

**R & D Project**

**Academic Year- 2023-24**

### PROJECT 07

**Unveiling Digital Trails: Forensic Extraction of Telegram Application Across Platforms**

|  |  |  |
| --- | --- | --- |
|  | Submitted by |  |
| Jatin Joshi | BT21GCS068 | BTech CSE |
| Yash Shinde | BT21GCS165 | BTech CYS |
| Suryansh Verma | BT21GCS042 | BTech CSE |

Supervisor: Dr. Anand Kumar Mishra



##### ====================================================================

**====================================================================**

### CERTIFICATE BY SUPERVISOR(S)

This is to certify that the present R&D project entitled “Unveiling Digital Trails: Forensic Extraction of Telegram Application Across Platforms” being submitted to NIIT University, Neemrana, in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology, in the area of BT/CSE/ECE/GIS, embodies faithful record of original research carried out by Jatin Joshi, Suryansh Verma, Yash Shinde. They have worked under our guidance and supervision and this work has not been submitted, in part or full, for any other degree or diploma of NIIT or any other University.

Place: Niit University Neemrana Name: Dr.Anand Kumar Mishra Date: 22/05/2024



========================================================================

### DECLARATION BY STUDENT(S)

We hereby declare that the project report entitled “Unveiling Digital Trails: Forensic Extraction of Telegram Application Across Platforms” which is being submitted for the partial fulfillment of the Degree of Bachelor of Technology, at NIIT University, Neemrana, is an authentic record of our original work under the guidance of Dr.Anand Kumar Mishra. Due acknowledgements have been given in the project report to all other related work used. This has previously not formed the basis for the award of any degree, diploma, associate/fellowship or any other similar title or recognition in NIIT University or elsewhere.

Place: Niit University, Neemrana Date: 22/05/2024

|  |  |  |  |
| --- | --- | --- | --- |
| Student Name | Enrollment ID | Branch Name | Signature |
| Jatin Joshi | BT21GCS068 | BTech CSE |  |
| Suryansh Verma | BT21GCS042 | BTech CSE |  |
| Yash Shinde | BT21GCS165 | BTech CYS |  |

#### Table of contents

|  |  |  |
| --- | --- | --- |
| **S.no** | **Topic** | **Pg.no** |
| 1 | Abstract | 5 |
| 2 | Keywords | 5 |
| 3 | Introduction | 6 |
| 4 | Related Work | 7 |
| 5 | Tools | 11 |
| 6 | Result | 13 |
| 7 | Analysis | 15 |
| 8 | References | 20 |

**Abstract**

Telegram utilizes cloud-based storage and end-to-end encryption, experts with specialized skill in digital forensic science are necessary to analyze data within Telegram's platform. Standard investigative protocols often are insufficient for accessing and decrypting encrypted communications and media. Consequently, we have focused on detecting and applying approaches that can overcome these obstacles without jeopardizing the integrity and reliability of the recovered information. This research analyzed the intricate process of performing digital forensics on the popular instant messaging platform Telegram. Our findings concerning the strategies and resources leveraged to surmount these barriers are condensed in this assessment.

A number of technologies have proven critical to our data forensic work concerning Telegram. For instance, Telepathy is a versatile instrument that collects data pertaining to user interactions, message history, and media files by interfacing with Telegram's application programming interface. In the field of digital forensics, extracting memory contents is a core step, particularly regarding transient data that may not be saved permanently. Windows Memory Extractor, records the contents of a system's memory, like session keys, decrypted contents, and other remnants related to active Telegram sessions, using the Windows Memory Extractor. Telegram Scraper tool automatically collects user identifiers and other associated metadata from Telegram groups.

Similarly to obtain metadata from multimedia files shared on the Telegram platform, we have ExifTool.

#### Keywords

Telegram, Digital Forensics, Instant Messaging Platform, User data Extraction, Forensics analysis tools, Telepathy, Windows Memory extractor, Telepathy, Telegram Scraper, ExifTool.

#### Introduction

Telegram is a cloud-based instant messaging platform known for its speed, security, and flexibility. The messaging platform Telegram was founded in 2013 by twin brothers Nikolai and Pavel Durov. It experienced rapid growth, reaching over 700 million active users globally by 2024. It allows users to send text messages, media files, and make voice and video calls, as well as support large groups and channels with thousands of members. Telegram's focus on privacy and encryption makes it a popular platform for both personal and professional communication. Telegram offers several unique functionalities, including self-destructing messages, end-to-end encryption for private conversations, and multi-device access to messages. These capabilities, along with a robust API, have made Telegram a popular choice for both personal and professional communication. Telegram's strong API allows for a large variety of bots and integrations to be built on the platform. Digital forensics, on the other hand, is the branch of forensic science that deals with recovering and investigating material found on digital devices, often related to computer crimes. This area focuses on identifying, preserving, analyzing and presenting digital evidence, which can be critical in legal proceedings. Performing digital forensics on Telegram requires using specialized tools and techniques to extract, analyze and interpret data. Digital forensics analyzes data from computers, mobile phones, and other digital devices to potentially discover evidence of wrongdoing, criminal acts, or security issues. As digital communication and data storage has become more pervasive, the importance of digital forensics has grown. Investigations may necessitate access to vital data, such as chat history, media files, and user activity, which digital forensic methods may help uncover.

Due to the platform's privacy and encryption features, performing forensic analysis on Telegram data can be challenging, yet it is critical to comprehend the online behavior and interactions of the individuals who are the focus of the investigation.This process can help uncover important information that may be needed for investigations, such as user interactions, message histories, and media files. This digital forensic analysis utilizes multiple tools, each meant to handle certain data types and provide insight into activities done through Telegram. Experts in digital forensics employ this systematic method to gather usable intelligence while preserving the integrity and

acceptability of any evidence gathered. Digital forensics experts follow a strict methodology to ensure the integrity and admissibility of any evidence collected, while also gathering pertinent intelligence. In such situations, meeting the moral and legal benchmarks required for forensic investigations relies greatly on using reliable and conforming tools.

#### Related Work

Yusoff et al. [1] have analyzed Firefox OS (FxOS), a distinct mobile operating system based on web standards, and the developing discipline of mobile forensics. Prior studies on cloud storage and network traffic analysis were conducted on other systems; FxOS did not receive this kind of attention. Using a simulated environment, this study looked at popular social media programs like Facebook and Twitter as well as an instant messaging service called Telegram on FxOS. Encrypted credentials were discovered despite the presence of valuable evidence such as texts and photographs. It's interesting to note that while others employed encryption, Telegram sent data without encryption. The study also showed that not all data is stored by service providers like Twitter. This study highlights the need for more research into FxOS apps and recommends using the approach for thorough mobile forensics with cloud storage and other mobile OSes.

Gregorio et al. [2] have proposed a forensic analysis technique for extracting digital evidence using the “Telegram Messenger” app for Windows Phone. Author concentrates on learning the app's structure and managing user data, chats, and conversations. The methodology includes decoding and interpreting a variety of data types, including messages, images, videos, and files. The study demonstrates how comprehensive information can be obtained by examining the application's source code, which can be useful in criminal investigation. The methodology is adaptable and works across different app versions and operating systems. The study evaluates the efficacy of various forensic analysis techniques and emphasizes the significance of validating forensic tools to ensure accurate data extraction. Overall, the study helps to improve forensic analysis practices, which benefits both open-source tools and forensic analysts.

Fernández et al. [3] have explored the analysis of Instant Messaging (IM) applications, specifically focusing on the Windows Telegram Desktop Application. The study is structured in three phases: message extraction, analysis, and reporting. Recognizing limitations in existing tools, the researchers developed two open-source forensic tools, namely “Windows Memory Extractor” (coded in C++) and “IM Artifact Finder” (coded in Python). These tools prove

effective in foundational analysis and extracting memory artifacts from IM application data, even after actions like deletion or logout. The research primarily delves into the analysis of “Telegram Desktop 2.7.1” and recommends updates for later versions. Future projects aim to delve deeper into forensic memory artifacts & adapt tools for non-Windows platforms.

Vasilaras et al. [4] have proposed an investigation of Telegram applications on android mobile devices. One of the ways they have used to retrieve the deleted data back is by using a tool called SDK Platform Tools 31.03 and ADB commands (CLI) for tools used for data acquisition, Image acquisition. They have conducted this research on two different devices (LG G6 Android 9 & Samsung A50 Android 11). Authors compared the data retrieved from both devices, no differences were found. Authors explored that in two scenarios the data could not be retrieved back: When a user has logged out and when the local database has been cleared out.

Iqbal et al. [5] have proposed an investigation for retrieving data back from Google Meet. Authors tells us about how memory and browser artifacts are being retrieved. Their research paper has aimed to provide us with: All the important clues that tell us how meetings are being conducted. Authors have developed a tool based on Python to retrieve artifacts, Authors have made test cases based on RAM (4 GB, 8 GB, 12 GB) which tell us about how much memory is required by the Google Meet application to work properly, Authors also executed their experiment based on web browsers and whether the data being retrieved was the same or not. The test cases that they have used are as follows: Starting an instant meeting and creating a meeting for later, scheduling a meeting in Google Calendar, join a meeting that is set up by another user. They Found in the research paper that retrieving the data is independent of the web browser but is dependent on the size of the RAM.

Heath et al. [6] have proposed an investigation on the ‘disappearing message’, or ephemeral message, which is a feature in applications for multimedia that automatically banishes messages from the recipient's screen once they have been viewed. This feature is enabled by default in some applications like Snapchat but some require manual enabling like WhatsApp and Telegram. The time period can be 24 hours, 7 days or 90 days. A forensic examination looked into messages that vanished from Telegram, Snapchat, and WhatsApp. Seven groups investigated pre and post disappearing behavior, desktop behavior of these applications, accessing the cloud, and rolling time and date back and forth to recover messages in both android and iOS. The findings indicated that while in Android, Snapchat and Telegram had high

recovery rates whereas for iOS, WhatsApp was robust. Rolling the dates back made data retrieval easier, however cloud access and desktop applications didn't work effectively. On Android, Snapchat was helpful, while Telegram required laborious techniques like manual hex editor.

#### Problem Statement:

*“Develop a robust methodology for conducting Telegram forensic analysis on Linux systems, focusing on extracting secret chat data, contact information, geographical locations, and videos.”*

#### Proposed Methodology:

##### COLLECTION:

**Find the Telegram Database Files:**

Locate and open the target Windows system's %AppData%\Telegram Desktop\tdata\ directory. Obtain the Telegram database file telegram.db from this directory by using FTK Imager or other forensic imaging tools.

##### Gather More SQLite Database Files:

For a more thorough examination, look for and obtain more SQLite database files, such as chatcache.db and groupcache.db, in the Telegram data directory.

##### Collect Media Files and Cache Data:

Examine the %AppData%\Roaming\Telegram Desktop\ directory to gather cached information and media files—such as documents, movies, and images—that have been shared via Telegram.

##### Obtain Telegram Client History:

To rebuild the timeline of user activity within the Telegram application, extract the log files located at %AppData%\Telegram Desktop\logs\.

##### Locate and Examine Windows Registry Entries:

To collect Telegram configuration settings and artifacts, navigate to the Windows registry keys located at HKEY\_CURRENT\_USER\Software\Telegram Desktop and HKEY\_CURRENT\_USER\Software\Microsoft\Windows\CurrentVersion\Uninstall.

##### EXAMINATION:

**Parse Files in the Telegram Database:**

Use forensic analysis software such as Autopsy, Magnet AXIOM, or X-Ways Forensics to parse and examine the telegram.db file.

Extract contacts, groups, media references, messages, and other database-stored metadata.

##### Check Out Further SQLite Database Files:

Examine more SQLite database files, such as groupcache.db and chatcache.db, to obtain more contextual data and metadata on Telegram usage.

##### ANALYSIS:

**Examine Media Files and Cache Data:**

Examine the media files and cached data gathered from %AppData%\Roaming\Telegram Desktop\ to find downloaded files, shared media, and attachments.

##### Examine the client logs for Telegram:

To gain insight into Telegram usage patterns, review log files that were collected from

%AppData%\Telegram Desktop\logs\. This will allow you to extract timestamps, user interactions, and other pertinent information.

##### Decipher Windows Registry Data:

To comprehend installation details, cached credentials, and configuration settings, interpret the Windows registry entries for Telegram located under.

HKEY\_CURRENT\_USER\Software\Telegram Desktop and HKEY\_CURRENT\_USER\Software\Microsoft\Windows\CurrentVersion\Uninstall.

##### REPORTING:

The forensic team documents its findings, methods of investigation, and conclusions in a thorough forensic report.

Details about retrieved messages, contacts, media assets, user behaviors, and registry entries pertaining to Telegram usage are all included in the report.

The report determines whether the data breach happened through Telegram and names possible culprits based on the forensic examination.

The report also offers suggestions for improving security protocols and lowering the likelihood of future attacks.

#### Tools:

##### Telegram Group Scraper

A Telegram scraper tool is a program that gathers user IDs from Telegram groups automatically. These tools gather the required data by utilizing Telegram's API or web scraping methods.

The tool streamlines the retrieval of user IDs, reducing the time and effort required when compared to manual gathering. It provides a significant advantage for collection of data and analysis of data

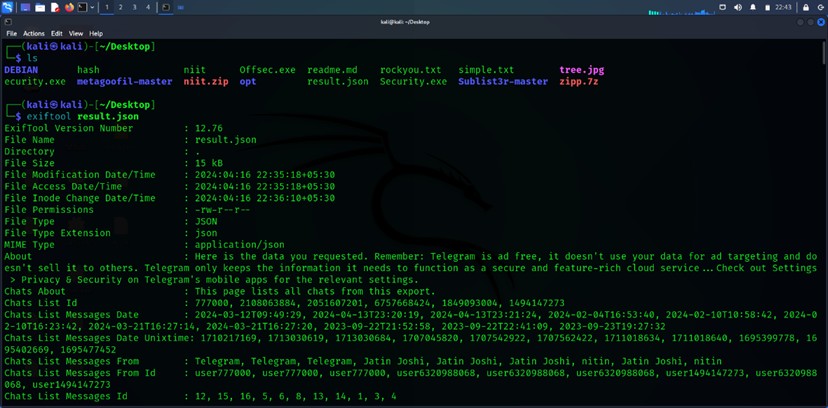


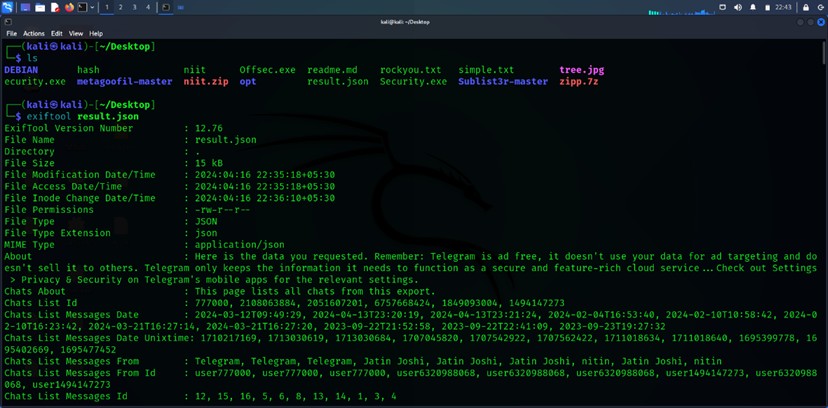
##### EXIF-Tool :

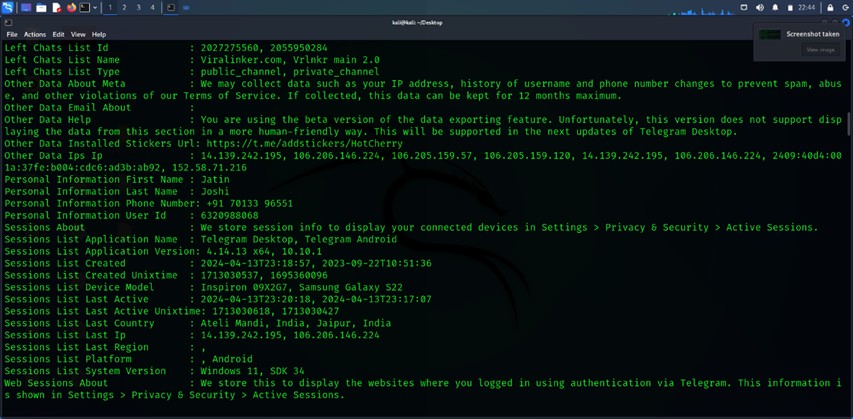
Phil Harvey developed the flexible command-line tool ExifTool to manage metadata in a variety of file formats. It was designed to handle photos, but it can now handle audio, documents, movies, and more. Its ability to analyze data in batches makes jobs like geotagging photos and renaming files according to metadata easier. It has a

command-line interface, but it has a lot of documentation and is still user-friendly.

ExifTool is a popular tool in the fields of research, journalism, photography, and law enforcement because of its effectiveness and versatility in managing information.







**ExifTool** is a valuable tool in the field of cyber forensics since it facilitates the extraction and analysis of information from various file formats, such as photos and videos. Crucial details for investigation reasons, such creation dates, GPS data, and device specifications, can be uncovered. It is an essential tool in forensic investigation because of its capacity to read proprietary formats and process massive datasets quickly.

* + 1. **Command-line Parsing:** To extract parameters like PID, memory protections, module name, etc., the tool parses command-line inputs.
    2. **Memory Extraction:** After the parameters have been processed, the program uses the given criteria to extract memory areas. It makes a distinction between areas that are not executable and those that have certain memory safeguards.
    3. **Module-Specific Extraction:** The tool concentrates on extracting memory areas unique to a given module if the module name is entered.
    4. **Optional Features:** The program may be used to get module version information (--file-version-info) and unite memory areas into a single file (--join).
    5. **Help and Error Handling:** The tool offers ways to handle errors when inputs are erroneous. Users can use the --help command to get further help.

##### Telepathy : A Comprehensive Analysis of Telepathy for Telegram Studies

Telepathy, an Open-Source Intelligence (OSINT) toolset created especially for in-depth analysis of Telegram discussions, was used in this study [7]. Beyond just watching public channels, telepathy enables researchers, with group admins' consent, to collect and examine important data from both public and private Telegram groups. The following is a thorough summary of Telepathy's features that were very helpful for this investigation:

##### Extensive conversation Archiving:

Telepathy has the capacity to save whole conversation histories, accumulating a multitude of data that extends beyond the essential message content. This contains information such as message responses, user reactions (likes, dislikes), and timestamps for every action [7]. Telepathy also has the ability to save media files (pictures, movies, and documents) linked to messages. Users must, however, take care to make sure they are not keeping any unlawful material on their systems.

##### Granular Message Analysis:

To comprehend discussion threads and sentiment inside a group, more analysis can be performed on the message data that has been extracted. Researchers may determine which topics are more interesting or contentious by counting the number of responses to each communication sent via telepathy. Data on reactions, such as the kind and frequency of responses, can reveal how a community feels about particular ideas or messages.

##### User Lookups:

Researchers can utilize telepathy to check for certain users by ID or username. This may disclose a user's involvement in other Telegram groups, which might offer context for their actions in the group that is being looked into [7]. But keep in mind that usernames might be repeated on several platforms, therefore further verification is required to prevent misidentification.

##### Location-based User Discovery (Limited Functionality):

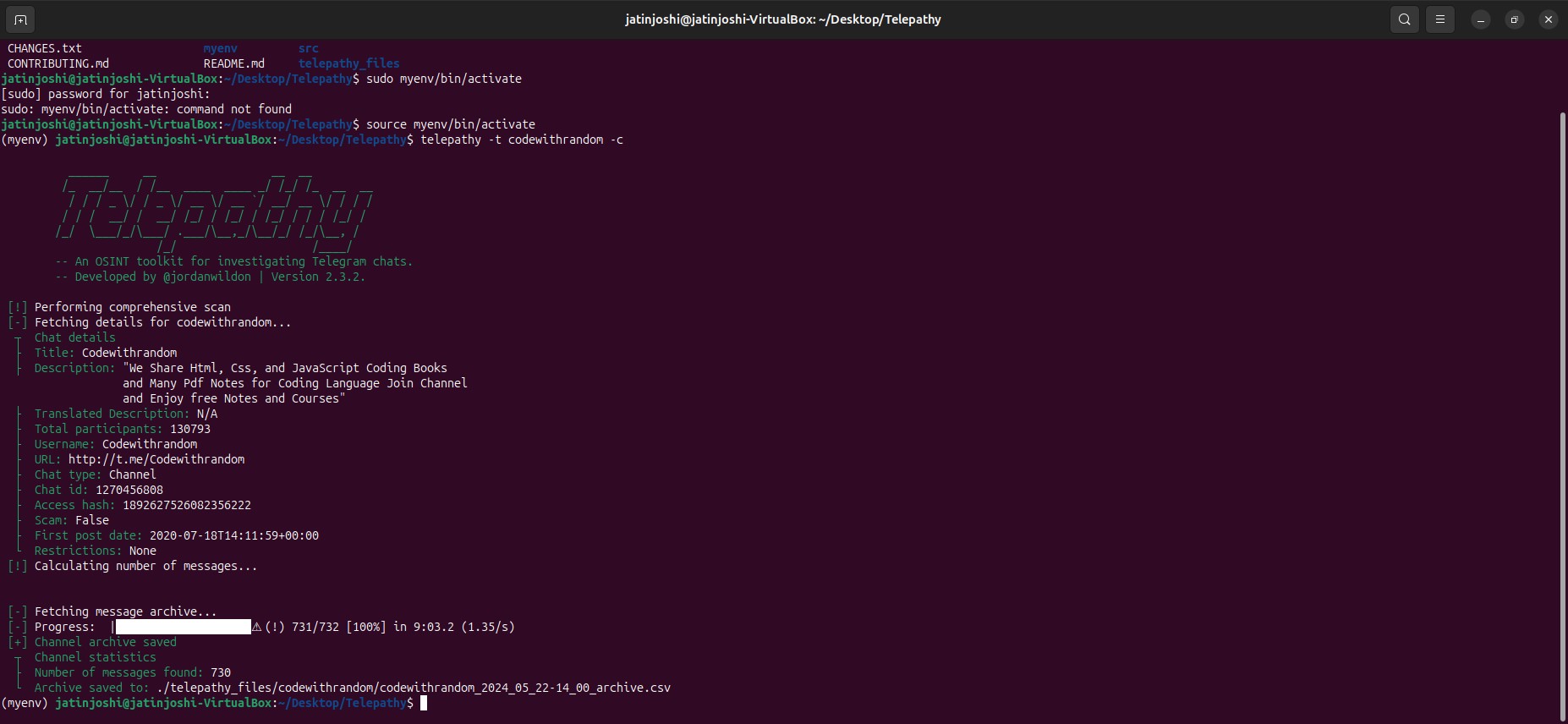
Telepathy allows you to look for users in certain places, however there are some restrictions [7]. In order to use this functionality, the Telegram account doing the search must have a publicly accessible profile photo. While it could occasionally be possible to look for users in various areas at once, Telegram seems to limit how rapidly an account can carry out these location-based searches. There probably is a ten-minute cooling delay as of this writing between location queries made with the same account. In a later version, the creators of Telepathy are looking into ways to provide location-based searches with numerous accounts.

##### Advanced Network Analysis :

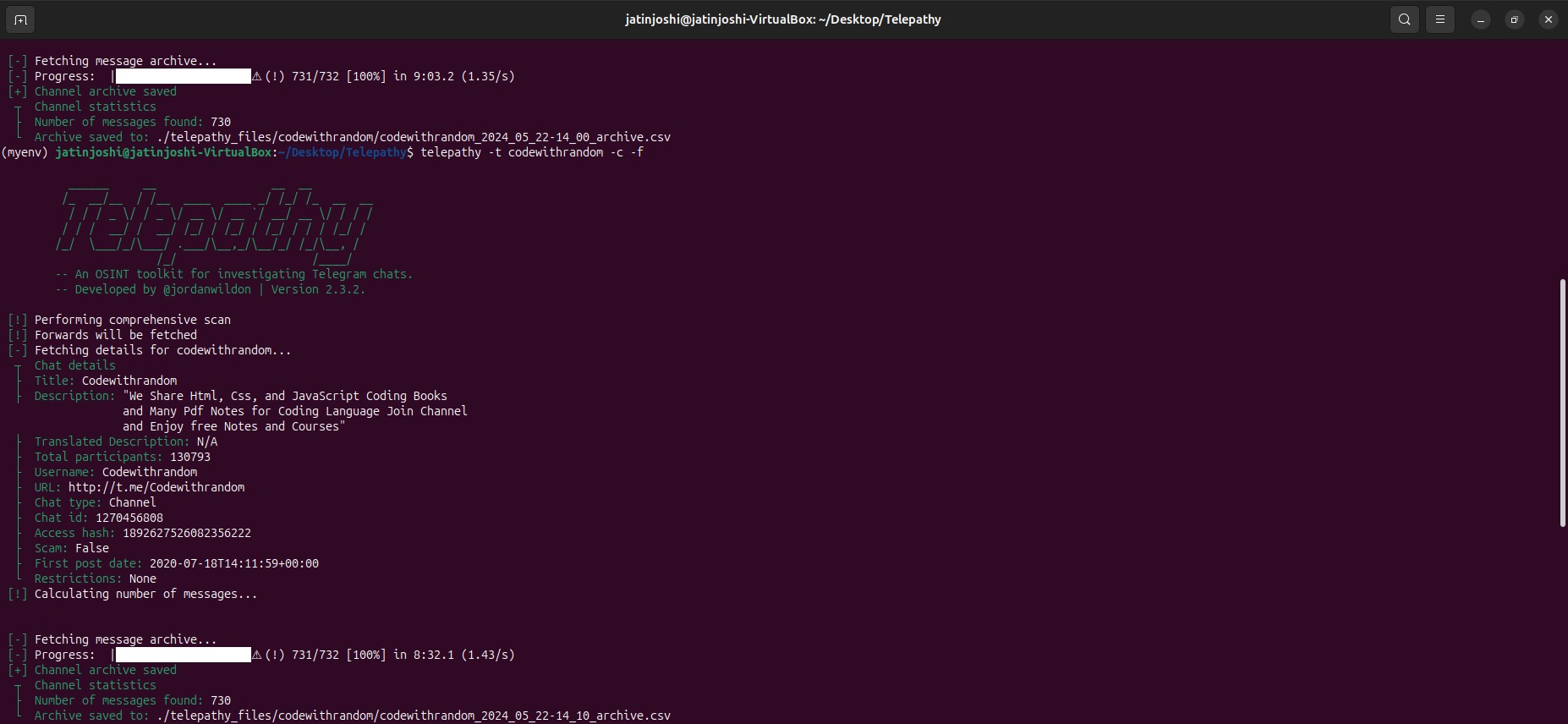
Telepathy may provide "edge lists" that show the routing of messages among users for academics who are interested in the flow of information inside a group [7]. Network analysis tools such as Gephi may be used to display this data, giving insights into the content dissemination mechanisms and users who are key players in the content sharing process.

#### Result:

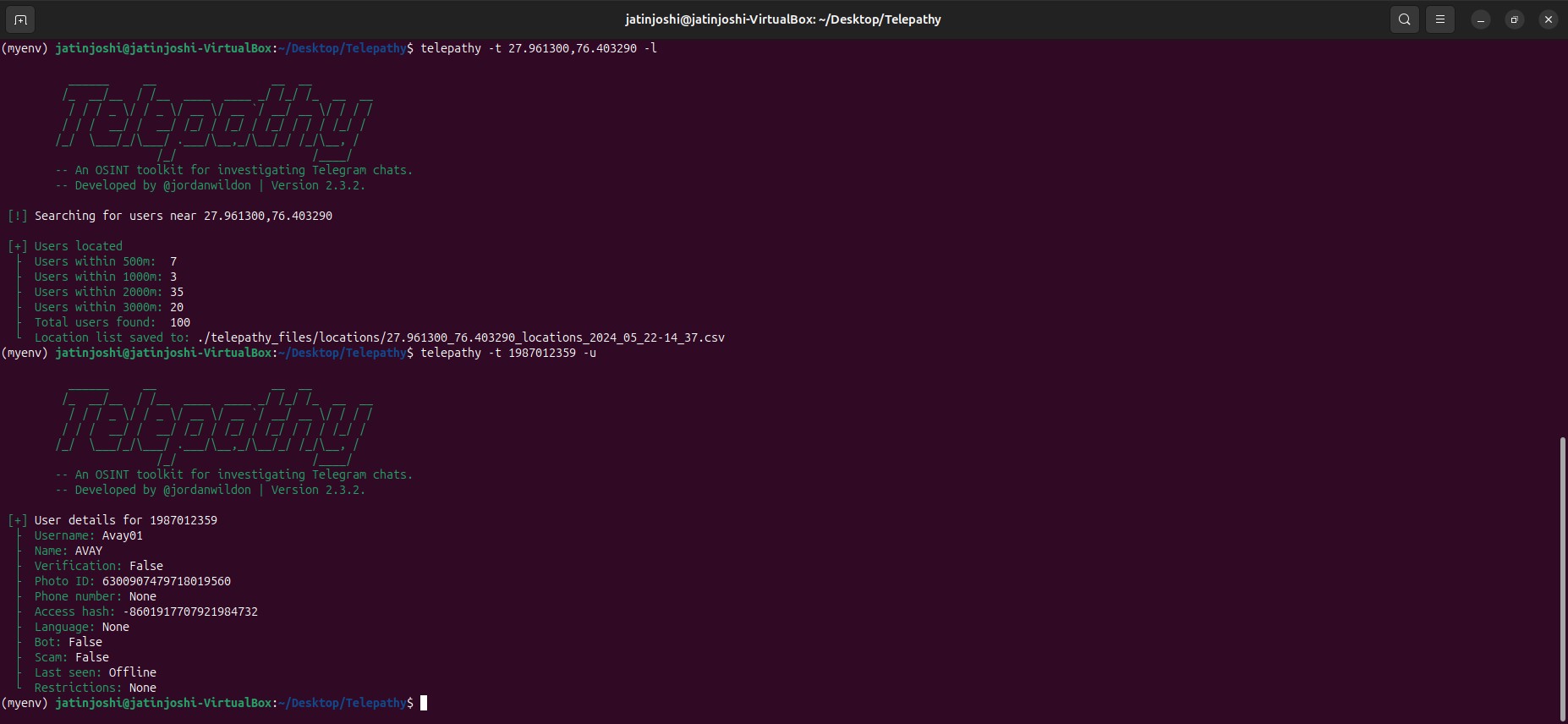
##### Comprehensive chat scan



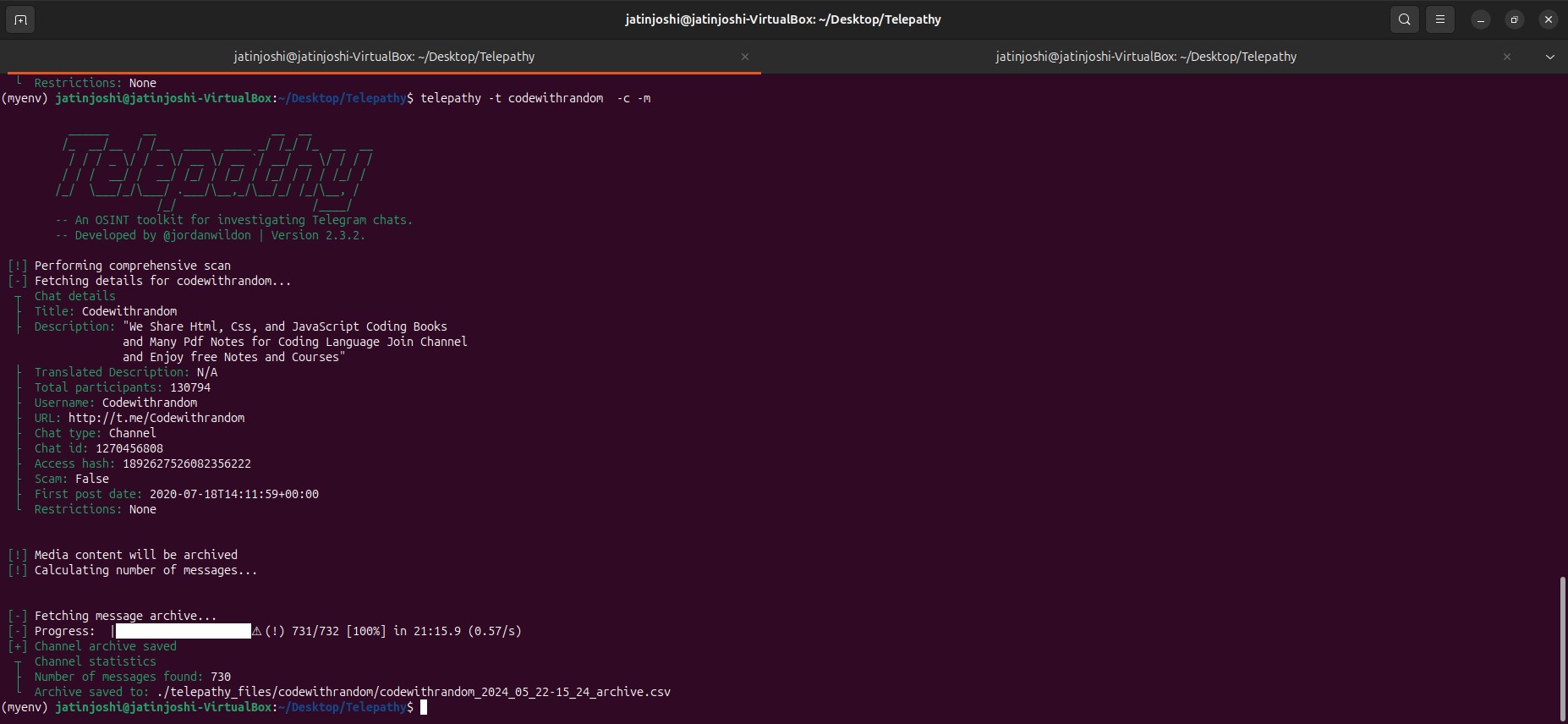
1. **Forwarded messages scan**



##### Finding user near location and Specific user scan



1. **Media scan**



##### Analysis

1. From the below result retrieved after providing the comprehensive scan of the group CodeWithRandom we were successfully able to retrieve the messages sent on the group , from whom they were sent, messageId, Time stamp, what message was sent to the group, reactions, emoji etc.



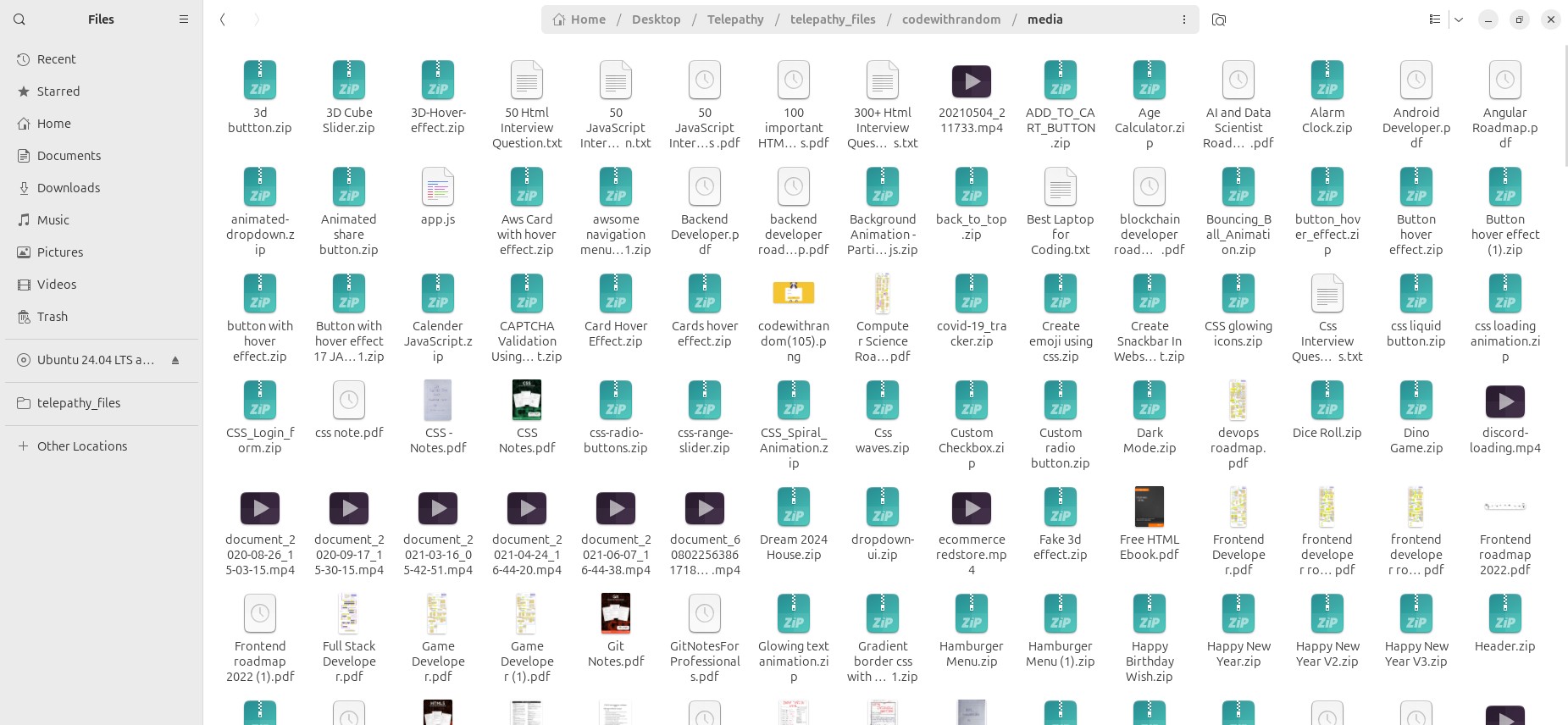
1. From the below result retrieved successfully this shows the messages that were forwarded by which user to which user and also at what time, channelid etc.



1. The below is the result that shows the users that are nearby a particular user here we successfully retrieved the user id, lat and long of the user, date when we retrieved it.From the above since we have retrieved the user id we can also scan a particular user to find whether the user is legitimate or not.



1. From the below result we were successfully able to retrieve the media files that were sent on the group and the files in all the formats were able to be retrieved.



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S.No** | **Tool Name** | **Environment Support** | **Data extraction** | **Limitations** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1 | Telepathy | Linux,MacOs | * Photo * Zipped files * Excel * Pdf * user id * Forwarded messages * Nearby users, timestamp * Message id * Videos * Location | * It can retrieve the memberlist for groups having less than 5000 members. * sqlite.db error on windows |
| 2 | Windows memory extractor | Windows | * Meta data * User Id * Time Stamp * Message Type | * Difficult to implement * Only works for Windows OS * Not user friendly |
| 3 | Telegram Scraper | Windows | User ID | * Only works on Windows. * Do not have a smooth user interface. * Does not mention the username in front of userID. |

Table 2. Summary of Tools used, Data they can extract and Challenges

##### Q. Can we use Telegram scraper on channels in telegram ?

Ans. Yes, you can use Telegram Scraper on channels in telegram. They use premium telegram APIs that help developers to improve telegram scraper software performance. Majority of other software tools use free telegram API or python. Using free APIs will not help you in extracting whole members.

##### Q. How fast does the telegram scraper extract telegram members from groups and channels ?

Ans. Premium version of Telegram Scraper exports Telegram members very quickly.Telegram scraper premium version scrapes members at a rate of approximately 20 to 40 telegram members per minute. However, trial version speed is slow because we are only providing a trial version of software for interface demonstration purposes.

##### Q. Is there a limit to the number of members that can be scraped from Telegram groups or channels using the Telegram scraper?

Ans. Yes, there is a limit for scraping members. It can only extract 50 members, but premium access has no restrictions. Telegram scraper software allows you to scrape all members.

Telegram scraper makes use of the Telegram User API. The Bot API is the recommended and officially supported method of interacting with Telegram programmatically. Telegram scraper uses the Telegram Bot API to send requests to groups and channels and retrieve data. It is not similar to Python scrapers. Python scripts scrape data without using the API, which will result in your Telegram account being blocked.

##### Q. Can we use the Telegram Scraper software tool on different OS other than Windows like Mac or Linux ?

Ans. No, currently this software is only available for Windows OS. It does not support Mac or Linux at this time.

#### Conclusion

Telegram is a well-known cloud-based instant messaging service that has gained popularity due to its security, speed, and adaptability. By 2024, Telegram—which was founded in 2013 by Nikolai and Pavel Durov—will have over 700 million active users worldwide. It provides a number of functions, including audio and video chats, text messaging, media file sharing, and support for big groups and channels. Telegram is a popular option for both personal and business communication because of its emphasis on privacy and encryption.

Understanding and looking into online behaviors and interactions requires knowledge in digital forensics, which focuses on the recovery and analysis of content from digital devices. Because of Telegram's privacy characteristics, forensic analysis has particular difficulties that call for specific tools and methods for data extraction, analysis, and interpretation.

This study highlighted the extraction of contact details, videos, geolocation, and secret conversation data as well as a strong approach for doing Telegram forensic analysis on Linux platforms. The process includes:

**Collection:** Locating and compiling registry entries, media files, cache data, client histories, and Telegram database files.

**Examining:** Extracting pertinent data from database files by parsing and inspecting them with forensic analysis tools.

**Analysis:** Examining registry data, media files, and logs to learn more about user interactions and activities.

**Reporting:** Outlining recovered messages, contacts, media assets, user activities, and registry entries in a thorough forensic report.

The usefulness of a number of tools for Telegram forensic analysis was assessed. The open-source intelligence (OSINT) toolkit Telepathy was emphasized for its capacity to do

advanced network analysis, detailed message analysis, user lookups, and conversation archiving. The limits and capabilities of Telegram Group Scraper and other tools were evaluated in relation to data extraction.

#### Future Scope:

Future developments and enhancements to Telegram forensic analysis will focus on the following areas:

* Improved Support Across Platforms
* Better Methods for Data Extraction
* Automated Reporting and Analysis
* Combining Other Forensic Instruments
* Advanced Privacy Analysis and Encryption
* Alerts and Real-Time Monitoring

#### References:

1. Yusoff, Mohd Najwadi, Ali Dehghantanha, and Ramlan Mahmod. “Network traffic forensics on firefox mobile OS: facebook, twitter, and telegram as case studies.” In Contemporary Digital Forensic Investigations of Cloud and Mobile Applications, pp. 63-78. Syngress, 2017.
2. Gregorio, Jesús, A. Gardel, and Bernardo Alarcos. “Forensic analysis of telegram messenger for windows phone.” Digital Investigation, vol.22, pp. 88-106, 2017.
3. Fernández-Álvarez, Pedro, and Ricardo J. Rodríguez. “Extraction and analysis of retrievable memory artifacts from Windows Telegram Desktop application.” Forensic Science International: Digital Investigation, vol. 40, 2022.
4. Vasilaras, Alexandros, Donatos Dosis, Michael Kotsis, and Panagiotis Rizomiliotis. “Retrieving deleted records from Telegram.” Forensic Science International: Digital Investigation, vol. 43, 2022.
5. Iqbal, Farkhund, Zainab Khalid, Andrew Marrington, Babar Shah, and Patrick CK Hung. "Forensic investigation of Google Meet for memory and browser artifacts." Forensic Science International: Digital Investigation, vol. 43, 2022.
6. Heath, Howard, Áine MacDermott, and Alex Akinbi. “Forensic analysis of ephemeral messaging applications: Disappearing messages or evidential data?.” Forensic Science International: Digital Investigation, vol. 46, 2023.
7. https://github.com/proseltd/Telepathy-Community