Comparison Analysis of Free Forensic Programs on SQLite Corpus

Thesis

Submitted to fulfill one of the

Requirements for obtaining a bachelor's degree

In the Informatics Study Program

Faculty of Informatics

Telkom University

1303174059

Indra Jaya Kusuma

Information Technology Study Program

Faculty of Informatics

Telkom University

Bandung

2022

LEMBAR PENGESAHAN

Analisa Performa pada Program Forensik di Corpus SQLite

Performance Analysis of Forensic Programs on SQLite Corpus

NIM :1303174059

Indra Jaya Kusuma

This final project has been accepted and ratified to partially fulfill the requirements for obtaining a degree in the Information Technology Study Program

Faculty of Informatics

Telkom University

Bandung, <19/01/2023>

Menyetujui

Advisor I,	Advisor II,
<Hilal Hudan Nuha, S.T., M.T., Ph.D. >	<Gandeva Bayu Satrya, Ph. D.>
<NIP : 13860093>	<NIP : 13850009>

Head of Informatics Bachelor Program

<Rio Guntur Utomo, S.T., M.T., Ph.D.>
NIP : 19900020

Statement Sheet

I, Indra Jaya Kusuma, hereby declare that my final project entitled "**Performance Comparison of Forensic Programs on SQLite Corpus**" along with all of its contents is of my own, and I do not commit any plagiarism that is not in accordance with the scientific ethics that apply in the scientific community. I am ready to bear the risks/sanctions given if in the future a violation of scientific ethics is found in the TA Report or if there are claims from other parties against the authenticity of the work.

Bandung, <06/02/2023>

The Declared

<Indra Jaya Kusuma>

Performance Analysis of Forensic Programs on SQLite Corpus

Indra Jaya Kusuma

[1,2,3]Faculty of Informatics, Telkom University, Bandung
Jl. Telekomunikasi No.1, Bandung, Indonesia
e-mail : wolferker@students.telkomuniversity.ac.id

Abstract
Tujuan penelitian ini adalah untuk mengevaluasi performa berbagai program forensik pada korpus dengan program SQLite. Penulis mencoba mewakili beragam ukuran dan struktur basis data. Program-program forensik dijalankan pada berbagai versi SQLite. Namun, perlu diperhatikan bahwa karakteristik kinerja spesifik dari program-program tersebut dapat bervariasi tergantung pada karakteristik spesifik dari database yang sedang dianalisis dan versi SQLite tersebut. Secara keseluruhan, penelitian ini bertujuan untuk memberikan wawasan yang berharga tentang kinerja program-program forensik pada basis data SQLite dan dapat digunakan untuk menginformasikan pemilihan alat untuk analisis forensic.
Kata kunci : forensik, korpus, performa, analisis

Abstract
The goal of this study was to evaluate the performance of various forensic programs on a corpus of SQLite databases. Author is trying to represent a diverse range of database sizes and structures. The forensic programs were run on various SQLite version. However, it should be noted that the specific performance characteristics of the programs may vary depending on the specific characteristics of the database being analyzed and versions of said SQLite. Overall, this study aim to provide valuable insights into the performance of forensic programs on SQLite databases and can be used to inform the selection of tools for forensic analysis.
Keywords: forensic, corpus, performance, analysis

1. Introduction

Forensic analysis plays a crucial role in the field of digital forensics, as it allows for the identification and analysis of digital evidence from computers, mobile devices, and other electronic media. The use of forensic software is a common approach to facilitate this process, as it provides a range of tools and techniques to extract, analyze, and report on digital evidence. However, with the increasing number of forensic programs available, it is important to evaluate their performance in order to select the most suitable tool for a given task.

Developments in forensic research, tools, and process over the past decade have been very successful and many in leadership positions now rely on these tools on a regular basis frequently without realizing it.[1]

As such, SQLite is one of the popular tools to make use. For instance, the mobile messaging program WhatsApp, which has over 2.24 billion users on June 2022.[2] and is one of the most popular communication channels worldwide, uses SQLite as its storage format.

SQLite, however, is not just for mobile applications. For instance, SQLite is used by well-known web browsers like Firefox.[3] and Chrome.[4]. In addition to storing application data, SQLite is also used to store meta data, including bookmarks, history, and login information.

Therefore, it is crucial to extract all data connected to a SQLite database from the perspective of IT forensics. The acquisition of deleted SQLite data is challenging, even though the extraction of allocated database information is simple.

As such, there are a few tools that will be introduced in this paper. Firstly, SQLite Expert Professional. SQLite Expert Professional is its intuitive and user-friendly interface, which makes it easy to create, edit, and manage SQLite databases. The software offers a variety of visual tools and wizards that simplify the process of designing and modifying database schema, as well as importing and exporting data. Secondly, SQLite Doctor has advantages that make it an ideal choice for working with SQLite databases. it has a user-friendly interface that allows users to easily navigate through the various functions and tools available. SQLite Browser as a open-source software is actively developed and maintained by a dedicated team of contributors, ensuring that the software remains up-to-date and responsive to the needs of its users. SQLite Studio have an intuitive graphical interface, which allows users to view and modify database schema with ease, The software provides a variety of tools for advanced SQL queries and a built-in query builder. Lastly, SQLite DB Recovery is a specialized tool uses advanced algorithms that can scan even corrupt database file and extract as much data as possible.

The goal of this study was to evaluate the performance of various forensic programs on a corpus with the use of SQLite. SQLite itself is a widely used database management system(DBMS) that is commonly found in wide range of applications, including mobile devices, various operating services, web browsers, and IoT devices[5]. The datasets used in the analysis represented a diverse range of database sizes and structures.

1.1 Problem Identification

Within this studies, we focus testing several SQLite versions in order to compare the capabilities of each versions in order to identify whether those versions are able to process corpus correctly. The problems that will be discussed in this research are listed as :

A. How to perform forensic analysis on SQLite Corpus using SQLite Expert Professional?
B. How to perform forensic analysis on SQLite Corpus using SQLite Doctor?
C. How to perform forensic analysis on SQLite Corpus using SQLite Browser?
D. How to perform forensic analysis on SQLite Corpus using SQLite Studio?
E. How to perform forensic analysis on SQLite Corpus using SQLite DB Recovery?

1.2 Contributions

In this paper, We are using some of the corpus used in "A standardized corpus for SQLite database forensics" (Garfinkel et al) [6], which follows a standardized corpus. The main contribution of this research are the following :

A. Performing forensic analysis on SQLite Corpus using SQLite Expert Professional
B. Performing forensic analysis on SQLite Corpus using SQLite Doctor
C. Performing forensic analysis on SQLite Corpus using SQLite Browser
D. Performing forensic analysis on SQLite Corpus using SQLite Studio
E. Performing forensic analysis on SQLite Corpus using SQLite DB Recovery

1.3 Scope

A. The SQLite Corpus that will be analyzed is from the paper made by Garfinkel et.al. which follows a standardized corpus dataset.
B. The comprehension of analyzing SQLite Corpus using SQLite Expert Professional
C. The comprehension of analyzing SQLite Corpus using SQLite Doctor
D. The comprehension of analyzing SQLite Corpus using SQLite Browser
E. The comprehension of analyzing SQLite Corpus using SQLite Studio
F. The comprehension of analyzing SQLite Corpus using SQLite DB Recovery

1.4 Research Plan

The research plan is depicted by the following figure.

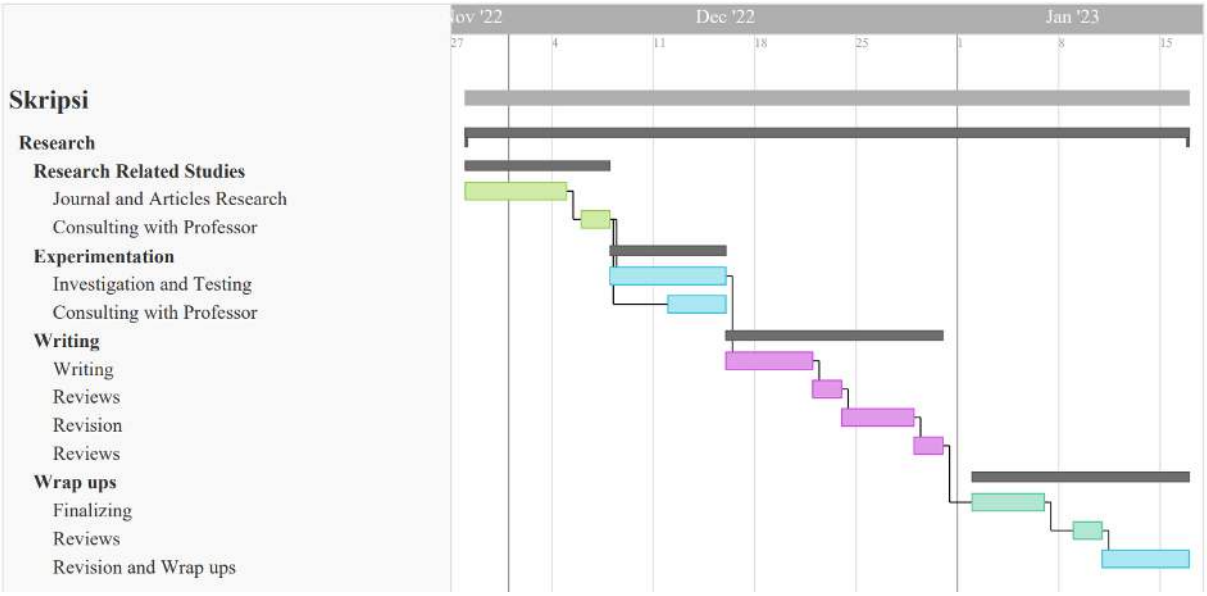

Figure 1. Research Plan Gantt Chart

2. Related Works

This section presents theoritical background and related studies. First section will be discussing about the theoritical background of this research. Related studies will be discussed in the second section.

2.1 Theoretical Background

A. Digital Forensic

The process of identifying, preserving, analyzing, and presenting evidence found in digital devices in a legally admissible manner is known as digital forensics. Digital forensics encompasses a wide range of technologies and techniques that are used in a variety of contexts such as criminal investigations, civil litigation, and incident response.

Digital forensics is a multidisciplinary field that combines computer science, electrical engineering, and forensic science expertise. The digital forensics process is divided into several stages, including the identification and preservation of evidence, the analysis of that evidence, and the presentation of the results of the analysis in a legally admissible manner.

B. DBMS

A Database Management System (DBMS) is a software system that provides a comprehensive set of tools and techniques for data management, storage, and retrieval. The database management system (DBMS) serves as a bridge between the physical data stored in a database and the applications and users who need to access that data.

Data manipulation, which includes operations such as inserting, updating, and deleting data in a database, is an important function of a DBMS. A database management system (DBMS) provides a high-level language, such as SQL (Structured Query Language), that enables users to perform these operations in a simple and efficient manner.

C. SQLite

SQLite is an open source, embedded relational database management system for a wide range of applications. It is lightweight, portable, and simple to use, making it ideal for embedded systems, mobile devices, and desktop applications. Its file-based storage method, reliability, and robustness make it an excellent choice for data management and storage, while its small footprint and low resource consumption make it an excellent candidate for resource-constrained environments and embedded systems.

2.2 Related Studies

Carolin Elisabeth et. Al.(2020)[7] Conduct an Forensic Analysis on Instagram by identifying artifacts left from the android application of said social media. They were able to reconstruct the messages based on the artifacts and locates where the application stores its resources.

Asma Majeed et. Al(2015)[8] Studied three social media apps in Windows 10. In said paper concluded that Facebook, Skype, and Viber has similar parent directory of their artifacts.

Shahzad Saleem et. Al(2017)[9] The research conducted by author of related paper produce a fruitful success. They were able to locate of said user(s) that they conducted along with their informations in form of user activity, text messages, timestamps, and sender/receiver names

Chih Ping Yen et al(2019)[10] Investigated both of web and mobile version with forensic analysis. There, they found out that there is differences on privacy control caused by different browsers. Meanwhile on the mobile application there are differences caused by framework spaces.

Nemetz et. Al.(2018)[11] Conduct a research of a standardized corpora and testing a set of forensic tools. They argued that that without standardized corpora, research efforts and results that are performed on individual data sets are not comparable and therefore less useful. They propose that a more scientific approach is to make data sets available to the public, so that they can be used by different people for different purposes over time. By using publicly available test dataset, new approaches and algorithms for forensic analysis can be directly compared to existing tools. This way, the acceptance of a tool by its users can be based on its quality, as measured by publicly available and reproducible test results. The authors suggest that this approach will make forensic tools more robust, reliable, and trustworthy.

Table I. Related Studies

No.	Case	Method	Comparison	Analysis
1. Elisabeth et.al. 2020	Instagram	SQLite parser GUI	-	Artifacts
2. Majeed et. Al. 2015	Windows 10 Facebook,Skype, Viber	SQLite DB Browser	-	Artifacts
3. Saleem et.al. 2017	Windows 10 MySpace, Twitter, Facebook	SQLite DB Browser	-	Artifacts
4. Chih Ping Yen et.al. 2019	Instagram	SQLite Editor	-	Artifacts
5. Nemetz et.al. 2018	SQLite	SQLite Parser, SQLite Doctor, Phoenix Repair, Undark, DB Recovery, Forensic Browser	-	Corpus elements
6. Our work	SQLite	SQLite Expert Professional, SQLite Doctor, SQLite Browser, SQLite Studio, SQLite DB Recovery	-	Corpus Elements

Table 1 summarized works related to this research, it can be noticed that the topic has not been discussed in other works.

3. Methodology

Before starting the test, it is necessary to confirm the forensic tools we are operating with.

3.1 Research Method

The goal of this research is to analyze the performance of the mentioned SQLite versions on the arranged corpus. To do this, some of the Corpus from Nemetz et.al. were used. Then, we'll evaluate each SQLite attempt to execute that corpus. Once we are done testing said SQLite version we will review the output of each of them. This procedure is illustrated by the following figure below.

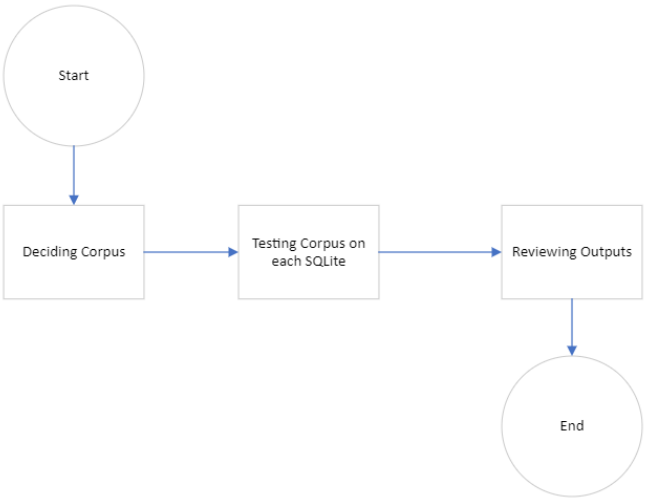

Figure 2. Diagram Flowchart

3.2 Forensic Programs

Using a combination of the SQLite tools SQLite Expert Professional, SQLite Doctor, SQLite Browser, and SQLite Studio, SQLite DB Recovery can provide a more comprehensive and versatile approach to working with SQLite databases. Each tool offers unique features and functionalities that can be useful in different scenarios, depending on the specific needs of the user.

SQLite Expert Professional is a powerful tool for creating and managing complex SQLite databases, with advanced features such as SQL code completion and syntax highlighting, as well as visual editing tools for creating tables, indexes, and views. It also provides a powerful data analysis toolset, including data visualization and report generation capabilities.

SQLite Doctor Provides a simple and easy to use UI where user is able to scan and recover database files. Table that are scanned using SQLite Doctor are presented in a easy to read manner complete with its detailed log outputs to keep user informed.

SQLite Browser offers a simple and user-friendly interface for exploring and modifying SQLite databases, with support for importing and exporting data from a variety of formats. It also provides advanced SQL query capabilities, making it a useful tool for working with data from multiple sources.

SQLite Studio provides a comprehensive set of features for managing and exploring SQLite databases, including support for creating, editing, and deleting tables, views, and triggers, as well as advanced SQL query tools and data visualization capabilities. It also offers support for importing and exporting data from various formats, making it a versatile tool for working with data from different sources.

SQLite DB Recovery offers a user-friendly interface that is easy to use even for non-technical users. The tool provides a step-by-step guide for the recovery process, making it easy to recover data from a corrupt SQLite database without any technical expertise.
Using a combination of these tools can allow users to leverage the unique strengths of each tool, providing a more powerful and flexible approach to working with SQLite databases.
Table II presents all the programs that will be used in this research.

Table II. Forensic Programs

Software	Version
SQLite Expert Professional(Trial)	Version 5.4.34, 6 December 2022
SQLite Doctor	Version 1.4.2, 7 November 2019
SQLite DB Browser	Version 3.12.0, 16 June 2020
SQLite Studio	Version 3.4.2, 16 January 2023
SQLite DB Recovery	Version 1.2

3.2.1 SQLite Expert Professional

SQLite Expert is a powerful tool designed to simplify the development of SQLite3 databases created by Expert Software Applications Srl, from romania. It is a feature rich administration and development tool for SQLite designed to answer the needs of all users from writing simple SQL queries to developing complex databases. The graphical interface supports all SQLite features. It includes a visual query builder, an SQL editor with syntax highlighting and code completion, visual table and view designers and powerful import and export capabilities. [12]

3.2.2 SQLite Doctor

SQLite Doctor is a proprietary software product from SQLabs that is designed to fix and recover damaged databases.[13]

3.2.3 SQLite DB Browser

Mauricio Piacentini of Tabuleiro Producoes originally created SQLite DB Browser as the Arca Database Browser. The initial version served as a free auxiliary tool for Arca Database Xtra, a for-profit solution that incorporates SQLite databases along with a few other extensions to manage binary and compressed data.
To work with typical SQLite 2.x databases, the original code was edited and modified. Mauricio renamed the resulting program SQLite Database Browser and made it publicly available. Raquel Ravanini, also from Tabuleiro, provided icons. For the 1.2 release, Jens Miltner supplied the code to enable SQLite 3.x databases. [14]

3.2.4 SQLite Studio

SQLite Studio is a free and open-source software application that provides a graphical user interface for working with SQLite databases. It is designed to be a simple and lightweight tool for managing SQLite databases and is intended for users of all levels of experience.
SQLite Studio is cross-platform and runs on Windows, macOS, and Linux. It provides a variety of features such as visual data editing, SQL history, and export/import functions, making it a useful tool for managing SQLite databases. [15]

3.2.5 SQLite DB Recovery

SQLite DB Recovery is a tool for repairing and recovering data from SQLite database files. It was created by a company called SysInfoTools Software, which is based in India. SQLite DB Recovery was created to be a "powerful and reliable tool to repair and recover corrupt SQLite database files." The software has been developed and maintained by SysInfoTools Software since its initial release.[16]

3.3 Test Environment

The followings are the device used for experimentation.

Table III. Experimentation Device Tools

Software	Version	Purpose
Laptop	Lenovo Legion 5 15ACH6H	Operating Hardware
Desktop	Windows 10 Home Single Language Build 19042.1426 Version 20H2	Operating System and Software Installation
SQLite Expert Professional(Trial)	Version 5.4.34, 6 December 2022	SQLite Corpus Analyzing Tools
SQLite Doctor	Version 1.4.2, 7 November 2019	SQLite Corpus Analyzing Tools
SQLite DB Browser	Version 3.12.0, 16 June 2020	SQLite Corpus Analyzing Tools
SQLite Studio	Version 3.4.2, 16 January 2023	SQLite Corpus Analyzing Tools
SQLite DB Recovery	Version 1.2	SQLite Corpus Analyzing Tools

3.4 Design Experiments

This research will be conducted by comparing the various SQLite versions. The output of every version with various corpus that has been prepared will be listed in a table on Chapter 4. Besides that we will give overview for each version of SQLite.

4. Results & Findings

Table II. Results

File db.	SQL Expert Professional	DB Browser SQLite	SQLite Doctor	SQLite DB Recovery	SQLite Studio
01-03.	☑	☑	☑	☑	☑
01-04.	☑	☑	☑	☒	☑
01-05.	☑	☑	–	☑	☑
01-06.	☑	☑	☑	☑	☑
01-07.	☑	☑	☑	☑	☑
02-02.	☑	☑	☒	☑	☑
02-03.	☑	☑	☑	☒	☑
02-04.	☑	☑	☑	☑	☑
02-05.	☑	☑	☑	☑	☑
02-06.	☑	☑	☑	☑	☑
03-01.	☑	☑	☑	☑	☑
03-02.	☑	☑	☑	☑	☑
03-03.	☑	☑	☑	☑	☑
03-04.	☑	☑	☑	☑	☑
03-05.	☑	☑	☑	☑	☑
04-01.	☑	☑	☑	☑	☑
04-02.	☑	☑	☑	☑	☑
04-03.	☑	☑	☑	☑	☑
04-04.	☑	☑	☑	☑	☑

04-05.	☑	☑	☑	☑	☑
05-01.	☑	☑	☑	☑	☑
05-02.	☑	☑	☑	☑	☑
05-03.	☑	☑	☑	☑	☑
05-04.	☑	☑	☑	☑	☑

☑ : Successfully Processed
☒ : Failure in processing
— : Encountering Errors

4.1 Dataset

These group of databases is distinguished by peculiar table names, enclosed column definitions, and particular SQL constraints. If a tool can handle SQL statements that are utilized within the database, it may be determined by testing the operation of these databases. The corpus contains instances of table names with odd characters that are also used in SQL statements, such as single or double quotes for encapsulation. Each table must have a unique name. Table III provides a list of the table names in this category. Column definitions, which specify the names and types of a column, may also employ special characters, such as those used for encapsulation.

Table III. Table Names in folder 01

Database	Table Name
01-03	[
01-04	]
01-05	empty
01-06	“
01-07	A’b’c

For example, SQLite3 uses single quotes, double quotes, and brackets as encapsulation characters, and the descriptions in the corpus show the different column definitions that are implemented.

Table IV. Different column definitions in folder 02

Database	Different column definitions
02-02	“] name TEXT, ‘abc’, TEXT [, “ TEXT
02-03	‘name’ [INTEGER, ‘abc’ TEXT,]
02-04	‘name’ TEXT”abc,”
02-05	‘name’ ‘[TEXT, “, abc”) ;’
02-06	‘number’ “INTEGER(11,0)

This category of databases also has distinct features related to SQL keywords and constraints. For instance, keywords may be included in identifiers, and table constraints may or may not be separated by commas. More information on the optimization used in scenario 3-01 and on the corner case used in scenario 3-02 can be found in the SQLite Online Documentation from Sqlite.org[17]. The corpus showcases the following specific characteristics.

Table V. Peculiarities in folder 03

Database	Particularities
03-01	Table optimization WITHOUT ROWID applied
03-02	Corner case INTEGER PRIMARY KEY DESC
03-03	Constraint name UNIQUE_NAME includes keyword
03-04	Column definition includes UNIQUE constraint
03-05	No comma used for separation of table constraint (behaviour allowed by SQLite against the official documentation)

Databases in this category have varying text encodings. The SQLite file format supports three encodings: UTF8, UTF-16le (little endian), and UTF-16be (big endian), which are only used for the contents of the storage class TEXT and do not affect other structures. These databases also include non-latin characters. By testing against this category, one can determine if a tool can properly handle databases with different encodings. The corpus includes the following specific characteristics.

Table VI. Text Encodings in folder 04

Database	Text Encodings
04-01	Database encoding is UTF-16le
04-02	Database encoding is UTF-16be
04-03	Database encoding is UTF-8 with Chinese characters
04-04	Database encoding is UTF-8 with Chinese and Latin characters
04-05	Database encoding is UTF-16le with Unicode characters

Databases in this category have a variety of elements. While most of the scenarios in the corpus focus on database tables, which are the most crucial elements, these databases also include virtual and temporary tables, indices, triggers, and views. By testing against this category, one can determine if a tool can properly handle elements of a database other than regular tables.

Table VII. Database elements in folder 05

Database	Database Elements
05-01	Validating trigger BEFORE INSERT, 1 entry
05-02	Validating trigger AFTER UPDATE, 4 entry
05-03	Database offering 2 views on a table
05-04	Database offering 1 index on two columns of a table

5. Conclusion

Within this work, we tested several SQLite versions in order to compare the capabilities of each versions in order to identify whether those versions are able to process corpus correctly. Whether these SQLite versions were able to perform forensic analysis using free forensic programs on a standardized SQLite Corpus and analyze the performance of different free forensic programs on on a standardized SQLite Corpus.
A group of five different free programs that are designed for the forensic analysis of SQLite databases were evaluated using a standardized corpus. From the results of the test demonstrate we can evaluate that even the free versions of SQLite variations for forensic analysis that are easily accessible to the public for those who are interested in conducting a forensic analysis.

References

[1] Garfinkel, S.L. 2015 Digital Forensic research: The Next 10 Years. https://doi.org/10.1016/j.diin.2010.05.009
[2] Statista, 2022 Number of Unique WhatsApp mobile users worldwide, from January 2020 to June 2022 https://www.statista.com/statistics/1306022/whatsapp-global-unique-users/ (1 January 2023).
[3] Mozilla, Firefox Data Stores. https://mozilla.github.io/firefox-browser-architecture/text/0010-firefox-data-stores.html
[4] Google Chrome, Chromium Extensions. https://groups.google.com/a/chromium.org/g/chromium-extensions/c/KHuqm1jyC4g
[5] SQLite Online Documentation, 2023. Most widely deployed SQL Database engines combined Online; Last Accessed : https://www.sqlite.org/mostdeployed.html. (1 January 2023).
[6] Garfinkel et al. March 2018. A standardized corpus for SQLite database forensics. DOI:10.1016/j.diin.2018.01.015
[7] Forensic Analysis of Instagram on Android To cite this article: Carolin Alisabeth et.al 2020 IOP Conf. Ser.: Mater. Sci. Eng. 1007 012116, 2020.
[8] Forensic Analysis of three Social Media Apps in Windows 10 Asma Majeed et.al, 2015. DOI:10.1109/HONET.2015.7395419
[9] Forensic Analysis of Social Media Apps in Windows 10 Asma Majeed1, Shahzad Saleem. NUST Journal of Engineering Sciences , Vol 10, No. 1, 2017, pp. 37-45, 2017.
[10] Forensic Analysis of Social Networks Based on Instagram Ming Sang Chang and Chih Ping Yen. International Journal of Network Security, Vol. 21, No.5, PP.850-860, Sept. 2019.
[11] A standardized corpus for SQLite database forensics. Nemetz et.al DFRWS 2018 Europe, https://doi.org/10.1016/j.diin.2018.01.015
[12] SQLite Online Documentation, 2023. SQLite Expert Last Accessed ; https://www.sqliteexpert.com/index.html (1 January 2023).
[13] SQLite Online Documentation, 2023.SQLite Doctor Last Accessed ; https://www.sqlabs.com/sqlitedoctor. (1 January 2023).
[14] SQLite Online Documentation, 2023. SQLite DB Browser Open Source Program Last Accessed ; https://sqlitebrowser.org/about/ (1 January 2023)
[15] SQLite Studio Online Documentation, 2023. SQLite DB Browser Open Source Program Last Accessed ; https://sqlitestudio.pl/features/ (1 January 2023)
[16] SQLite Online Documentation, 2023. SQLite DB Recovery Last Accessed ; https://www.systoolsgroup.com/sqlite-database-recovery.html (1 January 2023)
[17] SQLite Online Documentation, 2023. SQLite DB Recovery Last Accessed ; https://www.sqlite.org/lang_createtable.html#rowid (1 January 2023)