Enhancing Digital Forensic Investigation: A Focus on Compact Electronic Devices and Social Media Metadata

Ibnu Rohan Tuharea¹, Ahmad Luthfi², Erika Ramadani³

¹²³Master Program in Informatics, Faculty of Industrial Technology, Universitas Islam Indonesia
Email: ¹ibnu.tuharea@students.uii.ac.id, ²ahmad.luthfi@uii.ac.id, ³erika@uii.ac.id

Abstract

The rise of portable electronic devices and social media has led to new criminal activities, necessitating advancements in digital forensics. This paper introduces Small-Scale Digital Device Forensics (SSDDF), focusing on the forensic examination of miniature digital devices often used in crimes. SSDDF addresses the challenges posed by these devices, particularly in extracting and analyzing data from them. A key aspect of this research is exploring ontology in social media forensics, particularly within the Android operating system. This involves extracting digital evidence like user accounts, messages, and images from social media platforms. While the paper primarily focuses on social media data, it acknowledges the importance of the devices used in crimes. The integration of SSDDF and the analysis of Android system structures are highlighted as key advancements in digital forensic methodologies. These enhancements are expected to improve the process of collecting and analyzing digital evidence from both compact electronic devices and social media platforms. The study offers significant contributions to the field of digital forensics. It provides new strategies for more efficient and effective forensic investigations, especially in the context of extracting and analyzing digital evidence from small electronic devices and social media, thus paving the way for more robust digital evidence handling in future forensic inquiries.

Keywords: ontology, SSDDF, social media, digital forensic, smartphone

1. INTRODUCTION

The utilization of the internet and social media has witnessed consistent growth year after year. As of February 2022, Indonesia's digital resource usage statistics, reported by We Are Social and cited from datareportal.com, revealed a total population of 277.7 million, with 204.7 million individuals (73.7%) being internet users, and 191.4 million (68.9%) actively engaged in social media. This represents a substantial 12.6% increase from January 2021. Among individuals aged 16 to 64, the most popular social media platforms include WhatsApp (88.7%), Instagram (84.8%), Facebook (81.3%), TikTok (63.1%), Telegram (62.8%), and Twitter (58.3%) [1].
This surge in digital connectivity has given rise to various social media phenomena and behaviors over time, profoundly impacting personal lives, communication dynamics, and even criminal activities. These include trends such as self-disclosure, the emergence of new behaviors like taking selfies, cyber warfare, online shopping, user personalization, the culture of sharing, phubbing, social media for entertainment, cyberbullying, the spread of hoaxes, narratives of terrorism and radicalism, and various other criminal cases connected to social media [2]–[9].

Within the domain of digital forensics, ensuring the admissibility of evidence in a court of law involves several critical stages [10]. While the specific order and number of these stages can vary based on case specifics and expert opinions, four of them remain particularly essential: acquisition, examination, analysis, and presentation [11]. In the context of social media, two primary sources contribute digital evidence: the devices owned by victims or suspects (referred to as clients) and the service providers (servers). These sources assume a central role during the acquisition stage, serving as the cornerstone for all subsequent investigative and analytical efforts.

As we delve deeper into the realm of digital evidence, researchers frequently turn to ontological models to establish knowledge bases that support the analytical process. Notably, David Christopher Harrill and Richard P. Mislan introduced the Small-Scale Digital Device Forensics (SSDDF) ontology, a framework further integrated into the Device Forensic sub-ontology by Nickson M. Karie, M.Sc, and Hein S. Venter, Ph.D [12], [13]. These ontological models serve a pivotal role in structuring and organizing the landscape of digital evidence.

While some ontologies have been developed to address digital forensics in the context of social media, a notable gap persists. This gap becomes evident when examining the work of Edlira Kalemi, Sule Yildirim-Yayilgan, Elton Domnori, and Ogerta Elezaj, who developed the SMoNt ontology specifically related to social media metadata. Their focus revolves around the digital evidence found in social media metadata, which can be considered admissible evidence in court. Despite the wealth of information provided by social media metadata forensics, including diverse entities such as user profiles, messages, status posts, photos, friends, groups, and more [14], [15], these studies do not explore the crucial connection between digital evidence and the electronic devices used.

The field of digital forensics recognizes that acquisition, one of its three fundamental stages [11], involves the collection of electronic devices from suspects and/or victims as evidence, which is then transformed into digital evidence. Despite the availability of separate ontologies for digital devices and social media, a significant void exists when it comes to integrating these essential aspects. This gap impedes the comprehensive and efficient examination of digital evidence in cases involving both small-scale electronic devices and social media.
platforms. By bridging this divide through the development of a unified ontology, this research endeavors to address this critical gap, thus contributing to the advancement of digital forensics and enhancing our capabilities in investigating criminal activities in the digital age.

Therefore, the objective of this research is to construct a novel ontology model that encompasses both dimensions of forensics: the small-scale electronic device aspect utilizing the SSDDF subclass, which defines classes for device types (such as cell phones, smartphones, tablet computers, notebook computers, and others), and the social media data aspect within the social media forensics subclass. By amalgamating these two classes—the small-scale electronic device aspect defined in the SSDDF subclass and the social media data aspect within the social media forensics subclass—this research aims to create a unified ontology. The expected outcome of this research is to establish the interrelationship between mobile devices and social media metadata within the ontology's hierarchy of classes and objects.

2. METHODS

This study employed an experimental methodology to develop and map the ontology, basing its foundation on a comprehensive case study that explored the interaction between a suspect's account (Garry Swihart) with a Samsung Galaxy Mega 2 Android device and a victim's account (Norah Nolan) equipped with a Samsung Galaxy J1 Ace Android device.

2.1 Case Study Implementation

As previously mentioned, and illustrated in Figure 1:

1. Devices and Accounts
   a. Account A (Victim – Fictitious): Norah Nolan, a fictional character depicted as a regular social media user, utilizes a Samsung Galaxy J1 Ace Android device as her primary means of accessing and engaging with online content.
   b. Account B (Suspect – Fictitious): Garry Swihart, another fictional character portrayed with potential suspicious activities on social media, uses a Samsung Galaxy Mega 2 Android device as his primary tool for digital interactions.

2. Interaction within the Facebook Platform: The initial contact occurred when the fictitious suspect (Garry Swihart) initiated a friend request to the fictitious victim (Norah Nolan) on the Facebook platform. Subsequently, he
extended an invitation to join a group titled "Branded Bags & Accessories," where he held the position of group administrator. This marked the commencement of their communication within this social media ecosystem.

3. **Case:** Within the Facebook group "Branded Bags & Accessories," the fictitious suspect (Garry Swihart) strategically posted content designed to capture the attention of the fictitious victim (Norah Nolan). These posts aimed to pique her interest, leading to further engagement through comments. This initial interaction within the group eventually evolved into private messages.

4. **Investigation:** Responding to the escalating interaction between the fictitious victim (Norah Nolan) and the fictitious suspect (Garry Swihart) through private messages, investigators initiated the process of evidence gathering. This involved meticulous collection and preservation of all pertinent digital communications, including text messages, multimedia files, and associated timestamps. The investigation aimed to construct a comprehensive timeline of interactions, identify potential evidence of illicit activities, and scrutinize the intentions and actions of both fictitious parties.

2.2 Device Conditions and Limitations

This research utilized two Samsung Android smartphones, namely the Samsung Galaxy Mega 2 and Samsung Galaxy J1 Ace. These devices differed in software specifications. The Samsung Galaxy Mega 2 ran on Android version 4.4.4 (KitKat), while the Samsung Galaxy J1 Ace operated on Android version 5.1.1 (Lollipop). A detailed overview of device specifications is provided in Table 1.

<table>
<thead>
<tr>
<th>Brand &amp; Type</th>
<th>Device Code</th>
<th>Android Version</th>
<th>Username</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samsung Galaxy Mega 2</td>
<td>vasta3g</td>
<td>4.4.4 (KitKat)</td>
<td>Norah Nolan</td>
<td>Victim</td>
</tr>
<tr>
<td>Samsung Galaxy J1 Ace</td>
<td>j1acevelte</td>
<td>5.1.1 (Lollipop)</td>
<td>Garry Swihart</td>
<td>Suspect</td>
</tr>
</tbody>
</table>

Conditions observed for the two devices:
1) for the Android KitKat version (used by the victim): The Facebook application (com.facebook.katana or FB) is unavailable in the Play Store. Instead, Facebook Lite (com.facebook.lite or FBL) serves as a lightweight alternative for installation.

2) for the Android Lollipop version (used by the suspect): The Facebook application is available in the Play Store without obstacles encountered, in contrast to previous Android versions.
2.3 Data Acquisition

The following the implementation of the case study, data acquisition was performed using the following methods:

1) rooting both devices using Magisk, Magisk, a suite of open-source software, was utilized for comprehensive access to the device's data and applications almost all version of Android with more than 260 contributors on the project development [16], [17]. Rooting allows elevated privileges, enabling the extraction of meaningful data that might otherwise be inaccessible [18].

2) conducting the acquisition of both devices using the dd method, the dd method was employed to create a bit-by-bit copy of the device's storage, ensuring the acquisition of extensive data [19], including text messages, images, application data, and system files.

3) transferring the acquired data from the Android Debug Bridge (ADB) shell to the host computer using nc (Netcat) command installed by Busybox. The acquired data was transferred from the ADB shell to the host computer using the Netcat (nc) command installed by Busybox. ADB, an open-source tool, allows command-line operations, including installing and debugging apps on Android devices [19], a common tools like [18], [20], [21]. Busybox [22] was used to install and run the Netcat command on both devices for networking operations.

4) The transferred data acquisition results are in the .dd file extension. The transferred data, in the form of '.dd' files, represented the acquired data in raw disk image format, a standard extension in digital forensics.

2.4 Challenges and Limitations

The case study revealed several challenges and limitations:

1) Data Limitations for FB and FBL, one challenge was the limited scope of acquired data from Facebook (FB) and Facebook Lite (FBL). While FB provided contact and local media data as evidence, FBL had more restricted data access, offering only contact data.

2) Inclusion of Messenger (FBM), another challenge involved integrating an additional application, Messenger (com.facebook.orca or FBM). Extracting data regarding private messages or conversations between the two accounts presented challenges despite Messenger offering specific functions for sending messages between FB users.

These challenges and limitations significantly influenced the data collection process and should be considered when interpreting the research findings. They underscore the complexities and nuances involved in digital forensics and the acquisition of data from social media platforms.
2.5 Ontology Mapping

To understand the fundamental process of ontology mapping, it's essential to highlight that ontology mapping is the process of establishing relationships between concepts and terms in different ontologies. In this context, the Small-Scale Digital Device Forensic (SSDDF) Ontology was developed and mapped to the data structures and evidence categories obtained in the case study. The mapping process involved:

1) **Identification of Concepts**: The first step was identifying key concepts within the acquired data, such as storage blocks, file structures, device specifications, user profiles, messages, and various digital evidence categories.

2) **Creation of Classes and Properties**: Based on the identified concepts, appropriate classes and properties were established within the SSDDF Ontology. These classes and properties were crafted to represent the discovered data and its relationships accurately.

3) **Integration with Existing Ontologies**: To enhance the utility and relevance of the SSDDF Ontology, it was integrated with existing ontologies, such as the Social Media Evidence Ontology. This integration facilitated a more comprehensive and enriched contextual analysis of digital evidence.

4) **Mapping Classes and Properties**: The actual mapping process involved associating classes and properties in the SSDDF Ontology with equivalent or related concepts in the acquired data. For example, data related to internal storage and Facebook applications were mapped to corresponding classes and properties within the ontology.

By adopting these steps, the SSDDF Ontology effectively captured and organized digital evidence from small-scale digital devices, enabling structured analysis and enhancing the efficiency of digital forensic investigations. The mapping process harmonized the ontology with the practical needs of digital forensic analysis, allowing investigators to navigate and understand the intricate data structures within these devices more effectively. These are the steps and principles that guided the ontology mapping process in this study.

3. RESULTS AND DISCUSSION

3.1 Practical Application of SSDDF Ontology

The Small-Scale Digital Device Forensic (SSDDF) Ontology, depicted in Figure 2, plays a central role in categorizing and organizing digital evidence from small-scale digital devices. This ontology serves as a foundation for exploring digital forensic investigations on devices like smartphones, memory cards, and embedded systems.
In this section, we illustrate the practical utility of the Small-Scale Digital Device Forensic (SSDDF) ontology in real-world digital forensic investigations. The SSDDF ontology serves as a valuable tool for law enforcement agencies and forensic experts, enhancing their capabilities in the following ways:

1) **Streamlined Data Analysis**: The SSDDF ontology offers a structured framework for organizing and categorizing digital evidence from small-scale digital devices. Its predefined classes and relationships facilitate efficient data analysis, enabling investigators to gain quicker insights.

2) **Cross-Device Correlation**: In multi-device investigations, the SSDDF ontology enables the correlation of data from various sources. It allows investigators to link evidence from smartphones, memory cards, and embedded devices, reconstructing a comprehensive timeline of events.

3) **Enhanced Data Retrieval**: With well-defined classes and properties, the SSDDF ontology simplifies data retrieval. Investigators can quickly locate
relevant information, such as chat histories, media files, or user profiles, leading to more effective case resolutions.

4) Integration with Existing Tools: The SSDDF ontology is designed to be integrated seamlessly with existing digital forensic tools and software. This integration enhances accessibility for forensic experts, streamlining the investigative process without requiring extensive retraining.

3.2 Development of Small-Scale Digital Device Forensic Ontology

In developing the SSDDF ontology, the exploration focused on the Android system's storage blocks, specifically mmcblk0 for internal storage and mmcblk1 for external storage. Figure 3 illustrates these blocks on the Samsung Galaxy J1 Ace device.

![Figure 3. Block internal storage (mmcblk0) and external storage (mmcblk1) on the Samsung Galaxy J1 Ace device.](image)

Within the internal storage's userdata partition, encompassing "data," "media," and "system" folders, the investigation sought findings related to Facebook applications (FB, FBL, and FBM). Figure 4 presents the userdata partition structure and the list of Facebook application packages.

![Figure 4. Userdata partition structure and Facebook application package list.](image)
Based on the folder and file structures, extensive data related to the activities of both accounts was discovered. Table 2 outlines some of the key findings of digital evidence obtained from the implementation of the case study.

<table>
<thead>
<tr>
<th>App Code</th>
<th>File</th>
<th>File Path</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Samsung Galaxy Mega 2 (vasta3g)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>com.facebook k.lite</td>
<td>fblite_dat</td>
<td>/vol_vol30/data/com.facebook.lite/databases/fblite_database.db</td>
<td>This database contains critical data from Facebook Lite, including user profiles, messages, and activity logs.</td>
</tr>
<tr>
<td>com.facebook k.orca</td>
<td>msys_dat</td>
<td>/vol_vol30/data/com.facebook.orca/databases/msys_database_10000905414_1448</td>
<td>The Messenger system database, which holds chat histories, attachments, and contact information.</td>
</tr>
<tr>
<td><em>Samsung Galaxy J1 Ace (jacevelte)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>com.facebook k.katana</td>
<td>app_upload</td>
<td>/vol_vol30/data/com.facebook.katana/app_upload</td>
<td>A directory where the Facebook app uploads media files, including photos and videos shared by users.</td>
</tr>
<tr>
<td>com.facebook k.katana</td>
<td>contacts_</td>
<td>/vol_vol30/data/com.facebook.katana/databases/contacts.db2</td>
<td>A database that stores contact information from the Facebook app, aiding in contact tracing and connections analysis.</td>
</tr>
<tr>
<td>com.facebook k.katana</td>
<td>local_media</td>
<td>/vol_vol30/data/com.facebook.katana/databases/local_media_db</td>
<td>This database houses locally stored media files shared on Facebook, offering insights into user media preferences.</td>
</tr>
<tr>
<td>com.facebook k.katana</td>
<td>authentication</td>
<td>/vol_vol30/data/com.facebook.katana/app_light_prefs/com.facebook.katana/authentication</td>
<td>Authentication preferences data, crucial for understanding user login patterns and security measures.</td>
</tr>
<tr>
<td>system</td>
<td>accounts.db</td>
<td>/vol_vol30/system/users/0/accounts.db</td>
<td>The system-level accounts database that holds information about user accounts on the device.</td>
</tr>
</tbody>
</table>

From these findings, it is evident that the SSDDF ontology requires additional structured classes to effectively map the discovered data. The complexity of existing data structures within devices raises the competency question: "How to map the Small-Scale Digital Device Forensic ontology?" This question forms an
initial assumption considering the intricate nature of data structures within the devices.

The presence of storage blocks (mmcblk0 and mmcblk1) suggests that the primary focus of exploration should be on internal storage, with external storage being optional. Within internal storage, the data, media, and system folders indicate a need for separate classes to map these directories for clearer identification and categorization. As a result, the competency question can be refined into the following sub-questions:

1) **How to map based on the nature of Internal and External storage?**
   Justification: Distinguishing between internal and external storage elements is crucial for precise data categorization. This differentiation aids investigators in focusing their analysis on relevant data sources and ensures that the ontology accurately reflects the nature of the storage medium.

2) **How to map the important structures present in internal storage?**
   Justification: Internal storage contains various critical structures (e.g., data, media, system) that require separate mapping. These structures are key to organizing and categorizing digital evidence effectively.

3) **How to map the structures present in external storage?**
   Justification: While the primary focus is on internal storage, external storage remains a potential source of digital evidence. Including classes to map external storage ensures investigators can account for and analyze data stored on removable media when relevant.

By offering these justifications, our objective is to elucidate the reasoning behind selecting and incorporating particular classes and properties into the SSDDF ontology. These additions have been specifically crafted to harmonize with the practical needs of digital forensic analysis, streamlining investigations for greater efficiency and effectiveness.

### 3.3 Implementation of Competency Questions

Competency questions play a crucial role in shaping the SSDDF ontology. They guide the development process to ensure the ontology addresses the needs of digital device forensic analysis. The original class structure serves as a foundation, with additional class structures created based on the competency questions to enhance relevance and utility.

For the first competency question ("How to map based on the nature of Internal and External storage?") Figure 5 provides a visual representation of the added class for internal and external storage.
Figure 5. Addition of a class for the first competency question in the SSDDF ontology.

For the second competency question ("How to map the important structures present in internal storage?"), Figure 6 illustrates the added classes.

Figure 6. Class addition for the second competency question in the SSDDF ontology.

For the third competency question ("How to map the structures present in external storage?"), Figure 7 displays the added classes.

Figure 7. Class addition for the third competency question in the SSDDF ontology.

The complete SSDDF ontology resulting from the implementation of the competency questions is visualized in Figure 8.
Integration with existing ontologies, such as the Social Media Evidence ontology, contributes to a holistic approach to digital device forensic analysis. The SSDDF ontology seamlessly converges with the established Social Media Evidence ontology, creating a unified knowledge framework. This integration leverages the strengths of both ontologies, with SSDDF focusing on structuring digital device forensic data and Social Media Evidence specializing in capturing metadata and contextual information from social media platforms.

The methodological integration involves mapping relevant classes and properties from each ontology to ensure compatibility and data interoperability. Despite challenges such as reconciling differences in class definitions and handling overlapping properties, the benefits far outweigh these challenges. Figure 9 provides an overview of the resulting integrated ontology.
3.4 Class Structure

The ontology comprises two major classes: DeviceForensic and SocialMediaEvidence, each with its sub-classes. The DeviceForensic ontology organizes data acquired from small-scale digital devices, covering storage, files, and device characteristics. In contrast, the SocialMediaEvidence ontology handles evidence from social media platforms, including user profiles, messages, connections, and online activities. Each of them has its sub-classes:


This dual ontology approach offers a comprehensive solution for digital forensic investigations, facilitating a comprehensive and enriched contextual analysis of digital evidence.

3.5 Object Properties Structure

In the object properties structure, numerous properties have been added to enrich knowledge within the ontology. These properties cover a range of relationships, such as: administrator, advocate_of, as_account_in, author_of, bank_account_of, check_in, co_author, comment, current_work, dislike, education, eyewitness_of, family_relations, followed, geolocation, going_to, has_additional_item, has_advocate, has_attended, has_author, has_bank_account, has_blocked, has_brother, has_cousin, has_criminal_profile, has_device, has_eyewitness, has_father, has_husband, has_juror, has_mother, has_officer, has_political_status, has_religion, has_sister, has_wife, has_witness, hash_value, inspector_of, interested_in, is_closed, is_open, is_part_of, is_private, is_public, juror_of, like, location, member, mentioned, mentioned_by, officer_of, ownedby, owns, participate_same_riot, past_work, photo_of, published_emotions, published_status, same_organisation, share_via, stored_in, subscribe, tagged, tagged_by, talking_about, transferred_amount, transferred_by, transferred_to, uses_app, vality_url, video_of, visited, witness_to.

3.6 Data Properties Structure

In the data properties structure, several properties have been added, such as directory, has_app, and parent_directory. These properties contribute to mapping individuals related to DeviceForensic.

4. CONCLUSION

This research represents a significant leap forward in the application of ontological frameworks as vital tools for systematically categorizing digital evidence on both server and client endpoints. By employing the Social-Media Digital Evidence
Ontology and introducing the Small-Scale Digital Device Ontology (SSDDF), we've contributed to a nuanced procedural methodology for examining case files related to social media incidents. The first contribution is the meticulous organization of data from server-side service providers using the Social-Media Digital Evidence Ontology, proving to be a robust procedural methodology for scrutinizing case files related to social media incidents. The second contribution, SSDDF, explicitly delineates storage architecture within Android smartphones, playing a pivotal role in categorizing digital evidence obtained from mobile apparatuses or clients.

The practical implications of our research are profound. Implementing SSDDF equips forensic investigators and law enforcement agencies with a structured framework for organizing and categorizing digital evidence from small-scale digital devices. This results in more efficient data analysis, quicker insights, improved cross-device correlation, and enhanced data retrieval. Additionally, the seamless integration of SSDDF with existing digital forensic tools eliminates the need for extensive retraining, streamlining the investigative process. Looking forward, it is essential to expand research on mapping storage systems across various mobile device platforms, extending beyond Android to include iOS-based smartphones, smartwatches, Unmanned Aerial Vehicles (UAVs), and more. Furthermore, there's a need to explore the practical implementation of SparkQL to substantiate intricate interlinkages between server and client data, shedding light on the complex dynamics of data interactions within the digital ecosystem.

REFERENCES


