# Análise de Métodos para o Tratamento de Arquivos Falso-Positivos a partir de Ferramentas de Recuperação de Dados Digitais: Uma Revisão Sistemática da Literatura

**Eric Pereira**[1], **William Silva**[1], **Sandro Bezerra**[1], **Josivaldo Araújo**[1]

[1]Faculdade de Computação - Universidade Federal do Pará (UFPA)
Rua Augusto Corrêa, 01 - CEP 66075-110 - Belém - PA - Brasil

{eric.cbcc, williamsawaki}@gmail.com, srbo@ufpa.br, josivaldo@ufpa.br

***Abstract.*** *The development of technology has generated information that needs storage in the form of digital files. Therefore, the secondary devices are needed it. However, these files can be deleted intentionally or by accident, motivating their recovery. In the process of retrieving data in digital media, the so-called false-positive files are also generated in addition to retrieving the desired files. This paper aims to present the results of a Systematic Review of Literature on digital data recovery through analysis of tools used in this process as well as techniques and methods used in the treatment of false-positive files.*

***Resumo.*** *O desenvolvimento da tecnologia tem gerado informações que precisam de armazenamento na forma de arquivos digitais, necessitando, assim, de dispositivos secundários. No entanto, esses arquivos podem ser deletados de forma intencional ou por acidente, motivando a sua recuperação. No processo de recuperação de dados em mídias digitais, além de recuperar os arquivos desejados, também são gerados arquivos denominados de falso-positivos. Este trabalho tem o objetivo de apresentar os resultados de uma Revisão Sistemática da Literatura sobre a recuperação de dados digitais por meio de análise de ferramentas utilizadas nesse processo, bem como de técnicas e métodos usados no tratamento dos arquivos falso-positivos.*

## 1. Introdução

Com o avanço da tecnologia, ter um dispositivo móvel interligado à Internet ficou cada vez mais fácil. Dados divulgados pela Agência Nacional de Telecomunicações mostram que o Brasil registrou, em dezembro de 2018, 229 milhões de linhas móveis [Anatel 2019], ou seja, o país possui uma população de celulares maior que o número de habitantes. Essa facilidade de acesso faz com que imagens e vídeos possam ser gerados em tempo real e transmitidos via aplicativos de mensagens ou redes sociais. No entanto, com essa crescente atividade de gerar e receber arquivos, cresce também a necessidade de se armazenar cada vez mais mídias digitais.

Dispositivos de armazenamentos são fundamentais em sistemas de computação e, por conta disso, as informações armazenadas podem ser removidas de forma proposital, ou acidental. Contudo, na maioria das vezes, essas remoções são somente lógicas e não físicas, fazendo com que as informações, logicamente removidas, possam ser recuperadas a partir da análise do meio físico [Weber and Zorzo 2017].

A análise de dispositivos de armazenamento é uma das bases da Computação Forense, onde peritos utilizam softwares específicos [Nurhayati and Fikri 2017] para recuperar informações que foram removidas de mídias digitais sob investigação, com o intuíto de ocultar rastros de crimes que foram praticados pela Internet como, por exemplo, a pedofilia [Moreira and Fechine 2018], ou apagar arquivos que contenham alguma informação que comprometa os investigados. Isso se deve ao fato de que, em um processo de investigação criminal, deve-se considerar a utilização de métodos e protocolos bem definidos, visando a análise eficaz e ampla aceitação na esfera jurídica dos dados recuperados [Jesus and Couto 2017].

No entanto, quando é realizado um processo de recuperação de dados digitais, podem ser gerados, além dos arquivos reais, arquivos que são incorretos ou corrompidos, denominados de falso-positivos [Laurenson 2013]. Quanto maior a quantidade de falso-positivos ao final do procedimento de recuperação, maior será o esforço e o tempo necessários para separá-los dos arquivos reais, adicionando, dessa forma, dificuldades no trabalho de pesquisadores e profissionais da área Forense.

Assim, o objetivo deste trabalho é, por meio de uma Revisão Sistemática da Literatura (RSL), pesquisar e avaliar soluções propostas para o tratamento dos arquivos falso-positivos na recuperação de dados em mídias digitais, pois, segundo [Kitchenham 2007], por meio de uma RSL é possível identificar, avaliar e explanar as pesquisas relevantes para uma área ou para um problema específico.

Este artigo foi organizado da seguinte maneira: os trabalhos relacionados estão descritos na Seção 2; os procedimentos metodológicos da Revisão Sistemática da Literatura estão na Seção 3; em seguida, na Seção 4 estão descritos os resultados obtidos com a RSL, com o propósito de responder as questões principal e secundárias de pesquisa; e, por fim, na Seção 5 são apresentadas as conclusões e os trabalhos futuros.

## 2. Trabalhos Relacionados

Estudos sobre recuperação de dados não são novos, porém raros são os estudos que se preocupam com algum tipo de tratamento para os arquivos falso-positivos gerados.

O estudo de [Ninahualpa et al. 2018] consiste na restauração de dados em dispositivos de estado sólido (SSD) utilizando a técnica de *File Carving*. Os autores fazem uma breve explicação sobre as diferentes classificações das técnicas utilizadas, além de listarem algumas ferramentas de recuperação de dados. O trabalho relata, ainda, uma análise das diferentes técnicas de recuperação em dispositivos SSD afetados por quedas e, ao final, obtém resultados categorizados por métricas, sendo uma delas a quantidade de falso-positivos.

O trabalho de [Karresand and Shahmehri 2016] tem uma preocupação com o número de arquivos falso-positivos gerados, onde destaca uma taxa de falso-positivos entre 0.5% e 1.9%, porém todos os algoritmos e testes executados utilizam apenas arquivos no formato JPEG, não tratando outros formatos.

No trabalho de [Laurenson 2013] pode-se notar que o número de arquivos falso-positivos gerados após o processo de recuperação de dados é bastante elevado e, dependendo da ferramenta utilizada, esse número pode ser ainda maior. Para a área Forense essa é uma preocupação bastante relevante, pois a recuperação de dados é algo que os

profissionais da área lidam diariamente, seja recuperando fotos ou documentos em discos rígidos, ou outro tipo de armazenamento não volátil.

O trabalho de [Alherbawi et al. 2013] consiste de uma Revisão Sistemática da Literatura (RSL), onde suas questões de pesquisa são sobre a medição da qualidade de ferramentas *Data Carving*, as diferentes técnicas de *Carving* e as principais dificuldades encontradas pelos pesquisadores nessa área. As questões de pesquisa tratam de assuntos diferentes das que serão tratadas nesse artigo, onde se estará focado nos arquivos falso-positivos.

Além da geração de arquivos falso-positivos no contexto da recuperação de dados em mídias digitais, também pode-se ter esse tipo de arquivos no contexto de redes. Em [Beverly et al. 2011] são utilizadas técnicas de *Carving* para a recuperação de pacotes em uma rede de computadores.

Apesar de todos os estudos, citados anteriormente, tratarem da geração de arquivos falso-positivos no processo de recuperação de dados, este trabalho se diferencia dos demais, pois realiza uma Revisão Sistemática da Literatura direcionando, as suas questões de pesquisa, no sentido de encontrar técnicas e métodos que possam reduzir a ocorrência desses tipos de arquivos durante todo o processo de recuperação, independente do formato dos arquivo a ser recuperado.

## 3. Revisão Sistemática da Literatura

Uma Revisão Sistemática da Literatura (RSL) consiste em um estudo secundário, por utilizar como base estudos primários previamente publicados. Assim, pode ser feita a integração de diversos estudos experimentais, de forma a comparar seus resultados, visto que nenhum estudo individualmente produzido pode ser considerado definitivo [Mafra and Travassos 2006], sendo necessária a confirmação de resultados obtidos a partir da análise de um número maior de estudos.

A principal meta de uma RSL é realizar uma pesquisa exaustiva na literatura em busca de evidências que possam apoiar uma determinada hipótese, ou simplesmente a busca por conhecimento aprofundado acerca de certo fenômeno de interesse. Para isso, a Revisão Sistemática da Literatura faz uso de estudos previamente publicados e validados pertinentes ao tópico de interesse: os estudos primários são estudos de natureza experimental que envolvem hipóteses e resultados obtidos com pesquisas e experimentação a partir de diferentes métodos, como *surveys*, estudo de caso e experimentos [Mafra and Travassos 2006].

Diante disso, a questão levantada para este trabalho foi organizada conforme a estrutura *Population*, *Intervention*, *Context*, *Outcomes*, *Comparison* (PICOC), recomendada por [Kitchenham 2007]. Entretanto, apenas os itens População, Intervenção, Contexto e Resultados foram considerados relevantes para a pesquisa. Tal restrição, segundo [Santos 2010], caracteriza esta pesquisa como uma Revisão QUASI Sistemática da Literatura.

Na questão de pesquisa, objetiva-se identificar abordagens utilizadas por profissionais e pesquisadores da área Forense (População) para apoiar as atividades de Recuperação de Dados (Intervenção), na qual busca-se encontrar métodos, técnicas e práticas utilizadas (Contexto) para a recuperação de dados (Resultados).

### 3.1. Procedimentos Metodológicos

O processo de uma RSL visa encontrar e analisar os possíveis estudos que são capazes de responder as questões de pesquisa formuladas [Paviani et al. 2018]. Para identificar e selecionar estes estudos, foi definido e executado um protocolo de busca. O protocolo de busca definido visa responder a seguinte questão de pesquisa:

- (QP1) Quais são as abordagens existentes para o tratamento de arquivos falso-positivos que apoiam as atividades de Computação Forense no contexto da recuperação de dados, utilizando técnicas computacionais?

Um conjunto de questões secundárias, referentes à questão principal, foram estabelecidas e respondidas durante a fase de extração de informações. Tais questões têm o objetivo de esclarecer detalhes importantes que esta revisão procura identificar, e, com isso, colaborar com o projeto onde este se insere. Desta forma, as questões secundárias definidas são:

- QS1. Quais atividades da Computação Forense são apoiadas pelas abordagens encontradas?
- QS2. Caso a abordagem seja uma prática ou técnica, a mesma pertence a algum método computacional? Qual?
- QS3. Qual o contexto de aplicação da abordagem encontrada?
- QS4. Quais as técnicas e ferramentas usadas para a recuperação de dados no contexto da Computação Forense?
- QS5. Quais os tipos de dispositivos de armazenamento secundários foram envolvidos nestas abordagens?
- QS6. Quais os tipos de sistemas de arquivos envolvidos nestas abordagens?

### 3.2. *String* de Busca

As variações das *strings* de busca foram adaptadas segundo as máquinas de buscas de cada fonte de pesquisa. As variações utilizadas, com base nas palavras-chave definidas, podem ser acessadas no endereço: http://bit.ly/2V0kx8h.

### 3.3. Processo de Busca

Para responder as questões (principal e secundárias) levantadas neste trabalho, foram utilizados trabalhos publicados entre 01 de janeiro de 1995 e 31 de dezembro de 2017. Este período foi estabelecido a partir de uma informação relevante, uma vez que, em 1995 foi criada a Organização Internacional de Evidências de Computador (IOCE) para a troca de informações sobre investigação criminal e outras questões Forenses. Vale ressaltar que esta revisão foi realizada no período de julho de 2018 a janeiro de 2019, por isso, não contemplou os trabalhos de 2018, uma vez que durante a execução, este ano ainda não havia sido finalizado.

No protocolo desenvolvido para esta Revisão Sistemática foi definido que trabalhos publicados em inglês e em português seriam considerados, e o mesmo foi executado considerando as seis principais bases científicas: ACM Digital Library, Ei Compendex, IEEEXplore Digital Library, Science Direct, Scopus e ISI Web of Knowledge. A Tabela 1 apresenta o total de publicações retornadas de cada uma das seis bases pesquisadas, representando um total de 11.123 trabalhos.

**Tabela 1. Resultado de Publicações Retornadas por Base Pesquisada**

| Fontes | Estudos Retornados |
|---|---|
| ACM Digital Library | 2045 |
| Ei Compendex | 3100 |
| IEEEXplore Digital Library | 1064 |
| Science Direct | 2749 |
| Scopus | 743 |
| ISI Web of Knowledge | 1422 |

### 3.4. Critérios de Inclusão e Exclusão

Os critérios de inclusão e exclusão servem para avaliar a qualidade de um artigo científico e, assim, criar uma lista de possíveis artigos primários e outra com os artigos excluídos.

Inicialmente, foram retornados 11.123 artigos e após uma filtragem inicial para a remoção dos artigos repetidos restaram 7.730 estudos. Tais artigos tiveram seus títulos, resumos e palavras-chave lidos e, após esta etapa, restaram um total de 154 artigos. Tais estudos foram declarados potencialmente relevantes, pois passariam pelos critérios de inclusão e exclusão.

Posteriormente, restaram apenas 107 artigos incluídos, sendo que estes passaram pelos critérios de qualidade e pela etapa de extração de dados e, consequentemente, tiveram sua leitura feita em sua totalidade. Os critérios utilizados para inclusão, foram:

- CI-01: Estudos que apresentem primária ou secundariamente abordagens, no contexto das técnicas computacionais, que dão apoio às atividades de recuperação de dados;
- CI-02: Estudos que apresentem relatos de experiência da indústria, ou pesquisas de caráter experimental ou teórico, contanto que apresentem exemplos de aplicação, descrição de experimentos ou casos reais de uso de abordagens, no contexto da Computação Forense, para apoio às atividades de recuperação de dados.

Enquanto que na terceira etapa, artigos não relevantes, duplicados, inacessíveis ou em outro idioma que não fossem em português ou inglês foram excluídos. Os critérios que foram utilizados para excluir os artigos foram:

- CE-01: Estudos que não estejam disponíveis livremente para consulta ou *download* (em versão completa) a partir das fontes de pesquisa ou a partir das ferramentas de busca Google (http://www.google.com.br/) e/ou Google Scholar (http://scholar.google.com.br/);
- CE-02: Estudos que claramente não atendam as questões de pesquisa;
- CE-03: Estudos repetidos (em mais de uma fonte de busca) terão apenas sua primeira ocorrência considerada;
- CE-04: Estudos enquadrados como resumos, *keynote speeches*, cursos, tutoriais, *workshops* e afins;
- CE-05: Estudos que não mencionem as palavras-chave da pesquisa no título, resumo ou nas palavras-chave do artigo, salvo trabalhos que abordem o processo de recuperação de dados, nos quais seja observada possibilidade da Computação Forense e técnicas computacionais a serem tratados ao longo do trabalho;

- CE-06: Estudos que não estejam inseridos no contexto da Computação Forense;
- CE-07: Estudos que não sejam apresentado nas linguagens aceitas (Português e Inglês);

A Tabela 2, apresenta o número de trabalhos excluídos por cada um dos critérios.

**Tabela 2. Artigos Excluídos por Critério de Exclusão**

| Critério | Trabalhos Excluídos |
|---|---|
| CE-01 | 14 |
| CE-02 | 1721 |
| CE-03 | 31 |
| CE-04 | 78 |
| CE-05 | 132 |
| CE-06 | 5645 |
| CE-07 | 0 |

### 3.5. Processos de Avaliação dos Estudos Primários

Os estudos primários selecionados foram lidos em totalidade e, então, avaliados quanto aos critérios de qualidade. Para avaliar o grau de adequação aos critérios de qualidade foi adotada uma estratégia de avaliação semelhante à proposta por [Costa 2010], onde se utiliza a escala de Likert-5, permitindo respostas gradativas de 0 (discordo totalmente) a 4 (concordo totalmente). Porém, neste trabalho, foi utilizada a escala Likert-3, permitindo respostas gradativas de 0 (discordo totalmente), 1 (neutro) e 2 (concordo totalmente), pois com menos itens de Likert os critérios tornam-se menos subjetivos.

## 4. Resultados

Nesta seção será apresentada a condução das análises dos resultados dos estudos primários, visando responder as questões de pesquisa.

### 4.1. Avaliação da Qualidade dos Estudos Primários

A nota foi calculada baseada nos atributos avaliados nos critérios de qualidade e na escala Likert-3, que representa a adesão destes atributos aos critérios de qualidade.

Realizada a análise, poucos trabalhos ficaram na faixa Baixa, apenas 2 estudos (1,87%), 23 estudos (21,50%) estão na faixa Média, enquanto 24 estudos (22,4%) estão na faixa Boa, 46 estudos (43%) estão na faixa Muito Boa e 12 estudos (11,2%) na faixa Excelente. A lista dos trabalhos incluídos e selecionados, encontra-se disponível através do link: http://bit.ly/2JcwTJ1.

### 4.2. Extração dos Dados

Esta etapa consiste em organizar os dados extraídos para apresentação dos gráficos que serviram como panorama geral e base para futuras análises. A base para as respostas às questões de pesquisa faz parte da lista dos trabalhos incluídos e selecionados na RSL, que se encontra acessando o link: http://bit.ly/2JcwTJ1.

### 4.3. Respostas às Questões de Investigação ou de Pesquisa

Esta seção responde a questão principal levantada no protocolo desta Revisão Sistemática da Literatura como meio de investigar o estado da arte no âmbito de encontrar técnicas ou métodos que tratem os arquivos falso-positivos durante o processo de recuperação de dados em mídias digitais. A seguir são apresentadas as respostas das questões de pesquisa e as considerações dos autores a respeito delas, considerando os códigos do trabalhos selecionados disponíveis em: http://bit.ly/2JcwTJ1.

- (QP1) Quais são as abordagens existentes para o tratamento de arquivos falso-positivos que apoiam as atividades de Computação Forense no contexto da recuperação de dados utilizando técnicas computacionais?

Para um trabalho de recuperação de dados demandar menor tempo e esforço da parte do perito, é necessário que haja a menor quantidade possível de arquivos falso-positivos em seus resultados. Para que isso aconteça de uma forma eficaz, é imprescindível que a ferramenta utilizada gere a menor quantidade possível de falso-positivos ou então que tenha um pós-processamento confiável para que seja feita a exclusão automática desses arquivos.

Com isso, durante a etapa de extração de dados foi encontrado apenas um artigo que relata algum tratamento dos arquivos falso-positivos. O artigo [EP19] procura reduzir a quantidade de falso-positivos a partir de uma tabela, onde se define tamanhos máximos de arquivos para a recuperação. O autor afirma que a taxa de falsos-positivos é bastante reduzida se forem escolhidos os tamanhos de arquivos apresentados no trabalho.

Porém, essa é uma técnica que não se pode generalizar para outras pesquisas, visto que, na maioria dos processos de recuperação de dados, não é possível escolher o tamanho dos arquivos a serem recuperados, obrigando o profissional a trabalhar com tamanhos variados de arquivos.

- (QS1) Quais atividades da Computação Forense são apoiadas pelas abordagens encontradas?

Em todos os artigos pesquisados a atividade da Computação Forense apoiada é o *Data Carving*, ou também chamada de *File Carving*. Essa atividade baseia-se em uma técnica Forense para a recuperação de arquivos com base apenas na estrutura e conteúdo do mesmo, sem a necessidade de utilizar metadados [Pal and Memon 2009].

- (QS2) Caso a abordagem seja uma prática ou técnica, a mesma pertence a algum método computacional? Qual?

A abordagem encontrada no estudo [EP19], que trabalha para reduzir a quantidade de arquivos falso-positivos, não pertencente a nenhum método computacional como era esperado, dado que a abordagem apenas limita um tamanho máximo para o arquivo a ser recuperado.

- (QS3) Qual o contexto de aplicação da abordagem encontrada?

Grande parte dos trabalhados selecionados tem o contexto apenas voltado para fins acadêmicos. Apenas o artigo [EP19] tem o contexto direcionado para perícia de smartphone (iPhone) e o artigo [EP106] com o contexto direcionado para perícia em dispositivos USB (Pendrive).

- (QS4) Quais as técnicas e ferramentas usadas para a recuperação de dados no contexto da computação forense?

Há inúmeras ferramentas no mercado com o propósito de recuperar dados digitais, algumas delas são proprietárias, mas existem outras de código livre, cada uma com o seu método e sua particularidade, e dependendo da necessidade do usuário ou perito, pode-se escolher uma dentre essa variedade de opções. A partir da RSL foram identificadas 31 diferentes ferramentas utilizadas nos trabalhos para recuperação de dados.

A aplicação, bem como, os resultados produzidos por essas ferramentas, podem ser consultados em: Scalpel [EP1, EP19, EP32, EP40, EP46, EP48, EP49, EP58, EP75, EP76, EP83, EP85]; FTK [EP1, EP24, EP39, EP45, EP48, EP56, EP64, EP66, EP88]; Foremost [EP1, EP45, EP48, EP58, EP62, EP73, EP84, EP91]; WinHex [EP1, EP17, EP24, EP45, EP48, EP50, EP63, EP78]; e PhotoRec [EP39, EP45, EP72, EP73, EP75, EP79, EP85, EP91].

Vale ressaltar que nos estudos retornados foram detectadas também 19 ferramentas próprias, ou seja, desenvolvidas pelos próprios autores. Essas ferramentas podem ser consultadas nos trabalhos: [EP21, EP23, EP25, EP33, EP58, EP59, EP62, EP64, EP65, EP73, EP74, EP75, EP76, EP78, EP88, EP94, EP97, EP104]. No entanto, nenhuma dessas ferramentas próprias realizam algum tratamento dos falso-positivos, apenas recuperam dados.

Outras ferramentas foram citadas apenas uma única vez, como é o caso da: *Revivelt, Cohens PDF File Carving Method, Blade, Disktools, Simple File Carver, Phone Image Carver, Physical Analyser, SafeCopy, Revit, Recuva, Final Data, Gddrescue, FastScalpel, EseCarve, Defraser, Data Recovery, bulk Extractor, Vicarve, Thumb DB Viewer, DECODE, Cellebrite, Autopsy*.

- (QS5) Quais os tipos de dispositivos de armazenamento secundários foram envolvidos nestas abordagens?

A maior parte dos estudos retornados utiliza como mídia, em seus experimentos, o HD [EP17, EP24, EP32, EP58, EP79], por ainda ser um tipo de armazenamento comum e financeiramente mais viável, utilizado, por exemplo, em computadores desktop e notebooks. Mas, também, foram encontrados experimentos utilizando pen drives [EP32, EP41, EP56, EP84], cartão SD [EP47, EP77] e, mais recentemente, um aumento no uso de equipamentos com dispositivos SSDs [EP49, EP58, EP99, EP100], apesar da recuperação de dados, neste dispositivo, ser mais complexa.

Entretanto, pôde-se verificar experimentos em *Smartphones*, que apesar de utilizarem em sua grande maioria memória *flash*, acabam se diferenciando pelos diferentes sistemas operacionais, como o Android [EP30, EP31, EP83, EP101, EP105], iOS [EP19, EP40] e Windows Phone [EP01].

- (QS6) Quais os tipos de sistemas de arquivos envolvidos nestas abordagens?

Percebe-se, a presença de 14 sistemas de arquivos diferentes encontrados nos estudos selecionados. Apesar da grande variação de sistemas de arquivos, o mais utilizado em testes é o NTFS [EP11, EP20, EP41, EP72, EP75, EP84], sistema de arquivo principal do sistema operacional Windows. O segundo sistema de arquivo mais utilizado é o ext4 [EP13, EP16, EP27, EP30], sistema de arquivo da plataforma Linux.

## 5. Conclusão

O objetivo desta Revisão Sistemática da Literatura foi de investigar ferramentas utilizadas na área da Computação Forense, com ênfase no processo de recuperação de dados, em busca de métodos que fizessem algum tipo de tratamento, ou pelo menos, diminuíssem, a quantidade de arquivos falso-positivos que são gerados após um processo de recuperação de dados em mídias digitais. O propósito era identificar abordagens, técnicas e/ou ferramentas, para então catalogar e, valendo-se destas informações, realizar uma análise das vantagens e desvantagens de cada solução e, assim, propor o desenvolvimento de uma ferramenta que tivesse como prioridade a redução, ou completa eliminação de arquivos falso-positivos.

No entanto, observou-se a pouca, ou quase inexistente exploração desse assunto nos estudos, até então desenvolvidos na Computação Forense, mesmo diante da crescente demanda em recuperação de dados digitais. Contudo, especificamente sobre o tratamento de arquivos falsos-positivos, foi encontrado, durante o processo, apenas um único trabalho (como relatado na QP1), entre os 107 estudos primários pautados, exclusivamente, no processo de recuperação de dados, evidenciando que o referido processo ainda é insuficientemente explorado, o que encoraja uma maior investigação e a condução de estudos experimentais para desenvolver propostas que tenham como meta o tratamento deste tipo de problema.

Como trabalhos futuros, pretende-se conduzir estudos onde se possam apresentar propostas para o tratamento de arquivos falsos-positivos, seja combatendo a sua geração, ou analisando de forma eficiente os arquivos recuperados, com a possibilidade de uma exclusão automática após o processo de recuperação. Para isso, irá se utilizar os padrões de mídias e sistemas de arquivos levantados neste trabalho, além de uma melhor compreensão do funcionamento das ferramentas mais empregadas no processo de recuperação de dados, aqui listadas.

## Referências


Alherbawi, N., Shukur, Z., and Sulaiman, R. (2013). Systematic literature review on data carving in digital forensic. *Procedia Technology*, pages 86–92.

Anatel (2019). Agência Nacional de Telecomunicações. `http://www.anatel.gov.br/dados/acessos-telefonia-movel`. Online; acessado em 25 de Fevereiro de 2019.

Beverly, R., Garfinkel, S., and Cardwell, G. (2011). Forensic carving of network packets and associated data structures. *Digital Investigation*, pages S78–S89.

Costa, C. S. (2010). Uma abordagem baseada em evidências para o gerenciamento de projetos no desenvolvimento distribuído de software. *Dissertação de Mestrado – Programa de Pós-Graduação em Ciência da Computação, Universidade Federal de Pernambuco, Recife - PE.*

Jesus, C. O. and Couto, M. S. (2017). Forensicwork: Desenvolvimento de um framework para perícia forense em computação na nuvem. *XVII Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais, 6 a 9 de Novembro de 2017, Brasília - DF*, pages 596–601.

Karresand, M. and Shahmehri, N. (2016). File type identification of data fragments by their binary structure. *Proceedings of the IEEE Information Assurance Workshop*, pages 140–147.

Kitchenham, B. (2007). Guidelines for performing systematic literature reviews in software engineering. *Technical Report EBSE 2007-001, Keele University and Durham University Joint Report*.

Laurenson, T. (2013). Performance analysis of file carving tools. *IFIP International Information Security Conference*, pages 419–433.

Mafra, S. and Travassos, G. (2006). Estudos primários e secundários apoiando a busca por evidencia em engenharia de software. *Relatório Técnico: RT-ES-687/06 – Programa de Engenharia de Sistemas e Computação - COPPE/UFRJ – Rio de Janeiro-RJ*.

Moreira, D. C. and Fechine, J. M. (2018). A forensic nudity detector based on machine learning. *XVIII Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais, 22 a 25 de Outubro de 2018, Natal - RN*, pages 267–280.

Ninahualpa, G., Perez, C., Yoo, G. S., Guarda, T., Diaz, J., and Piccirilli, D. (2018). Restoring data in solid state devices damaged by crushing and falling, using file carving technique. *13th Iberian Conference on Information Systems and Technologies (CISTI), 2018*, pages 1–4.

Nurhayati and Fikri, N. (2017). The analysis of file carving process using photorec and foremost. *4th International Conference on Computer Applications and Information Processing Technology (CAIPT), August, 8-10, 2017, Bali-Indonesia*, pages 1–6.

Pal, A. and Memon, N. (2009). The evolution of file carving. *IEEE Signal Processing Magazine, 26(2)*, pages 59–71.

Paviani, O., Adriano, D. D., and Wangham, M. S. (2018). Revisão sistemática da literatura sobre autenticação anônima em redes veiculares. *Computer on the Beach, 22 a 24 de Março de 2018, Florianópolis-SC*, pages 170–179.

Santos, G. (2010). Revisão sistemática, mini-curso. *IX Simpósio Brasileiro de Qualidade de Software, 7 a 11 de Junho de 2010, Belém-PA*.

Weber, J. S. and Zorzo, A. F. (2017). Eliminação segura de arquivos em memória não-volátil. *XVII Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais, 6 a 9 de Novembro de 2017, Brasília - DF*, pages 56–69.