Politécnico de Leiria
Escola Superior de Tecnologia e Gestão
Departamento de Engenharia Informática
Mestrado em Cibersegurança e Informática Forense

AUTOPSY – ENHANCED DISTRIBUTED FORENSIC ANALYSIS

PEDRO HENRIQUE GASPAR CORDEIRO FERREIRA

Leiria, Fevereiro de 2020

[January 16, 2020 at 10:51 – version 1.0]

Politécnico de Leiria
Escola Superior de Tecnologia e Gestão
Departamento de Engenharia Informática
Mestrado em Cibersegurança e Informática Forense

AUTOPSY – ENHANCED DISTRIBUTED FORENSIC ANALYSIS

PEDRO HENRIQUE GASPAR CORDEIRO FERREIRA
Número: 2180078

Relatório de estágio realizado sob orientação da Professora Doutora Marisa da Silva Maximiano (marisa.maximiano@ipleiria.pt).

Leiria, Fevereiro de 2020

[January 16, 2020 at 10:51 – version 1.0]

ACKNOWLEDGEMENTS

I would like to express appreciation for the opportunity provided by “VOID SOFTWARE, S.A.“ to work in an enterprise environment. I’d also like to thank my faculty advisor “Marisa da Silva Maximiano“ who provided much needed help writing this report in LaTeX 2_ε. I’m also thankful that Antonio Branco took time off his work to help me with so many minor details in my CSS layouts, as well as everyone one else in the company who provided me with everything I needed to succeed in my internship.

RESUMO

TODO.

A B S T R A C T

TODO.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABC	A lista de acrónimos deve ficar ordenada alfabeticamente.
ADSL	Assimetric Digital Subscriber Line.
ASCII	American Standard Code for Information Interchange.
BIOS	Basic Input/Output System.
bit	Digito binário.
Byte	Unidade de informação digital composta por oito bits.
CODEC	COmpression/DECompression.
CPU	Central Processing Unit.
DHCP	Dynamic Host Configuration Protocol.
DLL	Dynamic Link Library.
DNS	Domain Name System.
FTP	File Transfer Protocol.
IP	Internet Protocol.
ISP	Internet Service Provider.
SO	Sistema Operativo.
TCP	Transmission Control Protocol.

1

INTRODUCTION

The scope of this internship concerns digital forensics, as it focuses on adapting an existing forensics platform into a collaborative client-server model.

In this chapter, a contextualization of digital forensics is given, an analysis of both The Sleuth Kit and Autopsy is made, it's given a brief description of the existing forensic platform alternatives and is described the proposed solution for the scope of the internship.

1.1 DIGITAL FORENSICS

1.1.1 *Contextualization*

Forensic science is the use of scientific methods or expertise to investigate crimes or examine evidence that might be presented in a court of law.

The definition of digital forensics is directly related to the definition of computer forensics which is the collection, preservation, analysis, and presentation (**daniels**) of evidence stemming from digital sources for use in a legal matter using investigative processes, tools, and practices.

Digital forensics is the application of computer technology to criminal cases where evidence includes items that are created by digital systems.

Digital forensics, according to **nist**, is the field of forensic science that is concerned with retrieving, storing and analyzing electronic data that can be useful in criminal investigations. This includes information from computers, hard drives, mobile phones and other data storage devices.

Digital forensic investigators face challenges such as extracting data from damaged or destroyed devices, locating individual items of evidence among vast quantities of data, and ensuring that their methods capture data reliably without altering it in any way.

Personal data should ultimately be attributable to an individual; however, making that attribution can be difficult due to the presence or absence of individualized user accounts, security to protect those user accounts, and the actual placement of a person at the same location and time when the data is created.

1.1.2 *Digital Evidence*

Digital evidence is any type of digital data with incriminating characteristics, which can result from any type of action preformed by a user, like transactions, recordings, or virtually any action preformed on a device.

Nowadays it's virtually impossible not to leave a digital track behind, since most of us carry and use devices capable of connecting to the internet.

The explosion of social media sites has created a whole new area of electronic evidence. Most people today are willing to share all kinds of information through social media platforms.

In order for electronic data to become digital evidence, it must be stored and be recoverable by a forensic examiner. One of the great challenges is not whether digital evidence may exist, but where the evidence is stored, getting access to that storage, and finally, recovering and processing that digital evidence for relevance within a civil or criminal action.

The potential storage options for electronic evidence has shifted from being only contained locally to being either located locally or remotely in what is called "The Cloud".

More and more everyday computing processes are moving to the Internet where companies offer software as a service. Software as a service means that the customer no longer has to install software on their computer, allowing access to the software remotely, and not storing any data locally.

1.1.3 *Processes and Procedures*

Digital forensics is the application of forensic science to electronic evidence in a legal matter.

While there are many different subdisciplines and many types of devices, communication, and storage methods available, the basic principles of digital forensics apply to all of them.

These principles encompass four areas:

1. Acquisition
2. Preservation
3. Analysis
4. Presentation

Each of these areas includes specific forensic processes and procedures.

1.1.3.1 *Acquisition*

Acquisition is the process of collecting electronic data. Seizing a computer at a crime scene or taking custody of a smartphone in a civil suit are examples of device acquisition, but the data must be extracted from these devices using specific procedures that equate to making a copy of the storage devices, while following strict rules to ensure the integrity of all the extracted data.

Since acquisition is the first interaction between the investigators and the evidence, it is the step where it's most likely to occur modifications of the contents of the seized devices, because turning on the device or extracting the data without following the right procedures can alter it's contents irreversibly.

1.1.3.2 *Preservation*

For evidence to be defendable in court it must be preserved properly. Preservation in the forensics context is the process of creating a chain of custody that begins before collecting the evidence and ends when the evidence is released. Any interference in the chain of custody can lead to issues regarding the validity of the evidence. Additionally, preservation includes maintaining the evidence in a safe environment, preventing intentional destruction with malicious purpose or accidental modification by unqualified people.

A chain of custody log allows proving that the integrity of the evidence has been maintained from seizure through presentation in court. It should contain entries for every time that a piece of evidence has been touched, including collection, storage transport, and any time the evidence is checked out for handling by any personnel.

1.1.3.3 *Analysis*

Analysis is the process of locating and categorizing items from evidence that has been collected in a case. Each case is unique as the circumstances surrounding each case can vary immensely, not only in the evidence being analyzed, but also in the approach used to perform the analysis. The analysis is the area where the individual skills, tools used, and the training of the forensic examiner have the greatest impact on the outcome of the examination. Considering that electronic evidence appears in so many forms and comes from diverse locations and devices, the training and experience of the examiner has a much greater impact on the results of the examination.

Analysis of digital evidence is more than just determining whether a file exists on a hard drive, it involves finding out how that file got on the hard drive, and if possible, who put the file on the hard drive.

1.1.3.4 *Presentation*

Presentation of the examiner's findings is the last step in the process of forensic analysis of electronic evidence. This includes not only the written findings or forensic report, but also the creation of sworn statements, depositions of experts, and court testimony. There are no exact rules or standards for reporting the results of an examination. Each entity may have its own particular guidelines for reporting. However, forensic examination reports should be written clearly, concisely, and accurately, explaining what was examined, the tools used for the examination, the procedures used by the examiner, and the results of the examination. The report should also include the collection methods used, including specific steps taken to protect and preserve the original evidence and how the verification of the evidence was performed.

1.2 THE SLEUTH KIT

The Sleuth Kit (TSK) is a library and collection of command line tools that allow the investigation of disk images. The core functionality of TSK allows volume and file system data analysis. The plug-in framework allows incorporation of additional modules to analyze file contents and build automated systems. The library can be incorporated into larger digital forensics tools and the command line tools can be directly used to find evidence.

1.2.1 *Description*

The original part of Sleuth Kit is a C library and collection of command line file and volume system forensic analysis tools. The file system tools allows examining file systems of a computer in a non-intrusive fashion. Because the tools do not rely on the operating system to process the file systems, deleted and hidden content can be shown.

The volume system tools allow examination of the layout of disks and other media. TSK supports DOS partitions, BSD partitions, Mac partitions, Sun slices, and GPT disks. With these tools, partition locations can be identified and extracted so that they can be analyzed with file system analysis tools.

When performing a complete analysis of a system, command line tools can become tedious. Autopsy is a graphical interface to the tools in TSK, which allows easier conduction of an investigation. Autopsy provides case management, image integrity, keyword searching, and other automated operations.

A complete analysis also requires more than just file and volume system analysis. However, a single tool can't provide support for all file types and analysis techniques. The TSK Framework allows tools to easily incorporate file analysis modules that were written by other developers.

TSK Analyzes raw, Expert Witness and AFF file system and disk images. It supports the NTFS, FAT, ExFAT, UFS 1, UFS 2, EXT2FS, EXT3FS, Ext4, HFS, ISO 9660, and YAFFS2 file systems.

1.2.2 *Search Techniques*

TSK allows listing allocated and deleted ASCII and Unicode file names, can display the details and contents of all NTFS attributes, can display file system and meta-data structure details, can create time lines of file activity, which can be imported into a spread sheet to create graphs and reports. TSK allows the lookup of file hashes in hash databases, it organizes files based on their type, and pages of thumbnails can be made from graphic images to facilitate quick analysis.

1.3 AUTOPSY

Autopsy is a digital forensics platform and a graphical interface to The Sleuth Kit along with other digital forensics tools. It is used by law enforcement, military, and corporate examiners to investigate what happened on a computer. It can even be used by anyone to recover photos from a camera's memory card.

Autopsy was designed to be intuitive out of the box. Installation is easy and wizards guide the user through every step. All results are shown in a single tree.

Autopsy was designed to be an end-to-end platform with modules that come with it out of the box and others that are available from third-parties. Some of the modules provide:

- Timeline Analysis - Advanced graphical event viewing interface;
- Hash Filtering - Flag known bad files and ignore known good;
- Keyword Search - Indexed keyword search to find files that mention relevant terms;
- Web Artifacts - Extract history, bookmarks, and cookies from Firefox, Chrome, Internet Explorer and Microsoft Edge;
- Data Carving - Recover deleted files from unallocated space;
- Multimedia - Extract EXIF metadata from pictures and videos and display these files;
- Indicators of Compromise - Scan a computer using STIX.

Autopsy runs background tasks in parallel using multiple cores and provides results as soon as they are found. It may take hours to fully search a drive, but the user will know in minutes if certain keywords were found in a specific folder.

Autopsy is free. As budgets are decreasing, cost effective digital forensics solutions are essential. Autopsy offers the same core features as other digital forensics tools and offers other essential features, such as web artifact analysis and registry analysis, that other commercial tools do not provide.

1.4 ALTERNATIVE SOFTWARE

1.4.1 *Nuix*

Nuix Lab allows investigators to work efficiently on gigabyte to terabyte-sized investigations and beyond. It's ideal for local or small regional forensic labs struggling with the expanding volume, variety, and complexity of digital evidence and looking to build or upgrade a dedicated digital forensics facility.

If the resources are spread geographically or companies are looking to facilitate greater collaboration across departments, Nuix Lab breaks down evidence silos and makes better use of existing team members and intelligence. Nuix software puts evidence into the hands of less technical reviewers or case officers sooner in the investigation.

The core technologies of the Nuix Lab, Nuix Workstation and Nuix Investigate, give digital forensic technicians and case investigators different lenses into the same case data. Investigators benefit from an easy-to-use browser experience where they can collaborate on the same data at the same time, creating efficiency and helping them share insights.

Implementing Elasticsearch as a data store for the Nuix Lab boosts evidence processing, investigation, and intelligence capabilities. It's appropriate for investigations that contain massive volumes of digital evidence and numerous digital exhibits; are conducted across multiple regions or jurisdictions, or need to cross-reference and correlate intelligence across multiple current and historical cases.

In addition, it contains powerful artificial intelligence, machine learning, and analytics to supercharge the investigations.

1.4.2 *EnCase Forensic*

EnCase Forensic enables quickly searching, identifying, and prioritizing potential evidence, in computers and mobile devices, to determine whether further investigation is warranted. This will result in a decreased backlog so that investigators can focus on getting to case closed.

EnCase Forensic helps acquire more evidence than any product on the market. Can collect from a wide variety of operating and file systems, including over 25

types of mobile devices with EnCase Forensic. Parses the most popular mobile apps across iOS, Android, and Blackberry devices so that no evidence is hidden. This is the flexibility needed to ensure teams can complete cases no matter where the potential evidence resides.

EnCase Forensic is unmatched in its decryption capabilities, offering the broadest support of any forensic solution. Encryption support includes products such as Dell Data Protection, Symantec, McAfee, and many more. Decryption power can be further expanded with Tableau Password Recovery — a purpose-built, cost-effective hardware solution to identify and unlock password-protected files.

The EnCase Forensic evidence processor provides industry-leading processing capabilities that can automate the preparation of evidence, making it easier to complete an investigation. Powered by an indexing engine built for scale and performance, it can automate complex queries across varied evidence sources in one step saving time and increasing efficiency.

The most important part of any investigation is your ability to analyze evidence. EnCase Forensic is built with the investigator in mind, providing a wide range of capabilities that enables performing deep forensic analysis as well as fast triage analysis from the same solution.

EnCase Forensic provides a flexible reporting framework that empowers tailoring case reports to meet specific needs. With comprehensive and triage reporting options built in, can create reports for a wide range of audiences and easily share them across an organization.

1.4.3 *Forensics Toolkit*

FTK is an award-winning, court-cited digital investigations solution built for speed, stability and ease of use. It quickly locates evidence and forensically collects and analyzes any digital device or system producing, transmitting or storing data by using a single application from multiple devices. Known for its intuitive interface, email analysis, customizable data views, processing speeds and stability.

All digital evidence is stored in one case database, giving teams access to the most current case evidence. It reduces the time, cost and complexity of creating multiple datasets. And best of all, there is continuous data transfer between AccessData's forensic and e-discovery solutions, allowing for true collaboration between all parties working on the case.

With customizable processing, teams have the ability to establish enterprise-wide processing standards, creating consistency for the investigations and reducing the possibility of missed data. Since evidence is processed up front, investigators don't have to wait for searches to execute during the analysis phase. FTK is designed to provide the fastest, most accurate and consistent processing with distributed processing and true multi-threaded/multi-core support.

Indexing is done up front, so filtering and searching are faster than with any other solution. FTK offers the flexibility to perform multipass data review and change indexing options without reprocessing data. Whether teams are in the investigating phase or performing document review they have a shared index file, eliminating the need to recreate or duplicate the file. Most importantly, they receive consistent search results regardless of whether they are searching in FTK or Summation. Social Analyzer allows viewing email communications at the domain level and drill down to the custodian level to see communications among specific individuals.

FTK allows users to create images, process a wide range of data types from forensic images to email archives and mobile devices, analyze the registry, crack passwords, and build reports—all within a single solution.

With the single-node enterprise, users can preview, acquire and analyze evidence remotely from computers on the same network.

Automatically constructs timelines and graphically illustrates relationships among parties of interest in a case. With Email, Social and File Visualization users can view data in multiple display formats, including timelines, cluster graphs, pie charts, geolocations and more, to help determine relationships and find key pieces of information. Then generate reports that are easily consumed by attorneys, CIOs or other investigators.

Almost every investigation involves the analysis of Internet artifacts. Web browsing caches store records of sites a suspect has visited, web-based emails may help to prove intent or correlate other events and instant message conversations or social media sites can contain evidence. When evidence is processed, artifact files are categorized and organized so they can easily be seen.

Available as an option to FTK, Cerberus is an automated malware triage platform solution designed to integrate with FTK. It's a first layer of defense against the risk of imaging unknown devices and allows identifying risky files after processing data in FTK. Then users can see which files are potentially infected and can avoid exporting them. Cerberus is one tool in the malware arsenal and helps identify potentially malicious files.

Teams can:

- Determine both the behavior and intent of security breaches sooner by providing complex analysis prior to a full-blown malware attack.
- Strengthen security defenses and prevent malicious software from running with state-of-the-art technology called whitelisting.
- Take action sooner when security breaches occur; unlike other competitors, Cerberus doesn't rely on a sandbox or signature-based solutions.

1.5 PROPOSED SOLUTION

As can be seen from the existing alternative software, collaboration is a major feature included in all these programs, and while Autopsy also allows the software to be configured in a manner to allow collaboration it involves a complicated setup and has certain limitations. The setup required for collaborative cases in Autopsy is as follows:

- Shared hard drive accessible for every computer setup for every machine using the same drive letter
- PostgreSQL Database server
- Solr indexing server
- ActiveMQ messaging server

The limitations of Autopsy's multi-user case feature are that it requires every user to be using a Windows O.S. computer and it also requires specific configuration on each machine involved in the case that is set up.

The proposed solution aims to cover 3 main aspects:

1. Accessibility
2. Collaboration
3. Organization

1.5.1 *Accessibility*

Given a client-server architecture, any client with access to the network where the server is located can access the contents provided by the server. The aim is to

condense the processing heavy features of Autopsy in a single server and provide any number of clients access to this information, requiring less resources from each client, allowing collaboration and removing any type of setup required for each of the client machines, while also providing a more modern and user friendly design.

1.5.2 *Collaboration*

In order to provide collaboration, all the information is maintained in a single server, or a collection of servers providing different functions (like exposing endpoints, storing data, and indexing searches), and every client can preform all the allowed actions whether they consist in consulting, generating, or removing information. Collaboration comes naturally with a client-server model, as the same server that provides the endpoints can also communicate with each client using websockets, and maintain information in a coordinated state along every connected client.

1.5.3 *Organization*

Digital forensics investigations are usually done by specialized organizations, who need to organize their human resources in an efficient and secure manner. Assigning investigators to teams, assigning teams to cases, allowing access to the platform and certain information is a critical part of the activities preformed by a company that specializes in digital forensics, so having these functionalities properly integrated into a digital forensics platform should be an important feature.

2

HOST ENTITY CHARACTERIZATION

To attain an internship, I was interviewed at "VOID SOFTWARE, S.A.", where I was welcomed to carry out my planned 9 month long curricular internship. The reasons behind choosing this company for the internship were related to it being, in my point of view, the most interesting place in my home town to begin my career as a software engineer.

This chapter gives some information about the company, like it's inception, areas of expertise, past projects developed and work conditions provided.

2.1 COMPANY INFORMATION

VOID is a privately held software development company established in Leiria, Portugal, in 2006, focused on building high-end products embodied in web, mobile and desktop applications, supported by creative software engineering tailored to each challenge's specific needs. It currently employs 30 High-End Professionals in several fields of expertise.

2.2 AREAS OF EXPERTISE

VOID mostly functions as a company that develops software tailored to the specifications provided by the client, although it can also provide services in different areas like cybersecurity and digital forensics.

The company is capable of comfortably providing services in the following areas:

- Blockchain
- Machine learning and data science
- Augmented Reality and virtual reality
- Mobile applications
- Web applications

- Desktop applications
- Cybersecurity and digital forensics

2.3 PAST PROJECTS

Throughout it's 14 years of being active in the software development industry, VOID has conducted some very interesting projects, such as:

- Yes Account - a suite of applications for automated digitization of accounting documents
- Web Portal - a large scale project for the European commission
- Digital Archive - a digital preservation application
- Dream Football - a social network along with web and mobile applications
- Fuel Write - a comprehensive platform for fleet management, data collection and route optimization
- Caspers - a mobile augmented reality customer experience and engagement
- PBCore Toolkit - a desktop application to support the creation, editing, and export of moving image-related inventory metadata as PBCore XML records
- Avenue Securities - a trading platform

2.4 WORK ENVIRONMENT

VOID prides itself in providing very good conditions to it's workers, making them feel like they are at home while working, and also to feel motivated to come to work every day. These conditions include the work environment itself, which is an open space where everyone can interact with each other, the "play areas" where people can relax while playing a game of pool or video games, the rooftop terrace where workers can relax on sunny days. the company also makes sure nothing is missing to provide the best work environment possible by always having food and drink to all it's workers at any time.

The first impressions for someone who first walks into the space are very positive, and while being part of this family for 9 months I can also say that I had everything I needed to succeed in my career as a software developer while here.

3

INTERNSHIP PROGRAMME

The internship started on September 2nd 2019, being planned to last for 9 months, ending on the 29th of May 2020.

The first day at the company was interesting, meeting most of the personalities of the workplace and getting set up with my work environment, consisting of a desk, chair, computer and 2 monitors.

As soon as the computer was properly configured with all the software needed the project became the main focus of my time at the company, making progress daily, sometimes just planning the architecture or analyzing the source code, sometimes just writing code, and sometimes a mix of both.

3.1 PROJECT PLANNING

Autopsy by itself is capable of providing a distributed solution for multi user collaboration, but it is very resource intensive and requires many complex configuration steps, and is also only fully supported on the Windows O.S.

The plan for this internship is to achieve the same kind of functionality provided by the original software, but without the dependency on harduous pre-configuration or hardware intensive requirements, resulting in needing only single capable server, and allowing the program to be used by multiple low capacity client devices using any web capable O.S.

To achieve that goal, the project will be an adaptation of the original Autopsy source code, into a client-server model, with the server developed in Java using the Quarkus framework, and the client developed in Javascript using the React framework.

The project is outlined to work in a multi user environment, allowing users to be assigned to teams and teams assigned to cases, and allowing multiple users to interact with a case simultaneously.

Autopsy has a major limitation, which is it only allows one case to be open at a time, ideally in this project we should find a workarround to allow working on multiple cases at once, but we feel like dedicating a server instance (which a single machine can have many virtualized within) per case is a good enough approach, though as future work the ability to spawn different containers as requested to work on different cases at once is an interesting challenge.

Given that the core features will be running in a remote server, it was decided that the addition of data sources to cases will be handled by an FTP client, allowing users to transfer files to their respective data source directories, and FTP access will be controlled according to each user's credentials on the platform.

3.2 AUTOPSY SOURCE CODE ANALYSIS

Autopsy is a digital forensics analysis software that is available as Open Source Software on GitHub.

With the goals set for this project, the source code was analyzed to understand which components need to be replicated and adapted in order to obtain the same logic flow.

The "Core" module is where the most important components are located, and after analysis it was concluded that the following directories contain relevant information:

- Actions: user interactions
- Casemodule: case class and other resources needed for the functioning of an autopsy case like data sources and artifacts
- Centralrepository: data persisted and accessed by multiple cases (Correlation Engine)
- Contentviewers: panels used for data representation
- Coordinationservice: configuration information distribution system
- Core: addition of command line options, system configurations and collaboration monitor
- Corecomponents: main user interface components
- Datamodel: all the entities needed to represent ingested data
- Datasourceprocessors: data processing utilities
- Directorytree: file explorer for ingested artifacts

- Ingest: utilities and events for data ingestion
- Keywordsearchservice: utility to search artifacts by keyword
- Modules: all the pre-included modules (data ingestion procedures)
- Progress: progress indicators and similar classes
- Python: resources needed for the functioning of the Jython language
- Rejview: resources used to analyse Windows registry
- Report: report generator
- Timeline: recent addition to Autopsy, allows visualization of artifacts in temporal chart, only available for Windows O.S.

The "KeywordSearch" module is also of critical importance as it provides one of the most meaningful features which is filtering all the artifacts in a case with a keyword search using the Apache Solr search platform, which indexes the text contents of all the artifacts and allows extremely fast searching through a large amount of data.

Another module that needs to be adapted is the "RecentActivity" module, which contains the tools needed to extract information from browsers, providing a great amount of critical evidence from data sources.

3.3 DEVELOPMENT

3.3.1 *Management Entities*

As a first step in the development, the different persisted entities were created, which are Users, Teams, and Cases. All the endpoints for actions involving these entities were created, resulting in the ability for the client program to interact with these entities and modify their relationships and other variables. For these functionalities there are two roles associated, the Manager role allows manipulation of the existing entities while the Investigator role only has access to his own information and the teams and cases he was assigned to. For these interactions, it was decided to create a drag and drop interface, which allows users to be dragged into teams and teams dragged into cases. All these entities are listed side by side and each have their own options and filtering input, as can be observed in Figure 1.

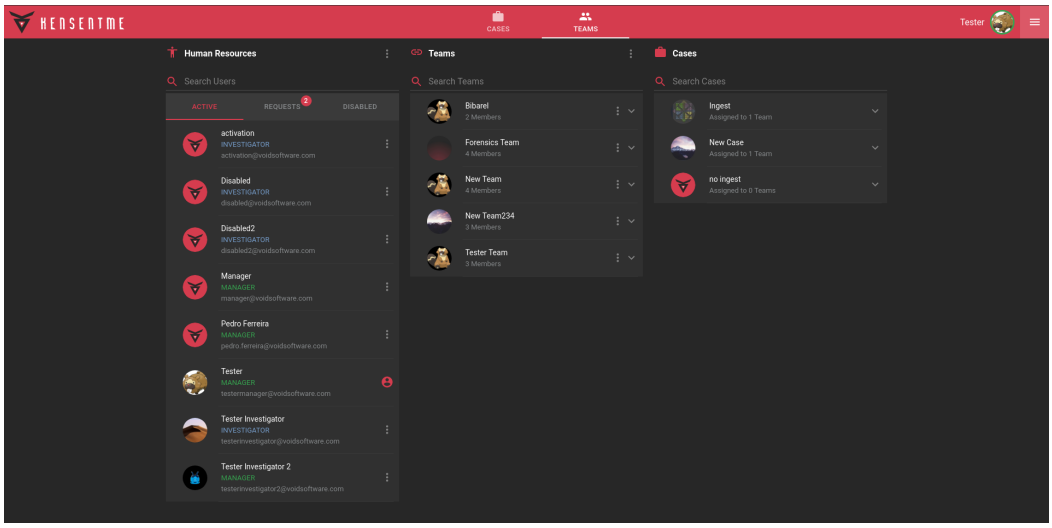

Figure 1: Entity Management Interface

Users can change their own profile picture, while Managers can also change any team or case's display picture. Managers can add new users to the platform, can approve membership requests, can enable/disable user accounts and can create new teams. When a user is added by a Manager or his membership request is approved he must define a password while activating his e-mail account.

3.3.2 *Basic Autopsy Functionalities*

Then the most basic functionality from Autopsy was adapted, the ability to open an autopsy case. For this some elements of the original Casemodule package were adapted, and after that all the other similar actions like closing, creating and deleting cases were also adapted.

Autopsy cases have a case file containing case metadata, which allows the program to connect to the right database when the case is open, this database is also present in a file inside the filesystem, which uses the SQLite database engine, so for the cases to be usable in the server these files must also be present in the server, which resulted in the creation of a directory within the server called "repository", containing all the different cases created within the application.

Later in the development there was the need to create an additional directory alongside the "repository" called "central-repository" which contains the database used by the Correlation Engine to ingest data that can be queried by any case.

3.3.3 *Ingested Results Presentation*

Ingested results are the items present inside the provided data sources, Autopsy runs multiple modules on each data source and extracts these results using The Sleuth Kit, extracted results can either be a file instance or an artifact (which is something that corresponds only to a piece of information inside a file).

The ingested results are presented in three different containers, one taking the shape of a file explorer, allowing exploration of the structure of all the results, one taking the shape of a table or thumbnail viewer, presenting all the contents of the results selected from the explorer, and one taking the shape of a content viewer, allowing visualization of the data contained inside the result selected from the table as can be seen in Figure 2.

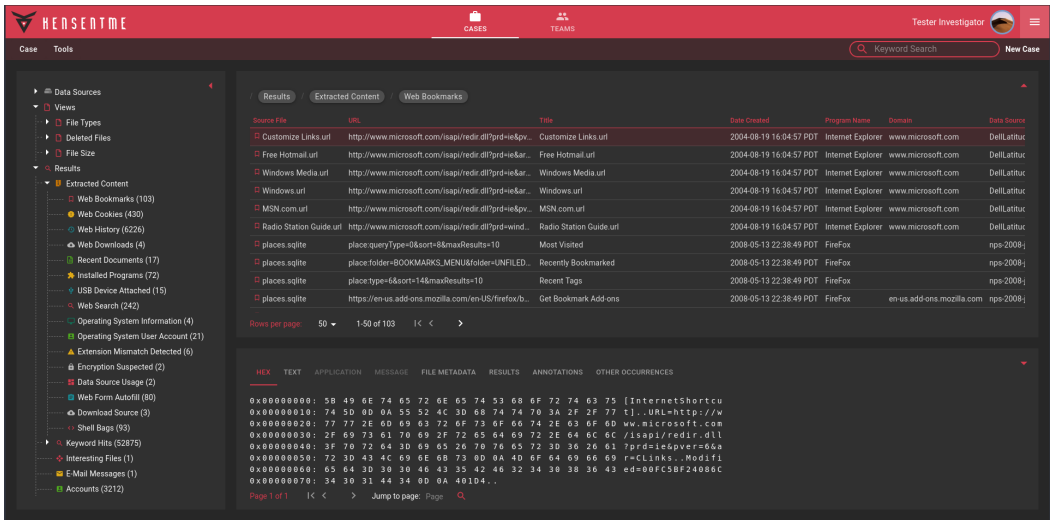

Figure 2: Ingested Results Presentation

The content viewer can display different kinds of information depending on the type of item selected, these can be some of the following:

- Text browser
- Media viewer
- Database browser
- Registry browser
- Key-value browser
- Table data viewer

The layout for the ingested results presentation, and for all the case related actions, was based on the original Autopsy layout, so that each container can be re-sized as needed, allowing the user to focus on the information that is most important to him.

3.3.4 *Data Sources*

Using the same credentials used to log-in to the platform, the user can also upload data source files into his folder located inside the server, using an FTP client like filezilla. Then the user can browse these directories using the web interface and select a data source to add to the case, as can be seen in Figure 3

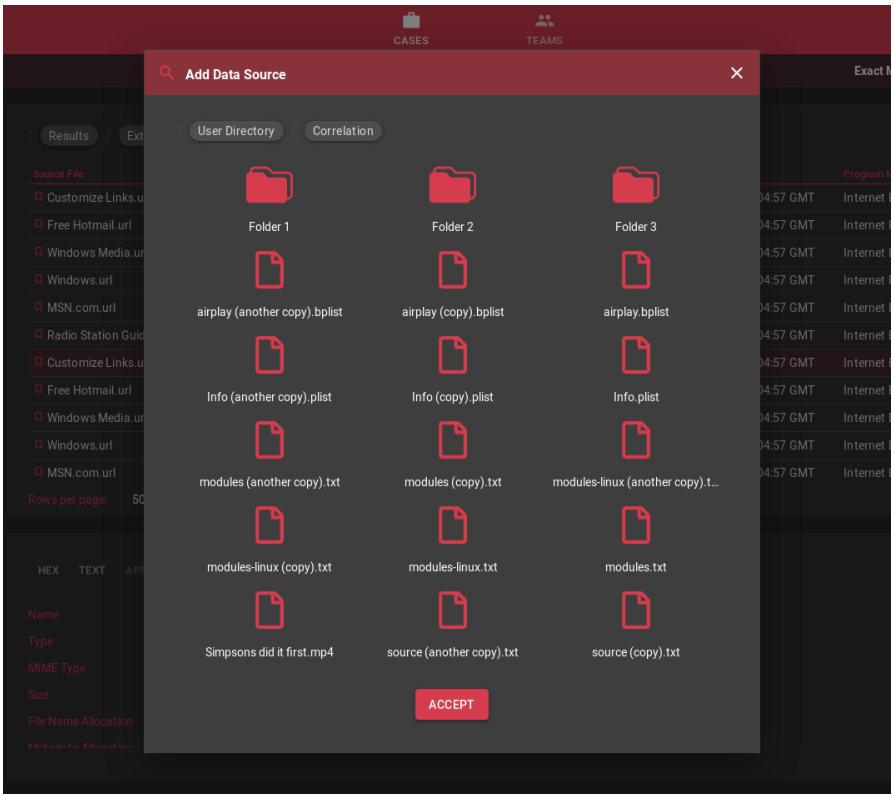

Figure 3: Data Source Selection

The procedure for adding a data source to a case was adapted from Autopsy's source code, and depends on the type of data source added, which can be one of the following:

- Disk image
- Virtual machine
- Logical file collection

- Unallocated space image file
- Autopsy logical imager results
- Memory image files from Volatility

Local disk data sources were also an option provided by Autopsy but since the local disks the software has access to belong to a server, this feature is undesired.

3.3.5 *Data Ingestion Modules*

Firstly the default modules included with Autopsy were adapted, and can be ran through the web interface. When the modules are running the server communicates to each client when they start, updates their progress, and informs that they finished, as can be seen in Figure 4.

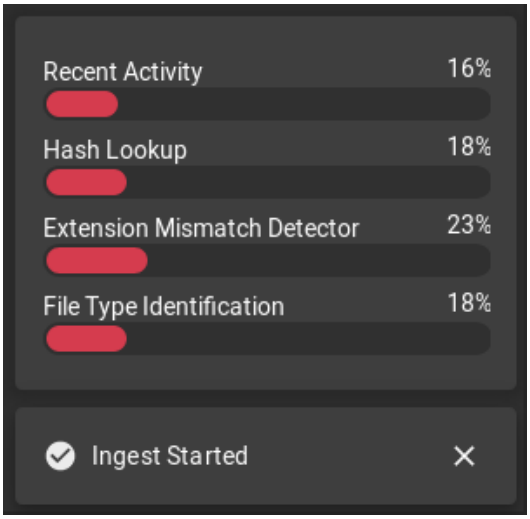

Figure 4: Ingest Modules Communication

When a module finishes the case explorer is updated with the extracted information.

3.3.6 *Report Modules*

TODO

3.4 OTHER PROJECTS

TODO

4

CRITICAL ANALYSIS AND PROPOSED IMPROVEMENTS

TODO

5

CONCLUSIONS

TODO

APPENDICES

A

APPENDIX A

B

APPENDIX B

D E C L A R A Ç Ã O

Declaro, sob compromisso de honra, que o trabalho apresentado nesta dissertação, com o título *"AUTOPSY – Enhanced distributed forensic analysis"*, é original e foi realizado por Pedro Henrique Gaspar Cordeiro Ferreira (2180078) sob orientação de Professora Doutora Marisa da Silva Maximiano (marisa.maximiano@ipleiria.pt).

Leiria, Fevereiro de 2020

Pedro Henrique Gaspar Cordeiro Ferreira