Investigating YouTube, TikTok, and Instagram Social Media Interactions on Chrome and Edge Browsers

Kubra Gundogan
Department of Computer Science
Sam Houston State University
Huntsville, Texas, USA
kxg067@shsu.edu

Jeffrey Berg
Department of Computer Science
Sam Houston State University
Huntsville, Texas, USA
jdb118@shsu.edu

Cihan Varol
Department of Computer Science
Sam Houston State University
Huntsville, Texas, USA
cxv007@shsu.edu

2024 12th International Symposium on Digital Forensics and Security (ISDFS) | 979-8-3503-3036-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISDFS60797.2024.10527315

***Abstract*—This paper presents an in-depth digital forensic analysis of user interactions on popular social media platforms, including YouTube, Instagram, and TikTok, through the usage of Chrome and Edge web browsers in both their standard and private browsing modes. The research methodically creates scenarios that include actions such as liking, posting, viewing, commenting, sharing, and direct messaging. We investigate physical and virtual machine environments. Key to our analysis were two approaches: firstly, an examination of data preservation in local, session, and indexedDB storage using the Developer Tools of the browsers; secondly, a thorough inspection of the contents in local storage directories using the Magnet Axiom tool. Despite employing private browsing modes, our findings were significant. The Magnet Axiom tool successfully extracted revealing user data, including posted content and usernames, from the specified local storage paths. The fact that data is kept in both browsing modes presents serious privacy issues. This paper clarifies these implications, particularly about social media interactions, and offers important insights into the limitations and effectiveness of current digital forensic practices.**

Keywords—Digital Forensics, Social Media Analysis, Web Browsers, Data Preservation, User Interaction Analysis, Virtual Machine

I. Introduction

Digital forensics has become increasingly crucial with the growth of social media use, requiring a deeper understanding of how user data is managed and stored by web browsers [1]. With platforms like YouTube, Instagram, and TikTok being integral to digital communication, the way these services handle user data under various browsing conditions is of paramount importance [2]. This study is positioned at the intersection of digital forensics and social media, aiming to uncover the specifics of data storage and retrieval in commonly used web browsers.

Numerous social media platforms, such as Instagram, TikTok, and YouTube, attract millions of daily users, who engage with these sites through various means, including apps and web browsers. Our research aims to investigate the extent and nature of the data acquired by these platforms specifically, Instagram, TikTok, and YouTube, when accessed through Chrome and Edge web browsers. We seek to understand the specific information these websites store on these browsers and unravel the processes triggered when users interact, such as liking, commenting, or sharing content from video creators.

The primary objectives of this research are to:

- Investigate the extent of data preservation in local, session, and indexedDB storage in Chrome and Edge browsers during interactions on YouTube, Instagram, and TikTok, in both normal and private browsing modes.
- Analyze the local storage paths on physical and virtual machines to uncover retained user data, specifically focusing on areas where browser Developer Tools showed no change.
- Determine the type and extent of user data retrievable from these local storage paths, with a special emphasis on information extracted using the Magnet AXIOM forensic tool.

This investigation concentrates exclusively on the Chrome and Edge web browsers, chosen due to their widespread usage and significance in the digital landscape. The focal platforms of this study include YouTube, Instagram, and TikTok, specifically selected for their widespread popularity and diverse user interaction functionalities. It is important to note that the research is confined to the exploration of data storage aspects and does not encompass the analysis of network traffic or server-side data management by the mentioned platforms.

This research project makes a significant contribution to the understanding of data dynamics within the context of popular social media platforms: YouTube, Instagram, and TikTok when accessed through the Chrome and Edge web browsers. By focusing on the intricate details of data storage, the study provides valuable insights into the user-interaction mechanisms and information-handling practices employed by these platforms.

The subsequent sections of this paper unfold as follows: Section II provides an overview of related work in the field, Section III delineates the adopted methodology for this study, Section IV presents the results of the data analysis, Section V discusses these findings in the context of existing literature, and Section VI concludes the paper by summarizing the key findings and offering recommendations for future research.

II. Related Work

In the realm of digital forensics, understanding user behavior, and gathering evidence from various sources is paramount. Several studies contribute to this understanding, with a focus on browser forensics, keyword retrieval, network traffic analysis, and more.

A. Browser Forensics

Dija et al. [3] conclude that their framework provides an effective means for acquiring and analyzing browser files during live forensics investigations in Windows systems. The framework is designed to minimize tampering with suspect machines, which aligns with the principles of digital forensics. The authors emphasize the significance of browser forensics in cybercrime investigations, particularly in gathering evidence related to internet activities. This aligns with the broader field of digital forensics, where understanding user behavior and online activity is essential.

Gupta et al. [4] explore the potential forensic value of artifacts left by Discord when used on the Google Chrome browser. They demonstrate that significant data, like payment information and sent messages, can be recovered from browser caches and logs. This data can link user accounts, interaction frequencies, and emotional states through emoji analysis, offering insights into criminal activities. This research underscores the importance of digital forensics in uncovering criminal activities on popular online platforms.

Suma et al. [5] explore the detailed analysis of cache files generated by Google Chrome, providing insight into the potential for gathering crucial cyber forensics information from frequently visited websites. The authors conclude that cache file analysis is a valuable component of cybercrime investigations, providing insights into visited websites and the objects loaded from them. This conclusion underscores the relevance of browser forensics in uncovering important evidence in various cybercrime scenarios.

Nalawade et al. [6] discuss the importance of forensic evidence collection from web browsers, including cache, history, cookies, and download lists. The authors conclude that web browser forensics plays a vital role in both criminal and civil cases involving evidence from internet activities. They evaluate various tools such as WEFA (Web Browser Forensic Analyzer), and ESECarve used in web browser analysis, emphasizing the need for an in-depth examination of web browser artifacts. This conclusion emphasizes the practical aspects of conducting browser-based digital forensics.

While Ohana et al. [7] primarily focuses on the analysis of residual artifacts from private and portable web browsing sessions, it raises important considerations related to privacy and traces left by browsers. The authors' conclusion challenges the notion that private browsing sessions, like Google Chrome Portable, can fully conceal user activity from forensic investigators. Their findings suggest that further data can be recovered on host machines even in the absence of a portable storage device. Although not directly tied to social media analysis, this paper underscores the complexity of web browser forensics and the potential implications for broader digital forensics investigations.

B. Keyword Retrieval

Dija et al. [8] concentrate on the retrieval of searched keywords from web browsers, a key aspect of user behavior analysis. The authors conclude that their methodology for reconstructing searched keywords is valuable for investigators. While the primary focus is on keyword searching, understanding user search behavior holds relevance in social media analysis, where keyword-based searches are often employed to discover content and user interactions.

C. Network Traffic Analysis

Muehlstein et al. [9] delve into the passive analysis of encrypted network traffic, aiming to identify the operating system, browser, and application used by a desktop or laptop computer. The authors conclude that their framework effectively classifies encrypted traffic, revealing user attributes that can be exploited in attacks. Although not directly tied to social media analysis, this paper highlights the capability to extract user information from encrypted traffic, which has broader implications in digital forensics investigations.

In summary, the reviewed literature provides valuable insights into various aspects of digital forensics, including browser forensics, keyword retrieval, network traffic analysis, and the challenges posed by private and portable web browsing. These studies have contributed significantly to our understanding of user behavior and the extraction of digital evidence from diverse sources.

Our study which focuses on the digital forensic examination of data from popular social media platforms such as YouTube, Instagram, and TikTok, forms a distinct category within digital forensics. In our work, we aimed to uncover evidence, patterns, and user interactions within these platforms, contributing valuable insights to the field of digital forensics, particularly in the context of contemporary online social behavior.

III. Methodology

This section presents the methodology adopted for investigating data preservation practices on social media platforms, particularly focusing on the digital forensic aspects of data stored by web browsers. This research explores the field of digital forensics, focusing on the data storage strategies of social media websites like YouTube, Instagram, and TikTok, as well as the forensic implications of these strategies. The central hypothesis is that various user interactions with these platforms leave behind digital traces that can be retrieved and analyzed through forensic methods. To investigate this, we simulated user activities on these platforms and employed a combination of manual and automated forensic techniques to analyze the data retained by web browsers.

A. Experimental Setup and Scenarios

Our study's experimental approach included creating user scenarios for three well-known social media sites: TikTok,

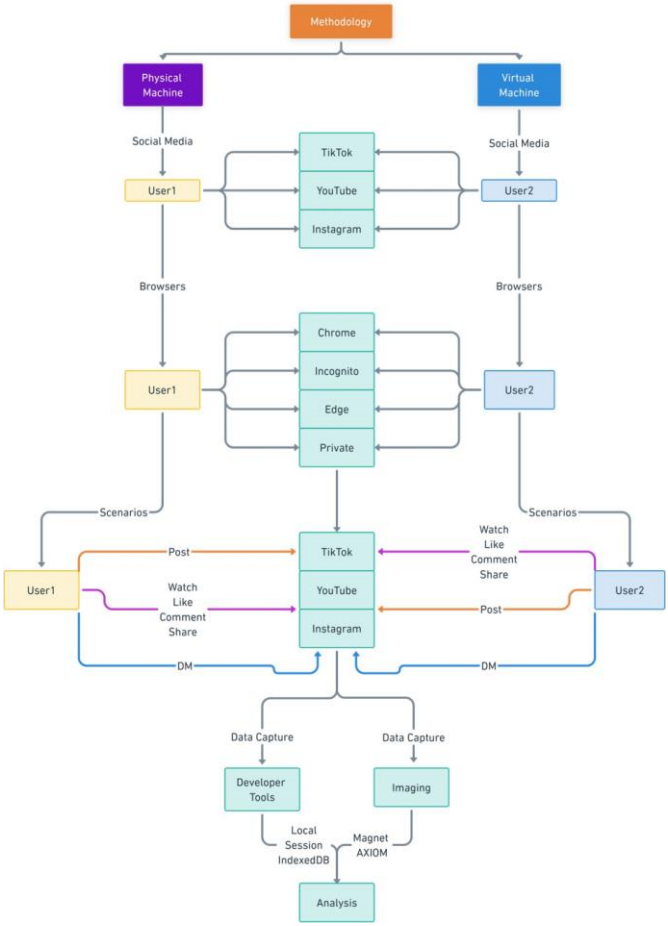

Fig. 1. Methodology

Instagram, and YouTube. These scenarios included various activities involving both content creation and user engagement, such as watching, commenting, liking, and sharing.

Figure 1, provides a comprehensive overview of the digital forensic examination methodologies applied in our study. To ensure a thorough analysis, we established two sets of Google accounts for each platform: one exclusively for content creation and another dedicated to engagement activities. To investigate the role of browser-based data preservation, we selected Google Chrome and Microsoft Edge as our experimental browsers due to their widespread use and distinct data handling mechanisms. Because each experiment was run in both the standard and incognito/private modes on these browsers, we were able to compare the patterns of data preservation in various browsing scenarios.

Our methodology comprehensively organizes different scenarios by detailing the social media platforms examined, the range of user actions performed on these platforms, and the respective digital forensic analysis techniques applied. We explore each scenario independently to offer insights into how various platforms handle user data, the nature of user interactions, and the effectiveness of forensic tools in data retrieval and analysis.

Adding to the complexity of our research, we conducted experiments in two distinct computing environments: a typical user's Windows personal computer and a controlled virtual machine setup using Parallels on a MacBook. This division enabled us to analyze how different computing environments impact data preservation and traceability.

Furthermore, we enacted the following specific procedures for a more comprehensive examination:

- Platform Interactions: We observed the same scenarios across the new TikTok, Instagram, and YouTube accounts, each linked to two different Google accounts. For YouTube, one account initiated the interaction by sharing a post, while the other account engaged by watching, liking, commenting, and sharing. A similar approach was employed for Instagram, with one account publishing a post and the other conducting the same actions. This procedure was repeated for TikTok. Throughout these interactions, we closely monitored and documented changes in local, session, and IndexedDB data storage.
- Direct Messaging (DM): We also examined data preservation during direct messaging interactions on Instagram. In this phase of the study, we sent DMs between the accounts and assessed changes in local, session, and IndexedDB data with each message exchange.
- Browser Modes: All the scenarios described above were executed in both normal and incognito/private modes of Chrome and Edge browsers to comprehensively analyze data preservation patterns under different browsing settings.
- Data Preservation: All data generated during these experiments was locally stored on our personal computers, preserving it for subsequent examination using the Magnet Tool.

B. *Procedures*

In this part, we provided a detailed account of the procedures followed in our research, encompassing data collection and forensic analysis, acquisition methods, and data analysis and interpretation.

1) Data Collection and Forensic Analysis: The goal of the systematic collection of data process was to record a range of user interactions. Regular engagement on the selected social media platforms was carefully recorded, with specific attention to the digital footprints left by these activities. The examination was not limited to visible browsing history but extended to the backend directories of the chosen web browsers. The key aim was to trace how, and to what extent, user activities were stored in these directories.

Our forensic analysis was conducted using a combination of browser Developer Tools and Magnet Axiom, which is a specialized forensic software. To find data traces of user activity, it is essential to examine IndexedDB, local and session storage, and other databases. In our systematic data collection process, significant emphasis was also placed on the analysis of .log and .idb files. The .log files, typically recording user interactions and system events, and the Indexed Database

(IndexedDB) files, storing large amounts of structured data, were an in-depth examination for their forensic value. By analyzing these file types, we were able to access a deeper layer of digital footprints, offering a more comprehensive understanding of user behavior and interactions within web browsers [10]. Complementing this, Magnet AXIOM provided an advanced forensic analysis framework, enabling us to delve deeper into the digital traces and extract related information for our study.

2) Acquisition: The acquisition process in our study was a critical step in the data collection phase. This process involved the systematic capturing of data from the web browsers used during the experiments. We focused on acquiring all relevant data that could potentially hold traces of user activities on the selected social media platforms. This included browsing history, cache files, cookies, and any local storage used by the browsers. The acquisition was performed using both manual methods and specialized forensic tools to ensure the completeness and integrity of the collected information data.

3) Data Analysis and Interpretation: In the data analysis phase of our research, we concentrated primarily on the changes observed within the browser Developer Tools and the insights gathered from the Magnet AXIOM forensic tool. This approach allowed us to precisely pinpoint the nature and extent of data preservation resulting from user interactions on social media platforms.

- Browser Developer Tools Analysis: Our initial focus was on analyzing the modifications within the browser Developer Tools, specifically after user interactions on YouTube, Instagram, and TikTok. We monitored and documented any changes in the local and session storage, as well as in the IndexedDB. This process enabled us to identify the types of data retained by the web browsers following user activities. The comparison of these findings between Google Chrome and Microsoft Edge provided insights into how each browser handles user data differently.
- Magnet AXIOM Forensic Analysis: The utilization of Magnet AXIOM played a crucial role in our research. This advanced forensic tool allowed us to delve deeper into the data remains and extract detailed information regarding the user activities on these platforms. We analyzed the data extracted by Magnet AXIOM for both the personal computer and virtual machine environments. Subsequently, a more in-depth analysis was conducted by examining the local storage directories on both environments from "C:\Users\UserName\AppData\Local\Microsoft\Edge\User Data" and "C:\Users\UserName\AppData\Local\Google\Chrome\User Data". This was to uncover any user data and activity traces that might not be visible or accessible through the browser's Developer Tools. We utilize local files on Magnet AXIOM data imaging.

The comparison of findings between these two environments was particularly informative. We assessed what specific types of information were available in one environment but not in the other, offering a unique perspective on how different computing environments influence data preservation and traceability in digital forensics.

IV. Result

Our comprehensive analysis of various browsers and modes has provided valuable insights into user digital footprints with Magnet AXIOM. Below, we present detailed findings for each platform and browsing mode, with a particular focus on data preservation specifics.

A. Physical Machine Results

1) Chrome: Chrome on the physical machine retained an extensive array of data. This included web visits, with comprehensive tracking of user navigation through transition types, web history displaying titles and visit times, and potential browser activities. Autofill data were extensively captured, containing usernames for social media and email accounts with both typed and suggested values. Cache records and cookies were thoroughly logged, alongside social media URLs and favicon information. Detailed Google search records, session storage logs with timestamps and URLs for social media websites, and local storage data, such as web push permission timestamps, were also observed; these details were specifically found in the "session and local storage .log files." Additionally, Chrome managed to log login credentials and tokens, identifying usernames and encrypted passwords, located in the "session .idb files." The forensic analysis revealed parsed search queries with webpage titles, Google keyword search terms, shortcuts, and text documents related to social media websites. Notifications for user interactions such as comments and likes, stored in the "platform notifications .log file," were also observed. Chrome's data preservation extended to include thumb snapshots of videos watched, liked, commented on, or shared, and network action predictor information which provides data on browser predictions based on user activities.

2) Edge: Edge's data preservation on the physical machine showcased a similar depth in recording web visits, history, and potential browser activities. Cache records and cookies were consistently logged alongside social media URLs, and favicons were present. The forensic analysis highlighted that Edge while maintaining a detailed account of parsed search queries with webpage titles, did not capture as extensive autofill information as Chrome. Encrypted files were still present, and images from each video were captured, but unlike Chrome, video snapshots were not available. Interestingly, one picture retrieved from Edge contained metadata revealing that it was edited with Adobe Photoshop CC 2015 on a Macintosh, a clue that the video originated from User 2, who operated on a virtual machine environment using a MacBook. This suggests potential cross-environment data interactions and may serve as a valuable clue in digital forensic investigations, particularly because User 1 does not have any known association with a MacBook, indicating the origin of the content from User 2.

Table I presents a comparison of data preservation capabilities between Chrome and Edge browsers on a physical

TABLE I
DATA ON PHYSICAL MACHINE
CHROME VS. EDGE

Data	Chrome (Physical)	Edge (Physical)
Web Visits and History	✓	✓
Autofill Information	✓	~
Cache Records	✓	✓
Cookies	✓	✓
Social Media URLs	✓	✓
Session Storage	✓	×
Local Storage Data	✓	~
Encrypted Files	✓	✓
User Preferences	✓	×
Network Predictors	✓	×
Video Thumbnails	✓	×
Image Metadata	×	~

machine. In the tables presented, the following symbols are used to represent the extent of data preservation: (✓) indicates that the feature is fully available or the data is extensively retained. (×) denotes that the feature is not available or the data is not retained. (~) shows limited availability or partial data preservation.

B. Virtual Machine Results

1) Chrome: Chrome's performance on the virtual machine mirrored its behavior on the physical machine, demonstrating a comprehensive data preservation strategy. Web visits were an in-depth log, capturing both the type of transition and the titles associated with web pages. The browser's history recorded visit times, painting a detailed picture of user interaction timelines. Autofill data were extensively preserved, including usernames for both social media and email accounts, with precision in capturing both typed entries and autofill suggestions. The analysis uncovered a strong collection of cache records, cookies, and a complete inventory of social media URLs visited by the user. Google search queries were kept, offering a full log of search terms and related page titles. Web push permission timestamps were stored in the local storage, and session storage provided a detailed view of user activity on social media platforms, including URLs and timestamps. Chrome on the virtual machine also stored thumb snapshots from videos, effectively capturing visual records of media that the user interacted with, whether watched, liked, commented on, or shared. This comprehensive data capture in a virtual environment underscores the extensive nature of Chrome's data preservation practices.

2) Edge: Edge on the virtual machine demonstrated a significant level of equality with Chrome in terms of data preservation capabilities. It accurately recorded web visits and captured a full breadth of autofill information, similar to its Chrome counterpart. Cache records and cookies were consistently logged, ensuring a persistent trail of user activities and preferences. Social media URLs were retained along with favicons, which could serve as visual identifiers for visited sites. A particularly notable finding was Edge's preservation of images associated with user activities on social media platforms. These images, potentially embedded with metadata, could reveal additional context about user interactions and the environment from which the content originated. In the context of a virtual machine, this ability to preserve such data highlights the potential for cross-environment analysis in digital forensic investigations. The preservation of parsed search queries within Edge provided further insights into the user's search behavior, complementing the detailed web history records. This data, combined with autofill entries, offered a comprehensive view of the user's online engagements. The similarity in data preservation between the virtual machine and the physical machine instances of Edge indicates a consistent approach by the browser, regardless of the underlying system architecture.

Table II details the data preservation comparison between Chrome and Edge when operated on a virtual machine.

Taking into account the differences in configuration and security settings between the virtual and physical instances of Edge, it might be possible to address the observed difference in data preservation. There may have been more permissive data storage policies or additional features enabled on the virtual machine, such as enhanced syncing or backup functionality, that are not typically active or configured differently on a physical machine. In the virtual environment, data capture is increased due to this variation. The importance of environmental factors in digital forensic analysis is highlighted by such differences.

TABLE II
DATA ON VIRTUAL MACHINE
CHROME VS. EDGE

Data	Chrome (Virtual)	Edge (Virtual)
Web Visits and History	✓	✓
Autofill Information	✓	✓
Cache Records	✓	✓
Cookies	✓	✓
Social Media URLs and Activities	✓	✓
Session Storage Details	✓	✓
Local Storage Data	✓	✓
Encrypted Files	✓	✓
User Preferences	✓	✓
Video Thumbnails	✓	×
Metadata from Images	×	✓

C. Incognito/Private Mode Findings

In Chrome's Incognito mode, we noticed partial data preservation for web visits, autofill information, cache records, cookies, and social media activities, which indicates that not all user data is completely untraceable. On the contrary, Edge's Private mode showed no preservation for these types of data, suggesting a more stringent privacy approach. In contrast, Edge's private mode did not retain any of the data types listed, indicating a potentially higher level of privacy protection. The reasons behind Chrome's partial data preservation could

include the need to maintain session continuity or accidental storage due to sync features. In contrast, Edge seems to eliminate these elements entirely, which might be attributed to different handling of temporary files or a more aggressive privacy policy. The comparison highlights differing privacy protections between the two browsers and underscores the need for users to be aware of the limitations of each browser's private mode. Table III contrasts the data preservation effectiveness of Chrome and Edge browsers in their respective private browsing modes: Incognito for Chrome and Private for Edge.

TABLE III
PRIVACY MODE DATA
CHROME INCOGNITO VS. EDGE PRIVATE

Data	Chrome (Incognito)	Edge (Private)
Web Visits and History	~	×
Autofill Information	~	×
Cache Records	~	×
Cookies	~	×
Social Media URLs and Activities	~	×
Session Storage Details	×	×
Local Storage Data	×	×

V. CONCLUSION

Our research investigated digital forensic aspects of data preservation across YouTube, Instagram, and TikTok. We set up scenarios where two accounts interacted: one publishing content and the other engaging with it. Direct messaging on TikTok and Instagram was also examined. These activities were conducted using Chrome and Edge, in both normal and private browsing modes, and replicated on a user computer and a Virtual Machine. Analysis using browser Developer Tools revealed no significant changes in local, session, and IndexedDB storage. However, local directories showed storage of user-related data like posted content and usernames. The forensic tool Magnet AXIOM was instrumental in revealing information that we otherwise would not be able to find. When we used Developer Tools, it didn't show everything since it doesn't open the database files directly like Magnet AXIOM does. We found different information when it comes to both web browsers: Edge did not store prefetch files or browser notifications and Chrome did not store pictures when it came to analyzing them in Magnet. Potential applications of this research are to see how we can limit the amount of information that the web browsers need to store to allow the user to have more privacy but won't lose the efficiency of the websites. Future research could entail analysis of more web browsers to see how they all differ from each other, which stores the most information, and which stores the least so we can optimize the privacy of the users.

REFERENCES

[1] P, Vivekananth., The Role of Social Media Forensics in Digital Forensics (August 28, 2022). http://dx.doi.org/10.2139/ssrn.4202753

[2] O. K. Dwivedi, E. Ismagilova, D. L. Hughes, J. Carlson, R. Filieri, J. Jacobson, V. Jain, H. Karjaluoto, H. Kefi, A. S. Krishen, V. Kumar, M. M. Rahman, R. Raman, P. A. Rauschnabel, J. Rowley, J. Salo, G. A. Tran, and Y. Wang, "Setting the future of digital and social media marketing research: Perspectives and research propositions," in Int. J. Inf. Manage., vol. 59, 102168, 2021.

[3] S. Dija, V. Indu, A. Sajeena and J. A. Vidhya, "A Framework for Browser Forensics in Live Windows Systems," 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Coimbatore, India, 2017, pp. 1-5, doi: 10.1109/ICCIC.2017.8524412.

[4] K. Gupta, C. Varol, and B. Zhou, "Digital forensic analysis of discord on google chrome," Forensic Science International: Digital Investigation, vol. 44, p. 301479, 2023.

[5] G. S. Suma, S. Dija and A. T. Pillai, "Forensic Analysis of Google Chrome Cache Files," 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Coimbatore, India, 2017, pp. 1-5, doi: 10.1109/ICCIC.2017.8524272.

[6] A. Nalawade, S. Bharne and V. Mane, "Forensic analysis and evidence collection for web browser activity," 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), Pune, India, 2016, pp. 518-522, doi: 10.1109/ICACDOT.2016.7877639.

[7] D. J. Ohana and N. Shashidhar, "Do Private and Portable Web Browsers Leave Incriminating Evidence? A Forensic Analysis of Residual Artifacts from Private and Portable Web Browsing Sessions," 2013 IEEE Security and Privacy Workshops, San Francisco, CA, USA, 2013, pp. 135-142, doi: 10.1109/SPW.2013.18.

[8] S. Dija, J. Ajana, V. Indu and M. Sabarinath, "Web Browser Forensics for Retrieving Searched Keywords on the Internet," 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 2021, pp. 1664-1668, doi: 10.1109/ICAC3N53548.2021.9725457.

[9] J. Muehlstein et al., "Analyzing HTTPS encrypted traffic to identify user's operating system, browser and application," 2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 2017, pp. 1-6, doi: 10.1109/CCNC.2017.8013420.

[10] Paligu, Furkan, and Cihan Varol. 2020. "Browser Forensic Investigations of WhatsApp Web Utilizing IndexedDB Persistent Storage" Future Internet 12, no. 11: 184. https://doi.org/10.3390/fi12110184