UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

ONTOLOGIA APLICADA NOS PROCESSO DE COMPUTAÇÃO FORENSE

EGBERTO VILAS BOAS LEMOS FILHO

ORIENTADOR: BRUNO WERNECK PINTO HOELZ
CO-ORIENTADOR: LAIS DO NASCIMENTO SALVADOR

DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA
ÁREA DE CONCENTRAÇÃO INFORMÁTICA FORENSE E SEGURANÇA DA INFORMAÇÃO

PUBLICAÇÃO: PPGENE.DM - 634/16

BRASÍLIA / DF: DEZEMBRO/2016

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

ONTOLOGIA APLICADA NOS PROCESSOS DE COMPUTAÇÃO FORENSE

EGBERTO VILAS BOAS LEMOS FILHO

DISSERTAÇÃO DE MESTRADO PROFISSIONAL SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA ELÉTRICA DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE.

APROVADA POR:

BRUNO WERNECK PINTO HOELZ, Dr., POLÍCIA FEDERAL
(ORIENTADOR)

FLÁVIO ELIAS GOMES DE DEUS, Dr., ENE/UNB
(EXAMINADOR INTERNO)

AUTO TAVAVRES DA CÂMARA JÚNIOR, Dr., POLÍCIA FEDERAL
(EXAMINADOR EXTERNO)

Brasília, 14 de Dezembro de 2016.

FICHA CATALOGRÁFICA

LEMOS FILHO, Egberto Vilas Boas
Ontologia Aplicada nos Processo de Computação Forense [Distrito Federal] 2016.
xiii, 91 p., 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2016).

Dissertação de Mestrado– Universidade de Brasília, Faculdade de Tecnologia. Departamento de Engenharia Elétrica.

1. Computação Forense 2. Ontologia
3. Exame Pericial

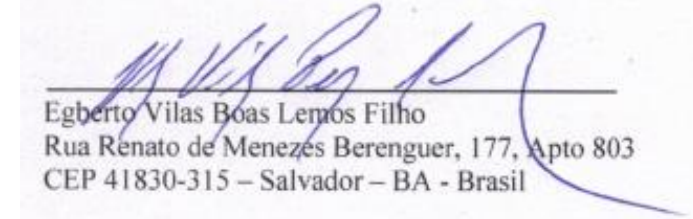
I. ENE/FT/UnB. II. Título (Série)

REFERÊNCIA BIBLIOGRÁFICA

LEMOS FILHO, Egberto Vilas Boas (2016). Ontologia Aplicada nos Processos de Computação Forense. Dissertação de Mestrado, Publicação PPGENE.DM - 634/16, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 91 p.

CESSÃO DE DIREITOS

NOME DO AUTOR: Egberto Vilas Boas Lemos Filho
TÍTULO DA DISSERTAÇÃO: Ontologia Aplicada nos Processos de Computação Forense.
GRAU/ANO: Mestre/2016.



Egberto Vilas Boas Lemos Filho
Rua Renato de Menezes Berenguer, 177, Apto 803
CEP 41830-315 – Salvador – BA - Brasil

Dedico às minhas filhas Clarissa e Alice, que sirva de inspiração para os voos delas.

AGRADECIMENTOS

À minha família, em especial meus pais e filhos, que estiveram ao meu lado, me incentivando a todo o tempo, compreendendo as ausências.

Ao meu orientador Prof. Dr. Bruno Werneck Pinto Hoelz, pelo constante apoio, incentivo, pelas inúmeras contribuições fundamentais para o desenvolvimento deste trabalho.

À minha Coorientadora Prof. Laís do Nascimento Salvador, pela oportunidade de participar das aulas sobre ontologia na UFBA e pelos ensinamentos e contribuições tão necessárias para realinhamentos e melhorias.

À minha colega da UFBA, Rebeca, que pela paciência em me passar preciosas dicas.

Aos meus amigos de turma do mestrado, Vitorino, Vitor, Cirilo e Gustavo, que compartilharam tantas conversas inspiradoras nas saídas das aulas e nos trabalhos em grupo.

Aos meus colegas da Coordenação de Computação Forense, em especial nosso coordenador Marcelo Sampaio, pela ajuda em diversos aspectos e palavras de incentivo.

O presente trabalho foi realizado com o apoio do Instituto de Criminalística Afrânio Peixoto do Departamento de Polícia Técnica do Estado da Bahia, com recursos do Programa Nacional de Segurança Pública com Cidadania - PRONASCI, do Ministério da Justiça.

RESUMO

ONTOLOGIA APLICADA NOS PROCESSO DE COMPUTAÇÃO FORENSE

Autor: Egberto Vilas Boas Lemos Filho

Orientador: Bruno Werneck Pinto Hoelz

Programa de Pós-graduação em Engenharia Elétrica

Brasília, dezembro de 2016

Esta dissertação apresenta uma proposta para o desenvolvimento de uma base de conhecimento no domínio da Computação Forense, suportada por ontologia, com o intuito de promover aumento da eficiência operacional do processo de realização dos exames periciais, por meio da definição de um vocabulário comum entre os atores, e da geração de conhecimento sobres as técnicas e ferramentas forenses aplicadas em cada tipo de evidência digital, dentro do contexto da investigação. A ontologia proposta, denominada Ontocf, foi especificada e conceitualizada a partir dos conhecimentos extraídos do processo de Computação Forense, no qual interagem dois principais atores: a Autoridade Solicitante e o Perito Criminal. O procedimento metodológico adotado para construir a ontologia foi baseado na metodologia METHONTOLOGY, complementado por um tópico preliminar sobre as "Questões de Competência", baseado no processo Ontology Development 101. A validação da ontologia foi feita por meio de análise das respostas das questões de competência e de questionário respondido por Peritos Criminais. Destaca-se que esta ontologia aborda prioritariamente os conceitos, relações e instâncias que podem ser identificados e obtidos nos documentos: pedido de exames periciais e Laudo Pericial. A implementação da base de conhecimento foi feita em um framework que disponibiliza um repositório RDF que pode processar consultas SPARQL. Ao final, as questões de competência foram respondidas pela base de conhecimento das consultas pré-definidas, demonstrando a relevância da ontologia na explicitação do conhecimento do processo de Computação Forense, sobretudo para apoiar a autoridade solicitante na formulação de quesitos e para guiar o perito no planejamento dos procedimentos periciais.

ABSTRACT

ONTOLOGY APPLIED TO THE COMPUTER FORENSICS PROCESSES

Author: Egberto Vilas Boas Lemos Filho

Advisor: Bruno Werneck Pinto Hoelz

Postgraduate Program in Electrical Engineering

Brasília, December of 2016

This dissertation presents a proposal for the development of a knowledge base in the field of Computer Forensics supported by ontology in order to promote an increase in operational efficiency of carrying out expert investigations, by defining a common vocabulary between the Actors and the knowledge production of forensic techniques and tools applied to each type of digital evidence within the investigation context. The proposed ontology, called Ontocf, was specified and conceptualized from the knowledge extracted from the Computer Forensics process, in which two main actors interact: the Requesting Authority and the Forensic Scientist. The methodological procedure adopted to develop the ontology was based on the METHONTOLOGY methodology, complemented by a preliminary topic on "Competency Questions", based on the Ontology Development 101. The validation of the ontology was done by analyzing the answers of competency questions and of a questionnaire answered by Forensic Scientists. It should be emphasized that this ontology focuses primarily on the concepts, relations and instances that can be identified and obtained in the documents: request for expert investigations and Expert Report. The implementation of the knowledge base was done in a framework that provides an RDF repository which can process SPARQL queries. In the end, competency questions were answered by the knowledge base of the predefined queries, demonstrating the relevance of the ontology in making explicit the knowledge of the Computer Forensics Process, mainly to support the requesting authority to formulate questions and to guide the Forensic Scientists in the planning of expert procedures.

SUMÁRIO

LISTA DE QUADROS

LISTA DE FIGURAS

LISTA DE SÍMBOLOS, NOMENCLATURA E ABREVIAÇÕES

Abox	*Assertional Box*
API	*Application Programming Interface*
BPD	*Business Process Diagrams*
BPKB	*Business Processes Knowledge Base*
BPM	*Business Process Modeling*
BPMN	*Business Process Modeling Notation*
CPP	Código de Processo Penal
CSV	*Comma Separated Values*
DAML	*Darpa Agent Markup Language*
DL	*Description Logic*
DPT-BA	Departamento de Polícia Técnica da Bahia
EAP	Estrutura Analítica de Projetos
ETL	*Extract Transform Load*
GT	Glossário de Termos Completo
GUI	*Graphical User Interface*
HTML	*Hyper Text Markup Language*
JSON	*Java Script Object Notation*
KIF	*Knowledge Interchange Format*
KR2RML	*Karma RDB to RDF Mapping Language*
OIL	*Ontology Inference Layer ou Ontology Interchange Language*
OWL	*Web Ontology Language*
PC-BA	Polícia Civil da Bahia
POP	Procedimentos Operacionais de Perícia
R2RML	*RDB to RDF Mapping Language*
RDB	*Relational Database*
RDF	*Resource Description Framework*
RIF	*Rule Interchange Format*
SGBDs	Sistemas Gerenciadores de Bancos de Dados Relacionais
SHOE	*Simple HTML Ontology Extensions*

SPARQL	*Simple Protocoland RDF Query Language*
SSDD	*Small Scale Digital Devices*
SWRL	*Semantic Web Rule Language*
Tbox	*Terminological Box*
TI	Tecnologia da Informação
USC	Universityof Southern California
XML	*EXtensible Markup Language*
XOL	*XML Ontology Language*

1 INTRODUÇÃO

Ciência Forense pode ser definida como a aplicação das ciências à matéria ou problemas legais cíveis, penais ou mesmo administrativos (GIALAMAS, 2000). Segundo Velho, Geiser e Espindula (2013), a utilização das Ciências Forenses para elucidação e constituição da prova das infrações penais e da identidade dos autores respectivos compõe o sistema de Criminalística.

Em uma análise atual, a Criminalística é uma ciência aplicada, com métodos e leis próprias, que utiliza conceitos de outras ciências firmadas nos princípios da física, da química e da biologia (INMAN; RUDIN, 2002). Dessa forma, Criminalística é a organização de conhecimentos oriundos de diversas ciências, com métodos e leis próprias, tendo finalidade última a geração de respostas as questões técnicas formuladas pela justiça, ou autoridade competente nas investigações de crimes, e transmissão destas para instruir um processo (VELHO; GEISER; ESPINDULA, 2013).

A Perícia Criminal define como os objetivos da Criminalística são alcançados seguindo um método científico e estruturado, representados, respectivamente, pelo Exame e Laudo Pericial. A Criminalística e, por conseguinte, a Perícia Criminal, busca estabelecer ou provar três questões fundamentais (VELHO; GEISER; ESPINDULA, 2013):

1. A existência de um crime (O que aconteceu?)
2. A identidade do criminoso (quem?)
3. Seu *modus operandi* (como?)

Os laudos periciais são realizados por meio de conhecimento advindo da Criminalística, que trata da pesquisa, da coleta, da conservação e do exame dos vestígios, ou seja, da prova objetiva ou material no campo dos fatos processuais, cujos encargos estão afetos aos órgãos específicos, que são os laboratórios de Polícia Técnica (GARCIA, 2002). Segundo o autor, perícia, de uma forma geral, é o conjunto de técnicas usadas, visando provar a materialidade do crime e apontar o autor. Uma das perícias realizadas trata-se do exame de corpo de delito. O corpo de delito, por sua vez, é o conjunto de vestígios deixados pelo criminoso. Esses vestígios ou evidências, quando coletados do local do crime, seguem uma cadeia de custódia, que é um conceito básico utilizado por todas as ciências forenses e consiste, em termos gerais, em um registro documental e cronológico sobre o manuseio de uma evidência ou prova. (VELHO; GEISER; ESPINDULA, 2013) .

Para a realização de Exame Pericial, quesitos sobre a evidência podem ser formulados pelas partes, como se infere da leitura do artigo 176 do Código de Processo Penal brasileiro (NUCCI, 2009). Esses quesitos são remetidos aos órgãos de Criminalística por meio de documento, o qual, neste trabalho, é definido como Solicitação de exame pericial ou Guia de exame pericial. O ato de elaborar quesitos para a criminalística não deve ser resumido ao uso de uma lista de perguntas pré-processadas. Os quesitos, válidos para qualquer exame forense, devem ser pensados para cada caso concreto, levando em conta as inovações peculiares do crime (BRASIL, 2012).

A Computação Forense, também ramo da Ciência da Computação, é um dos diversos ramos da ciência que integram as Ciências Forenses, conceituada, mais especificamente, como a aplicação da ciência à identificação, coleta, exame e análise de dados, preservando a integridade da informação e mantendo uma estrita cadeia de custódia dos dados (KENT et. al., 2006). Sendo assim, as perguntas comumente realizadas na Perícia Criminal de Computação Forense são, entre outras:

- O suspeito produziu o vestígio?
- A evidência digital produziu o vestígio?
- Existem arquivos ou fragmentos de determinado conteúdo nos dispositivos?
- Ocorreu violação de segurança no acesso aos dados do dispositivo?
- Existem vestígios de ação de softwares maliciosos no dispositivo?
- Existem vestígios de fatos delituosos no dispositivo?

Os procedimentos escolhidos pelo perito na elaboração das respostas dos quesitos do Exame Pericial são definidos por diversas variáveis, dentre elas: a tecnologia do artefato digital; o emprego da evidência no contexto do evento criminal, e até mesmo pelas ferramentas computacionais disponíveis para o exame. Normalmente, durante os procedimentos preliminares do exame, grande quantidade de dados são processados, antes do processo de análise dos dados para responder a quesitação postulada pelos investigadores.

O conjunto de crimes relacionados aos meios digitais mudam rapidamente, acompanhando a evolução e disseminação da tecnologia da informação (TI), sobretudo dos dispositivos de armazenamento e transmissão de dados. Nesse contexto, é cada vez mais difícil realizar a busca das evidências digitais, assim como manter a sua cadeia de custódia (ĆOSIĆ; ĆOSIĆ, 2012).

Segundo Alzaabi, Jones e Martin (2013), o domínio de Computação Forense está enfrentando uma série de grandes desafios com o aumento do tempo e do esforço necessário

para analisar dados a partir de dispositivos digitais. Esse é o resultado de uma tendência do aumento da capacidade de armazenamento, da diversidade e complexidade dos formatos de dados. Tais tendências influenciaram o processo de identificação de traços relevantes que normalmente são rodeados por um vasto volume de traços irrelevantes, o que faz o investigador passar a maior parte do tempo para a compreensão da estrutura dos dados em vez de localizar elementos relevantes. Todas essas características geram inúmeras possibilidades de métodos e estratégias empregados nos exames periciais, o que torna necessário sistematizar o conhecimento sobre esses itens para o melhor emprego das técnicas e das ferramentas e para reutilização desse conhecimento.

Uma estratégia para sistematizar o conhecimento sobre Computação Forense é por meio da criação de uma base de conhecimento sobre o tema. Nonaka e Takeuchi (1997) afirmam que as atividades criadoras de conhecimento são captadas e recontextualizadas na base de conhecimento da empresa como um todo, tanto para os conhecimentos explícitos, quanto para os tácitos. Moigne, Pinheiro e Paz (1977, p. 34) afirmam que "modelar é conceber, para um objeto, um modelo que permita conhecê-lo, compreendê-lo, interpretá-lo e auxilie na antecipação do comportamento dele", ou seja, o processo de conhecer passa pela construção de modelos de um domínio. Na visão dos autores, o observador é um sujeito ativo que procede a uma descrição comunicável do que percebe e do que concebe.

As atividades do processo de Computação Forense possuem o potencial de serem a fonte de dados para criação da base de conhecimento por meio de uma ontologia, definida em Gruber (1993, p. 199), como: "Especificação explícita de uma conceitualização", ou seja, uma conceitualização compartilhada de um determinado domínio de conhecimento. Ela é composta de um conjunto de conceitos dentro desse domínio, sendo esses organizados como uma taxinomia, e de relações entre esses conceitos. Uma ontologia pode também possuir axiomas, ou seja, regras pertinentes ao domínio em questão. Ela pode ser utilizada como um esqueleto formado por um conjunto de termos ordenados hierarquicamente para descrever um domínio, com o objetivo de construir uma base de conhecimento (SWARTOUT et. al., 1996).

Uma base de conhecimento que possa receber consultas na linguagem da ontologia (por exemplo SPARQL) promoverá o conhecimento compartilhado sobre crimes digitais e o processo de Computação Forense entre a autoridade solicitante e o perito, de forma a direcionar as questões a serem examinadas para a obtenção dos resultados esperados. Sem tal conhecimento, a autoridade solicitante pode ter uma visão limitada das possibilidades

investigativas de uma evidência digital e, consequentemente, da capacidade de produção de prova material.

Ontologias computacionais são um meio para modelar formalmente a estrutura de um sistema, isto é, as entidades relevantes e as relações que emergem a partir da sua observação, e que são úteis para os nossos objetivos (GUARINO;OBERLE; STAAB, 2009). Nesse contexto, emerge a relevância de utilizar ontologias relativas aos exames de perícia digital, organizando as diversas entidades relevantes do exame em conceitos e suas relações em um sistema de base de conhecimento.

Entre as principais motivações no contexto das perícias digitais para desenvolvimento de uma ontologia para suportar uma base de conhecimento podem-se destacar:

- falta de conhecimento por parte dos solicitantes das possibilidades de questionamentos, ou linha de investigação, sobre uma determinada evidência digital, endereçando à perícia quesitos inadequados ou com aspectos generalistas (VELHO; GEISER; ESPINDULA, 2013);
- Quesitos inadequados ou não recomendados provocam perdas de eficiência dos exames periciais criminais, pois, uma vez formalizados, aumentam o consumo de recursos da perícia criminal para assuntos não pertencentes à alçada técnica ou que demandariam muito esforço para a produção de resposta praticamente dispensável (BRASIL, 2012);
- diversidade das possibilidades de técnicas e ferramentas que podem ser empregadas em um determinado exame pericial em dispositivos eletrônicos (ALZAABI; JONES; MARTIN, 2013);
- rápida evolução da tecnologia que produz diversas evidências digitais relacionadas a crimes, assim como o crescimento do volume e complexidade das informações armazenadas (ĆOSIĆ; ĆOSIĆ, 2012).

Desse modo, este trabalho busca explorar o uso da ontologia na gestão do conhecimento dos exames periciais na área de Computação Forense, que tanto pode contribuir para a eficiência operacional das atividades periciais, sobretudo pela definição de um vocabulário comum para os atores envolvidos.

Sendo assim, o problema de pesquisa visa responder, se a utilização e reutilização de informação da base de dados suportada por ontologia nos processos de Computação Forense ajudam os Peritos Criminais na execução dos exames, como também, se essas informações

podem subsidiar a autoridade solicitante na investigação. As hipóteses decorrentes do problema apresentado podem ser descritas como:

- utilização de ontologia para construção de bases de conhecimento na área de Computação Forense promove melhor comunicação entre os atores, Perito Criminal e Autoridade Solicitante, pela utilização de vocabulário comum;
- utilização de ontologia para construção de bases de conhecimento na área de Computação Forense promove melhoria dos processos por meio da definição de procedimentos comuns aos atores;
- bases de conhecimento suportada por ontologia das evidências digitais podem auxiliar na busca de informações sobre procedimentos empregados nas evidências associadas a casos semelhantes dentro de um contexto criminal.

O uso da ontologia na formação de base de conhecimento sobre os exames periciais permite representar a diversidade de métodos de exames em diversos tipos de materiais e meios eletrônicos. Busca-se, com o emprego da ontologia, classificar os exames realizados de forma que se possa inferir quais as técnicas e metodologias que foram utilizadas para casos semelhantes. Essa conexão sugere a possibilidade de classificação dos crimes cibernéticos de acordo com as evidências coletadas no ambiente cibernético e com a aplicação das técnicas e ferramentas periciais correspondentes. Em resumo, busca-se o aumento da eficiência operacional por meio da otimização das requisições periciais e pela seleção inicial das técnicas e ferramentas forenses aplicáveis, dentro de um contexto criminal.

1.1 OBJETIVOS

O objetivo geral deste trabalho é desenvolver um modelo agregado de base de conhecimento baseado em ontologia sobre exames periciais de Computação Forense, que possam auxiliar os atores na produção de informações de apoio as atividades de solicitação e realização dos exames. Para tanto, faz-se necessário a coleta de dados dos exames das evidências materiais obtidas durante o processo pericial nos artefatos eletrônicos, sistemas computacionais, arquivos armazenados em meio digital, de forma que essa informação possa ser reutilizada em uma nova solicitação ou perícia de Computação Forense e, consequentemente, na investigação dos seus autores e das organizações criminosas.

1.2 RESULTADOS ESPERADOS

Espera-se dessa pesquisa obter diretrizes para desenvolvimento de base de conhecimento para apoio ao processo de perícia e de inteligência aplicada, correlacionando as informações que permitam:

- melhorar a comunicação dos atores por meio de um vocabulário comum e da explicitação do conhecimento sobre os procedimentos periciais;
- promover a reutilização de técnicas aplicadas na Computação Forense, considerando: os questionamentos da autoridade solicitante, as tecnologias utilizadas no incidente investigado, como, também, o fator temporal dos acontecimentos;
- aumento da eficiência operacional dos Exames Periciais;
- sistematizar o conhecimento sobre os procedimentos de Computação Forense por meio da classificação das solicitações de exames correlacionando com a metodologia e ferramentas empregadas.

Diante do exposto, este trabalho visa contribuir, neste contexto, com a construção de uma base de conhecimento por meio de ontologia do domínio do processo de Computação Forense, que envolve as atividades de solicitação e execução do exame pericial. Para direcionar as atividades desenvolvidas neste trabalho, utilizou-se o conceito de Estrutura Analítica de Projeto (EAP)[1], que explicita as cinco subdivisões de entrega deste trabalho, conforme apresentado na Figura 3.1, do Capítulo 3 desse trabalho.

Dessa forma, organizou-se o texto da seguinte maneira: o Capítulo 2 apresenta o estudo bibliográfico realizado sobre Computação Forense e Ontologia; o Capítulo 3 apresenta trabalhos relacionados a ontologia aplicada na Computação Forense; o Capítulo 4 apresenta a proposta da ontologia do processo de Computação Forense, inclusive a metodologia e os passos de implementação da solução tecnológica da base de conhecimento; o Capítulo 5 apresenta os resultados dessa implementação; e o Capítulo 6 finaliza com as conclusões e trabalhos futuros.

[1]EAP – Estrutura Analítica de Projetos definida pelo Project Management Institute**PMI**®

2 ONTOLOGIAS E COMPUTAÇÃO FORENSE

Neste Capítulo serão abordados os dois principais tópicos do referencial teórico que embasa a construção deste trabalho: ontologia e Computação Forense.

2.1 ONTOLOGIA

Uma ontologia é um conjunto de termos ordenados hierarquicamente para descrever um domínio que pode ser usado como um esqueleto para uma base de conhecimento (SWARTOUT; PATIL; KNIGHT, 1996). Segundo os autores, uma ontologia de domínio deve possuir um conjunto de termos organizados com uma hierarquia associada, ou seja, uma taxonomia, entretanto Guarino (1998) defenda que a ontologia é uma estrutura não exclusivamente taxonômica. Se dois especialistas do domínio constroem suas bases de conhecimento com uma mesma ontologia, os sistemas gerados irão compartilhar uma estrutura comum, o que facilitará o compartilhamento de conhecimento entre essas duas bases.

Uma ontologia com um conjunto de instâncias de classes individuais constrói uma base de conhecimento (ĆOSIĆ; ĆOSIĆ, 2012). Sendo assim, a ontologia fornece um conjunto de conceitos e termos para descrever um determinado domínio, enquanto a base de conhecimento usa esses termos para descrever uma determinada realidade. Na medida que essa realidade é alterada, a base de conhecimento acrescenta essa alteração. Já a ontologia, só será alterada se houver necessidade de refletir uma mudança no domínio.

A ontologia também é utilizada para compartilhar o conhecimento e reusar documentos da Internet de forma automática por aplicações ou agentes. Isso só é possível se e os documentos contém informações ontológicas codificadas, comumente denominada anotação semântica (*semanticmarkup*), que permite aos agentes de *software* interpretar precisamente o seu significado (GASEVIC; DJURIC; DEVEDZIC, 2006).

Oren et al. (2006) diferenciam três tipos de anotações semânticas: informais, formais e ontológicas. Anotações informais não são legíveis por máquina, porque não usam uma linguagem formal. Anotações formais são interpretáveis por máquina, mas não o fazem por si só. Nas anotações ontológicas, a terminologia tem um significado normalmente entendido que corresponde a uma conceitualização compartilhada chamada de ontologia. As descrições de

computador interpretável dos recursos são a base para a web semântica (OREN et al., 2006). O mecanismo mais usado de anotação para dados de computador é o "*tagging*", ou seja, associar ao dado algum rótulo (*tag*) compreensível pelos algoritmos. *Tags* são usadas em blogs, redes sociais e outros sites para associar termos aos seus recursos, como, por exemplo, postagens de blogs, fotografias e comentários (HOELZ; RALHA, 2013).

2.1.1 Definição

As ontologias são estruturas de representação do conhecimento que permitem o processamento semântico das informações, bem como a construção dos sistemas baseados em conhecimento, os quais fornecem uma maior efetividade em relação aos sistemas tradicionais.

Ontologia é uma especificação explícita da conceptualização de um processo (GRUBER, 1993). Uma Ontologia visa capturar o conhecimento declarativo do domínio e fornecer uma compreensão deste, possibilitando o reuso e o compartilhamento por meio de aplicações em grupos.

Para Studer et al. (1998) uma ontologia é uma especificação explícita e formal de uma conceitualização compartilhada. Os autores ainda definem esses termos:

- especificação explícita: definições declarativas de conceitos, instâncias, relações, restrições e axiomas;
- formal: declarativamente definida, sendo compreensível e manipulável para agentes e sistemas;
- conceitualização: modelo abstrato de uma área de conhecimento ou de um universo limitado de discurso;
- compartilhada: conhecimento consensual, seja uma terminologia comum da área modelada, ou acordada entre os desenvolvedores dos agentes que se comunicam.

Fundamentalmente, ontologias são utilizadas para melhorar a comunicação entre seres humanos ou computadores. (JASPER; USCHOLD, 1999). Em termos gerais, estas utilizações podem ser agrupadas nas três áreas seguintes:

- ✓ para ajudar na comunicação entre agentes humanos
- ✓ para promover a interoperabilidade

- ✓ para melhorar o processo e/ou a qualidade da engenharia dos sistemas de software.

Uma ontologia incluirá necessariamente um vocabulário de termos e a especificação de seu significado (definições e relações entre conceitos) que impõe uma estrutura ao domínio e restringe possíveis interpretações (JASPER; USCHOLD, 1999).

Ontologia define um vocabulário comum para uma área de estudo que precisa compartilhar informação dentro de um domínio. Isso inclui definições em linguagem de máquina interpretável dos conceitos dentro do domínio e as relações entre eles. (ĆOSIĆ; ĆOSIĆ, 2012).

Para garantir que uma ontologia seja construída com qualidade, é necessário definir o domínio de conhecimento com objetividade, descrevendo o conhecimento essencial ao domínio e definindo um vocabulário que evite interpretações ambíguas (GRUBER, 1993).

Ontologias são utilizadas como descrições formais e explícitas de conceitos dentro de um domínio. Para isso, devem-se definir (MORAIS; AMBRÓSIO, 2007):

- Classes: Normalmente organizadas em taxonomias, as classes representam algum tipo de interação da ontologia com um determinado domínio;
- Relações: Representam o tipo de interação entre os elementos do domínio (classes);
- Axiomas: São utilizados para modelar sentenças sobre classes e relações consideradas sempre verdadeiras;
- Funções: Eventos que podem ocorrer no contexto da ontologia;
- Instâncias: São utilizadas para representar elementos específicos das classes, isto é, os próprios dados da ontologia.

Uma ontologia em combinação com um conjunto de instâncias de classes individuais constrói uma base de conhecimento (NOY; MCGUINNESS, 2001).

Para Stevens (2000) as relações da ontologia podem ser categorizadas em dois grandes tipos:

- Taxionômicas, as quais são utilizadas para definir a estrutura de árvore das classes/conceitos dentro da ontologia, como, por exemplo, a relação de herança "é um tipo de" (*is a kindof*). Para uma instância, Livro de Ciências é um tipo de livro. Uma taxonomia é um sistema de classificação que agrupa e organiza o conhecimento num domínio usando relações de generalização/especialização por meio de herança

simples/múltipla, ou seja, identificação das classes de "tipo de" ou “é um”. Como exemplo a classe “Autor/Agente” ou “Vítima", é tipo de “Pessoa”.

- Associativas, as quais relacionam as classes/conceitos em toda a ontologia. Exemplos de relações associativas são diversos, e dependem do domínio abordado. Como exemplo a relação “possuiCrime” pode ser utilizada para descrever a relação entre as classes “Autor/Agente” e “Crime”.

Um exemplo de ontologia de alto nível que pode ser utilizada na área de Computação Forense é o PROTON *ontol-ogy*. Essa ontologia é um avanço do projeto KIMO, que consiste em 300 classes e 100 propriedades, e provê cobertura para os conceitos gerais para uma grande variedade de atividades, incluindo anotação semântica, indexação e recuperação de informações (HOELZ; RALHA, 2013).

2.1.2 Classificação da Ontologia

A principal classificação de uma ontologia é com relação ao propósito para o qual ela é definida (natureza do assunto) (USCHOLD; GRUNINGER, 1996). Nesta dimensão, as ontologias podem ser classificadas em três categorias:

a) Ontologias de domínio: expressam conceituações específicas para domínios particulares;

b) Ontologias de tarefa: especificam as conceituações necessárias para se definir uma tarefa;

c) Ontologias de representação: explicam as conceituações utilizadas nos formalismos de representação do conhecimento, sendo neutras em relação às entidades do mundo (ou seja, o domínio).

Já em Guarino (1998), é proposta outra abordagem, classificando as ontologias em níveis de generalidades:

a) Ontologia de alto nível (*Top-levelontologies*): conceitos muito genéricos, independentes de um problema ou domínio partícula;

b) Ontologia de domínio: descrevem o vocabulário relativo a um domínio específico, através da especialização de conceitos presentes na ontologia de alto nível;

c) Ontologia de tarefa: descreve o vocabulário pertinente a uma tarefa genérica ou específica através da especialização de conceitos presentes na ontologia de alto nível;

d) Ontologia de aplicação: descreve conceitos dependentes do domínio e da tarefa particulares.

Uschold e Gruninger (1996) também classificam as ontologias quanto ao grau de formalidade, que expressam os seus termos e os seus diferentes significados. São os seguintes graus de formalidade:

a) Altamente informal: quando é expressa livremente em linguagem natural;

b) Estruturada informal: quando é expressa em linguagem natural, de forma restrita e estruturada;

c) Semiformal: quando é expressa em uma linguagem artificial, definida formalmente;

d) Rigorosamente formal: quando é expressa com semântica formal, teoremas e provas.

O nível de formalidade do documento de ontologia depende da forma de especificação, que pode ser a linguagem natural, questões de competência ou a abordagem "*middle-out*" (USCHOLD; GRUNINGER, 1996). A abordagem "*middle-out*", em oposição as clássicas abordagens "*top-down*" e "*bottom-up*", mostra-se mais adequada a especificação de ontologia. A abordagem "*top-down*" identifica os conceitos mais gerais primeiro, já a abordagem "*bottom-up*" identifica primeiro conceitos mais específicos. A abordagem "*middle-out*", por sua vez, permite identificar os conceitos primários e posteriormente especializá-los ou generalizá-los, tornando os conceitos mais estáveis, além de exigir menos esforço e re-trabalho.

2.1.3 Metodologia de Desenvolvimento de Ontologias

Os tópicos seguintes apresentam duas metodologias, uma efetivamente empregada, a metodologia *METHONTOLOGY*, e a metodologia 101, que forneceu o conceito de "questões de competência" no processo de especificação da ontologia.

2.1.3.1 Metodologia 101

Embora algumas ideias da construção de ontologias sejam baseadas na literatura da desenvolvimento orientado a objeto, o desenvolvimento de ontologias é diferente do desenho de classes e relações da programação orientada a objetos. (NOY; MCGUINNESS, 2001). Os autores acrescentam que uma ontologia é um modelo do mundo real e os seus conceitos devem refletir essa realidade, sendo que não existe apenas uma forma correta de modelar um domínio, sempre há alternativas viáveis; a melhor solução quase sempre depende da aplicação e das extensões que se deseja.

Noy e McGuinness (2001), propõem a Metodologia 101, que é um processo interativo para construção de ontologias, que considera os seguintes passos do processo de construção de uma ontologia:

- Determinar o domínio e o escopo da ontologia;
- Considerar o reuso de ontologias existentes;
- Enumerar os termos importantes na ontologia, definindo as terminologias iniciais;
- Definir as Classes e suas hierarquias.
- Definir as propriedades das Classes (*slots*). Descreve a estrutura interna dos conceitos explicitando suas extrínsecas propriedades (ex. nome, duração, uso), suas intrínsecas propriedades (ex. peso), partes, e relações com outras classes e individuais dessa classe;
- Definir características das propriedades (*slots*), como, por exemplo, tipo de valor, valores permitidos (domínio e faixa), cardinalidade, entre outras características que possam ter;
- Criação de instâncias incluindo a inclusão do valor da propriedade de cada instância criada.

No desenvolvimento da ontologia proposta, o primeiro passo é definir o seu domínio e escopo, por meio das seguintes questões básicas: (NOY; MCGUINNESS, 2001)

- Qual é o domínio que a ontologia irá cobrir?
- Porque será utilizado essa ontologia?

- Quais as questões de informação que essa ontologia proverá respostas – "*Competency questions*" ou Questões de Competência?
- Quem irá usar e manter essa ontologia?

As Questões de Competência são perguntas que se pretende responder a partir de inferências feitas na ontologia. Cabe observar que as respostas dessas questões podem mudar com o desenvolvimento da ontologia, mas deve-se estar atento ao escopo definido do modelo.

2.1.3.2 Metodologia *METHONTOLOGY*

Segundo Fernández-López, Gómez-Pérez e Juristo (1997), O processo de desenvolvimento de uma ontologia descreve quais as atividades são necessárias para a construção da ontologia. Este processo não implica na definição de sequência de execução das atividades, mas define a lista de atividades que devem ser completadas. Os autores propõem a metodologia *METHONTOLOGY* para construção de ontologias, que contempla um conjunto de estágios de desenvolvimento, um ciclo de vida baseado em evolução de protótipos e técnicas para realizar atividades de planejamento, desenvolvimento e suporte.

Figura 2.1 - Componentes da metodologia *METHONTOLOGY*
Fonte: Fernández-López, Gomez-Perez e Juristo (1997)

A *METHONTOLOGY* utiliza o conceito de ciclo de vida incremental (McCRACKEN; JACKSON, 1982 apud FERNÁNDEZ-LÓPEZ; GÓMEZ-PÉREZ; JURISTO, 1997) que não obriga o levantamento de todos os requisitos no estágio de especificação. A ontologia é desenvolvida em camadas, permitindo a inclusão de novas definições somente quando uma nova versão é planejada. Desenvolver protótipos de forma evolutiva é mais apropriado para a desenvolvimento de ontologias.

No Quadro 2.1 estão descritos os principais componentes da metodologia *METHONTOLOGY* (FERNÁNDEZ-LÓPEZ; GÓMEZ-PÉREZ; JURISTO, 1997), considerando inserções de procedimentos adotados por este trabalho, nas atividades e fases previstas na metodologia:

Quadro 2.1 - Planejamento e gerenciamento do projeto *METHONTOLOGY*

ATIVIDADE	**OBJETIVO**
PLANEJAMENTO	Identificar os recursos básicos para desenvolvimento da ontologia. Definir um modelo de gerenciamento do projeto.
ESCALONAMENTO	Estabelecer um cronograma inicial que será ajustado sobretudo com a execução do estágio de especificação. Identificar as tarefas a serem executadas, a maneira como são organizadas e o período e os recursos necessários para realização de cada uma delas.
CONTROLE	Garantir que as tarefas escalonadas sejam contempladas de modo adequado.
GARANTIA DE QUALIDADE	Assegurar que a ontologia, o software e a documentação sejam concluídos com a qualidade devida.

Fonte: Adaptado de Fernández-López, Gomez-Perez e Juristo (1997)

O Quadro 2.2 descreve os estágios de desenvolvimento da ontologia prevista na *METHONTOLOGY*, que segue o que foi planejado e definido nas atividades de planejamento e gerenciamento de projetos.

Quadro 2.2 - Estágios de desenvolvimento *METHONTOLOGY*

ESTÁGIO	**OBJETIVO**
ESPECIFICAÇÃO	Produzir um documento (informal, semi-formal ou formal) de especificação da ontologia em linguagem natural, usando um conjunto de representações intermediárias ou usando questões de competência. As informações mínimas necessárias para a especificação são:

	a) Proposito da ontologia: intenção de seu uso, os possíveis cenários de uso e os usuários finais da ontologia; b) Grau de formalidade: conforme classificação de Uschold e Gruninger (1996); c) Escopo contendo o conjunto de termos que serão representados, suas características e granularidade.
CONCEITUALIZAÇÃO	Estruturar o domínio do conhecimento em um modelo conceitual que descreve o problema e a sua solução nos termos nos termos do vocabulário do domínio. Os termos incluem conceitos, instâncias, verbos e propriedades. Verbos representam ações no domínio. a) Construir o Glossário de Termos Completo (GT); b) Agrupar os conceitos como conceitos e como verbos, identificando os conceitos disjuntos; c) Construir diagrama GT contendo na primeira ramificação a Árvore de Classificação de Conceitos, que organizam o domínio de conceitos em taxonomias, e na segunda ramificação os Diagramas de Verbos; d) Descrever os conceitos usando: Dicionário de dados; Tabelas de atributos de instâncias; Tabela de classes de atributos; Tabelas de constantes; Tabelas de instâncias; e Árvore de classificação de atributos. e) Descrever os verbos usando: Dicionário de verbo para expressar o significado do verbo de forma declarativa; Tabela de condições para especificar um conjunto de condições que tem que ser satisfeitas antes da ação ou conjunto de condições que tem que ser garantidas depois da ação. f) Construir a Tabela de fórmulas e a Tabela de regras, considerando as duas representações intermediárias: Árvore de Classificação de Conceitos e Diagramas de Verbos. Nesse estágio a atividade de aquisição de conhecimento também é realizada no sentido de refinar o glossário preliminar de termos desenvolvidos na fase de especificação e de fornecer significados para os conceitos envolvidos. Em resumo, essa fase produz o modelo conceitual como um conjunto de entregáveis bens definidos para o usuário final que o permite avaliar se a ontologia é útil e usável e para comparar o escopo e completitude com outras ontologias, sua reusabilidade e sua capacidade de ser compartilhada.
FORMALIZAÇÃO	Formalizar o modelo conceitual em uma linguagem formal. *Web Ontology Language* (OWL)
INTEGRAÇÃO	Buscar conceitos em outras ontologias existentes (meta-ontologias) que podem ser reutilizados, como, por exemplo, a "*ontolingua*" e a "*Cyc*". Para tanto, sugere-se os seguintes passos: a) Selecionar nas outras ontologias um conjunto de termos básicos que melhor se aplica na conceitualização a ser desenvolvida; b) Buscar bibliotecas de ontologias que provê definição de termos, os quais a implementação semântica é coerente com os termos identificados na conceitualização proposta. Produzir um documento de integração com as seguintes informações: a) O nome do termo; b) A definição do termo na meta-ontologia; c) Nome da meta-ontologia;

	d) O nome do termo na ontologia em construção.
IMPLEMENTAÇÃO	Codificar a ontologia em uma linguagem por meio de um ambiente que suporte as meta-ontologias e ontologias selecionadas na fase de integração. Os requisitos para seleção do ambiente de suporte para a implementação a ontologia são: a) Analisador léxico e sintático para garantir correções de possíveis erros; b) Tradutores para garantir a portabilidade de definições de outras linguagens; c) Um editor para adicionar, remover ou modificar definições; d) Um aplicativo de pesquisa para examinar as definições mais apropriadas; e) Uma máquina de busca para examinar as definições mais apropriadas; f) Avaliadores para detectar conhecimento incompleto, inconsistente e redundante; g) Gerenciador automático para verificação de inclusão, remoção ou modificação de definições já existentes.
MANUTENÇÃO	Manter a ontologia e dar suporte para correções e ajustes decorrentes de novos requisitos ou evolução dos conceitos e relações existentes.

Fonte: Adaptado de Fernández-López, Gomez-Perez e Juristo (1997)

As atividades de suporte da *METHONTOLOGY* estão descritas no Quadro 2.3.

Quadro 2.3 - Atividades de Suporte *METHONTOLOGY*

ATIVIDADE	**OBJETIVO**
Aquisição do conhecimento	Identificar os conceitos do domínio e elaborar o primeiro glossário de termos potencialmente relevantes que servirá a fase de modelagem conceitual. Essa atividade é independente das outras atividades ou estágios do processo de desenvolvimento de ontologia, embora aconteça simultaneamente com as especificações de requisitos, sendo que a aquisição é mais intensa durante o estágio de conceitualização e decresce com o avanço do processo. As principais fontes de aquisição do conhecimento são: especialistas do domínio, livros, figuras, tabelas, outras ontologias e manuais. O conhecimento pode ser adquirido e explicitado utilizando as seguintes técnicas: brainstorming, entrevistas estruturadas, análise informal e formal de documentos e ferramentas de aquisição de conhecimento.
Documentação	Documentar as várias etapas durante todo processo de desenvolvimento da ontologia. Para cada fase deve ser desenvolvida uma documentação registrando todo o conhecimento pertinente ao que foi feito, conforme descrito abaixo: a) Documento de Especificação de Requisitos; b) Documento de Aquisição de conhecimento; c) Documentos de Modelo Conceitual;

	d) Documento de Formalização; e) Documento de Integração; f) Documento de Implementação; g) Documentos de Avaliação.
Avaliação	Executar um julgamento técnico das ontologias, do ambiente de software e documentação com relação aos requisitos de especificação, de cada fase e entre fases do ciclo de vida da ontologia. Para tanto, as seguintes tarefas devem ser realizadas: a) Verificação – Verificar tecnicamente a consistência da ontologia e dos seus artefatos com respeito ao conjunto de documentos de especificação. b) Validação – Garantir que a ontologia e seus artefatos correspondam ao sistema que supostamente a represente. Deve ser produzido um documento que descreve como foi feita a avaliação, as técnicas utilizadas, os tipos de erros encontrados e as fontes de conhecimento usada em cada avaliação, com isso evita-se a propagação de erros às etapas subsequentes.

Fonte: adaptado de Fernández-López, Gomez-Perez e Juristo (1997, p.37)

A atividade de suporte de Documentação é realizada paralelamente aos estágios de desenvolvimento descritos no Quadro 2.2.

2.1.4 Linguagens de Representação de Ontologias

As linguagens para implementação de ontologias começaram a surgir no início dos anos 90, tomando como base a Inteligência Artificial. Os paradigmas de representação de conhecimento adotados, em sua maioria, eram: a Lógica de Primeira Ordem, como na linguagem Knowledge Interchange Format (KIF); frames combinados com Lógica de Primeira Ordem, como as linguagens Ontolingua, OCML e FLogic; e Lógica de Descrição, como a linguagem Loom (CORCHO; FERNÁNDEZ-LÓPEZ; GÓMEZ-PÉREZ,2003).

O surgimento da Internet levou a criação de outros tipos de linguagens, as chamadas linguagens de marcação de ontologias, que exploravam as características da Web. A Figura 2.2 mostra a relação entre elas.

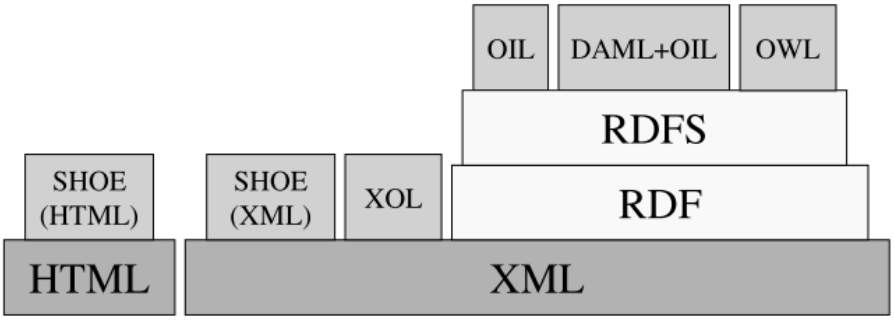


Figura 2.2 - Linguagens de Marcação de Ontologias
Fonte: (CORCHO; FERNÁNDEZ-LÓPEZ; GÓMEZ-PÉREZ, 2003).

- Simple HTML Ontology Extensions (SHOE) foi criada em 1996 como uma extensão do *Hyper Text Markup Language* (HTML) e faz uso de diferentes *tags* permitindo a inserção de ontologias em documentos HTML.
- O *EXtensible Markup Language* (XML) surgiu e foi adotado como linguagem padrão para troca de informações na Web, com isso, a sintaxe da SHOE foi modificada para fazer uso da XML. Outras linguagens de ontologia, como a XML *Ontology Interchange Language* (XOL), também fizeram uso do XML.
- O *Resource Description Framework* (RDF) foi desenvolvido pela W3C como uma linguagem baseada em redes semânticas para descrever recursos na Web e o RDF Schema surgiu como uma extensão do RDF. A combinação de ambos é denominada de RDF(S).

Essas linguagens estabeleceram a Web Semântica e permitiram o surgimento de outras derivadas e construídas como extensões dela, é o caso da *Ontology Inference Layer* ou *Ontology Interchange Language* (OIL), *Darpa Agent Markup Language* (DAML)+OIL e finalmente, a OWL.

2.1.4.1 OWL (*Web Ontology Language*)

OWL também é uma linguagem desenvolvida pela W3C sendo derivada da DAML+OIL e baseada no RDF (WELTY; MCGUINNESS; SMITH, 2004). Foi desenvolvida para ser a linguagem padrão de representação de ontologias e permitir a web semântica, portanto, a

linguagem prioriza itens como extensibilidade, mutabilidade e interoperabilidade, enquanto tenta alcançar uma boa combinação entre escalabilidade e expressividade. OWL permite a descrição de conceitos e também emprega novas funcionalidades, como por exemplo, operadores de interseção, união e negação. Outras funcionalidades de restrição podem ser aplicadas a relação entre os conceitos, isso é feito através da definição do domínio (*domain*) e alcance (*range*) de uma propriedade, determinando que instâncias da classe de domínio se relacionam com instâncias da classe de alcance (MIZOGUCHI, 2004)

Assim como o RDF, faz uso do modelo de triplas para representação do seu conteúdo. Uma tripla consiste em <sujeito, predicado, objeto>, onde sujeito é um recurso que representa algo no domínio modelado; predicado é a ligação entre o sujeito e o objeto; e objeto que pode ser um valor literal ou um outro recurso (MIZOGUCHI, 2004) . Por exemplo: Exame Pericial é feito por Perito Criminal, onde: "Exame Pericial" é o sujeito, "é feito por" é o predicado e "Perito" é o objeto.

As triplas RDF podem estar armazenadas em um banco de dados de triplas (*triple store*) e para que essa informação seja consultada é usado a linguagem *Simple Protocoland RDF Query Language* (SPARQL), a linguagem de consulta da Web Semântica (LAUFER, 2015).

Os bancos de dados de triplas geralmente oferecem pontos de acesso via Web que aceitam o protocolo SPARQL e sua linguagem para consulta. Esses pontos de acesso são chamados de *SPARQL endpoints* e são capazes de aceitar consultas e retornar resultados via HTTP (LAUFER, 2015).

A OWL faz uso de construtores presentes na Lógica de Descrição, do inglês *Description Logic* (DL). De acordo com Isotani e Bittencourt (2015) há duas formas de especificar semântica em OWL:

1) Semântica Direta (OWL 2 DL), que provê o significado através da Lógica de Descrição;
2) Semântica baseada em RDF (OWL 2 FULL), extensão da semântica presente no Esquema RDF e utilizada para visualização de grafos RDF.

Uma ontologia especificada em OWL consiste em um conjunto de afirmações (os axiomas) e possuem dois grupos principais: *Terminological Box* (Tbox) e *Assertional Box* (Abox). O Tbox descreve as classes e os relacionamentos entre elas, e o Abox, por sua vez, capta o conhecimento sobre os indivíduos que pertencem a essas classes (LAUFER, 2015).

A OWL 2 DL é uma versão mais simplificada e restrita que a OWL 2 FULL, mas é decidível, isto é, responde uma pergunta em tempo finito, enquanto que a OWL 2 FULL é indecidível.

A primeira versão da OWL, a 1.0, trazia propostas de OWL 1 DL e OWL 1 FULL, além da OWL 1 Lite, que pode ser vista como uma versão simplificada da OWL 1 DL. Já na segunda versão, a OWL 2, novos conceitos foram adicionados, sendo a maior diferença a definição de perfis, projetados para atender as necessidades de uso e capacidade computacional, tendo por esse motivo diferentes níveis de expressividade. A Figura 2.3 destaca a diferença entre as duas versões da OWL, bem como os diferentes níveis de expressividade.

Figura 2.3 - Diferentes níveis de expressividade da linguagem OWL
Fonte: (ISOTANI; BITTENCOURT, 2015)

Segundo a W3C[2], os três perfis da OWL 2 são (MOTIK; GRAU; HORROCKS, 2009):

- OWL 2 EL: perfil que trabalha com a família EL da Lógica de Descrição, pode ser usado para especificar classes complexas e axiomas sobre elas. Foi projetado para sistemas de alta complexidade e que precisam de maior poder de expressividade;
- OWL 2 QL: perfil que pertence à família DL-Lite da Lógica de Descrição. Seu foco é trabalhar com bancos de dados relacionais. Esse perfil permite fazer uso dos benefícios dos Sistemas Gerenciadores de Bancos de Dados Relacionais (SGBDs) para implementações robustas e tem o foco na reescrita de consultas em SQL;

[2]W3C (https://www.w3.org)

- OWL 2 RL: perfil voltado para aplicações que precisam de raciocínio de forma escalável. Pode ser implementado usando famílias de linguagens de regras ou através do *Rule Interchange Format* (RIF). Esse perfil é ideal para o enriquecimento de dados especificados e conectados via RDF.

2.1.4.2 Restrições, Regras e Raciocinadores

É possível definir restrições baseadas nas propriedades definidas em uma ontologia OWL, isto é, formas de descrever um grupo de indivíduos a partir das relações nas quais eles participam. Segundo Horridge et al. (2004), em OWL existem três categorias de restrições:

1. Restrições de Quantificador: dado um indivíduo, o quantificador adiciona restrição nas relações que aquele indivíduo pode participar. Pode ser de dois tipos:
   a. A Restrição Existencial (*some*): descreve um conjunto de indivíduos que possuem ao menos um tipo de relação específica com indivíduos que são membros de uma classe específica;
   b. A Restrição Universal (*only*): descreve um conjunto de indivíduos que, considerando determinada propriedade, apenas podem ter esse tipo de relação com outros indivíduos que são membros de uma classe específica.
2. Restrições de Cardinalidade: Usadas para determinar o número de relações que um indivíduo pode participar considerando uma propriedade. São três tipos:
   a. Cardinalidade Mínima (Min): especifica o número mínimo de relações que um indivíduo deve participar considerando determinada propriedade;
   b. Cardinalidade Máxima (Max): especifica o número máximo de relações que um indivíduo pode participar considerando determinada propriedade;
   c. Cardinalidade Exata (*Exact*): especifica o número exato de relações que um indivíduo deve participar considerando determinada propriedade.
3. Restrições tem Valor (Restrições *há sValue*): Descreve um conjunto de indivíduos que se relacionam com outros através de uma propriedade. Ao contrário da restrição de quantificador, aqui a classe desses indivíduos com quem a relação ocorrerá não faz parte da restrição, podendo eles pertencerem a qualquer classe.

Semantic Web Rule Language (SWRL) é uma linguagem de regras que combina cláusulas Horn com os conceitos definidos em OWL, podendo ser utilizada para aumentar a capacidade

de inferência sobre os indivíduos de uma base de conhecimento descrita em OWL (HORROCKS et al., 2004).

As regras seguem o formato em que um antecedente (corpo) implica em um consequente (cabeça). Isso significa que, sempre que as condições especificadas no antecedente forem verdadeiras, então, as especificadas na consequente também serão. Ambas as partes consistem em uma conjunção de zero ou mais átomos. O Quadro 2.4 mostra os tipos de átomos possíveis de expressar em SRWL (HASSANPOUR; O'CONNOR; DAS, 2009).

Quadro 2.4 - Tipos de átomos SWRL

Tipos de átomos SWRL	**Prioridade**	**Exemplo de átomo**
Classe do átomo	1	Pessoa(?p), Carro(?c)
Propriedade individual do átomo	2	Has_Driver_License(?p, ?d) Issued_in_State_of(?d,?s) Pode_dirigir(?p, ?c)
Átomo igual/diferente	3	mesmoQue(?x, ?y) diferenteDe(?x, ?y)
Propriedade do átomo	4	temIdade(?p, g?) numero_de_dias_visitados_na_CA(?p, ?x) tem_peso_em_libras(?c, ?w)
Built-in do átomo	5	swrlb:notEqual(?s, "CA") swrlb:lessThan(?g,18)
Data range átomo	6	Xsd:double(?x)

Fonte: (HASSANPOUR; O'CONNOR; DAS, 2009)

Uma sintaxe para representar regras, e de fácil compreensão para humanos, é a destacada no código 1, onde uma seta (->) separa o antecedente e o consequente, o acento circunflexo (^) representa a conjunção entre os átomos e a interrogação (?) distingue variáveis de nomes de indivíduos.

```
Pai (?x, ?y) ^ irmão (?y, ?z) -> tio (?x, ?z)
```

Código 1 – Exemplo de regra

No código 1, temos as variáveis x e y se relacionando pela propriedade pai e (^) as variáveis y e z se relacionando pela propriedade irmão. Essa regra define que nesse caso, como

consequente, têm-se que as variáveis x e z são relacionadas pela propriedade tio. Resumidamente, se y é pai de alguém x e y tem um irmão z, logo z vai ser tio de x.

Os raciocinadores (*reasoners*) são *software*s capazes de mapear uma base de conhecimento (seus conceitos, relações, fatos e regras) e a partir disso inferir consequências lógicas de um conjunto assertivas, fatos ou axiomas, gerando conhecimento adicional, mostrando informações implícitas. Alguns exemplos de inferência são, por exemplo, a classificação, computar de todas as classes existentes, quais um dado indivíduo é membro; e a realização, que é encontrar as classes mais específicas no qual um indivíduo pertence (HORTÊNCIO FILHO; LÓSCIO; CAMPOS, 2008).

Alguns dos raciocinadores mais comuns são: Hermit[3], Pellet[4], Fact++[5], entre outros.

2.2 COMPUTAÇÃO FORENSE

Uma das muitas definições de Computação Forense é a aplicação da ciência e engenharia ao problema jurídico da evidência digital (SAMMES; JENKINSON, 2000). Assim como outras ciências forenses, como, por exemplo, Balística, Documentoscopia, ou Fonética, Computação Forense é conjunto distinto de conhecimento que requer abordagens e ferramentas específicas para o seu objetivo, além de uma educação especializada e capacitação para seus peritos (BEM et al., 2008).

O foco de utilização, até os dias atuais, da Computação Forense para a justiça e forças policiais, tem sido a identificação do crime tradicional. Este foco tem mudado rapidamente para uma abordagem mais ampla, envolvendo os crimes digitais, mas ainda é, em grande parte, dentro da esfera de aplicação da lei e sua necessidade de analisar os sistemas de uma forma juridicamente aceitável, a fim de levar os culpados à justiça (CLARKE, 2010).

[3]http://www.cs.ox.ac.uk/boris.motik/pubs/smh08HermiT.pdf
[4]http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.461.3111&rep=rep1&type=pdf
[5]https://pdfs.semanticscholar.org/e788/f08148b4b06f4e537b27b51338f1c88ccea8.pdf

2.2.1 Crime e Crime Digital

Segundo conceito dogmático ou jurídico de crime, apelidado por muitos autores de "analítico", o crime, portanto, passou a ser definido como: “toda a ação, típica, antijurídica e culpável” (ELEUTÉRIO, 1997).

Em outras palavras, para Roxin (2012), o crime é composto de três requisitos: tipicidade; antijuridicidade e responsabilidade. A culpabilidade, sob essa visão, é elemento componente da estrutura do crime, estando dentro da responsabilidade. A antijuridicidade pode ser conceituada como a contrariedade da conduta com o ordenamento jurídico. A tipicidade representada pelo Tipo Penal, um dos elementos definidores do próprio crime, que na definição de Welzel (apud ROXIN, 2012) é a “descrição concreta da conduta proibida" e, também, é a “matéria da proibição das prescrições jurídico-penais”.

O crime digital ou crime cibernético, em geral, está conectado com os mais diversos crimes cometidos na sociedade, pela evidência digital e pelo ambiente cibernético (PARK; CHO; KWON, 2009). Esta conexão sugere a possibilidade de classificação dos crimes cibernéticos de acordo com as evidências coletadas no ambiente cibernético e a aplicação da lei correspondente.

No âmbito dos órgãos de criminalística brasileiros ressalta-se as iniciativas de produzir procedimentos e manuais que buscam relacionar o contexto criminal com evidência digital, visando orientar tanto o perito, quanto os solicitantes dos exames periciais, nas boas práticas da criminalística. Dentre essas iniciativas destacam-se:

- Procedimento Operacional Padrão Perícia Criminal (POP): que tem por finalidade produzir procedimentos operacionais relacionados às principais atividades periciais, inclusive Computação Forense – POP (BRASIL, 2012);
- Manual de Orientação de Quesitos de Perícia Criminal. Que orienta a formulação dos quesitos para a criminalística dentro de um contexto criminal (BRASIL, 2012);

Ambos os documentos, o POP e o Manual, possuem capítulo específico relacionado a Computação Forense, como também, a proposta de quesitos de solicitação do exame pericial dentro de um contexto criminal.

Tão importante quanto saber os quesitos apropriados no contexto criminal, é importante saber quais são os quesitos não recomendados, pois se revestem de grande importância para a otimização dos exames periciais criminais, pois, uma vez não formalizados, evitam o consumo

de recursos da perícia criminal para assuntos não pertencentes à alçada técnica, ou que demandariam muito esforço para a produção de resposta praticamente dispensável (BRASIL, 2012).

2.2.2 Processo de Computação Forense

A evolução e aumento da complexidade dos exames periciais, em meios digitais, não impacta apenas o trabalho do perito forense, mas todo o ciclo da investigação na qual o artefato tecnológico está inserido. O exame pericial inicia-se a partir do documento de solicitação de exames expedido pela autoridade solicitante, que é encaminhado aos órgãos da Criminalística, sendo designada ao perito forense executor dos exames. Como resultado do seu trabalho, o perito produz o documento denominado, no Brasil, "Laudo Pericial".

Um modelo de processo forense proposto em Kent et al. (2006) e aqui adaptado com a inserção da etapa de "solicitação do exame" é apresentado na Figura 2.4. A etapa de solicitação de exame reflete o processo na maioria dos órgãos de Criminalística no Brasil, quando a autoridade envia a solicitação dos exames. Cabe salientar que, em alguns casos, essa etapa pode anteceder a etapa de coleta, quando esta possuir especificidade técnica ou por procedimento operacional do órgão.

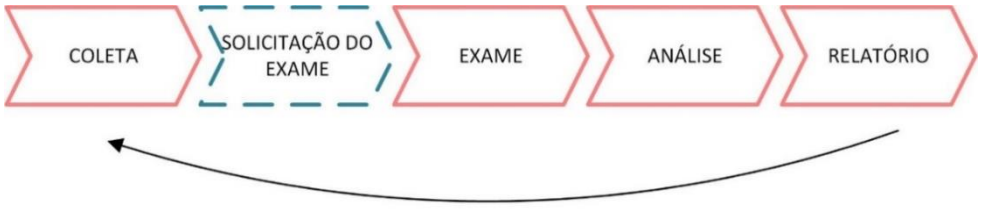


Figura 2.4 - Processo de Computação Forense
Fonte: adaptado de (KENT et al., 2006)

O processo de realização de exame pericial está contido em um processo investigatório de análise da informação digital (COSTA, 2012), conforme diagrama de processo apresentado na Figura 2.5. Esse processo foi realizado segundo as seguintes premissas:

- Descreve o processo de Computação forense no cenário brasileiro;

- Cada solicitação possui apenas um tipo de material com um conjunto de quesitos específicos;
- A aplicação de uma técnica forense, e respectiva ferramenta pode mudar a depender da fase do processo do exame pericial.

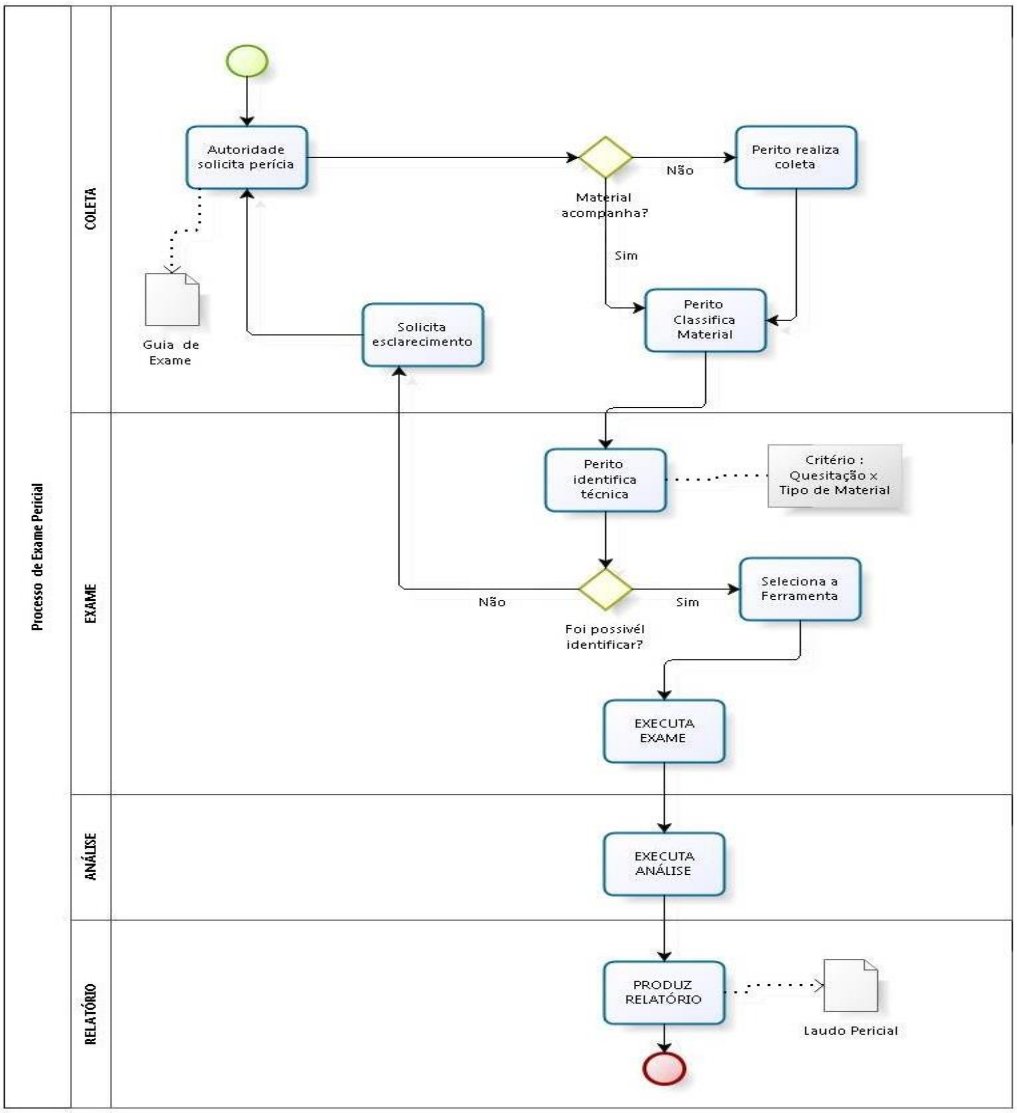


Figura 2.5 - Diagrama de Processo Computação Forense
Fonte: adaptado de (COSTA, 2012)

Durante as fases de Coleta, Exame e em alguns casos na fase de análise, o Perito utiliza procedimentos técnicos, neste trabalho denominada de Técnica Forense, suportados por ferramentas computacionais, denominadas Ferramentas Forenses.

Algumas das Ferramentas Forenses são ferramentas comerciais e outras são iniciativas não comerciais, fortemente voltadas à extração de evidências de um ambiente computacional (COSTA, 2012). Segundo o autor, as ferramentas forenses mais utilizadas no Pericia Criminal, no Brasil, carecem de características que as especializem no processo investigatório de análise da informação.

Um novo e maior desafio para o futuro da Computação Forense é modernizar os seus métodos e processos, para maximizar o rendimento do valor das provas que podem ser obtidas nos meios digitais (BEM et al., 2008). O autor ressalta que a Computação forense já está se movendo para além da análise de imagens de disco rígido. Análise de dados em memória, dos mais variados dispositivos e de redes de dados, e metodologia de investigação, por meio de sistemas em tempo real, estão se desenvolvendo, tanto em termos de pesquisa, quanto em ferramentas específicas de *software* forense.

2.3 ONTOLOGIAS SOBRE O TEMA DE ESTUDOS

Uma ontologia desempenha um papel fundamental na formação de ideias emergentes no campo da cibernéticos forense (HARRILL; MISLAN, 2007).

Um exemplo de ontologia que pode ser utilizada como conjunto de padrões e procedimentos aplicados a dispositivos digitais de pequena escala é a ontologia *Small Scale Digital Devices* (SSDD), desenvolvida por David Harril e Richard Mislan (HARRILL; MISLAN, 2007).

2.3.1 Ontologias de Computação Forense

De acordo com Brinson, Robinson e Rogers (2006), Computação Forense (*cyber forensics*) é uma parte significativa das investigações criminais e cíveis, portanto é de fundamental importância propor ontologias para essa área do conhecimento. Diante desse desafio, os autores propuseram uma ontologia voltada, sobretudo, a desenvolver formação profissional na área dos crimes cibernéticos (*cyber crimes*). O modelo apresentado possui cinco níveis de estrutura hierárquica para o campo da Tecnologia, que está desmembrada em *Hardware* e *Software*.

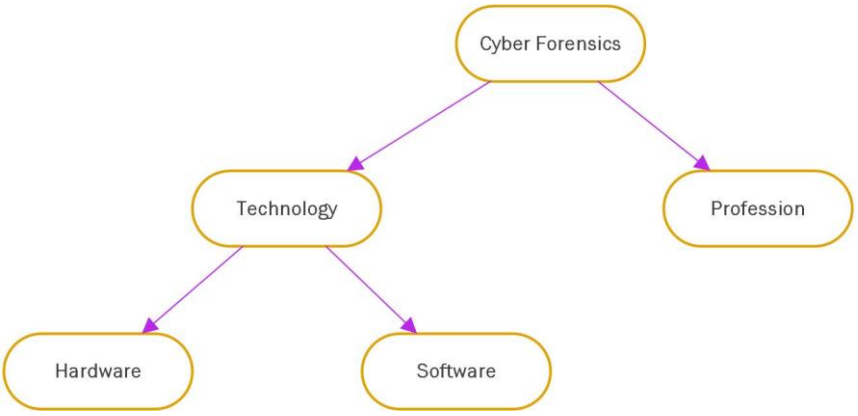


Figura 2.6 - *Cyber forensicsontology*
Fonte: (BRINSON;ROBINSON; ROGERS, 2006)

O modelo ontológico proposto utiliza informações sobre profissões que interagem no domínio dos crimes digitais. O modelo também apresenta uma taxionomia para as tecnologias relacionadas a Computação Forense, incluindo nessas tecnologias as evidências digitais.

Diante da dificuldade de correlacionar os tipos de crimes com as evidências coletadas, Park, Cho e Kwon (2009) desenvolveram uma ontologia para o domínio da Computação Forense, com o objetivo de utilizá-la nos processos de investigação dos casos criminais, e para mineração de dados; classificação, associação e detecção dos tipos de crimes, casos de crimes e evidências.

Durante a investigação de crimes digitais é necessário coletar evidências digitais e o equipamento digital, seja de armazenamento ou de rede de dados, para poder investigar as relações com o crime em geral, para classificar documentos, aplicar relações relevantes com as leis. Para ter eficiência nessa investigação é necessário integrar vários conceitos de crimes digitais (PARK; CHO; KWON, 2009). É necessário, também, identificar qual método, técnica e/ou ferramenta de software que será empregado para proceder buscas, extrair conteúdo, correlacionar artefato digital com o software ou sistema utilizado para produzi-lo.

Park, Cho e Kwon (2009) propõem o modelo de ontologia para crimes cibernéticos apresentado nas Figuras 2.7 e 2.8. Observa-se, nesse modelo, a separação dos crimes em geral com os crimes cibernéticos, e a representação dos conceitos do processo da Computação Forense, como, por exemplo, as classes “*Colection*” e “*Process*”.

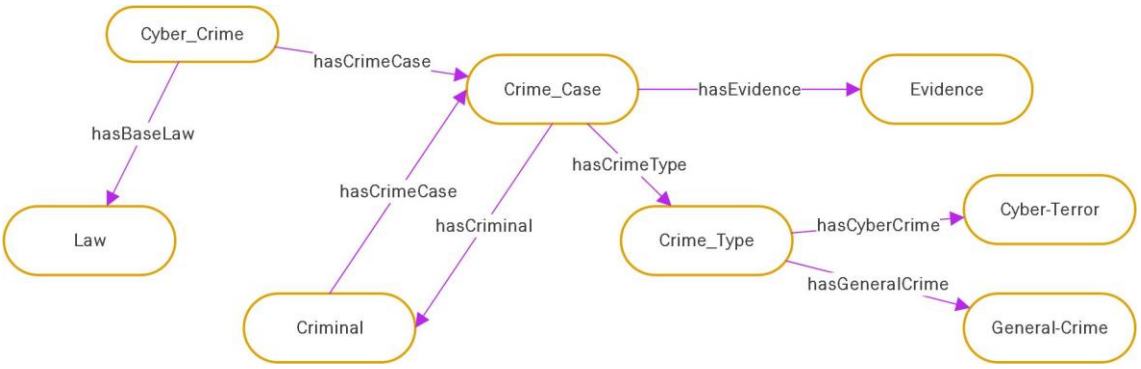


Figura 2.7- Diagrama conceitual de ontologia de crimes cibernéticos
Fonte: (PARK; CHO; KWON, 2009)

A Figura 2.8 apresenta a ontologia de crimes cibernéticos sob a perspectiva do processo de coleta de evidências.

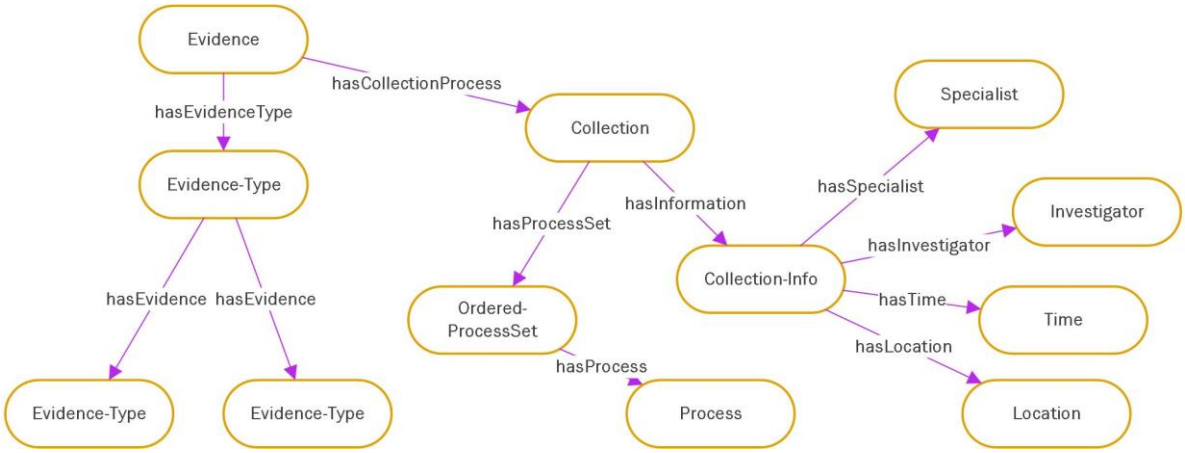


Figura 2.8 - Subclasses e Relações da Evidência e do Processo de Coleta
Fonte:(PARK; CHO; KWON, 2009)

Finalizando o estudo das ontologias de interesse, Azaabi, Jones e Martins (2013) propõem um *framework* baseado em tecnologias de Web Semântica onde ontologias são usadas no modelo do ambiente do exame. Com o conjunto de dados provenientes dos conceitos e relações que envolvem os dispositivos digitais, que atua como uma rede de dados, pode ser construída uma sólida base de conhecimento interligada de diferentes objetos de evidências. No entanto, na medida em que a rede semântica e a ontologia agem sobre as hipóteses, eles também podem direcionar o exame, logo, deve-se compreender tanto a sua natureza quanto as limitações desses métodos, a fim de evitar erros e omissões em seus resultados.

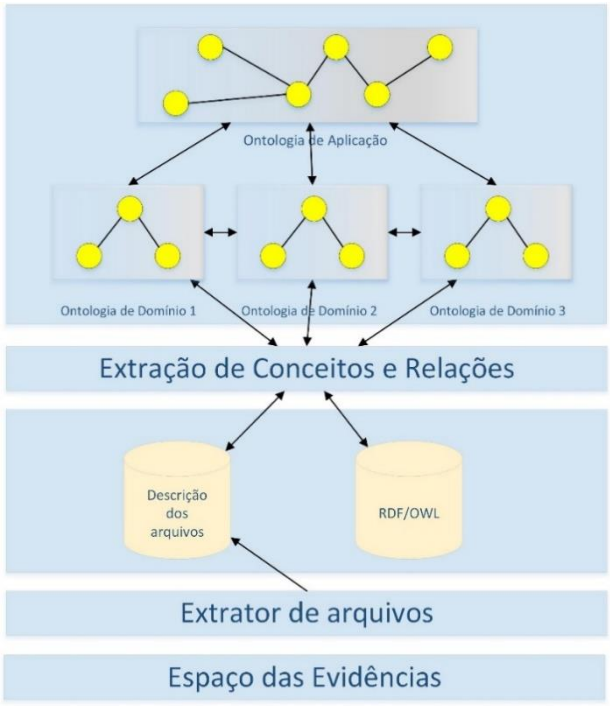


Figura 2.9 - *Framework*- Ontologia para ferramentas de análise forense
Fonte: (AZAABI; JONES; MARTINS, 2013)

A proposta do framework da Figura 2.9 é uma estrutura em camadas baseada em ontologia, que segue o princípio da sobreposição de metadados de evidências. Esses metadados são usados pela camada de "Conceito e Relações" que faz a extração *ondconfute* há vários conceitos e a inter-relação entre eles. Esta extração é realizada com a ajuda de uma representação estruturada de domínios específicos de interesse, que são referidos como ontologias de domínio.

O "Espaço das Evidências" é onde a evidência potencial de objetos como arquivos pretendia ser documentos, imagens, vídeos, e bases de dados estão localizados. Em termos forenses, este espaço pode ser uma imagem forense de um dispositivo sob exame. O "Extrator de Arquivo" é simplesmente um programa que extrai informações descritivas sobre vários tipos de arquivos a partir do "Espaço das Evidências", e essas informações são armazenadas no banco de dados de "Descrição dos Arquivos". O banco de dados "RDF/OWL" armazena, nas respectivas linguagens RDF e OWL, a base de conhecimento de uma potencial evidência sob investigação, proveniente da camada de "Extração de Conceitos e Relações", que tem como objetivos principais extrair conceitos do banco de dados "Descrição dos arquivos" e determinar a que classe este conceito pertence, com base nas camadas de ontologias superiores. Exemplos de tais conceitos, que podem ser obtidos a partir de um smartphone são: um contato que pertence

à classe de “contato”, uma imagem que pertence à classe “mídia”, e um arquivo de documento do *Word*, que pertence à classe de “documentos”. E por último, no framework, tem a camada de “Ontologias de Domínio e Aplicação”. Estas duas camadas, de maneira colaborativa, formam um modelo ontológico para um ambiente particular. Este modelo consiste em conceitos (ou classes) e a relação entre eles.

A ontologia de aplicação é utilizada para montar uma visão sistêmica do modelo e introduz uma nova camada de ontologias de domínio interligadas, por meio da adição de relações de alto nível entre ontologias de domínio. Por exemplo, uma relação “*hasSent*” pode ser utilizado entre a classe de “contato” e a classe de “mensagem” (o que significa que um contato enviou uma mensagem) como uma relação de alto nível. Esta relação constitui uma visão abstrata das relações entre a classe “Contato” e a classe “Mensagem” de ontologias de domínio diferentes.

2.3.1.1 Análise das Ontologias de Computação Forense

Os modelos de ontologia e *framework*, sobre o tema de estudo apresentado nesta seção, refletem a abrangência da dimensão e abordagens possíveis no campo da Computação Forense. Enquanto os dois primeiros modelos apresentados, o de Brinson, Robinson e Rogers (2006) e o de Park, Cho e Kwon (2009), buscam representar de forma mais sistêmica o campo dos crimes digitais, inclusive os seus processos, o modelo de *framework* de Alzaabi, Jones e Martin (2013) aborda os conceitos e conhecimentos mais específicos referente aos dispositivos digitais, cujas instâncias da ontologia são extraídas diretamente dos metadados das informações armazenadas no artefato digital. Cabe salientar, que esse último modelo é extremamente capilarizado, pois cada tipo de artefato digital agregado no *framework* (dispositivos, arquivos, logs, entre outros) terá uma ontologia de domínio própria.

A multidisciplinaridade do campo da Computação Forense está refletida nos três distintos trabalhos, dos autores supracitados, relacionados com ontologias sobre o tema apresentados nesta seção. Na Figura 2.10 tem-se um quadrante com quatro conceitos: Crimes Digitais, Exames, Processo e Evidência, os quais foram ressaltados para representar o direcionamento adotado por esses autores. Buscou-se representar comparativamente no quadrante, também, o tamanho da ontologia por meio das dimensões dos balões, contidos no domínio particular dos: Exames, Crime Digital, Processo Pericial e da Evidência. O tamanho do balão está relacionado

a quantidade de termos empregados nas ontologias e o posicionamento é direcionado pela correlação dos termos das ontologias com os quatro conceitos que definem o quadrante.

O posicionamento da ontologia proposta, quando analisado em relação aos modelos formulados pelos autores, possui uma menor dimensão e encontra-se na parte central do quadrante com uma relevância maior para o processo da computação, destacando-se as atividades de solicitação dos exames. Observa-se no trabalho de Alzaabi, Jones e Martin (2013) com foco na evidência digital e na automação dos seus exames. No trabalho de (BRINSON et al., 2006) observa-se o foco na caracterização dos exames relacionados ao crime digital para auxiliar na definição formação profissional dos atores envolvidos. Já (PARK; CHO; KWON, 2009) direciona seu trabalho para construir bases conhecimento para análise de dados de crimes digitais.

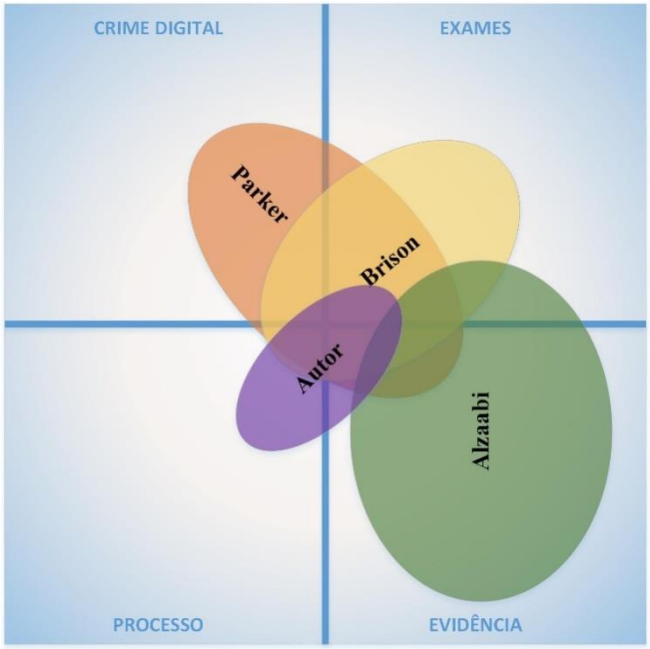


Figura 2.10 - Quadrante comparativo de abordagens das ontologias
Fonte: Elaboração própria

2.3.2 Ontologia BPMNO

As ontologias sobre o tema de estudo apresentada até então, nesta seção, abordam o domínio da Computação Forense sem aprofundar no domínio dos seus processos relacionados. Considerando que deseja-se abordar o processo de Computação Forense, cabe estudar ontologia, que trata das características intrínsecas dos componentes que definem o processo,

em especial o gerenciamento do processo de negócio, do inglês, *Business Process Modeling* (BPM).

BPMNO[6] é uma ontologia BPMN que provê a formalização da parte estrutural dos BPDs, isto é, quais são os elementos básicos do BPD e como eles são ou podem ser conectados. A BPMNO utiliza a *Business Process Model and Notation* (BPMN) versão 2[7]

BPMN é uma linguagem para especificação gráfica de diagramas de processo de negócio *Business Process Diagrams* (BPD). Critérios para a correta/incorreta anotação são afirmações que preenchem a semântica do BPMN e a semântica da ontologia de domínio (FRANCESCO MARINO et al., 2008). A proposta dos autores é codificar todas as informações sobre anotação semântica de processos em uma base de conhecimento, denominada *Business Processes Knowledge Base* (BPKB), apresentada na Figura 2.11.

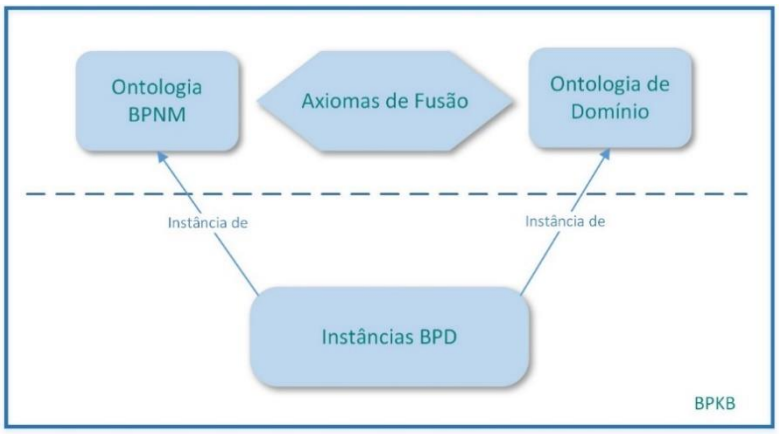


Figura 2.11 - Business Processes Knowledge Base (BPKB)
Fonte: (FRANCESCO MARINO et al., 2008)

O *framework* do BPKD é composto dos seguintes elementos apresentados :

- Ontologia BPMN formaliza a estrutura do BPD. Essa ontologia é a formalização do padrão BPMN e consiste em um conjunto de axiomas que descrevem os elementos do BPMN e o modo que eles são combinados para a construção do BPDs;
- Ontologia de Domínio *Bussiness Domain Ontology* (BDO). A ontologia desenvolvida ou existente que descreve o negócio específico de interesse;

[6]Disponível para download em: https://dkm.fbk.eu/bpmn-ontology
[7]Disponível para download em: http://www.omg.org/spec/BPMN/2.0

- Axiomas de fusão é a correspondência entre a Ontologia de Domínio e a ontologia BPMN. Formalizam o critério correta/incorreta anotação semântica;
- Instâncias BPD contém a descrição de um conjunto de BPDs em termos da ontologia BPMN/Domínio. Cada elemento do processo é representado como um indivíduo da classe.

Utilizando x para descrever um conceito na BPMNO e y para descrever um conceito na BDO, pode-se definir as quatro restrições mostradas no Quadro 2.5, formalizadas como Axiomas em *Description Logic* – DL, que representa critério de correção da anotação semântica.

Quadro 2.5 - Exemplos de restrições do BPKB

Restrição	**Significado intuitivo**	**Axioma DL[8]**
$x \overset{AB}{\rightarrow} y$	Um elemento BPMN do tipo x pode ter sua notação com um conceito equivalente ou mais específico que y	$x \subseteq$ y
$x \overset{nAB}{\longrightarrow} y$	Um elemento BPMN do tipo x não pode ter sua notação com um conceito equivalente ou mais específico que y	$x \subseteq \neg$y
$y \overset{A}{\rightarrow} x$	Qualquer conceito equivalente ou mais específico que y pode ser usado para denotar os elementos BPMN do tipo x	y $\subseteq x$
$y \overset{nA}{\rightarrow} x$	Qualquer conceito equivalente ou mais específico que y não pode ser usado para denotar os elementos BPMN do tipo x	y $\subseteq \neg x$

Fonte: (FRANCESCO MARINO et al., 2008)

Utilizando um raciocinador sobre BPKB, podem-se realizar pesquisas nas instâncias BPD podendo serem formuladas envolvendo a Ontologia do Domínio, a Ontologia BPMN ou ambas. Pode-se ainda realizar a verificação se a rotulagem semântica satisfaz as restrições especificadas utilizando os Axiomas de Fusão. Ainda será possível utilizar axiomas de fusão para sugerir os elementos do processo durante a sua anotação.

[8] Apesar do significado de $x \overset{nAB}{\longrightarrow} y$ e $y \overset{nA}{\rightarrow} x$ coincidirem, os dois casos podem ser usados; dependendo do caso a ser modelado, um pode ter um resultado mais intuitivo do que o outro.

3 ONTOLOGIA PROPOSTA

O escopo definido para o modelo ontológico tem como objetivo principal a definição de um vocabulário comum entre os atores: autoridade que solicita o exame e o perito dentro do exame pericial em Computação Forense. Esse vocabulário deve servir também como referência para guiar o perito na seleção das técnicas e ferramentas aplicáveis nos exames. Para atender esse objetivo, é necessária a coleta de dados dos exames periciais, das solicitações de perícia, dos dispositivos de *hardware* das evidências obtidas, dos sistemas computacionais, das ferramentas periciais, entre outros dados do processo pericial. Neste capítulo são descritos os passos para a construção da ontologia proposta, guiados pela metodologia adotada.

3.1 METODOLOGIA

A metodologia para desenvolvimento de ontologia que serviu de referência para o desenvolvimento deste trabalho foi a *METHONTOLOGY* (LÓPEZ, 1999), que contempla um conjunto de estágios de desenvolvimento, um ciclo de vida baseado em evolução de protótipos e técnicas para realizar atividades de planejamento, desenvolvimento e suporte, além de apresentar um método bem estruturado para a representação e detalhamento da ontologia. Buscou-se produzir os artefatos mais importantes da *METHONTOLOGY*, os quais são apresentados nesta Seção.

Complementando a metodologia adotada, acrescentou-se nesta Seção um tópico preliminar sobre as “Questões de Competência” (NOY; MCGUINNESS, 2001), que serviram como guia para o refinamento do escopo da ontologia, o qual foi inserido na fase de especificação da ontologia prevista pela *METHONTOLOGY*. No Quadro 3.1 são apresentadas algumas das questões de competência sobre seu respectivo conceito.

Quadro 3.1 - Questões de Competência

CONCEITO	NÚMERO	QUESTÕES
Dado uma **Evidência** coletada de um **Caso Criminal** pergunta-se:	QC1	Quais **Técnicas Forenses** podem ser aplicadas nesta **Evidência**?
	QC2	Quais **Crimes** envolvidos com um determinado tipo de **Evidência**?
	QC3	Existe incompatibilidade nos **Quesitos da Solicitação** propostos?
Dado um **Quesito da Solicitação** de exame, pergunta-se:	QC4	Quais **Quesitos da Solicitação** já foram respondidos diante de uma determinada **Evidência**?
	QC5	Quais **Quesitos da Solicitação** podem ser formulados para o exame de uma **Evidência** envolvida em um determinado **Crime**?
Dada uma **Ferramenta Forense**, pergunta-se:	QC6	A **Ferramenta Forense** aplica-se a determinada **Técnica Forense**?
	QC7	A **Ferramenta Forense** tem sido utilizada em uma determinada **Evidência**?

Fonte: Elaboração própria

Inicialmente foi feito o planejamento de desenvolvimento da ontologia resumido pelos itens descritos no Quadro 3.2, onde são descritos o modelo de gerenciamento do projeto, cronograma, recursos necessários para construção da ontologia e o local de realização das atividades. Os itens de planejamento serviram também para o processo de negociação para obtenção dos recursos necessários, inclusive com as devidas autorizações de acesso a informações.

Quadro 3.2 - Itens de Planejamento - *METHONTOLOGY*

ITEM DE PLANEJAMENTO	DESCRIÇÃO
Modelo de gerenciamento do projeto contemplando: escalonamento, controle e garantia de qualidade	Utilização da ferramenta de gerenciamento de projetos Estrutura Analítica de Projeto (EAP), que propõe uma decomposição hierárquica orientada às entregas do trabalho a ser executado pela equipe para atingir os objetivos do projeto e criar as entregas requisitadas. De forma resumida a EAP deste projeto, que também inclui a construção da ontologia está descrita na Figura 3.1.
Cronograma	Especificação: março a maio 2016 Conceitualização: maio a agosto 2016 Formalização: junho a setembro 2016 Implementação: setembro a outubro 2016 Documentação: outubro a novembro 2016
Recursos	Especialista no Domínio: Egberto Lemos Solicitações e Laudos de Exames periciais de 2015-2016 do Departamento de Polícia Técnica da Bahia (DPT-BA)
Local de realização das atividades	Laboratório da Coordenação de Computação Forense do Departamento de Polícia Técnica da Bahia

Fonte: Elaboração própria

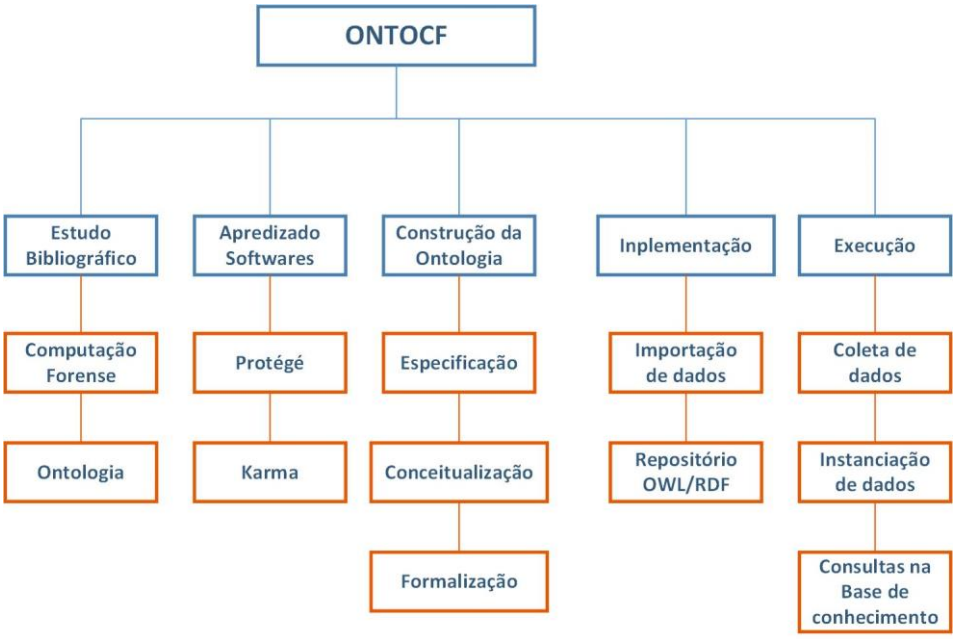


Figura 3.1 - Estrutura Analítica do Projeto

A parte considerada mais importante para este trabalho, na metodologia *METHONTOLOGY*, se refere aos estágios de especificação e conceitualização, onde são descritos os artefatos que devem ser gerados no seu desenvolvimento. Nessa fase os artefatos não são codificados em lógica formal e sim em uma representação intermediária que descreve a ontologia.

3.2 ESPECIFICAÇÃO DA ONTOLOGIA

Considerando as ontologias apresentadas na seção 2, cabe ressaltar que a ontologia proposta neste trabalho é voltada para o processo da Computação Forense ilustrado na Figura 2.4, desde a composição dos quesitos da solicitação de exames, passando pela escolha de uma técnica e ferramenta forense, até a apresentação dos resultados. Uma ontologia que abarca os principais conceitos da Computação Forense tendo como foco as evidências digitais coletadas em um contexto de crime e seus respectivos exames decorrentes de um pedido de perícia embasado por questões.

3.2.1 Propósito da ontologia

No campo da Computação Forense, os crimes demandam o exame pericial em uma evidência digital. As atividades que decorrem dessa afirmação definem os principais conceitos e relações da ontologia do processo dos exames periciais. A principal ideia de desenvolver uma ontologia para Computação Forense, aqui neste trabalho denominada Ontocf, é construir uma base de conhecimento para apoio a decisão dos atores. Para isso, os conceitos, relações e demais componentes da ontologia foram especificados de forma a permitir:

1. padronização de procedimentos;
2. melhoria da comunicação entre os atores do processo pericial, por meio de um vocabulário comum.

Além desses dois propósitos, essa ontologia tem o objetivo de ser empregada em outros cenários de uso, conforme apresentado no Quadro 3.3, que detalha o propósito da ontologia Ontocf.

Quadro 3.3 - Propósito da ontologia

PROPÓSITO	**DESCRIÇÃO**
Intenção de seu uso	A intenção de uso da ontologia Ontocf é aplicá-la no desenvolvimento de uma base de conhecimento sobre o tema Computação Forense, que possa ser utilizada para guiar tanto o solicitante quanto o perito, desde a formulação dos quesitos investigativos sobre uma determinada evidência, passando pela seleção das técnicas e ferramentas aplicáveis, até pela pesquisa de procedimentos adotados em casos similares.
Possíveis cenários de uso	• Busca de informações, a partir das solicitações de exames, sobre metodologias, técnicas e ferramentas empregadas aplicáveis, definindo assim um conjunto de procedimentos periciais. • Definição de vocabulário comum entre os atores do processo pericial. • Guiar autoridades requisitantes na decisão do que será questionado dentro dos procedimentos empregados na perícia de Computação Forense. • Guiar o perito na seleção das técnicas e ferramentas aplicáveis nos exames • Direcionar o aprendizado no âmbito da computação baseado em casos examinados.
Usuários finais da ontologia	Peritos em Computação Forense, autoridades requisitantes, investigadores, policiais e especialista em crimes digitais (*cyber crimes*)

Fonte: adaptado de Fernández-López, Gomez-Perez e Juristo (1997)

A Figura 3.2 representa o domínio do conhecimento a ser modelado pela ontologia proposta. Para tanto, as três principais áreas do conhecimento relacionadas com a Computação Forense e que fornecem informações para especificação dos conceitos e relações do modelo são:

1. Exame Pericial – Técnicas, métodos, ferramentas forenses, dados da solicitação;
2. Crimes Digitais – Tipos de crimes, caso criminal;
3. Evidência Digital – Tipos de evidências, suporte da evidência, caracterização da prova material.

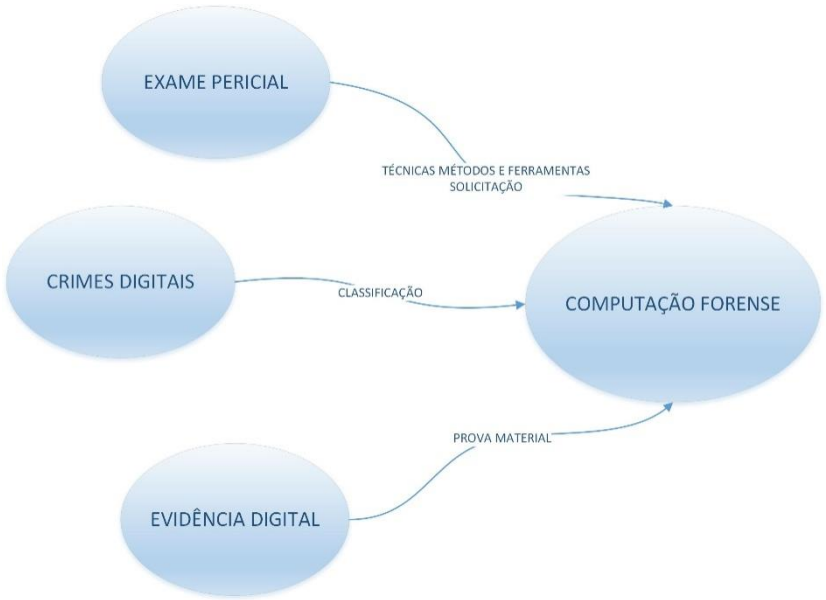


Figura 3.2 - Domínio do conhecimento a ser modelado pela ontologia
Fonte: Elaboração própria

Considerando os questionamentos básicos sobre o Exame Pericial apresentados na introdução, os quais são comumente apresentados nos exames periciais, a ontologia proposta pretende expandir o escopo das questões investigativas e permitir a recuperação da informação em base de conhecimento por meio de novos questionamentos:

- Quais quesitos (linha de investigação) podem ser formulados pelas autoridades solicitantes dado um material e crime investigado;

- Os quesitos formulados já foram utilizados para os materiais apresentados, considerando as possibilidades de técnicas de exames periciais?
- Diante dos quesitos de solicitação, qual o procedimento mais adequado para o exame de determinado artefato digital?
- Diante das características das evidências digitais encontradas, qual a metodologias que podem ser empregadas?
- Tais evidências são semelhantes às utilizadas em outras práticas criminosas que foram objeto de perícia?
- Quais técnicas ou ferramentas são mais utilizadas por um determinado perito?

As informações que alimentam a base de conhecimento, mais especificamente o conjunto das informações das instâncias da ontologia, são extraídas dos documentos periciais: "Solicitação de Exame Pericial" e "Laudo Pericial", encontrados no laboratório de Computação Forense. O diagrama da Figura 3.4 apresenta os principais conceitos e relações definidos nas fases de especificação e conceitualização da *METHONTOLOGY*. Do documento "Solicitação de Exame Pericial", são extraídas as informações pertencentes às seguintes classes: "Caso Criminal", "Ator", "Crime", "Solicitação de Exame", "Quesito de solicitação exame". Já no documento "Laudo Pericial", extraem-se as classes: "Evidência" e seus tipos, "Técnica forense" e "Ferramenta Forense".

Conforme classificação de Uschold e Gruninger (1996), a ontologia proposta neste trabalho é uma ontologia de domínio, onde serão realizadas conceituações do domínio particular do processo da Computação Forense. O grau de formalidade da ontologia é Semi-formal, quando é expressa em uma linguagem artificial, definida formalmente.

3.2.2 Escopo da ontologia

O escopo da ontologia é definido a partir das seguintes premissas:

- As informações das instâncias da ontologia (A-Box) serão obtidas pelos documentos de "solicitação de exames" e "Laudo Pericial";

- O nível de detalhamento dos conceitos (classes) representados na ontologia e suas relações será definido pelo conteúdo das informações geradas no processo entre "solicitante" – "perito";
- As informações entre os atores ("solicitante" – "perito") são as obtidas nos processos pericial do Departamento de Polícia Técnica da Bahia.

O Quadro 3.4 define o escopo contendo as fontes dos conjuntos de termos que serão representados, suas características e granularidade.

Quadro 3.4 – Escopo da Ontocf

ESCOPO	CARACTERÍSTICAS E GRANULARIDADE	FONTE DE CONHECIMENTO
Processo de Computação Forense	Principais conceitos e relações	Literatura sobre Computação Forense Procedimentos Operacionais
Técnicas e métodos de exames periciais	Técnicas e métodos comumente utilizados pelos profissionais da área de Computação Forense.	Literatura sobre Computação Forense
Lista de ferramentas forenses proprietárias	Ao menos as principais ferramentas utilizadas nas organizações periciais oficiais	Fórum dos Peritos Criminais no Brasil. Ferramentas utilizadas no local da pesquisa
Lista de ferramentas computacionais *opensource*	Ao menos as principais ferramentas utilizadas nas organizações periciais oficiais	Pesquisa nos fóruns de peritos forenses Ferramentas utilizadas pelo autor
Tipos de crimes e crimes digitais	Classificação dos crimes, seus tipos, inclusive crimes digitais	CPP, Livros sobre Computação Forense *Scientific Working Group on Digital Evidence* (SWGDE)
Solicitação de exames periciais para crimes digitais	Solicitações de crimes digitais contendo os quesitos.	Guias de exames pericias envidas para a CCF-DPT-BA. Manual de Orientação de Quesitos da Perícia Criminal.
Relação dos tipos de evidências digitais	Classificação dos principais artefatos digitais que compõe a evidência digital.	Scientific Working Group on Digital Evidence (SWGDE) Ontologias existentes sobre o tema de estudo

Fonte: adaptado de Fernández-López, Gomez-Perez e Juristo (1997)

3.3 CONCEITUALIZAÇÃO DA ONTOLOGIA

O estágio de conceitualização estabelece um modelo conceitual que descreve o problema e a solução nos termos do vocabulário do domínio, que é representado por um conjunto de entregáveis, principalmente tabelas e diagramas. É o momento de se construir os termos, ou seja, identificar conceitos e relações no documento gerado no passo anterior. Nesse momento,

os substantivos encontrados serão candidatos aos conceitos e os verbos serão candidatos às relações. Para atender os objetivos deste trabalho, este estágio foi resumido na definição dos conceitos e suas relações, por meio da produção dos seguintes artefatos de conceitualização:

- Árvore (grafo) de classificação de conceitos – Apêndice A;
- Glossário de Termos Completos – GT – Apêndice B;
- Dicionário de dados – Apêndice C;
- Tabela de fórmulas e regras – Apêndice D.

N grafo de classificação de conceitos (Apêndice A) são mapeadas as relações entre os conceitos da Ontocf, conceituadas no Glossário de Termos Completos (Apêndice B). O Dicionário de dados (Apêndice C) descreve as propriedades de dados, enquanto a tabela de regras (Apêndice D), contém as restrições da Ontocf.

No diagrama da Figura 3.3 estão relacionadas as principais classes da ontologia Ontocf. As relações taxionômicas estão representadas nas setas na cor lilás, indicando que uma classe é subclasse de outra classe. Diferenciou-se também as relações do tipo "tem" (has-a), com o objetivo de ressaltar as relações mais específicas entre os conceitos, representadas pela linha vermelha tracejada.

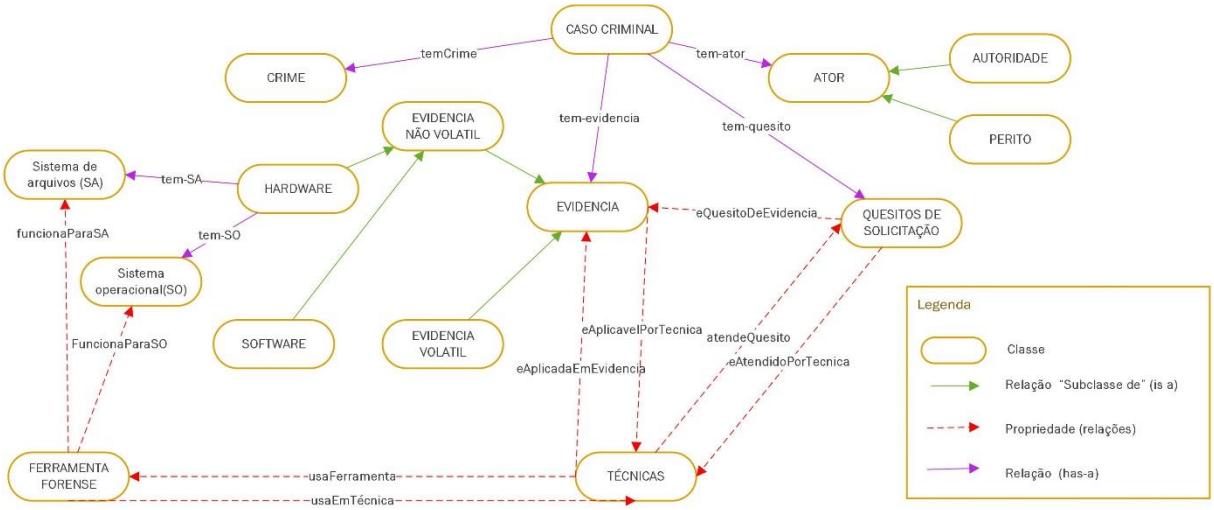


Figura 3.3 - Principais Classes e relações da Ontocf

Neste trabalho, busca-se relacionar o conceito de crime, ou seja, relacionar a ação do crime tipificado que produz a necessidade de um exame pericial, e consequentemente a quesitação da

autoridade sobre determinada evidência relacionada a esse crime. Diante das diversas formas de definição dos crimes e da sua tipificação encontrada na literatura, optou-se em defini-los de forma semelhante a empregada no Manual de Orientação de Quesitos da Perícia Criminal (BRASIL, 2012), no qual o crime é descrito como um contexto criminal sem seguir um formalismo jurídico. Por exemplo, o manual cita os seguintes crimes: "Sítio de internet com conteúdo ilícito". Sendo assim identificou-se um conjunto de Tipos Penais, a partir da coleta feita nos 50 casos criminais remetidos a Coordenação de Computação Forense entre os anos de 2015 e 2016, os quais formaram o conjunto dos dados desta pesquisa. Em seguida, agrupou-se esses Tipos Penais em os crimes relacionados, conforme lista a seguir:

- Crimes contra a pessoa;
- Crimes contra a honra;
- Crimes contra o patrimônio;
- Crimes contra a administração pública;
- Crimes contra a dignidade sexual;
- Crimes relacionados a drogas;
- Crimes econômicos;
- Crimes informáticos.

3.4 FORMALIZAÇÃO DA ONTOLOGIA

A Formalização da ontologia corresponde à especificação da ontologia em uma linguagem. A linguagem formal utilizada na implementação da Ontocf é a *Web* Ontology Language (OWL), mais especificamente OWL-DL 2.

Essa etapa é executada utilizando a ferramenta Protégé durante o estágio de implementação da Ontocf apresentada no capítulo 4.

4 IMPLEMENTAÇÃO

Neste capítulo são apresentados os resultados da utilização da base de conhecimento, como também descritos os passos de implementação da solução proposta, de acordo com os artefatos gerados nas fases de implementação e conceitualização da ontologia proposta.

4.1 REQUISITOS DE IMPLEMENTAÇÃO

A implementação da solução técnica da base de conhecimento foi orientada pelos seguintes requisitos:

- Definir uma base de dados de domínio da Computação Forense baseado em ontologias, fazendo uso da linguagem OWL, adotada como padrão pela W3C;
- Processar consultas na linguagem SPARQL;
- Gerar como resultado informações obtidas no formato URI como respostas às consultas.

Requisitos Funcionais:

- permitir cadastrar ontologias no formato OWL;
- disponibilizar uma interface para entrada de consultas em SPARQL.

Premissas:

- ter uma base de dados baseada em ontologias;
- ser implementada em *frameworks web*.

4.2 FERRAMENTAS DE IMPLEMENTAÇÃO

A seguir, são apresentadas as duas principais ferramentas utilizadas na implementação da Ontocf: Protégé e Karma.

A ferramenta Protégé é um *software* para desenvolvimento e edição de ontologia, desenvolvida pela Universidade de Stanford. Tem sido usada por especialistas em domínios como a medicina e a fabricação, para a modelagem de domínio e para a construção de sistemas de base de conhecimento (GASEVIC; DJURIC; DEVEDZIC, 2006). O Protégé facilita a

definição de conceitos (classes) em uma ontologia, propriedades, taxonomias e várias restrições, bem como instâncias de classe, tudo isso por meio de uma interface gráfica (GUI).

Protégé suporta várias linguagens de representação de ontológia, incluindo OWL e RDF(S). Com Protégé também são facilitadas algumas formas de raciocínio sobre ontologias desenvolvidas, por exemplo, uma vez que a OWL é baseada na *description logics* (DL), inferências como satisfatibilidade e testes de subsunção (GASEVIC; DJURIC; DEVEDZIC, 2006). Neste trabalho, utilizou-se o recurso raciocinador Hermit, por ser um plug-in incorporado ao Protégé que apresenta um comportamento estável na sua execução.

Karma[9]é uma ferramenta *open-source* para integração de dados, desenvolvida na Universityof Southern California (USC). Ela possibilita integração a partir de uma variedade de fontes de dados, incluindo bancos de dados relacionais, planilhas, arquivos de texto delimitados e arquivos em XML, JSON ou Web APIS.

Karma possibilita funcionalidades ligada a troca de dados (*data exchange*), com ela é possível reestruturar um dado de uma fonte de dados para o esquema de uma ontologia (HARTH et al., 2013). Ou seja, Karma apoia a transformação dos dados das fontes em RDF nos termos de uma dada ontologia. Esse processo compreende duas fases:

- A primeira fase (em tempo de modelagem) é responsável por mapear o esquema das fontes de dados para uma ontologia, a saída dessa fase é um arquivo de mapeamento;
- Segunda fase (em tempo de execução), o arquivo de mapeamento criado na fase anterior é usado para converter os dados das fontes de dados para o vocabulário da ontologia (RDF).

Alternativamente, esses mapeamentos podem ser interpretados dinamicamente por um mediador, que é responsável por receber consultas nas linguagens da ontologia (SPARQL *queries*) e recuperar os dados diretamente das fontes originais.

Como no cenário deste trabalho, não faz-se necessário uma integração de dados em tempo de execução, os RDFs gerados a partir da base serão carregados em um repositório de triplas RDF (*triple store*) que será disponível como SPARQL *endpoint* para eventuais consultas na linguagem da ontologia.

[9]http://usc-isi-i2.github.io/karma/

4.3 ARQUITETURA DA SOLUÇÃO

Conforme estabelecido, na Ontocf, as demais instâncias da base de conhecimento são extraídas dos documentos periciais: “Solicitação (ou guia) de Exame Pericial” e “Laudo Pericial”, que definem um Caso Criminal. Para viabilizar a coleta de dados desses documentos foi criada uma planilha Microsoft Excel com as abas apresentadas no Quadro 4.1 cujos dados são gravados no formato de arquivos *Comma Separated Values* (CSV).

Mais especificamente, o modelo de análise define que a principal fonte de informação para instanciação da Ontocf, que serão armazenadas no repositório RDF, são obtidas dos Casos Criminais que chegam na Coordenação de Computação Forense DPT-BA.

Quadro 4.1 - Tabelas da planilha de coleta de dados

Tabela	**Descrição**	**Fonte das informações**
Perícias	Tabela que representa em cada linha uma instância do Exame Pericial realizado em uma Evidência para responder a um Quesito de Solicitação, composta pelos campos das Classes: Caso Criminal, Quesito de Solicitação, data da solicitação, data do exame, Perito, Solicitante, Crime, Evidência, Sistema Operacional, Sistema de Arquivos, Tipo de Evidência, Técnica Forense, Ferramenta Forense.	Guia de solicitação de exame pericial e Laudo Pericial
Caso Criminal	Número utilizado pelo órgão oficial de perícia para identificar o conjunto: Solicitação de Exames – Laudo pericial, que integra uma investigação criminal	Guia de solicitação de exame pericial
Crime	Descrição do contexto criminal identificados pela leitura da guia de solicitação de exame pericial	Guia de solicitação de exame pericial
Tipo Crime	Enquadramento do “Crime” da lista anterior dentro de uma classificação adotada neste trabalho.	Guia de solicitação de exame pericial
Quesito de Solicitação	São os quesitos encontrados na guia de solicitação de exames ajustados de acordo com a terminologia técnica adotada na literatura e nos manuais operacionais.	Guia de solicitação de exame pericial
Evidência	São os materiais remetidos ou coletados para exame pericial de Computação Forense: equipamentos microprocessador, dispositivos, mídias digitas, artefatos digitais em memória, entre outros	Laudo Pericial
Tipo de Evidência	Classificação Taxionômica das evidências.	Laudo Pericial
Técnica Forense	Procedimentos técnicos utilizados nos exames definidos por uma metodologia.	Laudo Pericial
Ferramenta Forense	Ferramenta de *software* utilizada nas ações computacionais para realizar técnicas forenses.	Laudo Pericial
Perito	Profissional que realiza o Exame Pericial	Laudo Pericial
Solicitante	Autoridade que remete o pedido dos Exames Periciais contendo os quesitos de solicitação.	Guia de solicitação de exame pericial
Sistema de Arquivos	É a forma de organização de dados em algum meio de armazenamento de dados da evidência	Laudo Pericial

Sistema Operacional	É uma coleção de programas que inicializam o hardware do computador ou dispositivo microprocessado da evidência em exame.	Laudo Pericial

Fonte: Elaboração própria

Foram feitas as seguintes convenções para definir a coleta das informações:

- Cada linha define uma ação de Exame Pericial, definida como Perícia;
- Um Caso Criminal pode ter mais de uma Perícia;
- Uma Perícia possui uma Evidência;
- Cada linha contém uma Evidência e um Quesito de Solicitação correspondente;

Para auxiliar o entendimento e documentação do processo de construção do banco de dados da base de conhecimento, optou-se, como forma de representação do modelo de dados das fontes de informação, a Modelagem Dimensional por meio do Modelo Estrela (*Star Schema*) (KIMBALL; ROSS, 2011). Nesse modelo, uma tabela é usada, para cada dimensão, para armazenar dados sobre a dimensão (tabela de dimensão). A tabela de fatos armazena as instâncias com valores das dimensões descritivas para cada instância, e valores dos fatos, ou medidas, para aquela instância.

Figura 4.1 - Modelo dimensional do banco de dados
Fonte: Elaboração própria

No Modelo Estrela, todas as tabelas relacionam-se diretamente com a tabela de fatos. Seguindo esse modelo, a tabela de fatos "dominante" no centro do esquema, que no modelo de dados implementado é a tabela “Perícia”, e as tabelas de dimensões nas extremidades, que neste modelo coincide com as Classes da Ontocf, conforme apresentado na Figura 4.1.

A partir desse modelo utilizou-se a técnica de *Extract Transform Load* (ETL) para o trabalho de importação de dados arquivo *Comma Separated Values* (CSV) e posterior transformação de dados e para carga no repositório RDF da Ontocf.

A arquitetura de *software* da solução para coleta, armazenamento e recuperação da informação da base de conhecimento suportada pela ontologia Ontocf, está representada no diagrama da Figura 4.2. O diagrama ressalta o processo ETL da solução, onde os dados disponíveis nas fontes passam a ser definidos em RDF, seguindo a semântica da ontologia, construída no Protégé. Para isso, arquivos de mapeamentos são criados no ambiente do Karma, responsáveis por especificar as correspondências entre a fonte de dados e a ontologia. Esses mapeamentos são usados para a criação de triplas RDF e são utilizados para a criação de um repositório /SPARQL *endpoint.*

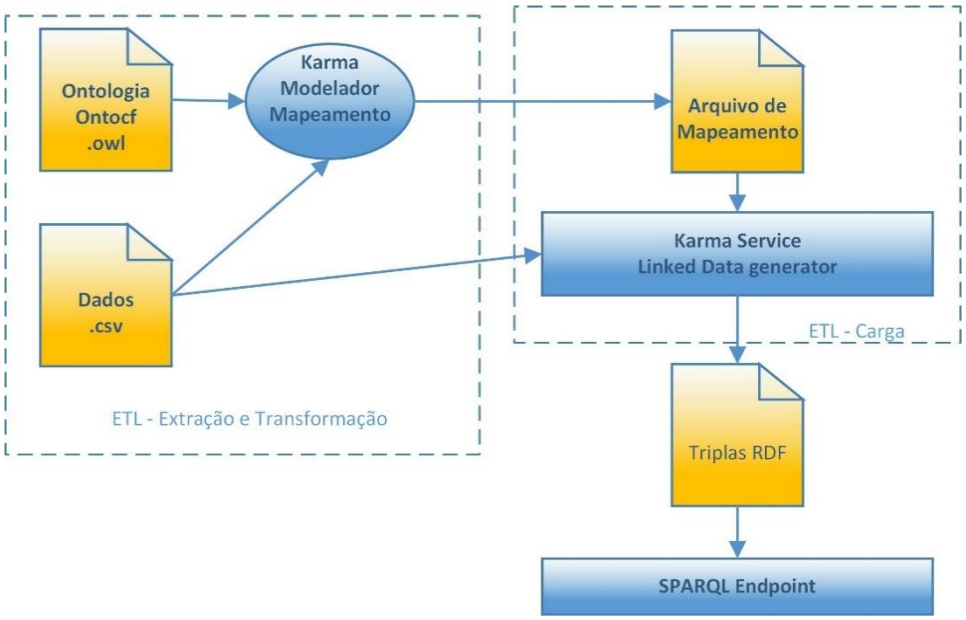


Figura 4.2 - Processo ETL do Repositório RDF da Ontocf
Fonte: adaptado de (HARTH et al., 2013)

As fases de Extração e Transformação consistem na modelagem dos dados presentes nos “arquivos csv” em “triplas rdf”, correspondentes ao vocabulário da ontologia e é realizada com

o apoio da ferramenta Karma. Usando o componente Karma Modelling, os seguintes passos são realizados:

1. A ontologia de domínio é carregada no componente – no caso deste trabalho, a Ontocf é carregada;
2. A fonte de dados a ser mapeada é carregada – no caso deste trabalho, os arquivos CSV correspondente aos laudos periciais;
3. Karma Modelling faz uso da fonte de dados e da ontologia para a criação dos modelos de transformação. Esses modelos realizam a correspondência entre o esquema da fonte de dados e a estrutura da ontologia. Isso é feito de forma semiautomática, enquanto o Karma consegue inferir algumas correspondências, se faz necessário a interferência do usuário para validar as propostas pela ferramenta e/ou adicionar e revisar algumas delas;
4. Finalizada a parte de modelagem, é gerado o arquivo de mapeamento que contém dados hierárquicos vindos da fonte de dados e as características da ontologia ligados pelos construtores de transformação. Esse mapeamento é especificado em KR2RML (SLEPICKA et al., 2015), uma variação da Karma para a linguagem R2RML[10].

A Figura 4.3 mostra o Karma Modeling durante a fase de modelagem. É possível notar que cada coluna do "arquivo csv" é associado a alguma propriedade da Ontocf e que estas, por sua vez, são ligadas a alguma classe da ontologia. É possível definir também as propriedades objeto, representada pelo grafo que liga uma classe a outra na Figura 4.3. Por exemplo, nesse recorte do mapeamento é possível observar a classe "CasoCriminal" cujo o valor da propriedade de dados "idCrime" é o valor vindo da coluna Caso Criminal no arquivo CSV, além disso, a classe "QuesitoSolicitação" está associada a classe "CasoCriminal" através da propriedade objeto "temQuesito".

Após esse processo de modelagem, a segunda fase é realizada com o apoio do componente Karma Service, um serviço HTTP, que receberá como entrada o arquivo de mapeamento criado na fase anterior e uma fonte de dados a ser convertida para o vocabulário da ontologia. Esse processo gera como resultado triplas RDF que representam os dados presente na fonte em termos da ontologia. O Código 2 é um exemplo de resultado gerado, ele mostra as triplas referentes a criação de dois indivíduos, um da classe "CasoCriminal" e um da classe Quesito e como eles se relacionam pela propriedade "temQuesito".

[10] https://www.w3.org/TR/r2rml/

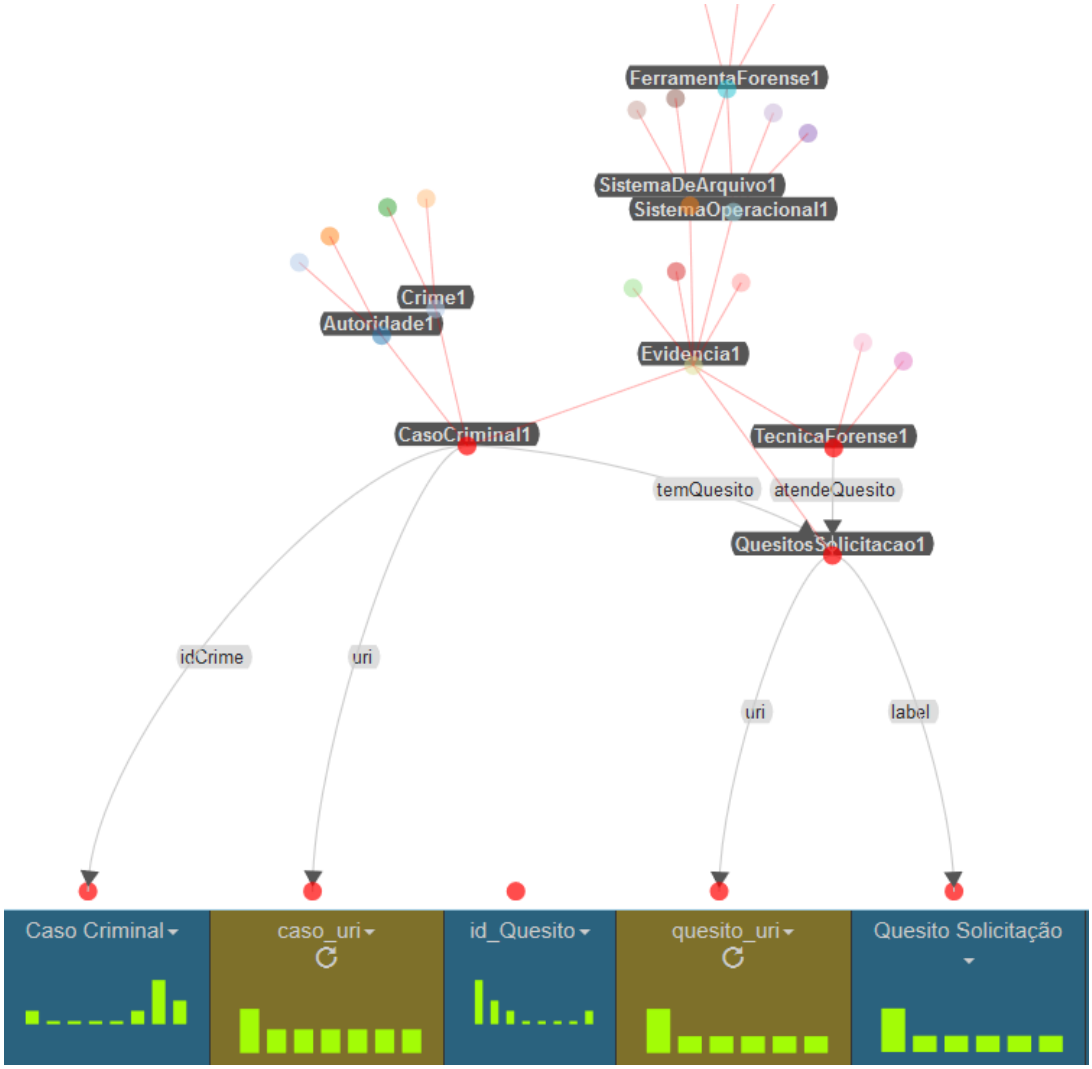


Figura 4.3 - Modelagem utilizando a ferramenta Karma Modeling
Fonte: Elaboração Própria

```
<http://ontologiaforense.esy.es/ontocf.owl#caso/2015001729><http://www.w3.org/1999/02/22-rdf-syntax-ns#type><http://ontologiaforense.esy.es/ontocf.owl#CasoCriminal> .
<http://ontologiaforense.esy.es/ontocf.owl#caso/2015001729><http://ontologiaforense.esy.es/ontocf.owl#idCrime> "2015001729" .
<http://ontologiaforense.esy.es/ontocf.owl#caso/2015001729><http://ontologiaforense.esy.es/ontocf.owl#temQuesito><http://ontologiaforense.esy.es/ontocf.owl#quesito/2> .
<http://ontologiaforense.esy.es/ontocf.owl#quesito/2><http://www.w3.org/1999/02/22-rdf-syntax-ns#type><http://ontologiaforense.esy.es/ontocf.owl#QuesitosSolicitacao> .
```

Código 2 - Exemplo de Código RDF

As triplas RDF geradas a partir dos arquivos CSV são adicionadas a um repositório RDF ou banco de dados de triplas implementado, usando a ferramenta Sesame[11], incorporada ao framework do Karma, que funciona como um *SPARQL endpoint*, podendo ser acessado via HTTP para a realização de consultas SPARQL.

4.4 DOCUMENTAÇÃO

A documentação das etapas do desenvolvimento desta ontologia que foram adotadas da *METHONTOLOGY*, estão inclusas no corpo dos tópicos, seguindo, quando possível, a mesma nomenclatura dessa metodologia. Sendo assim, fora gerados documentos de Especificação de Requisitos, Modelo conceitual, Implementação e Avaliação.

A Ontocf está disponível no endereço: http://ontologiaforense.esy.es/ontocf.owl e sua documentação produzida automaticamente a partir do arquivo ontocf.owl, por meio do *framework* de documentação para OWL denominado Parrot[12], que está disponível no endereço https://goo.gl/YodGgs. Dessa forma, pretende-se que durante o estágio de manutenção da Ontocf sua documentação mantenha-se atualizada.

[11] http://rdf4j.org/
[12]Disponível em http://ontorule-project.eu/parrot/parrot

5 AVALIAÇÃO

Na avaliação, busca-se verificar se a ontologia obtida a partir das etapas de especificação e conceitualização está de acordo com o projetado durante a etapa de definição de escopo, ou seja, deseja-se confirmar se a ontologia pode responder a todas as questões de competências estipuladas, se representa fielmente o domínio, se é facilmente compreendida por todos os atores envolvidos.

Seguindo a metodologia *METHONTOLOGY*, na atividade de suporte Avaliação busca-se atender as seguintes proposições:

a) Verificação – Verificar tecnicamente a consistência da ontologia e dos seus artefatos com respeito ao conjunto de documentos de especificação;

b) Validação – Garantir que a ontologia e seus artefatos correspondam ao sistema que supostamente a represente.

A avaliação da ontologia foi realizada em três etapas de verificação. Na primeira etapa foram realizadas consultas na base de conhecimento que possam atender as Questões de Competência. Na segunda etapa foram realizadas inferências com raciocinadores a partir das restrições. Na terceira e última etapa foi aplicado, para oito especialistas do domínio, um questionário de validação dos principais conceitos da Ontocf.

Implantado o ambiente tecnológico da base de conhecimento, foram feitas coletas de dados em um conjunto de 50 (cinquenta) Casos Criminais compostos, respectivamente pela sua Guia de Solicitação de Exames e pelos Laudos Periciais, produzidos na Coordenação de Computação do Departamento de Polícia Técnica da Bahia.

Cabe salientar que tanto a amostra de 50 Casos Criminais não representa uma amostra probabilística, como, também, a aplicação do questionário abordando apenas os principais conceitos não é suficiente para avaliação por especialistas, entretanto as etapas de verificação propostas destina-se a verificar se os caminhos trilhados na construção da Ontocf são coerentes com seus objetivos.

5.1 CONSULTAS NA BASE DE CONHECIMENTO

Os dados coletados, em um conjunto de 50 (cinquenta) Casos Criminais, alimentaram o sistema de ETL para formação do repositório RDF da Ontocf. Em seguida são apresentados os resultados das consultas SPARQL baseadas nas questões de competência.

As consultas foram realizadas no ambiente do *framework Sesame*, configurado como um *SPARQL endpoint.* A Figura 5.1 mostra uma dessas consultas SPARQL sendo executada dentro do repositório.

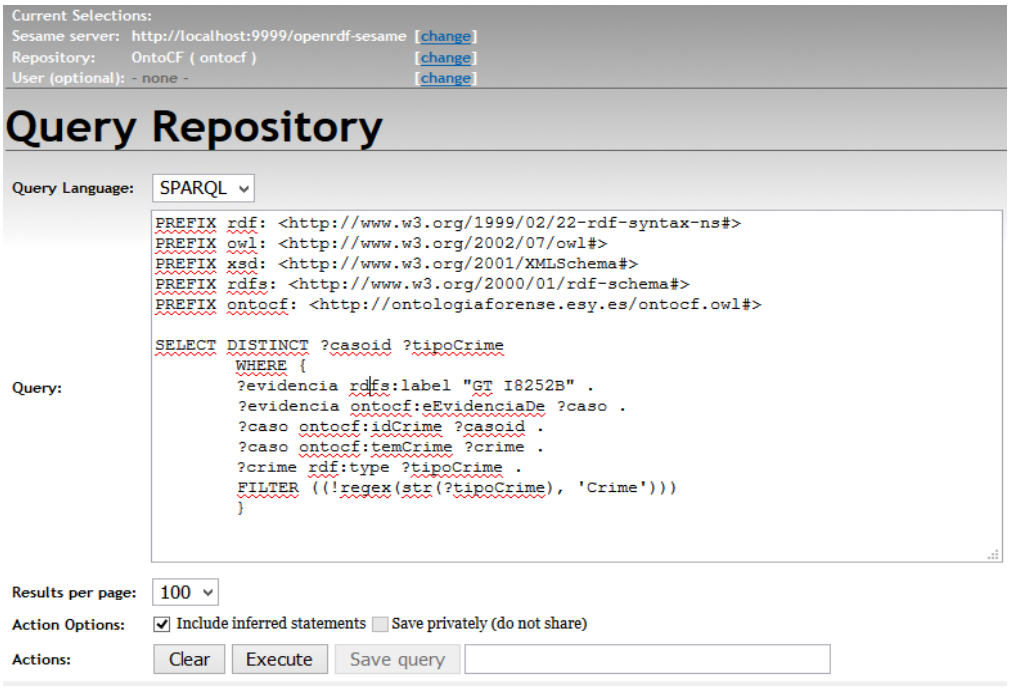


Figura 5.1 - Exemplo de consulta SPARQL no repositório

Nessa consulta, dada uma determinada evidência, busca-se saber quais tipos de crimes estão envolvidos com ela. A Figura 5.2 por sua vez, mostra o resultado dessa consulta executada na Figura 5.1.

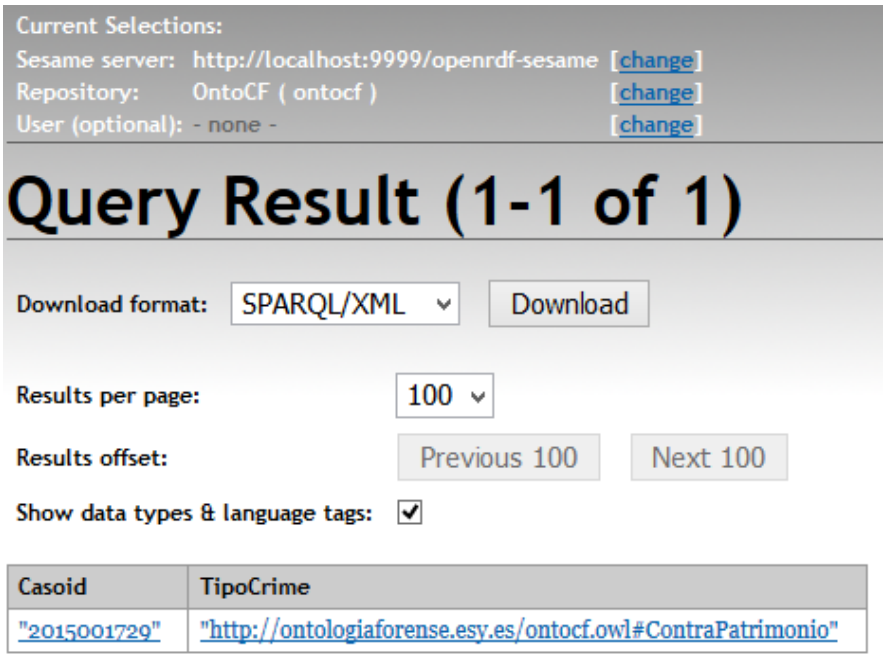


Figura 5.2 - Resultado da consulta SPARQL

O Quadro 5.1 descreve as questões derivadas das questões de competência elencadas durante a fase de especificação da ontologia, com suas respectivas consultas SPARQL. Nessa fase as questões são mais específicas, de forma que podem ser respondidas pela solução técnica implementada.

Quadro 5.1 - Consultas SPARQL para Questões de Competência Derivadas

ITEM	CONSULTA / QUESTÃO. DE COMPETÊNCIA	CÓDIGO SPARQL Obs. O prefixo da primeira consulta repete-se para as demais.
1	Quais Técnicas Forenses podem ser aplicadas nesta Evidência? (QC01)	PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> PREFIX owl: <http://www.w3.org/2002/07/owl#> PREFIX xsd: <http://www.w3.org/2001/XMLSchema#> PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#> PREFIX ontocf: <http://ontologiaforense.esy.es/ontocf.owl#> SELECT DISTINCT ?tipoTecnica WHERE { ?evidenciardfs:label "Samsung S6 SM-G920i" . ?tecnicaontocf:eAplicavelEmEvidencia ?evidencia . ?tecnicardf:type ?tipoTecnica . FILTER ((!regex(str(?tipoTecnica), 'TecnicaForense'))) }ORDER BY ?evidencia
2	Quais Crimes envolvidos com um determinado tipo de Evidência? (QC02)	SELECT DISTINCT ?casoid ?tipoCrime WHERE { ?evidenciardf:typeontocf:Notebook . ?evidencia ontocf:eEvidenciaDe ?caso . ?caso ontocf:idCrime ?casoid . ?caso ontocf:temCrime ?crime .

		?crime rdf:type ?tipoCrime . FILTER ((!regex(str(?tipoCrime), 'Crime'))) }
3	Quais Técnica Forense podem ser aplicadas para uma determinada ferramenta. (QC03)	SELECT DISTINCT ?ferramenta ?tipoTecnica WHERE { ?ferramenta rdfs:label "FTK". ?tecnicaontocf:usaFerramenta ?ferramenta . ?tecnicardf:type ?tipoTecnica . FILTER ((!regex(str(?tipoTecnica), 'TecnicaForense'))) } ORDER BY ?ferramenta

4	Quais Ferramentas Forense podem ser utilizadas em uma determinada Evidência considerando a Técnica X?
	(QC07)
		SELECT DISTINCT ?Ferramenta
		WHERE {
		?evidenciardf:typeontocf:Notebook .
		?tecnicardf:typeontocf:ExtracaoDados .
		?tecnicaontocf:eAplicavelEmEvidencia ?evidencia .
		?tecnicaontocf:usaFerramenta ?Ferramenta .
		
		}

5	Quais os quesitos da Solicitação já foram respondidos para uma determinada tipo de Evidência?
	(QC04)
		SELECT DISTINCT ?quesitoNome
		WHERE {
		?evidencia rdf:type ontocf:Notebook .
		?quesito ontocf:eQuesitoValidoEvidencia ?evidencia .
		?quesito rdfs:label ?quesitoNome .
		
		}
		}

6	Dado um quesito de solicitação, quais as evidências nas quais ele já foi utilizado?	SELECT DISTINCT ?evidenciaNome ?evidenciaTipo
		WHERE {
		
		?quesito rdfs:label 'Busca e extração de imagens pornograficas de criança ou adolescente' .
		?quesito ontocf:eQuesitoValidoEvidencia ?evidencia .
		?evidencia rdfs:label ?evidenciaNome .
		?evidencia rdf:type ?evidenciaTipo .
		FILTER ((!regex(str(?evidenciaTipo), 'Evidencia')))
		
		}

7	Quais quesitos sobre crime “Reprodução de material protegido por propriedade intelectual” já foram empregados para um tipo de evidência (classe da Taxionomia Evidência)	SELECT DISTINCT ?casoID ?quesitoNome ?tipoEvidencia
		WHERE {
		?crime rdfs:label 'Reprodução de material protegido por propriedade intelectual' .
		?casoCriminalontocf:temCrime ?crime .
		?casoCriminalontocf:idCrime ?casoID .
		?casoCriminalontocf:temEvidencia ?evidencia.
		?casoCriminalontocf:temQuesito ?quesito.
		?quesito ontocf:eQuesitoValidoEvidencia ?evidencia .
		?quesito rdfs:label ?quesitoNome .
		?evidencia rdf:type ?tipoEvidencia .
		FILTER ((!regex(str(?tipoEvidencia), 'Evidencia')))
		}

8	Quais ferramentas podem ser empregadas para uma determinada Técnica com Evidência que possui sistema operacional Android	SELECT DISTINCT ?ferramenta
		WHERE {
		?evidenciaontocf:temSO ?so .
		?sordfs:label 'Android' .
		?tecnicardf:typeontocf:ExtracaoDados .
		?tecnicaontocf:eAplicavelEmEvidencia ?evidencia .
		?tecnicaontocf:usaFerramenta ?ferramenta .
		}

Fonte: Elaboração própria

As consultas relacionadas no **Erro! Fonte de referência não encontrada.** 5.1 foram executadas, resultando nas respostas apresentadas a seguir.

As respostas apresentadas nas Figuras 5.3, 5.4, 5.5 e 5.6 são condizentes com o objetivo das questões elaboradas, tanto para atender o perito, quanto ao solicitante da perícia.

Observa-se na consulta 02 (Figura 5.3) que a resposta apresenta três técnicas já realizadas para uma evidência específica. Essa informação é útil para o perito durante o planejamento do seu trabalho.

tipoTecnica
http://ontologiaforense.esy.es/ontocf.owl#Carving
http://ontologiaforense.esy.es/ontocf.owl#AnaliseLog
http://ontologiaforense.esy.es/ontocf.owl#ExtracaoDados

Figura 5.3 - Consulta 2

A consulta 03 (Figura 5.4) traz relevante informação sobre as técnicas que podem ser empregadas com uma determinada ferramenta, sobretudo com o surgimento vertiginoso das ferramentas *open source.* Essa consulta serviria como uma busca no catálogo, podendo incluir na resposta atributos com os detalhes de uso da ferramenta.

ferramenta	tipoTecnica
http://ontologiaforense.esy.es/ontocf.owl#FTK	http://ontologiaforense.esy.es/ontocf.owl#Indexacao
http://ontologiaforense.esy.es/ontocf.owl#FTK	http://ontologiaforense.esy.es/ontocf.owl#Carving
http://ontologiaforense.esy.es/ontocf.owl#FTK	http://ontologiaforense.esy.es/ontocf.owl#AnaliseLog
http://ontologiaforense.esy.es/ontocf.owl#FTK	http://ontologiaforense.esy.es/ontocf.owl#ExtracaoDados

Figura 5.4 - Consulta 3

Já a consulta 05 (Figura 5.5) traz resultados úteis para o solicitante dos exames, pois lista as possibilidades de quesitos empregáveis na investigação.

quesitoNome
identificar arquivos de vídeo
Busca e extração de arquivos de imagens
Busca e extração de arquivos em geral
Busca e extração de imagens pornograficas de criança ou adolescente
Identificação e acesso de arquivos criptografados
Relacionar o histórico de eventos de rede e configurações do equipamento
Identificar softwares maliciosos de controle remoto no dispositivo
Identificar endereços web visitados a partir do dispositivo

Figura 5.5 - Consulta 5

A consulta 07 (Figura 5.6) refina a resposta da consulta 05 inserido o critério tipo de crime, assim obtendo-se uma lista menor de quesitos.

casoID	quesitoNome	tipoEvidencia
2015009305	Identificar arquivos de documentos semelhantes	http://ontologiaforense.esy.es/ontocf.owl#Desktop
2015009305	identificar arquivos de vídeo	http://ontologiaforense.esy.es/ontocf.owl#Notebook
2012005996	Verificar indícios de falsificação de de documentos	http://ontologiaforense.esy.es/ontocf.owl#Pendrive

Figura 5.6 - Consulta 7

5.1.1 Exemplo de Cenário de Uso

Considerando a utilização da base de conhecimento na rotina de trabalho de um laboratório de Computação Forense, elaborou-se um cenário de uso para os dois principais atores do processo: o solicitante e o perito, baseado em um caso criminal real.

Em um caso criminal específico, um delegado de polícia investiga um suspeito de armazenar e distribuir conteúdo pornográfico envolvendo criança e adolescente. O suspeito é especialista em informática, o que leva a crer que ele utilizava técnicas anti-forense como, por exemplo criptografia e acesso a *Deepweb*. Existia também relatos que o investigado acessava a internet em diversos locais da cidade violando a senha do sinal *Wifi*, inclusive para contatar possíveis compradores do material ilícito. No ato da apreensão do computador tipo *notebook*, o suspeito alegou que seu computador foi invadido e acessado remotamente por outra pessoa.

Neste cenário, são definidas duas perspectivas de busca de informações em uma base de conhecimento. Primeiro, o investigador deseja saber quais questões podem ser feitas para a evidência em função do crime relacionado. Na segunda, o perito deseja saber quais técnicas e ferramentas podem ser empregadas para responder o quesito.

Quadro 5.2 - Consultas SPARQL para cenário de uso - Autoridade

ITEM	CONSULTA / QUESTÃO. DE COMPETÊNCIA	CÓDIGO SPARQL

9	Autoridade pergunta, tendo um crime (pornografia infantil) e uma evidência (notebook) quais quesitos já realizados
	
	(QC05)
		SELECT DISTINCT ?casoID ?quesitoNome ?tipoEvidencia
		WHERE {
		?crime rdfs:label 'Armazenento e transmisão de dados de pornografia infantil' .
		?casoCriminalontocf:temCrime ?crime .
		?casoCriminalontocf:idCrime ?casoID .
		?casoCriminalontocf:temEvidencia ?evidencia.
		?evidencia rdf:typeontocf:Notebook .
		?casoCriminalontocf:temQuesito ?quesito.
		?quesito ontocf:eQuesitoValidoEvidencia ?evidencia .
		?quesito rdfs:label ?quesitoNome .
		}

	Quesitos realizados: quesitoNome Busca e extração de imagens pornograficas de criança ou adolescente Identificação e acesso de arquivos criptografados Relacionar o histórico de eventos de rede e configurações do equipamento Identificar softwares maliciosos de controle remoto no dispositivo Identificar endereços web visitados a partir do dispositivo	

Fonte: Elaboração própria

Como resposta (Quadro 5.2), obteve-se cinco quesitos que são pertinentes ao caso investigado, sendo que fica ao critério do investigador escolher qual dos quesitos serão escolhidos para compor o pedido de solicitação. Cabe ressaltar, que esse resultado apenas orienta a formulação de quesitos, e não limita a formulação de quesitos como são apresentados.

Considerando que o perito selecionou o primeiro quesito apresentado na resposta anterior, pergunta-se a base de conhecimento: “Para o quesito apresentado pela autoridade e para a evidência em questão, quais técnicas ou ferramentas”. As respostas são apresentadas no Quadro 5.3.

Quadro 5.3 - Consultas SPARQL para cenário de uso - Perito

ITEM	CONSULTA / QUESTÃO. DE COMPETÊNCIA	CÓDIGO SPARQL
10	O perito pergunta: para o quesito apresentado pela autoridade e para a evidência em questão,	SELECT DISTINCT ?tipoTecnica ?ferramenta WHERE { ?quesito rdfs:label 'Busca e extração de imagens pornograficas de criança ou adolescente' .

	quais técnicas ou ferramentas (QC01)	?evidencia rdf:type ontocf:Notebook . ?quesito ontocf:eQuesitoValidoEvidencia ?evidencia . ?tecnica ontocf:eAplicavelEmEvidencia ?evidencia . ?tecnica rdf:type ?tipoTecnica . ?tecnica ontocf:usaFerramenta ?ferramenta . FILTER ((!regex(str(?tipoTecnica), 'TecnicaForense'))) }

Técnicas com respectivas ferramentas:

tipoTecnica	ferramenta
http://ontologiaforense.esy.es/ontocf.owl#Carving	http://ontologiaforense.esy.es/ontocf.owl#IEF
http://ontologiaforense.esy.es/ontocf.owl#Carving	http://ontologiaforense.esy.es/ontocf.owl#FTK
http://ontologiaforense.esy.es/ontocf.owl#Carving	http://ontologiaforense.esy.es/ontocf.owl#Encase
http://ontologiaforense.esy.es/ontocf.owl#AnaliseLog	http://ontologiaforense.esy.es/ontocf.owl#FTK
http://ontologiaforense.esy.es/ontocf.owl#ExtracaoDados	http://ontologiaforense.esy.es/ontocf.owl#UFED
http://ontologiaforense.esy.es/ontocf.owl#ExtracaoDados	http://ontologiaforense.esy.es/ontocf.owl#Encase
http://ontologiaforense.esy.es/ontocf.owl#ExtracaoDados	http://ontologiaforense.esy.es/ontocf.owl#ManipulaÃ§Ã£odiretadoequipamento
http://ontologiaforense.esy.es/ontocf.owl#ExtracaoDados	http://ontologiaforense.esy.es/ontocf.owl#ferramentaAndroidDebugBridge(ADB)
http://ontologiaforense.esy.es/ontocf.owl#ExtracaoDados	http://ontologiaforense.esy.es/ontocf.owl#IEF
http://ontologiaforense.esy.es/ontocf.owl#ExtracaoDados	http://ontologiaforense.esy.es/ontocf.owl#FTK

Fonte: Elaboração própria

Os resultados das consultas SPARQL foram satisfatórios em responder as “Questões de Competência” demonstrado pela correlação entre o número da consulta e o código adotado para as questões de competência do Quadro 3.1, exceto pela questão QC03: “ Existe incompatibilidade nos Quesitos da Solicitação propostos?”.

A QC03 reitera o princípio de que o mundo da ontologia é aberto. Ela não sabe responder “não”, ele só sabe responder “sim ou não sei”. Ou seja, pode-se afirmar com quem o quesito está relacionado e não que ele não é compatível ou não pode ser utilizado. A não ser, claro, que regras fossem criadas nesse sentido.

Para resolver tais necessidades de informação formulou-se a proposta, para uma futura versão da Ontocf, de criar uma nova relação entre quesito e evidência: “quesitoNaoRecomedado”, baseado nos quesitos não recomendados no Manual de Quesitos da Perícia Criminal (BRASIL, 2012). A partir dessa nova relação a questão de competência seria reformulada para: Quais os quesitos não recomendados para determinada evidência, inserida em um determinado contexto criminal.

5.2 INFERÊNCIAS COM RACIOCINADORES

A inferência com raciocinadores tem o objetivo de verificar se a ontologia é ou não consistente, como também identificar relações não explicitas entre classes.

A Ontocf foi verificada pelo raciocinador Hermit 1.3.8, produzindo inferências com base nas regras SRWL previamente inseridas no ambiente do Protegé, conforme apresentado nos exemplos a seguir.

A Figura 5.7 mostra algumas regras SRWL especificadas através do Protégé, para quatro restrições selecionadas.

Figura 5.7 - Regras SRWL para a Ontocf

As regras SRWL da Figura 5.7 são descritas como:

- Regra 1: Na ocorrência do uso da ferramenta *Volatility* em alguma técnica Forense, a evidência na qual a técnica está sendo aplicado deverá ser uma Evidência Volátil;
- Regra 2: A regra 2 diz respeito a uma restrição em que na ocorrência da aplicação da técnica de Extração de Dados em uma evidência que possua Sistema Operacional Android, a ferramenta utilizada deve ser necessariamente a *Android Debug Bridge* (ADB);
- Regra 3: A regra 3 determina que na ocorrência do uso da ferramenta UFED em alguma técnica forense, a evidência na qual esta técnica está sendo aplicada deverá ser um Dispositivo de Pequena Escala.
- Regra 4: Na ocorrência do uso da ferramenta *Malwarebits* em alguma técnica Forense, a evidência na qual a técnica está sendo aplicado deverá ser um Sofware Malicioso;

A Figura 5.8 apresenta algumas inferências realizadas pelo raciocinador, a partir dessa regra e também a partir de outros axiomas OWL. Na Figura 5.8 verifica-se três quadros referentes a três indivíduos na ontologia, as seguintes declarações e inferências (declarações destacadas com fundo amarelado) podem ser observadas:

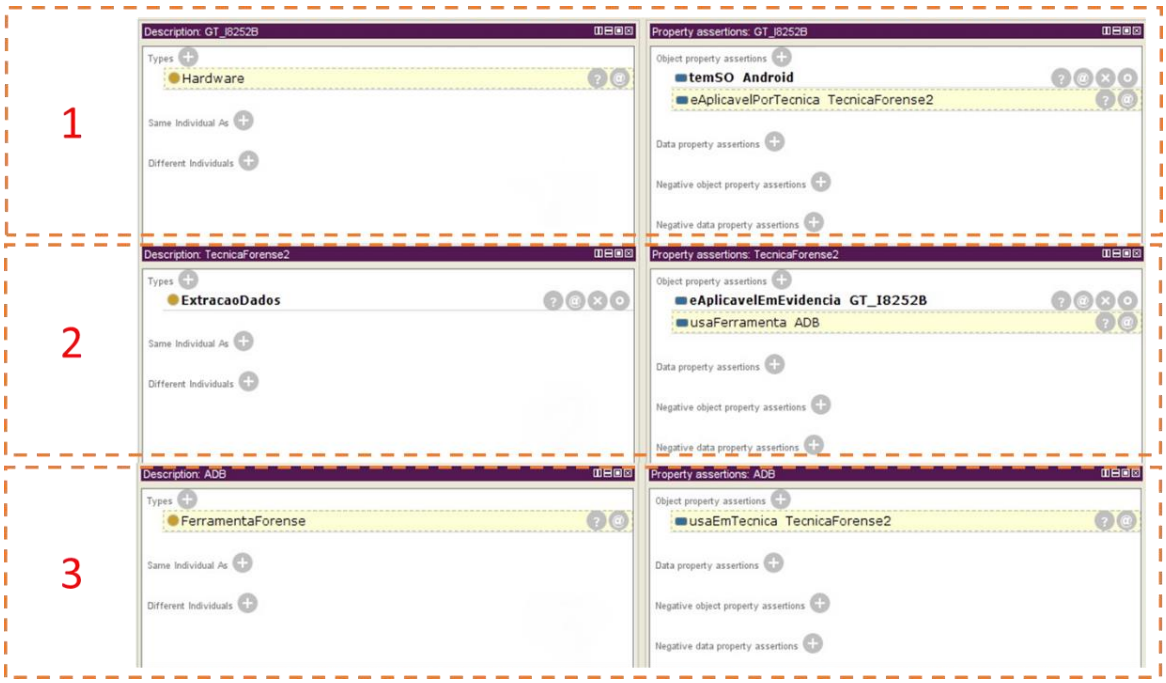


Figura 5.8 - Inferências do raciocinador

- Quadro 1 – Um Indivíduo GT_I8252B é declarado e a propriedade "**temSOAndroid"** é atribuída ao mesmo. Por inferência, em relação as propriedades de objeto, o raciocinador entende que esse indivíduo pertence à classe *Hardware*. Além disso, a propriedade "eAplicavelPorTecnica" "TecnicaForense2" é atrelada ao mesmo pois esta é inversa da propriedade declarada no Quadro 2.
- Quadro 2 – Um Indivíduo "TenicaForense2", membro da classe "ExtracaoDados", é declarado e a propriedade "eAplicavelEmEvidencia" GT_I8252B é atribuída ao mesmo. Pela regra 2, descrita na Figura 5.7, o raciocinador infere que essa técnica vai fazer uso da ferramenta ADB (propriedade de dados "usaFerramenta" ADB), pois a técnica será aplicada em uma evidência que tem Android como sistema operacional.
- Quadro 3 – Um Indivíduo ADB é declarado e nenhuma informação adicional é atribuída ao mesmo. Por inferência, também através da regra 2 (Figura 5.7), é possível saber que esse indivíduo está sendo usado pela "TecnicaForense2" e que por possuir essa propriedade de objeto (usaEmTecnica), ele é um membro da classe Ferramenta Forense.

A Figura 5.9 apresenta outras inferências realizadas pelo raciocinador a partir da regra 2 (Figura 5.7) e também a partir de outros axiomas OWL.

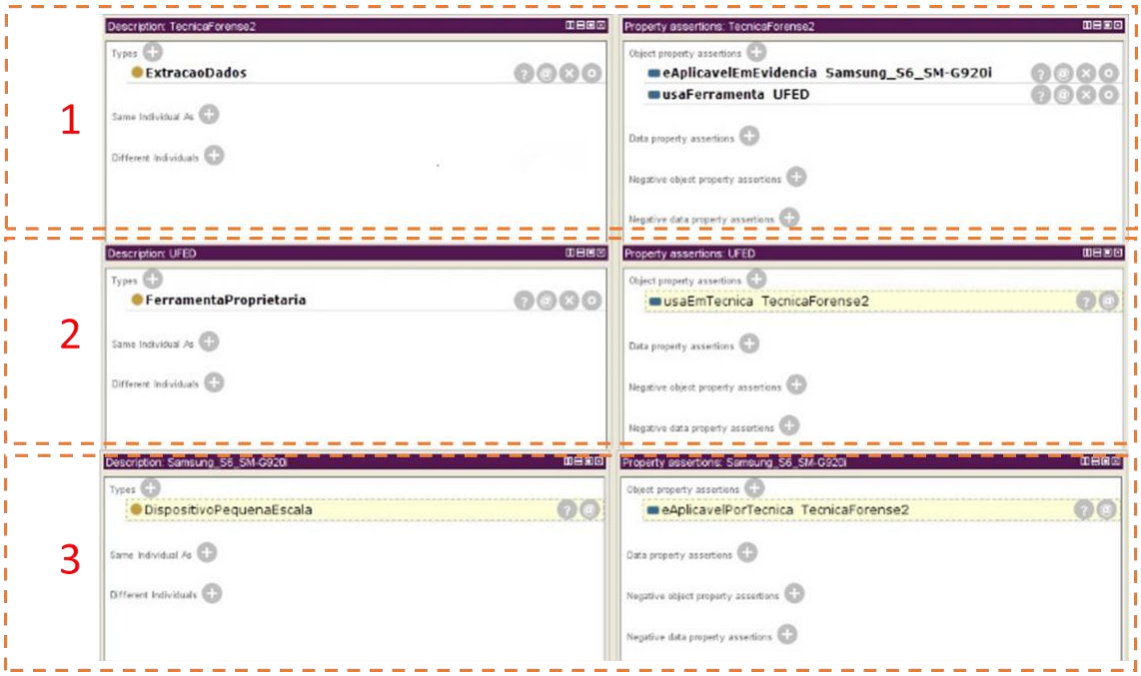


Figura 5.9 - Inferências a partir da regra 2

Na Figura 5.9 é possível notar três quadros referentes a três indivíduos na ontologia, as seguintes declarações e inferências (declarações destacadas com fundo amarelado) podem ser observadas:

- Quadro 1 – Um Indivíduo “TenicaForense2”, membro da classe “ExtracaoDados”, é declarado e as propriedades “eAplicavelEmEvidencia” Samsung_S6_SM-G920i; “usaFerramenta” UFED é atribuída ao mesmo.
- Quadro 2 – Um Indivíduo UFED, membro da classe “FerramentaProprietaria”, é declarado. Através de propriedades de objeto inversas, é possível inferir que esse indivíduo possui a propriedade de objeto “usaEmTecnica”, que o conecta a “TecnicaForense2” (Quadro 1).
- Quadro 3 – Um Indivíduo Samsung_S6_SM-G920i é declarado e nenhuma outra informação é atribuída ao mesmo. Por inferência de propriedades inversas, o raciocinador entende que esse indivíduo possui a propriedade “eAplicavelPorTecnica”, que o conecta a “TecnicaForense2” (Quadro 1). Por sua vez, devido a regra SWRL 3 na Figura 5.7, é inferido que esse indivíduo é membro da classe “DispositivoPequenaEscala”, visto que para a aplicação de uma técnica forense no mesmo, a ferramenta UFED está sendo utilizada.

5.3 VERIFICAÇÃO COM ESPECIALISTAS

Para verificação da ontologia, seus termos, relacionamentos definidos na Ontocf foi elaborado um questionário (Apêndice F), composto por 25 questões que foram respondidas por oito especialistas, peritos criminais e alunos do Mestrado Profissional, que atuam em variados órgãos de criminalística do Brasil. As questões objetivaram, em uma análise preliminar, verificar se a Ontocf estava condizente com relação aos critérios da clareza, coerência e completude.

Devido às limitações do questionário web, quando comparado a aplicação de um questionário presencial, em obter informações mais específicas e captar as impressões dos especialistas, a validação foi direcionada para os conceitos da ontologia. Apenas foi formulada uma questão apresentado o modelo conceitual da ontologia e suas relações.

Ao total, cada um dos 22 conceitos foi avaliado por 8 especialistas, resultando em 176 respostas. Conforme figura 5.10, 75% das respostas indicaram que os especialistas estavam plenamente de acordo com os conceitos apresentados.

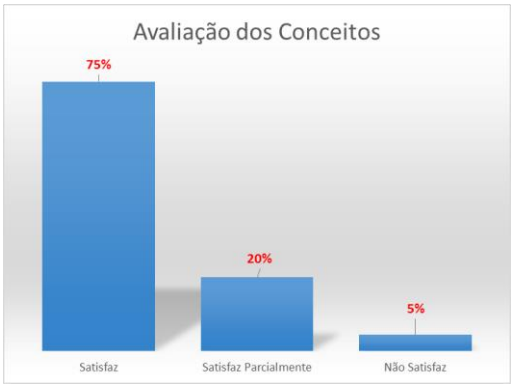


Figura 5.10 – Avaliação geral dos conceitos

Observa-se no Apêndice E que a soma da avaliação dos conceitos “satisfaz plenamente” ou “satisfaz parcialmente” representa 87% das respostas dos especialistas, exceto para o conceito evidência, que obteve este percentual de 75%, conforme apresentado na Figura 5.11.

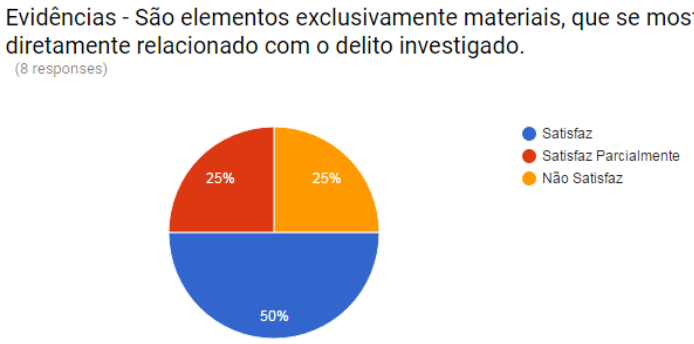


Figura 5.11 – Avaliação do conceito Evidência

Cinco dos oito especialistas fizeram contribuições escritas contendo sugestões para a definição dos conceitos, as quais foram analisadas e incorporadas, quando pertinente na versão final da lista de conceitos do Glossário de Termos Apêndice B.

Conforme anteriormente mencionado, foi elaborada uma questão do formulário de verificação contendo o modelo conceitual da Ontocf, com os principais conceitos e relações, para avaliação da correção dos relacionamentos entre conceitos.

Obteve-se nessa questão, conforme apresentado na Figura 5.12, um resultado de que 62,5% dos especialistas não sabem opinar sobre esta questão, o que expressa o desconhecimento dos especialistas sobre o tema ontologia, ou a necessidade de realizar uma explicação prévia sobre o assunto. Entretanto, um dos especialistas que identificou um relacionamento equivocado, fez em seguida o seguinte comentário:

> "A classe SOFTWARE também deveria se relacionar com EVIDENCIA VOLATIL. Um malware instalado exclusivamente em memória pode representar esse relacionamento."

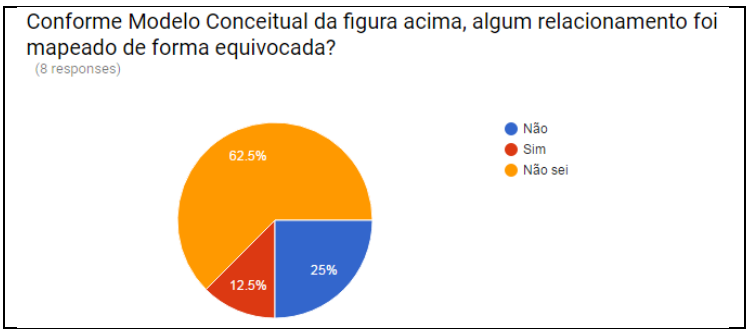


Figura 5.12 - Impressão dos especialistas sobre modelo conceitual da Ontocf

6 CONCLUSÕES

Mediante os estudos realizados neste trabalho, pode-se observar que o desenvolvimento da ontologia para representar conceitos e relações da Computação Forense representa uma solução para a falta de estruturação e padronização dos seus processos, sobretudo na comunicação, entre o solicitante dos exames e o perito. Com o uso da ontologia, pode-se fornecer informação objetiva no processo de realização da perícia, a fim de evitar erros e guiar os atores nos múltiplos caminhos do mundo da tecnologia da informação. Para integrar semanticamente tais informações em uma base de conhecimento, optou-se por conectar os conceitos-chave: caso criminal, crime, técnica forense, evidência e quesitos de exame.

Finalizado o estágio da especificação e conceitualização da Ontocf, os resultados parciais, que incluem a construção da ontologia no ambiente Protégé, com suas classes e relações, foram utilizados para a produção do artigo "Ontologia aplicada no processo de Computação Forense" [13](LEMOS FILHO, 2016), o qual foi apresentado no Seminário de Pesquisas em Ontologias - Ontobras), na edição de 2016. Os anais do Ontobras 2016 serão publicados no periódico online CEUR-WS[14].

Os resultados obtidos na primeira abordagem permitiram fazer correções na ontologia e as consultas realizadas trouxeram as respostas esperadas para as questões definidas, respondendo satisfatoriamente as mesmas. Embora restrito as opções de buscas ofertadas aos usuários da solução, os resultados satisfizeram a necessidade de informação que desejava ser recuperada.

Convém ressaltar, portanto, que a construção da Ontocf apresentou limitações na sua construção e avaliação. Primeiro, em decorrência tanto pequeno conjunto de dados (50 documentos), quanto a quantidade de questionários aplicados (8 questionários) que não representa uma amostra probabilística. Segundo, pelo reduzido escopo de avaliação do questionário, que aborda apenas os principais conceitos. Terceiro, pelo reduzido número de restrições implementadas, que limita as atividades de testes. Mesmo com essas limitações, foi possível utilizar os resultados das atividades de testes e avaliação com especialistas para aperfeiçoar a Ontocf, produzindo uma nova versão, como, também verificar tecnicamente a consistência da ontologia e dos seus artefatos.

[13]Disponível para download em: https://secplab.ppgia.pucpr.br/ontobras
[14]CEUR Workshop Proceedings (CEUR-WS.org)

Um efeito percebido da utilização da base de conhecimento foi que durante a utilização surgiram outros critérios para recuperação da informação que em alguns casos necessitavam apenas da implementação de novas consultas SPARQL, outras demandavam modificação do modelo ontológico pela criação de novas relações ou inclusão de novos conceitos na taxionomia.

O resultado das consultas na base de conhecimento do processo de Computação Forense trouxe ao cenário das atividades policiais investigativas, informação útil para tomada de decisão e para o aprendizado do amplo e complexo cenário do tema.

A base de conhecimento pode obter informações sobre o trabalho dos atores do processo, Perito e Autoridade Solicitante, individualmente, em função das suas escolhas dos quesitos, das técnicas e ferramentas empregadas. Tais informações podem ser utilizadas para explicitar conhecimentos por meio da disponibilização dos casos criminais, ou para promover a transferência de conhecimentos por meio de reuniões ou capacitação da equipe.

A proposta da linguagem única, operacionalizada pela Ontocf, possibilita abordagens à Computação Forense, através de questões organizacionais que vão além daquelas ligadas a TI, como:

1) questões relativas ao processo pericial: identificação da coleta da evidência digital, necessidade de novos quesitos de perícia diante da explicitação da evidência, dentre outros;
2) questões que envolvem a compreensão das autoridades solicitantes: possibilidades investigativas, conhecimento das evidências digitais, dentre outras;
3) questões que envolvem o conhecimento do perito: capacitação técnica, procedimentos alternativos, conhecimento do caso criminal e os respectivos pedidos de exames, procedimentos em geral, abordados nos limitados POPs ou manuais.

Diante do exposto, deseja-se que o processo de avaliação, extensão e melhoria da Ontocf seja contínuo e que a sua adoção no processo forense resulte em benefícios crescentes.

Enfim, considerando a multidisciplinaridade e complexidade conceitual, em diversas camadas de abstração da Computação Forense, a relevância desta dissertação está no fato de que abre novas perspectivas para o uso da ontologia como ponto de partida para sistematização do conhecimento dos processos da Computação Forense, que promova a sua disseminação entre os atores envolvidos, facilitando sua comunicação, aumentando a eficiência operacional dos exames periciais e. por conseguinte, das investigações criminais.

6.1 TRABALHOS FUTUROS

Inicialmente, pretende-se como trabalho futuro proceder uma avaliação sistemática da Ontocf, por meio de uma pesquisa quantitativa, baseada em critérios de avaliação. Conforme Vrandecic (2009) sugere, uma forma de avaliar uma ontologia é definir critérios de avaliação e escolher métodos atendem aos critérios definidos, minimizando dessa forma eventuais problemas na ontologia.

Em seguida, sugere-se disponibilizar o repositório RDF produzido na implementação deste trabalho, diretamente na web ou por meio de uma interface amigável de consulta para profissionais da área de segurança pública, que interagem no processo de Computação Forense. O sistema de consulta almejado poderá incorporar mecanismos de obtenção de novas instâncias, sugestão de alteração da ontologia, inclusive com a criação de novos conceitos e relações.

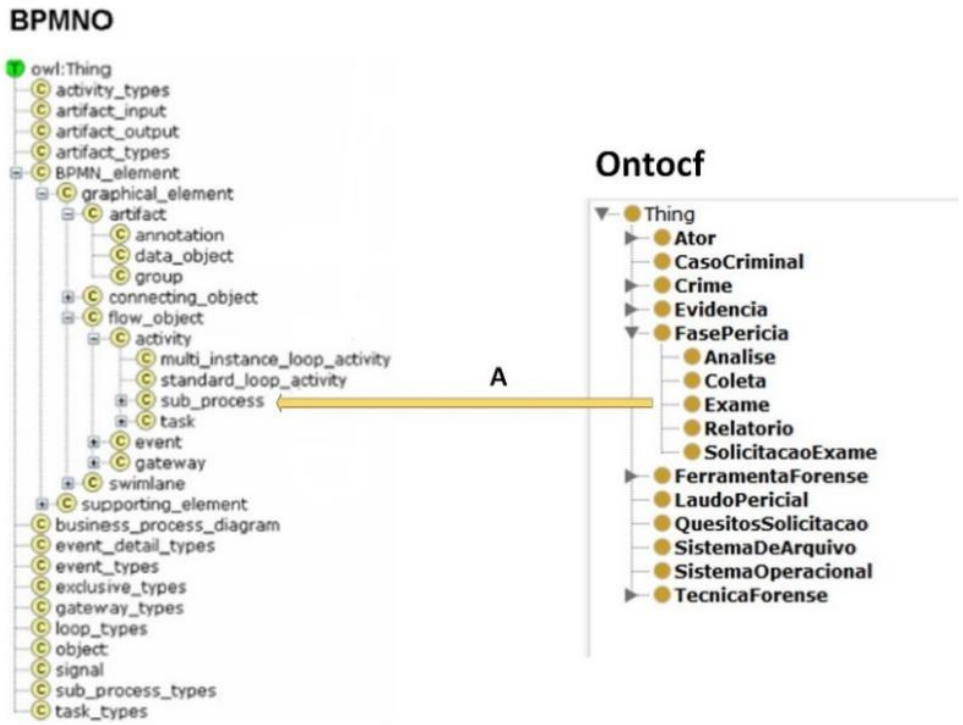


Figura 6.1 - Exemplo de utilização do BPKD na Ontocf

Posteriormente, como perspectiva para pesquisas futuras, pretende-se expandir a ontologia para representar o ciclo completo do processo de investigação dos crimes cibernéticos de forma flexível, adequando-se aos processos internos de cada departamento de investigação, por meio da integração da ontologia de modelagem de processos de negócios BPMNO, que provê a formalização da parte estrutural dos diagramas de processo de negócio *Business Process Diagrams* (BPD).

A Figura 6.1 ilustra a aplicação do *framework* BPKD no processo de Computação forense por meio da integração com a Ontocf. Neste exemplo, qualquer conceito equivalente, ou mais específico que Exame, pode ser usado para denotar os elementos BPMN do tipo Subprocesso (*sub_process*).

REFERÊNCIAS

ALZAABI, M.; JONES, A.; MARTIN, T. A. an Ontology-Based Forensic Analysis Tool. , v. 1, p. 121–136, 2013.

BEM, D. et al. Computer forensics-past, present and future. **Journal of Information Science and Technology**, v. 5, n. 3, p. 43–59, 2008.

BRASIL. **Manual de Orientação de quesitos da Perícia Criminal**. Brasília, 2012.

_____. **Procedimento Operacional Padrão Perícia Criminal**. Brasília, 2013.

BRINSON, A.; ROBINSON, A.; ROGERS, M. A cyber forensics ontology: Creating a new approach to studying cyber forensics. **Digital Investigation**, v. 3, n. SUPPL., p. 37–43, 2006.

CLARKE, N. **Computer forensics: A Pocket Guide**. Cambridgeshire: IT Governance Publishing, 2010.

CORCHO, O.; FERNÁNDEZ-LÓPEZ, M.; GÓMEZ-PÉREZ, A. Methodologies, tools and languages for building ontologies. Where is their meeting point? **Data & knowledge engineering**, v. 46, n. 1, p. 41–64, 2003. Elsevier.

ĆOSIĆ, J.; ĆOSIĆ, Z. The Necessity of Developing a Digital Evidence Ontology. **23th Central European Conference on Information …**, p. 325–330, 2012. Disponível em: <http://www.ceciis.foi.hr/app/public/conferences/1/papers2012/iss5.pdf>. Acesso em: 08 ago. 2015

COSTA, L. R. **Metodologia E Arquitetura Para Sistematização Do Processo Investigatório**, 2012. Universidade de Brasília.

ELEUTÉRIO, F. **Análise do conceito de crime.** Universidade Estadual de Ponta Grossa, Ponta Grossa, 1997.

FERNÁNDEZ-LÓPEZ, M.; GÓMEZ-PÉREZ, A.; JURISTO, N. METHONTOLOGY: From Ontological Art Towards Ontological Engineering. **AAAI-97 Spring Symposium Series**, v. SS-97-06, p. 33–40, 1997. Disponível em: <http://oa.upm.es/5484/>. Acesso em: 05 abr. 2016.

FRANCESCO MARINO, C. D. et al. Reasoning on semantically annotated processes. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 5364 LNCS, p. 132–146, 2008.

GARCIA, I. E. **Inquérito - Procedimento Policial**. 9ª ed. Goiania, 2002.

GASEVIC, D.; DJURIC, D.; DEVEDZIC, V. **Model Driven Architecture and Ontology Development**. 2006.

GIALAMAS, D. M. Criminalistics. In: J. Siegel; G. Knupfer; P. Saukko (Eds.); **Encyclopedia of Forensic Sciences**. p.471–477, 2000. Amsterdan: Elsevier.
GÓMEZ-PÉREZ, A. (FACULTAD D. I. Evaluation of Taxonomic Knowledge in Ontologies and Knowledge Bases. Banff Knowledge Acquisition for Knowledge-Based Systems, KAW'99. **Anais...** . v. 26, p.6.1.1--6.1.18, 1999. Disponível em: <http://sern.ucalgary.ca/KSI/KAW/KAW99>. Acesso em: 20 out. 2016.

GRUBER, T. R.. A translation approach to portable ontology specifications. **Knowledge Acquisition**, v. 5, n. 2, p. 199–220, 1993. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.101.7493>. Acesso em: 12 mar. 2016.

GUARINO, N. Formal Ontology and Information Systems. **Proceedings of the first international conference**, v. 46, n. June, p. 3–15, 1998. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.29.1776&rep=rep1&type=pdf>. Acesso em: 27 out. 2015

_____; OBERLE, D.; STAAB, S. What Is an Ontology. **Handbook on Ontologies**. p.1–17, 2009. Disponível em: <http://doi.org/10.1007/978-3-540-92673-3/>. Acesso em: 19 nov. 2016

HARRILL, D. C.; MISLAN, R. P. A Small Scale Digital Device Forensics ontology. **Small Scale Digital Device Forensics Journal**, v. 1, n. 1, p. 1–7, 2007.

HARTH, A. et al. On-the-fly integration of static and dynamic linked data. Proceedings of the Fourth International Conference on Consuming Linked Data-Volume 1034. **Anais...** . p.1–12, 2013.

HASSANPOUR, S.; O'CONNOR, M. J.; DAS, A. K. Exploration of SWRL rule bases through visualization, paraphrasing, and categorization of rules. International Workshop on Rules and Rule Markup Languages for the Semantic Web. **Anais...** . p.246–261, 2009.

HOELZ, B. W. P.; RALHA, C. G. A Framework for Semantic Annotation of Digital Evidence. **Proceedings of the 2013 ACM Symposium on Applied Computing**, 2013.

HORRIDGE, M. et al. A Practical Guide To Building OWL Ontologies Using The Protégé-OWL Plugin and CO-ODE Tools Edition 1.0. **University of Manchester**, 2004.

HORROCKS, I. et al. A semantic web rule language combining OWL and RuleML. **W3C Member submission**, v. 21, p. 79, 2004.

HORTÊNCIO FILHO, F. W. B.; LÓSCIO, B. F.; CAMPOS, G. A. L. DE. **Inferência sobre Ontologias no contexto da Web Semântica**. (2008). Disponível em: http://www.infobrasil.inf.br/userfiles/Infer%C3%AAncia%20sobre%20Ontologias%20no%20contexto%20da%20Web%20Sem%C3%A2ntica.pdf. Acesso em: 22 out. 2016.

INMAN, K.; RUDIN, N. The origin of evidence. **Forensic Science International**, v. 126, n. 1, p. 11–16, 2002. Elsevier.

ISOTANI, S.; BITTENCOURT, I. I. **Dados Abertos Conectados**. Novatec Editora, 2015.

JASPER, R.; USCHOLD, M. A framework for understanding and classifying ontology applications. **Proceedings 12th Int. Workshop on Knowledge Acquisition, Modelling, and Management KAW**, v. 99, p. 16–21, 1999.

KENT, K. et al. Guide to integrating forensic techniques into incident response. **NIST Special Publication**, , n. August, p. 800–886, 2006.

KIMBALL, R.; ROSS, M. **The data warehouse toolkit: the complete guide to dimensional modeling**. John Wiley & Sons, 2011.

LAUFER, C. **Guia de Web Semântica.** Disponível em: <http://ceweb.br/guias/web-semantica/>. Acesso em: 5 nov. 2016.

LEMOS FILHO, E. Ontologia aplicada no processo de Computação Forense. Seminário Brasileiro de Ontolgias (Ontobras). **Anais...** . p.1–6, 2016. Curitiba. Disponível em: <https://drive.google.com/file/d/0B8yxc3skHHbCM1JvUDRQRW1DTVE/view>. Acesso em: 15 nov. 2016.

LÓPEZ, F. Overview Of Methodologies For Building Ontologies. **Proceedings of the IJCAI99 Workshop on Ontologies and ProblemSolving Methods Lessons Learned and Future Trends CEUR Publications**, v. 1999, n. 2, p. 1–13, 1999. Disponível em: <http://iwayan.info/Research/Ontology/Tutor_Workshop/Tutorial_4_Analysis.pdf>. Acesso em: 12 fev. 2016 .

MIZOGUCHI, R. Tutorial on ontological engineering Part 2: Ontology development, tools and languages. **New Generation Computing**, v. 22, n. 1, p. 61–96, 2004. Springer.

MOIGNE, J.-L. Le; PINHEIRO, J.; PAZ, J. **A teoria do sistema geral:** teoria da modelização. 1977.

MORAIS, E. A. M.; AMBRÓSIO, A. P. L. **Ontologias: conceitos, usos, tipos, metodologias, ferramentas e linguagens**. Universidade Federal de Goiáis, 2007.

MOTIK, B.; GRAU, B. C.; HORROCKS, I.; et al. Owl 2 web ontology language: Profiles. **W3C recommendation**, v. 27, p. 61, 2009.

NONAKA, I.; TAKEUCHI, H. **Criação de conhecimento na empresa - como as empresas japonesas gerem a dinâmica da inovação**. 1997.

NOY, N.; MCGUINNESS, D. Ontology development 101: A guide to creating your first ontology. **Development**, v. 32, p. 1–25, 2001. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.136.5085&rep=rep1&type=pdf%5Cnhttp://liris.cnrs.fr/alain.mille/enseignements/Ecole_Centrale/What is an ontology and why we need it.htm>. Acesso em: 27 ago. 2015

OREN, E. et al. **What are Semantic Annotations?** , p. 14, 2006.

PARK, H.; CHO, S.; KWON, H.-C. Cyber Forensics Ontology for Cyber Criminal

Investigation. In: M. Sorell (Ed.); **Forensics in Telecommunications, Information and Multimedia**. p.160–165, 2009. Adelaide: Springer Berlin Heidelberg.

ROXIN, C. **Estudos de direito penal**. trad. Luis Greco. , 2012. São Paulo: Renovar.

SLEPICKA, J. et al. An Alternative Interpretation of R2RML for Heterogeneous Sources. Proceedings of the 6th International Workshop on Consuming Linked Data. **Anais...** , 2015

NUCCI, G. de Souza; BRASIL. **Código de processo penal comentado**. Editora Revista dos Tribunais, 2009.

SAMMES, T.; JENKINSON, B. **Forensic Computing: A Practitioner's Guide**. London, UK, UK: Springer-Verlag, 2000.

STEVENS, R. Ontology-based knowledge representation for bioinformatics. **Briefings in Bioinformatics**, v. 1, n. 4, p. 398–414, 2000. Disponível em: <http://bib.oxfordjournals.org/content/1/4/398.short>. Acesso em: 18 mai. 2016

STUDER, R.; BENJAMINS, V. R.; FENSEL, D. Knowledge engineering: Principles and methods. **Data & Knowledge Engineering**, v. 25, n. 1–2, p. 161–197, 1998.

SWARTOUT, B. et al. Toward Distributed Use of Large-Scale Ontologies. **Proc. of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems**, p. 138–148, 1996. Disponível em: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Toward+Distributed+Use+of+Large-Scale+Ontologies#0>. Acesso em: 12 abr. 2016.

USCHOLD, M.; GRUNINGER, M. Ontologies: principles, methods and applications. **The Knowledge Engineering Review**, v. 11, n. 2, p. 93, 1996. Disponível em: <http://www.journals.cambridge.org/abstract_S0269888900007797>. Acesso em: 17 nov. 2015.

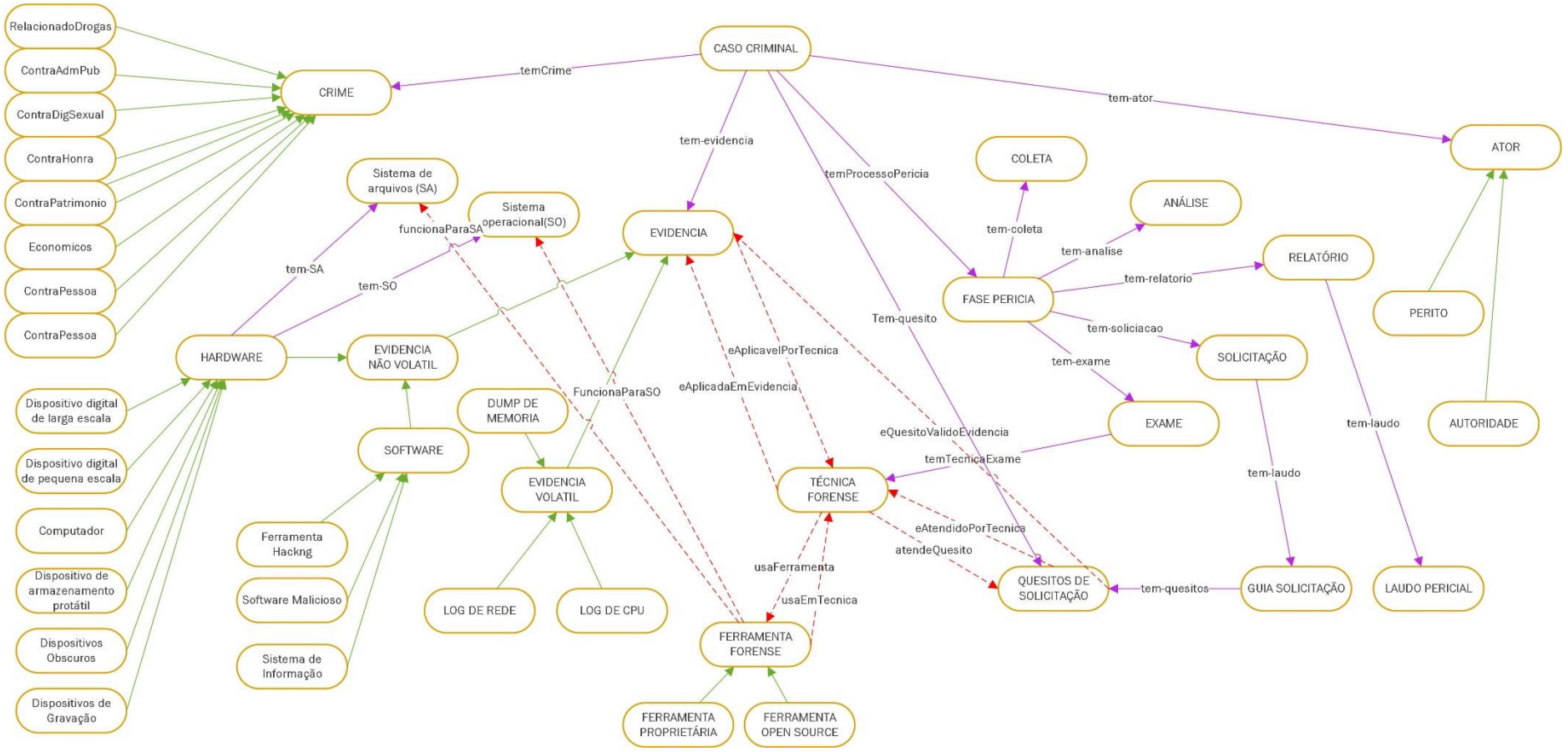
VRANDECIC, D., **Ontology Evaluation**, In: Handbook on Ontologies, Second edition, Staab, S., Studer, R. (Eds.), Springer, pp. 293 – 313, 2009.

VELHO, J. A.; GEISER, G. C.; ESPINDULA, A. **Ciências Forenses – Uma introdução às principais áreas da Criminalística Moderna**. 2ª ed. ed. Campinas: Editora Millenium, 2013.

WELTY, C.; MCGUINNESS, D. L.; SMITH, M. K. Owl web ontology language guide. **W3C recommendation, W3C (February 2004) http://www. w3. org/TR/2004/REC-owl-guide-20040210**, 2004.

APÊNDICES

APÊNDICE A – ÁRVORE (GRAFO) DE CLASSIFICAÇÃO DE CONCEITOS



APÊNDICE B – GLOSSÁRIO DE TERMOS COMPLETO - GT

CONCEITOS:

TERMO	DESCRIÇÃO
Analise (fase)	A fase de análise envolve a avaliação dos resultados obtidos na fase de exame para alcançar informações relacionadas aos fatos e alvos.
Exame (fase)	Processo execução de procedimentos periciais para obtenção de dados para análise
Coleta (fase)	Na fase de coleta os materiais referentes a um evento específico são identificados, rotulados, registrados e recolhidos, preservando-os em relação a sua integridade
Solicitação de Exame (fase)	Na fase de solicitação o investigador, autoridade, elabora o pedido de exames contendo quesitos orientados pela natureza dos materiais e pelo fato investigado.
Relatório (fase)	A fase de relatório tem por objetivo explicitar os resultados dos exames por meio do documento Laudo Pericial
Ator	Pessoa participante do processo de Computação forense
Autoridade solicitante	Solicitante do Exame Pericial com competências previstas em lei.
Perito Criminal	Profissional responsável pela materialização da prova do crime por meio do adequado exame e da interpretação correta dos vestígios materiais dessas infrações.
Caso Criminal	Reuni os elementos e atributos associados a uma investigação criminal, os quais são necessários para realização de um exame pericial.
Crime	Crime é toda a ação, típica, antijurídica e culpável. Tendo como um dos elementos definidores o Tipo penal, sendo a descrição concreta da conduta proibida
Evidência	Evidências são elementos exclusivamente materiais, que se mostram diretamente relacionado com o delito investigado
Evidência não volátil	Evidência que possui dados que podem permanecer em um dispositivo durante longos períodos de tempo e podem ser recuperados mesmo após o mesmo ser desligada.
Evidência volátil	Evidência que possui dados de natureza efêmera que podem ser recuperados do dispositivo enquanto a mesma estiver ligada.
Ferramenta Forense	Ferramentas computacionais utilizadas nas fases do processo de Computação Forense
Ferramenta Open Source	Ferramenta computacional com código aberto de uso livre utilizadas para realizar uma análise forense.
Ferramenta Proprietária	Ferramentas computacional adquiridas comercialmente com direitos exclusivos de distribuição e venda
Laudo Pericial	É um documento técnico-formal que exprime o resultado do trabalho do perito, realizado durante o processo de Computação Forense.
Fases do Processo Pericial	Fases do processo de Computação Forense que envolve os subprocessos de: coleta, solicitação, exame, análise e relatório.
Quesitos de Solicitação	quesitos apresentados pela autoridade solicitante, dentro de um contexto criminal, isto é, dentro de uma situação conhecida no bojo da investigação ou do processo penal que direciona a forma de examinar os materiais questionados.

Sistema de Arquivos	Um sistema de arquivos é um conjunto de estruturas lógicas que permite o sistema operacional o acesso aos dados armazenados
Sistema Operacional	Um conjunto de programas especialmente feitos para gerir recursos de hardware do computador, inclusive os dispositivos periféricos de um computador.
Sistema Operacional Proprietário	Sistema operacional adquirido comercialmente, com direitos exclusivos de distribuição e venda
Sistema Operacional Open Source	Sistema operacional de código aberto de uso livre
Técnica Forense	Técnica empregada para realização das atividades previstas no processo de Computação Forense

TERMO DA TAXIONOMIA EVIDÊNCIA:

TERMO (EVIDÊNCIA)	**DESCRIÇÃO**
Hardware	Evidência em suporte de hardware
Dispositivo digital de larga escala	Equipamentos corporativos de processamento para provimento de serviços computacionais
Grid computacional	Equipamentos interligados em processamento distribuído
Cluster	Associação de computadores com proposito de processamento
Dispositivo digital de pequena escala	Equipamentos microprocessados portáteis
Telefone Celular	Telefone móvel celular
Smartphone	Telefone móvel celular com procesamento de aplicativos e sistema operacional
Tablet	Tablet com procesamento de aplicativos e sistema operacional
Computador	Microcomputador
Desktops	Microcomputador de mesa
Disco Rígido	Disco Rígido retirado de computador
Notebook	Microcomputador portátil
Servidor	Microcomputador servidor de serviços em rede
Dispositivo de armazenamento protátil	Dispositivo para armazenamento de dados
Digital Music Player	Dispositivo para armazenamento de dados com tocador de dados
Pendrive	Dispositivo para armazenamento de dados portátil
Disco rígido Externo	Dispositivo para armazenamento de dados portátil de alta capacidade
Dispositivos Obscuros	Dispositivos microprocessados com armazenamento de dados sem identificação definida ou manufaturados
Dispositivos de jogos	Dispositivos para jogos microprocessados com armazenamento de dados
Dispositivos de Gravação	Dispositivos microprocessados com armazenamento de dados

Software	Evidência em suporte de software
Ferramenta Hacking	Software destinado a realizar um conjunto de ações hacker
Sistema de Informação	Software de sistemas de informações
Software Malicioso	Software com objetivo de causar danos ou violações de segurança em dispositivos eletrônicos
Banco de dados	Sistema de gerenciamento de banco de dados (SGBD)

TÉCNICAS FORENSES:

Extração de dados	Técnica para recuperar toda e qualquer informação presente no dispositivo questionado, inclusive arquivos apagados.
Carving de dados	Técnica de acesso de um arquivo independentemente da sua tabela de endereçamento (alocação), voltado para recuperação de arquivos baseando-se em seus headers, footers e suas estruturas de dados.
Análise de Log	Técnica de interpretação dos registros de log gerados por sistemas de equipamento, de forma a apresentar informações sobre um sistema de forma clara e compreensível, para determinar a posteriori ações realizadas e violações de segurança.
Indexação de arquivos	indexação é uma técnica de organização dos arquivos em um dispositivo de armazenamento que cria uma espécie de catálogo, contendo todas as cadeias alfanuméricas encontradas, assim como a localização de cada uma delas.
Análise de malware	Técnica de busca de assinatura de malwares e das suas ações realizadas
Análise de equipamentos eletrônicos	Análise de equipamentos eletrônico que possuem dispositivos microprocessados e de armazenamento de dados em mídias variadas, podendo possuir sistema de arquivos conhecidos ou proprietários.
Análise de criptografia	Técnica para identificar arquivos criptografados objetivando quebrar senhas de encriptação, como, por exemplo, busca de chave de criptografia, procedimentos de força bruta.
Análise de rede	captura de dados em redes ou seus dispositivos para análise e extração das informações, objeto da investigação, por meio de softwares especialistas ou por operação direta do sistema operacional dos dispositivos conectados.
Análise de artefatos web	Busca de artefatos de internet pelo processo de parsing e carving em espaços alocados e não alocados. Os artefatos que podem ser reconstruídos são: aplicativos de redes sociais, aplicativos de mensagens instantâneas e chat, atividade de Web browser, visualizar pesquisas realizadas, entre outros.
Análise de Email	Análise de encaminhamento e de conteúdo de mensagens de correio eletrônico (email) baseado nas informações de cabeçalho e de dados e log de servidores de email.
Análise de Registro Windows	Análise no banco de dados de configuração do sistema operacional Windows sobre aplicativos, sistema operacional, usuários e dispositivos de hardware.
Análise de video	Busca e exame de Conteúdo de vídeo, que consiste na verificação minuciosa de infrações, delitos ou crimes registrados em vídeo buscando, quando possível, a identificação dos autores, bem como da dinâmica dos eventos.
Análise de arquivos de imagem	Busca e exame de Conteúdo de vídeo, que consiste na verificação minuciosa de infrações, delitos ou crimes registrados em vídeo buscando, quando possível, a identificação dos autores, bem como da dinâmica dos eventos.

RELAÇÕES:

Relação	Conceito 1	Conceito 2	Inversa de:	Descrição
atende quesito	Técnica Forense	Quesito de Solicitação	é atendido por técnica	Relaciona os quesitos que podem se atendidos por uma determinada técnica.
é quesito válido para evidência	Quesito de solicitação	Evidência	Evidência aceita quesito	Quesitos aplicáveis em determinada evidência
é aplicável em evidência	Técnica forense	Evidência	é aplicável por técnica	Técnicas que são aplicáveis em determinada evidência
é aplicável por técnica	Evidência	Técnica forense	é aplicável em evidência	A evidência pode ser analisada pelas determinada técnica
é evidência de	Evidência	Caso Criminal	Tem evidência	Relaciona as evidências de um caso criminal
evidência aceita quesito	Evidência	Quesito de solicitação	É quesito valido Evidência	Se a evidência pode ser questionada com determinado quesito
funciona para SA	Ferramenta forense	Sistema de arquivo	N/D	As ferramentas que podem ser utilizadas para um sistema de arquivos
funciona para SO	Ferramenta forense	Sistema Operacional	N/D	As ferramentas que podem ser utilizadas para um sistema operacional
tem ator	Caso Criminal	ator	N/D	Atores do caso criminal
tem quesito	Caso Criminal	Quesito de solicitação	N/D	Relaciona os quesitos de um caso criminal
tem SA	Hardware	Sistema de arquivo	N/D	Sistema de arquivo do hardware
tem SO	Hardware	Sistema operacional	N/D	Sistema de arquivo do hardware

Usa Ferramenta	Técnica Forense	Ferramenta Forense	Usa em Técnica	A ferramenta usada pela Técnica forense
Usa em técnica	Ferramenta Forense	Técnica Forense	Usa Ferramenta	As técnicas que usam a ferramenta forense
tem técnica coleta	Coleta ou exame	Técnica forense	N/D	Técnicas destinadas a coleta que também podem ser do exame
tem técnica exame	exame	Técnica forense	N/D	Técnicas destinadas a fase do exame
tem coleta	Fase Perícia	coleta	N/D	Relaciona fase coleta
tem solicitação	Fase Perícia	solicitação	N/D	Relaciona fasesolicitação
Tem exame	Fase Perícia	exame	N/D	Relaciona fase exame
Tem análise	Fase Perícia	análise	N/D	Relaciona fase análise
Relatório	Fase Perícia	relatório	N/D	Relaciona fase relatório

APÊNDICE C - DICIONÁRIO DE DADOS

Classe	Atributo	Descrição	Tipo
Caso criminal	idCrime	Número de identificação do caso	string
Caso criminal	dataCaso	Data que ocorreu o fato	DateTime
Caso criminal	dataSolicitacao	Data de solicitação do exame pericial	DateTime
Caso criminal	dataExame	Data do início do exame das evidências do caso criminal.	DateTime
Evidência	chegouLab	Data que a evidência chegou Lab	DateTime
Evidência	saiuLab	Data que a evidência saiu do Lab	DateTime
Evidência	quantEvidencia	Quantidade de evidências iguais	Integer

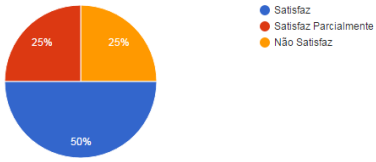
Obs. Utilizou-se a nomenclatura de tipo de dado utilizada no software Protégé.

APÊNDICE D - TABELA DE REGRAS OU RESTRIÇÕES

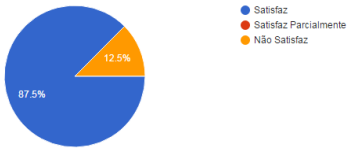
Item	Conceito/Relação	Restrição	Conceito/Relação
1	Técnica - Análise de Equipamento Eletrônico; Carving de dados; indexação de arquivos; extração de arquivos	Só se aplica para	Evidência Hardware
2	Ferramenta Forense UFED	Só se aplica	Dispositivo de pequena escala
3	Ferramenta Forense Malwarebits	Só se aplica	Software Malicioso
4	Ferramenta Forense Volatility	Só se aplica	Evidência volátil
5	ferramenta Android Debug Bridge (ADB)	Só se aplica	Sistema operacional Android

APÊNDICE E - RESULTADO DA AVALIAÇÃO DOS CONCEITOS E RELAÇÕES DA ONTOCF

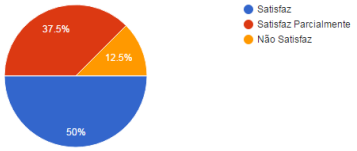
Evidências - São elementos exclusivamente materiais, que se mostram diretamente relacionado com o delito investigado.
(8 responses)



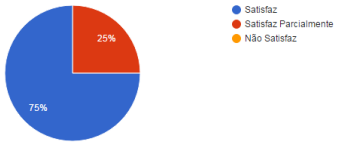
Evidência não volátil - É uma evidência que possui dados que podem permanecer em um dispositivo durante longos períodos de tempo e podem ser recuperados mesmo após o mesmo ser desligada.
(8 responses)



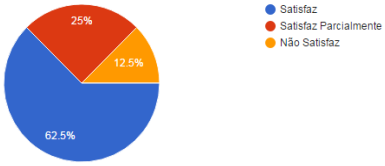
Evidência volátil - É uma evidência Evidência que possui dados voláteis que podem ser recuperados do dispositivo enquanto a mesma estiver ligada.
(8 responses)



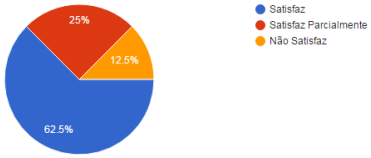
Ferramenta Forense - Ferramentas computacionais utilizadas nas fases do processo de Computação Forense.
(8 responses)



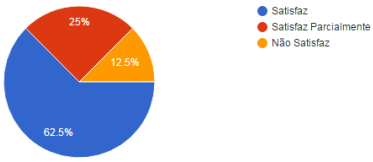
Ferramenta Forense Open Source - Ferramenta computacional em código aberto de uso livre utilizadas para realizar uma análise forense.
(8 responses)



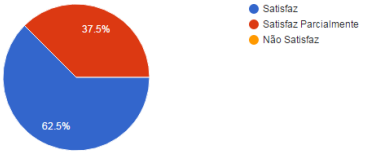
Ferramenta Forense Proprietária - Ferramentas computacional adquirid comercialmente e licenciada com direitos exclusivos para o produtor.
(8 responses)



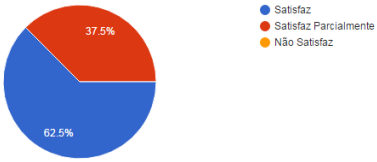
Ferramenta Forense Proprietária - Ferramentas computacional adquiridas comercialmente e licenciada com direitos exclusivos para o produtor.
(8 responses)



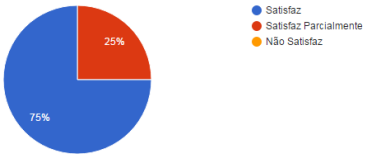
Sistema de arquivos - é um conjunto de estruturas lógicas que permite o sistema operacional controlar o acesso a um dispositivo de armazenam de dados.
(8 responses)



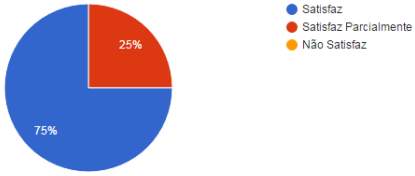
Sistema operacional - Um conjunto de programas especialmente feitos para a execução de várias tarefas, entre as quais servir de intermediário entre o utilizador e o computador. Um sistema operacional, tem também como função, gerir todos os periféricos de um computador.
(8 responses)



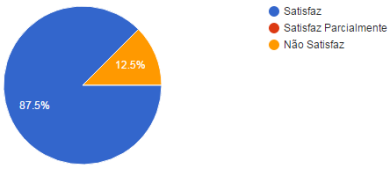
Quesitos de Solicitação - Quesitos apresentados pela autoridade solicit dentro de um contexto criminal, isto é, dentro de uma situação conhecic bojo da investigação ou do processo penal que direciona a forma de examinar os materiais questionados.
(8 responses)



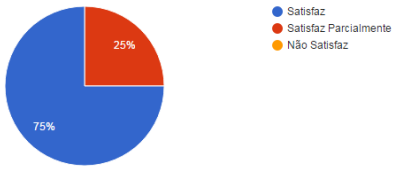
Técnica Forense - Técnica empregada para realização das atividades previstas no processo de Computação Forense.
(8 responses)



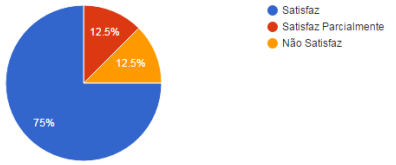
Perito Criminal - Profissional responsável pela materialização da pr crime por meio do adequado exame e da interpretação correta dos materiais dessas infrações.
(8 responses)



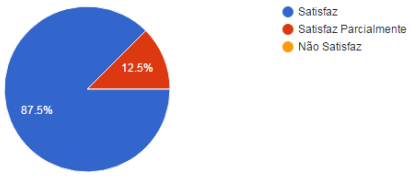
Autoridade Solicitante - Solicitante do Exame Pericial com competências previstas em lei.
(8 responses)



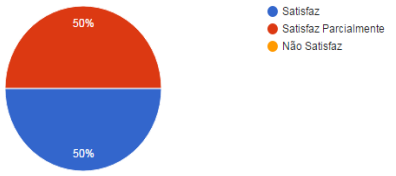
Caso Criminal - Caso que reuni os elementos e atributos associados investigação criminal, os quais são necessários para realização de u exame pericial.
(8 responses)



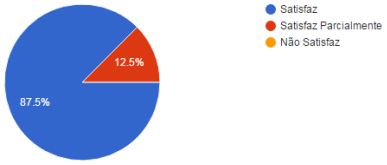
Crime - Crime é toda a ação, típica, antijurídica e culpável. Tendo como ur dos elementos definidores o Tipo penal, sendo a descrição concreta da conduta proibida.
(8 responses)



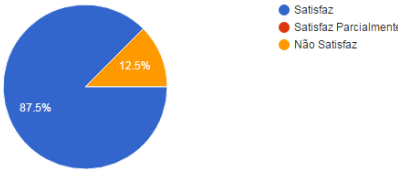
Fase do Processo Pericial - Fases do processo de Computação Fore envolve os subprocessos de: coleta, solicitação, exame, análise e rel
(8 responses)



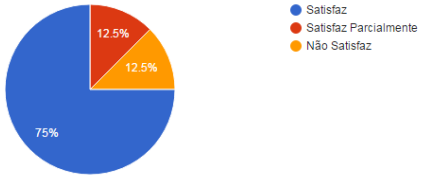
Fase de Solicitação do Exame - Na fase de solicitação do exame o investigador, autoridade, elabora o pedido de exames contendo quesitos orientados pela natureza dos materiais e pelo fato investigado.
(8 responses)



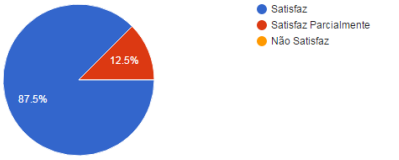
Fase de Coleta - Na fase de coleta os materiais referentes a um e específico são identificados, rotulados, registrados e recolhidos, preservando-os em relação a sua integridade
(8 responses)



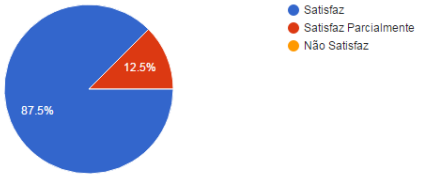
Fase de Exame - Processo execução de procedimentos periciais para obtenção de dados para análise.
(8 responses)



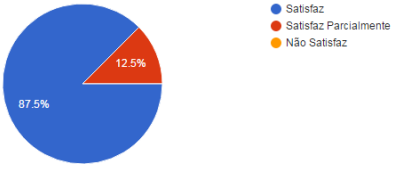
Fase de Análise - A fase de análise envolve a avaliação dos resultados obtidos na fase de exame para alcançar informações relacionadas aos e alvos.
(8 responses)



Fase de Relatório - A fase de relatório tem por objetivo explicitar os resultados dos exames por meio do documento Laudo Pericial.
(8 responses)



Crime - Crime é toda a ação, típica, antijurídica e culpável. Tendo c dos elementos definidores o Tipo penal, sendo a descrição concre conduta proibida.
(8 responses)



APÊNDICE F QUESTIONÁRIO DE AVALIAÇÃO DA ONTOCF

Questionário de Avaliação da Ontologia Domínio Sobre Processo de Computação Forense

Esta pesquisa faz parte de um trabalho acadêmico conduzido pelo aluno Egberto Vilas Boas Lemos Filho do Mestrado Profissional em Engenharia Elétrica área de concentração em Informática Forense e Segurança da Informação.

O principal objetivo desse questionário é obter um parecer de especialistas no domínio da Computação Forense em relação à definição, representação e relacionamentos e dos conceitos definidos na ontologia, afim de avaliar a acurácia, completude, coerência e correta cobertura.

Ontologia define um vocabulário comum para uma área de estudo que precisa compartilhar informação dentro de um domínio. Isso inclui definições em linguagem de máquina interpretável dos conceitos dentro do domínio e as relações entre eles.
Para garantir que uma ontologia seja construída com qualidade é necessário definir o domínio de conhecimento com objetividade, descrevendo o conhecimento essencial ao domínio e definindo um vocabulário que evite interpretações ambíguas.

Desde já agradecemos pela sua colaboração e disponibilidade em responder este questionário.

Evidências - São elementos exclusivamente materiais, que se mostram diretamente relacionado com o delito investigado. *

- ( ) Satisfaz
- ( ) Satisfaz Parcialmente
- ( ) Não Satisfaz

Evidência não volátil - É uma evidência que possui dados que podem permanecer em um dispositivo durante longos períodos de tempo e podem ser recuperados mesmo após o mesmo ser desligada. *

- ( ) Satisfaz
- ( ) Satisfaz Parcialmente

- ◯ Não Satisfaz

Evidência volátil - É uma evidência Evidência que possui dados voláteis que podem ser recuperados do dispositivo enquanto a mesma estiver ligada.

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Ferramenta Forense - Ferramentas computacionais utilizadas nas fases do processo de Computação Forense.

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Ferramenta Forense Open Source - Ferramenta computacional em código aberto de uso livre utilizadas para realizar uma análise forense.

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Ferramenta Forense Proprietária - Ferramentas computacional adquiridas comercialmente e licenciada com direitos exclusivos para o produtor.

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Sistema de arquivos - é um conjunto de estruturas lógicas que permite o sistema operacional controlar o acesso a um dispositivo de armazenamento de dados.

- ◯ Satisfaz

- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Sistema operacional - Um conjunto de programas especialmente feitos para a execução de várias tarefas, entre as quais servir de intermediário entre o utilizador e o computador. Um sistema operacional, tem também como função, gerir todos os periféricos de um computador. *

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Laudo Pericial - É um documento técnico-formal que exprime o resultado do trabalho do perito, realizado durante o processo de Computação Forense. *

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Quesitos de Solicitação - Quesitos apresentados pela autoridade solicitante, dentro de um contexto criminal, isto é, dentro de uma situação conhecida no bojo da investigação ou do processo penal que direciona a forma de examinar os materiais questionados. *

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Técnica Forense - Técnica empregada para realização das atividades previstas no processo de Computação Forense. *

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Perito Criminal - Profissional responsável pela materialização da prova do crime por meio do adequado exame e da interpretação correta dos vestígios materiais dessas infrações. *

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Autoridade Solicitante - Solicitante do Exame Pericial com competências previstas em lei. *

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

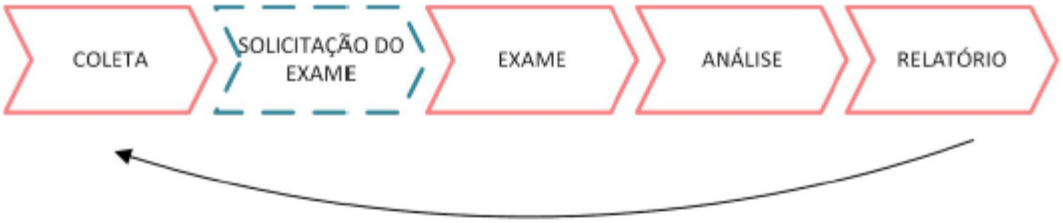
Caso Criminal - Caso que reuni os elementos e atributos associados a uma investigação criminal, os quais são necessários para realização de um exame pericial. *

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Crime - Crime é toda a ação, típica, antijurídica e culpável. Tendo como um dos elementos definidores o Tipo penal, sendo a descrição concreta da conduta proibida. *

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Fase do Processo Pericial - Fases do processo de Computação Forense que envolve os subprocessos de: coleta, solicitação, exame, análise e relatório. *



- Satisfaz
- Satisfaz Parcialmente
- Não Satisfaz

Fase de Solicitação do Exame - Na fase de solicitação do exame o investigador, autoridade, elabora o pedido de exames contendo quesitos orientados pela natureza dos materiais e pelo fato investigado.

- Satisfaz
- Satisfaz Parcialmente
- Não Satisfaz

Fase de Coleta - Na fase de coleta os materiais referentes a um evento específico são identificados, rotulados, registrados e recolhidos, preservando-os em relação a sua integridade

- Satisfaz
- Satisfaz Parcialmente
- Não Satisfaz

Fase de Exame - Processo execução de procedimentos periciais para obtenção de dados para análise.

- Satisfaz
- Satisfaz Parcialmente
- Não Satisfaz

Fase de Análise - A fase de análise envolve a avaliação dos resultados obtidos na fase de exame para alcançar informações relacionadas aos fatos e alvos. *

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Fase de Relatório - A fase de relatório tem por objetivo explicitar os resultados dos exames por meio do documento Laudo Pericial. *

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Crime - Crime é toda a ação, típica, antijurídica e culpável. Tendo como um dos elementos definidores o Tipo penal, sendo a descrição concreta da conduta proibida. *

- ◯ Satisfaz
- ◯ Satisfaz Parcialmente
- ◯ Não Satisfaz

Caso tenha respondido "Satisfaz Parcialmente" ou "Não Satisfaz", para alguma das questões apresentadas, qual contribuição você faria para auxiliar na melhoria da ontologia.

Long answer text

Modelo Conceitual da Ontologia

Conforme Modelo Conceitual da figura acima, algum relacionamento foi mapeado de forma equivocada? *

- Não
- Sim
- Não sei

Caso tenha respondido "Sim", qual ajuste você sugere?

Long answer text