BADAN PENDIDIKAN DAN PELATIHAN KEUANGAN
KEMENTERIAN KEUANGAN
REPUBLIK INDONESIA

JURNAL BPPK

TOWARDS AN OPTIMIZED DATA PIPELINE IN INDONESIAN CUSTOMS AND EXCISE FORENSIC DATA ANALYTICS: A PRACTICE-AS-RESEARCH INQUIRY

Agung Darono[a], Aldi Pratama[b]

[a] Tax Education and Training Center, Jakarta, Indonesia, Email: agungdarono@kemenkeu.go.id
[b] Customs and Excise Education and Training Center, Jakarta, Indonesia, Email: aldi.pratama@kemenkeu.go.id (penulis berkorespondensi)

ARTICLE INFORMATION

ARTICLE HISTORY
Received
02 September 2024

Accepted to be published
31 December 2024

KEYWORDS:
Analytics
Customs
Data Pipeline
Excise
Forensic
Technology Stack

ABSTRACT

Analitik data forensik adalah metode penerapan analitik data dalam forensik digital. Penelitian ini menggunakan metode praktik-sebagai-penelitian untuk menunjukkan cara membangun jalur data yang dioptimalkan untuk analitik data forensik guna mendukung pelaksanaan forensik digital dalam penegakan hukum bea dan cukai di Indonesia. Studi ini menawarkan wawasan tentang bagaimana prinsip-prinsip umum, proses, dan fase dalam forensik digital dapat diterapkan pada penegakan hukum bea dan cukai di Indonesia, dengan fokus pada pengembangan jalur data dalam analitik data forensik.

Studi ini juga menyarankan praktik-praktik untuk meningkatkan pelaksanaan forensik digital bea dan cukai, seperti menerbitkan panduan rinci untuk pelaksanaan forensik digital, meningkatkan infrastruktur teknologi informasi, dan menciptakan sistem manajemen kasus yang mendukung jalur data forensik. Upaya-upaya ini bertujuan untuk menjunjung tinggi prinsip rantai penjagaan dalam forensik digital sepanjang proses penegakan hukum bea dan cukai.

Forensic data analytics is a method of applying data analytics in digital forensics. This research uses the practice-as-research method to show how to effectively establish an optimized data pipeline for forensic data analytics in support of implementing digital forensics in Indonesian customs and excise law enforcement. The study offers insights into how general principles, processes, and phases in digital forensics can be applied to Indonesian customs and excise law enforcement, with a focus on developing data pipelines in forensic data analytics. It also suggests practices to improve the implementation of customs and excise digital forensics, such as issuing a detailed guide for implementing digital forensics, enhancing information technology infrastructure, and creating a case management system that supports the forensic data pipeline. These efforts are aimed at upholding the chain of custody principle in digital forensics throughout the customs and excise law enforcement process.

1. INTRODUCTION

The Directorate General of Customs and Excise (DGCE) has the authority to conduct law enforcement in customs and excise, including investigating and bringing criminal charges against violators. In 2021, The DGCE issued Regulation Number PER-19/BC/2021 (hereinafter PER-19) outlining the organization's investigation procedures (DGCE, 2021). According to PER-19, investigators are authorized to exercise digital forensic techniques during the investigations. This authority is supported by Article 51 paragraph (2) of PER-19, which refers to Article 5A of Law 10/1995 and its amendments regarding Customs, and Article 3A of Law 10/1995 and its amendments regarding Excise. These articles affirm that electronic data is considered valid evidence in customs and excise regulations in Indonesia. Thus, applying digital forensics techniques in customs and excise law enforcement activities is both legitimate and technically essential, as most customs and excise-related crimes usually leave digital footprints (WCO, 2022).

Digital forensics is a scientific investigative procedure used to identify, store, analyze, and present digital evidence in accordance with legal provisions (Palmer, 2001; Zatyko, 2007); ISACA ,2015; Sabillon et al., 2017). The procedure needs an operational framework to ensure that the principles of digital forensics can be fully implemented. Several institutions, academics, and practitioners have proposed frameworks that outline the workflow necessary to ensure that digital forensic activities comply with established principles of digital forensics (ACPO, 2012; CII, 2021; UNODC, 2020). Specifically for Indonesia, the National Standards Agency (BSN) has issued SNI 27037 as a guideline for identifying, collecting, acquiring, and preserving digital evidence. Adapted from ISO/IEC 27037:2012, the guideline acts as an implementation standard (BSN, 2014). According to the standard, digital forensics involve acquisition and preservation of digital evidence, processing and analysis using valid devices, and reporting the results of the evidence analysis, that should be carried out in accordance with applicable legal provisions while maintaining the objectives of implementing digital forensics itself (Palmer, 2001; Carrier, 2003; Rigby and Rogers, 2007; Yusoff et al., 2011; ISO, 2012).

The data analytics field has rapidly progressed as a knowledge framework that emphasizes extracting insights, whether descriptive, predictive, or prescriptive, from large volumes of diverse data through various data processing algorithms to help the decision-making process. As a result, digital forensic practices have incorporated this framework into their operational procedures (EY, 2013; EYIN, 2013; Nomorissa and Suryadithya, 2022; Clopton, 2015). Data analysis, in general, involves the use of quantitative analysis, fact-based management, explanatory and predictive models, and data to guide decision-making and actions. In a broader context, analytics is the application of computational and logical reasoning to the elements found during the analysis, looking for patterns and potential future applications (Davenport and Harris, 2007; Davenport, 2013; Power et al., 2018; Valchanov, 2018). The application of data analytics within digital forensics is referred to as forensic data analytics (Clary, 2015).

The use of forensic data analytics (or forensic analytics, which will be used interchangeably in this article) has created a need for a data pipeline. This pipeline serves as a tool for transferring data from its source for analysis. It helps draw conclusions and present results in line with the objectives of digital forensics work. In the context of digital forensics, having a data pipeline also supports the implementation of the chain-of-custody principle by ensuring the availability of traces for every activity in the analysis process. Understanding the establishment of a data pipeline is critical for implementing forensic data analytics within digital forensics, including its application, specifically in customs and excise law enforcement.

The study aims to investigate the most effective methods for establishing a data pipeline in forensic data analytics, particularly within the customs and excise field of law enforcement. The methodology used in this study is practice research (Candy, 2006); Bulley and Şahin, 2021), which tries to comprehend the organizational activities within a practice arena and generate new knowledge to enhance these practices. The primary focus is on the community of practice established through digital forensics training for law enforcement in customs and excise, specifically at the Indonesian Custom and Excise Education and Training Center. The study aims to contribute practice-based knowledge to benefit digital forensics practitioners in customs and excise.

The paper is organized as follows: The first part includes the background, problem formulation, and study objectives. The second part consists of a literature review on the role of digital forensics in customs and excise law enforcement in Indonesia, followed by a review of data pipelines in forensics data analytics. The

third part explains the research methodology. The fourth part discusses the research findings. Finally, the fifth part includes conclusions and recommendations.

2. LITERATURE REVIEW

The literature review mainly focuses on digital forensics in customs and excise investigations and using data pipeline in forensic data analytics. The author compiles a logical basis and then presents a proposition to discuss the findings. Finally, the author concludes the study by presenting the results and recommendations.

Digital Forensics in Custom and Excise Investigation

One of the considerations of PER-19 is that duties, functions, and authorities in criminal investigation must be carried out professionally, transparently, and accountably for every criminal case to realize the supremacy of law that fulfills a sense of justice. Also, PER-19 includes provisions related to digital forensics in the implementation of criminal investigations in the field of customs and excise as regulated in PER-19 Article 51, paragraph (2), letter c. The provisions in PER-19 state that the implementation of digital forensics for customs and excise investigations should be based on a digital forensics warrant with the procedures as stipulated in this regulation. However, PER-19 does not explicitly explain the principles and procedural frameworks for digital forensics activities in the context of customs and excise investigations. Therefore, digital forensics activities for customs or excise investigations can refer to generally applicable principles and procedural frameworks. The formal regulation that is closely related to the implementation of digital forensics is PER-35/BC/2017 concerning Customs Audit and Excise Audit Procedures. The regulation defines an Investigation Audit as an audit conducted to assist in investigating customs and excise crimes. One stage in this audit is the collection of electronic data, which can be financial reports, books, records, and documents that serve as primary evidence of bookkeeping. The electronic data also includes letters related to business activities, electronic data, and inventory of goods, as well as letters pertaining to activities in the field of customs and excise.

The discipline of digital forensics has emerged due to the widespread use of digital technology in almost all human activities, leading to the rise of “cybercrime”. This term encompasses “crimes against computers” and “crimes using computers”. Cybercrime includes offenses related to electronic data in business and government administration, cyberbullying, cyberstalking, spamming, and cyberterrorism. ISACA (2015), Sabillon et al (2017), and Zatyko (2007) define digital forensics as the application of computer science and investigative procedures for legal purposes, with principles that include the analysis of digital data as evidence, legitimate authority for data acquisition, and the use of certain mathematical functions (hashing) to ensure the validity of evidence through validated devices. Raharjo (2013) states that digital forensics is a part of forensic science that involves the discovery and investigation of electronic data/information found on digital devices such as computers, mobile phones, tablets, personal digital assistants, networking devices, and storage devices. ISACA (2015) states that digital forensics is the process of identifying, storing, analyzing, and presenting digital evidence that complies with legal provisions, especially for judicial purposes. Digital forensics encompasses computer forensics, network forensics, database forensics, and mobile forensics. Based on these definitions, digital data acquisition must adhere to specific rules (formal law) to ensure its use as legal evidence.

A methodological framework is necessary for digital forensics to provide clear evidence of whether a crime has occurred. Rigby and Rogers (2007) proposed a digital forensics practice framework called the General Digital Forensics Model (GDFM). This framework presents all the principles, processes, and phases required for a digital forensic investigation in a graphical representation. Digital forensics involves linking three components: processes, clients, and elements. For instance, the forensics preparation process needs support from people, technology, data, location, and time to serve the client’s purpose. Therefore, this framework will be implemented for each forensic process to ensure the expected results.

According to ISACA (2015), analyzing digital data obtained through legal procedures is critical to digital forensic work. This stage requires various testing techniques to gather evidence of the presence or absence of an unlawful incident. Implementation of forensic data analytics, as depicted in Figure 2, is gaining significance in this process. It is essential to ensure the existence of a data pipeline that can efficiently transfer data from its source to be processed and analyzed according to the objectives of the digital forensic assignment. The

next section of this literature review will focus on the data pipeline in forensic data analytics.

Data Pipeline in Forensic Data Analytics

The data pipeline is crucial in data analytics as it ensures that data from different sources and formats can be processed, distributed, and utilized automatically for the organization's benefit. It serves two main functions: defining how data is obtained and automating the processes of extraction, transformation, combination, validation, and visualization (Munappy et al., 2020; Foidl et al., 2023; Alley, 2019; Rad and Ghobaei-Arani, 2024). According to Foidl et al. (2023) and Rigby and Rogers (2007), the data pipeline consists of various task layers to uphold the principle of chain of custody, a primary concept in digital forensics. These layers include:

(1) Data Acquisitions: This involves collecting data from scenes of suspected actions requiring law enforcement, as well as other data supporting the evidence process

(2) Recovery and Extraction: This layer focuses on recovering data to its native form and transforming data across various storage locations into a single repository while maintaining the chain of custody

(3) Examination and Analysis: Various techniques are used to explore data with valid and reliable software, utilizing algorithms such as file-based or row-based data matching, regression analysis, time series analysis, named entity recognition, image recognition, optical character recognition, or natural language processing

(4) Presentation and Reporting: This involves presenting evidence using techniques such as chronological descriptions or disclosing of key evidence so that digital evidence from forensic analysis can support legal decision-making through trials or equivalents.

The main difference between generic and forensic data analytics is their connection to the chain of custody. Digital forensics involves maintaining the integrity of evidence through a chain of custody to ensure the authenticity of the data and document all steps taken. Forensic data analytics should incorporate features that ensure the chain of custody for the digital evidence under analysis. In Figure 3, the case (workflow) and data management layer can act as a guarantor for implementing the chain of custody.

3. RESEARCH METHODOLOGY

This research is considered practice research because it aims to acquire new knowledge through hands-on experience rather than just relying on desk research, surveys, interviews, or traditional research methods like case studies and ethnographies. While it is commonly associated with artistic contexts, practice research can be found in various fields such as science, business, and healthcare (Candy, 2006; Gherardi, 2019; Bulley and Şahin, 2021; Cronan and Douglas, 1990). According to Candy (2006), practice research can be categorized into practice-based and practice-led research. Practice-based research uses a creative artifact as the basis for contributing to knowledge, while practice-led research primarily leads to new understandings about practice. The explicit goal of practice-led research is to make theoretical contributions to the field of practice. This type of research moves from practice to theory, focusing on the characteristics of the practice and generating practical insights for the specific profession. As practice research, this research aims to enhance our understanding of the practice itself and produce information that will impact practice.

This research is based on a community of practice (CoP) created through digital forensics training activities for law enforcement purposes in the customs and excise field at the Indonesia Customs and Excise Education and Training Center. The author was part of this CoP. Data collection methods included documentation studies, interviews, and the practice of creating data pipelines for forensic data analytics using case studies from law enforcement training in customs and excise. Data analysis involved using interactive data analysis techniques proposed by Miles et al (2019), comparing the results of creating data pipelines for forensic data analytics with the relevant law enforcement provisions in the customs and excise field.

4. RESULTS AND FINDINGS

Context of Practices

Based on the activities of the established CoP, several use cases have highlighted the requirement to establish a data pipeline in the context of applying forensic data analytics. Digital forensics, as part of customs and excise law enforcement, is conducted using PER-19, which, in its implementation, utilizes GDFM (see Figure 2) as an analytical framework. The context of the practices related to the use of forensic data pipelines in customs and excise digital forensics activities is as follows:

(1) PER-19 regulates and put an emphasize on documentation and general workflow.

However, it does not determine the specific functions and data flow required in digital forensics work. This suggests that custom and excise digital forensics can utilize general principles, processes, and phases in digital forensics.

(2) It is considered to use digital forensics' laboratory managed by a unit outside the DGCE, such as a digital forensics laboratory operated by the police or the Directorate General of Taxes, for data processing and analysis purposes.

(3) The results of digital forensics are more likely to be used in criminal lawsuits rather than tax disputes in the tax court. To enhance the establishment of data pipelines for forensic data analytics in custom and excise forensic analysis, the author will present the results of the practice carried out through case studies used in the CoP.

The practice entails digital forensics training for law enforcement purposes in the field of customs and excise, focusing on the following use cases:

(1) Establishing an optimized data pipeline
(2) Dealing with various hardware and related digital forensics approaches
(3) Operationalizing data pipelines
(4) Chain of custody.

Practice 1: Establishing an optimized data pipeline

Based on the results of observations and documentation studies conducted, the optimized data pipeline in the context of customs and excise forensic data analytics includes:

(1) data flow support that is in accordance with the stages of digital forensics work.
(2) affordable devices in terms of availability, budget adequacy, and competence of the human resources used.
(3) equipped with a case and workflow management system to guarantee the implementation of chain-of-custody as one of the main principles of digital forensics and compliance with good data governance.

In establishing a forensic analytics pipeline, it is essential to distinguish between two key elements: generic and specific technical specifications. Generic specifications cover the broad technical needs of the data pipeline, whereas specific specifications focus on the data examination and analysis.

Generic requirements

The data pipeline in practice is illustrated in **Error! Reference source not found.**. This c ustoms and excise forensic data pipeline utilizes input data from data acquisition work as part of digital forensic work in customs and excise law enforcement fieldwork. The technology stacks used for data acquisition include FTK Imager, EnCase Imager, and Oxygen Forensic Detective, with the selection depending on the situation at the scene. The steps in forensic data analytics include the following:

(1) Data Recovery, Extraction, and Transformation (DRET):
   a. Data recovery, if necessary, for digital data that requires recovery
   b. Data extraction to retrieve data spread across multiple storage locations in various formats for storage in a single repository
   c. Data transformation
(2) Data Interrogation and Analysis (DIA):
   a. Data interrogation as an exploratory step to determine the type and analysis technique for further processing
   b. Data analysis to find digital data that can be used as evidence
(3) Data Presentation and Reporting (DPR): step to present evidence through reports or trial processes based on the results of the analysis using various techniques, such as chronological descriptions or storytelling, for the evidence to be used in determining the overall results of the legal action taken.

The procedure is designed to cover the entire process, beginning with data collection during fieldwork and extending to the entry of forensic data into the pipeline. This specialized digital forensic work entails case and workflow management to uphold the principle of chain of custody. Additional information about this will be available in the Practice 4 section.

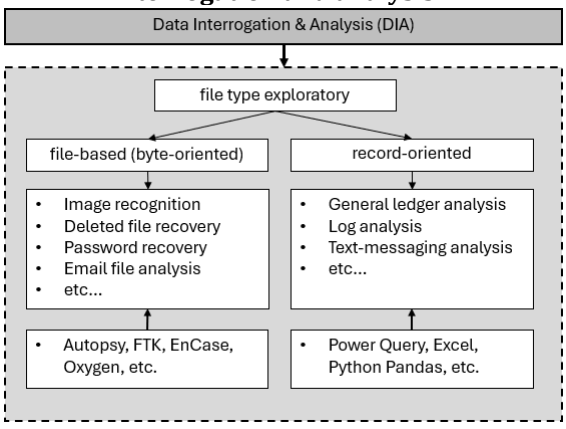
Specific requirements: dealing with data abstraction layers in Data Interrogation and Analysis

During the DIA stage of the forensic analytics pipeline, it is crucial to understand the file being analyzed. This understanding will determine the analysis techniques and tools required to achieve the objectives of the forensic tasks at hand. Carrier (2003) highlighted the significance of using abstraction layers techniques in relation to the data analysis stage for digital forensics. The primary purpose of this technique is to determine the appropriate tools for the data being analyzed and the purpose of the analysis itself. The digital data acquired is in

the form of a file. The term "file" refers to a collection of bytes (IBM, 2023) that represent the stored data, including images, sound, video, spreadsheets, documents, and other formats. These files can be grouped based on their abstraction layer: filesystem and non-filesystem. The filesystem describes data related to the system itself (meta-data), while a non-filesystem contains various user-stored data (images, sound, video, spreadsheets, documents, etc.) and can be retrieved according to the original purpose of its storage.

The development of data management needs has led to the concept of "records" (Mason, 2018)). This term refers to a file containing bytes of logically interconnected data using column and row techniques to represent data from real-world situations such as customer records, sales, asset ownership, etc. These record-oriented files, also known as structured data, can be in the form of a text file with a specific delimiter (e.g., comma-separated value / CSV), a spreadsheet (such as Microsoft Excel / XLSX), or a database (MySQL, SQLite, etc.). On the other hand, non-record files (unstructured data or byte-oriented) can include images, sounds, videos, or web pages. Forensic analytics involves the use of filesystems and non-filesystems, including byte-oriented and record-oriented data. Figure 1 outlines the data interrogation and analysis procedures, considering the type of file and the devices that may be utilized.

Figure 1 – File type exploratory in data interrogation and analysis



(source: authors' analysis from CoP observation and practice)

Practice 2: Dealing with various hardware and related digital forensics approaches

This practice focuses on creating a data pipeline for analyzing data collected from mobile devices running the Android operating system and computers (servers, desktops, or laptops).

Mobile phone forensics

1) Data Extraction and Transformation: Evidence of data acquisition, transformation steps, tools used.

Handling the mobile devices is the first procedures forensic auditor should consider. The procedure involves obtaining consent from the user to grant access to the device to the forensic examiner. Once permission is granted, the forensic auditor will assess the device to determine which data needs to be acquired, the method of acquisition to be used, and the tools that will assist in the acquisition process. Prior to the acquisition, the forensic auditor will create a backup of the messaging application database using the ADB (Android Debug Bridge) command via USB debugging. The messaging database is encrypted with a .crypt14 extension. Root access is required to extract the database, as the database file is located in the '\com.whatsapp\databases\msgstore.db' and '\com.whatsapp\databases\wa.db' directories, which can only be accessed with root permissions. The decryption key for the database is obtained from the Whatsapp application using the AVD (Android Virtual Device) tool, which is used to emulate the Android operating system. The encrypted database is then inserted into the AVD, and the WhatsApp application is installed. Once the installation is complete, a one-time password (OTP) code is sent by the WhatsApp Server to the registered phone number. Entering this OTP code decrypts the database, allowing it to be acquired using the ADB command.

Before proceeding to data acquisition, a forensic auditor should ensure that the devices are handled properly. Start by deactivating the cellular signal using airplane mode and storing the device in a Faraday bag to prevent any signal interference that could alter its condition. Additionally, turn off security features such as pins, patterns, and face recognition. When acquiring data from mobile devices, the forensic auditor may utilize logical extraction or physical extraction methods depending on some factors, such as the device's state (on/off, functioning normally/not, locked/unlocked), time constraints, and the urgency of acquiring the data. For customs and excise purposes, the acquired data from mobile devices typically consists of a messaging application database (e.g., WhatsApp) containing details of user conversations, including the other party's telephone number, messages, message delivery

time, images, and location. In this scenario, the logical extraction method is typically employed to obtain data from mobile devices.

2) Data Interrogation and Analysis: type of data analysis, case conclusions

The acquired and decrypted data is stored in SQLite format. To analyze the data, we first connect the SQLite database to an appropriate ODBC driver to create a data source on the Windows operating system. Then, we use the Power BI to connect to the data source using the Get Data from ODBC command. ODBC driver allows Power BI to read and parse the tables in the database for further analysis. Analysis in Power BI on the SQLite database can include categorization, text search, location extraction, and social network analysis.

Servers, Desktops or Laptops

1) Data Extraction and Transformation: Evidence data acquisition, transformation steps, tools used.

This forensic analytics practice involves the collection of investigation audit evidence, which can take the form of financial reports, books, records, and documents that serve as primary evidence of bookkeeping, business-related letters, electronic data, inventory of goods, and letters related to customs and excise activities. In this particular scenario, the forensic auditor obtains data in the form of a database from an accounting application used by the party undergoing audit. The accounting application database stores financial transaction information from the business entity. This information is then analyzed to obtain evidence of customs violations. The accounting data is stored on a computer that is regularly used in the company's daily activities. Once the computer has been identified, the forensic auditor conducts an initial handling and assessment of the computer. Based on the assessment results of the computer device, it was found that a bookkeeping application using the Firebird database is employed to record the company's daily economic activities.

In this scenario, data extraction is carried out in real time using the FTK Imager device. Forensic software is employed to maintain the integrity of the data being acquired. The entire acquisition process is documented in a report along with a hash value from the acquired accounting database file. Once the acquisition process is completed, the forensic auditor verifies the integrity of the acquired data through validation. The acquired data is then analyzed using a digital forensic framework. Initially, the forensic auditor creates multiple working copies for analysis. At the same time, the acquired data is securely stored as an original copy. After the working copy is made, the forensic auditor verifies that it matches the original copy precisely.

2) Data Interrogation and Analysis: type of data analysis, case conclusions

The next stage is Data Interrogation and Analysis. The steps involved in this stage are the following:

- Conducting data consistency tests, including gap and duplicate tests, and data master-detail relation tests, to ensure the validity and integrity of the data
- Performing financial ratio analysis and conducting data record searches using specific criteria or keywords to identify evidence indicating alleged violations

Practice 3: Data pipeline operationalization

When the data pipeline, as described in Practice 1, is executed to perform data analysis, several variants of the operationalization of the data pipeline are found. Figure 4 illustrates at least three variants encountered. These variants occur based on the hardware used to handle each case whose data will be analyzed.

(1) Pipeline 1: This data pipeline operates on a single computer device (PC/laptop) that works for all stages (DRET, DIA, and DPR) in forensic data analytics.
(2) Pipeline 2: This data pipeline operates using Network-attached storage (NAS) as a place to store source data (pristine-copy and working copy) for the DET, DIA, and DPR steps using the PC/laptop of the assigned data analyst.
(3) Pipeline 3: This data pipeline operates using Network-attached storage to store source data (pristine copy and working copy). It is also equipped with a GPU-based device to support specific jobs such as password recovery using brute-force techniques.

A change of custody repository is required to monitor all cases being worked on using their respective data pipelines, ideally in the form of a case (workflow) management application system. However, at the time this practice was carried out, the application was not yet available, so the author tried to compile it as explained in Practice 4.

Practice 4: Chain of custody

Chain of custody is a crucial principle in digital forensics. It ensures that the entire digital forensics process can be traced from start to finish. This allows other competent parties in the field to replicate the process and obtain the same results. Case and workflow management practices are chosen to implement the chain of custody principle using available tools in the compiled data pipeline. These practices involve processing various digital data in accordance with PER-19, which serves as a legal basis for digital forensics activities in customs and excise investigations. Currently, the practice is limited to data extraction and transformation from data sources (such as XLSX or PDF files) to produce a summary of digital forensics activities. From these practices, it is vital to urgently implement a case and workflow management system with key features, including a to-do list for each case, progress tracking for each case, and risk assessment to mitigate potential threats to implementing the chain of custody principle in each case.

5. CONCLUSIONS

Based on several use cases, in this study, we can draw the following conclusions regarding the implementation of practices as part of the practice-as-research approach for digital forensics in law enforcement for customs and excise:

(1) Competent human resources are needed in both digital forensics practices and legal provisions of customs and excise
(2) The technological infrastructure should be adjusted with the budget available for the project.
(3) The methodology to establish a data pipeline should consider the characteristics of the case being managed.
(4) Institutional cooperation is essential to support the entire digital forensic ecosystem to benefit customs and excise law enforcement.

This study proposes recommendations based on the practice carried out using a specific case as the practice arena. The recommendations are related to setting up data pipelines for implementing forensic data analytics for customs and excise law enforcement. The recommendations are as follows:

(1) PER-19 should be accompanied by technical guidelines (circulars for the scope of DGCE) that provide detailed explanations of the steps for implementing digital forensics, including the use of forensic data analytics
(2) Developing standards for the formation of information technology infrastructure that supports the availability of forensic data pipelines, both at the head office and vertical DGCE agencies, so that it can support the implementation of the chain of custody principle in digital forensics as part of the whole process of custom and excise law enforcement.

REFERENCES

ACPO. 2012. ACPO Good Practice Guide for Digital Evidence. The Association of Chief Police Officers (ACPO).

Alley, Garrett. 2019. “Data Integration vs. Data Pipeline: What’s the Difference?” 2019. https://dzone.com/articles/data-integration-vs-data-pipeline-whats-the-differ.

BSN. 2014. “SNI ISO/IEC 27037:2014 Teknologi Informasi - Teknik Keamanan - Pedoman Identifikasi, Pengumpulan, Akuisisi Dan Preservasi Bukti Digital (ISO/IEC 27037:2012, IDT) -.” Jakarta: Badan Standardisasi Nasional (BSN).

Bulley, James, and Özden Şahin. 2021. Practice Research - Report 1: What Is Practice Research? And Report 2: How Can Practice Research Be Shared? UKRI (Research England).

Candy, Linda. 2006. “Practice Based Research: A Guide.” Creativity and Cognition Studios Report 1 (November).

Carrier, Brian. 2003. “Defining Digital Forensic Examination and Analysis Tool Using Abstraction Layers.” International Journal of Digital Evidence 1 (4).

CII. 2021. “THE PRINCIPLES OF DIGITAL EVIDENCE.” Conference of International Investigators (CII). https://www.ciinvestigators.org/wp-content/uploads/2021/11/CII-General-Principles-for-Digital-Evidence-21stCII.pdf.

Clary, Scott. 2015. “Advanced Data Analytics in Fraud Identification.” https://chapters.theiia.org/houston/Documents/Government%20Auditors%20Conference%20Documents/EY%20-%20Advanced%20data%20analytics%20i

n%20investigations%202015%2002%2018.pptx.

Clopton, Jeremy. 2015. "All The Data: Integrating Data Analytics And Digital Forensics Into Fraud Examinations." In . Baltimore: ACFE. http://www.fraudconference.com/uploadedFiles/Shared_Content/Course_Materials/26th/cpp/9F-Jeremy-Clopton.pdf.

Cronan, Timothy Paul, and David E. Douglas. 1990. "End-User Training and Computing Effectiveness in Public Agencies: An Empirical Study." Journal of Management Litformation Systems I Spring 1990. Vol. 6. No. 4:21–39.

Davenport, Thomas H. 2013. "Introduction: The New World of Enterprise Analytics." In Enterprise Analytics Optimize Performance, Process, and Decisions Through Big Data, edited by Thomas H. Davenport. New Jersey: Pearson Education, Inc.

Davenport, Thomas H., and Jeanne G. Harris. 2007. Competing on Analytics: The New Science of Winning. Harvard Business Review Press.

DGCE. 2021. Regulation of Director General Customs and Excise Number PER-19/BC/2021. Investigation Procedures in Directorate General Customs and Excise.

EY. 2013. "Forensic Data Analytics." UK: Ernst & Young LLP (EY).

EYIN. 2013. "Forensic Data Analytics." Kolkata, India: Ernst & Young LLP (EY).

Foidl, Harald, Valentina Golendukhina, Rudolf Ramler, and Michael Felderer. 2023. "Data Pipeline Quality: Influencing Factors, Root Causes of Data-Related Issues, and Processing Problem Areas for Developers." Journal of Systems and Software 207 (September):111855. https://doi.org/10.1016/j.jss.2023.111855.

Gherardi, Silvia. 2019. How to Conduct a Practice-Based Study Introduction.

IBM. 2023. "An Introduction to the Byte File System." IBM. https://www.ibm.com/docs/en/zvm/7.3?topic=system-introduction-byte-file.

ISACA. 2015. "Overview of Digital Forensics." Rolling Meadows, IL, USA: ISACA. http://www.isaca.org/knowledge-center/research/researchdeliverables/pages/overview-of-digital-forensics.aspx.

ISO. 2012. "Information Technology — Security Techniques — Guidelines for Identification, Collection, Acquisition, and Preservation of Digital Evidence." International Organization for Standardization (ISO).

Mason, Tony. 2018. "Database versus File System By in Database, File Systems." https://fsgeek.ca/2018/01/09/database-versus-file-system/.

Miles, Matthew B., A. Michael Huberman, and Johnny Saldaña. 2019. Qualitative Data Analysis - A Methods Sourcebook 4th Edition. SAGE Publications.

Munappy, Aiswarya Raj, J. Bosch, and Helena Homström Olsson. 2020. "Data Pipeline Management in Practice: Challenges and Opportunities." In International Conference on Product Focused Software Process Improvement. https://api.semanticscholar.org/CorpusID:227129927.

Nomorissa, Telsy Aldemadra, and Chelsia Suryadithya. 2022. "Forensic Data Analytics Dalam Mendeteksi Fraud." Proceeding Accounting Skill Competition 1 (1): 161–80.

Palmer, Gary. 2001. "A Road Map for Digital Forensic Research By Collective Work of All DFRWS Attendees." In Report From the First Digital Forensic Research Workshop (DFRWS). New York.

Power, D.J., C. Heavin, J. McDermott, and M. Daly. 2018. "Defining Business Analytics: An Empirical Approach." Journal of Business Analytics 1 (1): 40–53. https://doi.org/10.1080/2573234X.2018.1507605.

Rad, Zahra Shojaee, and Mostafa Ghobaei-Arani. 2024. "Data Pipeline Approaches in Serverless Computing: A Taxonomy,

Review, and Research Trends." Journal of Big Data 11 (1): 82. https://doi.org/10.1186/s40537-024-00939-0.

Raharjo, Budi. 2013. "Sekilas Mengenai Forensik Digital." Jurnal Sosioteknologi 12 (29): 384–87.

Rigby, Steven, and Marcus K. Rogers. 2007. "The General Digital Forensics Model." In Annual ADFSL Conference on Digital Forensics, Security and Law.

Sabillon, Regner, Jordi Serra-Ruiz, Víctor Cavaller, and Jeimy Cano M. 2017. "Digital Forensic Analysis of Cybercrimes: Best Practices and Methodologies." International Journal of Information Security and Privacy (IJISP) 11 (April):25–37. https://doi.org/10.4018/IJISP.2017040103.

UNODC. 2020. "Standards and Best Practices for Digital Forensics." UNODC Teaching Module Series: Cybercrime. United Nations Office on Drugs and Crime (UNODC). https://sherloc.unodc.org/cld/en/education/tertiary/cybercrime/module-4/key-issues/standards-and-best-practices-for-digital-forensics.html.

Valchanov, Iliya. 2018. "Data Science vs Machine Learning vs Data Analytics vs Business Analytics." https://www.kdnuggets.com/2018/05/data-science-machine-learning-business-analytics.html.

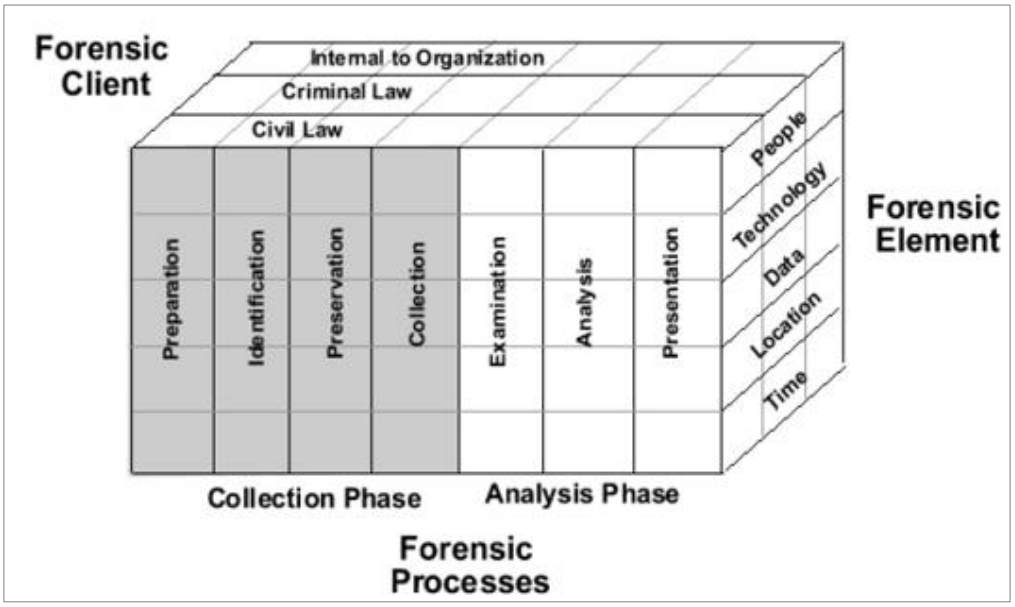
WCO. 2022. "WCO-WTO Study Report on Disruptive Technologies." WCO. https://www.wcoomd.org/en/topics/facilitation/instrument-and-tools/tools/wco-wto-study-report-on-disruptive-technologies-2022.aspx.

Yusoff, Yunus, Roslan Ismail, and Zainuddin Hassan. 2011. "Common Phases Of Computer Forensics Investigation Models." International Journal of Computer Science & Information Technology (IJCSIT) 3 (3): 17–31.

Zatyko, Ken. 2007. "Commentary: Defining Digital Forensics," 2007. https://www.forensicmag.com/article/2007/01/commentary-defining-digital-forensics.

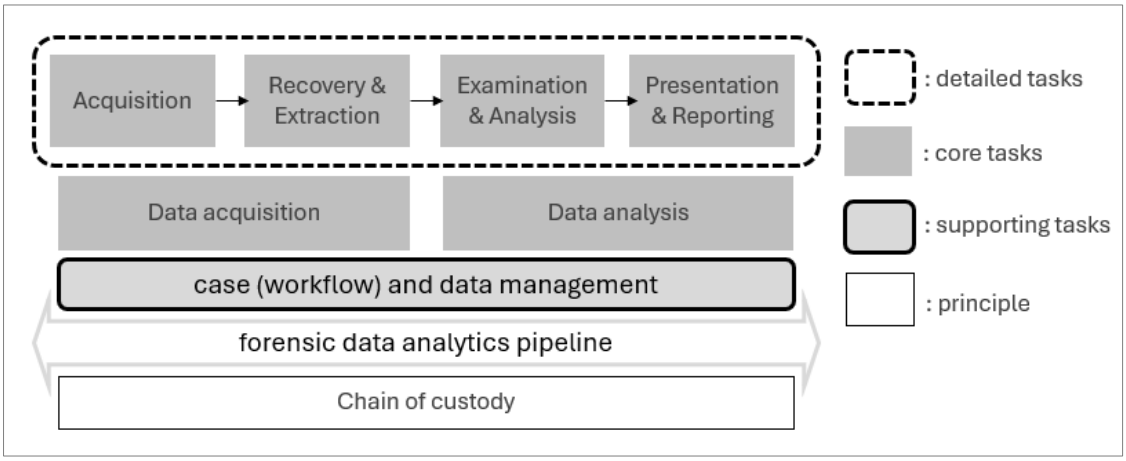
APPENDIX

Figure 2 - General Digital Forensics Model (GDFM)



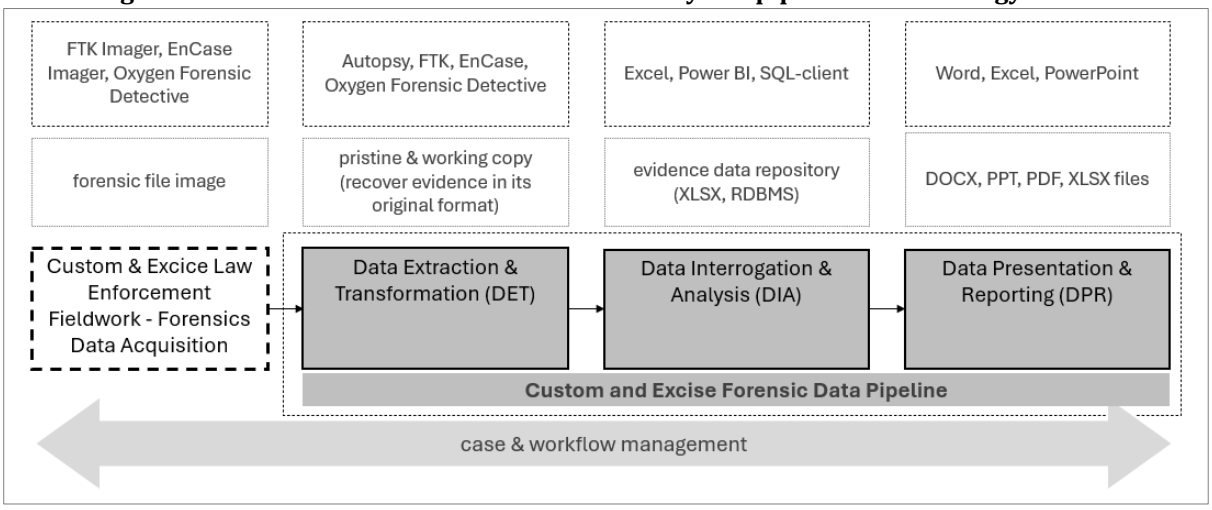
(source: Rigby and Rogers 2007)

Figure 3 – Forensic data analytics pipeline



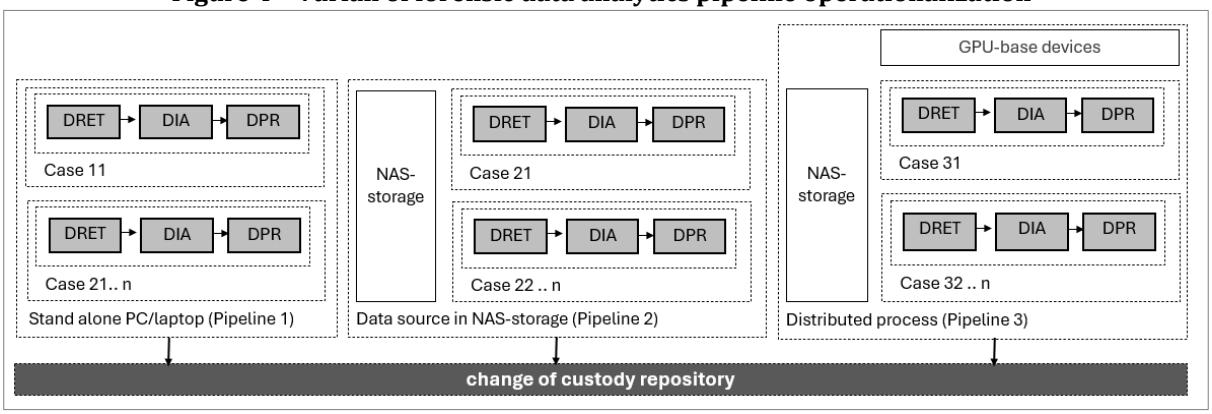
(source: adapted from Foidl et al. 2023; Rigby and Rogers 2007)

Figure 3 – Customs and Excise forensic data analytics' pipeline & technology stacks



(source: authors' analysis from CoP observation and practice)

Figure 4 – Varian of forensic data analytics pipeline operationalization



(source: authors' analysis from CoP observations and practices)