# Ontologia aplicada no processo de Computação Forense

**Egberto Lemos Filho**[1], **Bruno Hoelz**[2], **Laís Salvador**[3]

[1]Departamento de Engenharia Elétrica - UNB

[2]Instituto Nacional de Criminalistica - DPF

[3]Departamento de Ciência da Computação - UFBA

{egberto.lemos@gmail.com, werneck.bwph@dpf.gov.br, laisns@ufba.br}

***Abstract.*** *This paper presents a proposal to develop an ontology-supported knowledge base in the field of Computer Forensics. In order to improve operational efficiency of forensic examinations, a common vocabulary between the actors must be defined, as well as for the knowledge associated with forensic tools and techniques applied to each type of digital evidence.*

***Resumo.*** *Este artigo apresenta uma proposta para desenvolvimento de uma base de conhecimento, suportada por ontologia, no domínio da Computação Forense. Para promover ganhos na eficiência operacional dos exames periciais, um vocabulário comum entre os atores deve ser definido, bem como para a geração de conhecimento associado às técnicas e ferramentas forenses aplicadas em cada tipo de evidência digital.*

## 1. Introdução

O crime digital relaciona-se, por meio das evidências digitais e do ambiente cibernético, com os mais diversos crimes cometidos na sociedade [Park et al. 2009]. Consequentemente, a perícia criminal possui uma área dedicada à aplicação da ciência e da engenharia na produção de provas associadas a tais crimes, chamada de Computação Forense [Ćosić and Ćosić 2012].

Na Computação Forense, os métodos e técnicas utilizados pelos peritos criminais nos exames periciais dependem do tipo da evidência digital, do local dos exames, da data do evento, das ferramentas disponíveis e, sobretudo, das questões investigativas a serem respondidas. Todas essas características geram inúmeras possibilidades de métodos e estratégias empregadas nos exames, o que torna necessário sistematizar o conhecimento sobre esses conceitos para tornar os processos de Computação Forense mais eficientes.

Este trabalho visa contribuir com a construção de uma base de conhecimento por meio de ontologia do domínio da Computação Forense, que está em fase final de desenvolvimento, partindo dos principais conceitos que envolvem o processo do exame pericial. O texto está organizado da seguinte maneira: a Seção 2 descreve o processo de exame pericial na Computação Forense, a Seção 3 apresenta trabalhos relacionados; a Seção 4 apresenta a proposta da ontologia do processo de Computação Forense, inclusive a metodologia adotada e a Seção 5 finaliza com as considerações parciais e trabalhos futuros.

## 2. O Processo de Exame Pericial na Computação Forense

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 [Garcia 2002]. Para a realização de exame pericial, quesitos devem ser formulados pelas partes, como se infere da leitura do artigo 176 do Código de Processo Penal brasileiro. 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".

A evolução e aumento da complexidade dos exames periciais em meios digitais não impacta apenas o trabalho do perito forense, mas em 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 Periciais 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 1. A etapa de solicitação de exame reflete o processo na maioria dos órgãos de Criminalística no Brasil, quando a autoridade envia a referida requisição. 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.

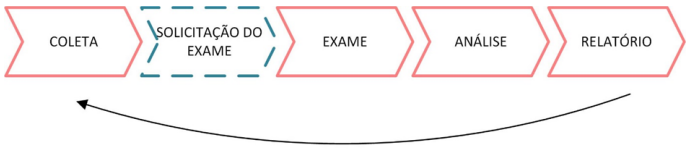


**Figura 1. Modelo de processo pericial, adaptado de [Kent et al. 2006]**

Em [Nonaka and Takeuchi 2004], os autores afirmam que as atividades criadoras de conhecimento são captadas e recontextualizadas na base de conhecimento da organização como um todo, tanto para os conhecimentos explícitos quanto para os tácitos. Logo, 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 et al. 1993] como: "Especificação explícita de uma conceitualização", ou seja, uma conceitualização compartilhada de um determinado domínio de conhecimento. A ontologia 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 (SPARQL)[1] promoverá o conhecimento compartilhado sobre crimes digitais e 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.

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

[1]SPARQL - Simple Protocol and RDF Query Language

- 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 et al. 2012];
- diversidade das possibilidades de técnicas e ferramentas que podem ser empregadas em um determinado exame pericial em dispositivos eletrônicos[Alzaabi et al. 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 and Ćosić 2012].

Diante do exposto, 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.

## 3. Trabalhos Relacionados

A multidisciplinaridade do campo da Computação Forense está refletida nos distintos trabalhos relacionados com ontologias sobre o tema. A investigação digital lida com uma massiva complexidade conceitual em várias camadas de abstração. Portanto, dificilmente existirá uma ontologia abrangente o suficiente para incluir todos os conceitos de interesse de cada investigador criminal [Huang et al. 2010].

Em [Alzaabi et al. 2013] é proposto um *framework* baseado em ontologias de domínio distintas para cada tipo de evidência digital conectadas por uma camada de extração de conceitos e relações.

Uma ontologia voltada sobretudo a desenvolver formação profissional na área dos crimes cibernéticos é proposta por [Brinson et al. 2006]. O modelo apresentado possui cinco níveis de estrutura hierárquica para o campo da Tecnologia, que está desmembrada em *Hardware* e *Software*.

Em [Park et al. 2009], os autores desenvolveram uma ontologia para o domínio da Computação Forense com o objetivo de utilizá-la nos processos de investigação dos casos criminais, e de aplicá-la na mineração de dados de crimes cibernéticos.

Os trabalhos sobre o tema de estudo supracitados refletem a grande abrangência e as possíveis abordagens do campo da Computação Forense. Enquanto os dois primeiros, buscam representar de forma mais sistêmica o campo dos crimes digitais, inclusive os seus processos, o terceiro em forma de framework aborda os conceitos e conhecimentos mais específicos referente aos dispositivos digitais. As ontologias desses trabalhos não estavam disponíveis para reutilização, mas os conceitos gerais da ontologia proposta neste artigo derivam das ideias dos seus respectivos autores. Cabe salientar que além desses conceitos gerais da Computação Forense, a ontologia proposta introduz o aspecto da comunicação dos demandantes dos exames, conforme apresentado na Figura 1.

## 4. Ontologia do Processo de Computação Forense

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

A Figura 2, representa o domínio do conhecimento a ser modelado pela ontologia. As três principais áreas do conhecimento que integram a Computação Forense: "Exame Pericial", "Crimes Digitais" e "Evidência Digital", fornecem informações para especificação dos conceitos e relações.

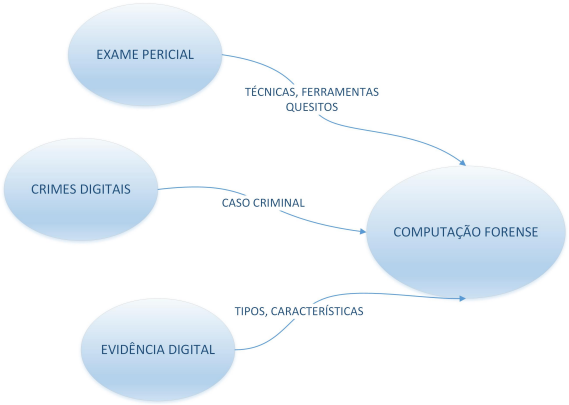


**Figura 2. Domínio do conhecimento a ser modelado pela ontologia**

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 junto com informações das técnicas e ferramentas aplicáveis nos exames, com suas respectivas relações, podem servir, também, como referência para guiar o perito no processo da Figura 1 . 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.

### 4.1. Metodologia

Após a definição dos objetivos da ontologia de Computação Forense pretendida, escolheu-se a metodologia METHONTOLOGY [Fernández-López 1999], por apresentar um método bem estruturado para a representação e detalhamento da ontologia, acrescida do tópico preliminar sobre as "Questões de Competência" do processo Ontology Development 101 [Noy et al. 2001], que serviram como guia para o refinamento do seu escopo. Na Tabela 1 são apresentadas algumas das questões de competência sobre seu respectivo conceito.

### 4.2. Proposta de Ontologia - Processo de Computação Forense

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

**Tabela 1. Questões de Competência**

| CONCEITO | QUESTÕES |
|---|---|
| Dada uma Evidência coletada de um Caso Criminal pergunta-se: | Quais Técnicas Forenses podem ser aplicadas em uma determinada Evidência?<br>Quais Crimes envolvidos com um determinado tipo de Evidência?<br>Existe incompatibilidade nos Quesitos da Solicitação propostos? |
| Dado um Quesito de Solicitação de exame, pergunta-se: | Quais Quesitos da Solicitação não podem ser respondidos para determinada Evidência?<br>Quais Quesitos da Solicitação podem ser formulados para o exame de uma Evidência envolvida em um determinado Crime? |
| Dada uma ferramenta pericial, pergunta-se: | A Ferramenta Forense aplica-se a determinada Técnica Forense?<br>A Ferramenta Forense tem sido utilizada em uma determinada Evidência? |

Computação Forense. O diagrama da Figura 3 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 as 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".

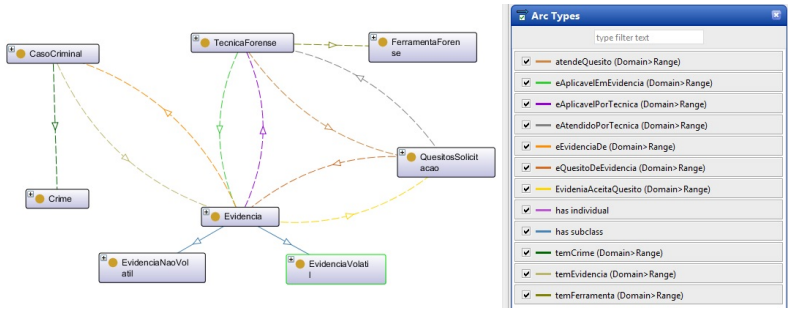


**Figura 3. Domínio do conhecimento a ser modelado pela ontologia**

A criação da ontologia foi feita com o apoio da ferramenta de edição de ontologias Protégé[2]. As informações sobre classes e relações coletadas dos documentos periciais são armazenadas em uma planilha no formato Microsoft Excel, que posteriormente servirá de fonte de dados para popular um repositório RDF criado a partir da ontologia.

## 5. Considerações Parciais

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 de 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.

Posteriormente, deseja-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[3], que provê a formalização da parte estrutural dos diagramas de processo de negócio (*Business Process Diagrams* –

[2]http://protege.stanford.edu/
[3]https://dkm.fbk.eu/bpmn-ontology/

BPD) [Di Francescomarino et al. 2008]. Outra proposta de aperfeiçoamento do modelo é implantar a conversão automática dos dados coletados no formato Microsoft Excel para um repositório RDF com o apoio da ferramenta Karma [Knoblock et al. 2012].

## Referências


Alzaabi, M., Jones, A., and Martin, T. A. (2013). An ontology-based forensic analysis tool. In *Proceedings of the Conference on Digital Forensics, Security and Law*, page 123. Association of Digital Forensics, Security and Law.

Brinson, A., Robinson, A., and Rogers, M. (2006). A cyber forensics ontology: Creating a new approach to studying cyber forensics. *digital investigation*, 3:37–43.

Ćosić, J. and Ćosić, Z. (2012). The necessity of developing a digital evidence ontology. In *the proceedings of the Central European Conference on Information and Intelligent Systems*, pages 325–330.

Di Francescomarino, C., Ghidini, C., Rospocher, M., Serafini, L., and Tonella, P. (2008). Reasoning on semantically annotated processes. In *International Conference on Service-Oriented Computing*, pages 132–146. Springer.

Fernández-López, M. (1999). Overview of methodologies for building ontologies.

Garcia, I. E. (2002). Inquérito: procedimento policial. *Goiânia: AB*.

Gruber, T. R. et al. (1993). A translation approach to portable ontology specifications. *Knowledge acquisition*, 5(2):199–220.

Huang, J., Yasinsac, A., and Hayes, P. J. (2010). Knowledge sharing and reuse in digital forensics. In *SADFE*, pages 73–78.

Kent, K., Chevalier, S., Grance, T., and Dang, H. (2006). Guide to integrating forensic techniques into incident response. *NIST Special Publication*, pages 800–86.

Knoblock, C. A., Szekely, P., Ambite, J. L., Goel, A., Gupta, S., Lerman, K., Muslea, M., Taheriyan, M., and Mallick, P. (2012). Semi-automatically mapping structured sources into the semantic web. In *Extended Semantic Web Conference*, pages 375–390. Springer.

Nonaka, I. and Takeuchi, H. (2004). *Criação de conhecimento na empresa*. Elsevier Brasil.

Noy, N. F., McGuinness, D. L., et al. (2001). Ontology development 101: A guide to creating your first ontology.

Park, H., Cho, S., and Kwon, H.-C. (2009). Cyber forensics ontology for cyber criminal investigation. In *International Conference on Forensics in Telecommunications, Information, and Multimedia*, pages 160–165. Springer.

Swartout, B., Patil, R., Knight, K., and Russ, T. (1996). Toward distributed use of large-scale ontologies. In *Proc. of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems*, pages 138–148.

Velho, J. A., Geiser, G. C., and Espindula, A. (2012). Ciências forenses-uma introdução às principais áreas da criminalística moderna. *Campinas: Millennium*.