informações, avaliação de qualidade, transformação e apresentação das informações para o usuário por meio de interface computacional. Visando elucidar a proposta de expansão da ontologia para uma fusão de informações com projeção, este trabalho estende, em forma de prova de conceito, o

estudo de caso apresentado por Botega *et al.* (2019), o qual apresenta três denúncias sobre um roubo e traz o resultado alcançado através de um processo de fusão de informações representado em uma instância, sendo esta adaptada e apresentada na Figura 13.

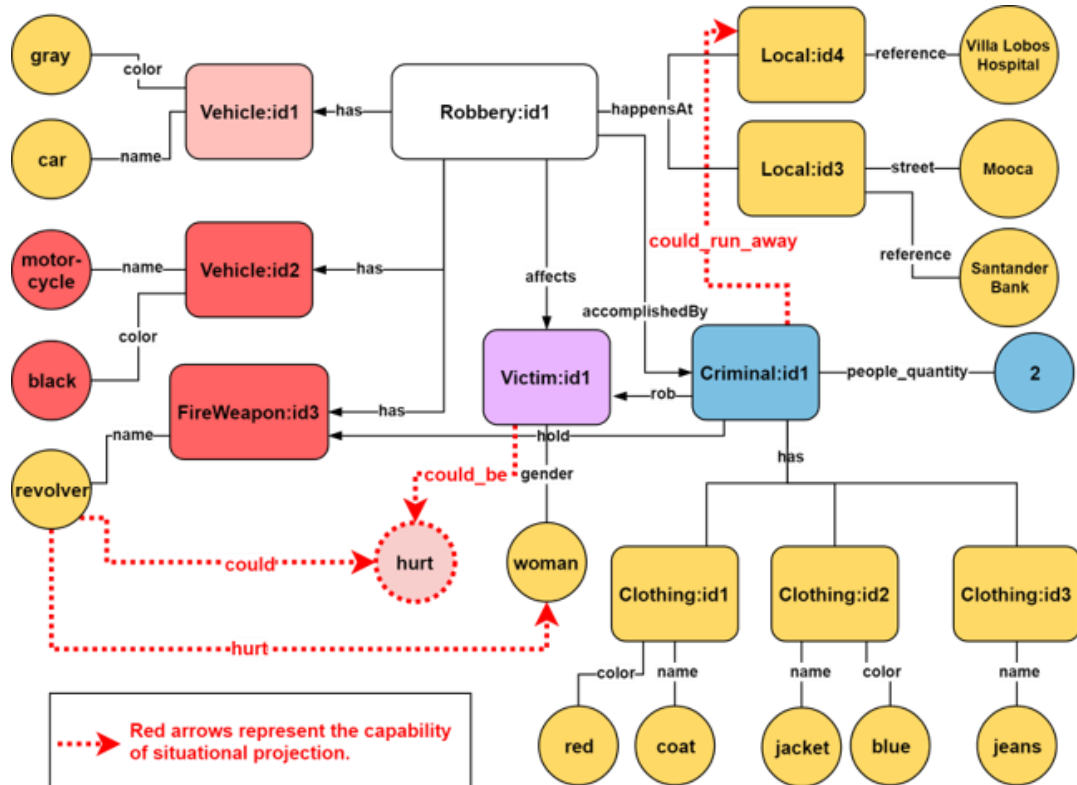


Figura 13 – Representação de um resultado de fusão com projeção informacional. Fonte: o próprio autor.

Na Figura 13, pode-se notar alguns elementos que compõem a situação, neste caso já definida e representada pelo retângulo branco, que apresenta a classificação e um identificador, Roubo (*Robbery*) com “id1”. A partir deste, são associados os elementos que constroem uma situação (retângulos coloridos), compostas por instâncias de classes da URBANITY, sendo dois veículos (*Vehicle*), uma vítima (*Victim*), um criminoso (*Criminal*) e dois locais (*Local*). Os círculos representam os valores literais de cada instância. As propriedades são representadas por setas, relacionando instâncias e valores, onde cada uma tem o nome da propriedade representado. As setas pretas são referentes à representação original e as vermelhas são propriedades para projeção situacional.

Com a elaboração deste ensaio, apenas acrescentando propriedades relacionadas a projeção situacional na URBANITY, é possível perceber um avanço na capacidade de representação das informações, assim como tornar viável as projeções situacionais. Pode-se considerar que ao aderir classes, propriedades, axiomas e instâncias direcionadas a projeções

situacionais em uma ontologia, torna-se factível representar a estimativa de um futuro na ontologia e apresentar essas informações para o humano, auxiliando assim no seu processo de construção de SAW e tomada de decisão, além da colaboração com o processo humano.

Conseguir representar a estimativa de uma informação de modo formal possibilita o aprofundamento da ciência, sendo possível e necessário novos métodos de coleta, transformação, processamento, inferência, apresentação e até mesmo transmissão de informações sobre o futuro, agregando outra disciplina como a qualidade e até mesmo estatística. Outro fator resultante da inclusão de tais elementos na ontologia é o aumento da confiabilidade e semântica das informações, gerando assim mais parâmetros e insumos para processos que venham manipular, processar e inferir tais informações.

Ao apresentar as dificuldades do domínio informacional em questão, o qual traz informações provenientes de inteligência humana, com falhas de qualidade e refletindo de forma negativa na SAW, foi possível identificar a problemática em lidar com tais informações dentro do meio informacional e computacional, revelando a necessidade de representar e avaliar informações de forma a auxiliar humanos no processo de SAW e tomada de decisão.

Este trabalho traz em uma temática antes não discutida com profundidade, a capacidade de representar a projeção situacional em modelos informacionais, possibilitando que diversos processos informacionais e computacionais se comuniquem para possibilitar o processamento das informações, como o uso de ontologias junto à índices de qualidade da informação e a fusão de informações.

O arranjo destes diversos processos junto ao arcabouço metodológico e tecnológico apresentado no decorrer desta proposta, guiaram a mesma para um modelo capaz de prover a transformação e representação de informação, de forma que gere melhores informações para contribuir SAW e projeções situacionais.

Além de apresentar novas contribuições para o ciclo de vida de dados e informações, principalmente por oferecer insumos informacionais para processos de transformação, inferência, avaliação de qualidade, apresentação das informações, a expansão da ontologia para lidar com projeções situacionais mostrou-se capaz de conseguir informações mais relevantes e completas para o consumo de humanos e sistemas em situações críticas que demandam a SAW e tomada de decisão .

Todavia, avanços no trabalho são necessários, tais como o levantamento e avaliação de novas classes, propriedades e axiomas, relacionadas ao domínio criminal e que auxiliem na representação de projeções informacionais. Outros pontos a serem discutidos são como

tais informações futuras serão apresentadas ao usuário, como serão transmitidas para modelos e processos externos de inferência, assim como buscar uma forma de validação para tais informações.

6. Prova de Conceito do Modelo Tree

Com o intuito de comprovar e validar o modelo proposto, foi desenvolvida uma prova de conceito, a qual seja capaz de executar os processos discutidos durante o trabalho apresentado. Elucidando as principais contribuições e limitações do modelo. A prova de conceito foi estruturada de maneira a comprovar a validade e aplicabilidade do modelo, todavia, não focando na produção de software, plataforma ou aplicativo que faça o uso do modelo, uma vez que não é proposta desta pesquisa.

Resumidamente, a prova de conceito vai lidar com um cenário de crime controlado, considerando o uso de três relatos sobre uma mesma situação de crime, porém com pontos de vista e informações diferentes, possibilitando demonstrar como ocorre a transformação, busca, associação, representação e inferência das informações. Para isso será estruturado um processo informacional baseado no modelo proposto, ou seja, colocar o modelo proposto em prática, assim é possível definir algumas etapas as quais as informações irão passar, indo desde a coleta até a representação das informações, a Figura 14 demonstra o processo informacional baseado no modelo.



Figura 14 – Etapas da prova de conceito para validar o modelo. Fonte: o próprio autor.

A prova de conceito foi dividida em cinco etapas, seguindo a estrutura do modelo, porém módulos que permeiam diversos processos, como a avaliação de qualidade que é um processo que está presente em mais de um módulo de modelo, se torna implícito neste processado, sendo destacado no detalhamento de cada uma destas, como por exemplo a etapa de transformação e fusão que faz uso direto da avaliação de qualidade.

Importante destacar que durante a elaboração da prova de conceito serão elencadas técnicas e tecnologias que se adequam ao modelo e são capazes de realizar os processos necessários, junto a isso serão demonstradas como as mesmas irão se comportar.

Primeira etapa é referente a coleta dos dados, uma vez que o modelo se propõem a lidar com dados e informações provenientes de humanos, principalmente dados em texto, tem-se como principal fonte de coleta dessas informações a internet, principalmente sites de

notícias, redes sociais como Twitter, Facebook, Instagram, entre outros. Todavia, tais dados presentes na internet não seguem um padrão de construção, formatação, representação e persistência, contendo gírias e/ou erros de português, junto a regionalismos.

Para as ferramentas de mineração e extração da internet (coleta) de dados isto não é o maior um problema, pois tal ferramental é preparado e seu foco é garantir a extração e não a formatação ou padronização. Porém isto geram problemas para os processos que irão lidar com tais informações coletadas, não apenas neste modelo de transformação e representação, mas em diversos outros como em bancos de dados que adotam uma estrutura relacional ou técnicas de análise como o *Machine Learning*, que buscam por padrões para estruturar e analisar informações.

Com as informações coletadas, a etapa seguinte é a transformação da informação, que tem a responsabilidade de agregar valor a informação coletada através da limpeza, padronização e transformar, principalmente dados não estruturados em dados estruturados, seguindo um modelo informacional, neste caso utilizando do modelo ontológico. Tal transformação permite que o dado coletado seja analisado por um número maior de processos, seja persistido, representado e transmitido dentro do modelo sem perder valor, representatividade e evitando erros e falhas que podem afetar a qualidade da informação final.

Já nesta etapa de transformação os processos podem contar apoio da avaliação de qualidade, que pode agregar índices de qualidade, assim como o suporte do modelo ontológico ou de representação, isto depende da tecnologia aplicada, como por exemplo a ferramenta Open Refine (<https://openrefine.org>) que consegue cruzar os dados com uma ontologia. O resultado desta etapa deve ser um conjunto de instâncias da ontologia utilizada, que irão caracterizar a situação a qual os dados se referem no mundo real. Importante ressaltar que o processo de instanciação pode ocorrer de algumas maneiras, conforme o ferramental aplicada, sendo principalmente de forma manual, inserindo as instâncias através da interface gráfica do Protégé ou gerando os arquivos RDF-S ou OWL que podem ser importados para a ontologia também através do Protégé.

Com os dados representados e disponíveis dentro da ontologia é possível realizar a Fusão Semântica das Informações que representa um processo de busca e associação de informações correlatas que juntas podem agregar mais valor as informações e torna-las mais representativas ao humano e supri a demanda informacional para compreensão do domínio. Como resultado da associação de informações correlatas dentro da ontologia, podem surgir novas instâncias e relações entra as informações, com isso altera-se a estrutura da informação

que reflete nos índices de qualidade da informação, sendo assim necessária uma nova avaliação de qualidade, que revela a melhora, ou não, na qualidade das informações.

Como resultado desta etapa do módulo pode ser um conjunto de novas instâncias ou revelar novas relações entre instâncias, ou seja, uma nova situação, é necessário a representação das situações, revelando quais informações de quais situações foram utilizadas para criar esse novo cenário, tem-se assim a etapa de Representação. Todo o fluxo informacional até essa etapa do modelo, focou na melhora da informação, afim tanto de gerar a melhor informação para o humano, assim como gerar insumos suficientes para realizar as inferências necessárias para alcançar a projeção informacional, ou seja, como a situação com as informações atuais poderá estar em um futuro próximo.

Ao processar e inferir informações em busca de projeções espera-se lidar diretamente com a variável de tempo, passado, presente e futuro, porém essa informação ou ao menos noções temporais podem ou estar presentes em dados provenientes de humanos, de maneira explícita ou implícita. Neste caso, devido ao domínio e ao cenário, por lidar com informações de bases histórias ou da internet, de base histórias (que não são em tempo real) e descritas por humanos, tem-se a percepção de tempo implícita.

Relacionando diretamente com a projeção, torna-se inviável definir o ponto exato no futuro o qual a informação está se aproximando. Com essa análise a etapa de Projeções foca na evolução das informações que podem auxiliar o humano na compreensão de cenários, trazendo noções implícitas de tempo. Com essa análise e considerações feitas, as próximas seções do trabalho se dedicam a mostrar passo a passo o fluxo informacional do modelo, através de uma prova de conceito em ambiente controlado, visando validar o modelo proposto.

6.1. Coleta de Dados

A primeira etapa da prova de conceito vai de encontro a camada de “Coleta de Dados Provindos de Inteligência Humana” em destaque na Figura 15. Quando se trata da coleta de dados provenientes por humanos há uma gama enorme de possíveis fontes, desde sites, redes sociais e bases oficiais, onde cada uma segue uma arquitetura, padrão, formatação, estrutura e até mesmo problemas quanto a qualidade e confiança dos dados. Por esse motivo a camada de coleta de dados é uma cama adaptável e dinâmica, podendo adaptar-se as necessidades ou requisitos da fonte que será utilizada.

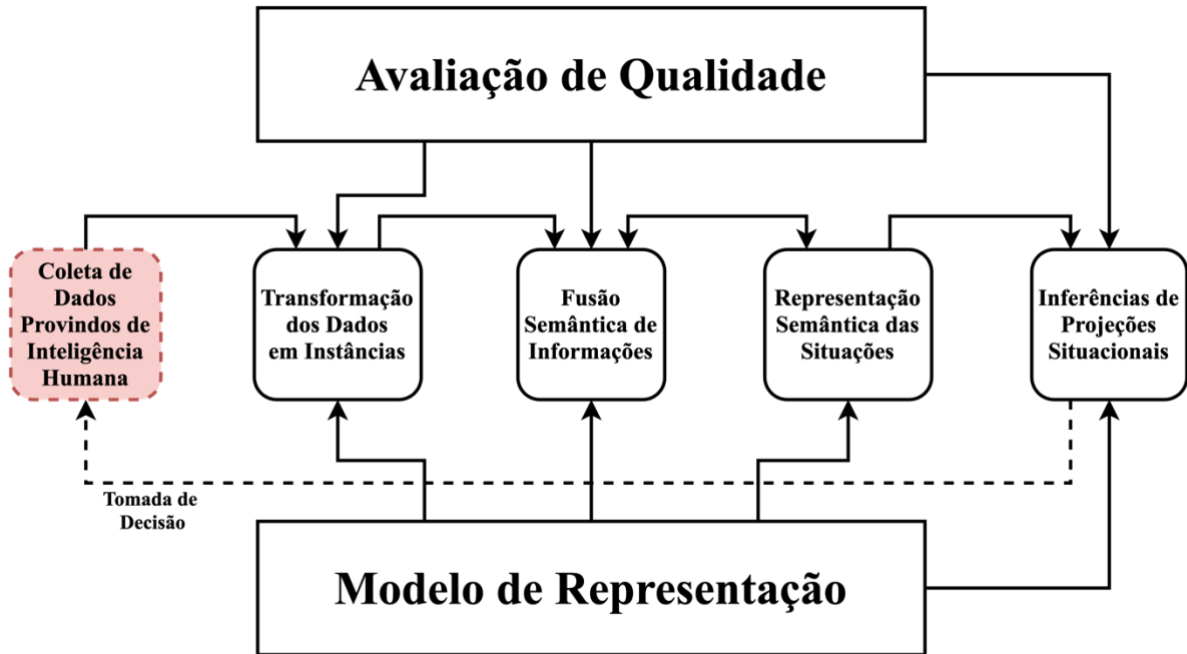


Figura 15 – Destaque para a camada Coleta de Dados Provindos de Inteligência Humana.

Fonte: o próprio autor.

Para suprir as demandas desta camada podem ser aplicadas técnicas computacionais de mineração e *scraping* de dados, independente de qual ferramenta seja aplicada, sendo possível fazer esse processo de forma semiautomatizada através de algoritmos ou automatizada aplicando algumas ferramentas. É válido ressaltar que o essencial é prevalecer a característica de extração de dados da web em formatos possíveis de manipulação.

Visando validar formas tecnológicas de se coletar a informação de diferentes fontes foi utilizada a ferramenta *open source* Scrapy (<https://scrapy.org>), que permite a mineração e *scraping* de dados, assim como uma breve manipulação, possibilitando exportar apenas os dados necessários, atentando-se que o uso dessa ferramenta está diretamente ao uso da linguagem de programação Python. Com o uso dessa ferramenta foi possível realizar o *scraping* de dados do Portal de Transparência da Secretaria de Segurança Pública do Estado de São Paulo (SSP-SP), focando nos dados disponíveis até março de 2020. A Figura 16 demonstra como os dados são apresentados no site.

Porém, com o foco em manter o ambiente controlado proposto para a prova de conceito, alguns relatos foram elaborados baseados nos dados reais coletados e analisados, os relatos elaborados, seguem os padrões e problemas encontrados no portal citado acima. Já considerando as próximas camadas do modelo proposto, foram desenvolvidos três relatos que se referem a uma mesma situação de roubo, trazendo informações correlatas e passíveis de

busca, associação, representação e inferência. A situação considerada foi o roubo do veículo de uma mulher por dois criminosos portando arma de fogo e que fugiram com uma moto.

A Figura 17, demonstra como o resultado da camada de coleta, um arquivo de texto em formato CSV (*comma-separated values*), a escolha desse formato de arquivo se deve a sua facilidade de manipulação das informações e integração com diversas outras ferramentas que podem contribuir para o processo.

Cidade	Delegacia Elaboração	Data Fato	Data Registro	Endereço Fato
S.PAULO	DELEGACIA ELETRONICA	29/02/2020	01/03/2020 00:02:26	RUA BARCELOS LEITE, 220
S.BERNARDO DO CAMPO	02° D.P. S.BERNARDO DO CAMPO	29/02/2020	01/03/2020 00:05:23	RUA MARIA BORDIN, 150
S.PAULO	DELEGACIA ELETRONICA	29/02/2020	01/03/2020 00:05:59	RUA ADELE ZARZUR, 643
CAMPINAS	PLANTÃO - 01° D.P. JUNDIAI	29/02/2020	01/03/2020 00:15:40	RUA HILÉIA, 60
CAMPINAS	DEL.SEC.2ª CAMPINAS-PLANTÃO	29/02/2020	01/03/2020 00:20:48	AVENIDA JOHN BOID DUNLOP CAM-465, 82
LORENA	DEL.POL.LORENA	29/02/2020	01/03/2020 00:31:22	RODOVIA BR 459, 20
S.PAULO	DELEGACIA ELETRONICA	29/02/2020	01/03/2020 00:32:15	RUA CONTOS GAUCHESCOS, 174
SOROCABA	PLANTÃO DE SOROCABA - Z. NORTE	29/02/2020	01/03/2020 00:55:14	AVENIDA elias maluf, 715
CAMPINAS	DEL.SEC.2ª CAMPINAS-PLANTÃO	01/03/2020	01/03/2020 01:20:52	RUA ROMEU CICONE, 300
S.PAULO	14° D.P. PINHEIROS	01/03/2020	01/03/2020 01:29:56	RUA PURPURINA, 1468
S.PAULO	53° D.P. PARQUE DO CARMO	29/02/2020	01/03/2020 01:46:36	RUA OTELO AUGUSTO RIBEIRO, 700
TABOAO DA SERRA	01° D.P. TABOÃO DA SERRA	01/03/2020	01/03/2020 02:12:32	RUA IRACEMA SENA CERQUEIRA SANTOS, 983
GUARULHOS	01° D.P. GUARULHOS	29/02/2020	01/03/2020 02:20:16	RUA DA FORTUNA, 470
S.PAULO	26° D.P. SACOMA	29/02/2020	01/03/2020 02:21:11	RUA TITO OLIANI, 91
S.PAULO	01° D.P. TABOÃO DA SERRA	01/03/2020	01/03/2020 02:26:06	AVENIDA PROFESSOR FRANCISCO MORATO, 5825
S.PAULO	16° D.P. VILA CLEMENTINO	29/02/2020	01/03/2020 02:41:02	RUA BARRETO MUNIZ, 250
GUARULHOS	07° D.P. GUARULHOS	23/02/2020	01/03/2020 02:42:23	RUA IJACI, 1

Figura 16 – Amostra de dados disponíveis no Portal de Transparência da SSP-SP. Fonte: o próprio autor.

Cidade,Data Fato,Endereço,Relato

São Paulo,17/03/2020,,	"A vítima parou o veículo em um semáforo. Então, ela foi surpreendida por dois indivíduos em uma motocicleta preta. O passageiro da moto bateu no vidro do carro com uma arma. Ameaçou a vítima e subtraiu seu veículo. Os criminosos escaparam. A vítima não tem condições para descrever os autores em detalhes."
São Paulo,17/03/2020,Rua Mooca com Taquari,	"Dois caras em uma moto escura roubaram um carro no semáforo no cruzamento entre as ruas Mooca e Taquari, apontou um revólver, e pegou o carro da mulher. Um deles estava de casaco vermelho."
São Paulo,17/03/2020,Rua Mooca,"	Dois homens roubaram um carro cinza na rua da Mooca para o lado do Banco Santander. Os homens estavam em uma motocicleta preta, um deles estava em uma jaqueta azul e jeans. Eles partiram na direção do Hospital Villa-Lobos."

Figura 17 – Saída da camada de coleta com dados. Fonte: o próprio autor.

6.2. Transformação

Após os dados coletados pela primeira camada do modelo, o processo continua com a camada de Transformação dos dados, que traz a capacidade tratar, limpar e padronizar os

dados coletados, podendo assim transformá-los no padrão aceitando pela ontologia e por consequência gerar as instâncias necessárias, se tornando uma camada de conexão e necessária para viabilizar a fusão de dados semântica, como destacado na Figura 18. Essa camada do modelo pode ser segmentada em quatro principais processos, os quais podem ocorrer com ferramentas e tecnologias distintas, que seriam a limpeza, padronização, transformação em instâncias e avaliação de qualidade, como ilustrado na Figura 19.

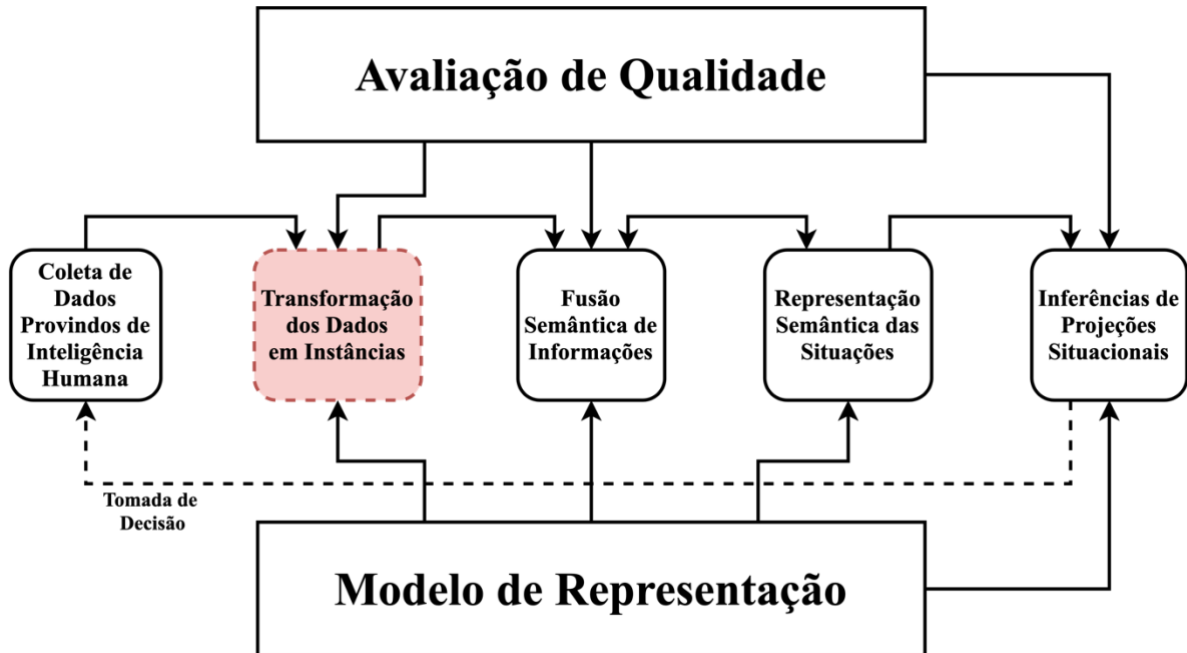


Figura 18 – Destaque para a camada Transformação dos Dados em Instâncias. Fonte: o próprio autor.

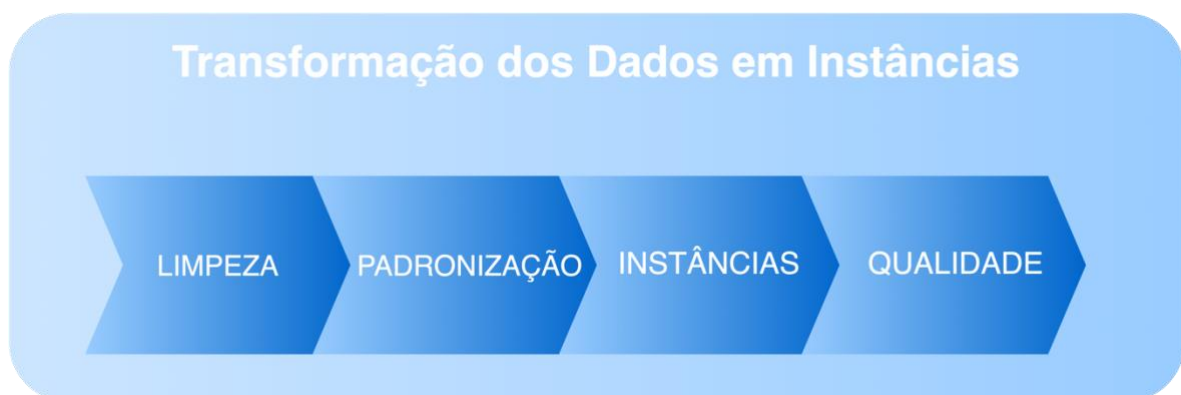


Figura 19 – Processos internos à camada de Transformação de Dados. Fonte: o próprio autor.

Quando trata-se de ferramentas tecnológicas capazes de tratar dados dessa forma existe com um conjunto enorme de possibilidades, principalmente com o volume de dados

disponíveis na web e serviços de análise de dados. Um conceito que tem se apresentado comum dentre esse conjunto de ferramentas é o ETL (*Extract, Transform and Load*), que disponibiliza um conjunto de funções que auxiliam a limpeza e padronização dos dados.

Tendo isso para realizar este primeiro processo de transformação dos dados, este trabalho aplica a ferramenta *open source* Pandas (<https://pandas.pydata.org>), que traz os conceitos de ETL e flexibiliza os procedimentos informacionais, importando e exportando dados em diversos formatos inclusive CSV (padrão adotado na primeira camada do modelo). Importante ressaltar que assim como a ferramenta apresentada para coleta, o Pandas também está diretamente ligado a linguagem de programação Python.

A seguir é demonstrado como é realizado os dois primeiros processos desta camada (Limpeza e Padronização) utilizando a ferramenta Pandas. Neste processo de limpeza e padronização, são retirados todos os caracteres especiais, acentos, realizada a normalização na formação da informação, possibilitando que sua propagação dentro de todas as camadas do modelo ocorra da melhor maneira possível se fragmentação ou perda de qualidade. A Figura 20 demonstra o processo aplica nesses procedimentos.

```
import pandas as pd
import csv
from unicodedata import normalize

path = 'Saida Camada Coleta - Defesa.csv'
with open(path,newline = '',encoding='utf8') as f:
    reader = csv.reader(f, delimiter=',')
    rg = []

    for row in reader:
        ln = {}
        ln['report'] = normalize('NFKD', row[3])\
            .encode('ASCII', 'ignore')\
            .decode('ASCII')\
            .strip('*,- ')\
            .lower()
        rg.append(ln)
df = pd.DataFrame(rg, columns = ['report'])
df.to_csv('data_clean.csv', sep=',', line_terminator='\n',header=False, index=False)
```

Figura 20 – Processos de limpeza e padronização das informações coletadas. Fonte: o próprio autor.

Como resultado desse procedimento de transformação é gerado um arquivo CSV, possibilitando a continuidade do processo independente do ferramental utilizado. Esse arquivo consta todos os dados coletados, ainda sem nenhuma segmentação clara ou em instâncias. A Figura 21, demonstra a estrutura dos dados e arquivo resultante.

"a vitima parou o veiculo em um semaforo. entao, ela foi surpreendida por dois individuos em uma motocicleta preta. o passageiro da moto bateu no vidro do carro com uma arma. ameaçou a vitima e subtraiu seu veiculo. os criminosos escaparam. a vitima nao tem condicoes para descrever os autores em detalhes"

"dois caras em uma moto escura roubaram um carro no semaforo no cruzamento entre as ruas mooca e taquari, apontou um revolver, e pegou o carro da mulher. um deles estava de casaco vermelho"

"dois homens roubaram um carro cinza na rua da mooca para o lado do banco santander. os homens estavam em uma motocicleta preta, um deles estava em uma jaqueta azul e jeans. eles partiram na direcao do hospital villa-lobos"

Figura 21 – Saída obtida dos procedimentos de limpeza e padronização na camada de Transformação. Fonte: o próprio autor.

Nesse momento do processo de transformação que são selecionados os dados relevantes para o processo informacional, no caso desta prova de conceito apenas o relato feito sobre o crime é relevante, apresentando elementos informacionais que compõe uma situação, inclusive o tempo de maneira implícita, fator importante a fusão das informações e projeção.

O processo subsequente ainda dentro da camada de transformação é a geração das instâncias para a ontologia, esse procedimento permite o desenvolvimento da construção das funções de busca e associação semântica. Este procedimento pode ocorrer de duas maneiras, manual ou automática com uso de ferramentas tecnológicas. Para o processo manual é necessário o apoio de um especialista do domínio adotado, onde são feitas análises entre as informações apresentadas e o modelo ontológico, selecionando quais são os dados relevantes e a partir destes são geradas as instâncias através de uma interface para gerenciar a ontologia, como o Protégé, por exemplo.

Para o processo automatizado de geração de instâncias podem ser utilizadas diversas ferramentas para contribuir com o processo que mantenham a premissa de permitir que o

especialista possa construir a relação entre os elementos informacionais identificados previamente em instâncias conforme as classes da ontologia.

Para este trabalho foi aplicado o procedimento de transformação automática em instâncias, contando com o auxílio da ferramenta *open source* Open Refine (<https://openrefine.org>). Dentro dessa ferramenta foi possível importar os dados a serem trabalhados (*dataset*), assim como o modelo ontológico a ser trabalhado, dando a possibilidade de cruzar o *dataset* com as classes da ontologia. Para viabilizar tal procedimento a ontologia utilizada deve estar em formato OWL ou RDF-S, neste caso foi utilizado o Protégé para exportar a ontologia URBANITY em OWL.

Neste momento do processo de transformação é feita a segmentação dos dados coletados, neste caso os três relatos, onde cada um poderá se transformar em uma instância de situação na ontologia ou em instâncias isoladas. Ao término do procedimento, terá como resultado três instâncias distintas em OWL, representando cada um dos relatos e com suas devidas instâncias sobre os elementos informacionais encontrados, ou seja, se existe informações sobre vítima, criminoso, objetos ou local. As Figuras 22, 23 e 24 representam respectivamente o primeiro, segundo e terceiro relatos demonstrados nos procedimentos anteriores, contendo trechos do OWL gerado pela ferramenta Open Refine.

```

### http://localhost:8090/crimevoc#roubo\_1
:roubo_1 rdf:type owl:NamedIndividual ,
        |
        | http://localhost:8090/roubo#Roubo ;
        |
        | :affects :vitima ;
        | :done :criminoso_1 ;
        | :has :arma_de_fogo_1 ,
        |     :veiculo_1 ,
        |     :veiculo_2 ;
        |
        | rdfs:label "roubo_1" .

### http://localhost:8090/crimevoc#criminoso\_1
:criminoso_1 rdf:type owl:NamedIndividual ,
        |
        | http://localhost:8090/pessoa#Criminal ;
        |
        | :has :arma_de_fogo_1 ;
        | :rob :vitima ;
        | :amount "2"^^xsd:int ;
        |
        | rdfs:label "criminoso_1" .

```

Figura 22 – Trecho do primeiro relato transformado em OWL para ser consumido pela URBANITY. Fonte: o próprio autor.

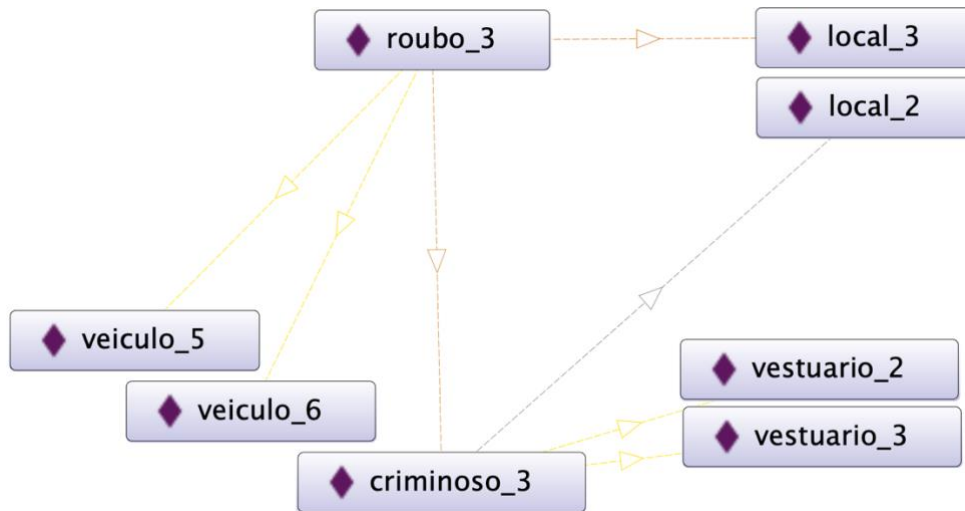


Figura 27 – Instância do terceiro relato e seus elementos já importados para a URBANITY.

Fonte: o próprio autor.

Com os relatos instanciados na URBANITY, é possível fazer o primeiro procedimento de avaliação das informações, neste caso para avaliar o nível de qualidade da informação. A metodologia utilizada para a avaliação de qualidade é a IQESA, junto ao vocabulário de qualidade DQV para representar as informações de qualidade na ontologia. Para realizar a prova de conceito e analisar aplicabilidade da avaliação de qualidade será apenas considerada a dimensão de completude da informação.

Para realizar o cálculo de completude da informação, será considerado o número de elementos presentes no relato em relação ao total esperado em uma situação. Neste caso, as situações são compostas por quatro elementos: Criminoso, Vítima, Local e Objetos, assim cada classe de elementos presente confere 25% de completude. Com isso foi elaborada uma função SPARQL que retorna os elementos informacionais presentes na instância do relato a ser analisado, possibilitando a análise de quão completa está a informação. A Figura 28 demonstra a função SPARQL para listagem dos elementos e a Figura 29 representa o resultado ao aplicar a função sobre o primeiro relato.

```
SELECT distinct ?situation ?criminal ?victim ?local ?objects
WHERE { ?situation rdf:type <http://localhost:8090/robbery#Robbery> .
  { ?situation voc:done ?criminal }
  UNION { ?situation voc:affects ?victim }
  UNION { ?situation voc:ocorreEm ?local}
  UNION {?situation voc:has ?objects} .
  ?situation rdfs:label 'relato_1' }
```

Figura 28 – Função SPARQL para listagem dos elementos informacionais por instância.

Fonte: o próprio autor.

situation	criminal	victim	local	objects
relato_1	criminoso_1			
relato_1		vitima_1		
relato_1				veiculo_2
relato_1				arma_de_fogo_1
relato_1				veiculo_1

Figura 29 – Resultado da Função SPARQL para listar elementos informacionais. Fonte: o próprio autor.

Ao analisar o resultado obtido pela função SPARQL com o primeiro relato, percebe-se que apenas três dos quatro elementos esperando estão sendo listados, o criminoso, a vítima e os objetos envolvidos, não tendo a ocorrência de um local. Considerando que neste cenário cada elemento equivale a 25% de completude, o relato tem um índice de qualidade relacionado à completude de 75%. Com essas informações é possível representar a qualidade na URBANITY junto a instância do primeiro relato. A Figura 30 representa o relacionamento entre a situação e a qualidade.

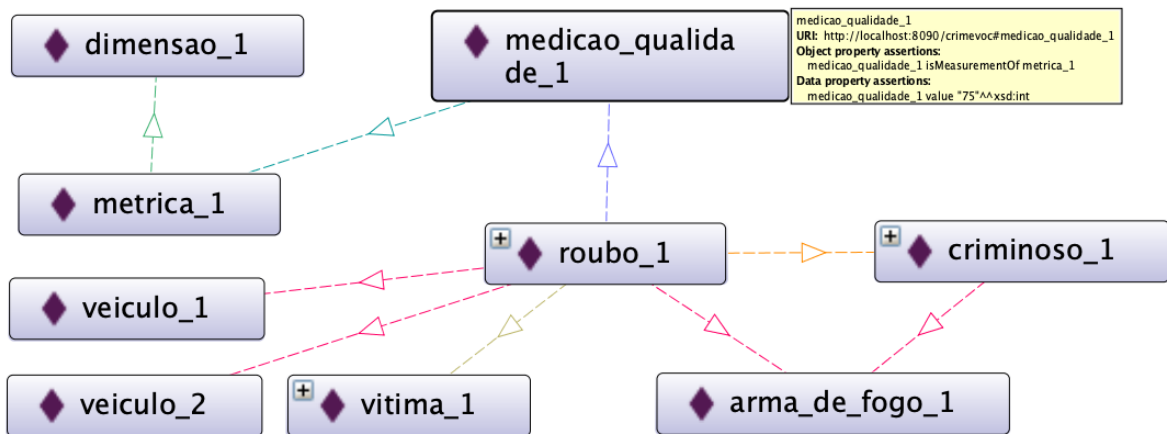


Figura 30 – Representação da instância da avaliação de qualidade com DQV e o relacionamento a situação. Fonte: o próprio autor.

Este procedimento de avaliação dos índices de qualidade e sua representação na ontologia finaliza todo o conjunto de processos relacionados à camada de Transformação de Dados em Instâncias, dando o suporte e gerando todos os insumos necessário para a próxima camada do fluxo informacional deste modelo, a Fusão Semântica de Informações que irá aplicar conjuntos de procedimento e funções de busca e associação de informações dentro da ontologia, como demonstrado na próxima seção.

6.3. Fusão

Após os dados tratados e transformados pela segunda camada do modelo, o processo continua com a camada de Fusão Semântica de Informações, destacada na Figura 31, que traz a capacidade realizar buscas por informações que apresentem algum grau de correlação ou covariância, ou seja, se aproximam através de algum de sus atributos, como por exemplo, instâncias dos elementos informacionais. Essa camada do modelo pode ser segmentada em quatro principais processos, os quais podem ocorrer com ferramentas e tecnologias distintas, a busca, associação, instâncias e avaliação de qualidade como demonstrando na Figura 32.

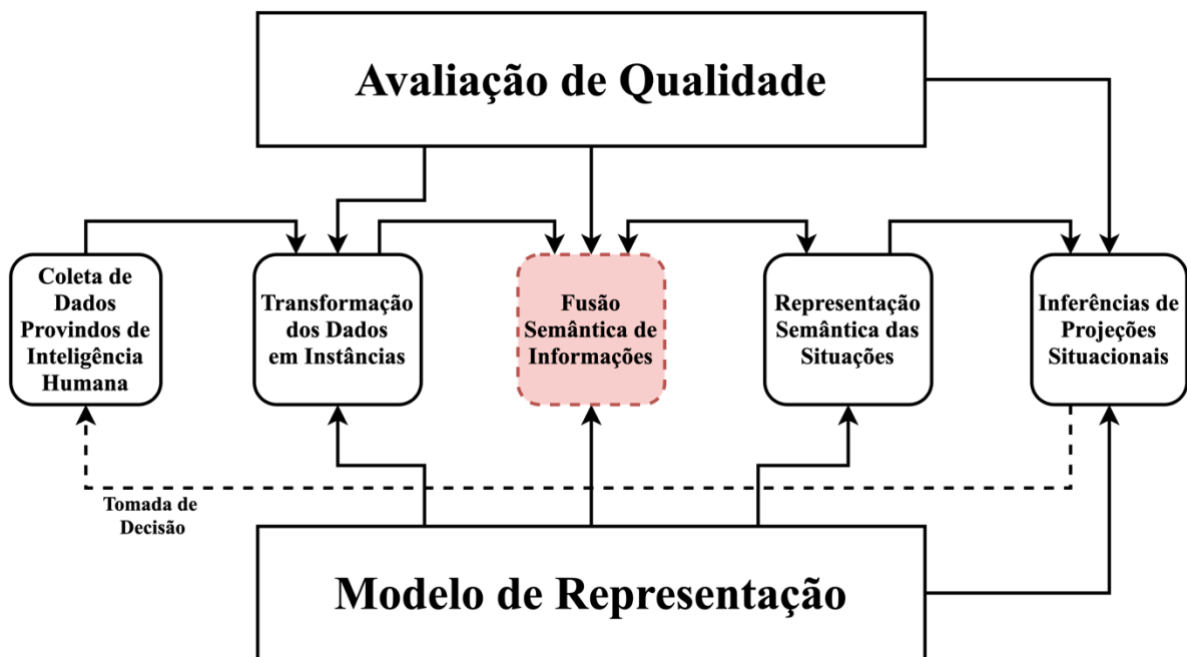


Figura 31 – Destaque para a camada Fusão Semântica de Informações. Fonte: o próprio autor.

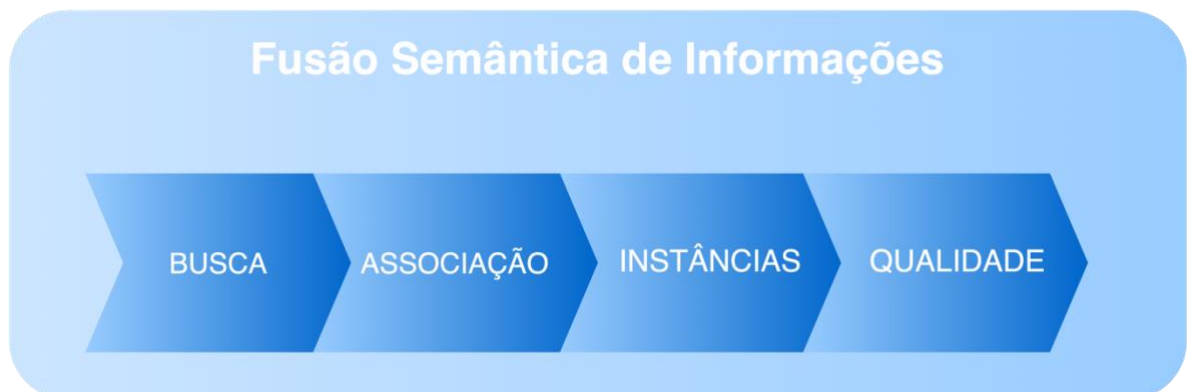


Figura 32 – Processos internos à camada de Fusão de Informações. Fonte: o próprio autor.

Para tornar possível os procedimentos internos da camada de Fusão de Informações, são aplicadas tecnologias como o SPARQL, principalmente para criar e executar a busca e associação das informações semânticas, ou seja, as instâncias que foram geradas na camada anterior, Transformação de Dados em Instâncias. Com o uso é possível manipular e criar condições entre as informações disponíveis, a fusão das informações deve conter pontos correlatos, ou seja, quais dos elementos mais informacionais se aproximam em alguns aspectos. Por exemplo, pode-se analisar os atributos de local, objetos encontrados, vítimas e criminosos.

Dando continuidade ao fluxo informacional da prova de conceito, considerando as instâncias dos três relatos citados nas Seções anteriores, o primeiro passo para o procedimento de buscar de informações é considerar a estrutura do modelo ontológico e as informações instanciadas, uma vez que as funções de busca e associação estão diretamente relacionadas ao domínio aplicado. Neste caso, o domínio criminal, para se criar a função de busca de informações foram consideradas as classes de Situação, diferenciando entre roubo e furto, Criminoso, Vítima, Local e Objeto. Sendo que cada uma dessas classes tem suas propriedades de objetos e dados, assim como atributos, que devem ser considerados em cada momento.

Assim, a função de busca semântica construída é capaz de listar situações de roubo e furto presentes na ontologia, junto a seus elementos informações, as Figuras 33 e 34, representam respectivamente a função SPARQL de busca semântica de informações e o resultado obtido.

```
SELECT distinct ?situation ?criminal ?victim ?local ?objects
WHERE {
  OPTIONAL {?situation voc:done ?criminal } .
  OPTIONAL {?situation voc:affects ?victim } .
  OPTIONAL {?situation voc:ocorreEm ?lc . ?lc voc:street ?local } .
  OPTIONAL {?situation voc:has ?objs . ?objs voc:name ?objects } .
}
```

Figura 33 – Função de busca semântica de informações para fusão. Fonte: o próprio autor.

Na Figura 33 é importante ressaltar que na primeira, dentro da cláusula SELECT, são selecionados os elementos informacionais a serem buscados e listados, neste referente às classes da ontologia, enquanto a cláusula WHERE monta as regras de associação das informações, nesta função apenas montando a estrutura base de uma situação de crime. Sendo que os elementos listados na cláusula SELECT compõem o resultado apresentado.

situation	criminal	victim	local	objects
roubo_1	criminoso_1	vitima_1		"arma"^^.
roubo_1	criminoso_1	vitima_1		"veiculo"^^.
roubo_1	criminoso_1	vitima_1		"moto"^^.
roubo_3	criminoso_3		"Mooca"^^<	"carro"
roubo_3	criminoso_3		"Mooca"^^<	"moto"
roubo_2	criminoso_2	vitima_2	"Rua Mooca".	"revolver"∨
roubo_2	criminoso_2	vitima_2	"Rua Mooca".	"carro"^^.
roubo_2	criminoso_2	vitima_2	"Rua Mooca".	"moto"^^.
roubo_2	criminoso_2	vitima_2	"Taquari"^^.	"revolver"∨
roubo_2	criminoso_2	vitima_2	"Taquari"^^.	"carro"^^.
roubo_2	criminoso_2	vitima_2	"Taquari"^^.	"moto"^^.

Figura 34 – Resultado busca semântica de informações para fusão. Fonte: o próprio autor.

Como resultado da função de busca é tida uma listagem de todas as situações possíveis de fusão, analisando o resultado é possível observar a repetição de algumas informações nas linhas, mas isso devido a configuração e apresentação dos resultados, onde cada nova informação distinta em uma coluna replica as demais informações comuns. Também é possível observar lacunas em branco, sem o preenchimento de nenhuma informação, isto indica a inexistência de instância de determinada classe para a instância da situação, por exemplo, ao analisar as linhas referente a instancia “roubo_3”, nota-se que não existe instâncias de vítima para essa situação, logo seus índices de qualidade também são afetados.

A partir da análise do resultado da função de busca é possível analisar informações que são correlatas, desde a presença de um elemento, como o criminoso ou detalhes como local e objetos podem ser considerados como correlatos, então a partir dessa análise é possível construir funções de associação semântica de informações. Assim como a função de busca está diretamente ligada ao domínio aplicada a função de busca também está e pode sofrer alterações em sua estrutura, principalmente nos filtros, que irão criar a associação entre os elementos correlatos.

Conforme os parâmetros correlatos de associação se alteram conforme a situação a função de associação também precisa ser alterada, assim não afetando a confiabilidade das informações apresentadas e a capacidade informacional do modelo utilizado. Tendo isso, a função SPARQL para associação de informações para o cenário desta prova de conceito, considerada os elementos informacionais de local, especialmente uma rua em questão, e a

quantidade de criminosos presentes, tornando possível a associação entre os três relatos analisados, informações em destacada na imagem. As Figuras 35 e 36 representam respectivamente a função SPARQL criada e o resultado obtido.

```
SELECT distinct ?situation ?local ?criminal ?criminal_amount
      ?criminal_objects ?victim ?details_victim ?street ?objs ?objects
WHERE { ?situation voc:done ?criminal .
      ?criminal voc:amount ?criminal_amount .
      OPTIONAL { ?criminal voc:has ?c_obj . ?c_obj voc:name ?criminal_objects}
      OPTIONAL { ?situation voc:affects ?victim . ?victim voc:gender ?details_victim } .
      OPTIONAL {?situation voc:ocorreEm ?local . ?local voc:street ?street} .
      OPTIONAL {?situation voc:has ?objs . ?objs voc:name ?objects } .
      FILTER (?criminal_amount >= 2 || regex(?street, "mooca", "i"))
}
```

Figura 35 – Função de associação semântica de informações correlatas. Fonte: o próprio autor.

situation	local	criminal	criminal_amount	criminal_objects	victim	details_victim	street	objs	objects
roubo_3	local_3	criminoso_3	"2"^^<http://ww	"jeans"^^<http://			"Mooca"^^<l	veiculo_5	"carro"
roubo_3	local_3	criminoso_3	"2"^^<http://ww	"jeans"^^<http://			"Mooca"^^<l	veiculo_6	"moto"
roubo_3	local_3	criminoso_3	"2"^^<http://ww	"jaqueta"^^<htt			"Mooca"^^<l	veiculo_5	"carro"
roubo_3	local_3	criminoso_3	"2"^^<http://ww	"jaqueta"^^<htt			"Mooca"^^<l	veiculo_6	"moto"
roubo_1		criminoso_1	"2"^^<http://ww	"arma"^^<http://				arma_de_fogo_1	"arma"^^<
roubo_1		criminoso_1	"2"^^<http://ww	"arma"^^<http://				veiculo_1	"veiculo"^^
roubo_1		criminoso_1	"2"^^<http://ww	"arma"^^<http://				veiculo_2	"moto"^^<
roubo_2	local_1	criminoso_2	"2"^^<http://ww	"revolver"^^<htt	vitima_2	"mulher"^^<h	"Rua Mooca"	arma_de_fogo_2	"revolver"^^
roubo_2	local_1	criminoso_2	"2"^^<http://ww	"revolver"^^<htt	vitima_2	"mulher"^^<h	"Rua Mooca"	veiculo_3	"carro"^^<
roubo_2	local_1	criminoso_2	"2"^^<http://ww	"revolver"^^<htt	vitima_2	"mulher"^^<h	"Rua Mooca"	veiculo_4	"moto"^^<
roubo_2	local_1	criminoso_2	"2"^^<http://ww	"casaco"^^<http	vitima_2	"mulher"^^<h	"Rua Mooca"	arma_de_fogo_2	"revolver"^^
roubo_2	local_1	criminoso_2	"2"^^<http://ww	"casaco"^^<http	vitima_2	"mulher"^^<h	"Rua Mooca"	veiculo_3	"carro"^^<
roubo_2	local_1	criminoso_2	"2"^^<http://ww	"casaco"^^<http	vitima_2	"mulher"^^<h	"Rua Mooca"	veiculo_4	"moto"^^<

Figura 36 – Resultado da função SPARQL de associação semântica de informações. Fonte: o próprio autor.

Com todas as informações associadas, é possível realizar o procedimento seguinte da camada desta camada, que é criação de uma nova situação através de instâncias da ontologia, estas novas instanciadas criadas a partir dos procedimentos de busca e associação semântica representam o resultado da fusão, ou seja, a partir de três situações presente na ontologia, criou-se uma única situação agregando informações de três relatos de maneira direta. Esse procedimento afeta diretamente a representativa das informações e seus índices de qualidade, que irá refletir procedimentos futuros de transformação, fusão e projeção, uma vez que estas novas instâncias estarão disponíveis para possíveis aplicações futuras do modelo.

O procedimento de instanciar o resultado da fusão pode ser realizado através de OWL, consultando o modelo ontológico ou gerando as instâncias através do Protégé e aplicando as devidas propriedades de objeto e dados. As Figuras 37 e 38 representam

respectivamente o OWL gerado a partir do resultado da fusão e a representação gráfica das instanciadas pela fusão.

```
### http://localhost:8090/crimevoc#roubo_fusao_1
:roubo_fusao_1 rdf:type owl:NamedIndividual ,
                <http://localhost:8090/robbery#Robbery> ;
                :affects :vitima_fusao_1 ;
                :done :criminoso_fusao_1 ;
                :has :veiculo_fusao_1 ,
                    :veiculo_fusao_2 ;
                :is_related_to :roubo_1 ,
                               :roubo_2 ,
                               :roubo_3 ;
                :ocorreEm :local_3 .
```

Figura 37 – Resultado da fusão de informação em OWL. Fonte: o próprio autor.

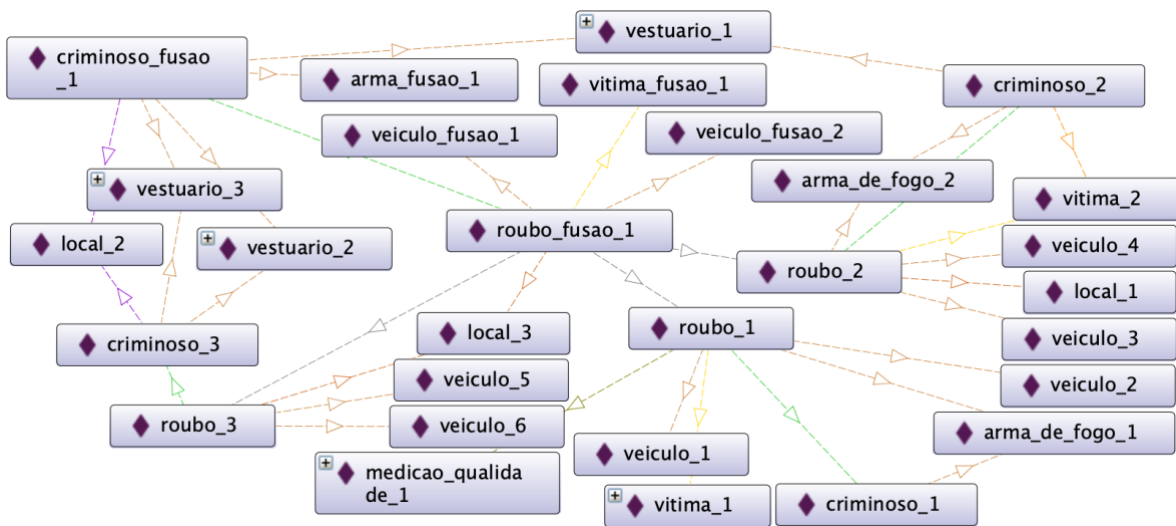


Figura 38 – Representação das instâncias. Fonte: o próprio autor.

Ao analisar a Figura 38, é possível notar elementos informacionais, instâncias, dos três relatos de maneira íntegra e com propriedades criando relacionamentos entre elas e todas se unindo à uma única instância de situação, a “roubo_fusao_1”, que amarra e descreve todo o contexto informacional da situação obtido, pelos processos de busca e associação semântica da fusão de informações.

O último procedimento na camada de Fusão Semântica é a avaliação dos índices de qualidade, neste caso a completude, como já descrito em Seções anteriores. Para o cálculo continua-se aplicando a metodologia IQESA e vocabulário DQV para a representação dos índices dentro da URBANITY. Ao comparar a avaliação de qualidade nesta camada do

modelo é possível perceber melhoras em relação em relação a completude analisada, pois no exemplo aplicada na camada de Transformação o relato analisado faltavam elementos informacionais, com o resultado da fusão, neste caso, não ocorreu, estando todos os elementos informacionais considerados presentes, atingindo o índice de completude de 100%.

A Figura 39, demonstra como estão relacionadas as instâncias de qualidade ao resultado da fusão, a figura em questão também demonstra como o DQV possibilita a reutilização de métricas e dimensão por diversas avaliações de qualidade. Neste caso compartilha a dimensão e métrica para completude.

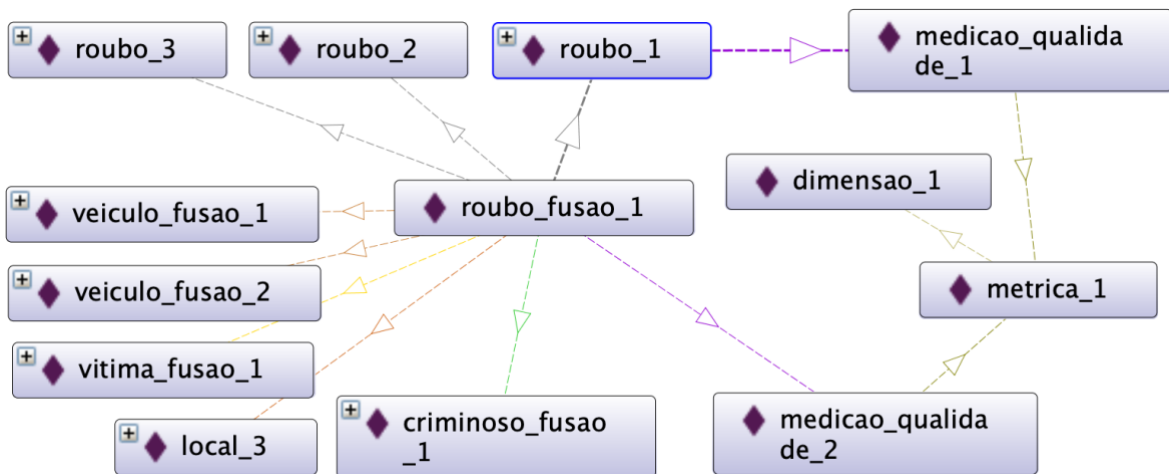


Figura 39 - Representação do resultado da fusão com avaliação de qualidade utilizando o DQV.

Fonte: o próprio autor.

A próxima camada do modelo descrita nesta prova de conceito é a camada de Representação, esta camada está como um ponto de auxílio para a camada de Projeção, pois esta reúne e representa todas as informações necessário, como descrito com mais detalhes na próxima Seção.

6.4. Representação

Após a camada de Fusão Semântica de Informações, que realizou os procedimentos de busca e associação semântica de informações, gerando novas informações e com maior relevância para a análise do cenário, isto através de novas instâncias e avaliação de qualidade. A camada de Representação tem um papel conectivo entre as de Fusão e Cama de Projeção, como demonstrado em destaque na Figura 40, organizando os insumos informacionais para

a representação da projeção informacional pela ontologia, ou seja, irá buscar e dispor instâncias da ontologia que podem ser aplicadas a projeção informacional.

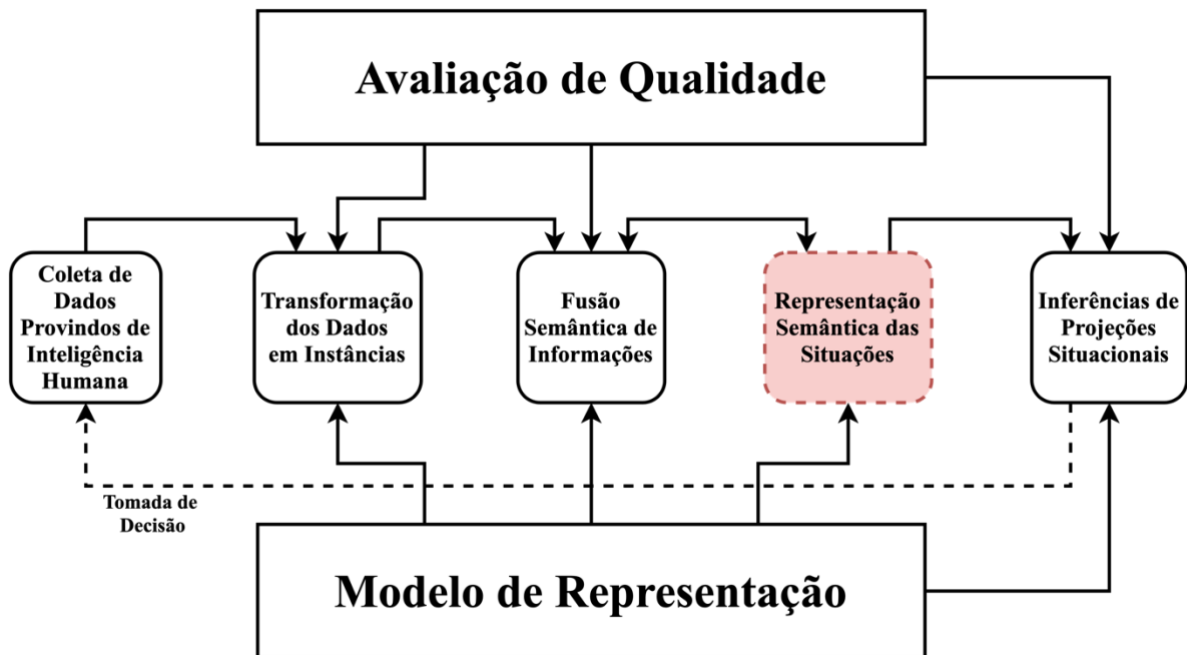


Figura 40 – Destaque para a camada Representação Semântica das Situações. Fonte: o próprio autor.

Tais insumos estão diretamente relacionados aos resultados da camada de Fusão de Informações, ou seja, aquelas informações que já foram associadas, criando assim um escopo informacional maior e mais completo para a projeção informacional. Para buscar e relacionar tais instâncias é novamente criado uma função SPARQL que lista todas as instâncias de situações e informações resultantes de fusão, o SPARQL criado segue o modelo ontológico aplicado, neste caso a URBANITY utilizando das propriedades OWL e do próprio vocabulário criado para pode buscar as informações, como demonstrado na Figura 41.

```
SELECT distinct ?situation ?criminal ?criminal_properties ?victim ?victim_properties
  ?local ?local_properties ?objects ?objects_properties
WHERE {
  ?situation rdf:type voc:Information_Fusion .
  ?situation voc:done ?criminal .
  ?criminal_properties rdfs:domain <http://localhost:8090/person#Criminal> .
  ?situation voc:affects ?victim .
  ?victim_properties rdfs:domain <http://localhost:8090/person#Victim>.
  ?situation voc:ocorreEm ?local .
  ?local_properties rdfs:domain <http://localhost:8090/local#Local>.
  ?situation voc:has ?objects .
  ?objects_properties rdfs:domain <http://localhost:8090/objeto#Object>.
}
```

Figura 41 – Função SPARQL para listagem de instâncias e propriedades como insumos para a camada de Projeção. Fonte: o próprio autor.

Ao executar essa função SPARQL o resultado obtido será uma listagem de todas as informações relacionadas aos resultados de fusão, como instâncias e propriedades. Conforme a complexidade da situação e informações resultantes dos procedimentos de fusão a complexidade do resultado da função SPARQL da camada de representação será afetada exponencialmente, uma vez que irá considerar todas as possíveis variações entre todas as instâncias e propriedades, como demonstrado na Figura 42, que lista parte do resultado dessa função quando aplicada ao cenário desta prova de conceito, com apenas três relatos.

situation	criminal	criminal_properties	victim	victim_properties	local	local_properties	objects	objects_properties
roubo_fusao_1	criminoso_fusao_1	has_projection	vitima_fusao_1	has_projection	local_3	cep	veiculo_fusao_2	nome_objeto
roubo_fusao_1	criminoso_fusao_1	has_projection	vitima_fusao_1	has_projection	local_3	estado	veiculo_fusao_2	nome_objeto
roubo_fusao_1	criminoso_fusao_1	has_projection	vitima_fusao_1	has_projection	local_3	rua	veiculo_fusao_2	nome_objeto
roubo_fusao_1	criminoso_fusao_1	has_projection	vitima_fusao_1	has_projection	local_3	latitude	veiculo_fusao_2	nome_objeto
roubo_fusao_1	criminoso_fusao_1	has_projection	vitima_fusao_1	has_projection	local_3	cidade	veiculo_fusao_2	nome_objeto
roubo_fusao_1	criminoso_fusao_1	has_projection	vitima_fusao_1	has_projection	local_3	bairro	veiculo_fusao_2	nome_objeto
roubo_fusao_1	criminoso_fusao_1	has_projection	vitima_fusao_1	has_projection	local_3	is_projection_local	veiculo_fusao_2	nome_objeto
roubo_fusao_1	criminoso_fusao_1	has_projection	vitima_fusao_1	has_projection	local_3	has_projection	veiculo_fusao_2	nome_objeto
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	numero_local	veiculo_fusao_2	is_projection_object
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	longitude	veiculo_fusao_2	is_projection_object
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	distrito	veiculo_fusao_2	is_projection_object
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	cep	veiculo_fusao_2	is_projection_object
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	estado	veiculo_fusao_2	is_projection_object
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	rua	veiculo_fusao_2	is_projection_object
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	latitude	veiculo_fusao_2	is_projection_object
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	cidade	veiculo_fusao_2	is_projection_object
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	bairro	veiculo_fusao_2	is_projection_object
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	is_projection_local	veiculo_fusao_2	is_projection_object
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	has_projection	veiculo_fusao_2	is_projection_object
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	numero_local	veiculo_fusao_2	has_projection
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	longitude	veiculo_fusao_2	has_projection
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	distrito	veiculo_fusao_2	has_projection
roubo_fusao_1	criminoso_fusao_1	has	vitima_fusao_1	is_projection_victim	local_3	cep	veiculo_fusao_2	has_projection

Figura 42 – Recorte do resultado obtido ao aplicada a função SPARQL para representar a situação.

Fonte: o próprio autor. Fonte: o próprio autor.

Ao analisar o resultado obtido nota-se alguns pontos essenciais que irão possibilitar e também definir a forma como será representada a projeção, como a aparição das propriedades *has_projection* e *is_projection* que irão acrescentar a representatividade da projeção dentro da ontologia. Outro ponto notado e que será abordado na seção de projeção é a relação dos relatos e situações com o tempo, uma vez que os dados aplicados não são de tempo real e as características relacionadas ao domínio mantem a questão da temporalidade implícita as informações. Estes pontos irão afetar diretamente a forma da projeção como demonstrado no próximo capítulo.

6.5. Projeção

Com os insumos gerados, a próxima camada do modelo é Inferência de Projeções Situacionais, o objetivo dessa camada é conseguir gerar e representar as projeções

informacionais baseadas na situação construída nas camadas de Transformação, Fusão e Representação. A camada em questão representa o final de um ciclo de processamento informacional, onde a partir dela o humano dispõe de um arcabouço informacional mais robusto e representativo para tomar uma possível decisão em relação ao cenário real. Ademais, a cama conta com suporte direta das camadas de Avaliação de Qualidade e Modelo de Representação, camadas inerentes em diversos procedimentos do modelo, como demonstrado na Figura 43.

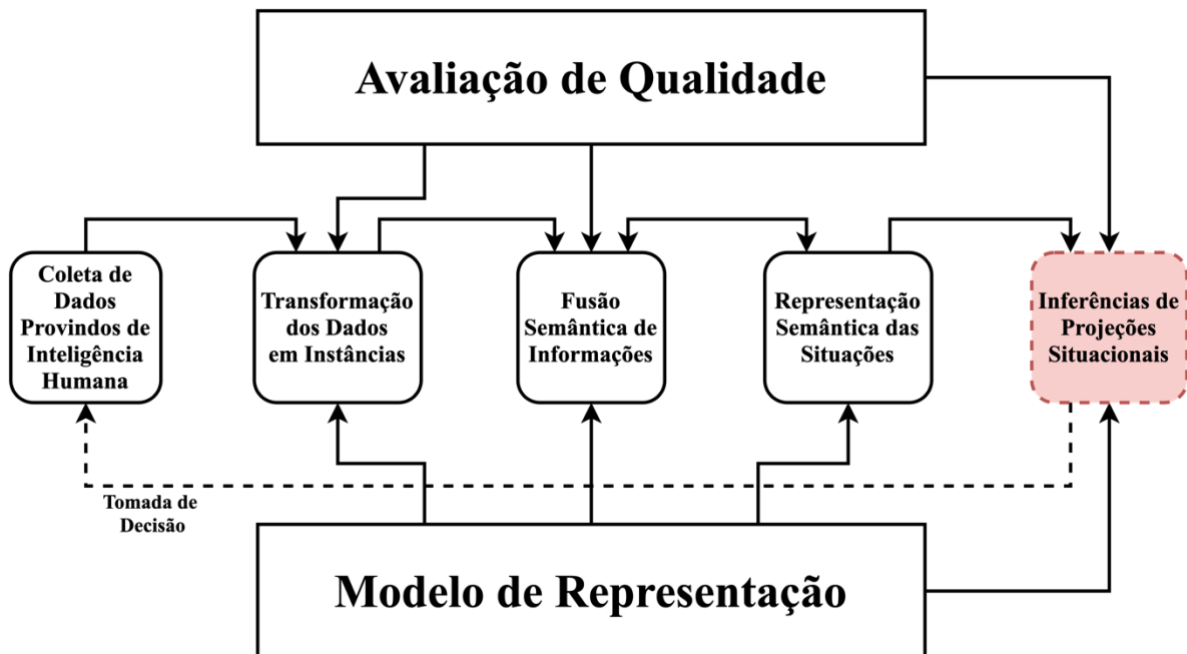


Figura 43 – Destaque para a camada Fusão Semântica de Informações. Fonte: o próprio autor.

Para cumprir com seu papel esta camada pode ser segmentada em quatro processo internos, que irão realizar os procedimentos necessários para conseguir produzir e representar as projeções. A Figura 44 demonstra quais os procedimentos e qual a sua ordem, passando por: Análise das Instâncias (insumo informacional produzido na cama de representação); Criação das Instâncias de Projeção (cada elemento informacional de maneira isolada); Projeção da Situação (criação da projeção utilizando as instâncias criadas no passo anterior); Avaliação da qualidade da situação projetada.

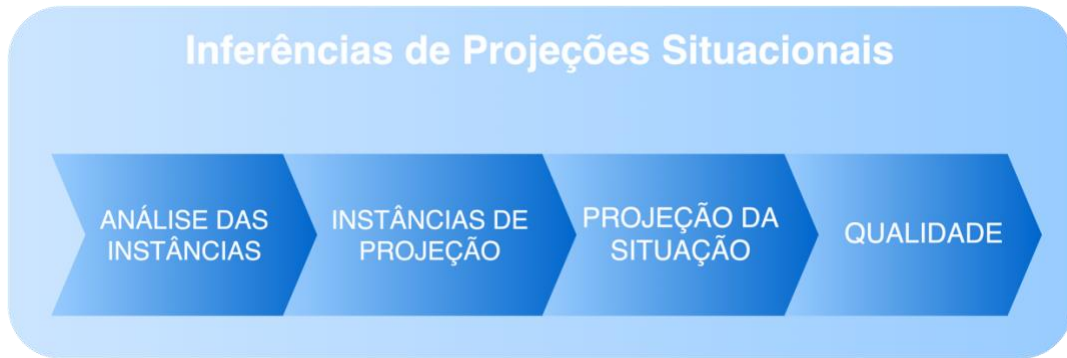


Figura 44 – Processos internos à camada de Projeções Situacionais. Fonte: o próprio autor.

Para os procedimentos de análise das instâncias é necessário cruzar os insumos informacionais gerados na camada anterior do modelo com o modelo ontológico em si, ou seja, analisar quais propriedades de dados e objetos as instâncias apresentam e quais todas as propriedades estão disponíveis no modelo ontológico. Tal comparação permite identificar as lacunas e falhas na informação, onde a partir destas falhas a informação pode evoluir e ser projetada.

Como exemplo, nessa prova de conceito é analisado uma instância de criminoso com o seu modelo informacional, a Figura 45 demonstra a função SPARQL para listar as propriedades da classe criminoso, essa função pode ser alterada conforme a classe a ser alterada conforme a demanda, para isso é necessário trocar o parâmetro em destaque.

```

SELECT distinct ?data_properties ?object_properties ?range_properties
WHERE {
  {
    ?data_properties rdfs:subPropertyOf owl:topDataProperty .
    ?data_properties rdfs:domain person:Criminal
  }
  UNION
  {
    ?object_properties rdf:type owl:ObjectProperty .
    ?object_properties rdfs:domain person:Criminal .
    ?object_properties rdfs:range ?range_properties .
  }
}

```

Figura 45 – Função SPARQL de listagem de propriedades. Fonte: o próprio autor.

A função escrita na Figura 45, considera tanto as propriedades de objetos e dados de uma classe, assim como com quais outras classes as propriedades podem ter relacionamentos, como demonstrado na Figura 46, onde a coluna *object_properties* relaciona as propriedades da classe e a coluna *range_properties*, quais classes cada propriedade se relacionada, isto respectivamente conforme as linhas.

data_properties	object_properties	range_properties
amount		
	rob	Victim
	is_projection_criminal	Criminal
	theft	Victim
	hurt	Victim
	kill	Victim
	has_projection	Projection

Figura 46 – Processos internos à camada de Projeções Situacionais. Fonte: o próprio autor.

Esse procedimento de análise das instâncias conforme o modelo ontológico se replica para todos os elementos informacionais contidos em uma situação, criminoso, vítima, local e objetos, seguindo as especificidades de cada ontologia. Após essa análise o próximo procedimento realizado é a criação das instâncias de projeção, relacionando cada elemento informacional com a classe de projeção da ontologia, agregando assim a representatividade da dar o suporte informacional necessário sobre a projeção da situação. A Figura 47 demonstra como a instância atual se relaciona com a instância da projeção.

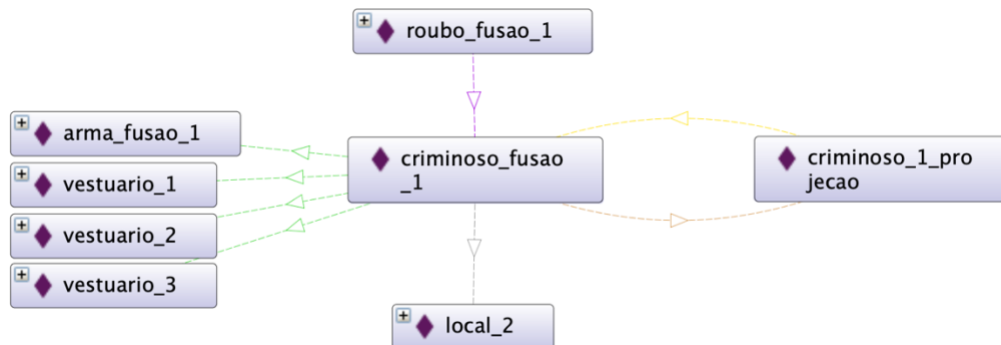


Figura 47 – Resultado do procedimento de criação de instâncias de projeção para a situação. Fonte: o próprio autor.

Esta representatividade da projeção entre as instâncias dos elementos informacionais, assim como a instância da situação, acontece através das propriedades *has_projection* e *is_projection*. Tais propriedades permitem representar a noção de evolução e temporalidade das informações dentro do modelo ontológico, porém, importante notar que neste ponto não é possível expressar e representar o ponto exato da informação no futuro, uma vez que as informações relacionadas ao domínio expressam uma noção implícita de tempo, não esclarecendo a ordem e ponto das informações na linha temporal.

Assim o foco em expressar tal projeção é criar indicadores da evolução informacional e sua representatividade para o humano, auxiliando a compreender o cenário e os possíveis

desdobramentos. Figura 48 explicita a relação de evolução e informação, considerando o tempo de forma implícita, demonstrando como ocorre o processo informacional do procedimento de Projeção da Situação.

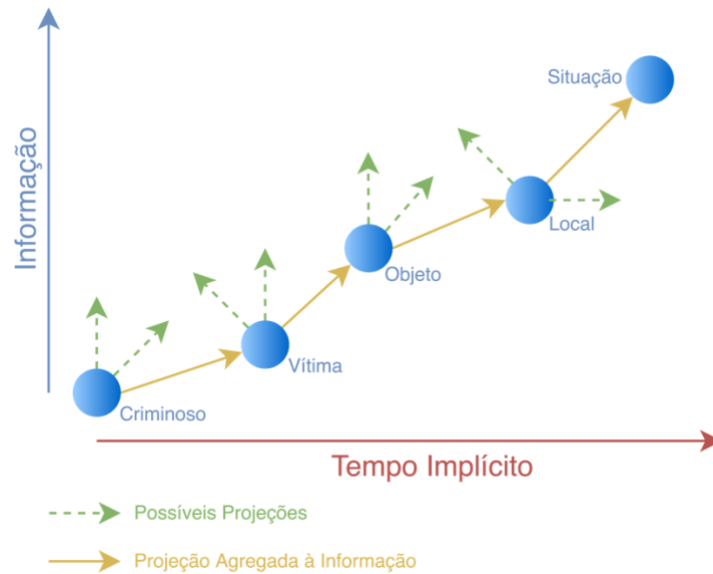


Figura 48 – Relação entre o nível informacional da projeção e temporalidade implícita. Fonte: o próprio autor.

Nesta figura a relação entre as informações e instâncias, geradas no procedimento anterior, são explicitadas através dos círculos, sendo cada uma instância. Conforme a projeção é criada entre as próprias instâncias as propriedades de relacionamento são criadas, representadas pelas setas contínuas. As setas pontilhadas representam as possíveis forma de projetar a informação.

Pode-se interpretar a projeção neste grafo, assumindo que cada instância existe de maneira isolada e a projeção é criada conforme a possibilidade de encontrar novas propriedades e relacionamentos entre tais informações, chegando assim na informação de mais alto nível e tempo mais avançado, sendo está a situação resultante dos procedimentos de projeção. O resultado deste procedimento da projeção da situação é demonstrado na Figura 49, que traz as instâncias projetadas e relacionadas com a projeção da situação.

Importante analisar que mesmo que ocorra todos as etapas da camada de projeção ocorram, não há garantias de que as informações irão evoluir ou se manter, isso devido a qualidade das informações e características de falta de detalhamento e parâmetros informacionais relacionados ao domínio.

O principal fator que pode revelar se houve melhora ou alterações nas informações durante os procedimentos de projeção informacional são os índices de qualidade, sendo este

que o modelo Tree consegue lidar com o fluxo informacional independente do conjunto tecnológico disposto.

Sob um olhar quantitativo, pode-se notar o aumento da quantidade de instâncias e relacionamentos entre as classes, esse aumento significativo demonstra dois fatores. O primeiro fator é o detalhamento maior entre as informações que são representadas, logo, uma quantidade maior de instâncias sobre uma classe, como por exemplo a de criminoso, pode trazer mais detalhes sobre o mesmo e como está sendo sua evolução. Como segundo fator, a adesão de uma classe da ontologia para relacionar possíveis cenários futuros (projeções informacionais) colabora para a solidez do modelo proposto junto ao arranjo de métodos e técnicas.

Todavia, ao mesmo tempo que a prova de conceito provou a eficácia do modelo, uma análise qualitativa e quantitativa demonstra algumas limitações, como por exemplo a capacidade de lidar com o tempo implícito relacionado às informações do domínio, revelando a necessidade de evolução do mesmo para suportar classes e inferências temporais, que podem representar grande parte das informações que auxiliam na compreensão da evolução da situação. O Capítulo 7, a seguir, traz as considerações sobre os resultados obtidos no trabalho, limitações e indica os trabalhos futuros, buscando consolidar o modelo e satisfazer os pontos limitantes identificados.

7. Conclusões

Tendo a Ciência da Informação como uma grande área das Ciências Sociais aplicadas e que se ramifica em diversas frentes e especificações, sempre buscando os melhores meios, modelos e técnicas relativas à informação e como esta pode contribuir no processo de construção de humanos em diversos meios. Considerando também sua abordagem interdisciplinar, perpassando por diversas ciências e áreas do conhecimento, porém um fator se destaca no cenário atual que é aproximação e alta adesão de meios tecnológicos como transmissão e consumo de informação.

A partir deste ponto, disciplinas como a Coleta, Transformação e Representação de Informações, desempenham um papel primordial na produção de insumos informacionais melhores. As demandas informacionais estão intrínsecas ao domínio e ambiente em que o humano está inserido, tendo nuances e questões particulares conforme as necessidades informacionais do domínio se alteram.

O presente trabalho adota para estudo os domínios críticos, especificamente o cenário criminal brasileiro, definindo o ambiente informacional baseado em seus parâmetros e requisitos informacionais. Ao considerar tal ambiente, no qual as informações irão afetar diretamente a imagem da situação criada pelo humano e sua futura tomada de decisão, isto baseado em informações de base histórica. Todavia, mesmo não sendo informações em tempo real estas apresentam grande dinamicidade e complexidade, primeiro por lidar com dados provindos de humanos e segundo pelas características do domínio. Assim, os processos informacionais devem ir de encontro as necessidades informacionais do humano para que consiga construir uma imagem da situação, através do desenvolvimento e manutenção de SAW.

Para suprir tais necessidades informacionais e as particularidades das informações do domínio, pode-se aplicar modelo semânticos, como ontologias, metodologias de avaliação, técnicas de análise e transformação de informações. A aplicação desse arcabouço de modelos, metodologias e técnicas, torna-se viável superar os desafios relacionados as informações, assim como a demanda informacional relacionada ao processo cognitivo humano.

A partir deste contexto e indo ao encontro aos objetivos do trabalho de gerar melhores insumos informações para humanos em situações críticas, de maneira a dar suporte para a SAW e tomada de decisão evitando riscos a vida e patrimônio, isto através de um modelo

que suporte processos semântica com apoio de modelos semântica e avaliação de qualidade o presente trabalho apresentou uma alternativa, o modelo Tree. Capaz de lidar os diversos processos informacionais, que cumprem com o objetivo do modelo de gerar melhorar insumos informacionais através da transformação de dados, avaliação de qualidade, representação semântica, fusão de dados e projeção de informações. Partindo coleta dos dados, passando pela transformação semântica, fusão de informações, até atingir a projeção das informações, sendo que as camadas recebem o suporte direto da avaliação de qualidade, com a metodologia IQESA, e de um modelo ontológico, a URBANITY. Sendo que o modelo se comporta de cíclica e iterativa.

Cabe ressaltar que o modelo construído está diretamente relacionado as características e especificidades do domínio em que está aplicado, no caso, o domínio crítico da criminalidade brasileira, focando em situações de roubo e furto. Sendo assim toda a sua estruturação foi feita baseada nos insumos informacionais analisados, mais especificamente relatos criminais descritos por humanos. Está análise guiou a seleção da URBANITY, os avanços feitos nesta ontologia para dar suporte a projeção, assim como a construção das funções SPARQL de busca e associação de informações. Esse direcionamento do modelo e seus processos conforme a demanda do domínio afeta diretamente os insumos oferecidos ao humano e as decisões tomadas, fator relevante considerando a criticidade do domínio que envolve fatores como vidas, patrimônios e meio ambiente.

Ao analisar os avanços, resultados e a prova de conceito descritos no decorrer do trabalho é possível perceber as contribuições do modelo e dos processos para com a Ciência de Informação, podendo listar:

- Elaboração de modelos ontológicos e vocabulários que suportem a evolução da informação durante os processos de transformação;
- Processos que contribuam avaliação e representação de índices de qualidade;
- Elaboração de funções SPARQL que contribuem para a busca e associação de informações semântica, sendo parametrizadas através do modelo semântico;
- Procedimentos com o apoio de ferramental tecnológico capaz de realizar a coleta, limpeza, padronização e transformação da dados da web em informações semântica.

Porém as contribuições mais relevantes cabem ao ganho na representativa das informações, isto pois a agregação de um modelo ontológico para representar as informações, junto a índices de qualidade e vocabulários específicos para fusão e projeção, resultando em

informações mais conexas e que podem expressar mais conhecimento e contexto ao final do processo. Como estas podem afetar diretamente a construção de SAW do humano em situações críticas, fornecendo-o melhores insumos informacionais para embasar sua tomada de decisão.

Com análises das contribuições percebe-se a dificuldade em lidar com as informações do domínio, as quais são seguem uma padronização ou modelo quando disponibilizadas na fonte, estando de maneira não estruturada e com problemas evidentes de qualidade, como a questão da completude. Para este ponto, nota-se a relevância das camadas de Coleta e Transformação de Dados em Instâncias, principalmente por conseguirem realizar a coleta de diversas fontes com a aplicação de tecnologias de mineração de dados e *scraping*, assim como os processos de limpeza e padronização dos dados, possibilitando agregar padrões e relacionamentos aos dados quando são transformados em instâncias da ontologia.

Outra contribuição observada e que contribuiu diretamente para os avanços informacionais foi a adesão contínua da avaliação de qualidade junto a URBANITY, possibilitando acompanhar o avanço da informação camada por camada do modelo, revelando como cada passo pode afetar a estrutura da informação. Esta adesão também levou a uma conclusão quanto a integração e dinamicidade dos processos, por mais que existe um fluxo informacional direto entre as camadas, este não é um processo totalmente independente e automatizado, pois mesmo a contribuição de diversas tecnologias como linguagens de programação e ferramentas, alguns procedimentos de transformação e transmissão demanda a interferência direta de um humano.

Em complemento, notou-se avanços quanto as questões de projeção informacional através de modelos semânticos, sendo possível fazer a análise e representação das projeções através de funções SPARQL e a construção de um vocabulário específico para o processo. Seguindo por esse aspecto, os avanços realizados na ontologia URBANITY se mostraram válidos, uma vez que foi possível representar como as informações podem se comportar em um futuro próximo, como demonstrado na prova de conceito, além de possibilitar a busca e associação semântica durante os procedimentos de transformação e fusão de informações, através de funções SPARQL e validando-se também a representação contínua dos índices de qualidade.

Essa integração entre as camadas de representação, transformação e avaliação garantem o acompanhando da representatividade da informação, pois a mesma já está associada a índices de qualidade e representada por instâncias da ontologia desde a segunda

camada, “Transformação de Dados em Instâncias”, onde a partir desse momento a semântica e índices de qualidade se mantem por todo o fluxo informacional.

Porém houveram limitações quanto a forma de analisar e inferir questões temporais e seus possíveis desdobramentos, isto dado por dois principais a capacidade de inferência do SPARQL, mas principalmente pela característica dos dados e informações do domínio que trazem a questão da temporalidade de forma implícita, tornando dificultoso o processo de identificar ponto temporais e a partir destes avaliar as possíveis evoluções das informações.

7.1. Trabalhos Futuros

Ao analisar os resultados obtidos com o trabalho, bem como as dificuldades e restrições encontradas durante a construção do modelo e realização da prova de conceito, possibilitou ampliar os horizontes do trabalho quanto a capacidade de coleta, análise, inferência e projeção de informações, em especial com informações de domínio crítico onde a sua influencia sobre o humano é extremamente relevante. Neste sentido, surgem as oportunidades de expansão do trabalho, bem como trabalhos futuros.

Um dos principais e mais relevantes pontos que demandam avanços e extrapolam o escopo do presente trabalho, refere-se a dificuldade enfrentada em encontrar ferramentas e tecnologias fora do cenário acadêmico que dão suporte ao processamento de informações semânticas, conseguindo lidar com dados e informações semântica como OWL e RDF (e suas variações). Essa escassez de tecnologias com uma abordagem semântica limita a escalabilidade e dinamicidade do trabalho, uma vez que determinados processos informacionais e tecnológico não tem integração, exigindo uma interferência manual.

No que se refere a ampliação do modelo ontológico, URBANITY, para dar suporte a projeções situacionais e fusão de informações, revelou uma demanda clara sobre a capacidade de ampliar o contexto e transformação de tais ampliações em um vocabulário, possibilitando assim integrar a representação e inferências de informações para projeção em outros cenários. Isto guardando os devidos estudos e análises do domínio para integrar os vocabulários.

Em relação a avaliação de qualidade, o uso da metodologia IQESA para avaliação dos índices de qualidade indicou a necessidade de sua adaptação conforme o contexto, especialmente a seleção das dimensões, principalmente em contextos que lidam com questões de temporalidade e projeção, trazendo questionamentos sobre como os índices de

qualidade podem se comportar uma vez que tenta-se representar a possibilidade de ocorrer uma situação.

Porém, uma das limitações e mais presente das indicações da necessidade de trabalhos futuros refere-se as funções SPARQL e seus motores, assim como possíveis motores de inferência dentro das ontologias que possam lidar com as questões de temporalidade implícita. Restringindo a capacidade de se consultar e associar dados que apresentam um determinado grau de relação no tempo, porém essa relação acontece de forma não explícita, prejudicando a capacidade de representação dessas informações. Levantando questões de como aperfeiçoar os métodos de consulta e inferência informacional dentro de ontologias.

Referências

- BAADER, F., HORROCKS, I., SATTLER, U., *Description logics as ontology languages for the semantic web. In: Mechanizing Mathematical Reasoning*, p. 228-248, 2005.
- BERNERS-LEE, T., HENDLER, J., LASSILA, O. *The semantic web. Scientific american*, v. 284, n. 5, p. 34-43, 2001.
- BLASCH, E., STEINBERG, A., DAS, S., LLINAS, J., CHONG, C., KESSLER, O., WALTZ, E., WHITE, F. *Revisiting the JDL model for information Exploitation. In: Proceedings of the 16th International Conference on Information Fusion. IEEE*, 2013. p. 129-136.
- BLASCH, E., PLANO, S. *DFIG level 5 (user refinement) issues supporting situational assessment reasoning. In: 2005 7th International Conference on Information Fusion. IEEE*, 2005. p. xxxv-xliii.
- BLASCH, E., VALIN, P., JOUSSELME, A. L., LAMBERT, D., BOSSÉ, É., *Top ten trends in high-level information fusion. In: 15th International Conference on Information Fusion*, p. 2323-2330, 2012.
- BORKO, H. *Information science: what is it?. American documentation*, v. 19, n. 1, p. 3-5, 1968.
- BOTEGA, L. C., OLIVEIRA, A. C. M., PEREIRA JUNIOR, V. A., SARAN, J. F., LADEIRA, L. Z., PEREIRA, G. M. C., ISOTANI, S. *Quantify: An Information Fusion Model Based on Syntactic and Semantic Analysis and Quality Assessments to Enhance Situation Awareness. In: Information Quality in Information Fusion and Decision Making. Springer, Cham*, 2019. p. 563-586.
- BOTEGA, L. C., SOUZA, J. O., JORGE, F. R., CONEGLIAN, C. S., CAMPOS, M. R., ALMEIDA NERIS, V. P., ARAÚJO, R. B. *Methodology for data and information quality assessment in the context of emergency situational awareness. Universal Access in the Information Society*, 16(4), 2017. p. 889-902.
- BOTEGA, L. C., VALDIR, A. P., OLIVEIRA, A. C., SARAN, J. F., VILLAS, L. A., ARAÚJO, R. B. *Quality-aware human-driven information fusion model. 2017 20th International Conference on Information Fusion (Fusion). IEEE*. 2017. p. 1-10.
- CARVALHO, R. N., MATSUMOTO, S., LASKEY, K. B., COSTA, P. C., LADEIRA M., SANTOS, L. L. *Probabilistic Ontology and Knowledge Fusion for Procurement Fraud Detection in Brazil. Uncertainty Reasoning for the Semantic Web ii. Springer, Berlin, Heidelberg*. 2010. p. 19-40
- COSTA, P. C. G., LASKEY, K. B., LASKEY, K. J. *PR-OWL: a bayesian framework for the semantic web*. 2005
- COSTA, P. C. G., Chang, K., Laskey, K. B., & Carvalho, R. N. *High level fusion and predictive situational awareness with probabilistic ontologies. George Mason University, Fairfax, VA, USA*. 2010.

COSTA, P. C. G., Laskey, K. B., CHANG, K. C., SUN W., PARK C. Y., MATSUMOTO, S. *High-level Information Fusion with Bayesian Semantics. Proceedings of the 9th Bayesian Modelling Applications Workshop*. 2012.

ENDSLEY, M. R. 1988. "Design and Evaluation for Situation Awareness Enhancement." *Proceedings of the Human Factors Society Annual Meeting*, 32(2): 97-101.

ENDSLEY, M. R. 2001. "Designing for Situation Awareness in Complex Systems." *Proceedings of the Second International Workshop on symbiosis of humans, artifacts and environment*.

ENDSLEY, M. R., JONES, D. G., *Designing for Situation Awareness: an approach to user-centered design, second edition, Taylor & Francis*, 2012.

ENDSLEY, M. R. 2016. "Designing for Situation Awareness: an approach to user-centered design" CRC Press.

ENGLISH, L. P. *Information quality applied: Best practices for improving business information, processes and systems*. Wiley Publishing, 2009.

GABA, D. M., HOWARD, S. K., SMALL, S. D., *Situation awareness in anesthesiology. In: Human factor*, v. 37, n. 1, p. 20-31, 1995.

GRUBER, T. R., *A translation approach to portable ontology specifications. In: Knowledge acquisition*, v. 5, n. 2, p. 199-220, 1993.

GUARINO, N. *Formal ontology in information systems: Proceedings of the first international conference (FOIS'98)*, June 6-8, Trento, Italy. IOS press, 1998.

HALL, D. L., MCNEESE, M., YEN, J., EL-NASAL, M. S., *A Three Pronged Approach for Improved Data Understanding: 3-d visualization, use of gaming techniques, and inteligente advisory agents*. Pennsylvania State Univ University., Park, Pa. *College Of Information Sciences and Technology*, 2006.

HALL, D. L., JORDAN, J. M. *Human-centered information fusion*. Artech House, 2010.

HITZLER, P., KRÖTZSCH, M., PARSIA, B., PATEL-SCHNEIDER, P. F., RUDOLPH, S. *OWL 2 Web Ontology Language Primer*, 2012. URL <http://www.w3.org/TR/owl2-primer/>.-Zugriffsdatum, v. 12.

ISOTANI, S., BITTENCOURT, I. I. *Dados Abertos Conectados: Em busca da Web do Conhecimento*. Novatec Editora, 2015.

KIM, H., SON, J., JANG, K., *Semantic Data Fusion: from Open Data to Linked Data. In: Proceedings of the European Semantic Web Conference*, 2013.

KIRYAKOV, A., OGNJANOV, D., MANOV, D. *OWLIM—a pragmatic semantic repository for OWL. In: International Conference on Web Information Systems Engineering*. Springer, Berlin, Heidelberg, 2005. p. 182-192.

KOKAR, M. M., ENDSLEY, M. R. *Situation Awareness and Cognitive Modeling*, IEEE Intelligent Systems, vol. 27, no. 3, pp. 91–96, 2012.

LLINAS, J., BOWMAN, C., ROGOVA, G., STEINBERG, A., WALTZ, E., WHITE, F. *Revisiting the JDL data fusion model II*. In: Space and Naval Warfare Systems Command Sand Diego CA, 2004.

MCGUINNESS, D. L., *Ontologies for information fusion*. In: *Proceedings of the Sixth International Conference on Information Fusion*, p. 650-656, julho 2003.

MELO, J. O. S. F., BOTEAGA, L. C., SANTARÉM SEGUNDO, J. E. Metodologia de avaliação de qualidade para dados conectados. In: XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (XVIII ENANCIB). 2017.

MIZOGUCHI, R. *Part 3: Advanced course of ontological engineering*. *New Generation Computing*, v. 22, n. 2, p. 193-220, 2004.

NOY, N. F., MCGUINNESS, Deborah L. *Ontology development 101: A guide to creating your first ontology*. 2001.

OLETO, Ronaldo Ronan. Percepção da qualidade da informação. *Ciência da informação*, v. 35, n. 1, 2006.

PEREIRA JUNIOR, V. A., PEREIRA, G. M. C., BOTEAGA, L. C. *Towards a Process for Criminal Semantic Information Fusion to Obtain Situational Projections*. In: *The Human Position in an Artificial World: Creativity, Ethics and AI in Knowledge Organization*. Ergon-Verlag, p. 51-72. 2019.

PEREIRA, V. A., SANCHEZ, M. F., SARAN, J. F., CONEGLIAN, C. S., BOTEAGA, L. C., ARAUJO R. B. (2016). *Towards Semantic Fusion Using Information Quality and the Assessment of Objects and Situations to Improve Emergency Situation Awareness*. *2016 Eleventh International Conference on Digital Information Management (ICDIM)* (pp. 260-265). IEEE.

PEREIRA, G. M. C., SILVA, J. N., PEREIRA JUNIOR, V. A., OLIVEIRA, A. C. M., BOTEAGA, L. C. Processo de Análise Quantitativa de Eventos Criminais Utilizando Abordagem Semântica. *Informação & Tecnologia - Especial I WIDaT*. V. 4, n. 1.2017.

RAMALHO, R. A. S. *Web Semântica: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da Ciência da Informação*. 2006.

ROGOVA, G. L., BOSSE, E. *Information quality in information fusion, 13th Conference on Information Fusion*, pp. 1–8, 2010.

ROY, J. *From data fusion to situation analysis*. *Proceedings of Fourth International Conference on Information Fusion*. Vol. 2. 2001.

SALERNO, J., *Information fusion: a high-level architecture overview*. In: *Proceedings of the Fifth International Conference on Information Fusion*, v. 1, p. 680-686, 2002.

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, p. 219-239, 2015.

SARACEVIC, T. Interdisciplinary nature of information science. *Ciência da informação*, v. 24, n. 1, p. 36-41, 1995.

SCHIESSL, M., BRÄSCHER, M. *Ontology lexicalization: Relationship between content and meaning in the context of Information Retrieval*. *Transinformação*, v. 29, n. 1, p. 57-72, 2017.

SHELTON, C. L., KINSTON, R., MOLYNEUX, A. J., AMBROSE, L. J. *Real-time situation awareness assessment in critical illness management: adapting the situation present assessment method to clinical simulation*. *BMJ quality & safety*, 2012.

SILVA, J. N., SOUZA, J., OLIVEIRA, Á. C. D., TAVARES, M. D. F., BOTEAGA, L. C. Desenvolvimento de ontologia ciente de qualidade de informações para o domínio de gerenciamento de emergências. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, Florianópolis, v. 23, n. 53, p. 184-200, set. 2018. ISSN 1518-2924.

SOUZA, J., BOTEAGA, L., SANTARÉM SEGUNDO, J.E., BERTI, C., *Conceptual Framework to Enrich Situation Awareness of Emergency Dispatchers, Human Interface and the Management of Information. Information and Knowledge*. In: *International Conference on Human Interface and the Management of Information*, p. 33-44, 2015.

STANTON, N. A., CHAMBERS, P. R., PIGGOTT, J., *Situational awareness and safety*. In: *Safety Science*, v. 39, n. 3, p. 189-204, 2001.

STEINBERG, A. N., CHRISTOPHER L. B., FRANKLIN E. W. 2008. "Revisions to the JDL Data Fusion Model." *Handbook of Multisensor Data Fusion* (pp. 65-88). CRC Press.

STEINBERG, A. N., BOWMAN, C. L., WHITE, F. E., Revisions to the JDL data fusion model. In: *Proceedings of SPIE*, v.3719, n.1, p. 430-441, 1999.

TODORAN, I., LECORNU, L., KHENCHAF, A., CAILLEC, J. L. A methodology to evaluate important dimensions of information quality in systems. *Journal of Data and Information Quality (JDIQ)*, v. 6, n. 2-3, p. 11, 2015.

VINCEN, D., STAMPOULI, D., POWEL, G. *Foundations for system implementation for a centralised intelligence fusion framework for emergency services*. In: *12th International Conference on Information Fusion*, p. 1401-1408, 2009.